Viharos, Zs. J.; Ciani, L.: Proceedings of the *16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"*, 2019.09.03-04., Berlin, Germany, ISBN: 978-92-990084-1-6, 2019, p. 165.

IMEKO International Measurement Confederation

Technical Committee TC10 Testing, Diagnostics & Inspection

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

Berlin, Germany, 3-4 September, 2019

Fraunhofer IPK, Institute for Production Systems and Design Technology

PROCEEDINGS

Organised by

IMEKO TC10

Technical Committee on Testing, Diagnostics & Inspection

EUROLAB aisbl

eurolab aisbl

Sponsored by

MTA SZTAKI Hungarian Academy of Sciences Institute for Computer Science and Control

CENTRE OF EXCELLENCE EPIC Production Informatics and Control

Fraunhofer IPK INSTITUTE PRODUCTION SYSTEMS AND DESIGN TECHNOLOGY

Editors: **Dr. Zsolt János Viharos and Prof. Lorenzo Ciani**

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

INVITATION

The International Measurement Confederation IMEKO, Technical Committee 10 on Testing, Diagnostics & Inspection and Eurolab aisbl kindly invites you to attend the 16th IMEKO TC10 Conference: "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety" to be held in Berlin, Germany, on September 3-4, 2019. It is the first conference jointly organized by IMEKO & EUROLAB. The Conference is a forum for advancing knowledge and exchange ideas on methods, principles, instruments, technologies and IT tools, standards, industrial applications, conformity assessment, quality management and measurement challenges on Testing, Diagnostics & Inspection as well as their diffusion across the scientific community. Participants have an excellent opportunity to meet top specialists from industry, the TIC Sector (Testing, Inspection & Certification) and academia all over the world and to enhance their international co-operation. The programme will feature scientists and experts as leading keynote speakers for selected presentations on the main topics of the Conference.

SPECIAL ISSUE

Selected papers of the conference will be invited to the Measurement and ACTA IMEKO Special Issues. All submitted papers will undergo a regular peer review process. The manuscript MUST be significantly extended beyond the IMEKO TC10 conference paper.

ISBN: 978-92-990084-1-6

http://www.imekotc10-2019.sztaki.hu/

Editors: **Dr. Zsolt János Viharos and Prof. Lorenzo Ciani**

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

SCIENTIFIC TOPICS

- T1 - Basic principles and development trends in testing, diagnostics and inspection
- T2 - Condition monitoring and maintenance of industrial processes, plants and complex systems: measurements and methods
- T3 - Testing, diagnostics, inspection and prognostics for maintainability, safety, risk assessment and management
- T4 - Regulatory framework (for safety of industrial products, food protection, environmental protection, health protection, customer protection, information security etc.)
- T5 – Data Analytics, artificial intelligence techniques and machine learning for testing, diagnostics and inspection
- T6 - Testing, diagnostics and inspection applications in industry, transportation, mechatronics, avionics, automotive, food and biomedical fields
- T7 - Decision support and IT solutions for testing, diagnostics and inspection
- T8 - Testing, diagnostics and inspection for the improvement of quality of life and environment
- T9 - Product conformity assessment and process analysis
- T10 - Certification (products, management systems, persons) and accreditation
- T11 - Traceability in testing, diagnostics and inspection
- T12 - Measurements and methods for TQM (Total Quality Management);
- T13 - Technical testing, diagnostics and inspection for Cyber security
- T14 – Non-destructive testing, diagnostics and inspection
- T15 –Diagnostics and Monitoring with Cyber Physical Systems
- T16 – Fog, Edge and Cloud infrastructure for distributed diagnostics, inspection and monitoring
- T17 – Industry 4.0 foundations, applications, trends and novelties

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

ORGANISATION

General Chairs

Dr. Zsolt János Viharos

Chairperson of the IMEKO TC10 on Testing, Diagnostics & Inspection
President of the Hungarian National IMEKO Committee
Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences (MTA SZTAKI), Budapest, Hungary
Research Laboratory on Engineering and Management Intelligence
viharos.zsolt@sztaki.mta.hu

Eckhard Hohwieler

Fraunhofer IPK, Institute for Production Systems and Design Technology, Berlin, Germany
Production Systems Division
Head of Production Machines and System Management department
eckhard.hohwieler@ipk.fraunhofer.de

Prof. Álvaro Silva Ribeiro

President EUROLAB aisbl, Brussels, Belgium
Head of Hydraulics Metrology Laboratory, National Laboratory for Civil Engineering (LNEC), Lisbon, Portugal
asribeiro@lnec.pt

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

TECHNICAL PROGRAMME CHAIRS

General Chair

Prof. Lorenzo Ciani

Scientific secretary of the IMEKO TC10 on Testing, Diagnostics & Inspection
University of Florence, Florence, Italy
School of Engineering
DINFO - Department of Information Engineering
lorenzo.ciani@unifi.it

Technical Chairs

Claudio Geisert
Fraunhofer IPK, Institute for Production Systems and Design Technology, Berlin, Germany
Production Systems Division
Production Machines and System Management department
claudio.geisert@ipk.fraunhofer.de

Mladen Jakovcic, MSc
President of HMD, Croatian Metrology Society, Zagreb, Croatia
mladen.jakovcic@hmd.hr

Dr. Rolf Kumme
Physikalisch-Technische Bundesanstalt (PTB),
Braunschweig and Berlin, Germany
Head of Department 1.2 Solid Mechanics
rolf.kumme@ptb.de

Prof. Kurt Ziegler
President of EUROLAB-Deutschland, Berlin, Germany
kurt.ziegler@gmx.de

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

International, Technical Programme Committee Members

MSc. Biserka Bajzek Brezak	**(CRO)**
Dr. Jozsef Beinschroth	**(HUN)**
Prof. Yakov BenHaim	**(ISR)**
Prof. Piotr Bilski	**(POL)**
Dr. Oleg Bushuev	**(RUS)**
Dr. Marco Carratu	**(ITA)**
Prof. Marcantonio Catelani	**(ITA)**
Prof. Wojciech Cholewa	**(POL)**
Prof. Marcello Colledani	**(ITA)**
Dr. Ana Cop	**(CRO)**
Prof. Loredana Cristaldi	**(ITA)**
Prof. David Delaux	**(FRA)**
Prof. Eduard Egusquiza	**(ESP)**
Prof. Diego Galar	**(SWE)**
Dr. Antonella Gaspari	**(ITA)**
Prof. Emilia Giulio	**(ITA)**
Prof. Lovorka Grgec Bermanec	**(CRO)**
Dr. Marijan Grgic	**(CRO)**
Ing. Giulia Guidi	**(ITA)**
Prof. Charaf Hassan	**(HUN)**
Dr. Yukio Hiranaka	**(JPN)**
Prof. Geza Husi	**(HUN)**
Dr. Justinas Janulevicius	**(LTU)**
Dr. Csaba Johanyak	**(HUN)**
Prof. Ivanka Lovrencic	**(CRO)**
Dr. Thomas Lundholm	**(SWE)**
Prof. Laszlo Monostori	**(HUN)**
Dr. Emanuela Natale	**(ITA)**
Dr. Michael Nitsche	**(DEU)**
Ing. Gabriele Patrizi	**(ITA)**

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

Editors: **Dr. Zsolt János Viharos and Prof. Lorenzo Ciani**

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

Local organisers

Gusztáv Hencsey
Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences (MTA SZTAKI), Budapest, Hungary
hencsey.gusztav@sztaki.mta.hu

Claudia Engel
Fraunhofer IPK, Institute for Production Systems and Design Technology, Berlin, Germany Head of Media and Public Relations / Event management
claudia.engel@ipk.fraunhofer.de

Financial coordination

Stefkóné Vermes Judit
Congress Rendezvényszervező Kft., Budakeszi, Hungary
stefko.judit@congress.hu

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

WORKSHOP AWARDS

An award will be given for the **Best Scientific Paper & Presentation of the Conference**.

To encourage the attendance of young researchers, an award will be given for the **Best Paper Authored and Presented by a Student**.

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

INVITED KEYNOTE LECTURERS

Industrial Keynote Lecturer

Dr. Álvaro Silva Ribeiro

PhD in Tech. Physics
President of the BoA of EUROLAB aisbl and RELACRE
Vice-President of UILI
Head of the Metrology Division and Quality Manager at LNEC

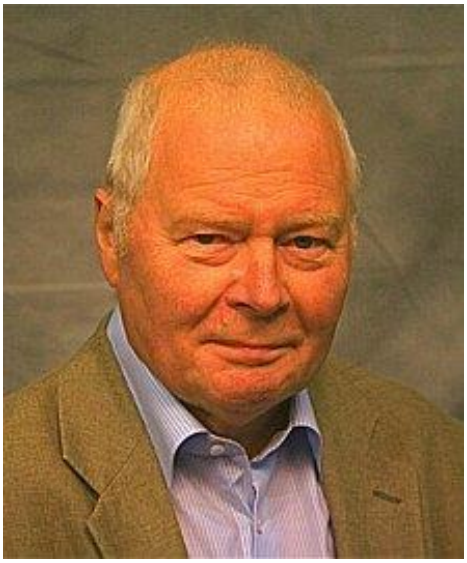
Presentation title: ***Measurement uncertainty added value for experimental research and testing in civil engineering***

Álvaro Silva Ribeiro, President of the BoA of EUROLAB aisbl and RELACRE (Eurolab Portugal), Vice-President of UILI, graduated in Tech. Physics, MSc in Instrumentation, Industrial Maintenance and Quality and PhD in Tech. Physics. Currently, he is Head of Metrology Division and Quality Manager at LNEC (National Institute for Civil Engineering, Research Institute in Lisbon), with interests in Metrology (flow, temperature and humidity, geometric dimensions, force, mass and pressure), quality management systems, accreditation, mathematical modelling, numerical simulation and measurement uncertainty, and measurement quality of testing in civil engineering domains (bridges, dams, seismic infrastructures and others). He is founder of the Portuguese Society for Metrology and member of the Portuguese Society of Physics, being author of a large number of publications in these fields of knowledge. He is member of several international committees and member of EURAMET Research Council.

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

Scientific Keynote Lecturer

Prof. Dr. Horst Czichos
Professor of Mechatronics
University of Applied Sciences, BHT Berlin, Germany



Presentation title: ***Fundamentals of Testing: the Role of Metrology, Sensors and Standards***

Horst Czichos, formerly President of BAM and EUROLAB, graduated in Precision Engineering and worked as design engineer in the optical industry. He holds degrees in Physics and in Materials Science. The University KU Leuven awarded him an honorary doctorate for his research in Tribology and in Technology Studies for European Research Programmes. Currently, he is Professor for Mechatronics at the University of Applied Sciences, BHT Berlin. He authored and edited several books, including the Springer Handbook of Metrology and Testing (Springer 2011), the Handbook of Technical Diagnostics (Springer 2013) and the Textbook Measurement, Testing and Sensor Technology (Springer 2018).

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

VENUE

The conference is held in **Fraunhofer IPK, Institute for Production Systems and Design Technology**, Pascalstraße 8-9, 10587 **Berlin, Germany**

IMEKO
International Measurement Confederation
Technical Committee TC10
Testing, Diagnostics & Inspection

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

Editors: **Dr. Zsolt János Viharos and Prof. Lorenzo Ciani**

16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"

GALA DINNER

Schnitzelei Charlottenburg

Röntgenstraße 7b, 10587 **Berlin**

CONTENTS

Invited lecture, industrial
Dr. Álvaro Silva Ribeiro
PhD in Tech. Physics, President of the BoA of EUROLAB aisbl and RELACRE, Vice-President of UILI, Head of the Metrology Division and Quality Manager at LNEC
Measurement uncertainty added value for experimental research and testing in civil engineering

Scientific Session: **Condition monitoring and maintenance of industrial processes, plants and complex systems: measurements and methods**
Chairperson: **Prof. Giulio D'Emilia**
IMEKO TC10 member, Department of Industrial and Information Engineering and Economics, University of l'Aquila, L'Aquila, ITALY

Scientific Session: **Data Analytics, artificial intelligence techniques and machine learning for testing, diagnostics and inspection, Traceability and Maintainability, safety, risk assessment and management**
Chairperson: **Prof. Eduard Egusquiza**
IMEKO TC10 member, Center for Industrial Diagnostics and Fluid Dynamics (CDIF), Polytechnic University of Catalonia (UPC), Barcelona., Spain

Closing & Award Ceremony

Dr. Zsolt János Viharos
Chairperson of the IMEKO TC10 on Testing, Diagnostics & Inspection, President of the Hungarian National IMEKO Committee, Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences (MTA SZTAKI), Budapest, Hungary, Research Laboratory on Engineering and Management Intelligence

Mladen Jakovcic, MSc
President of HMD, Croatian Metrology Society, Zagreb, Croatia, Organiser of the 17th IMEKO TC10 conference "Global trends in Testing, Diagnostics & Inspection for 2030", Dubrovnik, Oct 19 - 22, 2020 (https://www.imekotc10-2020.com/)

SCIENTIFIC PUBLICATIONS

Editors: **Dr. Zsolt János Viharos and Prof. Lorenzo Ciani**

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

RTOS aware non-intrusive testing of cyber-physical systems in HIL (Hardware In the Loop) environment

Balázs Scherer

Department of Measurement and Information Systems
Budapest University of Technology and Economics
Budapest, Hungary
scherer@mit.bme.hu

***Abstract* – Statistics show that more and more cyber-physical systems are using RTOS (Real-Time Operating System). RTOS based system software can introduce multitasking and real-time behaviour based errors. Therefore, the testing processes of such systems should address the detection of these possible errors. Unfortunately, the multitasking and real-time behaviour based errors are among the hardest to detect. The best way for the detection is to perform extensive testing in a very realistic environment like in a HIL (Hardware In the Loop) simulation. These environments provide the possibility to perform overloaded event simulation, that increase the chance of multitasking and real-time behaviour based error occurrence. Traditionally the RTOS aware measurement methods (providing information for the error detections) use software instrumentation, and are not integrated into HIL test environments. This paper introduces a novel integration of RTOS aware measurements into a HIL test development environment. Our integration also focuses on the non-intrusive measurements of multitasking behaviour of RTOS based software systems. These non-intrusive measurements are not widespread and their integration into a HIL based environment is also a novel solution.**

Keywords – Testing, HIL (Hardware In the Loop) tests, RTOS (Real-Time Operating System) aware testing, Non-intrusive software measurements.

I. INTRODUCTION

Statistics show that over two thirds of embedded systems use RTOS (Real-Time Operating System) [1]. The type and complexity of these RTOS versions are very divers, starting with Embedded Linux versions running on high performance application processors down to simple preemptive multitasking kernels like (FreeRTOS and uC/OS) running on small microcontrollers of typical cyber-physical systems. Our work will focus on the microcontroller based systems using simple preemptive multitasking kernels.

A. Typical testing environment of microcontroller based cyber-physical systems

One of the typical test method of microcontroller based cyber-physical systems is the HIL (Hardware in the Loop) test, where the behaviour of the integrated software and hardware of the DUT (Device Under Test) can be investigated in a simulated and stable environment. These tests are originated and widespread in the automotive and transportation industry, but using them for general purpose cyber-physical systems is getting more and more widespread [2], [3]. A typical HIL test setup is shown on Fig. 1.

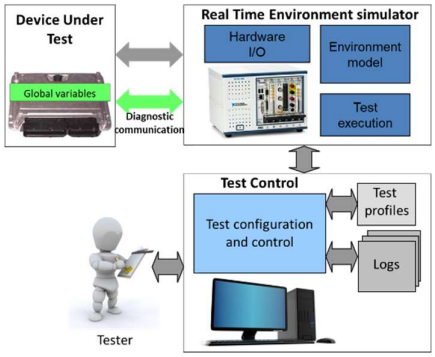


Fig. 1. Typical HIL test setup

HIL tests can reduce the time spent on real environment testing, because most of the failures can be detected earlier in the simulated environment. This can reduce the overall testing cost significantly, because real environment setups like prepared track with staff for autonomous robots, or real plant setups for industrial

controllers are very costly. HIL tests also have the advantage of repeatability, controllability and stability, which is usually not given in the real environment.

B. Challenges of detecting multitasking and real-time behaviour based errors

Cyber-physical systems using RTOS based software introduce new challenges for the testers. The failure model of such a real-time embedded system [4] can be divided into sequential and multitasking real-time behaviour based failures shown on Fig. 2.

Industrial experiences have shown that the test methods for detecting sequential failures are already elaborated. Traditional white box and HIL environment based tests catch these types of failures. The detection of multitasking and real-time failures is a much harder problem. Most of the traditional tests do not have sophisticated methods to do that. Therefore, the multitasking and real-time failure detection capabilities of test systems should be improved.

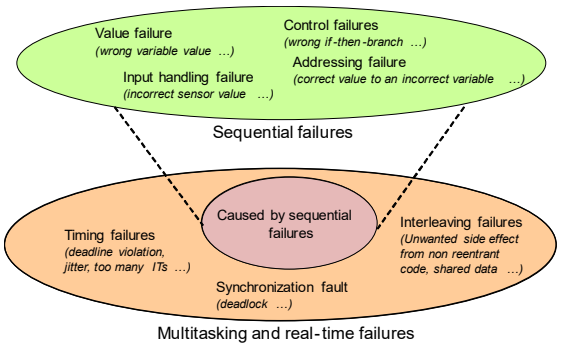


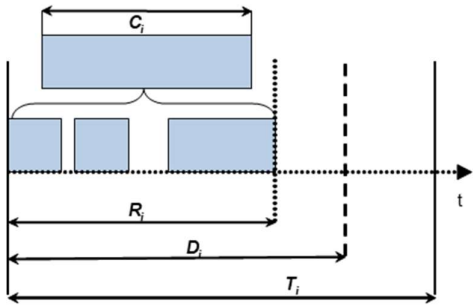
Fig. 2. Software failure models

There are three categories of multitasking and real-time failures: Timing failures (deadline violations of given tasks, unwanted jitter of periodically executed control functionality, too many and too long interrupts etc.), Synchronization failures (deadlock of tasks waiting for the same resource) and Interleaving failures (Unwanted side effect from non-reentrant code or shared data).

In our work we will focus on the Timing and Synchronization problems, because traditional tests leave these rather uncovered. Another reason to focus on these types of failures is that these can be measured well in HIL testing environments.

C. Methods of Timing and Synchronization failure detection

The most foundational thing of Timing and Synchronization failure detection is the static prediction or runtime measurement of software task WCETs (Worst Case Execution Time). These execution time predictions or measurements can be used to calculate the Worst Case Response time for each task, and therefore analyse whether the system will be able to behave in real-time in every situation by keeping the schedule of the tasks. Real-time systems out of their schedule can show symptoms like unwanted resets and strange behaviour for a short time. Fig. 3. introduce the notation and definitions of algorithms like Deadline Monotonic Analysis [5] used for Worst Case Response time calculation.



Ti is the period of task ***i***
Di is the deadline of task ***i***
Ci is the worst-case execution time of task ***i***
Ri is the worst-case response time to task ***i***

Fig. 3. Notations and definitions of task timing parameters

The static calculation of execution times and WCET is based on the pure source code of the tasks. The predictor calculates the task executional time for a given architecture by analysing the compiled assembly code without executing it, or by executing it in an emulator. Either method is used, the prediction of task execution times is a very complex problem, and many articles discuss its challenges [6].

Summary, there are tools for creating these static calculations, but their precision cannot reach the precisions of measurements made in HIL simulations or in real environments. Therefore, our paper focus on measurements made in HIL tests or real environments.

II. THEORETICAL BACKGROUND OF MEASURING TASK EXECUTION TIME

The value of ***Ti*** and ***Di*** used for Worst Case Response time calculation can be derived from requirements. While the value of ***Ci*** is need to be measured.

A. RTOS instrumentation based measurements

In some RTOS the built in instrumentation can provide the value of ***Ci***, but this requires a built in low level kernel measurement code. For example, FreeRTOS, which is currently the most widespread low weight preemptive multitasking embedded RTOS [1], [7] has the so called Trace hooks. These Trace hooks enables the user to perform measurements about the internal behaviour of the RTOS.

Trace hooks like traceTASK_SWITCHED_OUT (called when the RTOS switch out an old task), and

traceTASK_SWITCHED_IN (called when the RTOS switch in a new task) can be used to measure the ***Ci*** value of tasks.

There are professional solutions using these instrumentation capabilities: FreeRTOS+Trace is a run-time diagnostic and optimization tool provided by FreeRTOS's partner company Percepio. FreeRTOS+Trace captures valuable dynamic behaviour information, then presents the captured information in interconnected graphical views.

FreeRTOS+Trace is a very useful tool, but it is not integrated into a HIL testing environment. Therefore, its timing displays, are not synchronised with the HIL simulators stimulus signals and measurements, which could cause problems during testing.

RTOS instrumentation used by FreeRTOS+Trace also have the disadvantage, that for measuring the ***Ci*** values many data communication is required between the RTOS and the measuring software, which can cause significant overhead to the cyber-physical system.

RTOS instrumentation also have to deal somehow with the problems caused by interrupts. The interrupts can modify the measurements of ***Ci*** values unless the interrupt execution is not instrumented too, but instrumenting the interrupt execution can increase the overhead of the measurement significantly.

B. Non-intrusive measurements of task execution times

Most of the modern microcontrollers provide ways for high data rate non-intrusive data and instruction tracing. These tracing interfaces can be used to measure ***Ci*** values of tasks and interrupt timing too.

A typical example for such embedded trace support is the ARM's CoreSight trace architecture. The trace information in the CoreSight architecture is usually generated from three trace sources: Embedded Trace Macrocell (ETM), the ITM (Instrumentation Trace Macrocell), and the Data Watchpoint and Trace (DWT) blocks, shown on Fig. 4.

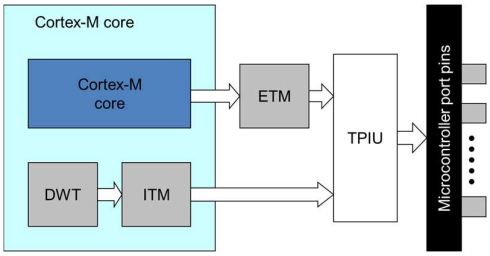


Fig. 4. ARM's CoreSight trace system in an ARM Cortex M core based microcontroller

The Trace Port Interface Unit (TPIU), formats the information form these sources into packets, and sends it to an external trace capture device.

The ITM has a capability to provide a "printf" style console message interface to the application software. The second purpose of the ITM block is to add timestamps to the DWT packets. The ETM block is used for providing instruction traces, therefore the whole software execution can be investigated. The drawback of using ETM is that it requires a huge amount of data transfer, and capturing and processing this amount of data is a challenging task [8].

From the task execution time (***Ci)*** measurement's point of view, the DWT block is the most interesting. The DWT has many of functionalities: among other things it can provide PC sampling at regular intervals and interrupt events trace. The DWT also includes several counters for measuring statistical parameters like interrupt overhead and sleep cycles. The DWT also has several comparators that can be used for data tracing on read or write accesses, or for triggering the ETM block. Summary the DWT outputs can be used for profiling and timing verifications.

The interrupt event trace function of DWT can be used to measure the interrupt execution with precise timestamps, so the interrupt overheads can be removed from the task execution times without software instrumentation. Statistics can also be created from interrupt occurrences and execution times, that can be used in schedule ability calculation.

The comparators of the DWT block also can provide useful help for execution time measurement. Every RTOS has a global variable, where the currently running task's property (usually a pointer to its task control block or TCB) is stored [7], [10]. By using the DWT's watch points, it is possible to follow every changes done to this variable, and the ITM block also can add timestamps to these events. This means that together with the interrupt tracing (every RTOS use software interrupts for task context switching) it is possible to measure the exact timing of tasks switching in and out with the information of which is the new task currently switched in. These tracing provides the same capabilities as software instrumented trace hooks, but without causing an overhead.

III. IMPLEMENTING THE NON-INTRUSIVE TASK EXECUTION TIME MEASUREMENT

To demonstrate the usability of the non-intrusive execution time measurement concept an experimental system was created. The unit under test is a Silicon Labs STK3700 development board containing an ARM Cortex-M3 based microcontroller, and running a FreeRTOS based multitasked demonstration software. The ARM Cortex-M microcontroller series is selected, because currently that is the most widespread 32-bit microcontroller core [1].

National Instruments LabVIEW development environment is used to perform the measurements. NI LabVIEW is selected because its popularity, and also because it can provide a straight forward integration to the NI-VeriStand HIL test development environment.

A. Interfacing to the trace port

The Trace Port Interface Unit (TPIU) supports two output modes, a synchronous double data rate clocked mode, using a parallel data output port with clock speed equals to the half of the system's core clock speed, and with a port width up to 4-bits (in the case of Cortex-M3 or M4 core microcontrollers) and a SWV (Serial Wire Viewer) mode, using a single-bit UART-like output.

The synchronous parallel interface has a high data rate, and it can be used for any trace measurements, including instruction trace. But, this high data rate port, which speed can easily reach the 100 Mbyte/sec requires a special FPGA based interface hardware. Such FPGA based hardware is costly and hard to integrate into a HIL environment [8].

The SWV's UART-like output is easily to interface directly through a serial port or through USB as virtual com port. The drawback of this solution is the possible message overflow because of the limited data rate. Usually the trace message outputs of the DWT blocks are short and not too frequent and therefore the DWT trace requires moderate or low data rate. An UART interface with its 1-2 Mbaud/sec usual maximum data rate can serve the average need of a normal DWT trace configuration. Unfortunately, the DWT message outputs are not to periodic, because event bursts can happen. Typical such event burst in our example is when an external interrupt happens, that cause a task activation. In this case the external interrupt sends at least 2 messages (entering and exiting), the task activation also done using a software interrupt (also 2 messages), and context switch will cause modification to the CurrentTCB global variable, which trigger another trace message. This means at least 5 trace messages with timestamps in a very short time interval (micro seconds range). However, these above trace messages are short messages, but such event burst can cause problems. The problem can be solved by integrated trace buffers specified in the CoreSight architecture, but the size of these buffers are implementation, and therefore microcontroller dependent. My measurement results show that the SWV interface can be used for RTOS aware non-intrusive testing, if the SWV data rate is higher than the 1/10 of the microcontroller's clock rate.

The implementation of the trace port interface is done using LabVIEW's OOP (Object Oriented Programing), which makes the implementation of trace interface extensible. Currently only a SWV based interface is created, but in the future the program is easily updateable with a synchronous high speed interface too.

B. Configuring the trace measurement

Configuring the trace measurement means the setting of the trace configuration registers [11]. This can be done inside the microcontroller, during the initialization phase, but that would cause software overhead and unnecessary trace output in every circumstances. More preferably this configuration also can be done through the debug port any time [3]. In this case the developers do not have to prepare the embedded software for the measurement, the test system will do this configuration.

What need to be set is the formatting and data port mode of the TPIU block (SWV wire output or parallel output). For RTOS timing measurement the DWT should be configured to have a comparator to the RTOS's Current Task Control Block global variable, and send messages if this variable is written by the microcontroller. To do so, the address of the Current Task Control Block global variable need to be known. Because this is a global variable, its address can be gathered from the .map or .elf file after building and linking the program.

The DWT block also needs to be configured to send trace messages on interrupt events.

Finally, the ITM block needs to be configured to add timestamps to the DWT messages. This timestamps can be local (relative time to the previous timestamp), or global timestamps. Global timestamps are more preferred for many reasons, but some microcontroller can only provide local timestamps.

C. Identifying the tasks of the system

The tasks running on the cyber-physical system can be directly specified by the operators before the test, or the test system can identify the software tasks during the test. For this identification the test system needs to know the type of the RTOS running on the cyber-physical system, and it also needs to know the current configuration setup of this RTOS.

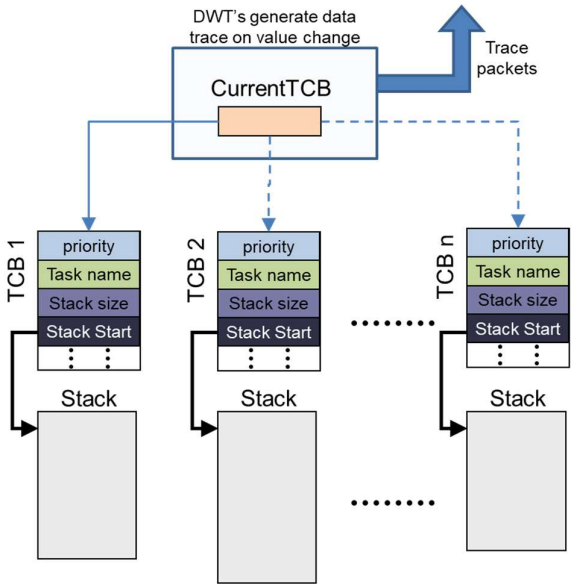


Fig. 5. Task identification through value change trace messages

Knowing the type of the RTOS and its configuration is important, because however the TCBs (Task Control Blocks) contain the same basic information regardless of RTOS type: task priority, task name, pointer to the task's

stack start, and last stacked item. But the structure of the TCB is RTOS dependent, and the TCB also can have configuration dependent values, like run time statistic counters, task tags for task identification etc.

After the test system knows the structure of the TCBs, then it can read the TCB information out though the debug port from the cyber-physical system for each task it switches on (Fig.5.).

D. Measurement results

Many measurements have been made to check the capabilities of the RTOS aware non-intrusive test system. The first measurement layer is simply an event activation graph with timestamps using LabVIEW's Digital Waveform Graphs. An example for this event activation graph is shown on Fig. 6. The figure shows an event timing capture of a sample system running two tasks: Low Priority Task (Low P. T.) and High Priority Task (High P.T.). The diagram presents the time interval, where the High Priority Task preempts the Lower Priority Task. It is also visible, that the Idle Task of FreeRTOS is running if both user task is waiting for an event. The interrupt overheads are also clearly visible: SysTick IT is the heart beat timer interrupt of the FreeRTOS system, and it runs in every 1ms, the PendSV IT (Pendable Service Call) is a software interrupt, and it is used by the FreeRTOS to perform task switching.

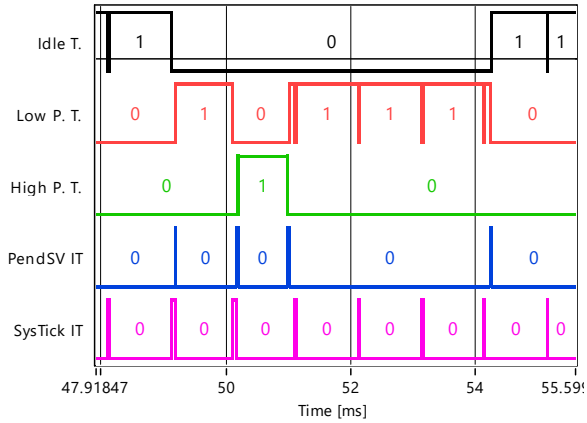


Fig. 6. Sample event activation view of a multitasked system

Based on the event activation measurements, higher abstraction level statistics and calculations can be created. From the Worst Case Response Time estimation's point of view, it is very important to have statistics for measured periods and execution times for every interrupts, and measured statistics of tasks execution times are also important (task periodicity should come from the specification). Fig. 7. shows an example of the measured parameters of the SysTick IT. SysTick IT is the heart beat timer of the FreeRTOS in this configuration, and therefore one of the most frequent interrupts. As the figure shows the periodicity of the SysTick IT is very stable as it can be expected from a high priority timer interrupt routine. The execution time results show a higher variability. The execution time depends on the functions needs to be performed during the interrupt. For example, Fig. 6. shows that the execution time increases, when the SysTick IT is executed before a task context switch. The reason for this, is that the SysTick IT notices that a task delay is elapsed, and then make that task ready to run.

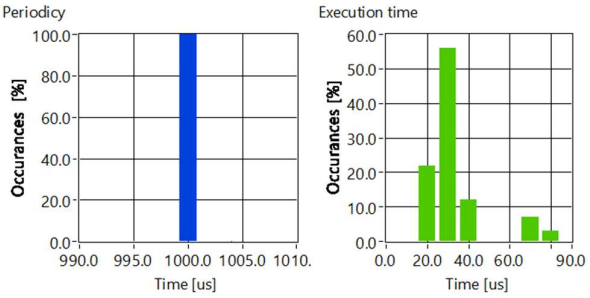


Fig. 7. SysTick IT periodicity and execution time statistics

Example for task execution time statistics is shown on Fig.8. These execution times show much higher variability, because the tasks usually perform much complex functions than interrupts, and their execution time highly depends on the actual control flow path.

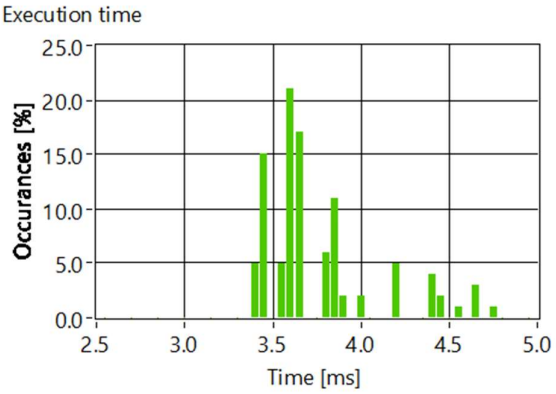


Fig. 8. Execution time statistics example of Low Priority Task

*E. Calculating the Worst Case Response Time (**Ri**) values*

Usually Deadline Monotonic Analysis (DMA) is used to calculate the worst-case response time of tasks (1). DMA is applying the following iterative formula, using the notations from Fig.3.:

$$R_i^0 = C_i$$

$$R_i^{n+1} = C_i + \sum_{\forall k \in hp(i)} \left\lceil \frac{R_i}{T_k} \right\rceil C_k \quad (1)$$

where

$$\sum_{\forall k \in hp(i)} \left\lceil \frac{R_i}{T_k} \right\rceil C_k$$

is the total interference from all higher-priority tasks, and **hp(i)** is the set of tasks with priority higher than **i**.

In this formula the worst case task execution times should be used as **Ci** and **Ck** parameters. This basic formula does not contain the IT effect compensation, but IT-s can be considered as very high level tasks, and in systems, where nested interrupt service (interrupts can preempt interrupts) is allowed their response time can be calculated in a same way.

Currently a basic calculation method using the maximum values from the execution time measurement statistics is used. But in the future it is possible, to implement a novel version replacing the single maximum values to statistical distributions from the task and interrupt timings.

IV. INTEGRATING THIS NOVEL MEASUREMENTS INTO A HIL TEST SYSTEM

The RTOS aware non-intrusive testing tool is written in LabVIEW; therefore, it is a straightforward decision to use NI VeriStand [9] as HIL environment form National Instruments.

To integrate the toolset into NI-VeriStand a so called Custom Device driver is needed. A NI VeriStand Custom Device functions as an interface to a special hardware. A Custom Device can have any number of input and output channels, and its functionality can be executed in every VeriStand engine cycle or can be implemented as a parallel executed task. In this integration the parallel executed task version is need, because the trace packet processing is an event based job. The Custom Device can provide event activation graph with timestamps for the test results display, and it also makes it possible to investigate the software timings in case of errors. The Custom Device also can provide Worst Case Response Time calculations to every tasks, which makes the signaling of Deadline violations possible.

V. SUMMARY AND CONCLUSIONS

This paper introduced the problem of multitasking and real-time failure detection in cyber-physical systems. The paper also described the ways of measuring RTOS task parameters important to detect this type of errors. A novel way of using non-intrusive tracing to measure these RTOS task parameters is introduced.

After the introduction of the concept the paper presented a LabVIEW based toolset, that performs these non-intrusive measurements. To demonstrate the operation of the toolset, measurement results using an ARM Cortex M microcontroller based test system running a FreeRTOS based multitasking software is presented.

These measurements are used to calculate and estimate the WCET for every tasks. With the estimated task WCETs the tool is able to determine the Worst Case Response Time of every tasks. This can ensure, that the test system notice Timing failures, like deadline violations of tasks, and interrupts.

The HIL test integration of these measurements also made it possible to provide information about the current state of the system's software in case of errors. For example, if the HIL system measure a non-correct response from the tested system, then our RTOS aware tool is able to show which tasks and interrupts have been executed and with what timings at a given time interval close to the error. Therefore, this tool can provide a great help for identifying the source and location of multitasking and real-time failures.

REFERENCES

[1] **ASPENCORE**: "2017 Embedded Market Study", *EETIMES, April 2017.*

[2] **A. Biagini, R. Conti, E. Galardi, L. Pugi, E. Quartieri, A. Rindi, S. Rossin.:** "Development of RT models for Model Based Control-Diagnostic and Virtual HazOp Analysis". *12th IMEKO TC10 Workshop on Technical Diagnostics. June 6-7, 2013, Florence, Italy.*

[3] **Balazs Scherer.**: "Hardware-in-the-loop test based non-intrusive diagnostics of cyber-physical systems using microcontroller debug ports" *ACTA IMEKO Volume 7 Issue 1* Pages 27-35.

[4] **H. Thane.**: "Monitoring, testing and debugging of distributed real-time systems" *In Doctoral Thesis, Royal Institute of Technology*, KTH, S100 44 Stockholm, Sweden, May 2000. Mechatronic Laboratory, Department of Machine Design.

[5] **Ken Tindell**.: "Deadline Monotonic Analysis", *Embedded System Progrming magazine*, June 2000.

[6] **R. Wilhelm, J. Engblom, A. Ermedahl, N. Holsti, S. Thesing, D. Whalley, G. Bernat, C. Ferdinand, R. Heckmann, T. Mitra, F. Mueller, I. Puaut, P. Puschner, J. Staschulat, P. Stenström**.: "The worst-case execution time problem - overview of methods and survey of tools" *ACM Transactions on Embedded Computing Systems*, Volume 7, Issue 3, April 2008.

[7] **Richard Barry.: "Mastering the FreeRTOS™ Real Time Kernel"** *Real Time Engineers Ltd.*

[8] **Balázs Scherer, Gábor Horváth.:** "Microcontroller tracing in Hardware In the Loop tests", *Proceedings of the 2014 15th International Carpathian Control Conference (ICCC)*, Velke Karlovice, Czech Republic, 2014.

[9] **National Instruments**, "NI VeriStand™ Fundamentals Coures Manual" *Part Number 325785A-01*, 2011.

[10] **Jean J. Labrosse,** "MicroC/OS-II: The Real Time Kernel" 2nd Edition, CRC Press, February 5, 2002.

[11] **ARM**, "ARM® v7-M Architecture Reference Manual", ARM DDI 0403E.b (ID120114), ARM Limited 2014.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Experimental analysis of IMU under vibration

D. Capriglione[2], M. Carratu[2], M. Catelani[1], L. Ciani[1], G. Patrizi[1], R. Singuaroli[1], P. Sommella[2]

[1]Dpt. of Information Engineering, University of Florence, via di S. Marta 3, 50139, Florence (Italy)
[2]Dpt. of Industrial Engineering, University of Salerno, via di G. Paolo II 143, 84084, Fisciano (Italy)

***Abstract* – MEMS-based Inertial Measurement Units are today widely employed in many contexts. Especially in the field of self-driving vehicles and navigation they provide important information to the electronic control units for implementing positioning, localization and tracking algorithms. As a consequence, it becomes important to analyse the accuracy, reliability and time to failure of such systems when operating in conditions as more as possible similar to ones experienceable in the practice. To these aims, in this paper we investigate on IMU performance under random vibration which can be thought of as a kind of stress to which IMUs are continuously interested during their common operating. The experimental results have proved that these devices are very sensitive to the considered conditions and that suitable measurement procedures and measurement setup should be designed for the IMUs performance analysis.**

Keywords – Testing, Reliability, Accelerometer, Gyroscope, Metrological Performance, Automotive.

I. INTRODUCTION

Today, Inertial Measurement Units (IMUs) are widespread in many application contexts. Cellular phones, cars, human motion, robotics, self-driving vehicles, navigation in transportation vehicles, military and aviation represent only a part of frameworks in which these kind of devices are more and more employed [1]-[5].

Consequently, due to their wide and even increasing use in several applications, the accuracy and reliability of such systems become fundamental for assuring the expected behaviour and the correct operating of all those systems based on IMUs.

Dependently on the complexity, costs, size and weight constraints of the specific application, IMUs could integrate triaxial accelerometers (for measuring the linear acceleration), triaxial gyroscopes (for measuring the angular rate) and triaxial magnetometers (for measuring the static magnetic field) or only a subset of them.

From a practical point of view, the common solutions available today on the market, are low cost systems based on Micro Electro-Mechanical Systems (MEMS) devices [6]. Thanks to their small size, these kind of devices are easily integrated in many systems and provide measurement information for algorithms of positioning, localization and tracking to cite a few [7]-[11].

Expected performance of such systems are provided in the related datasheets by their manufacturers which, generally, consider simplified operating conditions that are not well representative of actual way of operating of such devices. Indeed, typical information that can be found in datasheets deal with selectable ranges for measuring linear acceleration, angle rate and static magnetic field, as well as the related sensitivity (to each detected quantity) and the temperature operating range. Nevertheless, the dynamic metrological performance and how the actual operating conditions can affect the metrological performance and reliability of such systems is not adequately dealt with.

Recently, some analysis have been performed to verify the influence of temperature on the performance of accelerometers and gyroscopes embedded in IMUs [12],[13] , but currently, the literature of the field seems be still lacking of performance analyses of IMUs when operating in real scenarios characterized by the presence of significant vibrations. As an example, in automotive context, in navigation and industrial environments, IMUs are continuously interested by mechanical stresses as random vibrations, so it is expected that the effects of such vibrations could generally affect both the metrological performance, the reliability and the time to failure.

Starting from these considerations, in this paper, a suitable measurement setup has been designed for analysing the effects of random vibration on commercial and very popular IMU devices. To these aims, a suitable Printed Circuit Board (PCB) has been specifically designed and realized for hosting only the device under test (DUT), the electronic circuitry and connectors needed for powering the DUT and allowing digital data exchange with the required external Micro Controller Unit (MCU). By this way, it is expected that the experimental results will depend only on DUT performance and will not be affected by any auxiliary device (MCU, data logger, user monitor and so on) adopted in the test set-up. In particular, the DUT is an IMU including one triaxial accelerometer, one triaxial gyroscope and one triaxial magnetometer.

As for the vibration test profiles, since specific standards for testing IMUs are not yet available, they have been designed and carried out by taking into account the main international standards in force applicable to MEMS devices in automotive environment.

The paper is organized as follows: section II describes the adopted DUT and the realized electronic and user interface for managing the tests, section III describes the designed test profiles, section IV shows the experimental setup, section V the achieved measurement results, and finally conclusions are reported in section VI.

II. THE DEVICE UNDER TEST

In order to test the IMU, under conditions that will be reported in section III, a suitable experimental measurement setup has been developed. It is composed by the following main parts (see Fig.1): three commercial IMU adapter boards, three satellite boards (hereinafter *Module A*, *Module B* and *Module C*, respectively) hosting the IMU adapter boards, three STM32 Nucleo-64 boards and a Raspberry PI3 equipped with an LCD. The aim of the experimental setup is to acquire data coming from three identical IMUs equally stressed by a shaker.

More in detail, the inertial platform considered is typically adopted in several applications as indoor navigation, smart user interfaces, advanced gesture recognition and automotive; it includes a 3D digital linear acceleration sensor, a 3D digital angular rate sensor, and a 3D digital magnetic sensor. About the metrological performances, the LSM9DS1 has a linear acceleration full scale of ±2g/±4g/±8/±16 g, a magnetic field full scale of ±4/±8/±12/±16 gauss and an angular rate of ±245/±500/±2000 dps. In the experimental test bed considered, the LSM9DS1 has been configured to have a scale of 16g for the accelerometer, 2000 dps for the gyroscope and 16 gauss for the magnetometer, achieving respectively a sensitivity of 0.732 mg/LSB, 0.43 mgauss/LSB, and 70 mdps/LSB. To get data from the considered adapted board, different communication protocols can be used, but in this case the SPI serial standard interface has been used.

About the satellite boards, they have been specifically designed and developed to hold the adapter boards. More in detail, the three *Modules A, B* and *C* used for the experiments have been fixed on rigid support in order to be installed at the same time on a plane for the vibrations test (see section IV). The geometry of the satellite boards has been modeled to reduce the influence of the mechanical parts to the propagation of vibrations into sensors held on them, especially to avoid the excitation of resonance frequency able to destroy or disassemble the satellite boards with the respective IMU adapter boards. Particular attention has been devoted to the orientation of the satellite boards with respect to both the LSM9DS1 inertial sensor and the fixed support in order to align the sensor axes with the shaker.

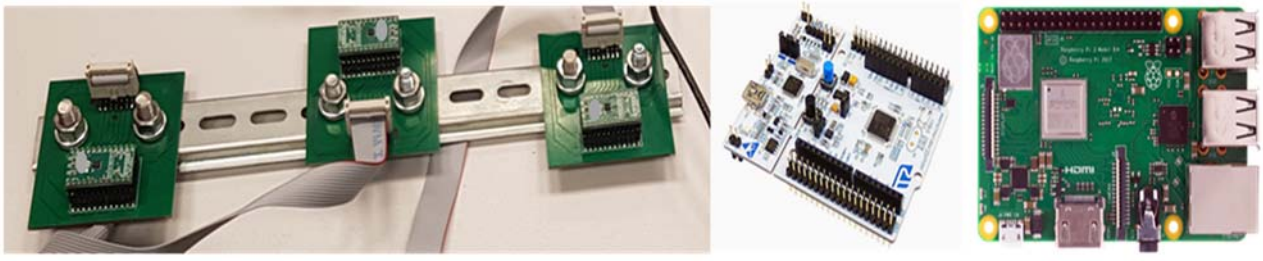
The STM32 Nucleo-64 boards [14] have been used to retrieve the data coming from the three *Modules A, B* and *C*, previously described, through a ribbon cable using an SPI communication configured to work at 10 kHz. More in detail, the STM32 Nucleo-64 board manages the configuration and communication over SPI with the inertial platform and sends the acquired data through USB to the Raspberry PI3. The firmware of the STM32 Nucleo-64 handles the data acquisition doing polling of the LSM9DS1 data-ready register. The data acquisition has been synchronized with the magnetometer ODR (Output Data Ready) since it represents the slowly peripheral among the gyroscope and accelerometer (80 Hz ODR).

The Raspberry PI3 [15] is a series of tiny single-board computers provided with all the features usually commons in a desktop computer; it has been useful in our case to acquire and store data from the inertial platform. The Raspberry PI3 is provided with a real-time operative system (Raspbian in our case) where suitable programming tools can be installed. For the aim, a Phyton console has been installed and used to contemporary store and acquire the data coming from the three Nucleo-64 boards. An SD card of 32 Gb has been used to enable the storing of a big amount of data according to the programmed test reported in section III.

A suitable LCD display with a touchscreen has been used to interact with the prototype and start the acquisition process. The data recorded during the experimentations can be extracted from the prototype thanks to variously available connections on the Raspberry PI3 as the USB port and WIFI/Ethernet port.

III. VIBRATION TEST FOR MEMS DEVICES

Both sensors and electronic devices inside automobiles and motorcycles are forced to endure extreme process and



a) ***b)*** ***c)***

Figure 1. The test bed experimented, from the left to right: a) Satellite Boards (Modules A, B and C), b) Nucleo Board, c) Raspberry PI3.

environmental conditions that can generate fatigue and fracture on the devices and affect their reliability performance [16]–[18].

For this reason, the environmental characterization of the MEMS Inertial Modules based on failure analysis and experimental tests is mandatory in order to prove the device validity for automotive fields. For the failure analysis, the work focuses on the mechanical physical domain. The typical failure mechanisms of MEMS sensors are [19]-[22]:

- *Fracture* caused by overload, fatigue, shock or stress corrosion;
- *Wear* due to adhesive or abrasive contact surface, corrosion or surface fatigue;
- *Creep and plastic deformation* caused by over vibration or intrinsic stresses;
- *Stiction* due to residual stress, chemical bonding, overload or electrical static force.

A common acceleration factor that influences all the MEMS failure mechanisms is the vibration. For this reason, a random vibration test profile was specifically designed for the device under test comparing several different vibration standards and reports: IEC 60068-2-64 (2012) [23], ISO 16750-3 (2003) [24], FIAT ENS0310 (2009) [25], MIL-STD-810G (2008) [26], ETSI EN 300 019-2-5 (2002) [27] and JESD22-B103B.01 (2016) [28]. The monitoring of critical components using experimental test also allow to achieve an improvement in term of system performances, reliability and availability [29]-[30].

The random vibration test is performed to verify that the MEMS Inertial Module will function in and withstand the vibration exposures of a life cycle in automotive application. This kind of test may be used to identify accumulated stress effects and the resulting mechanical weakness and degradation in the specified performance.

Gaussian random vibration has to be applied to the component's outer surface casing or leads in a manner to simulate classical motorcycle application or expected vibration during packaged shipment. The device case shall be rigidly fastened on the vibration platform and the leads adequately secured to avoid excessive lead resonance. The components will be mounted in such a manner so that they experience the full-specified vibration level at the component [28].

Generally, the random vibration test severity is described using the acceleration spectral density (ASD), which represents "the mean-square value of that part of an acceleration signal passed by a narrow-band filter of a centre frequency, per unit bandwidth, in the limit as the bandwidth approaches zero and the averaging time approaches infinity" [23]. The random profile developed for this application is reported in Fig. 2, where both frequency and ASD are reported in logarithmic scale.

At low-frequency, the test severity is described as follow:

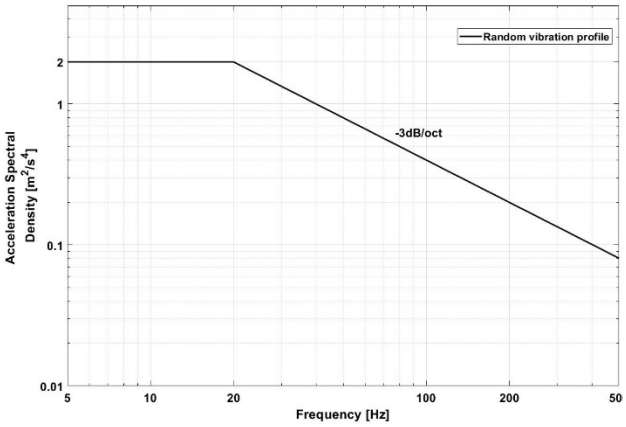


Fig.2. Random vibration profile

$$ASD = 2\frac{m^2}{s^3} = 0.02\frac{g^2}{Hz} \quad if \quad f \in [5Hz - 20Hz] \qquad (1)$$

Once overtaken the low-frequency range, the acceleration spectral density decreases as $-3dB/oct$ up to $500Hz$. The overall test level associated to the vibration profile is illustrated in Table 1. The key parameter is the Root Mean Square Acceleration G_{RMS} which represents the square root of the area under the ASD curve in the frequency domain. It is important to note that G_{RMS} contains no spectral information, therefore it cannot be used as the only constraint to define the test severity. G_{RMS} is useful in monitoring vibration tests since RMS can be monitored continuously, whereas measured spectra are available on a delayed, periodic basis.

The test will be performed for a duration of 30 minutes for each orthogonal axis, so 90 minutes in total to complete all the three axes. The ASD test level shall be applied within a tolerance $\pm 3dB$ of the nominal value at any frequency, allowing for the instrument and random error, referred to the specified ASD (see Fig. 2). The RMS acceleration levels shall not deviate more than $\pm 10\%$ of the nominal value included in Table 1. The given vibration profile was applied because it is a standard profile specifically tuned to reflect the operative condition of the device under test applied in the automotive field. Clearly, there are many other types of vibration test that could be used to characterize the performance of the IMU, like the sinusoidal profile or the step-test profile. The choice of the

Table 1. Overall measures of random vibration test level

Test Parameter	Test level
Profile Acceleration RMS	12.6197 m/s²
Profile Velocity RMS	0.0941 m/s
Profile Displacement RMS	1.8459 mm
Time for each axis	30 minutes

Table 2. Main Shaker parameters

Shaker Parameter	Max test level
Displacement Limit Peak	±25.5m
Maximum Velocity Peak	2 m/s
Maximum Acceleration Peak	2492 m/s²
Drive Frequency range	2Hz to 2000Hz

random vibration is basically due to the possibility to test the behavior of the system at different frequencies simultaneously, so other profiles are not considered in this work.

IV. EXPERIMENTAL SET-UP

The experimental tests were fulfilled at Analytical CETACE test laboratory using a Sentek M2232A shaker. The shaker main parameters are illustrated in Table 2.

Two identical *3056B2 General Purpose Accelerometer* by Dytron Instruments Inc. were used as input channel for the vibration controller. According to the manufacturer, the performances of the reference accelerometer are schematized in Table 3. The features of shaker and accelerometers allow to implement the vibration profile developed in the previous paragraph without restrictions on the severity level.

One accelerometer has to be located directly on the device under test, in such a way its output could be used as control input signal for the controller, while the other accelerometer has to be located on the shaker table. Using this configuration it is possible to control the amplitude of the vibration produced by the shaker allowing to apply the vibration profile directly on the DUT with high level of accuracy.

The fixing of the device to the shaker represents one of the most challenging steps of the vibration test, it must propagate the vibration equally to all the sections of the device under test without absorb it. Moreover, it must be safe and it must have the resonance mode out of the profile

Table 3. Accelerometer specifications

Accelerometer Parameter	Feature
Technology	Piezoelectric
Sensitivity, ±5%	100mV/g
Frequency range	1Hz to 10kHz
Electrical noise	0.0004grms
Linearity	±1% F.S.
Max vibration	±400g

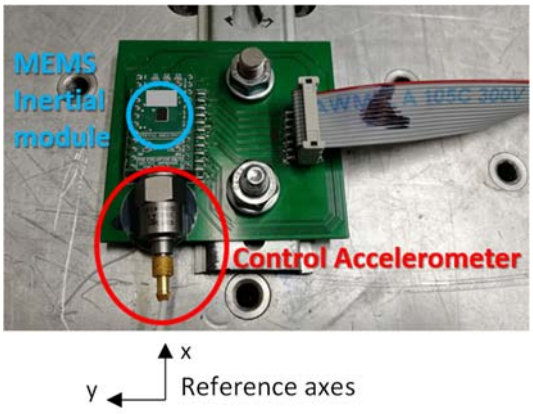


Fig.3. Experimental set-up

frequency range. Figure 3 shows the experimental set-up implemented for this test.

V. EXPERIMENTAL RESULTS

This section reports the results achieved through the experimental campaign described in the previous sections.

In particular, Figures 4 and 5 report the evolution of RMS values measured by the accelerometers and gyroscopes of *Module A* on both axes (*x* and *y*), when the shaker has actuated the vibration only along the *x*-axis (according to the reference system of Figure 3). Focusing the attention on Figures 4-5, we can identify three main zones: (i) Z_1 which corresponds to a "quite zone" in which no vibration is applied by the shaker, (ii) Z_2 which corresponds to the vibration applied according to the random profile described in section III, and (iii) Z_3 which corresponds to another "quite zone" achieved after the vibration has been stopped. As you can see, even if the vibration is actuated only on the *x*-axis, both accelerometer and gyroscope show a sensitivity also on the other axis (i.e. *y*-axis). Indeed, in Zone 2, the RMS values are significantly different from ones observed in Zone 1 and Zone 3 for both sensors axes. Similar trends have been observed for Modules *B* and *C*, here not reported for a sake of brevity.

Table 4 summarizes the achieved results for all

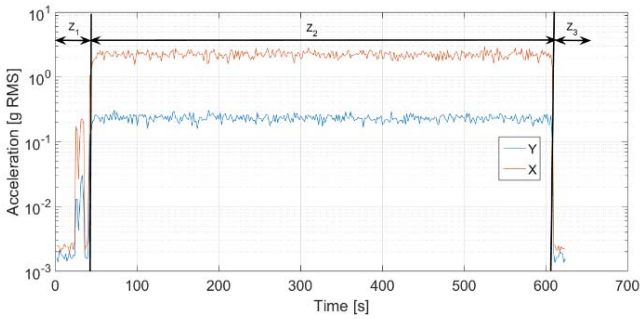


Fig.4. Module A accelerometer sensor output (the vibration is applied on the x-axis).

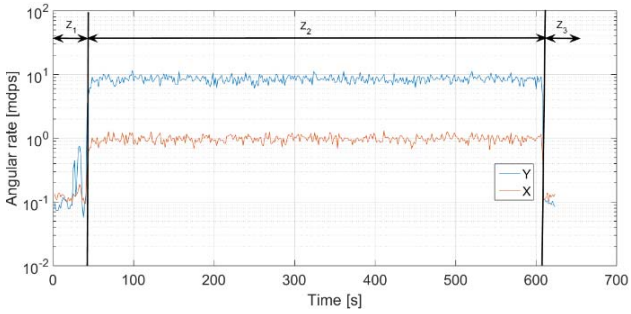


Fig.5. Module A gyroscope sensor output (the vibration is applied on the x-axis).

modules (*A*, *B* and *C*) and the two sensors axis by showing the means (μ) and standard deviations (σ) of RMS values. The following main considerations can be drawn:

- The results for Modules *B* and *C* confirm the same behaviours observed for module *A* (i.e. even if the vibration is actuated only on the *x*-axis, both Accelerometer and Gyroscope show a sensitivity also on the *y*-axis).
- As for the zones Z_1 and Z_3, whatever be either the sensor (Accelerometer or Gyroscope), the axis (*x* or *y*), and the Module (*A*, *B*, *C*) the achieved results show the experimental tests do not significantly affect the metrological performance of the considered systems. In particular, by comparing the mean RMS values (μ) observed in Z_1 and Z_3, the considered vibration tests have not lead to sensors losing calibration.
- For the zones Z_2, whatever be either the sensor (Accelerometer or Gyroscope) and the sensor axis considered (*x* or *y*), all modules show two-by-two compatible measurement results. Moreover, the RMS values measured by the sensors well-match the reference values imposed by the shaker.

VI. CONCLUSIONS

The experimental results achieved in this paper have confirmed how the vibration generally affects the operating of the considered IMUs. In particular, the application of the vibration along only one axis at a time (i.e. *x*-axis) has proved that both accelerometers and gyroscopes detect such vibration also on the unexpected axis (i.e. *y*-axis). For these reasons, measurement setup and measurement procedure for evaluating performance of such devices should take into account the need to analyse the rejection of undesired sensitivity on the not excited axis as well.

All these results also push toward the adoption of well-designed pre-processing techniques able to filter or compensate these behaviours mainly in applications where data fusions are adopted for object positioning and tracking purposes.

In addition, the experimental results have also highlighted how the considered test conditions have not significantly affected the general operating of such systems (in other words, a new calibration is not required after the test) and the stresses applied have not damaged the adopted devices.

Future developments will concern with the design of suitable test bed and measurement procedure for a deeper metrological performance analysis of IMUs devices with the aim of providing to final users of IMUs, important additional information about their behaviour under dynamic operating conditions.

REFERENCES

[1] J. Kim, N. D. Thang, and T. Kim, "3-d hand motion tracking and gesture recognition using a data glove," in 2009 IEEE International Symposium on Industrial Electronics, pp. 1013–1018, July 2009.

[2] G. Dissanayake, S. Sukkarieh, E. Nebot, H. Durrant-Whyte, "The aiding of a low-cost strapdown inertial measurement unit using vehicle model constraints for land vehicle applications" *IEEE Transactions on Robotics and Automation* Vol.17 , Issue: 5 , Oct 2001.

[3] K. Keunecke, G. Scholl, "Accurate Indoor Localization by Combining IEEE 802.11 g/n/ac WiFi-Systems with Strapdown Inertial Measurement Units", GeMiC 2014; German Microwave Conference.

[4] Z.Cao, S. Su, H. Chen, H. Tang, Y. Zhou, Z. Wang, "Pose measurement of Anterior Pelvic Plane based on inertial measurement unit in total hip replacement surgeries", 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

[5] S. Yosi, N. J. Agung, S. Unang, "Tilt and heading measurement using sensor fusion from inertial measurement unit", 2015 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC).

Table 4. Summary of sensors output for Modules A, B and C (the vibration is applied on the x-axis).

		Acc X [g RMS]			Gyro Y [mdps RMS]			Acc Y [g RMS]			Gyro X [mdps RMS]		
		Z_1	Z_2	Z_3	Z_1	Z_2	Z_3	Z_1	Z_2	Z_3	Z_1	Z_2	Z_3
A	**μ**	0.0024	2.2370	0.0022	0.091	8.492	0.095	0.0017	0.2329	0.0017	0.126	0.983	0.124
	σ	0.0002	0.2587	0.0002	0.017	1.025	0.007	0.0002	0.0267	0.0002	0.011	0.120	0.012
B	**μ**	0.0025	2.3456	0.0021	0.096	9.654	0.106	0.0023	0.2592	0.0019	0.106	1.054	0.136
	σ	0.0003	0.3652	0.0004	0.001	1.854	0.010	0.0003	0.0389	0.0003	0.012	0.253	0.014
C	**μ**	0.0031	2.5137	0.0030	0.079	11.179	0.114	0.0009	0.1975	0.0010	0.097	1.211	0.180
	σ	0.0004	0.4804	0.0005	0.007	2.1513	0.016	0.0001	0.0346	0.0001	0.009	0.224	0.016

[6] B. Alandry, L. Latorre, F. Mailly, P. Nouet, "A CMOS-MEMS Inertial Measurement Unit", SENSORS, 2010 IEEE.

[7] D. Capriglione, M. Carratu, P. Sommella, and A. Pietrosanto, "ANN-based IFD in motorcycle rear suspension," in 15th IMEKO TC10 Workshop on Technical Diagnostics 2017 - "Technical Diagnostics in Cyber-Physical Era", pp. 22–27, 2017.

[8] D. Capriglione, M. Carratù, A. Pietrosanto and P. Sommella, "Online Fault Detection of Rear Stroke Suspension Sensor in Motorcycle," in IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 5, pp. 1362-1372, May 2019. doi: 10.1109/TIM.2019.2905945

[9] M. Carratù, A. Pietrosanto, P. Sommella and V. Paciello, "Velocity prediction from acceleration measurements in motorcycle suspensions," 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Turin, 2017, pp. 1-6. doi: 10.1109/I2MTC.2017.7969943

[10] D. Capriglione, M. Carratù, A. Pietrosanto and P. Sommella, "NARX ANN-Based Instrument Fault Detection in Motorcycle," Measurement Journal, Volume 117, Pages 304-311, March 2018, doi: 10.1016/j.measurement.2017.12.026.

[11] D. Capriglione, M. Carratù, C. Liguori, V. Paciello, and P. Sommella, "A Soft Stroke Sensor for Motorcycle Rear Suspension," Measurement Journal, Vol. 106, pp. 46–52, doi: 10.1016/j.measurement.2017.04.011.

[12] S. Sabatelli, M. Galgani, L. Fanucci, and A. Rocchi, "A double-stage kalman filter for orientation tracking with an integrated processor in 9-d imu," IEEE Transactions on Instrumentation and Measurement, vol. 62, pp. 590–598, March 2013.

[13] D. Bereska, K. Daniec, W. Ilewicz, K. Jędrasiak, R. Koteras, A. Nawrat, M. Pacholczyk, "Influence of temperature on measurements of 3-axial accelerometers and gyroscopes: Embedded into inertial measurement unit", 2016 International Conference on Signals and Electronic Systems (ICSES).

[14] https://www.st.com/en/evaluation-tools/nucleo-f401re.html

[15] https://www.raspberrypi.org/products/raspberry-pi-3-model-b/

[16] Automotive Electronics Council, "Failure Mechanism Based Stress Test Qualification for Integrated Circuit." AEC - Q100 - Rev-H, 2014.

[17] Ahari, A. Viehl, O. Bringmann, and W. Rosenstiel, "Mission profile-based assessment of semiconductor technologies for automotive applications," *Microelectron. Reliab.*, vol. 91, pp. 129–138, Dec. 2018.

[18] K. Choi, D.-Y. Yu, S. Ahn, K.-H. Kim, J.-H. Bang, and Y.-H. Ko, "Joint reliability of various Pb-free solders under harsh vibration conditions for automotive electronics," *Microelectron. Reliab.*, vol. 86, pp. 66–71, Jul. 2018.

[19] W. Merlijn van Spengen, "MEMS reliability from a failure mechanisms perspective," *Microelectron. Reliab.*, vol. 43, no. 7, pp. 1049–1060, Jul. 2003.

[20] D. M. Tanner *et al.*, "MEMS Reliability : Infrastructure, Test structures, Experiments and Failure Modes," *Sandia Report*, no. January. Sandia National Laboratories, 2000.

[21] M. Tilli, T. Motooka, V.-M. Airaksinen, S. Franssila, M. Paulasto-Krockel, and V. Lindroos, *Handbook of Silicon Based MEMS Materials and Technologies*, Second. Elsevier, 2015.

[22] Y. Huang, A. Sai Sarathi Vasan, R. Doraiswami, M. Osterman, and M. Pecht, "MEMS Reliability Review," *IEEE Trans. Device Mater. Reliab.*, vol. 12, no. 2, pp. 482–493, Jun. 2012.

[23] IEC 60068-2-64, "Environmental testing - Part 2: Tests - Test Fh: Vibration, broadband random and guidance." International Electrotechnical Commission, 2012.

[24] ISO 16750-3, "Road vehicles - Environmental conditions and testing for electrical and electronic equipment - Part 3: Mechanical loads." International Organization for Standardization, 2003.

[25] FIAT Group, "Environmental Test Specification - Electronic Components," 2009.

[26] MIL-STD-810G, "Environmental Engineering Considerations and Laboratory Tests," no. October. US Department of Defense, Whasington DC, 2008.

[27] European Telecommunications Standards Institute, "Environmental Engineering (EE); Environmental conditions and environmental tests for telecommunications equipment; Part 2-5: Specification of environmental tests; Ground vehicle installations." ETSI EN 300 019-2-5, 2002.

[28] JEDEC Solid State Technology, "JEDEC STANDARD: Vibration, Variable Frequency." JESD22-B103B.01, 2016.

[29] A. Rossi et al., "A preliminary performance validation of a MEMS accelerometer for blade vibration monitoring," 22nd IMEKO TC4 Int. Symp. 20th Int. Work. ADC Model. Test. 2017 Support. World Dev. Through Electr. Electron. Meas., pp. 180–184, 2017.

[30] M. Catelani, L. Ciani, and A. Reatti, "Critical components test and reliability issues for photovoltaic inverter," in 20th IMEKO TC4 Symposium on Measurements of Electrical Quantities: Research on Electrical and Electronic Measurement for the Economic Upturn, Together with 18th TC4 International Workshop on ADC and DCA Modeling and Testing, IWADC 2014, 2014, pp. 592–596.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Magnetic Assisted Ball Burnishing of Magnetizable and Non-Magnetizable Materials

Zsolt Ferenc Kovács[1,3], Zsolt János Viharos[2,1], János Kodácsy[1], Roland Sándor[1]

[1]*John von Neumann University, GAMF Faculty of Engineering and Computer Science, Dep. of Vehicle Technology & Faculty of Economics and Business, Department of Economics and Law, Kecskemét H-6000 Hungary, kovacs.zsolt@gamf.uni-neumann.hu*
[2]*Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences (MTA SZTAKI), H-1111, Kende str. 13-17., Budapest, Hungary*
[3] *Department of Manufacturing Science and Technology, Budapest University of Technology and Economics, Budapest H-1111 Hungary*

***Abstract* – The goal of the reported research was to evaluate the machining conditions the magnetizable and non-magnetizable materials by the novel permanent Magnetic Assisted Ball Burnishing (MABB) tool. C45 steel, KO36 austenite steel, AA7075 aluminium and PA6 polymer materials were applied in the experiments. The main aim was to determine the optimal the technological parameters for these materials taking in consideration the hardness and roughness of the surface, too. Taguchi design of experiment methodology was applied in this study to simply look for optimal technology, compared to other kinds of technologies reported in various scientific papers. Surface quality is a complex feature that refers to the micro-geometrical characteristics of the machined surface. It includes roughness and waviness and gives a realistic picture about the top of the surface, while micro hardness and grains structure are especially important on sub-surface level. The results mirrored that all of the tested materials can be burnished by the MABB tool, however, the results from the economical viewpoints are diverse. The MABB tool was mainly designed to reduce the surface roughness but this cold metal forming process has further result like the surface hardening and further conclusions were also drawn that the MABB tool:**

- **capable to reduce the surface of non-magnetizable materials, too,**
- **increases the surface hardness of C45**

Keywords – magnetic assisted ball burnishing, technology optimization, material hardness, surface roughness.

I. INTRODUCTION

The MABB process is one of the cold-plastic finishing processes. It differs from other finishing solutions, such as hand scraping, lapping, grinding, etc., because it does not leave residual tension stresses on the machined surface. Furthermore, rolling is economically beneficial, because it is a simple and inexpensive process that requires short time and easy preparations. The introduced MABB process is unique, because the conventional burnishing is applied for finishing internal or external cylindrical surfaces, while the introduced MABB tool - thanks the special design - is suitable for flat or harmonically flat surfaces. The designed tool is shown in Fig. 1.

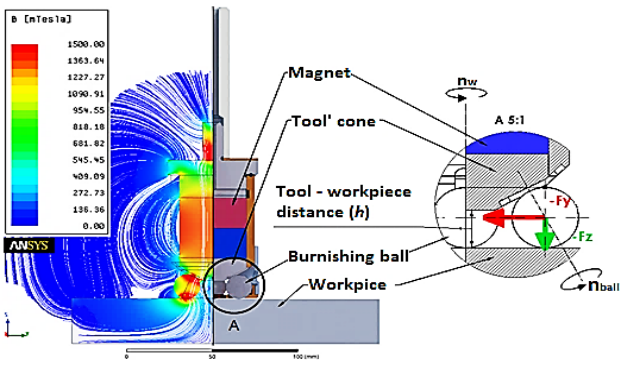


Fig. 1. Principle of MABB tool

The presented tool can be applied in conventional and CNC-controlled machines, too. During the analysis the tool is continuously cooled internally by Minimal Quantity Lubrication (MQL) oil. During this flat surface machining, while the tool moves on the planed path at the given feed rate, at the same time it rotates with a specific speed and as result it rolls down the surface.

In case of ferromagnetic materials, the required burnishing force is provided by the attractiveness of the balls, if the tool approaches the workpiece at a given *h* distance [1] (Fig. 1). However, this magnetic force cannot be established on non-magnetizable materials, because the magnetic attraction between the balls and the cone does

not allow the balls to rotate, so it results in the deterioration of the surface quality [2-3]. The reports also on such an experiment, using a magnetic chuck table to repel this phenomenon for the experiments.

II. RELATED RESULTS IN THE LITERATURE

There are many materials in our life, one of them are magnetizable and others are non-magnetizable, both types are important for the industry, however, in some cases the original state of the machined workpiece is it not suitable for the required usage. In such certain cases ball burnishing can change the material's roughness, hardness, corrosion and wearing resistance and decrease the incorporated stress values. In this paper, ball burnishing of flat C45 steel, KO36 austenite steel, AA7075 aluminium and PA6 polymer (Polyamide) are investigated.

A. C45 steel

The authors of this paper investigated in their former analysis the effects of ball burnishing on C45 steel applying the novel MABB tool [3]. But in this preliminary case the main aim was to determine the changes in ferrous materials' hardness and grain size. The C45 is a wildly popular structured steel, so there are several studies in this topic, but cylindrical workpieces were examined in all of them. E.g. Alberto Saldana-Robles at al. have explored that the burnishing force and feed have the main effects on the process (on surface roughness) in case of C45 steel as presented in Fig. 2. [4].

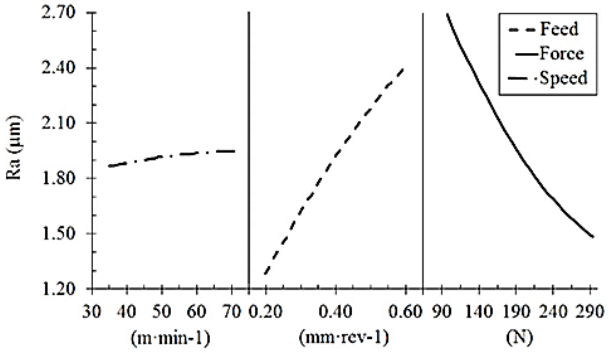


Fig. 2. Effect of technological parameters on roughness [4]

B. Stainless steels

Stainless steels are used in all areas of life, where parts exposed to hard environmental conditions and heavy loads, so it is important to have high strength and corrosion resistance. Lee at al. have studied the ball burnished AISI 316L stainless steel [5]. For the experiments, they used a 12 mm ball-ended tool after a milling process. They reported that the burnishing speed and the type of lubricant were found to affect the surface roughness most significantly, at a 99% level of confidence [5].

C. Aluminium

Aluminium and their alloys are widely used by the industry in large quantities because of low density and good mechanical properties. As other non-ferrous materials, the aluminium is also burnishable. Adel M. Hassan [6] and M.H. El-Axir et [7] al. explained the effects of ball burnishing of aluminium alloy. In both study AA2014 aluminium was applied, both of them burnished cylindrical workpiece, but Adel M. Hassan manufactured inner surface while M.H. El-Axir et al. machined the external surface., Their results are very similar, because they stated that the best results for average roughness is obtained when applying high depth of penetration. The reported also that the number of passes interacts with both burnishing speed and burnishing feed [6-8].

D. Plastics

Plastics and plastic-based raw materials play an increasingly important role in the industry. Lukasz Janczewski et al. (2016) have investigated the burnishing of PE500 polyethylene by diameter of 8 mm ball burnishing tool. Based on their results, the hardness of previously milled polyethylene after burnishing was increased only by 6%, while the wear was decreased by 58% [9]. Fig 3. reflects their results.

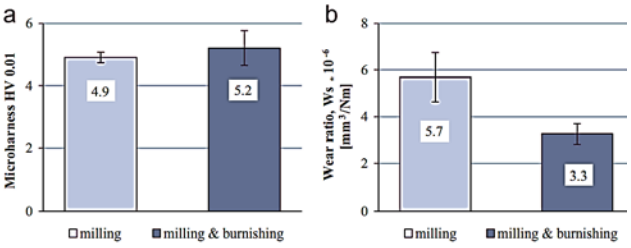


Fig. 3. a) Vickers hardness and b) wear rate for F=150 N and f=0,04 mm [9]

III. DESCRIPTION OF THE EXPERIMENTAL METHOD

The pre-machining process has also a very significant effect before any burnishing. It influences significantly the quality of the burnished's surface (e.g. accuracy and roughness), so analysing their effects is also a very important challenge. Different types of materials require different technological parameters for machining, in the given cases, the workpieces were pre-milled using technological parameters shown in Table 1-3.

Table 1. Technological parameters for C45.

No.	Feed-v_f [mm/min]	Cutting deep-a [mm]	Cutting speed-v_c [mm/min]
1	100	1	120
2	200	1	120
3	300	1	120

Table 2. Technological parameters for KO36.

No.	Feed-v_f [mm/min]	Cutting deep-a [mm]	Cutting speed-v_c [mm/min]
1	100	1	80
2	200	1	80
3	300	1	80

Table 3. Technological parameters for AA7075 and PA6.

No.	Feed-v_f [mm/min]	Cutting deep-a [mm]	Cutting speed-v_c [mm/min]
1	500	1	300
2	800	1	300
3	1100	1	300

For each workpiece, the feed rate was increased proportionally, indicating probably that the surface roughness values will also increase proportionally.

Because the KO36 austenite steel, AA7075 aluminium and PA6 polymer are non-magnetizable materials the necessary burnishing force cannot be generated. The solution was a magnetic table that was placed under the workpiece, as shown in Fig. 5.

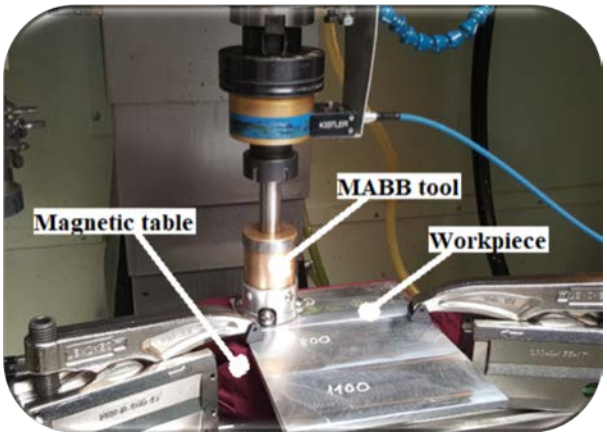


Fig. 5. Designation of experiments

Each specimen was cut to a dimension of 200×300×12 mm. For the addressed burnishing process experiments the standard Taguchi orthogonal array L9 (3^3) was designed which has three factors and three levels, the experimental results were analysed by MINITAB 17 software. The created Design of Experiments (DoE) table is shown in Table 4.

Table 4. DoE of ball burnishing.

No.	A	B	C
1	1	1	1
2	1	2	2
3	1	3	3
4	2	1	2
5	2	2	3
6	2	3	1
7	3	1	3
8	3	2	1
9	3	3	2

The factors in Table 4. are the same, expect the Ra roughness after milling, because there are four pre-milled different materials, see Table 5.

Table 5. Burnishing factors and levels for the Design of Experiments (DoE).

Factor		Level		
		1	2	3
A	Feed-v_f [mm/min]	10	30	50
B	Pre-milled surface roughness-Ra [µm]	See in Table 6.		
C	Burnishing speed-v_b [m/min]	20	40	60

For the analysis, n=36 experiments (n=4·9) were carried out, because of the 9 experiments per material and there were 4 materials available. The pre-milling process produced different Ra roughness, so, it must be handled by the levels of the B factor according to the tested materials, as represented in Table 6.

Table 6. Ra roughness values after milling.

Material	B-1 [µm]	B-2 [µm]	B-3 [µm]
C45	0,872	0,950	1,105
Ko36	1,455	2,553	1,249
AA7075	0,927	1,813	0,873
PA6	1,714	1,628	1,883

IV. RESULTS AND DISCUSSIONS

After the burnishing, the surfaces were evaluated by measuring the surface hardness, roughness and by microscopic pictures about the structure of the modified material layer. The following measuring equipment were used throughout the experimental work: for surface measurement MITUTOYO Formtracer SV-C3000, for hardness measurement Vickers microhardness tests have been performed with an optical microscope under a load of 100 g (Wilson-Wolpert 401 MVD microhardness $HV_{0,1}$ instrument), for microscopical evaluation a Zeiss Axio Imager.M2m light microscope and for SEM evaluation a Zeiss EVO MA10 SEM microscope was applied.

A. Surface roughness

After burnishing, the surface Ra roughness can be decreased even by 1/10 ratio compared to the original surface. This surface improvement is clearly visible in SEM images, see the milled surface in Fig. 6. and the burnished surface in Fig. 7.

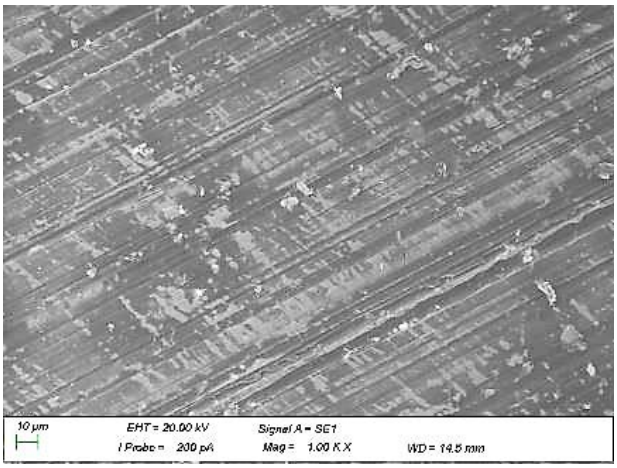


Fig. 6. SEM images of C45 surface after face milling, Mag=1k× (vf=200 mm/min, vc=120 m/min, ap=1 mm, Ra=1,105 µm)

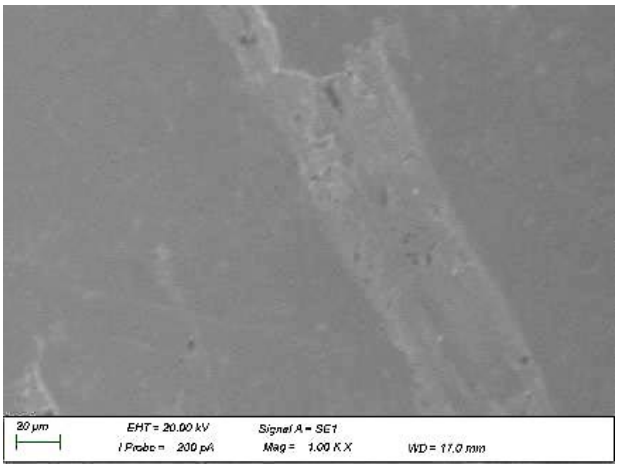


Fig. 7. SEM images of C45 surface after MABB burnishing, Mag=1k× (vf=200 mm/min, vb=120 m/min, pre-Ra=1,105 µm, burnished Ra=0,127 µm)

Figs. 8–11. show three-dimensional fitted curves by distance based interpolation method, as representation examples of the effects of various combinations of the selected ball burnishing parameters (burnishing speed, feed, pre-milled surface roughness) on the final Ra roughness of the C45, KO36, AA7075 and PA6 workpieces after burnishing by the novel MABB tool. It is worth mentioning that each curve represents the effects of two input parameters while the third (feed (v_f)) was kept at constant level where the resulted roughness was the smallest.: in case of C45 and PA6 polymer the v_f=10 mm/min while in case of AA7075 and KO36 the v_f=20 mm/min produced the lowest surface roughness.

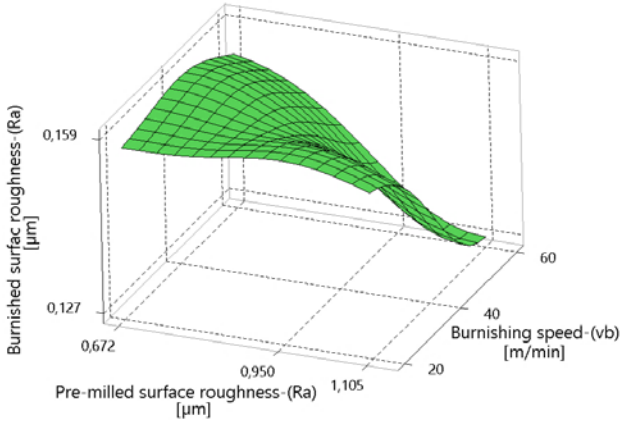


Fig. 8. Effect of burnishing speed and pre-milled surface roughness on surface average roughness of C45, vf=10 mm/min

A surprising result is shown in Fig 8. (material: C45), where the highest burnishing speed (v_b=60 m/min) and pre-milled roughness (Ra=1,105 µm) provided the lowest final surface roughness (Ra=0,127 µm) after burnishing. The suspicion is that the high pre-milling roughness provides relative high amount of material for plastic deformation. Furthermore, the high burnishing speed guarantees the high number of passes on the same material sub-surface.

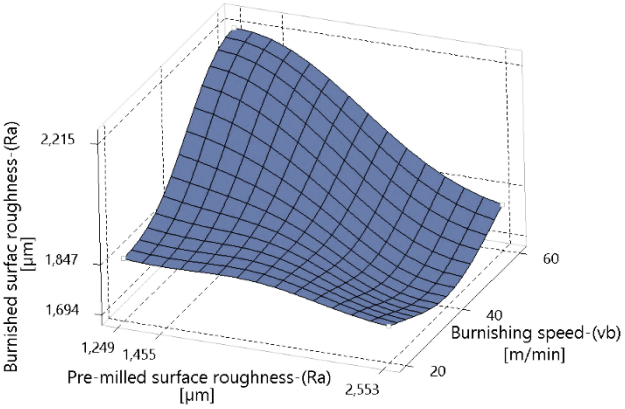


Fig. 9. Effect of burnishing speed and pre-milled surface roughness on surface average roughness of KO36, vf=20 mm/min

On the basis of the curve presented in Fig. 7. (KO36 austenite steel), it can be seen that the reduction in surface roughness is caused mainly by the pre-milled surface roughness, indicating that the burnishing speed has less effect.

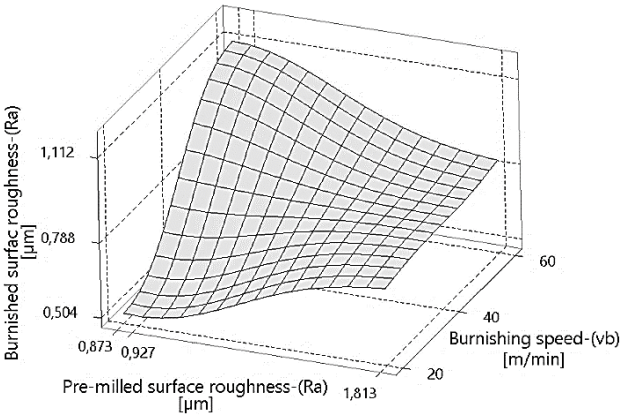


Fig. 10. Effect of burnishing speed and pre-milled surface roughness on surface average roughness of AA7075, vf=20 mm/min

According to Fig. 10. (AA7075 aluminium) it can be concluded that the initial low pre-milled surface produces low surface average roughness. The increase of the burnishing speed increases the surface average roughness, too. The plasticity of AA7075 should be the main reason of this phenomenon.

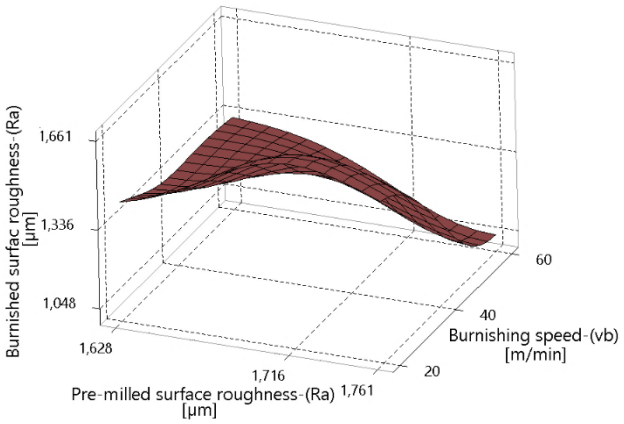


Fig. 11. Effect of burnishing speed and pre-milled surface roughness on surface average roughness of PA6 polymer, vf=10 mm/min

In Fig. 11. (PA6 polymer) the effects of the machining parameters are similar to the case of C45, so, the best results were obtained at the highest speed and pre-milling surface roughness.

B. Surface hardness

Figs. 12. and 14. show the resulted burnished surface hardness depending on the depth in the material and Figs. 13. and 15. show the microstructure images of C45 after burnishing with the novel MABB tool. The C45 surface material has reached the highest roughness reductions (from Ra=1,105 µm to Ra=0,127 µm), consequently, the highest material compaction was realized in this case. At all of the 9 tests the surface hardness of C45 steel get increased up to 220 $HV_{0,1}$ from 185 $HV_{0,1}$ (base material hardness without machining), see in Fig. 12.

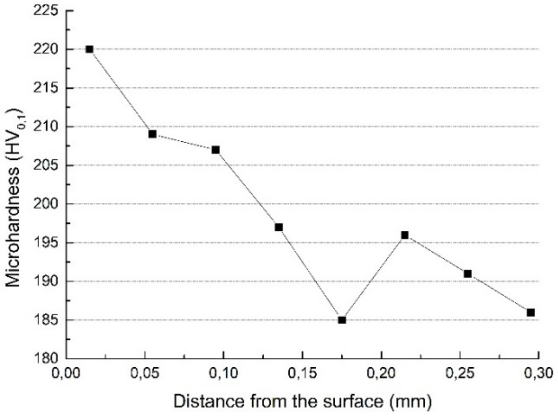


Fig. 12. Surface hardness of C45, No. 5 experiment

As Fig. 12. shows there is a valley in the curve between 0,15 and 0,25 mm depth from the surface, after it the material hardness reaches the base material hardness. This phenomena can be discernible at all of the hardness measurements. The rough grain layer which was generated between the upper fine grain layer by burnishing and the grain of the base material (Fig. 13. and 15.) could be the reason for that phenomena.

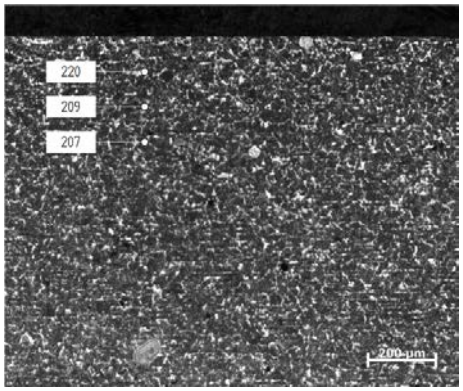


Fig. 13. Microstructure image of C45 surface and the hardness dispersion, No. 5 experiment

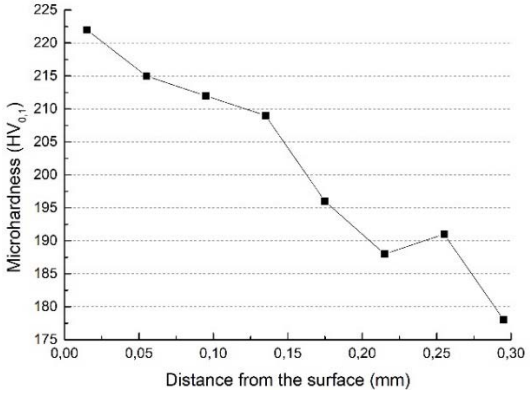


Fig. 14. Surface hardness of C45, No. 6 experiment

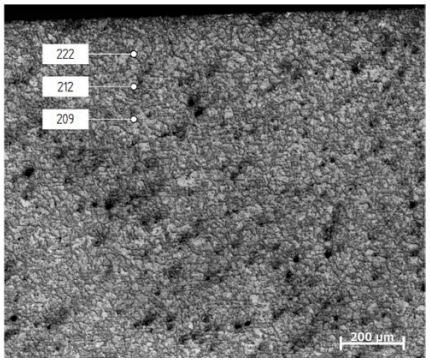


Fig. 15. Microstructure image of C45 surface and the hardness dispersion, No. 6 experiment

As the Fig. 12 and 14. shows the thickness of the hardening layer is about 0,1 mm. This value is more than the author's original expectation, similarly to the values of the hardness. In case of KO36 and PA6 after the burnishing the compaction of these materials were negligibly small, so it is not possible the evaluation them, because of the hardened layer thickness in comparison to the range of $HV_{0,1}$ microhardness measuring instrument. The measured Ra roughness of AA7075 mirrors that the decrease in it was similar to the decrease at C45, as shown in Fig. 16.

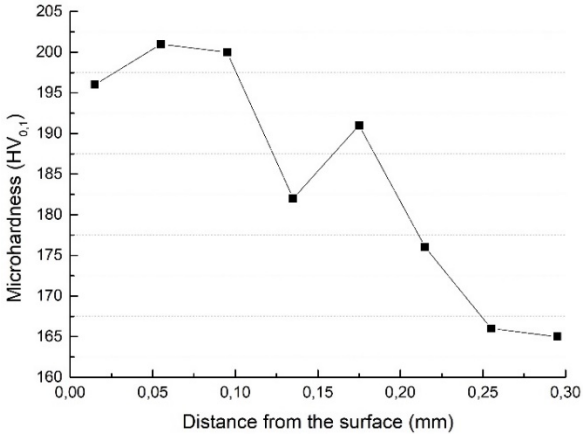


Fig. 16. Surface hardness of AA7075, No. 2 experiment

The surface hardness of AA7075 get increased up to 200 $HV_{0,1}$ from 165 $HV_{0,1}$ (base material hardness without machining).

V. CONCLUSIONS AND OUTLOOK

Analyses results on the ball burnishing technological parameters, on totally different behaving materials (steels, aluminium, PA6 polymer) and on the effects of the pre-milling surface roughness are reported in the given paper applying the novel Magnetic Assisted Ball Burnishing (MABB) tool for machining flat surfaces in which the balls rotate to generate high-speed and long-distance sliding under a constant burnishing force, generated by magnetic flux.

The conclusions of this work can be summarized as follows:

- The proposed MABB tool can burnish also non-magnetizable metals and other materials, supporting by a magnetic table.
- The best results for average roughness is obtained at low feed rate and if the pre-milled roughness value big enough to can be deformable.
- At AA7075 aluminium low average surface roughness can be reached with low technological parameters.
- The machining of KO36 austenite steel similarly to AA7075 aluminium requires low technological parameters.
- Burnishing of PA6 polymer has acceptable roughness reductions, but the process is very sensitive to the technological parameters.
- The hardness of the C45 steel after the proposed MABB machining can be upgraded by 20 percent.
- The hardness of the AA7075 aluminium also can be upgraded by 20 percent.
- The hardened layer thickness reached about 0,1 mm in case of C45 and A7075 machinings.

The results implicate further research and experiments. The PA6 polymer hardness measurement can be preformed, because the hardness measuring instrument is not suitable for polymer, but there exists specials hardness measurement instruments designed for polymers.

In case of KO36 austenite steel it is well known that the austenite microstructure is transformed into martensite by mechanical loads, it can be analysed, too. Another, short term plan is the measuring of the surface profile macro-geometrical change after burnishing, since currently, the effects of burnishing on the roughness are already known, but the effects for the surface topological accuracy are not yet explored.

VI. ACKNOWLEDGMENTS

This research is party supported by EFOP-3.6.1-16-2016-00006 "The development and enhancement of the research potential at John von Neumann University" project. The Project is supported by the Hungarian Government and co-financed by the European Social Fund.

The research in this paper was (partially) supported by the European Commission through the H2020 project EPIC (https://www.centre-epic.eu/) under grant No. 739592 and by the Hungarian ED_18-2-2018-0006 grant on an "Research on prime exploitation of the potential provided by the industrial digitalisation".

VII. REFERENCES

[1] **Kovács, Zs.F.; Viharos, Zs.J.; Kodácsy, J.**: Determination of the working gap and optimal machining parameters for magnetic assisted ball burnishing, *Measurement 118*, 2018, pp. 172-180.

[2] **Kovács, Zs.F.; Viharos, Zs.J.; Kodácsy, J.**: Determination of the optimal working gap for the magnetic assisted ball burnishing tool, *International Measurement Confederation (IMEKO)*, 2017, pp. 170-175.

[3] **Kovács, Zs.F.; Viharos, Zs.J.; Kodácsy, J.**: Making twist-free surfaces by magnetic assisted ball burnishing, *Solid State Phenomena 261*, 2017, pp. 159-166.

[4] **Saldaña-Roblesa A.; Plascencia-Morab H.; Aguilera-Gómezb E.; Saldaña-Roblesa A.; Marquez-Herreraa A.; Diosdado-De la Peña J.A.**: Influence of ball-burnishing on roughness, hardness and corrosion resistance of AISI 1045 steel, *Surface and Coatings Technology*, Vol. 339, 2018, pp. 191-198.

[5] **Lee, S.S.G.; Tam S.C.; Loh, N.H.**: Ball burnishing of 316L stainless steel, *Journal of Materials Processing Technology 37*, 1993, pp. 241-251.

[6] **HASSAN A.M.**: The effects of ball- and roller-burnishing on the surface roughness and hardness of some non-ferrous metals, *Journal of Materials Processing Technology 72*, 1997, pp. 385-391.

[7] **El-Axir, M.H.; Othmanb, O.M.; Abodiena A.M.**: Study on the inner surface finishing of aluminum alloy 2014 by ball burnishing process, *Journal of Materials Processing Technology 202*, 2008, pp. 435-442.

[8] **Kovács, Zs.F.; Viharos, Zs.J.; Kodácsy, J.**: The effects of machining strategies of magnetic assisted roller burnishing on the resulted surface structure, *Materials Science and Engineering 448*, 2018, pp. 1-8.

[9] **Janczewski, Ł.; Tobola, D.; Brostow, W.; Czechowski, K.; Lobland, H.E.H.; Kot, M.; Zagórski, K.**: Effects of bal lburnishing on surface properties of low density polyethylene, *Tribology International 93*, 2019, pp. 36-42.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Condition Monitoring of Wind Farm based on Wireless Mesh Network

Lorenzo Ciani, Alessandro Bartolini, Giulia Guidi, Gabriele Patrizi

Department of Information Engineering,
University of Florence,
via di S. Marta 3, 50139, Florence (Italy)
lorenzo.ciani@unifi.it

***Abstract* – Since the last years of twentieth century the use of wind energy rapidly increased and it is imposed as one of the best alternatives to burning fossil fuels. Nowadays, in renewable energy resources, wind energy is the leading candidate for electricity production due to its lower investment cost and well-developed technology in manufacturing high-power wind turbines. Basing on a Reliability Centred Maintenance (RCM) analysis, this paper focuses on the use of a Wireless Mesh Network in order to monitoring several condition parameters to identify possible incipient failures analysing the health-status of the turbine. The objective of the RCM is to identify the appropriate maintenance task for each item. Condition Monitoring is one of the possible task choices and represents a cost-effective solution to minimize the plant downtime. The case study identifies the most critical components of wind turbine and propose a set of possible sensors that could be included in the Wireless Mesh Network to monitor the condition parameters of the turbine.**

***Keywords* – Wireless Sensor Network, Wireless Mesh Network, Condition Monitoring, Wind Turbine, Maintenance**

I. INTRODUCTION

Nowadays, all the diagnostics process plays a fundamental role in industrial engineering representing an essential part of performance requirements. One of the easiest and widely used maintenance task is condition monitoring (CM): it is the process of monitoring one or more condition parameters in machinery to identify some changes that are indicative of an incipient fault or equipment health degradation [1]. In the past, condition monitoring was applied simply through routine manual diagnostic actions but, with the introduction of low-cost sensors and automated monitoring systems, online condition monitoring was adopted.

Condition monitoring is a type of condition-based maintenance used to select and survey parameters from the sensors placed in the system in order to detect a change in the health machine condition [2], [3]. CM, as well as reliability and availability, is mandatory in several manufacturing fields, such as energy and industry applications where products are forced to endure extreme process and environmental conditions [4]–[6].

II. RELIABILITY CENTERED MAINTENANCE

Reliability centred maintenance (RCM) is a method to identify and select failure management policies to efficiently and effectively achieve the required safety, availability and economy of operation. Failure management policies can include maintenance activities, operational changes, design modifications or other actions in order to mitigate the consequences of failure. RCM provides a decision process to identify applicable and effective preventive maintenance requirements, or management actions, for equipment in accordance with the safety, operational and economic consequences of identifiable failures, and the degradation mechanisms responsible for those failures.

The end result of working through the process is a judgement as to the necessity of performing a maintenance task, design change or other alternatives to effect improvements. The basic steps of an RCM programme are as follows [7]–[10]:

- Initiation and planning;
- Functional failure analysis;
- Task selection;
- Implementation;
- Continuous improvement.

Maximum benefit can be obtained from an RCM analysis if it is conducted at the design stage, so that feedback from the analysis can influence design. However, RCM is also worthwhile during the operation and maintenance phase to improve existing maintenance tasks, make necessary modifications or other alternatives.

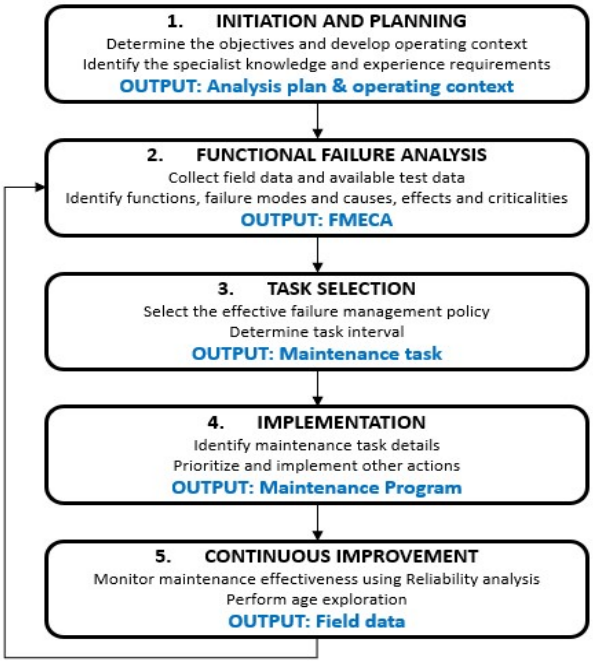


Fig. 1. Overview of the RCM process

According to [10], Fig. 1 presents a summary of the overall RCM process, divided into five steps. As can be seen from the figure, RCM provides a comprehensive programme that addresses not just the analysis process but also the preliminary and follow-on activities necessary to ensure that the RCM effort achieves the desired results[10].

The process starts with "Initiation & Planning" step, it is an important stage in order to analyse the operating context and to assess the optimal work order. A "Functional failure analysis" is crucial in the RCM process because understanding the effects of each failure mode leads to an optimization of the maintenance policies. The output of the second step is the Failure Mode, Effect and Criticality Analysis (FMECA), which is a systematic procedure for the analysis of a system to identify the potential failure modes, their causes and effects on system performance. FMECA is a method to identify the severity of potential failure modes and to provide an input to mitigating measures to reduce risk. In some applications it includes a means of ranking the severity of the failure modes (S), the frequency of occurrence (O) and the possibility of diagnose the failure mode before its effects are manifested on the system (D) to produce a metric called Risk Priority Number (RPN) [11]–[13].

By using the FMECA report it is possible to implement the third step ("task selection") in order to select the most appropriate and effective failure management policy. There are several different maintenance tasks that could be selected using an RCM decision diagram [10]:

- Condition monitoring: it is a continuous or periodic task to evaluate the condition of an item in operation against pre-set parameters in order to monitor its deterioration. It may consist of inspection tasks, which are an examination of an item against a specific standard.
- Scheduled restoration: it is the work necessary to return the item to a specific standard. Since restoration may vary from cleaning to the replacement of multiple parts, the scope of each assigned restoration task has to be specified.
- Scheduled replacement: it is the removal from service of an item at a specified life limit and replacement by an item meeting all the required performance standards.
- Failure-finding: it is a task to determine whether or not an item is able to fulfil its intended function. It is solely intended to reveal hidden failures. A failure-finding task may vary from a visual check to a quantitative evaluation against a specific performance standard. Some applications restrict the ability to conduct a complete functional test: in such cases, a partial functional test may be applicable.
- No preventive maintenance: in some situations no task is required, depending on the effect of failure. The result of this failure management policy is corrective maintenance or no maintenance at all. Sometimes, this approach is called "Run-To-Failure".
- Alternative actions: it can include redesign; modifications to existing equipment, such as more reliable components; maintenance procedure changes; pre-use or after-use checks; additional operator or maintainer training.

III. WIRELESS SENSOR NETWORK

Wireless Sensor Network (WSN) is a network characterized by a distributed architecture of small nodes (or motes). Each node can host multiple sensors and is equipped with computational and wireless communication capability. Recently, Wireless Mesh Networks (WMNs) have become an optimal solution to provide broadband internet access to large geographical areas [14]. The difference between a traditional WSN and Mesh network is [15], [16]:

- A traditional WSN (Fig. 2) is a point-to-multipoint (star) network where a single central node, known as the access point (AP), is directly connected to all other nodes known as stations. Traditional infrastructure Wi-Fi networks have the disadvantage of limited coverage area because every station must be in a range directly connected with the AP. Furthermore, overloading is another relevant drawback of traditional Wi-Fi networks because the maximum number of stations permitted in the network is limited by the capacity of the AP.

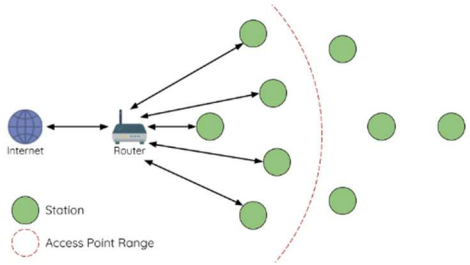


Fig. 2. Traditional Wireless Network

- A WMN (Fig. 3) is a self-organized and self-configured network. Nodes are mutually responsible for relaying each others transmissions. This allows MESH networks to have larger coverage area because nodes can still achieve interconnectivity, even if they are out of the central node cover range. Therefore, MESH network is also less susceptible to overloading. Another advantage of this network is the fault tolerance, when a node stops working, the whole network does not fail, but the access point can be reached by different paths [14].

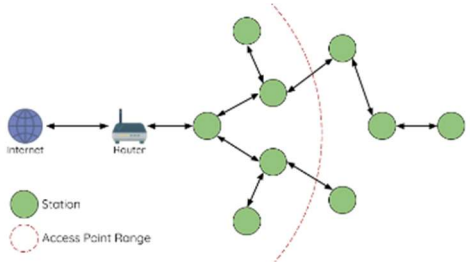


Fig. 3. Wireless Mesh Network

The block diagram of the typical sensor node is shown in Fig. 4. The picture highlights the main components that make up this system, as follow:

- *Sensing Unit:* It could include several different kind of transducer, such as thermal, vibration, magnetic sensor, etc. based on the physical quantities under analysis. If one or more sensors provide analogue signals, then they will be converted to digital form using ADC (Analog to Digital converter).
- *Processing Unit:* It process the sensors output using pre-defined instructions or programs loaded.
- *Communication Unit:* Once the elaboration of sensor information is completed, the data are relayed to the central Base Station using the communication unit. It consists of RF transceiver and antenna for transmitting and receiving packets. In order to optimize the energy efficiency of the sensor node, in addition to the traditional transmit and receive modes the communication unit could be set also as idle and sleep mode.
- *Power Unit:* It provides power to all the other units of sensor mote using one or more batteries.

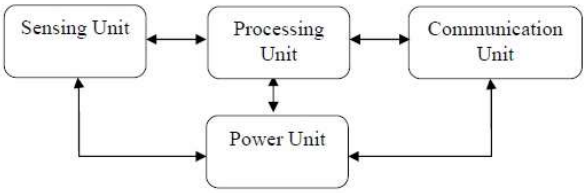


Fig. 4. Block Diagram of Sensor Node

WSN monitoring provides continuous and near real-time data acquisition and autonomous data acquisition. It can be used to [17]:

- maintain process tolerances;
- verify and protect machine, systems and process stability;
- detect maintenance requirements;
- minimize downtime;
- prevent failures saving Operation&Maintenance cost and time;
- request maintenance based on the prediction of failure rather than maintenance running to a standard schedule or being requested following an actual failure.

IV. WIND TURBINE COMPONENT FAILURES

A wind turbine (WT) is a device that converts the wind's kinetic energy into electrical energy, it has evolved from generating a few kilowatts in the 1980s to several megawatts today [18]–[20].

A wind farm, sometimes called wind park, is a group of wind turbines in the same location used to produce electricity. A large wind farm may consist of several hundred individual wind turbines and cover an extended area of hundreds of square kilometres, but the land between the turbines may be used for agricultural or other purposes [21].

WT is a very complex system composed by both mechanical and electrical/electronic components. The efficiency and practicality of horizontal-axis, three-blade turbines have resulted in domination of the market and, in almost all the new installations of wind farms, this is the only type that is employed for energy generation. This kind of wind turbine is comprised of four major components: Rotor, Tower, Blades and Nacelle, that is an enclosure which contains the electrical and mechanical components needed to produce electricity (e.g. gearbox, brake, yaw mechanism, generator, control system, etc.).

The functional failure analysis carried out on the wind turbine highlighted several different type of failure modes. Many causes can influence on electronic and mechanical components that make up the wind turbine; for instance, the failures of the mechanical items may be affected by the high vibration level on the shafts, the very extreme

temperature generated during the rotation speed conversion by the gearbox, corrosion and fatigue.

While the damage of the electrical and electronic components may be affected by the environmental conditions, such as the temperature, humidity and vibration.

V. CM NETWORK FOR WIND FARM

When the wind turbine is damaged, the maintenance costs and the economic losses caused by downtime are enormous, so repairs must be carried out quickly [22]. For this purpose, it is necessary to use a real-time monitoring system. In particular, thanks also to the greater capacity of the wind farm in the distribution network, an excellent solution is given by the mesh sensor networks [23].

After a preliminary Reliability Centred Maintenance analysis, the work focuses only on the blocks for which the procedure suggests to implement the Condition Monitoring as the optimal maintenance task. The CM is implemented using a Wireless Mesh Network, where each sensor node is provided by several different type of sensing element (Fig. 5). The use of a Wireless Sensor Network to monitor a wind farm is a powerful and low-cost solution that allows infrastructure modification and increasing the amount of sensor nodes after the plant installation compared to the higher cost of the changes in classical wired network.

Fig. 6 shows the architecture of the Wireless Mesh Network used to monitor the wind farm under test, where the blue lines stands for the direction of the wireless communications in standard configuration. In case a node fails, the other nodes have to identify a different communication path in order to ensure a complete monitoring of the farm. The LAN protocol used in the CM network is the IEEE 802.11n that represents a trade-off between geographical coverage, transmission of great amount of data and power dissipation [24].

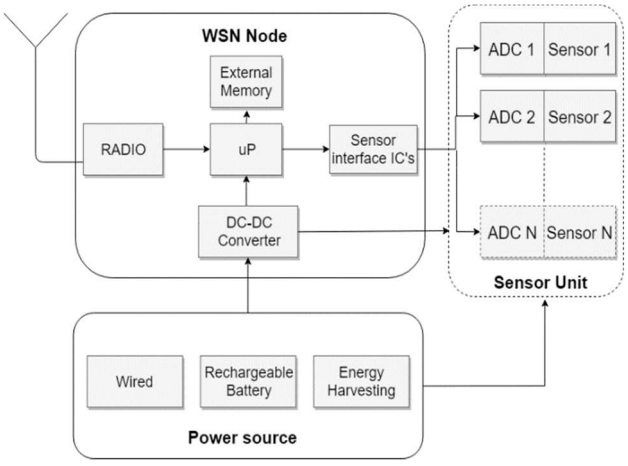


Fig. 5. Architecture of WMN node used for condition monitoring

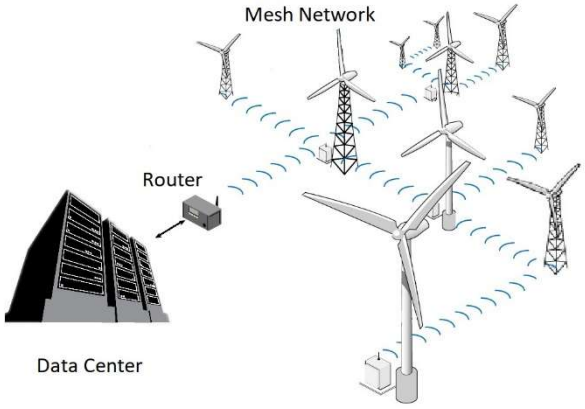


Fig. 6. Architecture of the network used to monitor the wind farm

Table 1 presents an extract of the CM assessment designed for the wind farm under test. The table contains:

- a description of the functionality of the items useful for the functional failure analysis;
- a list of the main failure modes (how we can observe the inability of the item to perform a required function) and mechanisms (the physical, chemical, or other processes that have led to a failure);
- The types of measurements used to monitor the item;
- The proposed sensor technology for the relative measurement.

In compliance to [1], [25] the electrical and electronic components are characterized by several failure modes, all due to electrical causes, that are mainly subjected to variation according to the environmental condition of the nacelle. For this reason, a Resistance Temperature Detector (RTD), a humidity sensor and a Piezoelectric accelerometer are used to monitor this kind of components.

The other types of considered items, like blades, gearbox and tower are subjected to mechanical failure mechanisms, such as fatigue, corrosion, sticking etc. that could be monitored using different types of sensing units [26], [27]. For instance, a large amount of heat is generated in the gearbox due to the friction between the gears, therefore it is advisable to use a thermocouple in spite of an RTD because the thermocouples have characterized by a larger measurement range. The acceleration, pitch, roll and yaw of the blades are monitored using a MEMS inertial module that integrates one triaxial accelerometer, one triaxial gyroscope and one triaxial magnetometer to provide a measure of the acceleration and of the angular position with a unique small and cheap device. The other vibration measurement (such as the vibration on the gearbox or on the electronic boards) are implemented using a single triaxial MEMS accelerometer that provides information only on the acceleration.

Table 1. Extract of Condition Monitoring assessment for Wind Turbine

Item	Description/Function	Failure modes and mechanisms	Measurement	Sensors
Blades	Power source of the wind turbine that catch the wind energy and convert it in mechanical motion.	.Fatigue failure .Lightning damage .Corrosion .Delamination	Crack/Fatigue detection	Acoustic Emission
			Wind speed/direction	Anemometer
			Acceleration, pitch, roll and yaw	Inertial module
			Microdeformation	Extensometer
Gearbox	Increases the rotor speed up to the appropriate generator speed using gears, pinions and shafts.	.Binding/Sticking .Excessive Wear .Fails to move .Mechanical damage	Vibration	MEMS accelerometer
			Temperature	Thermocouple
			Motion	Displacement transducer
			Stress	Strain gauge
Electrical/Electronic system	Contains all the electrical components in the turbine, such as power converter, PFC system, transformers, contactors etc.	.Parameter drifts .No output .Short/Open circuit . Error in data elaboration/storage	Temperature	Resistance temperature detector
			Humidity	Humidity sensor
			Vibration	MEMS accelerometer
Hydraulics system	Moves the nacelle towards the yaw axis (wind direction)	.Leakage .Contamination .Improper flow .Stuck valve	Pressure	Piezoelectric pressure transmitter
			Flow	Ultrasonic meter
			Level	Magnetic level meter
			Contamination	Contamination sensor
Tower	Supports the other parts and holds them off the ground.	.Fatigue failure .Corrosion .Fracture	Crack/Fatigue detection	Acoustic Emission
			Microdeformation	Extensometer
UPS	Provides emergency power when the input power source or mains power fails.	.Aging battery	Temperature	Resistance temperature detector
			Current	Hall effect integrated current sensor

The tower is monitored using an extensometer to detect the micro-deformation of the structure and an acoustic emission sensor which uses ultrasounds to investigate the fatigue of the construction.

Piezoelectric pressure transmitter, ultrasonic flow meter, magnetic level meter and contamination sensor are used to investigate the characteristics of the oil in the hydraulic system.

VI. CONCLUSION

Wireless sensor network and condition monitoring represent the optimal maintenance solution for extended wind farm because of the limited installation and management cost. The use of only human employee in turbine monitoring becomes a high-cost solution when the wind farm covers a very large area. Using CM network preventive maintenance tasks are minimized because the maintenance and management policies are addressed to the network. The operator performs corrective maintenance only in case the system identifies an intervention request. WMN provides a great amount of reliable and low-cost wearing data, allowing significant decreasing in the Operation&Maintenance cost and turbine downtime.

REFERENCES

[1] MIL-HDB-338B, "Electronic Reliability Design Handbook." US Department of Defense, Whasington DC, 1998.

[2] Y. Wang, Z. Zhu, H. Song, and K. Shi, "Wind turbine gearbox condition monitoring based on extreme gradient boosting," in *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, 2017, vol. 2017-Janua, pp. 6017–6023.

[3] K. Singh, H. Malik, and R. Sharma, "Condition monitoring of wind turbine gearbox using electrical signatures," in *2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS)*, 2017, vol. 2017-Janua, pp. 1–6.

[4] A. Reatti, M. K. Kazimierczuk, M. Catelani, and L. Ciani, "Monitoring and field data acquisition system for hybrid static concentrator plant," *Measurement*, vol. 98, pp. 384–392, Feb. 2017.

[5] M. Catelani, L. Ciani, and A. Reatti, "Critical components test and reliability issues for photovoltaic inverter," in *20th IMEKO TC4 Symposium on Measurements of Electrical Quantities: Research on Electrical and Electronic Measurement for the Economic Upturn, Together with 18th TC4 International Workshop on ADC and DCA Modeling and Testing, IWADC 2014*, 2014, pp. 592–596.

[6] L. Albertoni *et al.*, "Analysis and design of full-bridge Class-DE inverter at fixed duty cycle," in *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016, pp. 5609–5614.

[7] J. Moubray, *Reliability Centered Maintenance*, Second. Butterworth-Heinemann, 1997.
[8] N. B. Bloom, *Reliability Centered Maintenance: Implementation made simple*. McGraw-Hill Companies, 2006.
[9] NASA Headquarters, "Reliability Centered Maintenance guide for facilities and collateral equipment." 2000.
[10] IEC 60300-3-11, "Dependability management – Part 3-11 - Application guide – Reliability centred maintenance." International Electrotechnical Commission, 2009.
[11] M. Catelani, L. Ciani, and M. Venzi, "Failure modes, mechanisms and effect analysis on temperature redundant sensor stage," *Reliab. Eng. Syst. Saf.*, vol. 180, pp. 425–433, Dec. 2018.
[12] IEC 60812, "Failure modes and effects analysis (FMEA and FMECA)." International Electrotechnical Commission, 2018.
[13] Mil-Std-1629A, "Military standard procedures for performing a failure mode, effects and criticality analysis." US Department of Defense, Whasington DC, 1980.
[14] Karthika K.C, "Wireless mesh network: A survey," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2016, pp. 1966–1970.
[15] A. Hać, *Wireless Sensor Network Designs*. Chichester, UK: John Wiley & Sons, Ltd, 2003.
[16] Y. Zhang, L. Jijun, and H. Hu, *Wireless Mesh Networking: Architectures, Protocols and Standards*. Tailor & Francis Group, 2007.
[17] V. J. Hodge, S. O'Keefe, M. Weeks, and A. Moulds, "Wireless Sensor Networks for Condition Monitoring in the Railway Industry: A Survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1088–1106, Jun. 2015.
[18] I. Munteanu, A. I. Bratcu, N. A. Cutululis, and E. Ceanga, *Optimal Control of WInd ENergy Systems Towards a Global Approach*. Springer, 2008.
[19] T. Burton, D. Sharpe, N. Jenkins, and E. Bossanyi, *Wind Energy Handbook*. John Wiley & Sons, 2001.
[20] A. González and D. Galar, "Condition monitoring of wind turbine pitch controller: A maintenance approach," in *15th IMEKO TC10 Workshop on Technical Diagnostics: Technical Diagnostics in Cyber-Physical Era*, 2017, pp. 200–206.
[21] A. Hemami, *Wind Turbine technology*. Cengage Learning, 2012.
[22] Z. Fu, Y. Luo, C. Gu, F. Li, and Y. Yuan, "Reliability analysis of condition monitoring network of wind turbine blade based on wireless sensor networks," *IEEE Trans. Sustain. Energy*, vol. 3029, no. c, pp. 1–1, 2018.
[23] C. Na, Z. Haixiang, and D. Huizhu, "Dynamic Behavior of Integrated Wind Turbines during Fault Condition and Impact on Relay Settings of Distribution Network Feeders," in *2006 International Conference on Power System Technology*, 2006, pp. 1–8.
[24] J. Jansons and T. Dorins, "Analyzing IEEE 802.11n standard: outdoor performanace," in *2012 Second International Conference on Digital Information Processing and Communications (ICDIPC)*, 2012, pp. 26–30.
[25] IEC TR 62380, "Reliability data handbook – Universal model for reliability prediction of electronics components, PCBs and equipment," vol. 2004. International Electrotechnical Commission, 2005.
[26] NSWC-11, "Handbook of Reliability Prediction Procedures for Mechanical Equipment," no. May. Naval Surface Warfare Center Carderock Division, West Bethesda, Maryland, 2011.
[27] W. Denson, G. Chandler, W. Crowell, and R. Wanner, "NPRD-91 Nonelectronic Parts Reliability Data." Reliability Analysis Center, p. 632, 1991.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Evaluation of material structure changing after ultrasonic milling of aluminum foam by Computed Tomography (CT)

Janos Líska[1], Zsolt Ferenc Kovács[1], Ladislav Morovic[2], Ivan Buransky[2], Marcel Kuruc[2], Zsolt János Viharos[3], Michaela Kritikos[2]

[1]*John von Neumann University, GAMF Faculty of Engineering and Computer Science, Dep. of Vehicle Technology, Kecskemét H-6000 Hungary, liska.janos@gamf.uni-neumann.hu; kovacs.zsolt@gamf.uni-neumann.hu*
[2]*Slovak University of Technology in Bratislava, Faculty of Materials Science and Technology in Trnava, ladislav.morovic@stuba.sk; ivan.buransky@stuba.sk; kuruc.marcel@stuba.sk; michaela.kritikos@stuba.sk*
[3]*Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences (MTA SZTAKI), H-1111, Kende str. 13-17., Budapest, Hungary, viharos.zsolt@sztaki.mta.hu*

***Abstract* – In our everyday lives, we spend more and more time in our vehicles, where we are exposed to growing accident hazards. That is why we are increasingly focusing on security. This is particularly true for electric cars, where the size of the power source and the firewall are not the same as for conventional vehicles. The so-called energy-absorbing zone has a key role during an accident. In the era of modern materials the aluminum foam can works as an energy-absorbing structure. However, the machining of this material is hard at this time. This article attempts to investigate, through Computed Tomography-CT technology, whether the material has a permanent deformation or dimensional variation in the structure of the material can be achieved by the ultrasonic milling process.**

Keywords: Ultrasonic machining, Computer Tomography (CT), Industrial CT, 3D Measurement, Aluminum foam.

I. INTRODUCTION

Recently, several scientific journals dealt with the problems of metal foam machining. Metal foams belong to the group of so-called cellular materials. Currently can create cellular materials from several materials, such as polymers, ceramics, glasses, composites, aluminium, copper, metals or these mixture [1]. Recently, metalic foam has been considered for use as impact energy-absorption material for cars, trains and airplanes to take advantage of its lightweight and unique compression deformation characteristics [2].

In this study a special type of foam, namely the aluminium alloy foam was investigated. This material has many advantages thanks to the cellular structure. For example, the heat-transfer and impact energy-absorption properties are excellent [1, 3, 4]. But about this good properties the machinability is bad. Haipeng Qiao at al. are investigated the subsurface damage of cells. Based on their results the depth of deformation can reach the 5 mm, see in Fig. 1. [5].

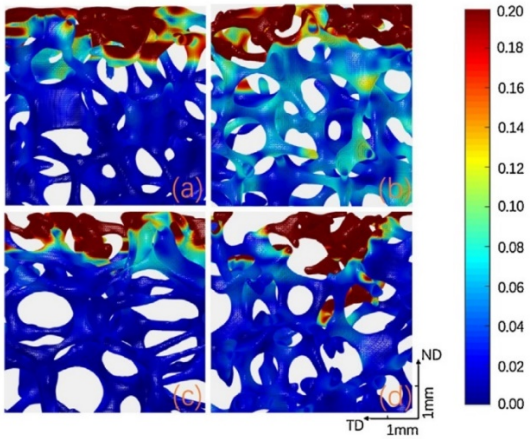


Fig. 1. Subsurface strain field from DVC of samples machined at vc = 2.6 m/s, feed rates of (a and c) vf = 0.102 mm/min and (b and d) vf = 0.204 mm/min, axial depths of cut of (a and b) ap = 1.5 mm and (c and d) ap = 1.0 mm [5]

The problem of the machinability is that, while a simple aluminium alloy can be machinable with a standard carbide tool, of course which is specially

designed for aluminium, but this tool insert not fit for aluminium foam because of the cellular structure and special alloys contents are require different carbide type and also tool geometry.

This also means problem for the tool manufacturer companies because they need to develop special tools which not only suitable for a particular material type, material properties (e.g. tools for Al machining).

To create this type of special tool we have determine the machinability of aluminium foam by various experiments. However, there are not many methods that can be used, because mostly the aluminium foams contains Al_2O_3 grains and made by air bubbles technology, this means that the workpiece structure is not compact and hollow. It is practically a cellular light structure where the thickness of the walls between the cells is very small (depending on the size of the cells). In this regard, it is hard to machining because there are several accompanying phenomena at the same time. Mainly the very intensive tool wear, negative burr formation, edge burr and damaged cells occurs the problems. These accompanying phenomena also destroy material surface, size and accuracy.

II. RELATED RESULTS IN THE LITERATURE

In order to better understand the causes of the phenomena, we believe that it is necessary to get to know the process of chip formation. In our opinion, ultrasonic high-speed machining technology can be used to minimize, possibly eliminate, the edge burr defect, tribological behaviour, reduce machining force and the subsequent cell destruction [6-8].

For the experiment an ultrasonic vibration assisted milling machine was used (DMG Sauer Ultrasonic 20 linear). It is a five-axis rotary ultrasonic and high-speed cutting machine tool. The ultrasonic tool vibration is generated by a piezo crystal (Fig. 2.), which can vibrate in range 20 to 30 kHz (its value depends on the tool diameter and length of acoustic assembly). High speed cutting is ensured by spindle, which can reach up to 42,000 rpm. Feed rate acceleration is over 2g and its value can reach value 40,000 mm/min.

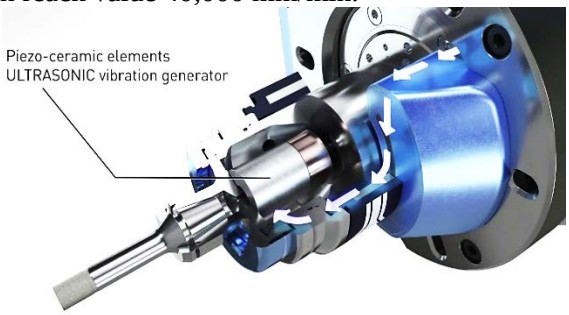


Fig. 2. Principle of ultrasonic milling [9]

This machine tool has a 5-axis gantry construction with an integrated NC swivel rotary table; therefore, it can machine complex shapes in one clamping. Its kinematic scheme is shown in the Fig. 3. The tool performs all three linear movements (X, Y, Z). Range of motions in these axes are: 200 mm in the X axis, 200 mm in the Y axis and 280 mm in the Z axis. Workpiece performs two remaining rotational movements (A, C). Rotation around X axis (A) is ensured by the cradle construction. Movement around this axis is limited by range -10° to +130°. Rotation around Z axis (C) is ensured by the rotary table. Movement around this axis is not limited. Velocity of rotational motions could reach the value up to 200 rpm. Performance of the machine is 15 kW. Torque momentum can achieve the value 6 Nm. It has high precision of positioning (± 2.5 μm). The cooling system can supply the process liquid or the pressured air by four outer nozzles and by the core of the tool. The tool magazine has 24 slots for different types of tools. In front of the tool magazine is also the laser measuring device for measuring the tool characteristics (like length and diameter). This machine tool operates under Sinumerik 840 solutionline control system [10, 11].

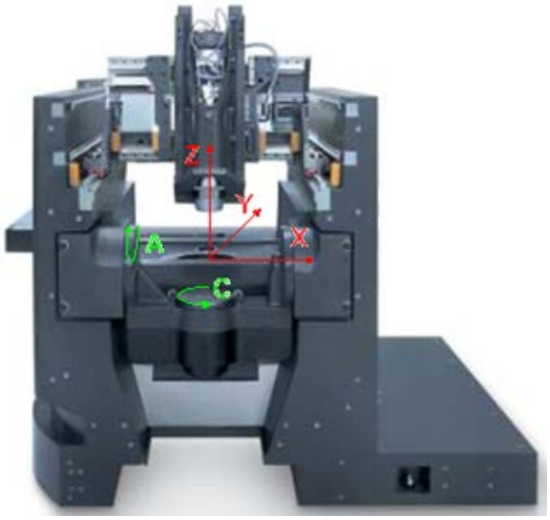


Fig. 3. Kinematic scheme of the controlled axes for DMG Ultrasonic 20 linear [12]

Of course there are other solution for aluminium foam machining for example the chemical milling. This process can produces smooth surfaces without damage but the machining time requirement so high [13, 14].

III. DESCRIPTION OF THE METHOD

As any machining process an experiment is performed, it is necessary to evaluate the data obtained with a design method or software. In this article the authors use the

Taguchi method to evaluate their results.

During of the experiment the parameter of cutting speed, cutting depth and feed per tooth were changed, so were selected as three experimental factors (parameters) and designated as factors A–C (Table 1).

Table 1. Milling conditions

	Parameters	Experimental conditions 1	2	3
A	Cutting speed, v_c (m/min)	400	500	600
B	Cutting depth, a_p (mm)	2	3	5
C	Feed per tooth, f_z (mm/tooth)	0,02	0,06	0,1

Three levels for each factor were configured to cover the range of interest and identified by the digits 1, 2, and 3 respectively. The L9 orthogonal array was selected to conduct the matrix (Table 2).

Table 2. Design matrix of the experiment

Exp. No.	Parameter combinations A	B	C
1	400	1	0,02
2	400	3	0,06
3	400	5	0,10
4	500	3	0,02
5	500	5	0,06
6	500	1	0,10
7	600	5	0,02
8	600	1	0,06
9	600	3	0,10

IV. RESULTS AND DISCUSSIONS

Due to the structure of the aluminium foam the extent of structural changes is not easy because the degree of lesion is not linear in each cross section of the material.

In this respect, before and after the ultrasonic milling industrial Computed Tomography (CT) technology is used. The measuring equipment is an industrial CT, Zeiss Metrotom 1500 (Fig. 4). Specification of the used equipment is in the Table 1.

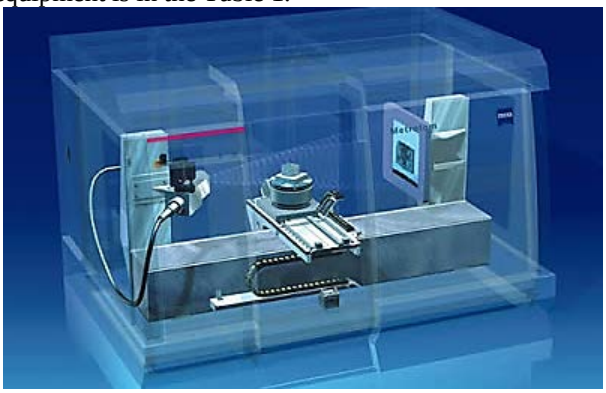


Fig. 4. Zeiss Metrotom 1500 industrial CT [15]

With this equipment it is also possible to examine the internal structure of the material. Our assumption is that, even during cutting directly below the machining zone the cells suffer deformation in the layers [16-18]. Knowledge of this can be important information for future use (e.g. formation of energy-absorbing zones).

Table 3. METROTOM 1500 specification

Tube power (W)	500
Tube voltage (kV)	225
Measuring range (mm)	dia. 350x300
Detector resolution (Pixel)	1024x1024; 2048x2048
Source/detector distance (mm)	1500
Tube type	Open X-ray

This experimental investigation was examined at a distance of 360 mm from X-Ray lamp with voxel size of 98 µm. Software used for scanning was METROTOM OS 2.8 (Fig. 5).

Integration time was 1000 ms and it was used 1050 pictures for evaluation. It took 45 minutes of scanning. Voltage value was used of 190 kV and current was used 600 µA.

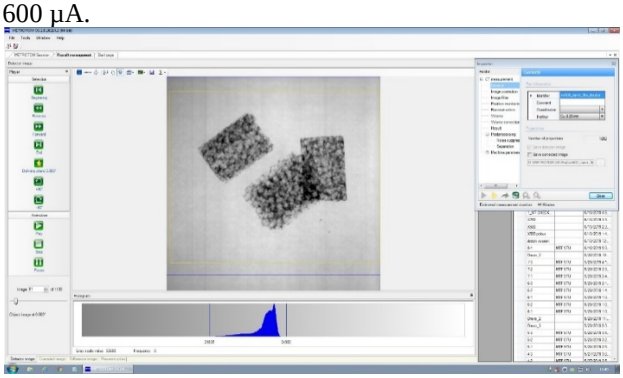


Fig. 5. METROTOM OS 2.8 interface

As it is seen in the figure above 3 pieces were scanned in one measurement. After they were separated and evaluated.

Before and after ultrasonic milling the specimens was 3D digitized by CT. This made it possible to compare the digitized 3D models, which was realized by software GOM ATOS Professional software.

The comparison aimed to determine the change in the geometry of the aluminium foam specimen, which we assume due to the cutting forces and possibly by the specimen clamping. At the beginning of the comparison process, the correct and accurate alignment of both samples is very important. The above mentioned comparison was evaluated respectively visualized by colour deviation map of shapes (and dimensions) (Fig. 6.) as well as by inspection sections in defined locations.

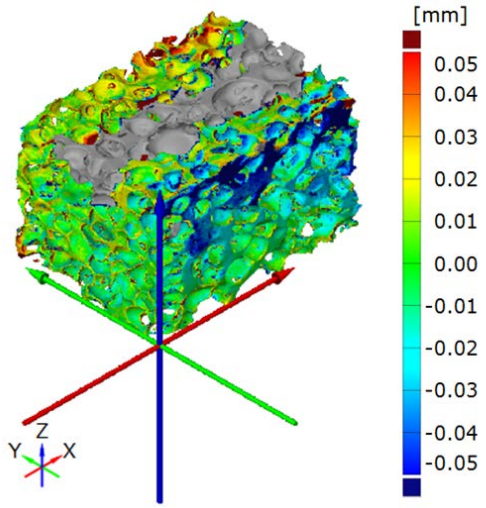


Fig. 6. Colour deviation map of shapes (and dimensions)

For the quantification of deviation evaluation of shapes and dimensions, inspection sections were used. For each comparison, 3 inspection sections were created: 2 inspection sections perpendicular to the X-axis at a distance of 2 mm from the side walls of the specimen and 1 inspection section at the centerline of the specimen along the slot (Fig. 7.).

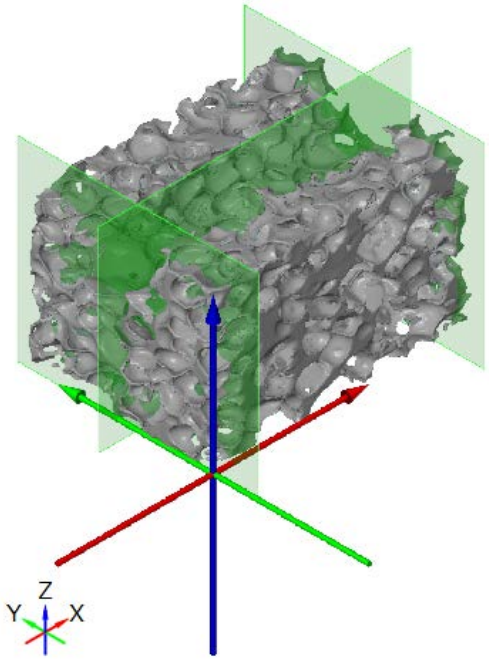


Fig. 7. Positions of the inspection sections

V. CONCLUSIONS

After the CAD comparison we get a surface deviations on the milled surfaces (Fig. 8). This model is very similar to Fig. 5, but here we can see the original part deleted.

The Fig. 8 clearly show the max. and minimum deviation from the original model. We can see also (on the right side of the model) the deformation of the clamping system. This deformation is not very high, but it is important fact (ho fixing and clamping of the foam materials).

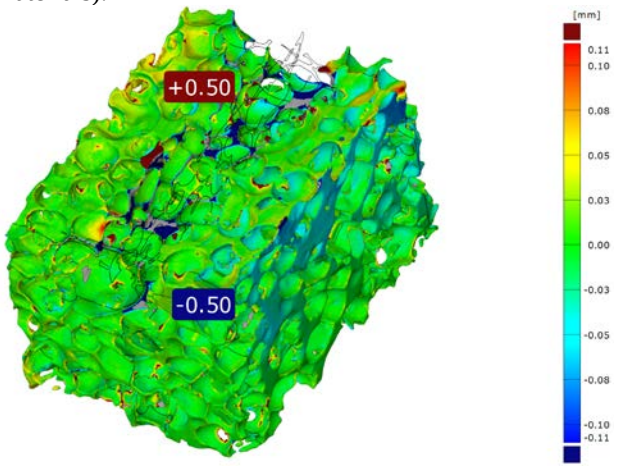


Fig. 8. Result of the milled surface deviations

After the inspection sections (Fig. 7), we can get the section planes from the deformed models (Fig. 9). Here we can see 3 planes. The first was the input, the third was the output of the tool during milling. And we have second also. This section represent across of the toolpath. This section plane was the most important, because here was biggest deformation. It is also very important, under of the milled zone was also cell deformations.

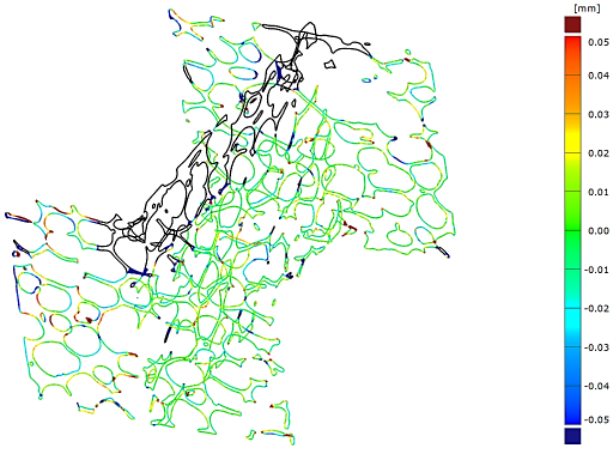


Fig. 9. Deformed parts of the inspection section planes

After the inspection section evaluations we observed the max. and min. deviations in the section planes. By the help of the statistical software we get a regression of the measured values. The regression was very wrong (because the material was very inhomogeneous), but we get some good results. On the base of the regression we can say: the feed per tooh (f_z) was the most important factor of this experiments. The F-significant value was 0,051 and also the p-value have similar number. This factor (f_z) was the only one from the DOE which have

more significant values. On the future we plan make more experiments, and we plane also work out the new measuring method (after cutting) for this type of materials.

VI. ACKNOWLEDGMENTS

This research is party supported by EFOP-3.6.1-16-2016-00006 "The development and enhancement of the research potential at John von Neumann University" project. The Project is supported by the Hungarian Government and co-financed by the European Social Fund.

REFERENCES

[1] **Guarino, S.; Rubino, G.; Tagliaferri, V.; Ucciardello, N.**: Thermal behavior of open cell aluminum foams in forced air: Experimental analysis, *Measurement*, Vol. 60, 2015, pp. 97–103.

[2] **Vesenjak, M.; Krstulovic-Opara, L.; Ren, Z.; Öchsner, Z.; Domazet, Z.**: Experimental study of open-cell cellular structures with elastic filler material, *Exp. Mech.*, Vol 49, 2009, pp. 501–509.

[3] **Beer, M.; Rybár, R.; Kal'avsky, M.**: Experimental heat transfer analysis of open cell hollow ligament metal foam at low Reynolds number, *Measurement*, Vol. 133, 2019, pp. 214–221.

[4] **Tanaka, S; Hokamoto K.; Irie, S.; Okano, T.; Ren, Z.; Vesenjak, M.; Itoh, S.**: High-velocity impact experiment of aluminum foam sample using powder gun, *Measurement*, Vol. 44, 2011, pp. 2185-2189.

[5] **Qiao, H.; Basu, S.; Saladana, C.; Kumara, S.**: Subsurface damage in milling of lightweight open-cell aluminium foams, *CIRP Annals - Manufacturing Technology*, Vol. 66, Issue 1, 2017, pp. 125-128.

[6] **Xing, D.; Zhang, J.;Shen, X.; Zhao, Y.; Wang, T.**: Tribological properties of ultrasonic vibration assisted millingaluminium alloy surfaces, *Procedia CIRP*, Vol. 6, 2013, pp. 539-544.

[7] **Ni, C.; Zhu, L.; Liu, C.; Yang, Z.**: Analytical modeling of tool-workpiece contact rate and experimental study in ultrasonic vibration-assisted milling of Ti–6Al–4V, *International Journal of Mechanical Sciences*, Vol. 142-143, 2018, pp. 97-111.

[8] **Verma, G. C.; Pandey, P. M.**: Machining forces in ultrasonic-vibration assisted end milling, *Ultrasonics*, Vol 94, 2019, pp. 350-363.

[9] https://images.app.goo.gl/da4jz9tRhvuKbbhN8 (01.05.2019)

[10] **Kuruc, M.; Zvončan, M.; Peterka, J.:** Investigation of ultrasonic assisted milling of aluminum alloy AlMg4.5Mn. In Procedia Engineering, vol. 69 (2014), pp. 1048 - 1053, 24th DAAAM International Symposium on Intelligent Manufacturing and Automation, https://www.sciencedirect.com/science/article/pii/S1877705814003 35X, (01.05.2019).

[11] **Kuruc, M.; Zvončan, M.; Peterka, J.:** Comparison of conventional milling and milling assisted by ultrasound of aluminum alloy AW 5083. In IN-TECH 2013: Proceedings of International Conference on Innovative Technologies, Budapest, Hungary, 2013. Rijeka: Faculty of Engineering University of Rijeka, 2013, pp.177-180. ISBN 978-953-6326-88-4.

[12] DMG Ultrasonic 20 linear. https://en.dmgmori.com/products/machines/ultrasonic/ultrasonic-linear, (01.05.2019)

[13] **Matsumoto, Y.; Brothers, A.H.; Stock, S.R.; Dunand, D.C.**: Uniform and graded chemical milling of aluminum foams, *Materials Science and Engineering*, Vol. 447, 2007, pp. 150-157.

[14] **Li, Q.; Wang, J.; Hu, W.**: Optimizations of electric current assisted chemical milling condition of 2219 aluminum alloy, *Journal of Materials Processing Tech.*, Vol. 249, 2017, pp. 379-385.

[15] **Plastic Fantastic**: Carl Zeiss Metrotom 1500 industrial CT scanner http://plasticszone.blogspot.com/2008/09/carl-zeiss-metrotom-1500-industrial-ct.html, (01.05.2019)

[16] **Tian, W.; Han, N.**: Analysis on meso-damage processes in concrete by X-ray computed tomographic scanning techniques based on divisional zones, *Measurement*, Vol. 140, 2019, pp. 382–387.

[17] **Hiller, J.; Reindl, L. M.**: A computer simulation platform for the estimation of measurement uncertainties in dimensional X-ray computed tomography, *Measurement*, Vol. 45, 2012, pp. 2166–2182.

[18] **Chen, S-W.; Chu, C-Y.**: A comparison of 3D cone-beam Computed Tomography (CT) image reconstruction performance on homogeneous multi-core processor and on other processors, *Measurement*, Vol. 44, 2011, pp. 2035–2042.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Experimental Investigation of Single Point Incremental Forming of Aluminium Alloy Foils

Imre Paniti[1], Zsolt János Viharos[2], Dóra Harangozó[3]

[1]*Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences, H-1111 Budapest, Kende u. 13-17., Hungary, imre.paniti@sztaki.mta.hu*
[2]*Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences, H-1111 Budapest, Kende u. 13-17., Hungary, viharos.zsolt@sztaki.mta.hu*
[3]*Széchenyi István University, H-9026 Győr, Egyetem tér 1., Hungary, harangozo.dora@sze.hu*

***Abstract* – Incremental Sheet Forming is a prosperous process to manufacture sheet metal parts that is well adapted for prototypes or small batch production. Compared to traditional sheet forming technologies this relatively slow process is only profitable for the above mentioned production types but it can be used in different applications in automotive and aircraft industries, in architecture engineering and in medical aids manufacturing. In this paper indirectly obtained axial forming force on Single Point Incremental Forming (SPIF) of variable wall angle geometry were studied under different process parameters. The estimation of the forces on AlMn1Mg1 sheets with 0.22 mm initial thickness is performed by continuous monitoring of servo motor currents. The deformation states of the formed parts were analysed using the ARGUS optical strain measurement system of GOM. Interaction plot of forming speed, incremental depth, tool diameter and lubrication were also reported.**

Keywords – Incremental Sheet Forming, Rapid Prototyping, Optical Strain Measurement, Design of Experiments.

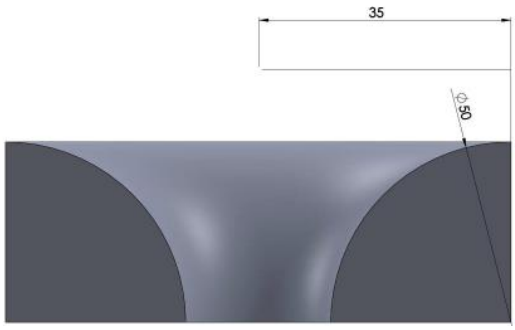
I. INTRODUCTION

Incremental Sheet Forming with its main groups (Single Point Incremental Forming – SPIF – and Two Point Incremental Forming – TPIF) is still an interesting research topic in Material Science because of its extreme and complex mode for deformation, the flexibility of the process and the high forming limits compared to traditional forming processes. Several articles are dealing with experimental study on force measurements for SPIF like [1] or [2], but only a couple of them are focusing on sheets with initial thickness less than 0.5mm [3-6].
Furthermore, as Gatea et al. [7] highlighted regarding the Fracture Forming Limit Curve (FFLC) in a review, that further investigation should be carried out into the effect of initial sheet thickness to tool radius ratio (t_0/R) on the FFLC, and whether it is enough to describe FFLC in SPIF. The above mentioned facts initialized this study to conduct experiments on AlMn1Mg1 sheets with 0.22 mm initial thickness. Former results of this research with the same material explained what kind of control system have been used to execute the tool path on this part [11], and how flat end tools can improve e.g. the accuracy of the part [12], so, the aim of this research work is twofold. At first, the goal is to apply a non-traditional force monitoring on AlMn1Mg1 foils and secondly, to experimentally validate main process parameters on sheets thinner than 0.5mm. The first part of this paper focuses on the material characterization introducing a Forming Limit Curve measured by Nakazima test; the second part of this paper presents preliminary investigations on the formability of truncated conical shapes with a continuously increasing wall angle as a function of major operating parameters. In addition, forming forces have been investigated experimentally with servomotor acquisitions.

II. DESCRIPTION OF THE METHOD

Single Point Incremental Forming experiments were carried out on a Rieckhoff CNC milling machine. The forming tool and a fast-clamping system is shown in Fig. 1(a). The number of experiments required to determine the forming limit of a sheet can be reduced by using a part geometry with variable wall angle as claimed in [8]. For this reason, a conical frustum with circular generatrix (model generating curve) design was used as shown in Fig. 1(b).
The CNC Machine control was realised with an open-source Real-Time Control Software called LinuxCNC. This control allowed to send the tool coordinates to a data acquisition program which collected also the Servomotor Current data of the Z-axis.

(a)



(b)

Fig. 1. (a) Set-up of the experiments, (b) Section view of the test geometry

The chemical composition of AlMn1Mg1 used for this study is given in Table 1. The tensile tests were carried out according to EN ISO 6892-1:2010 standard at room temperature using INSTRON 5582 universal testing machine. Specimens were cut from sheet in 0°, 45° and 90° to rolling direction. The planar anisotropy values (r) were evaluated from longitudinal and transversal strains measured by AVE video extensometer.

The mechanical properties are listed in Table 2. Result of Erichsen Cuping Test is IE=6.79 and Limiting Draw Ratio obtained from Cup Drawing Test according to Swift is LDR=1.7.

Table 1. Chemical composition of the sheet material.

Al	**Si**	**Fe**	**Cu**	**Mn**
96.90	0.201	0.448	0.212	0.807
Mg	**Zn**	**Cr**	**Ni**	**Others**
1.260	0.071	0.022	0.006	0.073

Table 2. Results of tensile tests.

Direction	**Rp0.2, MPa**	**Rm, MPa**	**Ag, %**
0°	88.3	183.0	16.44
45°	90.0	155.5	9.27
90°	86.3	170.3	12.48
A50, %	**n5**	**r10**	
16.88	0.297	0.554	
10.45	0.266	0.580	
12.98	0.268	0.594	

Fig. 2. illustrates the Forming Limit Curve (FLC) and Fracture Forming Limit Curve (FFLC) of tested sheet which were constructed from results of Nakazima test using a hemispherical tool of 50 mm in diameter and GOM ARAMIS digital optical measuring system. FLC curve was evaluated according to EN_ISO_12004-2-2009 standard. It is well known from the literature, that SPIF can achieve more times higher strains than it can achieve during conventional forming process like deep drawing. Therefore, forming limit in SPIF should be represented by Fracture Forming Limit Curve. Fracture limit strains of tested sheet were determined also from Nakazima test using ε_2-ε_1 plots of GOM evaluation system.

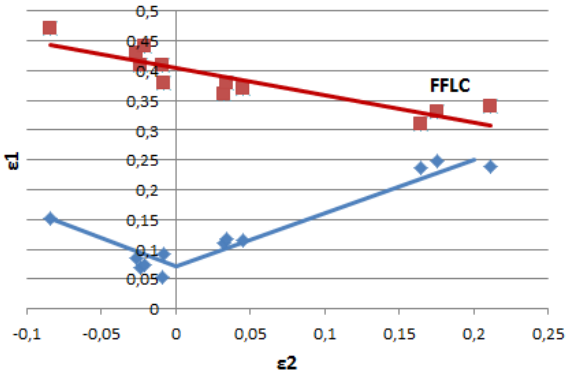


Fig 2. Forming and Fracture Forming Limit Curve

It should be noted that this FFLC seems not realistic as limit of Incremental Sheet Forming for many reasons. The main argue is that during Nakazima test local stretching deformation causes positive stress triaxiality factor while at SPIF the tool generates compressive stress in the sheet metal which might influence the stress triaxiality in negative direction and therefore higher strain limits can achieve. This effect is stronger at Double Point Incremental Forming where compressive load caused by tools is more significant. To realise this enhanced process two Parallel Kinematic Machines (PKMs) [13] or two Industrial Robots should be used [14]. In both cases the synchronisation of the two tools have to be solved. However, the same results can be achieved with one industrial robot, a C-frame, a linear actuator and a mechanical copying device [15].

More realistic values for limit strains can be found in literature. At plane strain (ε_2=0) the limit strain (ε_1=FLD_0) reaches 2.3 for AA1050-O (Filice et al., [16]) and 0.84 for AA6114-T4 while 3.0 for AA3003-O (Micari, [17]). These values show that significant scatter is among empirical values.

Applying the classical equation $t_f=t_0 \cdot \sin(90-\phi)$ (where t_f and t_0 are final and initial thicknesses, ϕ is wall angle) it can be calculated that if $\phi=60^o$ then logarithmic thickness reduction strain (FLD_0) is 0.9 and if $\phi=70^o$ then FLD_0=1.08. Initial wall thickness influences the limit strain, for example Kim et al. [18] showed that if thickness decreases from 0.5 mm to 0.3 mm the limit strain also decreases by 23% to FLD_0=0.92.

Jeswiet et al. [19] elaborated empirical formulae for

calculation of maximum wall angle for truncated conical specimens. The (1a) and (1b) equations show the influence of sheet thickness on wall angle as function of material.

$$\phi_{max}=8.5\ t_0 + 60.7 \quad \text{for AA3003-O} \quad (1a)$$
$$\phi_{max}=3.3\ t_0 + 58.3 \quad \text{for AA5754-O} \quad (1b)$$

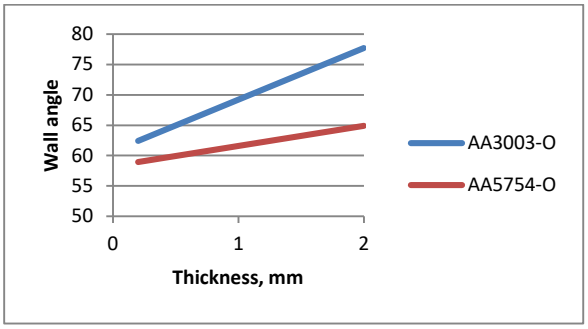


Fig. 3. Wall angle vs. sheet thickness

Fig. 3. shows that lower alloy content of AA3003 enables higher wall angle limits than AA5754 with 3.5% Mg content. At the same time, it is also visible that as the sheet thickness decreases the formability is also decreasing.

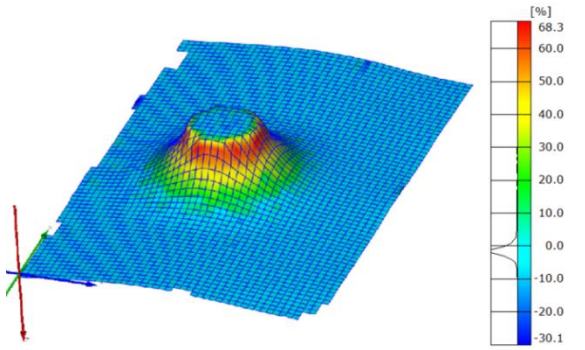
Empirical evaluation of local deformations was measured by GOM ARGUS system [20]. This technique uses a regular mesh on the surface of blank material (Fig. 4(b)). After forming process, the local deformations are calculated using ARGUS software. The results from the ARGUS system provide full-field information about major-minor strain, thickness reduction and geometric parameters of the sheet metal part.

III. RESULTS AND DISCUSSIONS

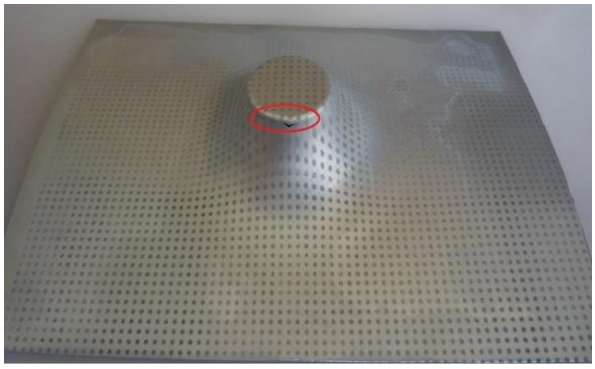
Table 3 shows the process parameters (F: forming speed, △Z: incremental depth, d: tool diameter) applied in the Design of Experiments (DOE, using L9 orthogonal array of Taguchi) and the examined output parameter, the Z coordinate where fraction occurred on the formed part (-Z frac.).

Table 3. Process parameters and results.

Exec. order	F (mm/min)	△Z(mm)	d (mm)	Lubricant	-Z frac. (mm)
1.	500	0.1	2.381	#3	22.32
5.	500	0.3	4	#2	20.16
9.	500	0.5	6	#1	19.60
4.	1750	0.1	4	#1	19.81
8.	1750	0.3	6	#3	19.93
3.	1750	0.5	2.381	#2	20.01
7.	3000	0.1	6	#2	18.95
2.	3000	0.3	2.381	#1	20.01
6.	3000	0.5	4	#3	20.10
1.	500	0.1	2.381	#3	22.32



(a)



(b)

Fig. 4. (a) Thickness reduction of the first part, (b) Picture of the first formed part, indicating the fracture

Z axis loads were obtained to monitor the necking and fracture as in [9]. Fig. 4(a) shows the thickness reduction of the first formed part (from the GOM ARGUS system), while Fig. 4(b) shows a picture of the first formed with a fracture caused by necking.

By using the data of the motor and the drive train, the force applied by the axle as a function (2) of the motor current is the following:

$$F_{Z-current} = \frac{2\pi}{h} \cdot i_s \cdot \left(k_M \cdot I - M_R\right) \quad (2)$$

where:

- $F_{Z\text{-}current}$ – axial force applied by the ball screw nut [N]
- h – ball screw pitch [mm]
- i_s – transmission ratio of the belt drive
- k_M – motor constant [mNm/A]
- I – motor current [A]
- M_R – torque loss due to friction in the motor [mNm]

Similar methodology was used by Rauch et al. [10] to evaluate tool loads in a Parallel Kinematic Machine.

Fig. 5 shows the validation of the measurement concept, by comparing measurement results from a Force

Cell (Fz) with the calculated forces (Fz-current) from the motor current measurements.

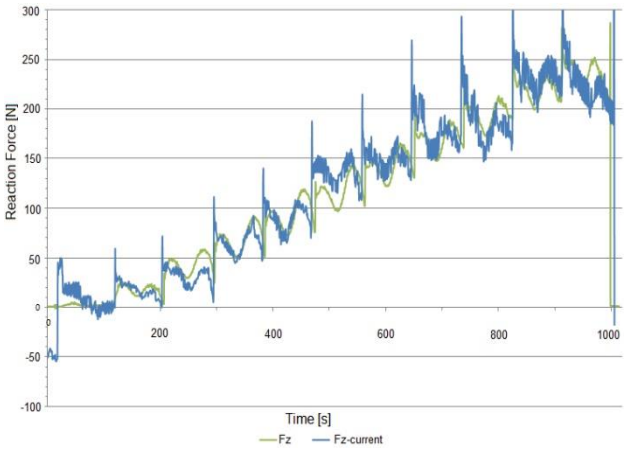


Fig. 5. Validation of the measurement concept

Current measurements were realised with a 0.33 Ohm electrical measurement resistance. From Ohms law the voltages on the CNC's Z axis can be obtained. The peaks in Fz-current are indicating the Z-level changes, where the tool pushes the sheet to reach the next depth level up to the fracture. Fz values are pre-filtered in quasi real-time to get a smoother value change. The reaction force (and the current) increases in the first phase of the forming as the sheet becomes harder to form. Fig. 6 shows the results of the first forming, indicating the values of the fracture (oval mark). In case local necking or fracture occurs, the voltage increases up around the starting value.

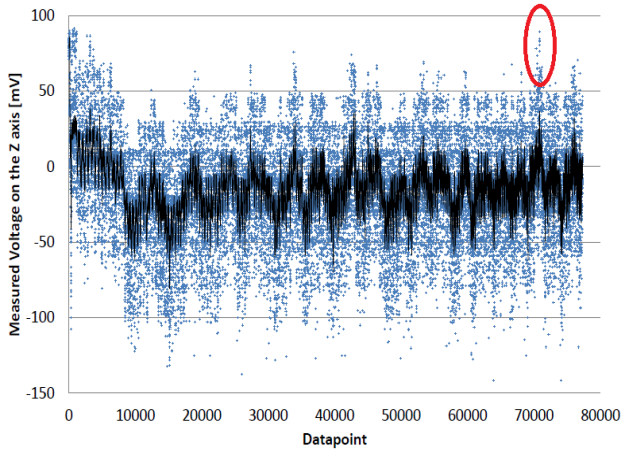


Fig. 6. Measured voltage on the Z axis by executing the forming of the first part and indicating the values of the fracture [11]

The major strain distribution in a dedicated section of the same part can be seen in Fig. 7. Major strain increased up to 126% (1.26) in the area of the fracture. Similar phenomena occur with thicker sheets.

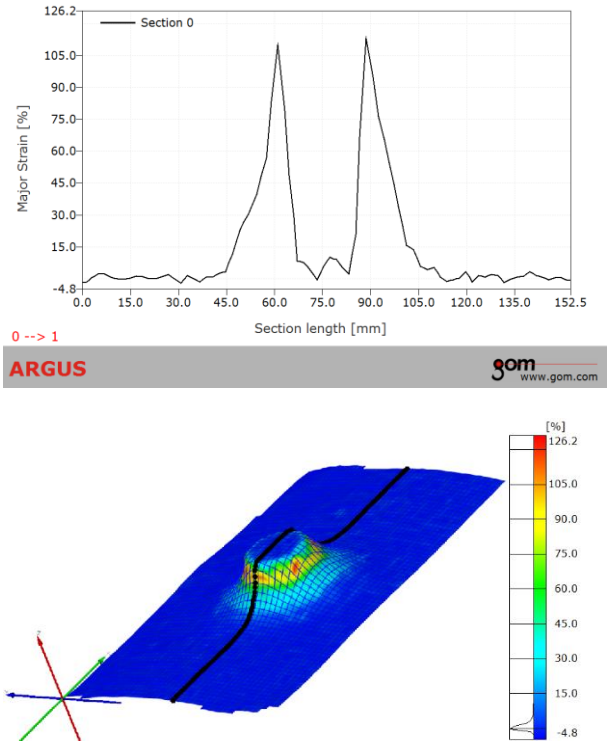


Fig. 7. Major strain distribution of the first part in a section

This value is similar to that of cited from the literature. As it can be seen from Table 3. the Z coordinate where fracture occurred on the formed part (-Z frac) is about 20 mm. Using the part dimensions from Fig. 1(b) the final wall angle can be calculated, giving a value of 78.46°. This is higher than the published wall angle values of regular truncated conic shape which is used as reference geometry for evaluating the limit strains.

To summarize the result of the experiments an Interaction Plot of the factors for $-Z_{fraction}$ is given in Fig. 8.

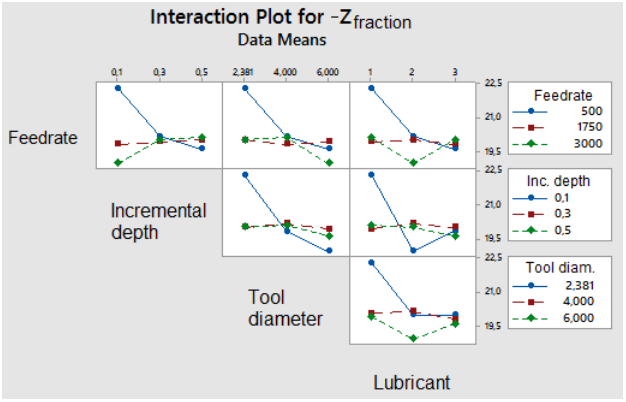


Fig. 8. Interaction of the factors for –Z fraction

Experimental results showed that tool diameter has the main influence on the forming depth in case of SPIF on AlMn1Mg1 sheets with 0.22 mm initial thickness, which reflects the importance of the t_0/R (initial sheet thickness to tool radius) ratio. The second factor in the line of the

influencing parameters is the lubrication (respect to Sommerfeld number) which is followed by the feedrate and the incremental depth.

Further analysis can be carried out by correlation matrix of parameters, which is displayed on Table 4. From that it follows, that highest correlation index is between tool diameter and forming depth, negative sign indicates that as lower the diameter as higher the depth. This fact supports the hypothesis that compressing stresses play a key role in the increased formability of SPIF, as smaller diameter induces compressive stresses more effectively in the surface of sheet.

Second highest indices can be regarded to feedrate and lubrication but the effect of △z incremental depth is less significant. This ranking is in good agreement with the conclusions derived from Fig. 8.

Table 4. Correlation matrix of parameters

	F	Δz	*d*	*Lubr.*	*-Z frac.*
F	1				
⊿z	0.130	1			
d	0.126	0.126	1		
Lubr.	-0.130	-0.130	-0.126	1	
-Z frac.	-0.574	-0.385	-0.661	0.564	1

As lubricant is not a quantified parameter, only tool diameter and feedrate have been analysed by multiple regression. Equation (3) shows the result of numerical calculation:

$$-Z_{frac.} = 22.8 - 0.42 \cdot d - 0.0005 \cdot F \qquad (3)$$

The coefficient of regression is 0.826, which is acceptable for further estimation of forming depth if the same sheet metal is used.

IV. CONCLUSIONS AND OUTLOOK

In this paper, the characterization of AlMn1Mg1 and Single Point Incremental Forming of the same material with 0.22 mm initial thickness have been conducted, applied a Design of Experiments using L9 orthogonal array of Taguchi.

The monitoring of servo motor currents allowed the estimation of the forming forces. All results regarding the estimation of fracture caused by necking are consonant with the results obtained in SPIF of thicker sheets.

Further investigations could be conducted in order to study the influence of lubrication.

V. ACKNOWLEDGMENTS

The research presented in this paper was carried out as part of the EFOP-3.6.2-16-2017-00016 project in the framework of the New Széchenyi Plan. The completion of this project is funded by the European Union and co-financed by the European Social Fund.
The support provided by T. Gyoger and Sz. Szalai in the experiments is also greatly acknowledged.

REFERENCES

[1] **Duflou, J.; Tunçkol, Y.; Szekeres, A.; Vanherck, P.**: Experimental study on force measurements for single point incremental forming, *Journal of Materials Processing Technology*, Vol. 189, Issues 1-3, 2007, pp. 65-72.

[2] **Bagudanch, I.; Centeno, G.; Vallellano, C.; Garcia-Romeu, M.L.**: Forming Force in Single Point Incremental Forming under Different Bending Conditions, *Procedia Engineering*, Vol. 63, 2013, pp. 354-360.

[3] **Thibaud, S.; Ben Hmida, R.; Richard, F.; Malécot, P.**: A fully parametric toolbox for the simulation of single point incremental sheet forming process: Numerical feasibility and experimental validation, *Simulation Modelling Practice and Theory*, Vol. 29, 2012, pp. 32-43.

[4] **Ben Hmida, R.; Thibaud, S.; Gilbin, A.; Richard, F.**: Influence of the initial grain size in single point incremental forming process for thin sheets metal and microparts: Experimental investigations, *Materials & Design*, Vol. 45, 2013, pp. 155-165.

[5] **Obikawa, T.; Satou, S.; Hakutani, T.**: Dieless incremental micro-forming of miniature shell objects of aluminum foils, *International Journal of Machine Tools and Manufacture*, Vol. 49, Issues 12–13, 2009, pp. 906-915.

[6] **Sekine, T.; Obikawa, T.**: "Micro Incremental Forming Characteristics of Stainless Foil", *Key Engineering Materials*, Vols. 447-448, 2010, pp. 346-350.

[7] **Gatea, S.; Ou, H.; McCartney, G.**: Review on the influence of process parameters in incremental sheet forming, *The International Journal of Advanced Manufacturing Technology*, 2016, pp. 1-21.

[8] **Hussain, G.; Gao, L.**: A novel method to test the thinning limit of sheet metal in negative incremental forming, *International Journal of Machine Tools and Manufacture*, 47, 2007, pp. 419-435.

[9] **Ambrogio, G.; Filice, L.; Micari, F.**: A force measuring based strategy for failure prevention in incremental forming, *Journal of materials processing technology*, Vol. 177., No. 1., 2006, pp. 413-416.

[10] **Rauch, M.; Hascoet, J. Y.; Hamann, J. C.; Plenel, Y.**: Tool path programming optimization for incremental sheet forming applications, *Computer-Aided Design*, Vol. 41., No. 12., 2009, pp. 877-885.

[11] **Paniti, I.; Viharos, Zs. J.**: Fracture diagnostics for Single Point Incremental Forming of thin Aluminum alloy foils, Proceeding of 15th IMEKO TC10 Workshop on Technical Diagnostics in Cyber-Physical Era, Budapest, Hungary, June 6-7, 2017, pp. 34-38.

[12] **Najm, S. M.; Paniti, I.**: Experimental Investigation on the Single Point Incremental Forming of AlMn1Mg1 Foils using Flat End Tools, IOP Conference Series: Materials Science and Engineering, Vol. 448, No. 1, 2018, November, p. 012032

[13] **Johnson, C. F.; Kiridena, V. S.; Ren, F.; Xia, Z. C.**, Ford Global Technologies LLC: System and method for incrementally forming a workpiece, 2012, U.S. Patent 8,322,176.

[14] **Meier, H.; Magnus, C.; Smukala, V.**: Impact of superimposed pressure on dieless incremental sheet metal forming with two moving tools. *Cirp Annals*, 60(1), 2011, pp. 327-330.

[15] **Paniti, I.; Somlo, J.**: Device for two sided incremental sheet forming, 2011, EU Patent EP2505279A1

[16] **Filice, L.; Fratini, L.; Micari, F.**: Analysis of Material Formability in Incremental Forming, *CIRP Annals*, Vol. 51/1/2002, pp. 199-202.

[17] **Micari, F.**: Single Point Incremental Forming: recent results, Seminar on Incremental Forming, 22 October 2004. Cambridge University. CD-ROM.

[18] **Kim, T.J.; Yang, D.Y.**: Improvement of formability for the incremental sheet metal forming process, *International Journal of Mechanical Sciences*, Vol. 42, 2001, pp. 1271-1286.

[19] **Jeswiet, J.; Hagan, E.; Szekeres, A.**: Forming Parameters for Incremental Forming of Aluminum Sheet Metal, *IMECHE part B, J. of Engineering Manufacture*, Vol. 216, 2002, pp. 1367-1371.

[20] https://www.gom.com/metrology-systems/argus.html, last visited: 26.06.2019.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

The role of laboratories in the international development of accreditation

A. S. Ribeiro[1], J. Gust[2], A. Vilhena[1], J. Wilson[3]

[1]*AffiliationNational Laboratory for Civil Engineering, Lisbon, Portugal, asribeiro@lnec.pt, avilhena@lnec.pt, (+351) 218443000*
[2]*Fluke Calibration, Everett WA, USA, jeff.gust@flukecal.com, +1 (425) 293 6035*
[3]*National Laboratory Association, Pretoria, South Africa, jgpwilson@xsinet.co.za, +27(83)6520770*

***Abstract* – Laboratories are the larger group of stakeholders of ILAC, the international organization for accreditation. The impact of laboratory activities as independent entities that ensure safety and quality of products and services is growing and becoming a key issue in many fields. New challenges of globalization and international trade, technology developments and increasing expectations of citizens and consumers requires a robust implementation of accreditation to promote the fair competition and the sustainability of laboratories. For this purpose, harmonization and consistency of accreditation is a major task to be achieved through cooperation of all stakeholders.**

***Keywords* – Accreditation, Laboratory, Conformity Assessment, Testing.**

I. INTRODUCTION

There are many reasons for understanding the importance of accreditation for the economy. One of the most relevant is that global markets require a growing level of confidence in goods and services. Trade operations are supported in trust regarding the quality of goods and services provided by suppliers and agents of the transactions (namely, buyers and sellers) [1].

To provide the assurance of goods and services, a quality infrastructure is needed, bringing an added value for the citizens and consumers through the development and use of tools able to fulfil the expectation of safety and quality of life. Quality infrastructure is [2] *"The system comprising the organizations (public and private) together with the policies, relevant legal and regulatory framework, and practices needed to support and enhance the quality, safety and environmental soundness of goods, services and processes."*.

According to the BIPM [2] *"The quality infrastructure is required for the effective operation of domestic markets, and its international recognition is important to enable access to foreign markets. It is a critical element in promoting and sustaining economic development, as well as environmental and social wellbeing."* considering that it relies on five domains:

- metrology;
- standardization;
- accreditation;
- conformity assessment; and
- market surveillance in regulated areas.

The relationship between conformity assessment and accreditation has a large impact on the global economy as conformity assessment is a way of ensuring that the suppliers adhere consistently to the standard. This means that they can be relied on and accreditation is a reinforcement of conformity assessment using external, independent entities (National Accreditation Bodies, NAB) that validate competence of Conformity Assessment Bodies (CAB) with services such as, calibration, testing, inspection and certification.

Accreditation of CAB's is implemented byusing system standards suitable to recognize, at international level, competence of conformity assessment providers. In the case of laboratories, ISO/IEC 17025 [3] is the reference and in the case of Inspection Bodies, ISO/IEC 17020 [4] supports the accreditation. The framework of these standards is called the ISO Casco toolbox (Fig. 1).

Fig. 1. ISO Casco Toolbox (Credits: ISO)

The recent revision of ISO/IEC 17025, introduced relevant changes, namely, a new vision of harmonization with other system standards (ISO 9000 and others), higher focus of the client, the interpretation of sampling activities, the implementation of decision rules associated with statements of conformance, the concept of impartiality, highlight of the process-based approach, risk-based thinking and flexible structure of organizations.

II. NEW TRENDS

Today laboratories face the same challenges as many organizations, overwhelmed by the growing development of technologies, products, materials, services which create new expectations that require innovation as new tool applied to conformity assessment. The dynamic and sometimes unpredictable world of today, described by the acronym VUCA (volatility, uncertainty, complexity and ambiguity), requires strong leadership, vision and knowledge to face daily shifts in economy, trade and international relations.

This development is strongly related to the digital transformation, being in the spotlight of government strategies worldwide. At European level, the digital single market (dsm) [5] – see Fig. 2 – is expected to have a huge impact on economic growth, boosting jobs, competition, investment and innovation [6].

Fig. 2. EU Digital Single Market Strategy, in [6]

To achieve the goals related to the digital economy, smart cities concept [6], data-driven Society, reshape of communication and businesses through 5G, artificial intelligence, and many other challenges means that laboratories need to be involved in the digital transition and to be able to adapt to many daily changes that happen.

Urban development and smart cities are key concepts for the understanding of the relations to be established between digital technologies, disruptive innovation and society environment [7,8].

Performing conformity assessment in this new framework is becoming a dynamic challenge that will need flexibility, responsibility, knowledge (scientific and technical) with more specialized and competence in new fields (e.g., data science and data management, artificial intelligence, nano and bio materials, large sensor networks using big data approaches and quantum computing).

III. ILAC FORUM AND THE ROLE OF ILAC LABORATORY COMMITTEE

ILAC (International Laboratory Accreditation Cooperation) was created in 1977 in Copenhagen "with the aim of developing international cooperation for facilitating trade by promotion of the acceptance of accredited test and calibration results", being the international organization for accreditation. It includes the accreditation bodies operating in accordance with ISO/IEC 17011, able to support the accreditation of entities such as, conformity assessment bodies (CABs, using ISO/IEC 17025 [3]), medical laboratories (using ISO 15189 [9]) and inspection bodies (using ISO/IEC 17020 [4]).

The main added value is given by the Mutual Recognition Arrangement (ILAC MRA) and the evolution of accreditation of CABs in recent years (Figure 3) shows an increasing interest due to globalization as a way to standout from competitors.

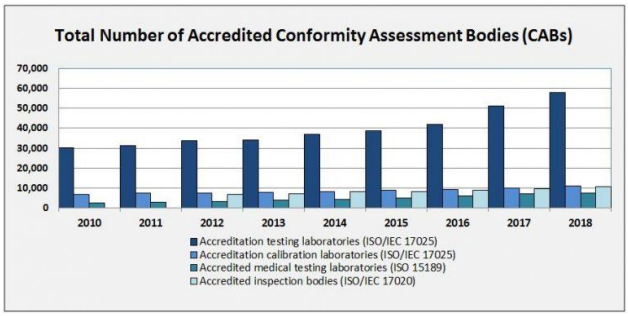


Fig. 3. Evolution of accredited CABs worldwide (2010 – 2018),in https://ilac.org/about-ilac/facts-and-figures/

This mutual recognition gives a unique approach that allows the recognition of competence of laboratory activities providers from the signatory Countries, based on mutual evaluation and acceptance of the national laboratory accreditation systems. The Global network of services rely on this system which expect that they can assure safety and quality of life to citizens and consumers by testing, inspection and certification (TIC).

The TIC Sector market is a large contributor to the economy (see Figure 4), as mentioned in [10] "*It provides greater consumer protection, safer products and industrial installations, reduces compliance costs for SMEs and increases brand reputations and consumers' trust and confidence in a product by ensuring that products, infrastructures and processes meet the required standards and regulations in terms of quality, health and safety, environmental protection and social responsibility and can, therefore, also be a facilitator to international trade".*

According with ILAC *"in 2018, almost 76,500 laboratories and over 10,500 inspection bodies were accredited by ILAC MRA Signatories".* This gives a special role to laboratories as stakeholders in the ILAC organization, recognized by including in its structure the Laboratory Committee, which provides a platform for interaction with the laboratory community.

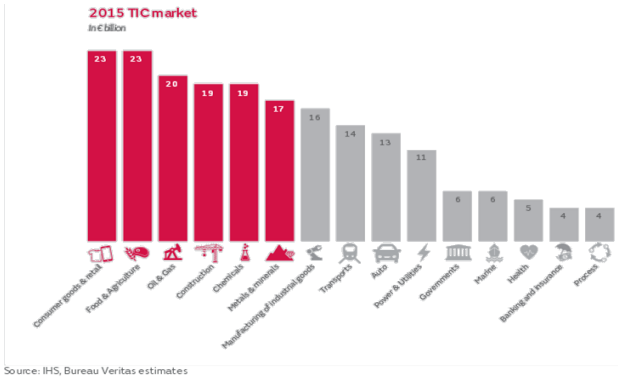


Fig. 4. 2015 TIC market (SGS Report)

IV. CONCERNS, EXPECTATIONS AND CHALLENGES OF LABORATORIES: THE 2017 INTERNATIONAL SURVEY

As described above, accreditation is essential for modern trade which requires harmonization and consistency in the application of its rules worldwide to have a fair and balanced competition. However, in some circumstances, different practices are found at national and regional levels. To support the development of ILAC strategy, the ILAC Laboratory Committee decided in 2017 to develop an international survey on harmonization and consistency of accreditation to provide a global view of the concerns, expectations and challenges of Laboratories and Conformity Assessment Bodies. 421 responses from 35 countries and economies were received, allowing an analysis divided in three parts:

- characterization and validation of the sample of participants;
- accreditation performance analysis; and
- satisfaction analysis regarding the services provided by AB's and improvement analysis.

Results are found in the ILAC LC report [11] and include the analysis of the following issues identified by stakeholders of the ILAC Laboratory Committee regarding the lack of consistency observed: Scope definition; Cycle and frequency of assessment; Quality of Assessment; Translation barriers; Non uniform interpretation of ISO 17025 at national level by Accreditation Bodies; Accreditation Body policies; Use of PT/ILC and similar quality control tools in assessment; and Lack of recognition of ILAC Mutual Recognition Arrangement.

The main conditions to measure the impact of the survey were [11]:

- number of replies should be statistical relevant;
- ideally geographical representation by countries and by regions;
- cover a wide range of activities;
- represent a range of small, medium and large companies and institutions;
- cover a range of small to large number of accreditation parameters (both testing methods and calibration parameters);
- capture respondents experience regarding accreditation process.

A critical issue regarding the development of the survey was how the sampling of laboratories that participated in the survey would represent the whole community and their activities.

The distribution of laboratories by activities showed that a large spectrum of domains were included, although there were higher participation in some fields, namely, agro-food, environment and metrology. Figure 5 shows the distribution of replies obtained by activity.

Another relevant feature of the survey was to develop a performance analysis of ABs, considering the main conditions for accreditation, namely, time to get accreditation; use of extra requirements for accreditation; cooperation with stakeholders; impact of quality control tools; surveillance time interval; and cycle of accreditation.

Finally, a section of the survey aimed to establish a procedure to evaluate the degree of satisfaction of the laboratories regarding the accreditation process in different aspects, such as:

- Quality and competence of the auditors;
- Quality of the interpretation guides;

- Accreditation process;
- Complaints procedure.

The classification obtained regarding aspects to be improved and best performed aspects of accreditation are presented in fig. 6 and Fig 7, respectively.

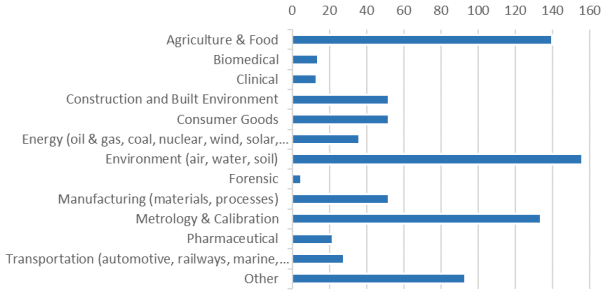


Fig. 5. Distribution of replies by types of activities [11]

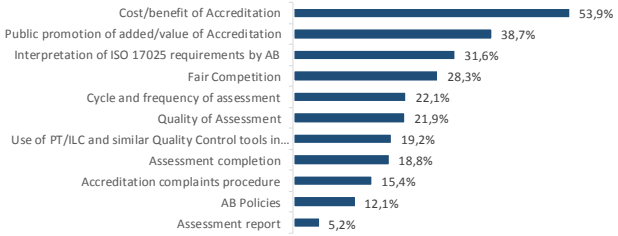


Fig. 6. Major aspects of accreditation needed to be improved in countries and economies [11]

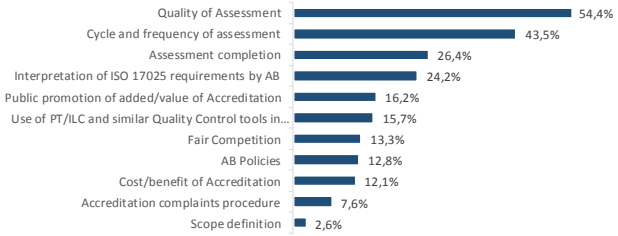


Fig. 7. Best performance aspect of accreditation in countries and economies [11]

V. CONCLUSIONS AND OUTLOOK

Accreditation plays a fundamental role in the economy and is expected to be an engine of economic development on a fair trade basis. In order to achieve this objective, it is necessary to ensure that the application and assessment to the standard requirements of ISO/IEC 17025 which is the reference standard for Laboratory accreditation, is applied worldwide in a balanced, harmonized and consistent manner.

The collaboration of the laboratory community for this purpose is necessary and the role of entities such as ILAC Laboratory Committee is particularly relevant, as demonstrated by the action of promoting a survey to this community in order to understand the current state of implementation of accreditation, and its positive aspects and opportunities for improvement.

This approach has also enabled an independent evaluation of NAB's performance as a way of supporting the evolution of accreditation and its recognition in the international context.

VI. ACKNOWLEDGMENTS

To the members of ILAC Laboratory Committee, co-authors of the 2017 International survey on harmonization and consistency of Accreditation. A global view of Laboratories and Conformity Assessment Bodies concerns, expectations and challenges.

To the laboratories and other stakeholders, national and international laboratory and conformity assessment bodies associations and international organizations that supported and contributed for this study.

REFERENCES

[1] **Frenz, M.:** The Economics of Accreditation, Marion Frenz, Ray Lambert, UKAS, March 2013

[2] **2017 definition Quality Infrastructure**, INetQI, adopted by BIPM, IAF, IEC, ILAC, ISO, ITC, ITU, OIML, UNECE, UNIDO and the World Bank.

[3] **ISO/IEC 17025:2017** General requirements for the competence of testing and calibration laboratories, ISO, Geneva, Switzerland.

[4] **ISO/IEC 17020:2012** Conformity assessment -- Requirements for the operation of various types of bodies performing inspection, ISO, Geneva, Switzerland.

[5] **A Digital Single Market Strategy for Europe**. Communication from the Commision to the European Patliament, the Council, the European Economica and Social Committee and the Committee of the Regions, 6 May 2015, Brussels, Belgium,

[6] **European Commission Digital Strategy – a digitally transformed, user-focused and data-driven Commission,** European Commision, 21 Nov. 2018, Brussels, Belgium.

[7] **Smart Cities. How rapid advances in technology are reshaping our economy and society**, Deloitte Report, Vers. 1.0, Nov. 2015.

[8] **Digital Agenda for the EU. Digital transition action plan.** EC, 28 Jun. 2018, Brussels, Belgium.

[9] **ISO 15189:2012 Medical Laboratories- requirements for quality and competence.** ISO, Geneva, Switzerland.

[10] **White Paper: Trends & Challenges for the third party TIC Sector, and its implications for CEOC International**, September 2015, CEOC International, Brussels, Belgium.

[11] **2017 International survey on harmonization and consistency of Accreditation. A global view of Laboratories and Conformity Assessment Bodies concerns, expectations and challenges.** ILAC Laboratory Committee report, April 2018.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

A New Perspective of the Cyber-Physical Production Planning System

Viola Gallina[1], Lukas Lingitz[1], Matthias Karner[1],

[1]*Fraunhofer Austria Research GmbH, Theresianumgasse 27, 1040 Vienna, Austria, viola.gallina@fraunhofer.at, +43 676 888 616 46*

***Abstract* –The following three research fields and levels can be identified in the field of production planning and control (PPC): production planning, production scheduling and production control. These topics were thoroughly investigated from different points of view by several authors in the recent years. However, the connection and interaction of these fields in the era of cyber-physical production systems (CPPS) has not received so many attention from the research community. The objective of this conceptual paper is to overview the current status of these fields; to examine the possible connection possibilities considering the different artificial intelligence solutions developed for the separate fields and to give an outlook regarding future research activities in PPC in the context of industry 4.0.**

Keywords – production planning and scheduling, artificial intelligence, machine learning.

I. INTRODUCTION

One of the biggest challenges of today`s production (planning) systems in the era of Industry 4.0 and CPPS is to be flexible and adaptive whilst being robust, resilient and efficient. This challenge is present on all levels of a production planning system. Mid-term production planning includes production order scheduling, lot-size calculation and capacity planning [25]. The term "production scheduling" refers to planning of dates, sequences (combined) and routing of products. It provides answers to the question: "When is order O processed on machine M" [20] In [25] production scheduling is described as short-term planning on the operational level and includes occupancy and sequence planning. Production control takes all actions into account, that are needed to guide a production order through the production system after its release. During the execution of a production schedule several unforeseen disturbances may occur (e.g. machine breakdowns, illness of the workers) that could result in discrepancies between planning and reality. Production control includes monitoring and controlling activities in order to implement a production schedule. Controlling is responsible for providing transparent and interpretable information with the evaluation and regulation of the production system in its basic settings – without a concrete order reference [25].

In the recent years several new topics have appeared in the field of production planning and control (PPC), such as:

- The high variety of product variants [24] leads to complex systems. Because of the flexibility of the production system the expenditures are getting higher. Today's production systems are no longer tradable by humans.
- Companies are facing the challenge of achieving short delivery times and a high ability to deliver despite increasingly shorter planning horizons, a large number of internal and external planning changes and increasing planning complexity.
- Today, many technological possibilities are available (AI, OR) whereas only few practical approaches exist in the manufacturing environment. [1]
- The topic of energy and resource efficiency has been on the rise for several years now. The pressure on producers to save CO_2 equivalents is also increasing significantly as a result of public interest. [26]

The above mentioned issues might be investigated with different artificial intelligence (AI) - based solutions. Mid-term production plans commonly neglect capacity restrictions or dynamic effects as described above, which can influence lead times. This often leads to high deviations between the plan and the later execution. Dynamic models that incorporate stochastic effects and interdependencies instead of static average values from the enterprise resource planning system (ERP) or manufacturing execution system (MES) could increase the planning quality of a mid-term production plan. If the deviations between the real and planned (time) data is smaller, more reliable production schedules could be created. Reliability of production schedules can be further increased if different aspects such as condition monitoring,

energy efficiency, predictive maintenance etc. are considered. Resilient, decentralized and adaptive production control can be achieved with the help of modern AI or information-communication technologies (ICT) (e.g. Industrial Internet-of-Things, cloud computing, sophisticated sensors, robotics etc.) solutions.

The current reference model for the PPC (Aachen PPC model) [25] dates from 2006 and is based on scientific findings and facts at the turn of the millennium. Industry 4.0 and CPPSs were hardly researched at that time.

The main objective of the paper is to examine the possibilities of reducing the production costs and coping with the prior mentioned issues through the use of ICT (machine-to-machine communication), AI and in particular methods of machine learning (ML) at all mentioned levels of production planning.

The paper is structured as follows: First the stat-of-the-art results are revised for the above mention three fields of production planning: mid-term planning with ML, condition-based scheduling and adaptive, decentralised control with reinforcement learning. In the subsequent chapter the connection and interaction of these levels and a vision of a new production planning system are discussed. In the conclusion, challenges and potential research fields are identified.

II. RELATED RESULTS IN THE LITERATURE

In this Section the state-of-the-art results of the three research fields of PPC are presented.

A. Mid-term production planning

Gyulai et al. [10] define the robustness of a production plan by comparing the performance indicators of the production system achieved, with those expected according to the planning. If the indicators are still at an acceptable level, he speaks of a robust plan that already anticipated unforeseen disruptions that occur during execution. Beside robustness, resilience – the ability of a system to cope with changes of all kinds – can take into account disturbances of a production system.

ML is one possibility to learn from stochastic fluctuation caused by different influencing factors and is therefore a common approach to deal with uncertainties. ML has gained increasingly high attention in the context of production in the recent past. Cheng et al. [4] found in their review article that most publications focus on production scheduling and other applications (defect analysis, quality improvement and fault diagnosis), and rarely explore the prediction of production planning relevant times (flow time, lot cycle time or lead time). The topic of production planning relevant prediction was investigated by several authors (see e.g. [11, 18, 22, 23, 33]), however, only a limited number of the analysed papers achieved a high technology readiness level. Moreover, data generated with the help of simulation models is dominating within the applied methods, suggesting that the usage of predictive data analytic techniques is seldom in PPC.

The overall objective of PPC is the creation of reliable production plans, so as their realization on the shop floor should be close to – or ideally the same as – the production plan as originally planned. With the help of ML data used for mid-term production planning might be dynamised in order to get more reliable production plans. However none of the related literature gives states a method how to apply the dynamised time prediction model within the planning to create more reliable production plans.

B. Condition- based scheduling

Condition Monitoring (CM), the measurement of the health of machines and tools, is a topic which is mainly discussed in the context of maintenance [2, 9, 12]. However, CM information is an important factor in PPC as well. This is especially true in manufacturing as an equipment's condition influences its capability of producing qualitative products. This influence can be described as a bidirectional correlation between the production program and an equipment's condition. In [15] an example concerning the food industry is presented: "In the food industry [...] the contamination [...] with allergens plays a fundamental role. Once an equipment is contaminated with e.g. gluten, it cannot produce gluten-free products anymore. PPC therefore needs to consider this change in condition [...]."

Traditionally, the topics PPC and maintenance are considered separately [8]. When it comes to CM there is even less literature available which deals with the integration of CM data within PPC. When CM is considered within scheduling or sequencing the condition is often modelled binary – a machine is either operative or non-operative [30]. However, there exists some research in the semiconductor industry. In [7] the production yield is considered as correlating with the machine conditions. The authors conclude, that wafer with a bigger diameter should be produced on machines in a better condition in order to increase yield. Another paper from the semiconductor industry [13] highlights the fact that the production quality drops, if the equipment's condition decreases. Nevertheless, approaches considering CM data in PPC are scarce, especially when it comes to applicability in different industries.

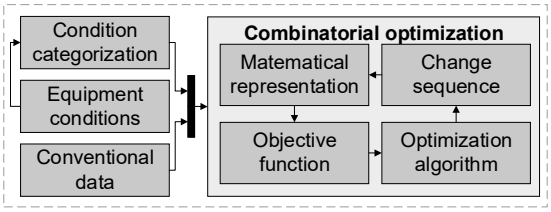


Fig. 1. Condition-based scheduling

In past research [15, 16] a condition-based scheduling was introduced. Based on a categorization of products' condition demands using a single-digit health parameter a decision supporting process can be modelled. Based on the decision supporting process a sequencing technique can be developed in order to optimize the production program by considering the condition of the equipment. Figure 1. shows the concept of condition-based scheduling.

Condition-based scheduling builds on the mid-term production plan (MTPP) and can be used for process and sequence planning. The basic dates from the MTPP are used as main constraints. Furthermore, the machine conditions (capabilities of certain resources to produce certain products) are considered as well. Taking these auxiliary conditions into account, condition-based scheduling creates a short-term production plan that resolves machine assignments as well as sequences.

C. Production control with adaptive, self-optimizing agents

Different agent-based and event-driven approaches were applied in PPC in order to answer challenges of the technology development of recent time [5, 17]. With the increase of computation power there is a possibility to combine these methods with AI techniques and solutions. Reinforcement learning (RL) seems to be a promising approach for self-optimizing, adaptive production control. RL is one category of ML – next to supervised and unsupervised learning. In case of RL there is no dataset in which a function should be found to describe the relationship between inputs and output(s) (supervised learning) or elements should be grouped (unsupervised learning) but the dataset is generated during the learning phase. In RL a learning agent learns the best behaviour, the so-called best action to choose without any prior knowledge and through a sequential and situational decision-making process. If the agent makes a good decision gets a good reward, however in case of a bad decision it can be punished. With the help of this feedback the agent is able to learn a strategy that guides the agent toward achieving the desired goal. The RL process is shown on Figure 2.

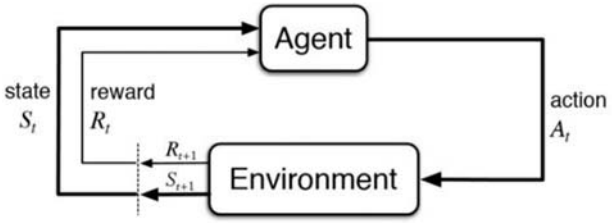


Figure 2.: Representation of RL

It must be mentioned here, that the RL agent needs a feedback from the environment during the learning phase. During the learning the agent has to test and learn the interrelation between the features and the consequences of its decision. In case of a production related problem this environment can exclusively be a simulation model, in case of a real production environment there is no possibility to try out different previously known bad outcomes.

Depending on the current state of the environment – described by different features of the system – the learning agent chooses always the action that results in the maximal expected reward. In this way, the learning agent is able to make real-time decisions in a dynamic environment that is a crucial ability in a cyber-physical production planning system (CPPPS).

Several production related topics were investigated with the help of RL in the recent decades. Das et al. [6] compared two well-known heuristics and a RL based solution for a *maintenance* problem and they received that maintenance policy learned by the RL agent was more flexible and could result in a better maintenance policy within a dynamic environment. With a modification of Das` model Mahadevan and Theocharous [19] were able to maximize the throughput of a transfer line while minimizing WIP and failures of the machines. Their model outperformed a Kanban heuristic. Kara and Dogan [14] studied the *ordering policies* of an inventory system and found that RL gives better result in case of a high demand variance. Rana and Oliveira [21] showed that RL can be used for dynamic *pricing* of interdependent products. In the recent years, RL has been applied to *motion planning* for industrial robots as well.

Production scheduling is the field of manufacturing that has been most widely investigated with the help of RL in the last decades. Several authors have studied this topic and showed that in a dynamic environment a strategy developed by a RL agent is able to outperform the conventional solutions currently applied in the industry. One of the first applications to a static job shop scheduling problem was presented by Zhang and Dietrich. [31] Wang and Usher [29] found that a RL agent is able to learn the best rules for different system objectives to a dispatching rule selection problem for a single machine. Bouazza et al. [3] applied two different agents in their simulation model: intelligent products – who chose a best machine – and intelligent machines – who select the most suitable dispatching rule. The successful implementation and application are presented in Sticker et. al and Waschneck et. al. [27, 28] They showed that RL can be used for adaptive production control in a dynamic, complex production environment. In the review article of Zhang et al. [32] it was concluded, that future research efforts should be shifted to smart distributed scheduling modelling and optimization. The vision of industry 4.0 for production planning and control to a decentralized, self-learning and adaptive production system can be achieved with the application of RL in production control.

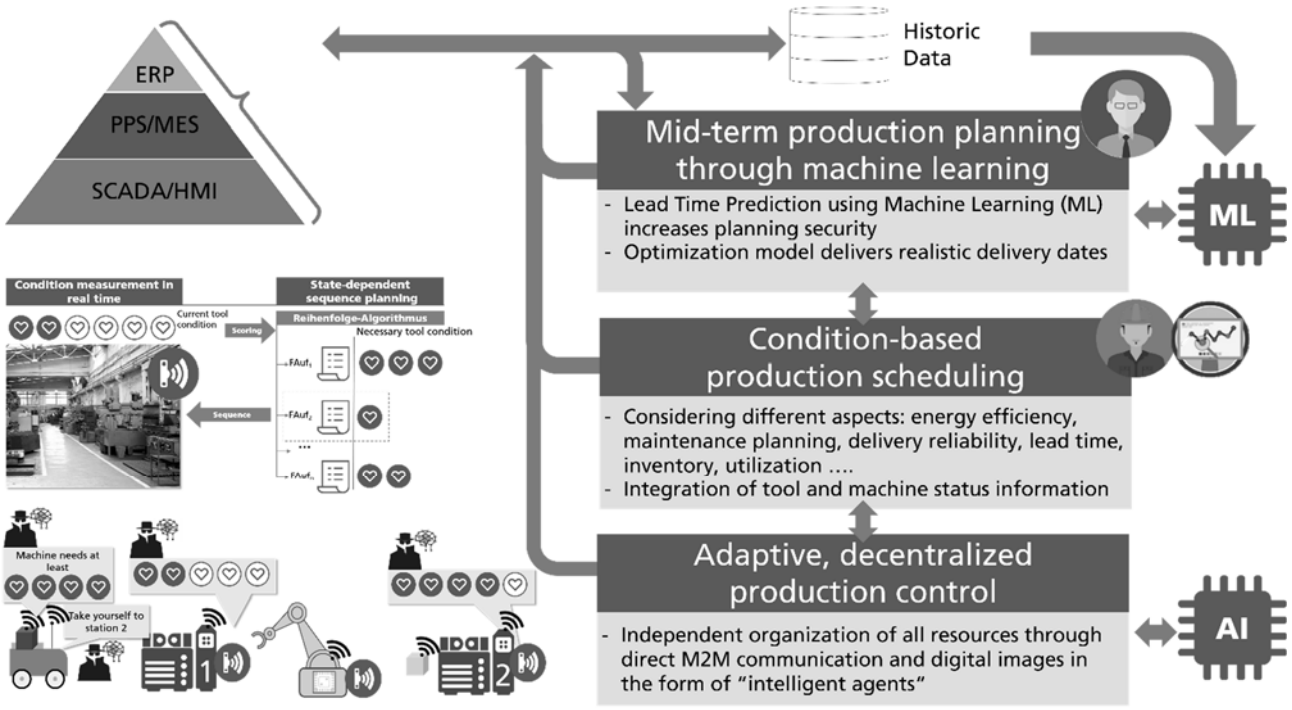


Figure 3. A new perspective of production planning and scheduling

III. DISCUSSION

As can be seen from the literature review several approaches have been researched so far to face the current challenges of manufacturing industries. These approaches use ML, AI, agent-based systems and traditional operations research methods. Still, they have been researched on the one hand only in isolation and on the other hand only to a certain extent. Therefore, only isolated solutions exist. To the best to our knowledge, there is no holistic view or model that combines all above mentioned methods into a single model for PPC. According to the authors, a radical new perspective of PPC is needed, in order to respond to the challenges PPC will phase in the near future when trying to fully utilize the potential of CPPSs. Figure 3. shows the different levels of the production system that need to be adapted as well as their related data sources. On the right hand side of the picture the three levels of the PPC system can be seen: mid-term planning, scheduling and controll. Traditionally, there could only be a direct connection or interaction between two adjacent levels. That means there is strictly hierarchical information flow between mid-term planning and controll level. However, the objects network of industry 4.0 enables this possibility. The different data sources – used as inputs on the given levels – are illustrated on the left hand side. At this point the authors want to mention, that the automation pyramide is just a simplified, yet very common representation for the different systems within the production environment. With this simplification we also address all intelligent subsystems of modern CPPSs. These different intelligent elements are connected in a network shaped structure and exchange information directly among each other and with the PPC system. However, the intelligent elements of the production system (machine, product, order) control the production based on different sensor data. It is a characteristic of such a decentralised system, that the intelligent elements can make sub-optimal decision for themselves in order to achieve the global optimum of the whole production system. The interaction and communication between the autonomous elements enables the creation of a more reliable and realistic production schedule. This new perspective of PPC leads to several research questions that need to be answered. In the following this research topics are discussed.

How can machine learning be used to create mid-term production plans? Although we list several approaches that use machine learning methods for time prediction e.g. for lead time or process time prediction, there are hardly any applications where such models are used to calculate a production plan. High prediction accuracy is dependent from the features that describe the production system best. It is very likely that at least a few of the feature, that are needed for a high prediction accuracy are not available as long as the planning doesn't reach a certain state.

How can different aspects be integrated into a production schedule? Although several examples can be found where different energy or resource efficiency aspects or the health condition of tools or machines are

considered in case of creation of a production schedule, mostly only one aspect is investigated at the same time and the different aspects are not combined.

How could production control be efficiently performed? In the era of industry 4.0 – where everything is automated and through various sensors a lot of data about the current status of the production is available – how could disturbances be handled? In case of a machine brake down an intelligent product or an intralogistic element transporting the products might be able to choose an alternative machine on their own. Is RL the most suitable method to deal with such disturbances or are there other adaptive, self-optimizing approaches?

How is the information exchange between the different planning levels (mid-term planning, condition based scheduling and production control with RL) of the planning system organized? Traditional planning systems follow a successive planning approach. The planning result of one planning level is the direct input of the next level. The possibilities through embedded systems and modern communication technologies have the possibilities to achieve a faster response in case of disturbances. Therefore, information from the shop-floor can be distributed among all relevant parties and systems. On the other hand it is not yet clear how the systems will interact exactly among each other. Hence, it is clear that the constant information exchange is needed. One important question of this topic is the frequency of the information exchange, as (quasi) real-time data can have different granularity, resolution, transition time depending on the industry branch or company.

Is the architecture of the future PPC centralised, decentralised or hybrid? One conclusion of industry 4.0 might be the decentralised production system, that brings flexibility into the supply chain, as well as interactions and collaborations into problem-solving and decision-making. The remaining question is to what extent the production planning process can be decentralised.

How can a simulation model of the production system be generated automatically? Whereas time-prediction to some extent can be done from confirmation data, scheduling and RL need a simulation model that depicts the production system. These models need high effort to build and to maintain. Therefore, a method for automatic generation is needed. Especially for new machines, where there is no historic data available for the start this is still a big challenge in the industrial practice.

How can condition monitoring data be utilized in order to enhance production scheduling and how does it interact with the other two levels of production planning?

In [15,16] it was shown, how condition monitoring data can be integrated within production sequencing and scheduling. However, the topic was considered in an isolated view, where the auxiliary conditions, such as basic dates, where considered as given. Further research should consider a holistic view of condition-based-sequencing and consider its bidirectional interactions with mid-term production planning (on the upper level) and short-term production control (on the lower level).

What role should the human factor play in future CPPPS? As system become more intelligent and to some point autonomous, the role of the human worker will change. On the one hand, one important consequence of decentralisation might be the empowering of the employees. This results in increased responsibilities and skills of the human worker, as he is dealing with highly complex systems. However, on the other hand it is an interesting question how the human worker will react when their decision making scope is limited by an intelligent system. Furthermore, the same question arises for the role of the production planer.

IV. CONCLUSION OUTLOOK

It is obvious, that the PPC will be affected by industry 4.0. As a consequence, the current reference model of PPC, the Aachen PPC model is not well suited towards the needed changes. However, it might be adapted to the requirements of the recent time.

In this concept paper the authors highlighted some key questions that are relevant for future research in the domain of production planning in the paradigm of CPPSs. It was shown, that there do exist different isolated solutions for mid-term planning, production scheduling and production control. When it comes to an integrative consideration of these topics, relevant research is scarce. It can be concluded, that current changes lead to an increase in complexity and it is the responsibility of the research communities to provide solutions which can elevate the production planning to the next level.

REFERENCES

[1] **Abedinnia, H.; Glock, C. H.; Schneider, M. D.**: Machine scheduling in production: A content analysis. Applied Mathematical Modelling, Vol. 50, 2017, pp. 279–299.

[2] **Alaswad, S.; Xiang, Y.**: A review on condition-based maintenance optimization models for stochastically deteriorating system. In: Reliability Engineering & System Safety, Vol. 157, 2017, pp. 54–63

[3] **Bouazza, W.; Sallez, Y.; Beldjilali, B.**: A distributed approach solving partially flexible job-shop scheduling problem with a Q-learning effect, IFAC PaperOnLine, Vol. 50(1), 2017, pp. 15890-15895

[4] **Cheng,Y.; Chen, K.; Sun, H.; Zhang, Y.; Tao, F.**: Data and knowledge mining with big data towards smart production, Journal of Industrial Information Integration, Vol. 9, 2018, pp.1-13.

[5] **Cupek, R.; Ziebinski, A.; Huczala, L.; Erdogan, H.**: Agent-based manufacturing execution systems for short-series production scheduling. Computers in Industry, Vol. 82, 2016, pp. 245-258.

[6] **Das, T.K.; Gosavi, A.; Mahadevan, S; Marchalleck, N.**: Solving semi-Markov decision problems using average

reward reinforcement learning, Management Science, Vol. 45(4), 1999, pp. 560-574.

[7] **Doleschal, D.; Klemmt, A**: Yield integrated scheduling using machine condition parameter. In: Proceedings of the 2015 Winter Simulation Conference, S., 2015, pp. 2953–2963.

[8] **Frost, T., Mc Carthy, J.**: New Level of Performance with dynamic maintenance Management:Achieving Excellence in Four Dimensions. In: *Journal of Maintenance Engineering;* Vol. 1, 2016, pp. 413–424

[9] **Glawar, R.; Ansari, F.; Kardos, CS.; Matyas, K.; Sihn, W.** : Conceptual Design of an Integrated Autonomous Production Control Model in association with a Prescriptive Maintenance Model (PriMa). Procedia CIRP,Vol. 80, 2019, pp. 482-487.

[10] **Gyulai, D.; Kádár, B.; Monostori, L.**: Robust production planning and capacity control for flexible assembly lines, IFAC-Papersonline, Vol. 48, 2015., pp. 2312-2317.

[11] **Gyulai, D.; Pfeiffer, A.; Nick, G.; Gallina, V.:** Lead time prediction in a flow-shop environment with analytical and machine learning approaches, IFAC-Papersonline, Vol 51, pp. 1029-1034

[12] **Irfan, M.; Saad, N.; Ibrahim, R.; Asirvadam, V. S.**: Condition monitoring of induction motors via instantaneous power analysis. In: Journal of Intelligent Manufacturing 28, Vol. 6, 2017, pp. 1259–1267

[13] **Kao, Y.-T.; Dauzère-Pérès, S.; Blue; J. ; Chang, S.-C.**: Impact of integrating equipment health in production scheduling for semiconductor fabrication. Computers & Industrial Engineering, Vol. 120, 2018, pp. 450–459.

[14] **Kara, A.; Dogan, I.**: Reinforcement learning approaches for specifying ordering policies of perishable inventory systems, Expert Systems with Applications, Vol. 91, 2018, pp 150-158.

[15] **Karner, M.; Glawar, R.; Sihn, W.; Kurt, M.**: An industry-oriented approach for machine condition-based production scheduling (in print). In: CIRP Manufacturing Systems Conference 2019.

[16] **Karner, M.; Sihn, W.**: Zustandsüberwachungssysteme zur Entscheidungsunterstützung in der Produktionsplanung. Einsatz von Zustandsüberwachungssystemen zur Entscheidungsunterstützung in der Produktionsplanung der variantenreichen Fertigung. In: GITO-Verlag, 2018.

[17] **Kasakow, G.; Menck, N.; Aurich, J. C.**: Event-driven production planning and control based on individual customer orders. Procedia CIRP, Vol. 57, 2016, pp. 434-438.

[18] **Lingitz, L.; Gallina, V.; Ansari, F.; Gyulai, D.; Pfeiffer, A.; Sihn, W.; Monostori, L.**: Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. Procedia CIRP, Vol. 72, 2018, pp. 1051–1056.

[19] **Mahadevan, S.; Theocharous, G.**: Optimizing production manufacturing uising reinforcement learning, The Eleventh International FLAIRS Conference, AAAI Press, 1998, 99. 372-377.

[20] **Pinedo, M. L.** : Scheduling. Cham: Springer International Publishing, 2016

[21] **Rana, R. ; Oliveira F. S.**: Dynamic pricing plicies for interdependent perishable products or services using reinforcement learning, Expert System with Application, Vol. 42, 2015, pp. 426-436.

[22] **Ringsquandl, M.; Lamparter, S.; Lepratti, R.**: Estimating Processing Times within Context-aware Manufacturing Systems, IFAC-PapersOnLine, Vol. 48, 2015, pp. 2009-2014

[23] **Schuh, G. (ed.).**: ProSense : Ergebnisbericht des BMBF-Verbundprojektes ; hochauflösende Produktionssteuerung auf Basis kybernetischer Unterstützungssysteme und intelligenter Sensorik, 2015.

[24] **Schuh, G.; Salmen, M.; Kelzenberg, C.; Lange, J. de**: . Managing technological entrepreneurship: the engine for economic growth. PICMET'18 : Portland International Conference on Management of Engineering and Technology, 2018.

[25] **Schuh, G.; Stich, V.** : Produktionsplanung und -steuerung 1: Springer Vieweg Wiesbaden, 2012

[26] **Sobottka, T. ; Kamhuber, F. ; Sihn, W.**: Increasing energy efficiency in production environments through an optimized, hybrid simulation-based planning of production and its periphery, Procedia CIRP, Vol. 61, 2017,pp. 440-445.

[27] **Stricker, N.; Kuhnle, A.; Strurm, R.; Friess, S.**: Reinforcement learning for adaptive order dispatching in the semiconductor industry, CIRP Annals, Vol 67, 2018, pp. 511-514

[28] **Waschneck, B; Reischstaller, A.; Belzner, L; Altenmüller, T., Bauernhansl, T.; Knapp, A.; Kyek, A**.: Optimization of global production scheduling with deep reinforcement learning, Procedia CIRP, Vol 51, 2018, pp. 2164-1269

[29] **Wang, Y.C.; Usher, J.M.**: Application of reinforcement learning for agent-based production scheduling, Engineering Applications of Artificial Inteligence, Vol 18., 2005, pp. 73-82

[30] **Yugma, C.; Blue, J.; Dauzère-Pérès, S.; Obeid, A.**: Integration of scheduling and advanced process control in semiconductor manufacturing. Review and outlook. In: *J Sched*, Vol. 18., 2015, pp. 195–205

[31] **Zhang, W.; Dietterich, T.G.**. A reinforcement learning approach to job-shop scheduling. In: IJCAI; vol. 95. 1995, pp. 1114–1120.

[32] **Zhang, J.; Ding, G.; Zou, Y.; Qin, S.; Fu, J.**:. Review of job shop scheduling research and its new perspectives under industry 4.0. In: Journal of Intelligent Manufacturing, Vol. 30(3), pp. 1809-1830.

[33] **Öztürk, A., Kayalıgil, S., Özdemirel, N.E.:** Manufacturing lead time estimation using data mining, In: European Journal of Operational Research 173, 2006, pp. 683-700.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Review of Methodologies for the Assessment of the Technological Capability of RTOs

Fabian Hecklau[1], Florian Kidschun[2], Sokol Tominaj[3], Holger Kohl[4]

[1]Fraunhofer IPK, Berlin, Germany, fabian.hecklau@ipk.fraunhofer.de, 030 39006 158
[2]Fraunhofer IPK, Berlin, Germany, florian.kidschun@ipk.fraunhofer.de, 030 39006 469
[3]Fraunhofer IPK, Berlin, Germany, sokol.tominaj@ipk.fraunhofer.de, 030 39006 158
[4]Technical University of Berlin, Berlin, Germany, holger.kohl@tu-berlin.de, 030 39006 233

Keywords – IMEKO, T10 - certification (products, management systems, persons) and accreditation, technological capability, assessment

I. INTRODUCTION

In a continuously fluctuating, dynamic environment, organisations in all areas must quickly adapt to changes, such as the digital transformation, in order to ensure their sustainable development and growth. Furthermore, the capability of generating and using new technologies effectively and efficiently has a direct impact on the organisations competitiveness and long-term success. Organisations with a high technological capability are able to achieve a higher differentiation through more innovative products and increase efficiency or reduce costs through process innovations.

Moreover, an advancing liberalisation of markets intensifies competition not only for regular businesses, but also for Research and Technology Organisations (RTOs) that cooperate with these businesses. RTOs face the challenge of shorter innovation cycles while at the same time the demand for advanced and market-ready product and process solutions rises.

In order to be successful, RTOs have to pursue technical solutions that eventually lead to state-of-the-art products and services. It is therefore necessary for RTOs to use technological resources and therefore acquire or maintain technological capabilities that make it possible to achieve the aforementioned technical solutions, which ultimately end in high-quality and innovative products. Hence, sustainable success and a competitive edge are the consequence of technological capabilities that are being optimally cultivated and used. [15]

Subsequently, RTOs have to be able to analyse and assess their technological capability, given that an advanced technological capability is imperative for their competitiveness. A tool or method that assesses the technological capability of RTOs in an objective and practical way, while being suitable to the special requirements of RTOs, is needed. While several other authors have previously developed models for the assessment of technological capability in organisations, there is yet to be developed a method that fully meets the specific requirements of RTOs. Therefore, in this paper, existing methodologies will be analysed and reviewed to identify the necessary requirements that a methodology needs to fulfil in order to be applicable for the analysis and assessment of the technological capability of RTOs.

II. RELATED RESULTS IN THE LITERATURE

Having motivated the necessity of a methodology for the analysis and assessment of the technological capability of RTOs, an overview about other existing research in the areas of RTOs in general as well as technological capability of organisations and the assessment of technological capability will be outlined in the following chapter.

RTOs are generally defined as organisations "that mainly provide research and development, technology and innovation services to enterprises, governments and other clients" [4]. Therefore, RTOs carry out activities between the "Technology Readiness Levels" 3 (detection concept of critical functions and/or properties) to 6 (model demonstration of critical functions in relevant environment) [17]. While universities concentrate on teaching and fundamental research, RTOs conduct applied research and use the generated knowledge in industrial innovation and development projects. Moreover, three key work areas of RTOs can be defined [1]:

- Satisfying the demand of public institutions for knowledge-related services
- Providing user- or problem-oriented research to solve societal challenges
- Supporting and developing the technological capacity of the industry

In contrast to the unambiguously definition of RTOs, there is no clear and consensual definition of technological capability in the literature. Thus, various definitions of technological capability and its embedding in a business or RTO context are prevalent. A reason for this may be the different currents within competence research [18]. Moreover, the analysis of related research work found that authors often define the term as a function of their

respective research goals. However, it is noticeable that the definitions mostly use the central terms "ability", "skill", "knowledge" and "resource". For example Figueiredo [5] defines technological capability exclusively as resources needed to generate and manage improvements in product and process organization, equipment and technical projects. Wang et al. [16] also refer to the ability of technological capability to develop and design new products and processes. Other authors such as Panda and Ramanathan [13] define technological capability more comprehensively as a set of functional capabilities that are reflected in the company's performance through various technological activities. In this paper various definitions will be considered.

Currenty, there are multiple methodologies and models for the assessment of technological capability in existence. However, they mainly focus on the application in production companies. An unreflected application of these methologies for the analysis and assessment of the technological capability of RTOs is not possible without adaptation to the specific requirements.

In addition, some of the methodologies mainly focus only on selected aspects (e.g. focus on resources - example: „Technology ATLAS" [2] or focus on technological capabilities in a sense of intangible skills and knowledge and largely disregard resources [14] – example: Phaal's „Technology Management Process Assessment"). The existing methodologies will be analysed and assessed according the the requirements of RTOs in this paper.

III. DESCRIPTION OF THE METHOD

To identify the requirements for the analysis and assessment of the technological capability of RTOs, the following approach will be followed within this paper. Firstly, research and development (R&D), technology, technological capability as well as RTOs will be defined on the basis of previous research work. Secondly, requirements of a methodology that could be used in the context of RTOs and was also developed in previous research will be presented. Thirdly, existing methodologies and their assessment according to the pre-set requirements will be elaborated. Lastly, the results will be discussed and a conclusion whether the existing methodologies are suitable for the use in RTOs will be given.

IV. RESULTS AND DISCUSSIONS

As mentioned above, requirements to the specific use of technological capability methodologies have been derived in a previous paper. Therefore, these requirements are introduced in the following table 1. Existing methodologies will be analysed and reviewed on the basis of these requirements.

Table 1: Requirements for a methodology for the analysis and assessment of the technological capability of RTOs

Require-ments	**Description**
Applicability in Research and Technology Organi-sations	
Project focus	RTOs perform their work in projects, so that it must be possible to analyse and evaluate projects using the method.
Openness	Various RTOs work on a very broad spectrum of topics, which must not restrict the application of the method.
Applied re-search	RTOs carry out activities between the "Technology Readiness Levels" 3 (detection concept of critical func-tions and/or properties) to 6 (model demonstration of critical functions in relevant environment), which have to be analysed by using the method.
Analysis & evaluability of technological capability	
Technology-independent	Different technologies are used in ap-plied research, so that the application of the method cannot be limited to in-dividual technologies.
Focus on technological resources & capabilities	The implementation of applied re-search activities requires technologi-cal resources and capabilities which must be analysed and evaluated using the method.
Targeted application	
Comprehensi-bility	The method must be understandable so that it can be applied with little preparation time.
Low effort	The application of the method must be accompanied by little effort, so that the daily business of the organi-sation to be analysed and evaluated is not disturbed more than necessary.
Degrees of freedom in application	Due to the wide range of topics that RTOs deal with using a variety of different technologies, a certain de-gree of freedom in the application of the method must be possible without distorting the results.
Practical results	
Comparability	The method of evaluation must be standardised in order to ensure the comparability of evaluation results between RTOs.

Implementation- oriented	The application of the method is intended to identify weaknesses in the RTOs, identify potential for improvement and provide practical assistance.

The existing methodologies that could be considered as suitable approaches for the analysis and assessment of the technological capability of RTOs will be described in the following paragraphs. As a conclusion the following methodologies will be assessed according to the above mentioned requirements that need to be fulfilled.

The *Technology ATLAS*, developed by the Asian and Pacific Center for Technology Transfer [2], assesses technological capability on the basis of four factors that are ranked according to their degree of maturity from 1 to 6. The factor "Technoware" assesses physical resources such as machines and tools. The factor "Infoware" assesses the existing knowledge of using hardware for the organisation's activities and processes. "Humanware" assesses the human capabilities and skills level that is needed to use hardware and infoware for the organisation's activities. "Orgaware" describes organisational and managerial structures to coordinate the aforementioned factors and the organisation's activities. This model is inappropriate for the use in RTO's due to the lack of concise definitions of each degree of maturity. It can therefore not be used to compare RTO's.

The *Technology Audit Model* by Garcia-Arreola [8] is an extensive three-layer (categories, valuation area, elements) model for the assessment of technological capability. On the highest layer, six categories are defined:

- Technological Environment
- Technologies Categorization
- Markets and Competitors
- Innovation Process
- Value-added Functions
- Acquisition and Exploitation of Technology

These six categories comprise 20 valuation areas which are then subdivided into 43 elements. A questionnaire serves the sake of assessment and is conducted by internal experts or auditors by the use of a five-step likert-scale.

Despite its extensiveness, the Technology Audit Model exhibits several weaknesses. As the model tries to assess the technological capability of organisations, it must incorporate all elements that are directly or indirectly involved in the relation of organisations and the use of technology. This is very complex in the case of large-scale organisations and therefore results in a disproportionally high effort. Furthermore, non-technological assessment criteria that contribute to an organisation's success such as markets and competitors, innovation processes etc. are also analysed. As those elements are not in the focus of the methodology for the analysis and assessment the Technology Audit is not meeting the above mentioned requirements.

The *Technology Management Process Assessment* by Phaal et al. [14] is a top-down assessment model that proceeds in three phases that are conducted in employee-workshops. In the first phase, the "Strategic Overview", the organisation is subdivided in technology areas and business units. The impact of each technology area on the business units is then assessed according to its delivered value to the unit, effort that is made and risk using a four-point scale ("high", "medium", "low", "not significant"). Naturally, these factors correlate positively. If they fail to do so, the technology area is then evaluated in the second phase, the "Process Overview". In this phase key technologies as well as current and future technology management activities of the technology areas identified in phase 1 are mapped. In order to facilitate the identification of activities, workshop participants are encouraged to identify significant events over a period of time (e.g. the last five years) related to the respective key technologies. The associated activities will then be categorised using Gregory's Five-process model (identification, selection, acquisition, use and protection of technologies, cf. [8]) and used as a basis for assessing technology management effectiveness on the basis of inputs, processes and outputs. Each of these elements is linked to a statement:

- Input: "The requirements for this activity were always clearly defined".
- Process: "The activity was always well controlled"
- Output: "The results of the activity were always used"

These are then evaluated on a five point scale (1 ="agree entirely",…,5="disagree at all"). In the third phase, the "Process Investigation", specific process areas are mapped in detail to identify good-practices and improvement-worthy areas. The approach is entirely qualitatively and rather concentrates on functions and strengths and weaknesses of technology management than on technology contents. It is therefore only conditionally usable to assess technological capability. [7]

The *Technological Capability Audit Tool* [19] is an assessment tool that aims to combine knowledge of key competences for the realisation of technological innovations with development stages of technological capabilities that enable an organisation to select and deploy technologies to gain a strategic competitive advantage. It uses nine main components: "Awareness", "Searching", "Core Competencies", "Strategy", "Assessment", "Acquisition", "Implementation, Absorption and Operation", "Learning", "Exploiting external linkages and incentives" that are self-explaining to build a survey to assess an organisations technological capability. 4 different statues are used to evaluate the condition of a main component: "Unaware or passive", "Reactive", "Strategic", "Creative". These four types allow each major component to be assigned a specific

number of points. Although the result of the assessment delivers numeric results, the model leads to a highly subjective evaluation. It is therefore not appropriate due to the comparability aspect of the requirements.

The following table 2 summarizes the four analysed methodologies and reviews the suitability of those methodologies for the application for the analysis and assessment of the technological capability of RTOs according to the requirements described in table 1.

Table 2: Summary of analysed methodologies assessed by the requirements for a methodology for the analysis and assessment of the technological capability of RTOs

Requirements	**Technology ATLAS**	**Technology Audit Model**	**Technology Mgmt. Process Assessment**	**Technology Capability Audit Tool**
Applicability in Research and Technology Organisations				
Project focus	◕	◑	◑	◑
Openness	○	●	○	●
Applied research	◔	◔	●	◔
Analysis & evaluability of technological capability				
Technology-independent	◕	◕	◕	◕
Focus on technological resources & capabilities	◕	◑	◑	◑
Targeted application				
Comprehensibility	◔	◑	◕	◕
Low effort	◔	◔	◔	◔
Degrees of freedom in application	◔	◔	◔	◔
Practical results				
Comparability	◕	◕	◔	◑
Implementation-oriented	◔	◔	◔	◔

Legend:

○ The model / method does not meet the requirement.

◔ The model / method hardly meets the requirement.

◑ The model / method partially meets the requirement.

◕ The model / method meets the requirement mostly.

● The model / method completely fulfils the requirement.

V. CONCLUSIONS AND OUTLOOK

In a first step the requirements for the analysis and assessment of the technological capability of RTOs have been identified based on previous research work of the authors as well as the definition and interpretation of the terms "RTO" and "technological capability". In a second step, possible methodologies for the analysis and assessment of the technological capability have been identified and briefly been described. As a summary, the identified methodologies have been assessed according to the identified requirements that needs to be fulfilled. As a result, no of the analysed methodologies fully meet all the defined requirements and therefore, the need for the creation of a new methodology for the analysis and assessment of the technological capability of RTOs could be pointed out within this paper. Further research work of the authors will focus on the creation of a methodology that fully meets the requirements described in this paper.

REFERENCES

[1] **Arnold, E.; Clark, J.; Jávorka, Z.:** Impacts of European RTOs: A Study of Social and Economic Impacts of Research and Technology Organisations. *A Report to EARTO. Hg. v. Technopolis Group*, 2010

[2] **Asian Center for Transfer of Technology (APCTT)**: Atlas Technology: A framework for technology planning, 1989.

[3] **Birke, F.:** Technologische Kompetenz und Erfolg junger Unternehmen. Wiesbaden: Gabler, 2011.

[4] **EARTO:** Knowing your innovation ecosystem actors: data on European RTOs. *European Association of Research and Technology Organization*, 2015.

[5] **Figueiredo, P. N.:** Does technological learning pay off? Inter-firm differences in technological capability-accumulation paths and operational performance improvement. *Research Policy* (31), 2002, pp.73–94.

[6] **Figueiredo, P. N.:** Beyond technological catch-up. An empirical investigation of further innovative capability accumulation outcomes in latecomer firms with evidence

from Brazil, *Journal of Engineering and Technology Management* 31, 2014, pp.73–102.

[7] **Garcia-Arreola, J.:** Technology Effectiveness Audit Model (TEAM): A Framework for Technology Auditing. Dissertation. University of Miami, Miami, FL, 1996.

[8] **Gregory, M.J.:** Technology management- a process approach. *Proc Instn Mech Engrs* (029), 1995, pp. 347–356.

[9] **Kröll, M.:** Methode zur Technologiebewertung für eine ergebnisorientierte Produktentwicklung. Dissertation. Universität Stuttgart, Stuttgart. Institut für Arbeitswissenschaft und Technologiemanagement. 2007.

[10] **Lee, M.; Lee, S.:** Evaluating Internal Technological Capabilities in Energy Companies. *Energies* 9 (3), 2016, pp. 145.

[11] **Mohammad, A. P.; Razaee, S.; Shayegh, F.; Torabi, F.:** A Model for Technological Capability Assessment in R&D Centers. *Proceedings of the 14th international Oil, Gas & Petrochemical Congress*, 2010.

[12] **Mohammadi, M.; Elyasi, M.; Kiasari**, **M.M.:** . Developing A Model For Technological Capability Assessment — Case Of Automotive Parts Manufacturers In Iran, *International Journal of Innovation and Technology Management)*, World Scientific Publishing Co. Pte. Ltd., vol. 11(02), 2014, pp. 1-19.

[13] **Panda, H.; Ramanathan, K.:** Technological capability assessment of a firm in the electricity sector. *Technovation* 16 (10), 1996, pp. 561–588.

[14] **Phaal, R.; Farrukh, C.J.P.; Probert, D. R.:** Technology management process assessment. A case study, *Int Jrnl of Op & Prod Mgmt* 21 (8), 2001, pp. 1116–1132.

[15] **Wang, C.H.; Lu, I.Y.; & Chie, C.**: Evaluating firm technological innovation capability under uncertainty. *Technovation.* 28., 2008, pp.349-363.

[16] **Wang, Y; Lo, H.-P.; Zhang, Q.; Xue, Y.:** How technological capability influences business performance. *Jrnl of Tech Mgmt China* 1 (1), 2006, pp. 27–52.

[17] **Martinez-Vela, C.:** Benchmarking Research and Technology Organizations (RTOs): A Comparative Analysis. Cambridge, 2016

[18] **Lierow, M. A. (2006):** Competence-Building und Internationalisierungserfolg. Wiesbaden: DUV.

[19] **Rush, H.; Bessant, J.; Hobday, M. (2007):** Assessing the technological capabilities of firms: developing a policy tool. In: R & D Management (37), S. 221–236.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Industry 4.0 in Germany, Austria and Hungary: interpretation, strategies and readiness models

Gábor Nick[1,3], Viola Gallina[2], Ádám Szaller[1,4], Tamás Várgedő[1], Andreas Schumacher[2]

[1]Centre of Excellence in Production Informatics and Control,
Institute for Computer Science and Control, Hungarian Academy of Sciences, Budapest, Hungary
(e-mail: {gabor.nick, adam.szaller, tamas.vargedo}@sztaki.mta.hu, phone: +36 1 279 {6139, 6262, 6110}).
[2]Fraunhofer Austria Research GmbH, Wien, Austria, (e-mail:
{viola.gallina, andreas.schumacher}@fraunhofer.at, phone:{+43 676 888 616 46, +43-1-58801-33046})
[3]Széchenyi István University, Research Center of Vehicle Industry, Győr, Hungary
[4]University of Technology and Economics, Department of Manufacturing Science and Engineering, Budapest, Hungary

***Abstract* – Global economy increasingly responds to trends of industrial digitization and start to develop their own digitalization strategies e.i. through application of Industry 4.0 concepts. Thereby, initiatives and developments are often carried out on a company-, industry- or country-level, which can result in a lack of international co-operation and utilization of synergy-effects. To analyse the Industry 4.0 development status on a macro level and to enhance transnational developments, in the first part of this study we introduce the national Industry 4.0 platforms of three EU member states – the frontrunner Germany, the solidly rising Austria and Hungary that is in the process of catching up. Our analysis are carried out in four dimensions: economic background, country-specific Industry 4.0 interpretation, and local platforms as well as initiatives. To analyse the actual Industry 4.0 development status of industrial enterprises and of development platforms, existing Industry 4.0 maturity models form all three countries are presented and compared. Our analysis helped to capture the different phases of digitization and Industry 4.0 in Germany, Austria and Hungary in regard to defined goals and strategies and further of resulting problems with respect to Industry 4.0. Building on our comparison-studies and captured problems, decision makers should be enabled to carry out transnational Industry 4.0 projects and initiatives more effectively.**

Keywords – Industry 4.0, Digitization, Maturity models

I. INTRODUCTION

A. The economics of Industry 4.0

Roland Berger states that industrial production plays a central role in the EU economy, and industry is the main driver of research, innovation, productivity, job creation and export. Industry creates 80 % of the European innovation and 75 % of the exports. Through its relationship with the service sector, industry can be regarded as Europe's social and economic engine. Still, the European industry has lost a lot of jobs in the last decade, while it is facing tough competition from the emerging economy. The ghost of "deindustrialisation" currently haunting most European governments and the European Commission is galvanizing them into action. [1][2]. The contribution of production to GDP at EU level was barely more than 15 % in 2013 (currently around 17 %), which represents a major stepping stone to the US and Southeast Asian industrial regions, and is far from the development strategy adopted in 2012: the target is 20 % to by 2020. All these fact have to be evaluated in the background of the fourth industrial revolution which is already on its way. According to Roland Berger, *"revolutions are fast, disruptive and destructive. And there is no going back. Industry 4.0 will be an answer to the challenges lying ahead."* This has led to the reindustrialization strategy announced by the EU. In its implementation, the Industry 4.0 organisations in the member states, the establishment of which is strongly recommended by the European Commission are deemed to play a key role. The Roland Berger study [1] points out that the Return of Capital Employed (ROCE) is the proper measure of progress both at individual company and country level. ROCE is the

economic index governmental industry development policies may have a direct influence on. ROCE is 15 % on average in Europe. When comparing the different strtegic options we find three basic alternatives (Figure 1):

(1) Automation (not to be confused with Industry 4.0): increases capital employed through investments. Profits increase due to automated activities, with ROCE remaining unchanged.

(2) Obsolescence: investments tend to drop to a level below depreciation due to decreasing profits. Asset turnover increases artificially enabling ROCE to remain constant.

(3) Industry 4.0: industrial policies implemented pursue the main objective of increased competitiveness by growth in value added production. Additionally, high margins and flexible production result in higher ROCE.

Alternatives (1) and (2) are iso-ROCE policies, thus only option (3) represents "good" progress. The second chart of Figure x shows the evaluation of some selected countries' ROCE over the period 2000 to 2014. Germany's outstanding performance in this context is evident.

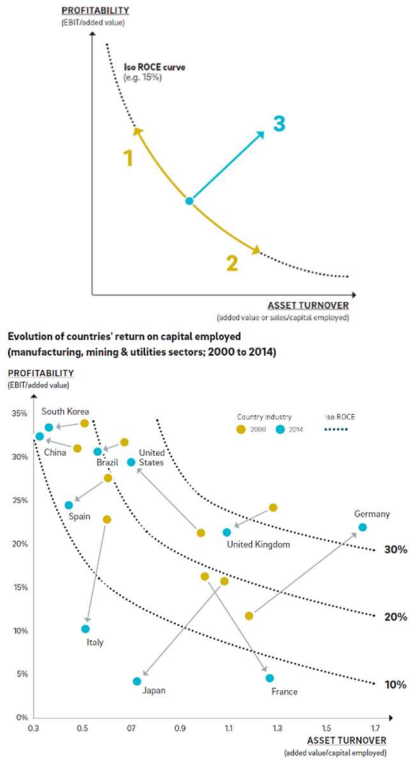


Figure 1: Industry development strategic options [3]

B. Competitiveness and maturity qualified by indices

In the literature different indices and rankings are used to qualify the individual countries' ecosystem in terms of the digitisation status and Industry 4.0 maturity, starting with the most comprehensive and global ones, and arriving at the specific level of the local industrial companies. Thus, the annual Global Competitiveness Report of the World Economic Forum will be discussed first which is a research product on the individual countries' macroeconomic level. The research work done by the EU is represented by the prestigious Digital Economy and Society Index, DESI and the European Innovation Scoreboard, EIS. The most frequently cited source of Industry 4.0 readiness evaluation is the annually published Roland Berger Industry 4.0 Readiness Index [1]-[3]. Although the companies and their digital transformation are always in the focus of Industry 4.0, the issue can only be addressed in conjunction with their environment where the individuals, the social organisations and the institutions of the State have a dominant role. The European Union deals in many aspects with industry policy. It gives directives and defines expectations towards its member states in respect to the digitalisation of economy and society.

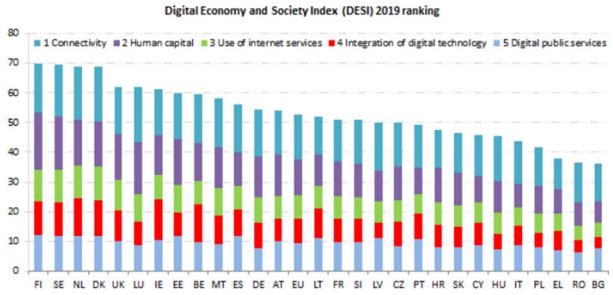


Figure 2. DESI 2019 ranking

It is evident form the chart that Germany and Austria are very close to each other in the overall ranking and Hungary is still lagging behind with a rating under the EU average. We will come back to some of the details when discussing the individual counries.

C. Industry 4.0 readiness models

In the literature, several Industry 4.0 readiness evaluation methods and maturity models can be found. In a recent review of these models, the authors enumerate 15 different maturity models in terms of the method they use, the focus, and also the gaps found in them [4]. These models assess the Industry 4.0 maturity in different dimensions that contain questions or maturity items. The organization or company is evaluated based on these elements by choosing an appropriate level of scale, which contains in general four to ten levels. In terms of these dimensions, the most common ones are the following: strategy and organization, technology, IT, smart factory, smart products, data utilization and employees. In some models, additional dimensions appear, too, for example security policies [5], performance [6] and customers [7]. The number of maturity items vary on a wide scale, e.g. the VDMA's online self-assessment questionnaire [8] contains altogether 18 items, but in another maturity model, the authors present a survey with 65 questions [7].

II. GERMANY – AT THE FOREFRONT OF DEVELOPMENT

A. Economic background

Germany is globally the only country that has managed to significantly improve the industry's ROCE index over the last 15 years, mainly due to its conscious industrial development program. Despite a slight decline in employment (9%), industrial surplus increased to 80%, and industry-generated profits by 158%. Investment and depreciation remained broadly unchanged over the period, while asset utilization rates improved and equipment utilization rose from 85% in 1998 to 95% in 2014. This shows the importance of organization, conscious behaviour and operation. Capital invested, while remaining roughly constant, produces much more than 15 years ago. This is the German industrial wonder.

B. Industry 4.0 interpretation

The solution was the recognition of the potential of Industry 4.0, which was first introduced at the 2011 Hanover Fair, and has since become the supplier of a vast number of innovative Industry 4.0 solutions. In Germany's strategy, Industry 4.0 solutions provide the means to maintain its dominance in global markets and keep its industrial production constant. Consequently, Germany's Industry 4.0 strategy is both defensive - to maintain domestic production, to deal more flexibly with crises on international markets - and offensive - to ensure that expertise and know-how remain in Germany for the export model to function [9]. In Germany, much attention is given to the societal aspects of Industry 4.0: there are various calls, too, focusing on the integration of humans in digital and AI technologies as well as the social impact of those technologies for human life.

C. Platform, initiatives

The long-term strategy developed by the federal government and supported by Plattform Industrie 4.0 is based on five pillars [10]:

- identifying and prioritizing research and innovation topics that ensure the economy's performance, quality of life, and wealth sources;
- promotion of national and international co-operation of knowledge-raising actors in the economy to strengthen their innovation potentia;
- launching of different programs to boost innovation dynamics, with the aim of increasing SMEs' contribution to high R&D costs;
- and creating an innovation-friendly framework and to promote social participation.

A survey conducted by German company managers in 2014 [11] highlighted areas where political support is expected for the Industry 4.0 initiative. On the basis of the results, the continuous training of the qualified workforce, the promotion of international standardization, the creation of competitive data protection laws and the granting of tax incentives for corporate investments are prominent. In addition to Plattform Industrie 4.0, the government has recently introduced a number of significant measures to promote the digital transformation of industry. These include the preparation of SMEs, the provision of technology transfer to small and medium-sized enterprises in connection with Industry 4.0, which is reinforced by the establishment of test laboratories and centers of excellence. The national IUNO project, which focuses on information technology security in the Industry 4.0 process, should also be highlighted. The project aims to address IT security challenges in an industrial application area, thus removing the fears of SMEs of the economic risks inherent in digital transformation.

Germany performs well in most DESI dimensions, especially in the digital skills, where it is among the top performers despite the fact that there is still a lack of ICT specialists in the country. Interestingly, as for the integration of digital technology by businesses, it ranks only slightly above the EU average. Many of the initiatives for the digitisation of the country target SMEs. These include the Mittelstand 4.0 competence centres and the Go-Digital programme.

D. Measurement of Industry 4.0 readiness

The VDMA's online self-assessment questionnaire [8][1] contains altogether 18 items separated into 6 dimensions, where the responding organization has to choose from options corresponding to 6 maturity levels (outsider, beginner, intermediate, experienced, expert, top performer) in case of each question. The dimensions are strategy & organization, staff, smart factory, smart operations, smart products and data driven services. The questionnaire was developed with the aim of self-assessment, thus an organization can check its status regarding Industry 4.0 and take the first step towards digitization.

III. AUSTRIA – SOARING ECONOMY

A. Economic backgorund

On the basis of the relatively low inflation and unemployment rate and the extensive network of social benefits, Austria is among the leading countries in the EU. Following the halt in the economic growth of 2012-2014, the Austrian economy has been growing at an increasing pace since 2015 (1.1% in 2015, 1.5% in 2016 and 2.9% in 2017) and this trend is likely to continue in the coming years [12]. Austria's economic development is driven by increasing industrial production: in June 2017, industrial production was nearly 5% higher than a year earlier. One of the engines of industrial growth is strong export activity (€ 141.9 billion in 2017, up 8.2% from the previous year). Another key to the growth of industrial production is the

competitiveness of Austrian products, which can be explained by the fact that Austria traditionally spends a lot on investment (Austria 23% of GDP in 2016 compared to the only 20% in Germany). According to a survey conducted in 2015, Austrian companies plan to invest 3.8% of their annual turnover on Industry 4.0 solutions on average by 2020, from which they expect an average annual 3.7% efficiency increase and a 2.6% reduction in production costs [13].
In addition to investment intensity, it is also important to mention the high level and continuous increase in R&D expenditure that has been higher than the European Union target of 3% of GDP to be attained by 2020. In 2017, EUR 11.4 billion was spent on R&D, representing 3.16% of GDP, estimated at 3.76% by 2020 [14].

B. Industry 4.0 interpretation

The Austrian definition of the term Industry 4.0 is essentially the same as the German one, with the difference in the approach that social effects of digitalisation are more emphasized. This way, the Austrian Industry 4.0 interpretation often involves topics such as creating new customer values, developing new business models, opening new markets, and changing the structure of employment.

C. Platform, initiatives

The Austrian Ministry of Transport, Innovation and Technology was an active initiator of the Austrian digitalization, recognizing the importance of the Internet in 2014, announced the launch of the Broadband Strategy 2020 (Breitbandstrategie 2020), which aims to make Austria one of the leading ICT countries with spreading the Internet gradually across the country (99% coverage). The government has launched targeted programs such as Production of the Future (Produktion der Zukunft), which has been launched 28 times since 2011, to ensure the long-term competitiveness of Austrian companies and to secure jobs in Austria. The program has strategic objectives such as promoting product innovation (such as supporting the establishment of demonstration and logistics systems), building the research expertise of research institutes (e.g. doctoral and professor scholarships), and supporting European and international cooperation and networks. Under these and similar programs, the Austrian Research Support Office supports € 120 billion annually in initiatives related to Industry 4.0.
The Austrian national industry 4.0 platform is an alliance of important social, economic, political and scientific actors. It was launched in 2015 and focuses on the following activities: ensuring the dynamic development of the Austrian production sector; promoting research, innovation and training; developing highly skilled workforce; and to achieve a high level of employment.

Austria has increased its DESI score only slightly due to a limited performance in some of the DESI dimensions: there is a growing lack of skilled IT workers in the economy and the country performs below average in Connectivity, Use of Internet services and integration of digital technologies. The Digital Roadmap Austria programme was launched early 2017 by the government.In Human capital, however, Austria ranks among the top EU countries, a relatively large proportion of the population has at least basic or above average digital skills.

D. Measurement of Industry 4.0 readiness

In [7] the authors stated that former maturity models are mainly technology focused, and neglect other organizational dimensions, therefore, they present a model, which includes these aspects as well. They define indicators with in 9 organizational dimensions, provide a model with 62 questions (maturity items) measured in 5 dimensions to measure the Industry 4.0 readiness of manufacturing enterprises – used in various cases in Austria. They also present a case study with an Austrian company. In [15], the authors introduce a 10-step Industry 4.0 realization model, which builds on they maturity model and Industry 4.0 strategic guidance [16].

IV. HUNGARY – AT THE GATEWAY TO DIGITIZATION

A. Economic background and Industrial Policy

In order to meet the challenges of Industry 4.0 and taking into account local specificities, Hungary has also developed its own reindustrialization strategy and its complex set of tools, which the government identified as the Strategic Program for Innovative Industrial Development, named after János Irinyi [17]. The ambitious goal of the Irinyi Industrial Policy Plan is for Hungary to become one of the strongest industrial performers in the EU. To attain this goal, innovation-driven economy, highly skilled and active workforce, greater added value, export-oriented value chains, and sustained balanced development are necessary. In line with the European Union's industrial policy, the focus of the Iriny Plan is on machinery and vehicle manufacturing, the health and green economy, the food industry, the defense industry and the ICT sector. The state takes an active role in innovation processes in a number of areas – with orders and investments. The main aim is the following: the GDP growth rate up to 2020 and beyond, should be in the range of 4-5% annually if industrial growth rates remain similar to those of the previous years. To reach the target for 2020 in connection with GDP increase rate – if the rate of growth remains the same as in previous years – an annual increase in industrial production of at least 7% should be achieved.

The implementation of the sectoral development strategies is being carried out by the Irinyi Industrial Policy Plan's 5 horizontal aspects: the application of new and digital technologies, the production of energy- and

material-efficient devices, the elimination of territorial inequalities, the expansion of employment and the efficient use of domestic resources are the central issues. This requires a more conscious entrepreneurial approach, adequate standardization, transparent regulations, and the establishment of predictable, regulated support systems (non-refundable grants or combined loan programs) as a prerequisite.

A significant proportion, 99,1% of enterprises registered in Hungary are SMEs, which is much higher than the EU average. Their role in employment is also higher than in the EU, while their capital strength, export capacity and contribution to GDP are below the EU average. SMEs represent a very significant economic factor at local level, their development is highly important, thus the development strategy has targeted interventions at all stages of the business life cycle (start-up, growing, mature, renewable, declining) and the associated financial resources. Knowledge is greatly appreciated, the accumulation of knowledge in SMEs and their incorporation into high-tech processes is an important objective.

B. Industry 4.0 interpretation

In Hungary, the term Industry 4.0 stands for the fourth industrial revolution based on cyber-physical systems, i.e. the formerly never seen integration of the physical and virtual worlds, and represents a new level of organising and controlling the entire value chain across product lifecycles. This cycle focusses on increasingly personalised customer wishes and extends from the concept to the order, development, production, and shipping of a product to the end customer and ultimately to its recycling, including all associated services.

The foundation is the real-time availability of all relevant information through the integration of all objects in the value chain and the capacity to determine the optimal value flow at any time from the data. The interconnection of people, objects, and systems produces dynamic, real-time-optimised, self-organising, cross-enterprise value-adding networks that can be optimised according to various criteria such as cost, availability, and resource consumption.

C. Platform, initiatives

The Industry 4.0 strategy is based on 5 pillars in Hungary. Although the whole pillar structure is strongly permeated by the government with its stimulating, regulatory activity and the operation of its institutional system, we do not consider the government sector itself to be an independent pillar. The pillars are: digitization and business development, manufacturing and logistics, labour market, research and development and innovation, and the ecosystem. The Hungarian platform was founded in 2016 with the full support of the relevant Ministry, it is operated as a legal entity called Industry 4.0 National Technology Platform Association (NTP). The participants are Hungarian Academy of Sciences' Institute for Computer Science and Control and over 60 economic actors. The main goal of NTP is to stimulate the exchange of information and development in the key areas of Industry 4.0 and to strengthen the competitiveness position of the national economy. Unfortunately, the development of the Hungarian innovation system and the culture of cooperation are not yet fully prepared for the effective application of this type of partnership model. There is a serious problem with the lack of trust capital needed to communicate between key players.

Hungary performs best in the broadband Connectivity dimension of DESI. However, it has not managed to improve its position in the overall ranking. In Human capital, although it has a high proportion of ICT graduates and a close to average share of IT specialists, there is room for improvement: Hungary is unfortunately among the worst performing EU countries in the integration of digital etchnology in businesses. Among the government initiatives launched to close the existing gaps, the National Info-communication Strategy 2014-2020, the Modern Enterprises Programme and the recently founded Artificial Intelligence Coalition should be inter alia highlighted.

D. Measurement of Industry 4.0 readiness

In 2016, NTP launched its first questionnaire project that explores the Hungarian Industry 4.0 ecosystem. The first results are published in [19], where the authors present the structure of the questionnaire (containing 99 questions) in detail, and the results in connection to data application in Hungarian organizations. The questionnaire comprises three main parts: firstly, some general questions with the aim of collecting management and statistical data about the organization. Secondly, company level questions about their individual Industry 4.0 chaarcteristics and capabilities (macro level). Lastly, relevant issues regarding Industry 4.0 with respect to the national economy policy, with the aim of having a comprehensive understanding of the situation in Hungary. This way, it provides results related to the Hungarian national economy as well, not only about one organization, in clear difference to the German and Austrian Indusrty 4.0 readiness models. The empirical research conducted when processing the responses aimed to create a map of the Hungarian Industry 4.0 ecosystem. In the paper, the main goal was to explore to what extent specific Industry 4.0 criteria appear in the operation and business model of Hungarian companies. What are the differences in the expectations of individual companies towards the national economy policy? In order to get an adequate picture on both the current and future situation, it is necessary to assess the needs and expectations of the industrial digital ecosystem in terms of strategic economic governance and also to explore the current status of Industry 4.0 awareness, acceptance and implementation at the individual (company, academy,

social organization) ecosystem level and the national economy level, as well. The open attitude of a company for R&D&I cooperation with other actors, the competitiveness potential and conditions, the situation of education and training are also important aspects. The persistent challenges to human resources, the local expectations and effects of industrial digitisation, and above all, the unique Industry 4.0 capabilities of companies are taken into consideration here, too.

The focus of the questionnaire is twofold as previously described: assessing the individual abilities and the level of readiness of enterprises at *micro* level, and gaining a comprehensive understanding of the situation in Hungary at *macro* level. According to the survey, the majority of industrial companies in Hungary understood the importance of this fact: 78% of them had at least partial data collection both on the production and the product usage throughout its life time, primarily for quality control and production statistics purposes. Although data is being collected, in most cases its actual use has not yet become the integral part of the manufacturing and production processes. Based on the survey, Hungarian companies are striving to use the collected product data, but they are yet far from fully utilizing its revenue generating potential through additional services developed on this basis. They are thus ahead of a long familiarisation, technical development and innovation process affecting their entire operation in terms of business model and integration of the new technologies.

V. CONCLUSIONS AND OUTLOOK

In the paper, the authors present the Industry 4.0 approaches in Germany, Austria and Hungary. The countries are investigated and compared through several aspects: economic background, country-specific Industry 4.0 interpretation, local platforms and initiatives, and Industry 4.0 maturity models. The goal for the countries at different levels of digitization is different: for Hungary, the aim is to catch up the frontrunners and join the international value chain. For Germany, the most important thing is to preserve its leading role by following an offensive and defensive strategy at the same time. Austria is in a prosperous phase of development, the main goal for the country is to ensure the dynamic development of the industry. The main benefit of our work is that understanding the differences in the approach and strategies between the three countries as consequences of country specific facts enables us to use the lessons learnt and adapt them in the the project "Centre of Excellence in Production Informatics and Control" (EPIC) funded by the European Union's Horizon 2020 research and innovation programme, in the definition of the business and operational objectives and rules of the said Centre.

As one of the most significant outcomes of our in-depth analysis, we are going to develop our proprietary Industry 4.0 maturity model, to be iteratively tested in a proof-of-concept process with real industrial clients. As the next steps of our research, the position of China, and the other two from the Triad countries – USA and Japan – is planned to be explored in a similar way.

ACKNOWLEDGEMENT

This research has been co-supported by the GINOP-2.3.2-15-2016-00002, ÚNKP FIKP 2017 and the NKFIA ED_18-22018-0006 grants of Hungary.

REFERENCES

[1] **Roland Berger**: Industry 4.0 – The new industrial revolution, How Europe will succeed. *Think Act March*, Munich, 2014

[2] **Roland Berger**: Industry 4.0 – The role of Switzerland within an European manufacturing revolution, *Think Act March*, Zurich, 2015

[3] **Roland Berger**: The Industrie 4.0 Transition Quantified – *Think Act*, Munich, 2016

[4] **Mittal, S., Ahmad Kahn, M., Romero, D., Wuest, T.** A critical review of smart manufacturing & Industry 4.0 maturity models: Implications for small and medium-sized enterprises. Journal of Manufacturing Systems, 2018, vol. 49, pp. 194-214.

[5] **Rockwell Automation**: The connected enterprise maturity model, 2014

[6] **Jung, K., Kulvatunyou, B., Choi, S., Brundage, M. P.** An overview of a smart manufacturing system readiness assessment. APMS 2016. IFIP International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology, 2016, vol. 488, pp. 705-712.

[7] **Schumacher, A, Erol, S, Sihn, W.** A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises. Procedia CIRP 2016, 52: pp. 161-166.

[8] **VDMA:** Industrie 4.0 Readiness. *VDMA's IMPULS-Stiftung*, Aachen, Köln, 2015

[9] **Huawei-HRI:** *Industrie 4.0 im internationalem Vergleich*, Handelsblatt R. Institute, 2016

[10] *Digital Economy and Society Index,* European Union, https://ec.europa.eu/digital-single-market/en/desi, Downloaded: 2019.06.15.

[11] **May-Strobl, E., Welter, F.,** *Das Zukunftspanel Mittelstand - Herausforderungen aus Unternehmersicht*, IfM Bonn: IfM-Materialien Nr. 239, 2015

[12] www.wko.at, accessed on 2019. 04. 30.

[13] www.pwc.at, accessed on 2019. 04. 30.

[14] www.scence.orf.at, accessed on 2019. 04. 30.

[15] **Schumacher, A, Nemeth, T, Sihn, W.** Roadmapping towards industrial digitalization based on an Industry 4.0 maturity model for manufacturing enterprises. Procedia CIRP 2018, 79: pp. 409-414.

[16] **Schumacher, A, Nemeth, T, Sihn, W.** Roadmapping towards industrial digitalization based on an Industry 4.0 maturity model for manufacturing enterprises. Procedia CIRP 2018, 79: pp. 409-414.

[17] **Irinyi Industrial Policy Plan** (Irinyi terv, in Hungarian) http://www.kormany.hu/download/d/c1/b0000/Irinyi-terv.pdf , Downloaded 2017.02.06.

[18] www.plattformindustrie40.at, accessed on 2019. 04. 30.

[19] **Nick, G., Szaller, Á., Bergmann, J., Várgedő, T.** Industry 4.0 readiness in Hungary: model, and the first results in connection to data application. IFAC MIM 2019, accepted.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Application of modern software engineering principles for continuous quality (CQ) management in research

Sascha Eichstädt[1], Bram van der Waaij[2], Björn Ludwig[3], Erik Langius[4]

[1]Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany, sascha.eichstaedt@ptb.de, +49 30 3481 2008
[2]TNO, Groningen, Netherlands, bram.vanderwaaji@tno.nl
[3]Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany, bjoern.ludwig@ptb.de, +49 30 3481 7625
[4]TNO, Groningen, Netherlands, erik.langius@tno.nl

***Abstract* – Transparency of changes, automated testing of new developments against documented criteria, combination of developments from different sources and automated publication of validated results are key aspects of modern collaborative software engineering. Various infrastructures, guidelines and tools are available for continuous integration (CI) and continuous deployment (CD) of software. The same principles could be used to design infrastructures for continuous quality (CQ) management in research and the implementation of the Open Science paradigm. In this work we provide a vision of how such a CQ workflow could look like to make scientific research traceable and reproducible.**

Keywords – software development, automated quality assurance, continuous integration, continuous deployment, digitalisation, Open Science

I. INTRODUCTION

Results of scientific research need to be reproducible in order to make actual use of their findings and developments. This simple rule holds for all areas of science – from psychology to materials science and metrology. Reproducibility of research requires a clear traceability from the result reported in a publication to the raw data from measurements or other means. Quality management in research thus should ensure reproducibility by providing measures for the verification (error free) and validation (fit for purpose) of the steps that lead from the raw data to the findings reported. Therefore, most scientific institutes have documented quality guidelines [10] and are using quality management processes for their research. However, these are typically paper-based documentations of principles and infrequent tests. Often inadequate quality management in research only becomes visible, when other scientists question specific results or report flaws. These days, scientific work is usually mostly digital, and a wide range of tools for each step of the research process is available. Hence, traceable and reproducible digital research results should be feasible with available building blocks.

Access to research results includes access to the steps, software and measures that led to the results reported. This is part of the Open Science paradigm [1]. For the open research data, the FAIR principles require the data published to be findable, accessible, interoperable and reusable [2]. In order to realise these principles, processes and infrastructures need to be adapted. To this end, several scientific groups have started developing guidelines and infrastructures to support researchers in the management of research data in accordance with the FAIR principles. It can thus be expected that research data management itself will become more reliable in the near future.

In order to achieve the same for the whole research process, further measures are required. At the same time though, it must be ensured that scientists can focus on what matters most: research. To this end, principles and infrastructures from modern software engineering practice can be employed as template for a mostly automated continuous quality management in research. This enables CQ throught the whole process, not only at the end. CQ is an integrated part of the development.

II. MODERN SOFTWARE ENGINEERING

Software engineering is the application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software; that is, the application of engineering to software [3]. Development teams usually work asynchronously and geographically distributed and to this end employ tools and workflows assisting the special needs arising from that, like dealing with version conflicts for instance. The starting point for every professional software engineering activity is to ensure the utilization of proper source code management (SCM) enabling all subsequent activities. The main purpose of SCM is tracking file changes linearly over time and non-linearly over different features and bug-fixes kept in so-called "branches". The inverse action to branching is the so-called "merging" and requires modern SCM to provide tools visualizing the differences in diverging versions of the same files and to simplify the process. All wide spread SCM nowadays tackle these issues very successfully by offering well integrated diff- and merge-tools. In conjunction with detailed tracking information about who and when applied what changes and for what purpose those diff-tools allow for transparency about the whole development process.

Regarding workflows the most crucial practice modern teams follow is continuous integration (CI) which lays the foundation for asynchronous and distributed engineering [4]. It is characterized by sharing the output of every team member early and often. Frequently integrating written code into the joint code base fosters a corporate understanding of the current project's state which in turn reduces the number of misunderstandings between developers, discrepancies between developments and therefore minimizes the need and the required effort to resolve conflicts. The early integration of code into the complete process is in contrast to the traditional approach of fine-tuning every detail of a certain module or code snippet before making it available to the collaborators. Hence, this is not only a technology or code development technique but requires a cultural change as well.

Another common practices in modern software engineering is permanent testing and perceiving test writing as being part of the process. These tests range from so called "unit tests" addressing the purpose and boundary conditions of small portions of source code. Up to tests which address the functionality of the final product and its deployment into the intended operating environment altogether, often referred to as the test suite. Adding tests from early stages on and testing new features against integrability ensures quality and robustness of the produced code.

To this end, automated tools are utilized which perform these tests and keep track of the project's current state. Such testing toolchain consists of elements incorporated locally into the team members' IDEs, centralized SCM components and distributed webservices to continuously aid every step from before integrating the code into the code base to final production deployment for why they are commonly referred to as continuous integration / continuous deployment (CI/CD) pipelines. Those pipelines naturally place some demands on the environment in which they are established like quasi-permanent physical and logical access to the SCM holding the code including the test suite and interconnections between the elements themselves. The latter enforces the automation aspect, as pipeline elements can communicate with each other, e.g. trigger the next element in the toolchain or abort further build stages in case of error and inform the developer.

The goal of these efforts is to produce reliable software avoiding all undocumented and unreproducible interventions thus enabling quality assurance and management during and after the development. This holds true especially for open-source software where not only the product itself is delivered but access to the SCM and CI/CD pipeline's outputs is granted in whole or in part to everybody, with access to the final code version being the minimum requirement.

III. COMPLEX RISK CALCULATIONS USE CASE

In the northern part of the Netherlands one of world's largest gas fields is operated. This field causes induced seismicity that is linked to the gas subtraction and damages the build environment, e.g. houses. To quantify the public risk that is associated with the gas production yearly risk estimations (Hazard and Risk Assessment) are being made. These calculations are made up of a complex chain of models which are constantly being improved and modified according to the latest insights.

As policy is based on these type of calculations, Quality Assurance and traceability is of essential importance. Can we reproduce the calculations of two years ago with the software – and data versions that have been used at that time?

These calculations can be characterized as data intense complex chain of models. On a high abstraction level, the model chain consists of three main components: Seismic Source Module (SSM),

Ground Motion Module (GMM) and Damage Module (DM).

The workflow of executing and developing components of the model chain is divided in three steps. A new analysis question starts with the *development phase*, which is a playground for new features. No versioning or traceability is kept for a long period. Soon the process moves to a *staging phase* working dedicated on specific functionality. Continuous testing becomes an integral part containing unit tests for error-free code and model verification and integration tests for validation of (parts of) the entire final pipeline of models. Finally the last phase is *production phase*, using only verified models and pipelines. This environment is kept persistent for a long time, containing all used data, configurations and model software. Because the whole process is automated these cycles can be very sort in time, allowing iterative development because of quick feedback.

In the staging phase the software runs on an exact mirror of computer infrastructure that is used in de production step. When modules have passed the staging phase, modules are built to small self-executable Docker-images. This results in a package that contains the operation environment (e.g. Linux), depending software libraries and the written software itself. This package is identifiable with an unique id by committing is to a CI/CD environment (GIT).
The staging- and production phase both need support for full automation of testing, running and traceability. For this the tool Pachyderm [11] is selected which can manage analysis pipeline runs, provides full traceability of each run and can reproduce previous (old) runs.
With Pachyderm a pipeline of models can be created by putting each model in a separate docker image and connects them through data sets. See figure 1. Each data set has its own full version history, including the pipeline specifications. Running a specific pipeline results in the storage of not only the end result data set, but also all intermediate and external data sets, the originating data sets and the configuration of all models. When a change is made to any of these, for instance a parameter in the configuration of a model is updated, automatically the pipeline is rerun. But only those parts which are effected by the parameter update. For instance when the parameters of the GMM model is updated, the SSM model will not be recalculated. Only the GMM and DM model resulting in an update of those corresponding data set.

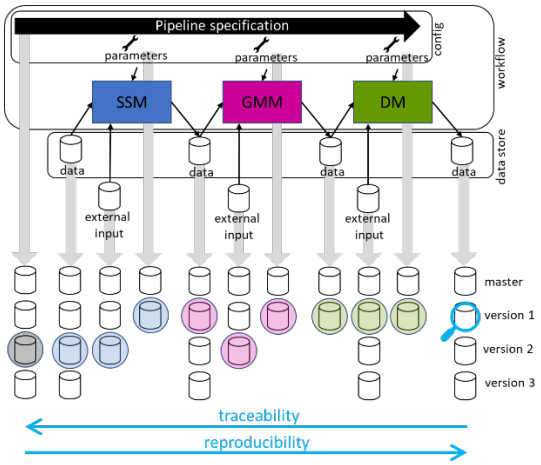


Figure 1: Pipeline versioning and traceability.

For each run full traceability information is generated (indicated by the blue arrow from right to left at the bottom of figure 1) linking the result data set to the models, their configuration, the intermediate data sets and finally back to all originating data sets.
When a question or complaint is raised, a result can be reproduced (indicated by the blue arrow from left to right at the bottom of figure 1) by rerunning the exact old run specification which was used to produce the result originally. That is precisely the traceability information generated with the original run. Therefore this must be stored as long as questions and complaints are expected.

IV. TRANSPARENCY AND OPEN SCIENCE

The shift to Open Science means sharing research results by making them available in a way that they can be accessed and utilized by the whole society [1]. However, this does not mean to provide a scientific paper as Open Access, because then only the financial barrier to access the publication itself would be removed. Access to the research itself also requires access to the relevant data and to the methods (software and configurations) that were employed to gain the results reported. This is also one important step towards reproducibility in science, because other scientists can easily verify and reproduce the steps that led to the conclusions. An excellent example for this concept is the LIGO "Gravitational Wave Open Science Centre" website. It offers the scientific publications and relevant research data as well as documented source code that was used to analyse the data [5]. In this way, reproducibility of the results and further development of the methods by the scientific community is greatly improved.
The prerequisite for a transparently documented

research process is a standardised set of minimum requirements, and the availability of appropriate research infrastructures. For research data this concept is given by the FAIR principles:

- Findable data requires persistent identifiers for the data sets
- Accessible data requires that the data can be downloaded in some way
- Interoperability requires using standardised metadata and annotation of the data
- Reusability requires that the permissions, i.e. license, for the data allows others to use the data

It is important to note, though, that the FAIR principles do not require the data to be publicly available for free as Open Access. That is, implementing a FAIR data management does not mean to make all data available for free. Instead, it means that research processes that generated the data need to be documented and structured in a way that supports the *I* of FAIR.

The *F* and *A* parts require a searchable data base for research data as well as persistent identifiers to individual data sets. Therefore, a couple of development projects have been initiated to provide infrastructures for FAIR research data management. For instance, the project CONQUAIRE aims at an infrastructure that uses git version control for research data [6]. In this way, changes to data sets are transparently documented and earlier versions of the data can be accessed easily. This provides advantages for the scientists performing the research as well as for others. The *TIB Datamanager* [7] uses a similar approach, but also provides basic visualisation tools for data sets. These kinds of infrastructures aim at providing a simple toolset for scientists to more or less automatically document, store and provide their research data.

In the same way, modern electronic lab journals provide an infrastructure for a transparently documented research process. For instance, software like *SciNote* [8] or *labfolder* [9] provide a server-based system that can be accessed via any web browser to create and edit entries. Each editing step automatically generates a new version of the entry, creating a transparent process of the research process.

V. CONTINUOUS QUALITY IN RESEARCH

Quality in research means reproducibility, reliability and transparency from the relevant data to the published conclusions and results. By the researcher during the process of doing his research as well as by other people (long) after publication of the results. In the following we discuss the meaning of CI and CD for research workflows consisting of (i) development of an idea/concept, (ii) acquisition of data, (iii) analysis of data and (iv) reporting the results as a scientific publication, presentation or other means. Based on this we then propose a concept for a continuous quality (CQ) for research.

Phase 1: Development of a concept or hypothesis

In general, a research process starts with a distinct problem, issue, question or challenge to address. Usually, this is the part of the research workflow that is the least documented. However, a good documentation and also publication of this step could be of benefit in several regards: supporting later re-iteration in case the results do not match the expectation; involvement of and feedback from other experts in the field; prevention of fraud. For instance, several scientific journal papers are now requesting their authors to pre-register their hypothesis before conducting the actual study. This is intended to prevent fraud by matching the hypothesis to the actual findings afterwards. At the same time some journals are encouraging authors and reviewers to also consider negative research results for publication, e.g. the journal F1000Research.

Technology-wise, this step can be supported, for instance, by using electronic lab notebooks (ELN) from the start of the research workflow. Therefore, the ELN needs to be able to include plain text, handwritten notes as well as documents with annotations. Furthermore, the ELN should be able to track changes in these elements. Thereby, later changes to the hypothesis, problem description, etc. are made traceable automatically. For research areas where an ELN may not be suitable, digital text-based documents can be used in conjunction with an automated versioning system like git or subversion.

Continuous integration (CI) for this initial step in the research workflow can consist, for instance, of using the ELN collaboratively with other researchers. In a sense, this documents what is common practice in science anyway: joint development of ideas, hypothesis and concepts.

Continuous deployment (CD) then means that this documentation is automatically made available to others, for instance, on a pre-print server.

Phase 2: Data acquisition

The step of data acquisition in terms of measurements, surveys or other means is typically digital as data is being captured in digital formats from the start. Thus, versioning systems can be directly applied. Ideally, this should happen

automatically by implementation of a corresponding data acquisition pipeline. Systems like CONQUAIRE, Datamanager and others are offering technological solutions for this. The researcher itself, though, has to make sure that the raw data streams are actually fed into these systems in an automated way.

With the data being available on an accessible platform and under version control, automated pipelines for data quality checks can be implemented. These quality checks can be very basic, such as, testing for NaN, Inf or other entries that indicate errors in the data acquisition process. Furthermore, tests against certain assumptions can be implemented easily. For instance, spectral measurements can be checked for non-negativity. More sophisticated quality checks may contain plausibility tests against reference data or automated data curation methods. The data quality system could then automatically inform the researchers about issues and annotate the metadata with details about its findings.

Not all issues detected need to result in withdrawal of the data set. For instance, a combination of the metadata about the issues with the subsequent analysis results could lead to unexpected findings and open up new research questions. In addition to the acquired data itself, the software for the data acquisition needs to be put under version control. Changes in the data can then also be traced to changes in the data acquisition process. Quality assurance of this software can be carried out following the principles described above. The version control and quality assurance of data acquisition software extends to a documentation of the data acquisition parameters. When the measurement setup is also software-based or other means for the automated access to the acquisition parameters are available, this part of the research workflow quality system can be automated as well. For example, measurements with an instrument with a programming interface (API) can be acquired together with metadata containing the information about the measuring device's parameters being used. Continuous integration (CI) at this step of the workflow could mean the combination of individual data sets within a larger system. For instance, in a particle accelerator many individual measurements are made and need later to be combined to form an analysis. As another example consider the sensor network example from above. The CI pipeline could, for instance, consider only measurements that satisfy the quality checks to be included in the final data set made available for the next step in the workflow.

The CD part can store the data sets to be used in analysis, the next phase. As before also publishing the data sets to a pre-print server or other place for reporting the research.

Phase 3: Data analysis

The workflow step (iii) data analysis is typically carried out in software, either in terms of self-written source code or using available software tools. For the case of self-written source code, the same versioning system approach as for the data acquisition part can be used. For quality assurance purposes the software testing principles described above can then be applied. Regarding CI principles, an interplay with the previous step in the workflow could be considered. For instance, the software for data analysis may require certain properties in the data, or the pipeline of software modules is depending on specific data properties. Moreover, continuous integration for this step could be realised by means of applying each analysis version, committed to an analysis repository, to the data sets after quality-checking the software itself. Then, the CI pipeline would automatically provide an updated analysis result for the latest software version. This, of course, is more challenging when commercial or other closed-source software tools are applied, only black box testing remains in those cases.

Phase 4: Reporting the findings

The last step in the here considered research workflow is the reporting of the research findings. This is also typically carried out in a digital format that can be put under automated version control easily. Quality assurance in this step, however, is usually limited to grammar, orthography and other basic errors. With more sophisticated machine learning tools, though, complex semantic checks are possible, too. For instance, a paper draft could for a medical survey could be checked whether it contains a well-defined hypothesis, information about the method used and sufficient details about the study.

Continuous integration for this step in the workflow can be carried out, for instance, for the case of LaTeX documents. The LaTeX source code can be put under version control and automated compiling into a PDF document can be carried out on the server. This also allows continuously integrating contributions from different co-authors. Moreover, for the automatically published data sets and data analysis code an automated citation in the report can be integrated.

Continuous deployment in the reporting step can

consist of automatically submitting the compiled document to a pre-print server. This publication could then contain also references to the exact analysis workflow trace, references to all analysis software versions, raw input data sets and configurations. With the automated and continuous quality assurance in every step, the authors can have reasonable confidence in the validity of the report being made publicly available. For intermediate work, this publication could provide automatically generated indications to what quality checks may have failed or what parts are still missing. This is common practice in open source software repositories, and it allows the users of the software to assess the maturity of the published work easily.

These steps all together can result in the following example research workflow. Assume a server-based electronic lab journal is used that allows access to its entries via document APIs. An external software module may then be configured to automatically check specified lab journal parts for consistency w.r.t. units, uncertainties as well as for completeness and other quality measures. This corresponds to unit tests in software engineering. If then all checks for that lab journal are positive, a data analysis algorithm is applied to produce a specific result for that lab journal data. Its output could again be verified by another software module against some reference or pre-specified criteria. If all checks are positive, an entry is created for the output in a research data infrastructure with metadata pointing to the persistent identifiers of the journal entry and the software version being used. Furthermore, a pre-specified documentation of the research, data analysis and findings could be created similar to automated software documentation based on comments in the code.

VI. CONCLUSIONS

Continuous quality (CQ) for research means to develop and implement automated checks for a set of measurable quality factors throughout the research workflow and lifecycle of the analysis. Similar to CI/CD pipelines in software development, automated processes can then be defined to implement FAIR principles for research data or specific aspects of the Open Science paradigm.

Continuous deployment based on CQ could strongly support the Open Science paradigm. For instance, a public dashboard for research project could summarize the CI/CD pipeline outcomes of every step in the above described research workflow. Starting from the original idea, developed collaboratively; to the acquired data; to the data analysis source code and the reporting draft: all elements could be made available in a transparent way and automatically created quality assessments.

The full implementation of CQ is – right now – more a vision than an actual project. However, the principles from software engineering are well established, several infrastructures for research data are being developed and more and more scientific communities are defining quality measures for their work. With CQ pipelines, the administrative overhead for scientists to satisfy FAIR principles and the Open Science paradigm can be minimised. At the same time, direct advantages for the scientists in their work can be realised.

VII. ACKNOWLEDGMENTS

Part of this work has been developed within the Joint Research project 17IND12 Met4FoF of the European Metrology Programme for Innovation and Research (EMPIR). The EMPIR is jointly funded by the EMPIR participating countries within EURAMET and the European Union.

REFERENCES

[1] **Vicente-Saez, R., & Martinez-Fuentes, C.** (2018). Open Science now: A systematic literature review for an integrated definition. Journal of Business Research, 88, 428-436. DOI: 10.1016/j.jbusres.2017.12.043

[2] **Wilkinson et al. "**The FAIR Guiding Principles for scientific data management and stewardship", Nature Scientific Data, vol. 3, 160018 (2016). DOI: 10.1038/sdata.2016.18

[3] IEEE 24765-2017 - ISO/IEC/IEEE International Standard - Systems and software engineering—Vocabulary (https://www.iso.org/obp/ui/#iso:std:71952:en)

[4] **Linping Chen** "Continuous delivery: overcoming adoption challenges", Journal of Systems and Software, vol. 128 (2017). DOI: 10.1016/j.jss.2017.02.013

[5] https://www.gw-openscience.org/start/ Acc. 2019-03-14.

[6] **CONQUAIRE "**Expanding the Research Data Management Service Portfolio at Bielefeld University According to the Three-pillar Principle Towards Data FAIRness" DOI: 10.5334/dsj-2019-006

[7] **TIBD** https://projects.tib.eu/datamanager/ Accessed 2019-03-14.

[8] https://scinote.net – Accessed 2019-03-14

[9] https://www.labfolder.com – Accessed 2019-03-14

[10] The **AQUA Book**: guidance on producing quality analysis for governement. HM Treasury. March 2015, isbn 978-1-910337-67-7

[11] https://www.pachyderm.io - Accessed 2019-6-19

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"
Berlin, Germany, on 3-4 September, 2019

The long way to reliable inline production measurement in metal wire products manufacture

Dr.-Ing. Wilfried Hinrichs
Materialprüfanstalt für das Bauwesen Braunschweig, Braunschweig, Germany, w.hinrichs@ibmb.tu-bs.de, +49(0)531 391 5902

***Abstract* – Optical inline measurement is common in textile production, but not in the manufacture of some metal wire products using comparable techniques. The reason is that some ISO standards require highly precise measurements which are hardly feasible in the running production. Major factors for getting results of the aperture - the most important value in this respect - are the time needed to focus the camera, strong vibrations of the weaving machines as well as the light on reflecting surfaces and reference objects. Even in offline measurement, the comparability of the results is limited due to more restricted requirements especially in ISO 3310-1:2016 and ASTM E11-17 (test sieves). In these documents the conformity assessment methods have been updated on the basis of metrological recommendations. An ISO working group decided to pick up this issue and to look into the possibilities to improve the situation. That addresses the number of apertures to be measured, the coverage factors used as well as the overall statistics applied in the standard. Some research project are going to start accordingly in order to make the standards usable in the future.**

Keywords – metal wire products, optical measurement, aperture, test sieves, measurement uncertainty, conformity assessment

I. INTRODUCTION

In industrial textile manufacture, inline production measurements are important for high-quality level products. The measuring instruments used are often highly sophisticated and needed to maintaining specified levels of a continuous regularity of the product. This situation is – from a measurement point of view – quite different in the manufacture of metal wire cloth and screens.

Picture: Manufacture of metal wire cloth

Metal wire products are world market products for a wide range of applications. An important share of these products is used for sieving and filtering as well as for the classification and grading of solid matter. In addition, screens for printing purposes are also often made of metal wire with more or less precise requirements as regards the wire diameters and apertures.

Metal wire cloth and screens are mainly produced in East and South Asia, in Europe and North America. Important standards are ISO 3310-1:2016 (Test sieves) and ASTM E11-17 (Standard Specification for Woven Wire Test Sieve Cloth and Test Sieves) as well as ISO 9044:2016 (Industrial woven wire cloth) and ISO 14315:1997 (Industrial wire screens).

In the last twenty years in the revision processes of these ISO standards the focus was on improving the measurement techniques. This aimed mainly at the precision of the products and/or the evaluation of measurement results with statistical methods. At the same time, users asked for improved documentations of the measurement results. Finally, it had to be clarified whether test sieves according to ISO 3310-1 are measuring devices or not. In sum, the revisions led to more stringent product specifications.

II. NESSECITY OF REASEARCH

The kick-off for the research on measurement procedures for wire screen and woven wire cloth products was the general observation that the comparability of measurement results of the mean aperture sizes of test sieves for nominal apertures of $w_N \leq 2.0$ mm is poor. This can be illustrated with the comparison test of the results from six laboratories of a product with a nominal aperture of $w_N = 0.1250$ mm. The same test sieve claimed to be in accordance with ISO 3310-1 was sent to the participating specialized laboratories. Three of these conformity assessment bodies came to the result that a certificate of compliance can be issued, whereas the three others stated that the test sieve did not comply with the standard. This example is specific but it confirmed a general experience from practice.

The simple reason why the producers have not paid attention before to this phenomenon is that measurement results were mostly generated under repeatability conditions and they were rarely documented. There was no substantial reason to further analyze this kind of data because that rarely raised problems neither in production nor when used in bilateral agreements where one party carried out the task. However, under reproducibility conditions the situation is quite different and problems are emerging with increasing multilateral trade relations.

The analysis of test data from the production sites in the research project has made clear that the general assumption of a Gaussian distribution of apertures is not correct. For the apertures in the weft direction in Fig. 1 both the apex and the average value are very close to each other and normal distribution can be stated using the Shapiro–Wilk test.

However, for the warp direction apex and average value are different, the distribution is obviously asymmetric showing a distinctive negative skew and there is no such tail of larger apertures than it is for smaller ones. This curve can be described as a specific shape of a beta-distribution, an interpretation which is highly plausible from a production point of view: In the warp the reeds limit rather stiffly the positioning of the wires. But as both the dimensions of the reeds and the wires are not invariable the bounds of the beta distribution are variable too. Hence it is assumed that the test results in the warp direction are beta-distributed with an overlapping Gaussian distribution due to the construction elements.

Fig. 2: Distributions in warp and weft direction

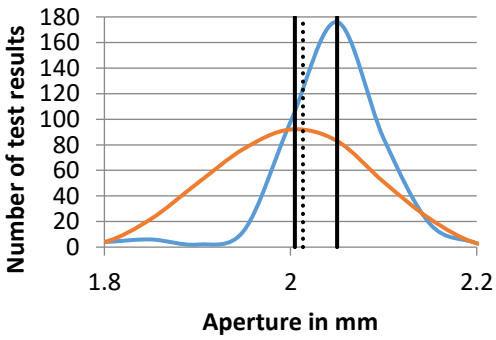


Fig. 2: Distributions in warp and weft direction

As the product standards for wire screen and wire cloth assume normal distributions the calculations in the research project have been made both for Gaussian and beta-distribution in order to see the differences between the outcomes based on the status-quo procedure and that on the basis of the new research results. Table 2 contains an example for the underlying problem in conformity assessment when the distribution of the measurement result is not clear.

Parameter	**Gaussian distribution**	**Modified beta distribution**
Without a bias		
Probability of conformity P_C	0.805	0.967
With a bias		
Probability of conformity P_C	0.764	0.942

Table 2: The probabilities of conformity for T_U and T_L for two distributions at $\sigma = 0.997$

III. BASIS OF CONFORMITY ASSESSMENT

WILLIAMS and HAWKINS, ROSSI and CRENNA, PENDRILL and the author of this contribution have published on conformity assessment in industrial production. Further documents used were the Joint Committee for Guides in Metrology 106 and ASME B89.7.4.1.

Their results have been applied to the product standards mentioned above. They all encompass requirements for mean aperture sizes $\hat{w}$ and maximum aperture sizes w_{max}.

Generally speaking, a profound impact of measurement uncertainty on conformity decision can be excluded if the uncertainty is small compared to the tolerance zone. A usual method for estimating this possible influence is the test uncertainty ratio (TUR). Another one is the calculation of a process capability index C_P which is derived from statistical quality control.

$$C_P \equiv \frac{T_U - T_L}{n \cdot \sigma_P} = \frac{T_{-2}}{n \cdot \sigma_P}$$

The choice of n is arbitrary, it usually differs in the range from 3 to 10. T_{-2} is the full width of a tolerance zone with T_U as the upper and T_L as the lower limit. σ_P is the standard deviation as a convenient measure of the process variability. With an assumed frequency distribution – mostly Gaussian – C_P quantifies that fraction of products that is within the tolerance zone T_{-2}. A similar index value is defined as the measurement capability index C_m which is defined for two-sided tolerances as

$$C_{m-2} \equiv \frac{T_U - T_L}{4 \cdot u_m}$$

Similarly to C_P the figure C_m characterizes the quality of the measurement system. A large C_m indicates low measurement uncertainty.

ISO 3310-1 specifies a maximum acceptable standard deviation σ_0. With the data on the dispersion σ_m of single measurement results collected in the interlaboratory comparison it is possible to calculate an index value for this tolerance figure as well:

$$C_{m-1\sigma} = \frac{\sigma_0}{2 \cdot \sigma_m}$$

The calculation of a measurement capability index focuses on the measurement itself. The index helps in decisions whether instruments or procedures are generally suitable for conformance or conformity testing. In order to quantify risks the above mentioned draft paper goes more into detail. According to the paper the model of the measurement process for two-sided tolerances is assumed to follow the Gaussian probability density function

$$p(x \mid I_m) = \frac{1}{u_m \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x - x_m}{u_m} \right)^2 \right]$$

I_m stands for information available after performance of the measurement, x_m is the result of a measurement and u_m is the associated standard uncertainty. The transformation to a cumulative probability Φ(z) leads to the conformance probability

$$P_C = \Phi\left(\frac{T_U - x_m}{u_m} \right) - \Phi\left(\frac{T_L - x_m}{u_m} \right)$$

or

$$P_C = \Phi\left[4 \cdot C_m \cdot (1 - \hat{x}) \right] - \Phi(-4 \cdot C_m \cdot \hat{x})$$

The expression P_C is the probability of conformance. With the use of uncertainty data from reproduction it can also be defined as the probability of conformity and can be calculated for all levels of confidence.

The assumption for u_σ is numerically done using the mean values σ_w ($\sigma_w = f(w)$) from an interlaboratory comparison. Again with $T_U = 0$ the probability of conformity runs as follows:

$$P_{C,\sigma_0} = \Phi\left[\frac{\sigma_0 - \sigma_m}{\sigma_w} \right] - \Phi\left[-\frac{\sigma_m}{\sigma_w} \right]$$

The requirements for σ_0 are based on a confidence level of about 68 % ($\sigma = 1$). It has therefore not been calculated for different confidence levels.

IV. RESULTS

In wire screen products the variance of the apertures is usually much larger than the variance caused by testing in conformity assessment, i.e. the uncertainty of individual test results is usually negligible. However, if test results from different stations of production are used it must be considered that they are afflicted with larger uncertainties than those from the final conformity assessment itself. It is also important to highlight that the quantification of risks

with a focus on tolerances does not allow neglecting even small influences, if the number of test results close to the tolerance limit is not small.

With large dispersions of a property of a measurand and a position that the uncertainty of test results is usually negligible it is only possible to get at data by reducing the effects of the variance of the measurement object. This has been done by creating wire screen samples with different apertures for use in an interlaboratory comparison test where all participants get the same sample, one after another, and with a detailed prescription for the performance of the test.

ISO 3310-1 (test sieves) sets the toughest requirements. Statements on conformity are not possible for products with aperture sizes of $w \leq 0.5$ mm on a 68 % confidence level and aperture sizes of $w \leq 2$ mm on a 99.7 % confidence level with the given state-of-the-art measurement procedures. This statement is generally in line with the C_m results.

In contrast to woven wire cloth products for wire screens according to ISO 14315 the whole range of the scope can be used for conformity assessment.

For the maximum aperture size w_{max} the provisions of ISO 3310-1 are such that statements on conformity are possible down to $w_N = 0.06$ mm. The specifications for the maximum standard deviation σ_0 do not cause a general conflict.

The measurement capability indices $C_{m\text{-}2w}$ (mean aperture size $\hat{w}$), $C_{m\text{-}1w}$ (maximum aperture size w_{max}) and $C_{m\text{-}1\sigma}$ (maximum standard deviation σ_0) have been calculated for a very wide range of aperture sizes (32 mm … 0.06 mm) with the assumption of favourable measurement conditions (calibrated instruments with a readability of 0.01 mm, trained personnel, well described measurement procedures etc.).

Wire screen samples with apertures between 4 mm and 32 mm and wire diameters between 1 mm and 5 mm were sent to 19 participants. For the tests a vernier caliper was used which is a major test instrument in production and in conformity assessment. The results are given in table 1. The standard deviation for the aperture is $\sigma \approx 0.04$ mm. It is necessary to adapt the measurement uncertainty to the purposes of risk calculation, i.e. the standard deviation as a measure for the measurement uncertainty u_m should be larger because

- interlaboratory comparison tests are usually done both by specialized personnel and with particular care.
- the instruments as used in production have usually a smaller resolution.
- it is necessary to pay attention to possible systematic effects. The results of some participants lay all above or below the median which underlines a general impression from the visits on the production sites.

Hence an assumption of $u_m = 0.1$ mm is regarded as appropriate for general tests in production and $u_m = 0.05$ mm for conformity tests in internal laboratories. If the applicant cannot reasonably exclude a bias of the results which is not negligible, it is deemed appropriate to assume a possible bias of $u_{m,b} = \pm 0.03$ mm.

The probability of conformity is strongly affected by the application on the basis of the modified beta-distribution. An example for the distribution can be seen in Fig. 2.

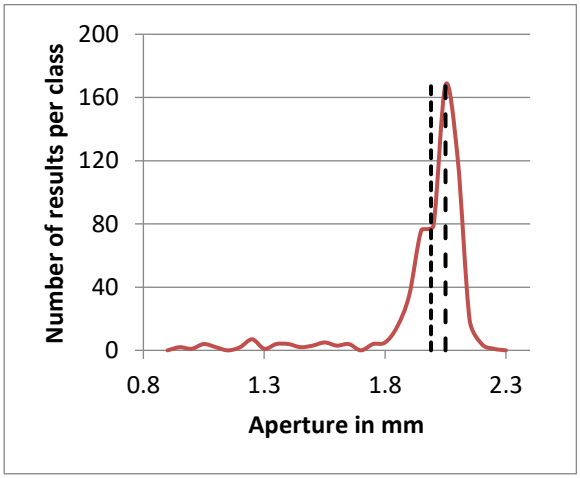


Fig. 2: Example for the distribution of the apertures

This result was not surprising because the number of test results within the tolerance limits is much larger than with the normal distribution. It becomes already clear that the application of a general procedure is not appropriate for these specific products. If the probability of conformity P_C would really be at a level of 0.8 testing would be a major issue in production and conformity assessment. In order to maintain a probability of conformity of $P_C = 0.967$ on the basis of the Gaussian distribution the tolerances would have to be expanded from $w_{max} = 8.40$ mm to $w_{max} = 8.68$ mm. However, this would not be accepted by the clients and it is not necessary as in practice it is already implicitly known by the producers. Possible measures are regular checks of the wear of the reeds, increasing the awareness for the need of some safeguarding limits and tests of the apertures when setting up the machines. This reflect a situation that is much more understandable from a high probability

of conformity point of view. The situation becomes more important when there is a significant bias in testing. This is a relevant issue as the producers are mainly acting in a bilateral structure. Here both partners can adapt their testing procedures according to their needs. From this point of view a bias is much less important but it would become a problem when entering a multilateral market. Here the bias may lead to discussions on the validity of test results.

III. INLINE MEASUREMENTS

The assessment of conformity of wire screen and woven wire cloth products on the basis of ISO 14315, ISO 9044 and ISO 3310-1 is not always possible with state-of-the-art measurement procedures. This statement does not necessarily include conformance testing.

Inline measurements in the production would be most useful for the aperture of woven metal wire cloth or of screens. However, this is still a problem: Ongoing production operations require measurements on the object being in motion. These operations work to assess the regularity of the product and help to identify some kinds of non-compliances. However, it is already difficult to produce reliable results for the aperture dimensions under (offline) 'laboratory' conditions. Therefore, in general, results from inline measurements have not been satisfactory. An important reason is that usually geometric pattern on applied on glass surfaces serve as 'two-dimensional' reference objects and the results are different from those produced with 'three-dimensional' reference objects made of wire.

Optically based online measurement techniques are applicable for controlling the regularity of the product but that would not be very helpful. There are two major reasons for this situation:

- The apertures are usually measured using bright-field microscopy. Here the time needed for focusing of the camera is an essential issue. Even when, compared to industrial textile manufacture, the production speed in wire cloth and screen manufacture is quite low the focusing process during the ongoing operations within a few microseconds in this heavily vibrating environment is still an unsolved problem.
- The measurement results strongly depend on the light intensity. It is possible to protect the measured objects in the ongoing production although they have light-reflecting surfaces. Nevertheless, a constant distance from the light source can hardly be maintained.

The situation is that for production control reasons the producer needs information whether the product meets the requirements especially as regards the aperture. But the uncertainty of these results is still too large for reliably monitor the production. Therefore additional, more precise offline measurements are necessary. In the long run, a combination of online and offline measurement provides data for a better production control. An example of a screen shot is Fig. 3.

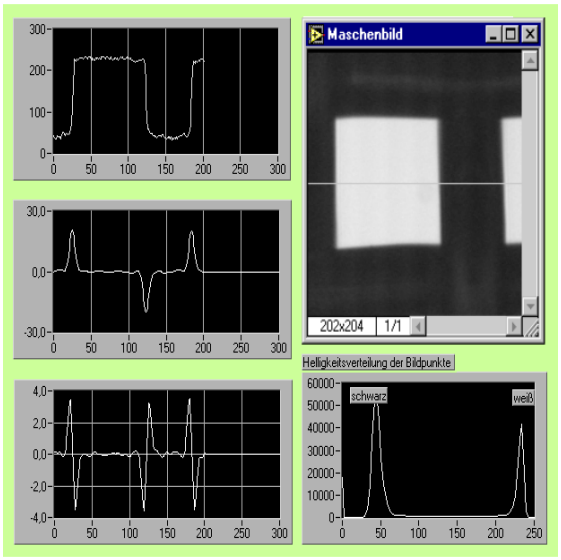


Fig. 3: An example of a screen shot in off-line measurement

Suppliers of measurement instruments claim that their products can cope with this situation. In principle, this claim is not wrong, at least from a production control point of view. However, the comparability of results arrived at using different industrial optical measurement techniques is usually quite poor. So the problem remains the applicants when they calibrate and/or check the products.

IV. REVISION OF THE STANDARD

Apart from problems with reliable results both in the production and in use the requirements for the measurement of the product in the standard for test sieves are quite extensive. The ISO working group are busy to improve them in order to reduce the number of single measurement items, keeping in mind that the confidence levels which are now 99 % or 99.73 %, must be maintained.

Moreover, there are other problems with ISO 3310-1:

- As the probability distribution functions applied in the standard does not reflect the distributions found in the measurement results, the influence on conformity assessment results must be somehow quantified.
- The coverage factors k used in the standard are not justifiable.
- The term 'calibration' is not in line with the VIM.

The research results of the last about ten years have been used to adapt the 2016 versions of the standards. Now the ISO working group is asking new questions calling for additional research projects.

- Is it possible to stop the measurement process when the product meets the requirements on a certain conformity level?
- Under which conditions is it possible to refer to a normal distribution even when it is known that the measurement results are beta-distributed? This would make the application of commonly used statistical instruments much more probable.
- How can the measurement data be used for improving the production process (process capability index)? For example, can be used on this basis operators like the acceptable quality limit or rejectable quality level for testing incoming goods?

REFERENCES

[1] ISO 3310-1:2016 (Test sieves - Technical requirements and testing – Part 1: Test sieves of metal wire cloth)

[3] ASTM E11-17 (Standard Specification for Woven Wire Test Sieve Cloth and Test Sieves)

[3] ISO 9044:2016 (Industrial woven wire cloth - Technical requirements and tests)

[4] ISO 14315 (Industrial wire screens - Technical requirements and testing).

[5] Williams, RH; Hawkins, CF, The economics of guardband placement, Proceedings of the 24th IEEE International Test Conference, Baltimore, 1993

[6] Hinrichs, Product-specific adaption of conformity assessment criteria and their financial consequences, Production Engineering Research & Development, 5 (2011) 5, 549–556

[7] Hinrichs, W, Sampling procedure for optical measurements in woven wire production, SENSOR+TEST Conferences 2011, 7-9 June 2011, Nuremberg; OPTO Short Proceedings

[8] Rossi, GB; Crenna, F, A probabilistic approach to measurement-based decisions, Measurement 39 (2006) 101-119

[9] Hinrichs, W, Linking conformity assessment and measurement uncertainty – an example, tm – Technisches Messen 73 (2006) 10, 571-577

[10] Pendrill, L, Operating 'cost' characteristics in sampling by variable and attribute, Accreditation and Quality Assurance (2008) 13:619-631

[11] Hinrichs, W, The impact of measurement uncertainty on the producer's and user's risks, on classification and conformity assessment: an example based on tests on some construction products, Accreditation and Quality Assurance (2010) 15:289-296

[12] Joint Committee for Guides in Metrology 106:2012. Evaluation of measurement data — The role of measurement uncertainty in conformity assessment

[13] ASME B89.7.4.1-2005; Measurement Uncertainty And Conformance Testing: Risk Analysis (Technical Report)

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Hybrid approach and sensor fusion for reliable condition monitoring of a mechatronic apparatus

Giulio D'Emilia, Antonella Gaspari, Emanuela Natale

DIIIE, University of l'Aquila, L'Aquila, ITALY, Giulio.demilia@univaq.it, +39 0862434324

***Abstract* – In this paper the results are presented of a methodology for identification of the best setting of a mechatronic device at its set-up and/or of fault occurrence in working conditions. The condition monitoring approach merges the information deriving from experimental data of sensors and from a simulation model. The methodology is used to identify the most suitable quantities to be measured, independently on location for sensors. The hybrid information suggests a synthetic set of features, able to identify specific statuses of the mechatronic system, different as for initial setting and/or for the occurrence of faults, to be processed by advanced algorithms like Artificial Neural Networks. The tests show that the methodology is able to realize a resolute and reliable condition monitoring of a mechatronic system in real scale, with reference to both the parameters of setting and to the occurrence of wear and lubrication problems in the kinematic linkage. Finally, the information content of data deriving from internal sensors to the controller is maximized, so reducing the need of external ones for reliable condition monitoring applications.**

Keywords – Sensor fusion, Measurement accuracy, Repeatability, Condition monitoring, Diagnostics, Simulation.

I. INTRODUCTION

Condition Monitoring (CM) of assets, in particular automated production lines, is a topic involving more and more interest in modern industrial scenarios, due to many advantages, which can be reached. To only cite the most important aspects, the following results are possible:

- the optimization of set-up and working conditions;
- the improvement of maintenance strategies;
- the increase of the reliability of operations.

Nevertheless, many aspects should be considered and approached if effective, accurate, reliable methods have to be provided in the field of real industrial applications, which are also able to be of general validity.

The main topics to be solved can be summarized as follows:

- purely data driven methods are often expensive from the computation and experimentation point of view [1]; the use of advanced analytics and algorithms, like Artificial Neural Networks (ANN) makes possible achieving interesting results, even though they have to be optimized for the specific application [2];
- simulation and modelling, merged to experiments of mechatronic systems can improve the monitoring and prognostic capability, in particular when reliability considerations are of concern; nevertheless, the so-called hybrid approach should be simple to implement and of general validity [3];
- networking of sensors is more and more feasible, due to the increased possibility of miniaturizing and interconnecting them, even though trade-off between amount of data and the load in terms of time and computation resources has to be considered. Furthermore, the possibility of using internal sensors at the controllers, which are already present for automation purposes, is often preferable to the installation of external sensors for monitoring purposes;
- features for status description should be synthetic, reliable, accurate, and, above all, meaningful, in order to make easy to the post processing algorithms, the acknowledgement of statuses, which are only slightly different. Features' behaviour depends on many factors, like the physical quantity, the sensor type, the position and so on. Many attempts to select the most suitable features according to general rules have been made, which are able to work for features' selection and/or ranking [4], [5], [6], in an automatic [7 - 8] or unsupervised way [9]; how much these methods are general is still an open discussion.

In this paper, the Authors aim at presenting the generalization of a methodology, which has been already discussed in [10] and [11], with reference to a preliminary application, which is both "hybrid" and based on sensor fusion.

The main steps of the methodology are described in the

following.

Features in the time and frequency domain, which are both typical and/or specifically defined for CM applications [12-15], are evaluated by merging experimental data and computations by a simulation model of the mechatronic system, concerning its kinematic and dynamic behaviour. The model is not used for a preliminary design of the system [16], neither to optimize the positioning accuracy of a manipulator [17]; its main usefulness, is to allow us to indirectly evaluate quantities of interest for CM, also at different positions with respect to the ones the sensors are installed on. Features of interest also refer to differences between measured and estimated values of the same quantities, "fusing" the whole information. The model, by a sensitivity analysis with reference to the effects of interest, indicates also the time windows of experimental and theoretical data, which are more physically related to the phenomenon of interest, improving the physical meaning of the feature.

Based on these considerations, in this paper a set of features will be identified, to be post-processed by ANN, for optimum setting recognition and/or the occurrence of a fault in the kinematic chain.

In Section II, the experimental test case is described, which is a mechatronic system in real scale, together with the steps of the data flow, able to make possible the evaluation of many features in both time and frequency domain.

In Section III, the main results will be presented, concerning the behaviour of the different features, when operating conditions are varied, changing the *jload* setting, or when the friction characteristics are modified by wear or changing the status of lubrication. In particular, the most suitable features will be identified, to selectively detect the different statuses. This will allow us to also reduce the size of the feature's data set, according to metrological and physical criteria, keeping a reliable and resolute procedure.

Conclusions end the paper, indicating also possible developments of the approach, based on the indications the present work makes available.

II. MATERIALS AND METHODS

The mechatronic apparatus under study is depicted in Fig. 1; it realizes a packaging action by an alternate linear and rotating motion.

The scheme of Fig. 2 shows the measurement apparatuses, with reference to both internal sensors to the controller, i.e. angular position and motor currents, and to the external ones, Laser Doppler Vibrometer (LDV) and Micro Electro-Mechanical System (MEMS) tri-axis accelerometer.

The methodology is based on the following main steps:

- Realization of a kinematic and dynamic model of the system, whose outputs are: angular position at the motor, angular velocity at the motor, linear acceleration at the ball screw shaft and current at the motor.
- Realization of experiments, corresponding to different settings, different *jload* parameter, or different operating conditions, varying the lubrication conditions of the ball screw shaft.
- Multiple runs of the model, considering as the input one quantity among the measured ones:
  - Angular position from the encoder, *Mpos*,
  - Angular velocity from the motor controller, *Mvela*,
  - Electric current at the driving servomotor, *Mcur*,
  - Linear acceleration from the accelerometer, *Maccl*,
  - Linear velocity from the LDV, *Mvell*,
- Definition, calculation and selection of the most suitable features for the system status identification. The calculation of the features is carried out on the basis of the comparison between outputs of the model and data deriving from measurements. In addition, some more features are obtained by processing Tracking Deviation (TD) data, supplied by the PLC, and related to the difference between the actual and the theoretical angular position of the motor axis.
- Post-processing of features by ANN, in order to experimentally optimize the whole procedure for CM.

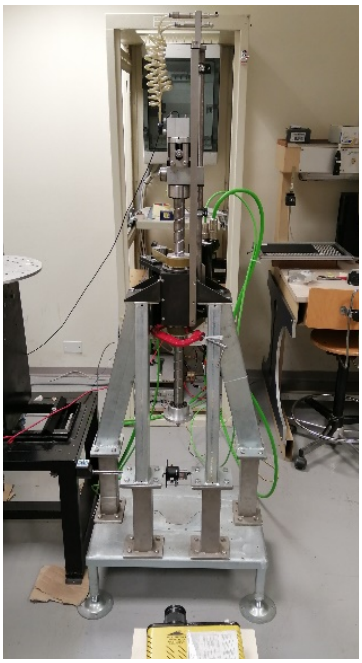


Fig. 1. Picture of the mechatronic device

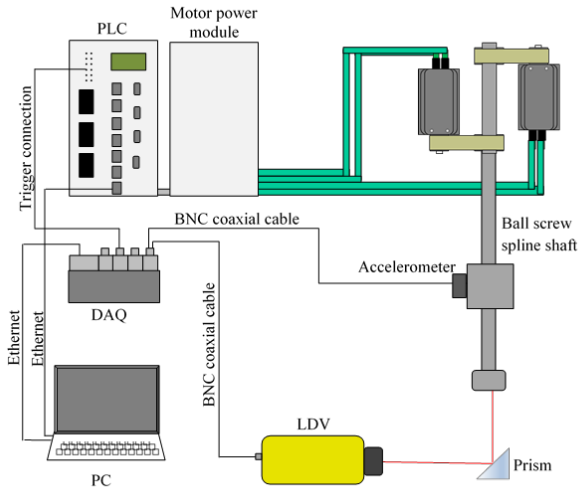


Fig. 2. Scheme of the measurement and data acquisition apparatus

III. RESULTS AND DISCUSSIONS

The tests have been realized varying the *jload* setting in the range 3 to 9 kg cm^2 at three different lubrication conditions:

- *NOG,* corresponding to uncontrolled status of lubrication
- *G1*, corresponding to a only few grease quantity applied on the ball screw
- *G2*, corresponding to the right amount of grease applied on the ball screw.

The sinusoidal motion frequency is 3 Hz, which is a challenging operating situation. The acquisition data rate is 1 kHz, for all the measured quantities.

Although simple and only coarsely validated, the model shows a satisfactory agreement with experimental data, as illustrated in Fig. 3. The comparison is between experimental data of current, *Mcur*, and simulation ones, based on accelerometer measurements, *Scur-Maccl*.

It has to be pointed out that model is used mainly to achieve the following results:

- Evaluation of quantities, even though measured at different points of the system or by other transducers;
- Sensitivity analysis to the specific effects;
- Identification of time windows during the operating cycle, where the effects of a phenomenon of interest are more remarkable.

All these actions are envisaged to improve the physical meaning of the features.

As an example, Fig.4 shows the effect of varying the friction coefficient in the ball screw assembly, with reference to the motor current.

Preliminary analysis has been also carried out with reference to the definition of the set of features to be proposed for analysis by ANN.

More than 90 features have been defined and evaluated in both the time and frequency domain and referred to both measured and simulated quantities: they will be selected, in order to identify a set of accurate and reliable features, of very limited number, to be processed by ANN.

The selection procedure considers metrological aspects, like repeatability and uncertainty of data, and physical likelihood, related to the suggestions by the simulation model. With respect to the ranking and selection proposals, which are available in literature, this approach could be classified as a supervised method; it predicts the feature relevance by physical considerations, without measuring previously the feature correlation with the class label [18].

As an example, the data of Fig. 5 and Fig. 6 show differences of amplitude 98th percentile (*PRC98*) of measured and simulated values of linear acceleration of ball screw and motor current, respectively, when the setting of the system, *jload*, is changed. The values are normalized with respect to the ones obtained when the *jload* is set to 5.5 kg cm^2 and the lubrication condition is *NOG*.

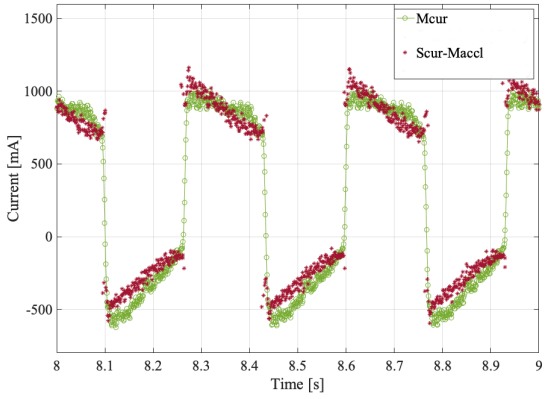


Fig. 3. Model (red) and measured (green) time series

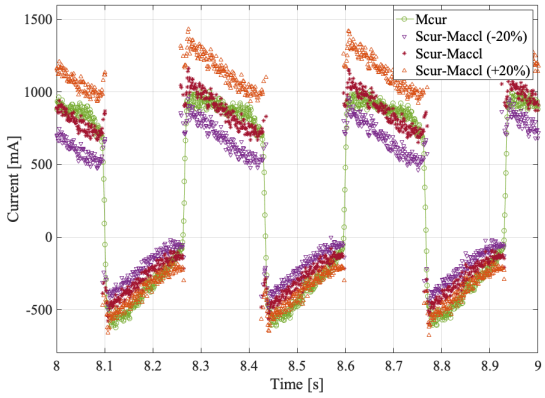


Fig. 4. Sensitivity analysis

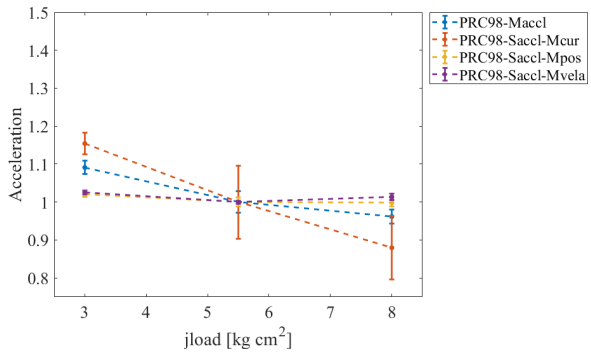


Fig. 5. 98th percentile, linear acceleration, NOG. The values are normalized.

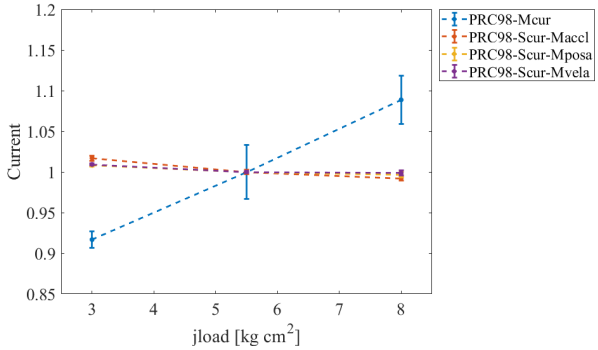


Fig. 6. 98th percentile of measured and simulated currents vs jload, NOG. The values are normalized.

It is quite evident that differences among operating conditions with different settings are remarkable only when linear accelerations or currents are considered, both measured and simulated by the model, (i.e. *PRC98-Saccl-Mcur* and *PRC98-Mcur*).

The temporal trends of TD and motor current in Fig. 7 and Fig. 8, show that differences arise, at the same *jload*, when the lubrication condition changes; in particular, TD seems the most sensitive parameter with respect to variations in the lubrication state. That is confirmed also by features calculated with reference to TD, as the 98th percentile level shows (Fig. 9). It can be seen that the three different lubrication conditions are clearly distinguished, if their variability is considered. The repeatability of data is represented as error bars.

The lubrication effects are shown also by features based on measured current (Fig. 10 and Fig. 11 in comparison with Fig. 5 and Fig. 6), with reduced gap with respect to TD.

The calculated features have been used to train a classification ANN, built to distinguish among the following 9 classes, obtained by combining the different *jload* and lubrication conditions:

1. *jload 3 – NOG*
2. *jload 3 – G1*
3. *jload 3 – G2*
4. *jload 5.5 – NOG*
5. *jload 5.5 – G1*
6. *jload 5.5 – G2*
7. *jload 8 – NOG*
8. *jload 8 – G1*
9. *jload 8 – G2*

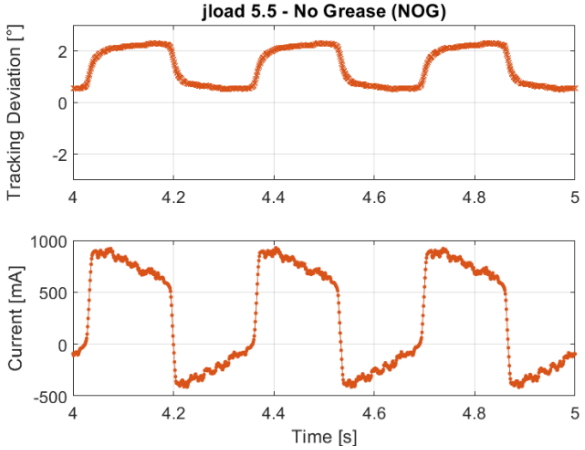


Fig. 7. Temporal trend of TD and current, for the NOG lubrication condition.

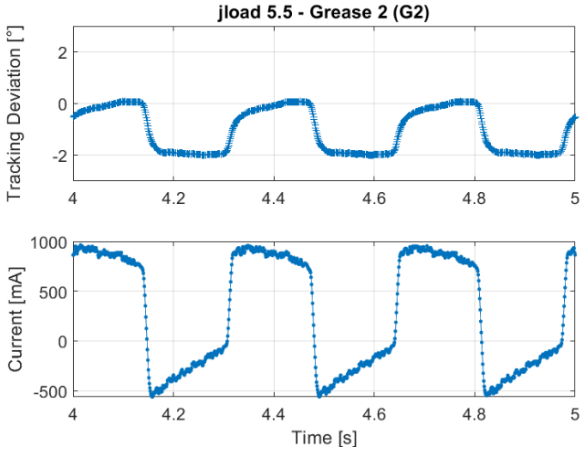


Fig. 8. Temporal trend of TD and current, for the G2 lubrication condition.

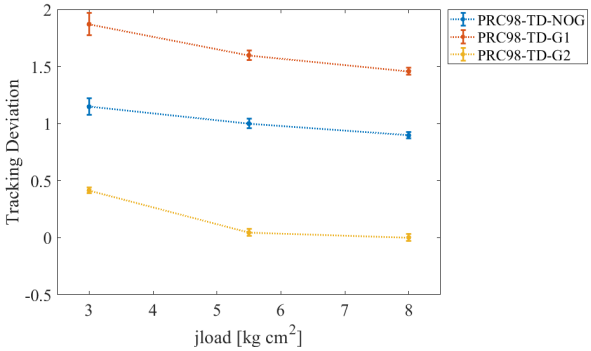


Fig. 9. TD 98th percentile vs jload, for different lubrication conditions. The values are normalized.

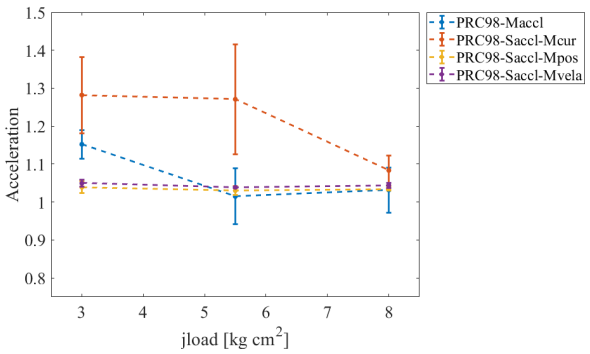


Fig. 10. 98th percentile of measured and simulated accelerations vs jload, for the G2 lubrication condition. The values are normalized.

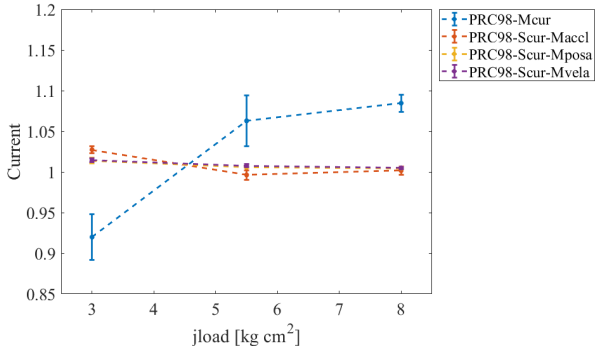


Fig. 11. 98th percentile of measured and simulated currents vs jload, for the G2 lubrication condition. The values are normalized.

The ANN results are summarized in Fig. 12 and Fig. 13, comparing the case when the whole set of features is used for training the ANN, and that related to a selected and optimized group of features, respectively. Fig. 12 shows that some classes are correctly recognized, when the output of the ANN is close to 1; some others are not recognized or are identified with a low output value.

The results clearly improve, as shown in Fig. 13, when the limited set of features is considered, based on TD or measured current and on a mixing of measured and simulated values of current.

The good performance of the classification algorithm is assessed by low values of the Cross-Entropy (CE), that is a performance index for classification ANNs [19]. The CE is reduced down to 0.007, when the selected group of features is used for training.

IV. CONCLUSIONS AND OUTLOOK

In this paper, the results are described of the application of a CM methodology for mechatronic systems, which is based on the parallel use of a kinematic and dynamic model of an automated device and experimental data.

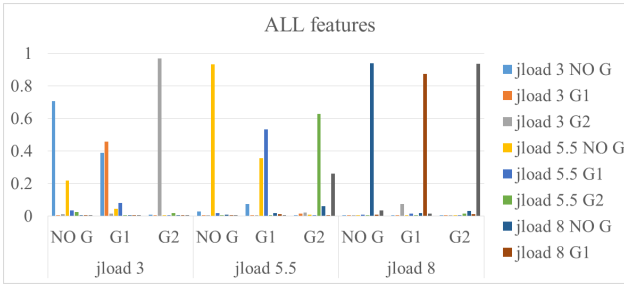


Fig. 12. Classification results, when the neural net is trained using all features.

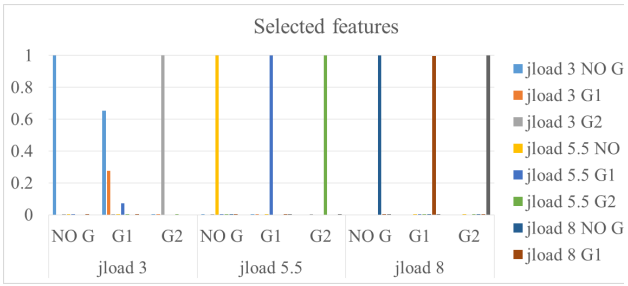


Fig. 13. Classification results, when the neural net is trained using only selected features.

The hybrid approach allowed us to define many features in the time and frequency domain, which are based also on statistics. Selection criteria of features have been applied, based on metrological and physical considerations, in order to supervise the automatic process of identification of the condition of interest. In particular, repeatability was considered as a good criterion for selectivity, together with aspects connected to the mechanical behaviour of the system.

The features mostly able to recognize differences between different *jload* settings and lubrication conditions, appear to be those related to the motor current and the TD.

Based on these consideration, a limited set of features has been identified and proposed, to be automatically processed by ANN, showing a better detection capability for status identification with respect to the whole set.

The obtained results confirmed experimentally the capability of this CM methodology to resolve even little changes of the operating status of the device, due both to settings and lubrication conditions.

Future work will be devoted to apply this methodology to the identification of defects occurrence and propagation in more complex mechatronic systems. This aim will require further improvement of selectivity of the method, proving its accuracy and robustness.

REFERENCES

[1] **Leturiondo U., Salgado O., Galar D.,** Validation of a physics-based model of a rotating machine for synthetic data generation in hybrid diagnosis, *Structural Health monitoring* Vol. 16(4), 2017, pp. 458-470.

[2] **Polok B., Bilski P.,** Analysis of the RBF ANN-based classifier for the diagnostics of electronic circuit, Acta Imeko, Vol. 7(1), 2018, pp. 42-49.

[3] **Leturiondo U., Salgado O., L. Ciani, Galar D., Catelani M.,** Architecture for hybrid modelling and its application to diagnosis and prognosis with missing data, Measurement, Vol. 108, 2017, pp. 152-162

[4] **Zhang, B., Katinas, C., Shin, Y.C.,** Robust tool wear monitoring using systematic feature selection in turning processes with consideration of uncertainties, *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, Vol. 140, Issue 8, 1 August 2018, Article number 081010

[5] **Lipsett, M.G.,** How system observability affects fault classification accuracy, and implications for sensor selection andplacementfor condition monitoring, *WCCM 2017 - 1st World Congress on Condition Monitoring,* London; United Kingdom; 13 to 16 June 2017.

[6] **Shao, C. Paynabar, K., Kim, T.H., Jin, J., Hu, S.J., Spicer, J.P., Wang, H., Abell, J.A.,** Feature selection for manufacturing process monitoring using cross-validation, *Journal of Manufacturing Systems*, Volume 32, Issue 4, October 2013, Pages 550-555

[7] **Ferreira B., Silva R.G., Pereira V.,** Feature selection using non-binary decision trees applied to condition monitoring, *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA4* January 2018, Pages 1-ETFA 2017; Limassol; Cyprus; 12 September 2017 through 15 September 2017

[8] **Möhring, H.-C., Maier, W., Werkle, K.,** Increasing the Accuracy of an Intelligent Milling Tool with Integrated Sensors, *Proceeding of 18th International Conference and Exhibition, EUSPEN 2018*, Pages 235-236, Venice; Italy; 4 - 8 June 2018; Code 138400

[9] **Silva,R.G., Wilcox,S.J**, Feature evaluationand selection fo r condition monitoring using a self-organizing map and spatial statistics, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM* Volume 33, Issue 1, 1 February 2019, Pages 1-10

[10] **D'Emilia G., Gaspari A., Natale E.,** Sensor fusion for more accurate features in condition monitoring of mecha-tronic systems, *I²MTC, 2019 IEEE international Instrumentation & Measurement Conferenc*e, Auckland, New Zealand, May 20-23, 2019.

[11] **D'Emilia G., Gaspari A., Natale E.,** Integration of model and sensor data for smart condition monitoring in mecha-tronic devices, *Conference "Metrology for Industry 4.0",* June 4-6, 2019, Napoli (Italy).

[12] **Lauro C.H., Brandão L.C., Baldo D., Reis R.A., Davim J.P.,** Monitoring and processing signal applied in machining processes – A review, Measurement, 58 (2014), 78-86.

[13] **Diez Olivan A., Del Ser J., Galar D. and Sierra B.,** Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0, *Information Fusion*, vol. 50, 2019, pp. 92-111.

[14] **Zhang Y., Duan L. and Duan M.**, A new feature extraction approach using improved symbolic aggregate approximation for machinery intelligent diagnosis", *Measurement*, Vol. 133, 2019, pp.468-478.

[15] **D'Emilia G. and Natale E.,** "Instrument comparison for integrated tuning and diagnostics in high performance automated systems", *Journal of Engineering and Applied Sciences*, 2018, Vol. 13, No.3.

[16] **Saxena S. and Sharma P.,** An approach to the analysis of higher linkage mechanisms and its validation via matlab, *Proceedings of 3rd IEEE International Conference on "Computational Intelligence and Communication Technology"* (IEEE-CICT 2017)

[17] **Ozgur E., Dahmouche R., Andreff N. and Martinet P.**, A vision-based generic dynamic model of PKMs and its experimental validation on the Quattro parallel robot" *Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM),* Besançon, France, July 8-11, 2014.

[18] **Sheikpour R., Sarram M.A., Gharaghani S., Chahooki M.A.Z.,** A Survey on semi-supervised feature selection methods, *Pattern Recognition*, 64 (2017) 141–158.

[19] **Dreiseitl S., Ohno-Machado L.**, Logistic regression and artificial neural network classification models: a methodology review, J. Biomed. Inform., vol. 35, n. 5, pp. 352-359, 2002.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"
Berlin, Germany, on September 3-4, 2019

Product quality and cutting tool analysis for micro-milling of ceramics

L Móricz[1], Zs. J. Viharos[2,3], A. Németh[4], A. Szépligeti[4]

[1]*University of Pannonia, Faculty of Engineering, Mechatronic Education and Research Institute H-8900, Gasparich Márk str. 18/A., Zalaegerszeg, Hungary*
[2]*Centre of Excellence in Production Informatics and Control, Institute for Computer Science and Control of the Hungarian Academy of Sciences (MTA SZTAKI), H-1111, Kende str. 13-17., Budapest, Hungary*
[3]*John von Neumann University, H-6000, Izsáki str. 10, Kecskemét, Hungary*
[4]*AQ Anton Kft., H-8900, Sport str. 16., Zalaegerszeg, Hungary*

***Abstract* – Based on their favourable mechanical features, applications of ceramics are continuously spreading in industrial environment. Such a good feature is their resistance against heat shock, so, currently they are applied e.g. as coating material for gas turbines. The main aims of the paper to follow the wearing process of the micro-milling tool during machining of ceramics and to compare it against the geometrical changes of the machined ceramic workpiece, applied as a special monitoring technique.**

Keywords: cutting of ceramics, regular cutting edge geometry, micro-milling, product quality, tool wear

I. INTRODUCTION

Machining of rigid materials with regular cutting edge geometry is one of the main trends in the 21th century. Ceramics are such rigid materials that are employed more and more widely as raw materials thanks to their high hardness and thermal resistance [1][2]. There are various options for machining them, e.g. using water, laser or abrasive grinding [3][4][5], however, their high costs and complex setups are important drawbacks of these technologies. Therefore, the machining of ceramics with a classical, regular cutting edge geometry is still a promising solution, however, considering the relative quick wearing process of the cutting tool without an appropriate technological optimisation, this methodology will be economically not acceptable.

Optimizing a technology is typically a multicriteria assignment, like here, the main aim is to find the smallest production cycle time, and in the same time the tool life has to be maximal, too. The tool wear state serves with information about the actual wearing stage of the tool.

Bian et al. [6] used ultra-miniature diamond coated tools to mill fully sintered ZrO_2. Wear of the tool was described by three stages:
1. Early coating delamination;
2. Extended coating delamination with slight wear of the exposed cutting edge;
3. Severe wear of the tool blank.

Romanus et al. [7] analysed the tool wearing process and the resulted surface quality during cutting of also ZrO_2 by diamond and cBN coated milling tools. They observed the same three stages of the wearing.

Bian et al. described experimental results in another publication [8] using WC milling tool coated by different sizes of diamond grains with the aim for analysing the peeling process of the diamond coating. They concluded that when the machining in the classical "X-Y" direction the cutting force is significantly smaller as in the "Z" direction. The reason for that is, in case of "X-Y" they managed to keep the cutting in the plastic deformation area, while in the "Z" direction the brittle area was dominating. They identified also that with the increase of the feed per tooth and axial cutting depth the cutting force will grow significantly.

Based on the literature review one can identify that the research on cutting of hard and brittle materials by advanced tools is still in an early stage, especially as far it concerns micro scale applications. Based on the analysis conducted, the main challenges can be summarized as follows:

a) availability of ultra-miniature tools having appropriate cutting edge geometry to ensure proper compression/tensile stress distribution in

the chip zone;

b) appropriate tool stiffness and sufficient wear resistance;
c) identification of suitable cutting parameters to ensure ductile material removal and damage-free machined surfaces. [7]

The main aims of the current paper are - continuing a previous research - to follow the wearing process of the micro-milling tool during machining of ceramics and to compare it against the geometrical changes of the machined ceramic workpiece, applied as a special monitoring technique.

II. FACTORS INFLUENCING THE CUTTING TOOL LIFE

There are plenty of technological parameters that influence the micro-milling of ceramics. However, the literature shows that the most influencing parameters are the following:

Technological parameters

- axial cutting depth
- radial cutting depth
- cutting speed
- feed(rate)
- tool tilt angle

Geometry of cutting tool

- flank angle
- rake angle
- number of teeth

Tool material, coating

- basic tool material
- type of coating
- thickness of coating
- grain size
- number of coating layers

Type of cooling

Other parameters

- combined shaping technology

An important assignment for the cutting process with regular geometry tool is to keep the cutting in plastic deformation area instead of in brittle area [8].

K. Ueda and his colleagues cut plenty of ceramics materials to find solution for this question, and they found that materials that have high fracture toughness can be cut easier in plastic deformation area with optimal cutting speed and feed [8]. However, with low fracture toughness they did not find a parameter combination that can cut in the ductile deformation zone.

The other author of the mentioned publication who conducts some research into the ductile-brittle range in shaping method is Muhammed A, who tried to find what those cutting parameters are, and how they influence the ductile range [9]. It was found that there are critical parameters belonging to the cutting depth and feed values which affects the chip-removal mechanism. These results are summarized in Figure 1.

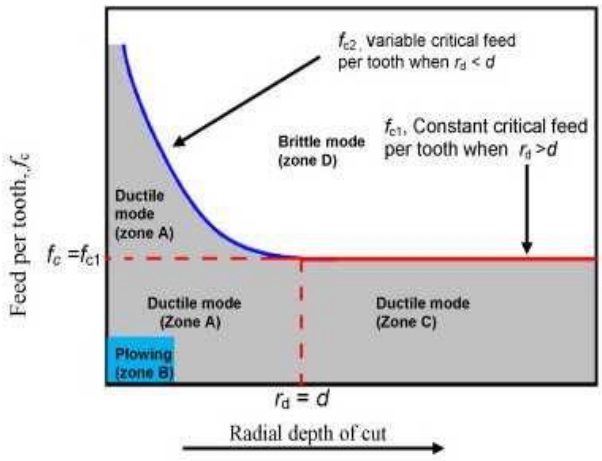


Fig. 1. Critical value of feed-depth to reach the plastic deformation range [9]

During their experiments it was found that there is a critical value of cutting depth and feed, where the analysed material can be shaped in the ductile removal area. Under a certain cutting depth, the value of feed can be increased without the removal of the cutting mechanism from the brittle area. This means that the plastic deformation strongly depends on the value of the cutting depth.

Differences should be considered between two types of cutting depths, axial cutting depth and radial cutting depth [7][9]. Another important aspect of the mechanism is the tilt angle of the tool [10].

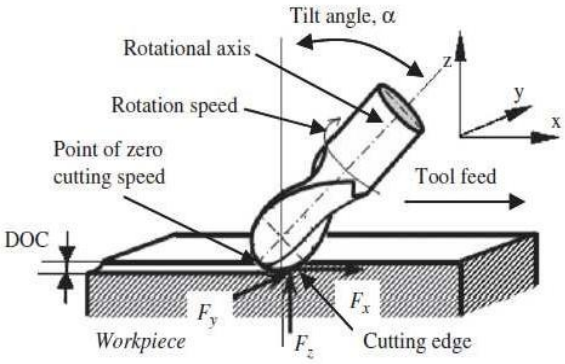


Fig. 2. Tilted milling tool [10]

This parameter is highly important in cutting with ball milling tool, as the cutting speed depends on the tilt angle. The scientific literature represents that the optimal tilt angle is between 40-60° that depends mainly on the ceramics' material [10].

A combined technology is a further interesting method for enhancing the lifetime of a cutting tool: Toru et al. [11] combined the conventional cutting technology with laser technology. J. Feng et al. [12] performed machining with irregular cutting tool geometry while they made conclusions on the cutting tool's wear based on cutting force and vibration measurements. Xiaohong and his colleagues [13] preformed similar research with laser-assisted cutting. They found that the laser has a strong influence on the grinding characteristics. The normal and tangential grinding forces for laser assisted grinding is 15% lower than that in the conventional grinding method.

In previous researches of the authors the linear, Taguchi based Design of Experiment (DoE) were applied to analyse the effect of the varying technological parameters on the tool wearing [14]. The results served with a quasi (linear) optimal technology concerning the maximisation of material removal speed; however, the analysis of the wearing process is an open issue, because a novel indirect measurement could be applied before [15]. The current paper goes beyond the previously applied methods, because the complete wearing behaviour of the tool is described.

III. EXPERIMENTS WITH MACHINING USING SMALL PITCH ANGLES

Micro-milling tools were applied for the ceramics cutting experiments.

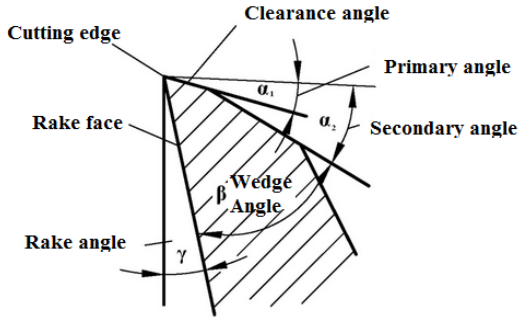


Fig.3. Tool geometry and angles

According to the available literature, the cutting edge and the flank edge (Fig. 3.) are loaded on the highest level during such machining [7][8].

Parameters of the milling machine

The basis of the experiments was the milling machine that was planned and built by the CncTeamZeg (student) group. The machine is operated partly by the Mechatronics Institute of University of Pannon in Zalaegerszeg (Fig. 4.). During the planning the aim was to cut metal material but the preliminary calculations and first tests on ceramic material removal proved that it is able do machining on ceramic materials, too.

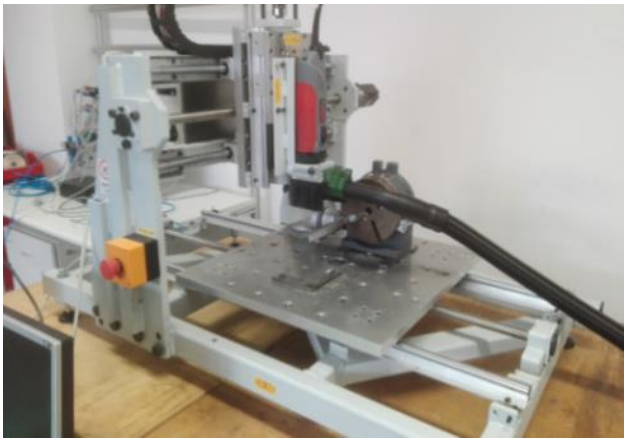


Fig.4. The applied three axis cutting machine

Parameters of the laboratory milling machine:

- Power: 1050 W
- Maximum tool diameter: 8 mm
- Spindle speed: 5000-25000 1/min
- Work area: 500x250x180 mm

Tool path

The path which was used during the experiments was created using the of EdgeCam software. A very small part of the tool path is presented in Fig. 5. representing mainly of linear movements and machining along circle slices having various radiuses.

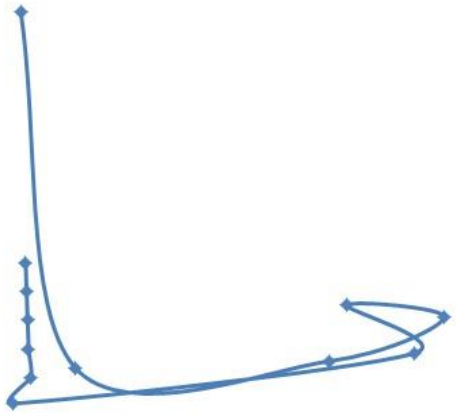


Fig.5. One simple, small part of the generated tool path

The applied cycloid form is a special milling technology where the milling tool is going along an arc, avoiding sharp changes in the direction.

The wearing-out process

The optimum cutting conditions determined in previous research [13] were applied in the current analysis, while the tool wear was measured by microscope in predefined intervals. These microscope measurements are presented in the next figures, representing the wearing-out process (Fig. 6-9.). From the previous experiments the approximate lifetime of the tool was also identified. Using this information, the specific stages of the tool lifetime are known.

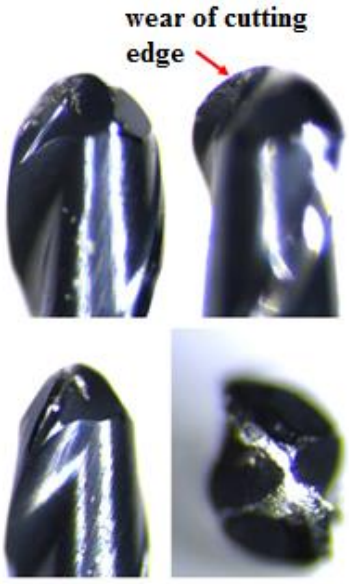


Fig. 6. Tool wear at 20% of the tool life

The first inspection point was defined for the 20. percent of the complete lifetime. The first wear form was observed on the cutting edge.

Fig. 7. Tool wear at 40% of the tool life

The second inspection point was defined for the 40 percent of the complete lifetime, the Fig. 7. shows the wearing of chisel edge and an early wear of flank.

Fig. 8. Tool wear at 60% of the tool life

Strong flank wear was observed at 60 percent of the lifetime and the first burning-out with geometrical distortion arose.

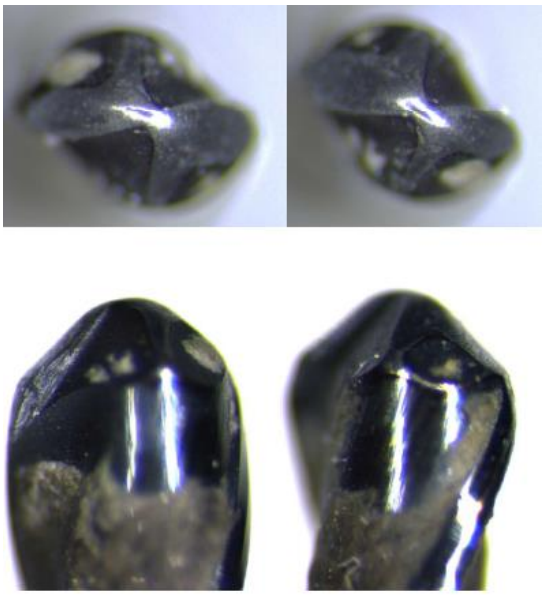


Fig. 9. Tool wear at the end of the tool life

Through the optimisation of the technological parameters more than 40% increase in the cutting tool life was achieved. On the final point the chisel edge was fully wearing out, and the tool geometry was strongly distorted, as shown in Fig. 9. Consequently, there were no sense to continue these experiments.

In the current paper only the cutting tool wearing forms were examined together with the specification in which section of tool lifetime they appear. To give nominal, directly measured values for tool degradation will be one of the next steps in the research.

Analysis of the cutting tool wear process based on geometrical changes of the machined workpiece

The wearing process can be followed also by indirect measurements of the resulted workpiece geometry, see the related details in paper [14]. Two machined geometrical features (Geometry I. and II.) were measured (Y axes), represented in Fig. 10., while the horizontal axes represents the number of machining cycles.

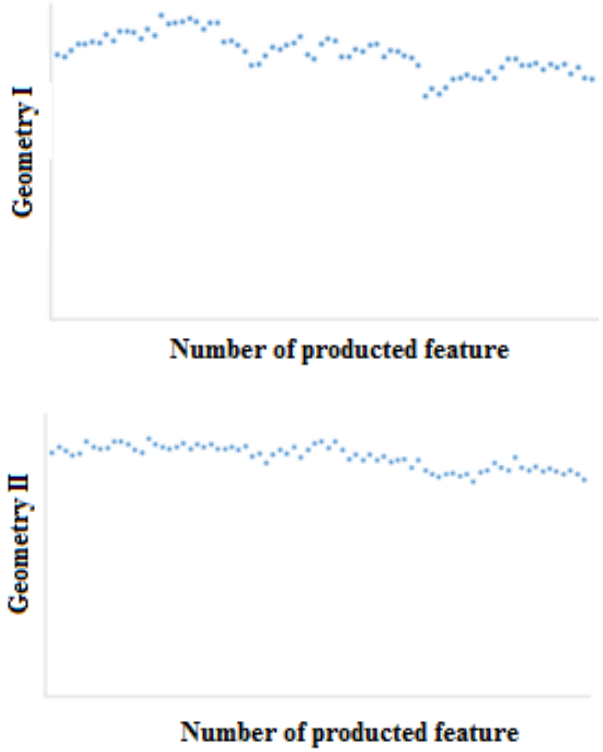


Fig 10. Changes in the Geometry I. and II. of the machined features

Fig 10. mirrors some repeated errors based on the measurements (increasing parts in the y values), however these are in relation to the uncertain behaviour of the workpiece re-positioning after the separated microscope measurements of the tool wearing. Consequently, it can be eliminated. Finally, the tool wearing trends were identified, since the wearing process started at the tool's chisel edge.

IV. RESULTS AND DISCUSSIONS

Based on the experiments the wear of the tool can be described in four stages:

1. Cutting edge and rake face early delamination.
2. Wearing of the chisel edge, flank face early delamination.
3. Strong flank face and chisel edge wearing.
4. Wear-out of the complete geometry.

The length of the stages can be influenced by technological optimisation:

- One of the most influencing factors is the cutting speed. Under a critical value consequent tool breakage was observed within a very short time period.
- Fast wearing of the cutting tool was observed when selecting an inappropriate value for the tilt angle. In this case the same wearing stages can be identified as described above, but much quicker.
- Comparing the wearing forms of the cutting tool and the geometrical changes of the machined workpiece, it was found that the wearing of chisel edge has the strongest influence on geometrical measures.

Further research directions

Research directions for the close future can be also formulated:

- Deeper insight into the relations between tool wearing and technological parameters.
- Optimisation of technological parameters to eliminate the negative effect of the chisel edge.
- Measurement method for nominal values of the factors of tool degradation.
- Monitoring and diagnostics of the cutting tool using vibration sensors and advanced data analysis.

V. CONCLUSIONS

Based on their favourable mechanical features, applications of ceramics are continuously spreading in industrial environment, however, there are many open issues in their machining, e.g. in cutting them with regular tool geometry.

During the reported research, by optimising the technological parameters, more than 40% increase in the cutting tool life was achieved At the experiments the wearing progress was analysed in parallel to the changes in the machined workpiece geometry and it was identified that the tool wearing is the key factor influencing the resulted workpiece accuracy and quality. Scientific progress has been reached in three fields:

- Increase in the cutting tool life by more than 30%.
- Monitoring and diagnostics on the cutting tool

wear directly and indirectly, and comparison of them.

- Analysis of the cutting tool also after the damage and wear-out of the tool coating.

The reported research allowed to improve and stabilise the cutting of ceramics significantly by micro-milling with regular tool geometry.

VI. ACKNOWLEDGMENTS

First of all, thanks have to be expressed to the AQ Anton Kft., serving with materials, equipment, challenges, experts any many other resources enabling and supporting the research. The research in this paper was partly supported by the European Commission through the H2020 project EPIC (https://www.centre-epic.eu/) under grant No. 739592 and by the Hungarian ED_18-2-2018-0006 grant on a "Research on prime exploitation of the potential provided by the industrial digitalisation".

REFERENCES

[1] **Peter, J.:** Modern gas turbine systems, High Efficiency, Low Emission, Fuel Flexible Power Generation. *Woodhead Publishing Series in Energy*: No 20, 2013., pp. 8-38.

[2] **Móricz, L.; Viharos, Zs. J.:** Trends on applications and feature improvements of ceramics, *Manufacturing 2015 Conference*, Budapest University of Technology and Economics, Vol. 15, No., 2, 2015, pp. 93-98

[3] **Lingfei, J.; Yinzhou, Y.; Yong, B.; Yijian, J.**: Crack-free cutting of thick and dense ceramics with CO2 laser by single-pass process. *Optics and Lasers in Engineering 46*., 2008, pp. 785-790.

[4] **Jiyue, Z.; Thomas, J. K.**: An erosion model for abrasive waterjet milling of polycrystalline ceramics. *Wear* 199 1996, pp. 275-282.

[5] **Móricz, L.; Viharos, Zs. J.**: Optimization of ceramic cutting, and trends of machinability, *17th International Conference on Energetics-Electrical Engineering - 26th International Conference on Computers and Educations,* Hungarian Technical Scientific, Society of Transylvania, 6-9 October, 2016, pp. 105-110.

[6] **Bian, R.; Ferraris, E.; Qian, J., Reynaerts, D.; Li, L.; He, N.**: Micromilling of fully sintered ZrO_2 ceramics with diamond coated end mills, *Conf on Prec Eng - Key Eng Mater, 523(524)*, 2012, pp. 87-92.

[7] **Heleen, R.; Eleonora, F.; Jan B.; Dominiek, R.; Bert, L.**: Micromilling of sintered ZrO2 ceramic via CBN and diamond coated tools, *Procedia CIRP 14, 2014*, pp. 371 – 376.

[8] **Ueda, K.; Sugita, T.; Hiraga, H.**:A J-integral Approach to Material Removal Mechanisms in Microcutting of Ceramics, *CIRP Annals - Manufacturing Technology 40*, 1991, pp. 61–64.

[9] **Muhammad, A.; Mustafizur, R.; Wong, Y. S.**: Analytical model to determine the critical conditions for the modes of material removal in the milling process of brittle material. *Journal of Materials Processing Technology 212*, 2012, pp. 1925– 1933.

[10] **Kevin, F.; Zhi, W.; Takash,i M.; Yong, H.**: Effect of tilt angle on cutting regime transition in glass micromilling. *International Journal of Machine Tools & Manufacture 49*, 2009, pp. 315–324.

[11] **Toru, K.; Ogasahar, T.; Sugita, N.; Mitsuishi, M.**: Ultraviolet-laser-assisted precision cutting of yttria-stabilized tetragonal zirconia polycrystal, *Journal of Materials Processing Technology 214*, pp. 267– 275.

[12] **Fen J, Kim B S, Shih A, Ni J** 2009 Tool wear monitoring for micro-end grinding of ceramic materials. *Journal of Materials Processing Technology 209*, pp. 5110–5116.

[13] **Xiaohong Z, Zhicheng Z, Zhaohui D, Si L, Qiaoping W, Zhongxiong K** 2019 Precision grinding of silicon nitride ceramic with laser macro-structured diamond wheels *Optics and Laser Technology 109*, pp 418–428.

[14] **Móricz, L.; Viharos, Zs. J.; Németh, A.; Szépligeti, A.**: Efficient Ceramics Manufacturing through Tool Path and Machining Parameter Optimisation, *15th IMEKO TC10 Workshop on Technical Diagnostics*, Budapest, Hungary, June 6-7., 2017, pp. 143-148.

[15] **Móricz, L.; Viharos, Zs. J.; Németh, A.; Szépligeti, A.**: Indirect measurement and diagnostics of the tool wear for ceramics micro-milling optimisation, *XXII IMEKO World Congress*, September 3-6, Belfast, United Kingdom, *Journal of Physics: Conference Series (JPCS)*, Vol. 1065, 2018., paper ID: IMEKO-OR-135.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Metrological Performances of Current Transformers Under Amplitude Modulated Currents

Y. Chen[1], G. D'Avanzo[2], A. Delle Femine[2], D. Gallo[2], C. Landi[2], M. Luiso[2], E. Mohns[1]

[1]*Physikalisch-Technische Bundesanstalt, Braunschweig, Germany*
[2]*Dept. of Engineering, University of Campania "Luigi Vanvitelli", Aversa (CE), Italy*

***Abstract* – All the measurement tasks, f.i. metering, protection and diagnostics, in electricity distribution and trasmission grids need instrument transformers (ITs) to scale voltage and current down to levels compatible with measuring instruments. Inductive voltage and current transformers (VTs and CTs) are still the most used ITs; however, their behaviour in presence of disturbances is not fully assessed. This paper analyses the behavior of CTs in presence of amplitude modulations, which is one of the disturbance under which the performance of Phasor Measurement Units (PMUs) has to be assessed. A measurement setup, able to test current transformers in accordance to the IEEE Std. C37.118.1 about the PMU performance testing, has been realized. Experimental tests on a commercial CT, changing the modulation frequency, are presented.**

Keywords* – *Current Transformer, Power System Diagnostics, Phasor Measurement Unit, Amplitude Modulations.

I. INTRODUCTION

All the measurement tasks, from metering to protection, from energy saving to diagnostics, that have to be performed in the electricity Transmission and Distribution Systems (TSs and DSs) need the measurement of grid voltage and current 1. Voltage and current levels spans from hundreds of volt and tens of ampere up to hundreds of kilovolt and tens of kiloampere [2]. This is the reason why all the measuring instruments for TSs and DSs need voltage and current instrument transformers (ITs) to make voltage and current levels compatible with their input range.

Although sensor technologies advance and new kind of voltage and current instrument transformers (active or passive, generally referred to as Low Power Instrument Transformers, LPIT) are available on the market [3,4], inductive voltage and current transformers (VTs and CTs) are still the most used ITs in TSs and DSs. Therefore, as the modern TSs and DSs experience a wide growth of conducted disturbances [5–9], (generally referred to as Power Quality, PQ), inductive ITs are used for PQ monitoring, as input stage of PQ measuring instruments [10].

Inductive ITs are designed to work at power frequency (i.e. 50/60 Hz) and, moreover, some recent literature [11, 12] show that inductive ITs suffer from an intrinsic non-linearity which is responsible for uncertainties up to some percent when measuring PQ phenomena. However, currently, the behavior of the inductive ITs in PQ measurement is not fully addressed.

The work presented in this paper is inserted in the framwork of the European project EMPIR 17IND06 Future Grid II: one of the objectives of this project is to define a set of test waveforms, representative of TSs and DSs real signals and that goes beyond the existing international standards, with which evaluate the performance of ITs in conditions similar to those that can be encountered in an actual power system.

In particular, one of the disturbance that can can be encountered in TSs and DSs are the amplitude modulation of voltage and current signals. Therefore, this paper is focused on the verification of the performance of CTs when their input currents are affected by undesired amplitude modulation.

Some paper in scientific literature face the issue of testing the performance of CT in non-sinusoidal conditions 13, 14. However, to the best of the author's knowledge, no paper reports experimental on CTs when its input is affected by amplitude modulation.

The international standard IEEE C37.118.1 [15] and its amendment [16], focused on synchrophasor measurement methods and performance verification of Phasor Measurement Units (PMUs), deals with amplitude modulations and proposes some tests to verify the performance of PMUs in their presence. PMUs have crucial role in modern TSs and DSs because they are used to implement remote diagnostic system of power grid by focusing on the synchronized phasor of fundamental component. In fact, the scope of a PMU is to measure the phasors of the voltage and current fundamental component, synchronized to the Universal Time Coordinated (UTC), even when they are affected by PQ disturbances. The same considerations made for PQ instruments are valid also for PMUs: inductive VTs and CTs are frequently used as input stage of PMUs for the measurement of TSs and DSs voltage and current fundamen-

tal synchrophasor. In this context, it would be very useful to know the additional uncertainty introduced by VTs and CTs to the fundamental synchrophasor, other than that of the PMU itself. Therefore, the approach followed in this paper is to evaluate the uncertainty contribution of CTs to the measurements performed by a PMU, when the currents are affected by amplitude modulation. The paper is organized as follows: Section II deals with the definitions and the testing procedure. Section III describes the measurement setup whereas Section IV presents the preliminary experimental results of the characterization of a commercial CT. Finally, Section V draws the conclusions.

II. DEFINITIONS AND TESTING PROCEDURE

A. Definitions

The current signals, $i(t)$, used for the scope of the paper are composed by a fundamental component (i.e. the carrier) at power frequency and a sinusoidal amplitude modulating component, which can be written as in equation (1):

$$i(t) = \sqrt{2}X_m(1 + k_a \cos(2\pi f_a t)) \cos(2\pi f t) \quad (1)$$

where X_m and f are, respectively, the root mean square (rms) amplitude and the frequency of the fundamental component, whereas k_a and f_a are, respectively, the amplitude modulation factor and the modulation frequency.

According to [17] the CT ratio and phase errors are evaluated as in equation (2) and (3).

$$\epsilon = \left(\frac{K I_2}{I_1} - 1 \right) * 100 \quad (2)$$

$$\Delta\phi = \phi_2 - \phi_1 \quad (3)$$

where K is the rated ratio of the CT, I_1 (I_2) and ϕ_1 (ϕ_2) are the magnitude and phase angle of the fundamental primary (secondary) current phasor, respectively.

According to [15, 16], the Total Vector Error (TVE) is evaluated as in equation (4).

$$TVE = \sqrt{\frac{(\Re(K\boldsymbol{I_2}) - \Re(\boldsymbol{I_1}))^2 + (\Im(K\boldsymbol{I_2}) - \Im(\boldsymbol{I_1}))^2}{\Re(\boldsymbol{I_1})^2 + \Im(\boldsymbol{I_1})^2}} \quad (4)$$

where $\Re(\boldsymbol{I_1})$ ($\Re(K\boldsymbol{I_2})$) and $\Im(\boldsymbol{I_1})$ ($\Im(K\boldsymbol{I_2})$) are, respectively, the real and imaginary parts of the fundamental primary (secondary, multiplied by CT rated ratio) current phasor.

B. CT Testing Procedures

As it is written in the introduction, the aim of the paper is to evaluate the uncertainty contribution of CTs to the measurements performed by a PMU, when the currents are affected by amplitude modulation. Therefore, the aim of the paper is graphically represented in Fig. 1, where the functional block diagram of the measurements is depicted. The same current is sensed by a reference CT (CT_{REF}) and by the CT under test (CT_{UT}). Their outputs are acquired and processed by two identical PMUs (PMU_1 and PMU_2), which returns the fundamental synchrophasors of the outputs of CT_{REF} and CT_{UT}, i.e. $\boldsymbol{I_{REF}}$ and $\boldsymbol{I_{UT}}$. After the calculation of $\boldsymbol{I_{REF}}$ and $\boldsymbol{I_{UT}}$, the quantities in equations (2) - (4) are evaluated: in fact, the equations (5) and (6) apply:

$$\boldsymbol{I_1} = K_{REF}\boldsymbol{I_{REF}} \quad (5)$$

$$\boldsymbol{I_2} = \boldsymbol{I_{UT}} \quad (6)$$

where K_{REF} is the ratio of the reference CT. In this way, since two identical PMUs measure the fundamental synchrophasor of the same current, but using two different CTs, the sole contribution of the CT under test to the uncertainty on ϵ, $\Delta\phi$ and TVE is evaluated. In particular, since in the realized setup (as it is explained in Section III) generation and acquisition are synchronized, the synchrophasors $\boldsymbol{I_1}$ and $\boldsymbol{I_2}$ are evaluated by performing the Discrete Fourier Transform (DFT) on the output samples of the CT_{REF} and CT_{UT}. The DFT is executed on an observation window of four cycles of the fundamental frequency.

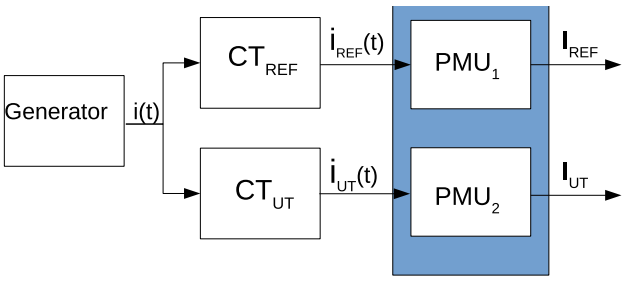


Fig. 1. Functional block diagram of the measurements

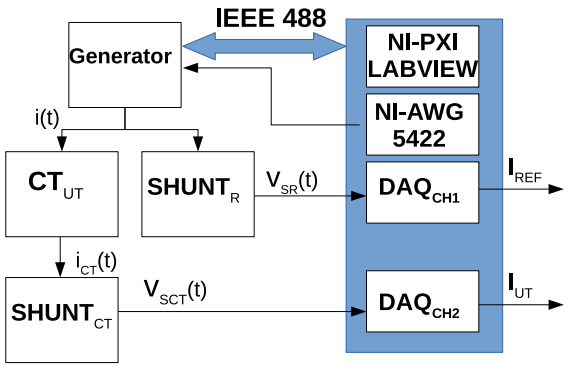


Fig. 2. Block diagram of the measurement setup

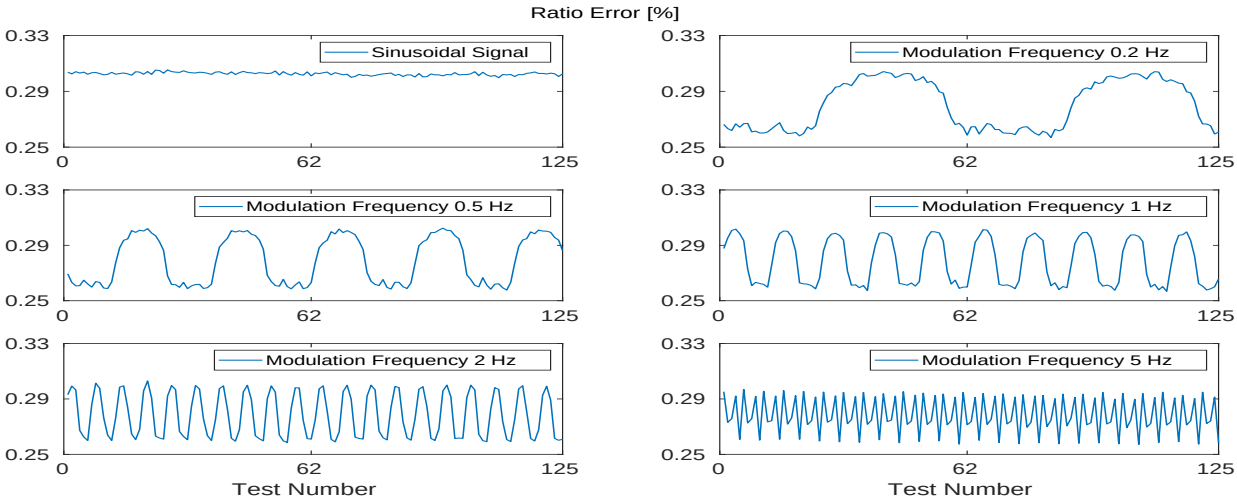


Fig. 3. Ratio Error, in sinusoidal conditions and with amplitude modulation

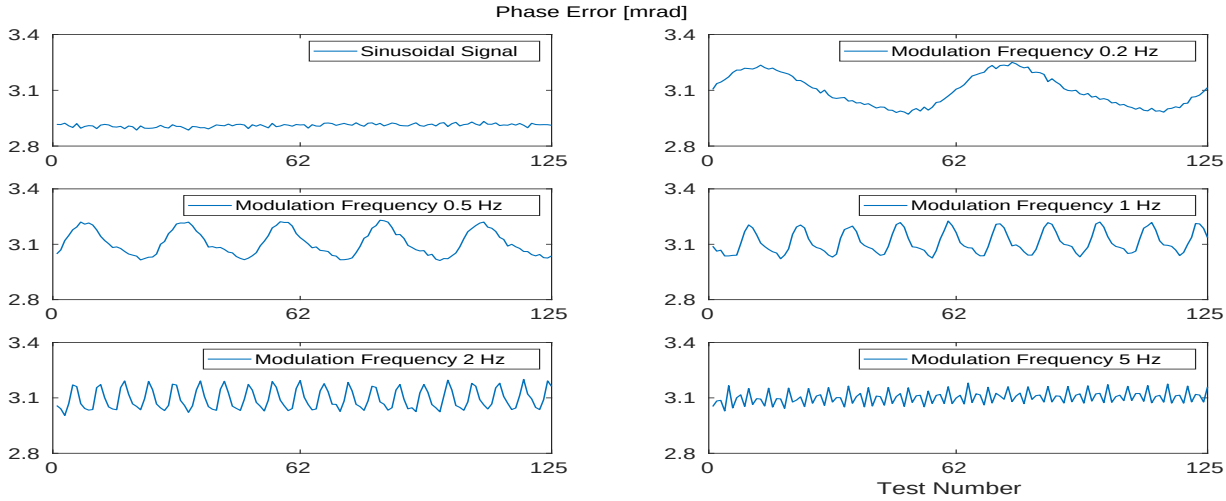


Fig. 4. Phase Error, in sinusoidal conditions and with amplitude modulation

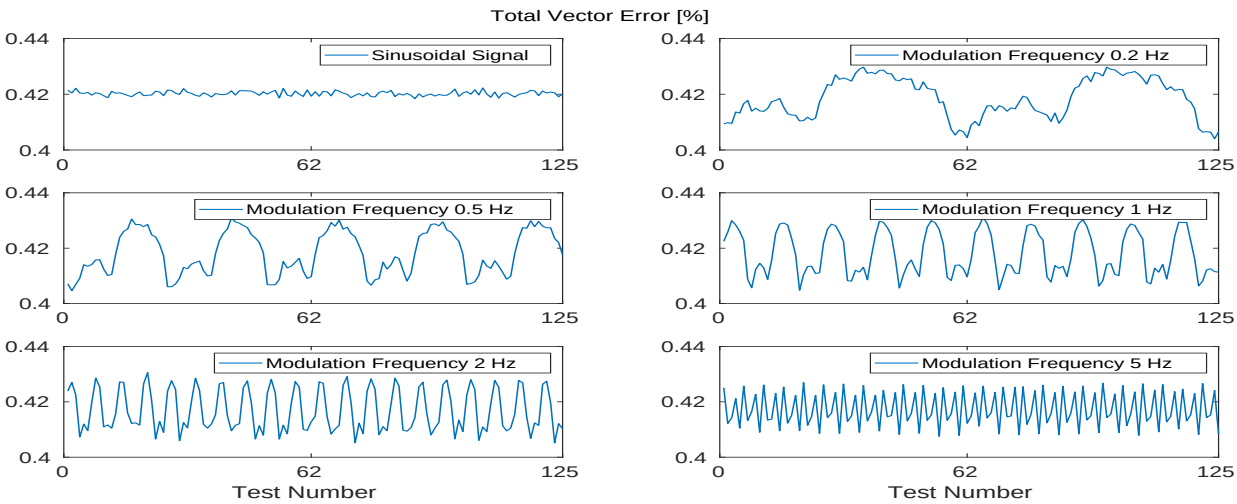


Fig. 5. Total Vector Error, in sinusoidal conditions and with amplitude modulation

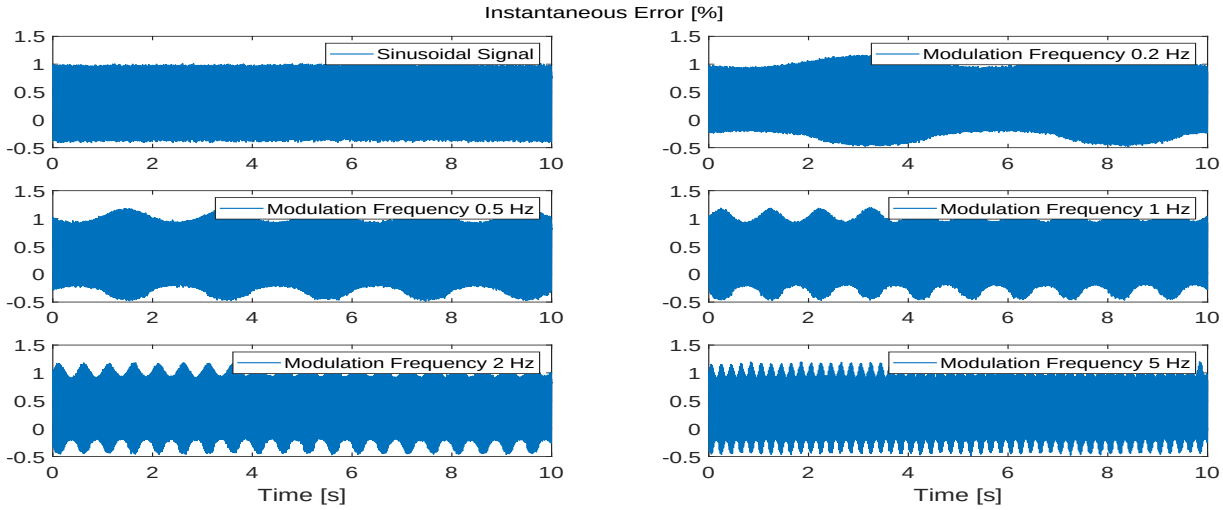


Fig. 6. Instantaneous Error, in sinusoidal conditions and with amplitude modulation

III. MEASUREMENT SETUP

The measurement setup is shown in Fig. 2. Current generation is obtained through the transconductance amplifier Fluke 52120 A (up to 120 A, up to 10 kHz). It is driven by the National Instrument (NI) PXI 5422 arbitrary waveform generator (AWG, $\pm$ 12 V, 200 MHz maximum sampling rate, 16 bit, onboard memory 256 MB).

The used data acquisition board (DAQ) is the NI PXI express (PXIe) 6124 (4 synchronous channels, 16 bit, 4 MHz maximum sampling rate). The DAQ and the AWG are housed in a NI PCI eXtension for Instrumentation (PXI) chassis which provides the same reference clocks to both of them: in such a way acquisition and generation are accurately synchronized. Current reference value is obtained by means of the current shunt FLUKE A40B 20/0.8 A/V (100 kHz, 30 μA/A). The CT under test is wound type and it has rated ratio of 10/5 A/A, rated burden of 5 VA and accuracy class of 0.5. Its output current is converted into a voltage by means of the current shunt LEM NORMA TRIAX Shunt 10/0.1 A/V (100 kHz, 0.03 %). The test bench is shown in Fig. 2. The software for data processing and instrument control was developed in Labview. A large choice of signal types can be generated, f.i. sinusoidal, fundamental plus a harmonic tone, fundamental with N harmonics, fundamental with an inter-harmonic, etc. Every signal generation is forerun and followed by a fade-in and a fade-out signal, in order to avoid generating large currents with a step variation. The samples at the output of the current shunts are acquired at a frequency of 10 kHz and stored into files. They are then processed in MatLab environment to extract the fundamental phasors and to evaluate ϵ, $\Delta\phi$ and TVE.

Table 1. Peak to peak values of the Ratio Error, Phase Error and TVE, in sinusoidal conditions and with amplitude modulation

Modulation Frequency $[Hz]$	Ratio Error [%]	Phase Error [mrad]	TVE [%]
Sine Wave	0.006	0.05	0.004
0.2	0.047	0.28	0.026
0.5	0.045	0.22	0.026
1	0.045	0.20	0.026
2	0.045	0.20	0.025
5	0.040	0.15	0.019

Table 2. Peak to peak values of the Instantaneous Error, in sinusoidal conditions and with amplitude modulation

Modulation Frequency $[Hz]$	Instantaneous Error [%]
Sine Wave	1.424
0.2	1.670
0.5	1.677
1	1.685
2	1.666
5	1.688

IV. EXPERIMENTAL RESULTS

The CT has been tested with currents having a fixed fundamental component, amplitude of 10 A, frequency of 50 Hz and zero phase, and an amplitude modulating component having fixed phase (0 rad) and amplitude (10 % of fundamental component) and variable frequency, from 0.2 Hz to 5 Hz. For each modulating frequency, 125 measurements have been performed. Fig. 3, Fig. 4 and Fig. 5 show, respectively, ϵ, $\Delta\phi$ and TVE in all the tests performed. Each figure shows six curves, referring, respectively, to sinusoidal conditions, amplitude modulation with frequency of 0.2 Hz, 0.5 Hz, 1 Hz, 2 Hz and 5 Hz. With respect to the sinusoidal case, the tests with amplitude modulation imply oscillations of the evaluated indexes, which probably, depend on the relative phase anlge between the fundamental and the modulating components within the observation window. Moreover, for ϵ and $\Delta\phi$, the averages of the oscillating values differ from the mean value observed for sinusoidal conditions, whereas for TVE the averages are more similar to the sinusoidal case. In all the cases, the peak to peak values of the oscillations decrease when the frequency increases. The peak-to-peak values of the evaluated indexes are reported in Table 1 and their maximum values are 0.047 % for ϵ, 0.28 mrad for $\Delta\phi$ and 0.026 % for TVE. It is worth to note that the oscillations, in presence of modulation, is about one order of magnitude higher than in sinusoidal case. The instantaneous error has also been evaluated and it is shown in Fig. 6: it is expressed in percentage of the fundamental amplitude. Table 2 shows the peak-to-peak values of the instantaneous error: it can be seen that, in presence of modulation, the istantaneous error is higher than in sinusoidal case of about 0.25 %.

V. CONCLUSION

This paper analyzes the uncertainty contribution of inductive current transformers to the synchrophasor measurement performed by a PMU, in the presence of amplitude modulated currents. A high-performance measurement configuration capable of generating up to 120 A was presented. Experimental tests, performed on a commercial CT, changing the frequency of the amplitude modulating component, showed that an additional error is introduced by the CT on the fundamental phasor, with respect to the sinusoidal case. However, other research activities, about the testing of different types of current transformers and with different types of input disturbances, are still in progress. The results of these activities, which are inserted in the framework of the research project EMPIR Future Grid II, will be also supplied to working group 47 "Evolution of Instrument transformer requirements for the modern market" of the IEC (International Electrotechnical Commission) Technical Committee 38 "Instrument Transformers". In fact, some of the authors participate to the standardization activity of this working group, which actually is dealing with the preparation of a document on the characterization of instrument transformers to be employed for power quality and synchrophasor measurements.

VI. ACKNOWLEDGEMENT

The work presented in this paper was funded by EMPIR, 17IND06 Future Grid II project, which is jointly funded by the EMPIR participating countries within EURAMET and the European Union.

REFERENCES

[1] **E. Mohns**, **G. Roeissle**, **S. Fricke**, and **F. Pauling**. An ac current transformer standard measuring system for power frequencies. *IEEE Transactions on Instrumentation and Measurement*, 66(6):1433–1440, June 2017.

[2] **G. Crotti**, **A. Delle Femine**, **D. Gallo**, **D. Giordano**, **C. Landi**, **P.S. Letizia**, and **M. Luiso**. Calibration of current transformers in distorted conditions. *Journal of Physics: Conference Series*, 1065:052033, aug 2018.

[3] **B. Kerr**, **L. Peretto**, **N. Uzelac**, and **E. Scala**. Integration challenges of high-accuracy lpit into mv recloser. *CIRED - Open Access Proceedings Journal*, 2017(1):260–263, 2017.

[4] **E. Houtzager**, **E. Mohns**, **S. Fricke**, **B. Ayhan**, **T. Kefeli**, and **H. Cayci**. Calibration systems for analogue non-conventional voltage and current transducers. pages 1–2, July 2016.

[5] **G. Crotti**, **D. Giordano**, **P. Roccato**, **A. Delle Femine**, **D. Gallo**, **C. Landi**, **M. Luiso**, and **A. Mariscotti**. Pantograph-to-ohl arc: Conducted effects in dc railway supply system. *9th IEEE International Workshop on Applied Measurements for Power Systems, AMPS 2018 - Proceedings*, 2018.

[6] **G. Crotti**, **A.D. Femine**, **D. Gallo**, **D. Giordano**, **C. Landi**, and **M. Luiso**. Measurement of the absolute phase error of digitizers. *IEEE Transactions on Instrumentation and Measurement*, 68(6):1724–1731, 2019.

[7] **A. Delle Femine**, **D. Gallo**, **D. Giordano**, **C. Landi**, **M. Luiso**, and **D. Signorino**. Synchronized measurement system for railway application. *Journal of Physics: Conference Series*, 1065:052040, aug 2018.

[8] **G. Crotti**, **D. Giordano**, **A. Delle Ferninc**, **D. Gallo**, **C. Landi**, and **M. Luiso**. A testbed for static and dynamic characterization of dc voltage and current transducers. *9th IEEE International Workshop on Applied Measurements for Power Systems, AMPS 2018 - Proceedings*, 2018.

[9] **A. Delle Femine**, **D. Gallo**, **D. Giordano**, **C. Landi**, **M. Luiso**, and **D. Signorino**. Synchronized mea-

surement system for railway application. *Journal of Physics: Conference Series*, 1065(5), 2018.

[10] **M. Faifer**, **C. Laurano**, **R. Ottoboni**, **S. Toscani**, and **M. Zanoni**. Harmonic distortion compensation in voltage transformers for improved power quality measurements. *IEEE Transactions on Instrumentation and Measurement*, pages 1–8, 2019.

[11] **A. Cataliotti**, **V. Cosentino**, **G. Crotti**, **A. Delle Femine**, **D. Di Cara**, **D. Gallo**, **D. Giordano**, **C. Landi**, **M. Luiso**, **M. Modarres**, and **G. Tine'**. Compensation of nonlinearity of voltage and current instrument transformers. *IEEE Transactions on Instrumentation and Measurement*, 68(5):1322–1332, May 2019.

[12] **A. J. Collin**, **A. Delle Femine**, **D. Gallo**, **R. Langella**, and **M. Luiso**. Compensation of current transformers' non-linearities by means of frequency coupling matrices. In *2018 IEEE 9th International Workshop on Applied Measurements for Power Systems (AMPS)*, pages 1–6, Sep. 2018.

[13] **T. De Carvalho Batista**, **B.A. Luciano**, **F. Das Chagas Fernandes Guerra**, and **R.C.S. Freire**. Errors in current transformers submited to nonsinusoidal conditions. *18th IMEKO TC4 Symposium on Measurement of Electrical Quantities 2011, Part of Metrologia 2011*, pages 113–116, 2011.

[14] **K. Draxler**, **J. Hlavacek**, and **R. Styblikova**. Determination of instrument current transformer errors in frequency range up to 1 khz. *Journal of Physics: Conference Series*, 1065(5), 2018.

[15] Ieee standard for synchrophasor measurements for power systems. *IEEE Std C37.118.1-2011*, pages 1–61, Dec 2011.

[16] Ieee standard for synchrophasor measurements for power systems – amendment 1: Modification of selected performance requirements. *IEEE Std C37.118.1a-2014*, pages 1–25, April 2014.

[17] Instrument transformers - part 2: Additional requirements for current transformers. *IEC EN 61869-2*, Sep 2012.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Design Approach for a Stand Alone Merging Unit

G. Cipolletta, A. Delle Femine, D. Gallo, C. Landi, M. Luiso

Dept. of Engineering, University of Campania "Luigi Vanvitelli", Aversa (CE), Italy,
antonio.dellefemine@unicampania.it

***Abstract* – In order to enhance levels of security and reliability of power systems, allowing for advanced remote diagnostics, Merging Units (MUs) play a key role. Some of the benefits are a more efficient transmission of electricity and a better integration with renewable energy systems. In this paper an implementation of a Stand Alone Merging Unit (SAMU), compliant with the IEC 61850-9-2 standard and based on a low cost ARM microcontroller, is described. It acquires two signals, one voltage and one current, and it sends the samples over the Ethernet connection. A high-resolution analog to digital converter (ADC), synchronized to the Universal Time coordinated (UTC) through a Global Positioning System (GPS) disciplined oscillator, is used. The results of the characterization of the ADC are presented.**

Keywords* – *Power System Measurement, Power System Diagnostics, Stand Alone Merging Unit, Time Synchronization, Analog To Digital Converter.

I. INTRODUCTION

In recent years, within the electricity supply industry there has been a growing awareness of the need to reinvent Europe's electricity networks in order to meet the demand of 21st century customers. This justifies the growing interest in smart grids [1,2] and it increases, in turns, the interest in smart substation development. According to [3], Merging Units (MUs) play a crucial role in the design of substations communication and automation systems. Substation automation using the IEC 61850 suite of protocols is an established reality for the purpose of enabling fast-acting protection and control application.

The Digital Substations (DSs) are electrical substations where operations are managed between distributed Intelligent Electronic Devices (IEDs) at different levels of automation, interconnected by communication networks. For a reliable information exchange among devices at different levels, data accuracy is needed not only on amplitude but in time as well.

The Stand Alone Mering Unit (SAMU) [4] acts as digital interface of the instrument transformers. It converts to digital the analog input signals from instrument transformers (inductive or low power, i.e. LPIT) and provides a time-coherent combination of the current and voltage samples, which are used to create a data stream of Sampled Values (SVs). SVs are used for transmitting digitized instantaneous values of power system quantities, mainly primary voltages and currents of one or multiple phases, to protection relays (PRs) and bay controllers (BCs) through either Ethernet or optical communication channels based on the IEC 61850-9-2 protocol.

IEC 61850-9-2 handles this publication of SVs over the ethernet that are sent to microprocessor-based IED. The benefits of digital process buses include a reduced complexity of Current and Voltage Instrument Transformers (CT and VT) secondary cabling and a simplified connection for centralized substation automation functions such as the disturbances recording and the power quality monitoring.

The presented work is part of EMPIR 17IND06 Future Grid II research project: one objective of this project is to develop metrological grade Stand Alone Merging Units and traceable time synchronization techniques.

Special care has to be taken to have accurate time synchronization between units. Multiple devices will have their local independent clock which can not be used as time reference for the published data. Otherwise, samples coming from different devices are not suitable to reconstruct the power network state [5, 6].

Therefore SVs need to be accurately time stamped, providing synchronization information to put in relation data from different IED placed in different locations of the smart grid. For this reason Local clocks must be synchronized to each other via a universal source. The standard [7] states that, to ensure deterministic operation of critical functions in the automation system, a precise time distribution and clock synchronization in electrical grids with an accuracy of 1 μs must be used.

The Precision Time Protocol (PTP)[8], that is a network-based time synchronization standard that can achieve clock accuracy in the sub-microsecond range, is suggested as time distribution mechanism. Such a synchronization method implies considerable efforts by microcontroller firmware to generate an ADC clock phase locked to the absolute time. In fact, continuous calibration of the internal microcontroller clock must be performed, with a proper closed loop digital control system that follow the reference imposed by PTP input.

To avoid this computational overhead and to reach higher synchronization accuracy, the proposed SAMU, instead, makes use of a GPSDO (GPS Disciplined Oscillator), that allows a long term stability and does not need

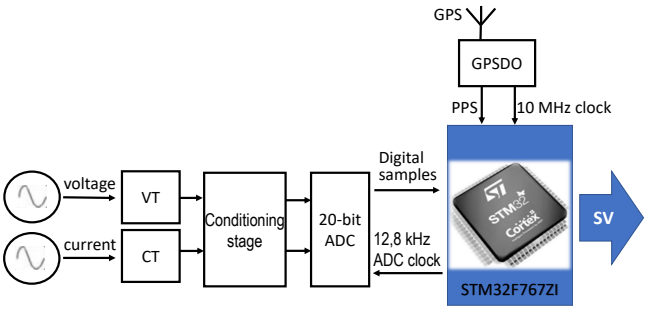


Fig. 1. Block diagram of the proposed SAMU

firmware calibration; the long term stability relies on the high accuracy of GPS Caesium references. The adopted solution allows to get synchronization on each sample with higher time resolution.

In the following, Section II describes the hardware implementation of the proposed SAMU, whereas Section III focuses on the embedded measurement firmware. Section IV shows experimental results of the ADC characterization. Finally, Section V draws the conclusions.

II. HARDWARE ARCHITECTURE

The block diagram of the proposed SAMU is shown in Fig. 1. The Connor-Winfield FTS125-CTV-010.0M was adopted as GPSDO. It produces a 1-PPS (1 Pulse Per Second) signal from the GPS timing receiver and generates both a 10 MHz CMOS levels square wave and a 10 MHz sine wave output from a low jitter VCTCXO (Voltage Controlled Temperature Compensated Crystal Oscillator). It accomplishes the task of keeping the system synchronized to absolute time: in fact it adjusts the frequency of the VCTCXO to be an integer multiple of the PPS and also the relative phase, that is the VCTCXO is adjusted to have 10 million oscillations in the PPS period, with the first pulse having the rising edge coincident with the PPS pulse.

The 10 MHz square wave is used as reference clock for the STM32F767ZI, ARM-Cortex M7 microcontroller with 2 MB Flash, 512 kB of SRAM, clock frequency of 216 MHz , L1 cache, art accelerator, Digital Signal Processor (DSP) with Floating-point Unit (FPU). The microcontroller is interfaced via Serial Peripheral Interfaces (SPI) with the external analog-to-digital converter (ADC) MAX11960, which is a differential Successive Approximation Register (SAR) ADC with resolution of 20 bit, maximum sampling frequency of 1 MHz, dual simultaneous sampling, Signal to Noise Ratio (SNR) of 99dB, Total Harmonic Distortion (THD) of -123 dB, differential non linearity of ± 1 LSB. The ADC has two independent SPIs with a shared clock, while the two SPIs of the microcontroller have two independent serial clock signals. Thus, in order to avoid electrical problems, one clock has been used for both the channels.

Data acquisition is performed in accordance to [3], using a sampling frequency of 12.8 kHz. Thus, the microcontroller supplies a 12.8 kHz sampling clock to the ADC, in order to have 256 samples for each rated period of the fundamental component (i.e. 50 Hz). The device has a differential $\pm 3\ V$ $(\pm\ V_{REF})$ input range. Thus, the signal conditioning stage produces a differential input signal centered around a common mode DC voltage of $\frac{V_{REF}}{2}$. The adopted conditioning stage has been presented in [9]. The two input stages for the respective transducers are followed by the optical insulation stages, that are identical for both channels and guarantee an insulation up to 2500 V and a bandwidth of 100 kHz. Also the output stages are identical for the two input signals. The conditioning ciruits reproduce signals for voltage and current channels that lay within the operating range of the ADC. Once that the signals are scaled down, accurately referred in time, sampled and converted to digital, microcontroller has to transmit the SVs. As regards the communication hardware interface, STM32F767ZI features a 10/100 Mbit/s EMAC (Ethernet Media Access Controller) peripheral that consists of a MAC 802.3 controller. It is in compliance with both Media Independent Interface (MII) and Reduced Media Independent Interface (RMII) to interface with the Physical Layer and supports ethernet frame time stamping as described in [8]. Furthermore it has a dedicated Direct Memory Access (DMA) controller that interfaces with the core and memories through the Advanced High-performance Bus (AHB) Master and Slave interfaces. The AHB Master Interface controls data transfers while the AHB Slave interface accesses Control and Status Registers (CSR) space. The Transmit First-In-Fisrt-Out (FIFO) buffers the data read from the system memory by the DMA, before the transmission by the MAC (media access control). Similarly, the Receive FIFO stores the Ethernet frames received until they are transferred to the system memory by the DMA.

III. EMBEDDED MEASUREMENT FIRMWARE

The microcontroller has been programmed using Standard C Language. The firmware implements the opportune drivers for all the described hardware devices and performs their coordination. Starting from the disciplined 10 MHz clock, the 12.8 kHz sampling has to be produced for the ADC. Various solutions are possible to solve this problem; most of them, however, involves firmware latency. The only one that completely avoids this issue is the use of the 10 MHz clock as the main oscillator for the entire microcontroller. In this way, in fact, all the internal clocks of the microcontroller are locked in frequency and phase with the absolute time and it is possible to construct the time base for the ADC using only internal hardware devices, without involving execution time latency due to firmware routines. In particular the time base for the ADC has been built up by making use of an internal timer as key peripheral. Sum-

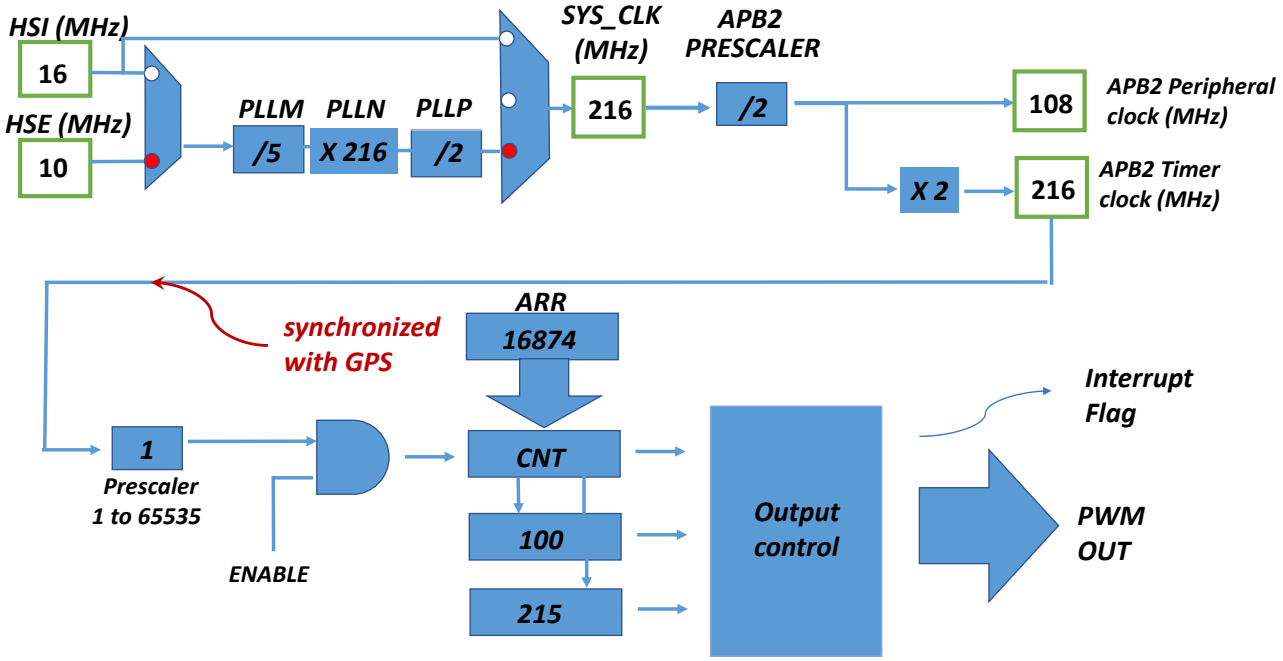


Fig. 2. Clock configuration.

marizing, the 10 MHz clock reference signal from GPSDO has been used as external source and given to the internal microcontroller PLL (Phase Locked Loop) circuit which allows to increase the frequency up to 216 MHz. The timer peripheral acts as divider, automatically controlling external PWM (Pulse Width Modulation) line with 50 % duty cycle. This line is used to clock the ADC. To maintain a high level of accuracy, Interrupt Service Routines (ISR) have been avoided and the timer is configured to control the output in hardware, without latency. Fig. 2 shows the block diagram of sample clock generation. System clock (SYS_CLK) is obtained by the equation (1) and is equal to 216 MHz.

$$SYS\_CLK = \frac{HSE \cdot PLLM \cdot PLLN}{PLLP} \qquad (1)$$

The timer accurately divides by 16875 the system clock, to generate a 12.8 kHz square wave using the PWM output. The standard PWM mode is programmed with the Auto-Reload Register (ARR) and the Capture Compare Register (CCR) to define period and duty cycle respectively. When the counter matches the content of CRR1, the output is turned off. When counter matches ARR, the counter is set to 0, the output is turned on and the counter starts counting up again. As mentioned in section 1, the aims of the presented work is to realize a reference SAMU. According to [10], the requested accuracy for the highest time performance class (i.e. T5) must be $\pm$ 1 μs. The use of GPSDO offers significant advantages in terms of synchronization accuracy.

Moreover, the microcontroller firmware receives and parses NMEA (National Marine Electronics Association) sentences from GPS receiver to obtain timestamp. Since the SAMU must associate a timestamp to each sample, a hard real-time mechanism has been implemented to identify the precise instant in which a 1-second-frame starts, using PPS signal external interrupt. Obviously this synchronization is needed only on the first PPS frame at startup. From this moment, the system proceeds counting up each sample time and following eventual corrections carried on by GPSDO. The firmware checks this correction process monitoring the difference between internal counters (locked on disciplined 10 MHz from DSO) and NMEA time strings (locked on PPS signal from GPS system). This difference must converge in the interval $\pm$200ns within a fixed timeout.

The microcontroller is responsible for IEC61850-9-2 communication via Ethernet, too. To fulfill this task, integration between the Lightweight Internet Protocol stack (LwIP) and libiec61850[11] (an open source implementation of the standard) has been realized. This integration has been conducted through Berkley Software Distribution (BSD) socket interface in LwIP stack. IEC 61850-9-2 is based on RAW socket, not fully supported by LwIP; it has been necessary to extend that support, opportunely modifying the whole library code.

For the data-acquisition from external ADC, a producer-consumer design pattern has been chosen. The samples' reading is achieved through the event-based state machine shown in Fig. 3. The transition between the states is interrupt-based. This way, the microcontroller can perform

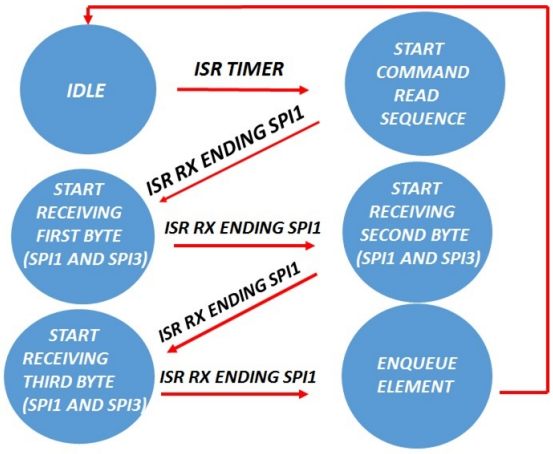


Fig. 3. ADC samples collection via SPI.

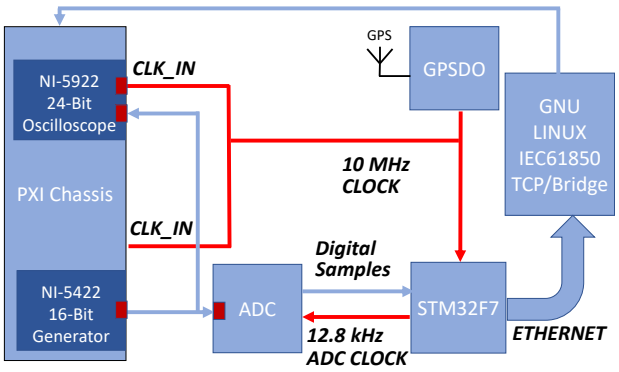


Fig. 4. Experimental test setup.

other tasks like sending the sample frames through the ethernet. Once the timer is enabled and started in PWM mode a 12800 Hz square wave synchronized with GPS is produced and used as conversion starter for both MAX11960 channels A and B in order to have a simultaneous sampling between voltage and current signals. The rollover timer ISR has been used to start the finite state machine; from this moment on, it evolves by interrupts from the Serial Peripheral Interface (SPI) used for data transfer from the ADC.

IV. EXPERIMENTAL TEST

The block diagram of the measurement setup for SAMU characterization is shown in Fig. 4 [12–14]. A characterization of the system has been conducted. In particular the Effective Number Of Bit (ENOB) of the ADC and the synchronization accuracy were evaluated.

For the evaluation of the ENOB, a NI PXI (National Instruments PCI eXtension for Instrumentation) chassis, housing the NI-5422 (16 Bit, $\pm$ 12 V, 200 MHz Arbitrary Waveform Generator, AWG) and the NI-5922 (24 Bit, $\pm$ 5 V, 500 kHz Data Acquisition Board, DAQ) have been used.

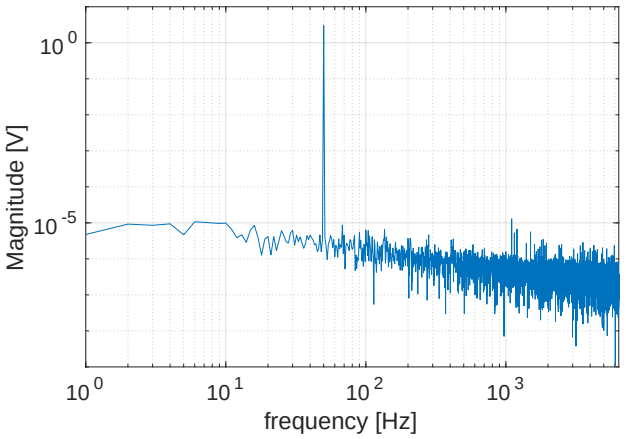


Fig. 5. Magnitude spectrum of compensated signal.

The PXI chassis has been supplied with a 10 MHz external clock source from the GPSDO. Both boards have been configured to use the PXI clock as reference clock in order to have coherent sampling between NI-5422, NI-5922 and proposed reference SAMU under test.

A sinusoidal signal, with amplitude equal to the ADC full scale and frequency of 50 Hz, has been used as test signal and directly supplied (excluding conditioning stages) to both the SAMU ADC and to the DAQ.

Using the DAQ as reference device, a closed loop compensation has been implemented, to enhance the test signal spectral purity. The sine wave obtained with this technique reaches a SINAD (Signal to Noise And Distortion Ratio) over 92 dB; its magnitude spectrum, obtained without the parallel connection of the SAMU ADC, is reported in Fig. 5). Unfortunately, when parallel connecting the SAMU ADC, the closed loop compensation does not give the same performance. In fact, as depicted in Fig. 6, which shows the magnitude spectrum of the signal present at the ADC input measured during the test, the spectral purity becomes lower as the SINAD drops to about $85dB$. Obviously, such spectral purity is not enough to characterize the rated ADC amplitude resolution of 20 bits. Further studies are still in progress to solve load effect issue and to enhance the spectral purity of the test signal.

The software for the measuring setup has been developed in LabVIEW under Windows OS (Operating System). Unfortunately, Windows reveals to have poor support for RAW socket. For this reason, the measurement results, i.e. the sampled values of the SAMU, are sent via ethernet to a IEC 61850-9-2/TCP (Transmission Control Protocol) bridge (realized using [11]) and subsequently to the PXI controller.

In Fig. 7 the magnitude spectrum of the sampled values is reported. The presence of the same harmonics of the input test signal can be seen. Moreover it is important to

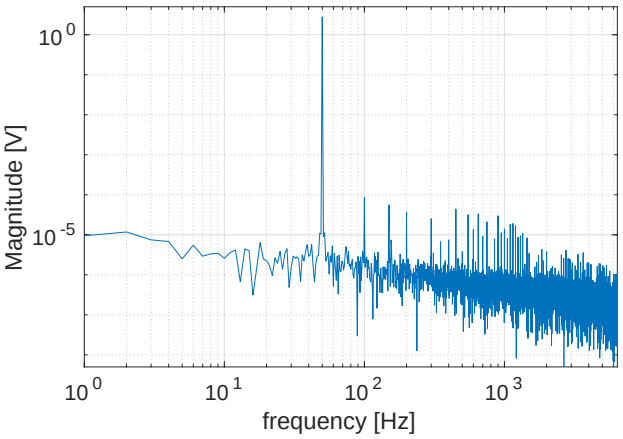


Fig. 6. Magnitude spectrum of test signal with load effect.

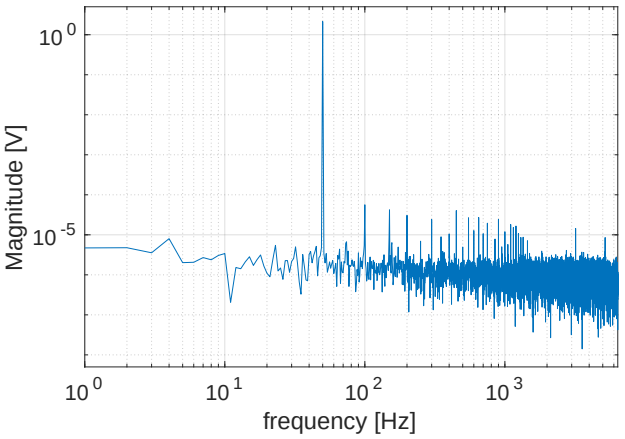


Fig. 7. Magnitude spectrum of ADC samples.

highlight that there is an undesired spectral component at exactly 3200 Hz, not present in the input signal. For this reason the SINAD at the ADC output never exceeds 84 dB.

This spurious spectral component is uncorrelated with the input signal; to prove that consideration, another test has been conducted putting zero volt at the input of the ADC (with a short circuit). As it can be seen from Fig. 8, the undesired tone is still present, located at the frequency $f_d = 3.2\ kHz$. Investigations about this issue (changing the sampling frequency and/or changing the clock signal amplitude and frequency), led to realize that the disturbance is due to capacitive coupling with the 10 MHz clock signal. This last frequency is aliased according to equation (2).

$$f_d = |f_{clock} \pm k \cdot f_s| \tag{2}$$

where f_{clock} is the clock signal at 10 MHz and f_s is the the sampling frequency of 12.8 kHz; choosing $k = 781$ the f_d results exactly equal to 3.2 kHz.

Despite its small amplitude, the spurious component is considerably higher than the noise floor. Further efforts, to enhance the performance of anti-aliasing filter at the input of ADC, are in progress.

In order to evaluate the time synchronization accuracy of the proposed SAMU, other experimental tests have been conducted. An Agilent 53230A Universal Frequency Counter has been used to determine the conversion starter stability. The experimental results show that a standard deviation of 10 μHz on the sampling frequency has been obtained. Furthermore the maximum error on sampling period, measured on 10 minute observation time, is about 200 ns [10]. Further experimental tests to evaluate the synchronization error with respect to the absolute time are still in progress.

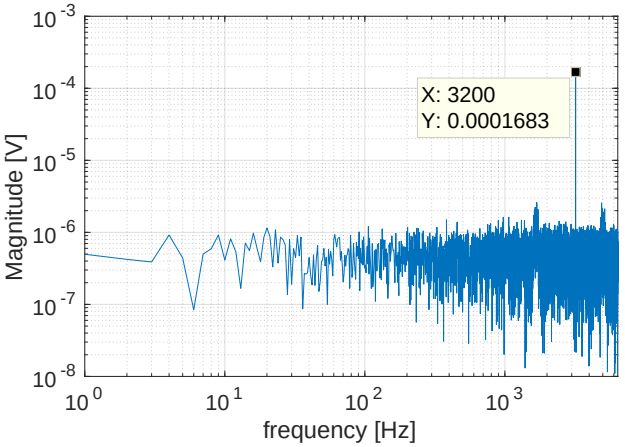


Fig. 8. Magnitude spectrum of ADC samples with a short circuit at the input.

V. CONCLUSION

This paper discusses the implementation of a reference Stand Alone Merging Unit, based on an ARM microcontroller. It also includes an external high resolution ADC, a conditioning stage, to adapt voltage and current signals coming from VTs and CTs to ADC input range, and a GPSDO. Experimental results to evaulate the ENOB and the stability of the timebase have been presented.

Further experimental tests of the realized SAMU are still in progress, in order to obtain a complete metrological characterization of the proposed reference instrument. Moreover, various solutions are being evaluated to reduce or possibly eliminate spurious spectral component due to capacitive coupling.

VI. ACKNOWLEDGEMENT

The work presented in this paper was funded by EMPIR, 17IND06 Future Grid II project, which is jointly funded by

the EMPIR participating countries within EURAMET and the European Union.

REFERENCES

[1] **G. Crotti**, **D. Giordano**, **A. Delle Ferninc**, **D. Gallo**, **C. Landi**, and **M. Luiso**. A testbed for static and dynamic characterization of dc voltage and current transducers. *9th IEEE International Workshop on Applied Measurements for Power Systems, AMPS 2018 - Proceedings*, 2018.

[2] **G. Crotti**, **D. Giordano**, **P. Roccato**, **A. Delle Femine**, **D. Gallo**, **C. Landi**, **M. Luiso**, and **A. Mariscotti**. Pantograph-to-ohl arc: Conducted effects in dc railway supply system. *9th IEEE International Workshop on Applied Measurements for Power Systems, AMPS 2018 - Proceedings*, 2018.

[3] Communication networks and systems for power utility automation - part 9-2: Specific communication service mapping (scsm) - sampled values over iso/iec 8802-3. *IEC EN 61850-9-2:2011*, Sep 2011.

[4] **M. Agustoni** and **A. Mortara**. A calibration setup for iec 61850-9-2 devices. *IEEE Transactions on Instrumentation and Measurement*, 66(6):1124–1130, June 2017.

[5] **A. Delle Femine**, **D. Gallo**, **D. Giordano**, **C. Landi**, **M. Luiso**, and **D. Signorino**. Synchronized measurement system for railway application. *Journal of Physics: Conference Series*, 1065(5), 2018.

[6] **G. Crotti**, **A. Delle Femine**, **D. Gallo**, **D. Giordano**, **C. Landi**, and **M. Luiso**. Measurement of the absolute phase error of digitizers. *IEEE Transactions on Instrumentation and Measurement*, 68(6):1724–1731, 2019.

[7] Communication networks and systems for power utility automation - part 9-3: Precision time protocol profile for power utility automation. *IEC EN 61850-9-3:2016*, May 2016.

[8] Ieee standard profile for use of ieee 1588 precision time protocol in power system applications. *IEEE Std C37.238-2017 (Revision of IEEE Std C37.238-2011)*, pages 1–42, June 2017.

[9] **S. Del Prete**, **A. Delle Femine**, **D. Gallo**, **C. Landi**, and **M. Luiso**. Implementation of a distributed stand alone merging unit. *Journal of Physics: Conference Series*, 1065:052042, aug 2018.

[10] Communication networks and systems in substations - part 5: Communication requirements for functions and devices models. *IEC EN 61850-5:2004*, Nov 2014.

[11] Official repository for libiec61850, the open-source library for the iec 61850 protocols, availabble online on 01.06.2018 at http://libiec61850.com/libiec61850.

[12] **A. J. Collin**, **A. Delle Femine**, **D. Gallo**, **R. Langella**, and **M. Luiso**. Compensation of current transformers' non-linearities by means of frequency coupling matrices. In *2018 IEEE 9th International Workshop on Applied Measurements for Power Systems (AMPS)*, pages 1–6, Sep. 2018.

[13] **A. Cataliotti**, **V. Cosentino**, **G. Crotti**, **A. Delle Femine**, **D. Di Cara**, **D. Gallo**, **D. Giordano**, **C. Landi**, **M. Luiso**, **M. Modarres**, and **G. Tine'**. Compensation of nonlinearity of voltage and current instrument transformers. *IEEE Transactions on Instrumentation and Measurement*, 68(5):1322–1332, May 2019.

[14] **G. Crotti**, **A. Delle Femine**, **D. Gallo**, **D. Giordano**, **C. Landi**, **P.S. Letizia**, and **M. Luiso**. Calibration of current transformers in distorted conditions. *Journal of Physics: Conference Series*, 1065:052033, aug 2018.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Multiple Heat Source Estimation by using Backward Simulator

Yukio Hiranaka

Yamagata University, Yonezawa, Japan, zioi@yz.yamagata-u.ac.jp

***Abstract* –The exact inverse solutions for real measurements are becoming realistic with the progress of computational performances. As an example, heat generation and release transients can be estimated from the temperature measurement if the thermal model is known. Although the inverse solver needs the correct knowledge of the target model, the method of backward simulation has a capability of determining the model structure and the model parameters by itself. A simple case of one heat source model has been shown as a successful example. In this paper, we describe an application of the method to a two heat source case, and show its result for a simulated temperature data.**

Keywords – Heat Source, Heat Transfer, Inverse Problem, Backward Simulation.

I. INTRODUCTION

For machine diagnostics, it may be effective to know the thermal structure and to identify heat sources inside the machine. It is common practice to estimate heat generation amount, heat source position and heat transfer characteristics from temperature measurements [1,2], including deconvolution and super resolution techniques [3,4]. However, such estimation is difficult if the temperature measuring point is far from the heat source or the number of measuring points is insufficient. Inverse heat transfer estimation is usually an ill-conditioned problem [5,6].

However, inverse calculation methods are becoming practical with the progress of computer performances [5,6]. If the thermal model is known, the heat input can be precisely estimated by inverse solvers with applying the measured temperature data [2].

Also, artificial neural network (ANN) method performing optimum estimation using a large amount of measurement data has become practical [7-14]. Even in the field of heat transfer problem, there exist a method to estimate the amount of heat generation from temperature change [12], a method to estimate the amount of heat generation from temperature distribution [13], a method to estimate heat source position from temperature distribution [14]. Mostly, those ANN methods estimates a few number of values such as heat generation in Watt by assuming the thermal structure.

The authors have been conducting researches by using the backward simulation technique [15-17] to estimate the time change of heat generation from the temperature change at one point in the experiment room. The method supposes the heat conduction model as a simple model and uses one shot measurement data without requiring a large amount of data for learning like ANN.

The backward simulation, one of inverse solvers, estimates heat generation in time sequence based on the energy conservation law and the proportionality of heat transfer to the amount of heat [15-17]. It was shown that the heat transfer coefficients between energy storage elements can also be determined by backward simulation, so that the heat generation can be estimated even if the model structure is not exactly known [17]. Of course, the certainty and the adaptability of the thermal model are important. As the heat phenomenon can be simplified by heat capacities with proportional heat transfers, they can be basically satisfied.

In practical applications, we have to deal the cases of heat generation at one point and heat leakage at another point or surroundings. In such cases, we have to install multiple temperature sensors to estimate both of the time changes of heat sources, positive or negative. This paper describes a method and results for such estimation by using transient temperature changes at two measuring points.

II. BACKWARD SIMULATION

Fig.1 shows a room which has a heat source (the human in the figure), local heat accumulation by air mass, and heat leak through the wall. The temperature changes (y_2, y_3) are measured in two places in order to calculate the two heat sources inside and outside. To match the situation, our model consists of four heat storage nodes corresponding to the air mass surrounding the heat source (u_1), its vicinity's air mass u_2 (the first temperature measuring point), the whole room air mass u_3 (the second temperature measuring point), the wall u_4, and heat sources of the inside generator x_1 and the heat source x_4 which corresponds to the outflowing negative heat source.

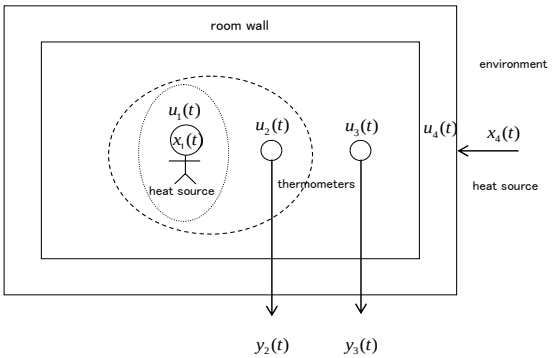


Fig. 1. Inverse problem of heat transfer estimation.

Fig. 2 shows a simplified model of heat storages of heat capacities C_1, C_2, C_3 and C_4. The transfer coefficients between them are k_{12}, k_{21}, k_{23}, k_{32}, k_{34} and k_{43}. Assuming that the heat transfer between the heat storages is proportional to the amount of heat, the following equations in relation to time t and t + 1 holds,

$$u_1(t+1) = (1-k_{12})u_1(t) + x_1(t) - k_{21}u_2(t), \quad (1)$$

$$u_2(t+1) = (1-k_{21}-k_{23})u_2(t) + k_{12}u_1(t) + k_{32}u_3(t), \quad (2)$$

$$u_3(t+1) = (1-k_{32}-k_{34})u_3(t) + k_{23}u_2(t) + k_{43}u_4(t), \quad (3)$$

$$u_4(t+1) = (1-k_{43})u_4(t) + x_4(t) + k_{34}u_3(t), \quad (4)$$

$$y_2(t) = u_2(t) / C_2, \quad (5)$$

$$y_3(t) = u_3(t) / C_3. \quad (6)$$

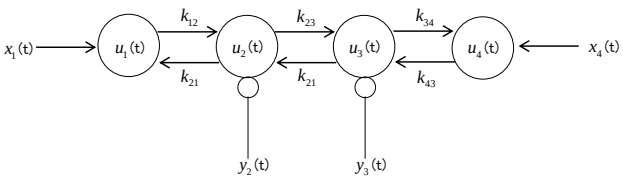


Fig.2. Four heat storage model.

Our objective is to obtain $x_1(t)$ and $x_4(t)$ by calculating $u_1(t)$, $u_2(t)$, $u_3(t)$ and $u_4(t)$ from the measured temperatures $y_2(t)$ and $y_3(t)$. The heat amounts $u_2(t)$ and u_3 (t)are directly known from $y_2(t)$ and $y_3(t)$. The remaining values of $u_1(t)$, $u_4(t)$, $x_1(t)$, $x_4(t)$ are calculated as follows,

$$u_1(t) = \frac{u_2(t+1) - (1-k_{21}-k_{23})u_2(t) - k_{32}u_3(t)}{k_{12}}, \quad (7)$$

$$u_4(t) = \frac{u_3(t+1) - (1-k_{32}-k_{34})u_3(t) - k_{23}u_2(t)}{k_{43}}, \quad (8)$$

$$x_1(t) = u_1(t+1) - (1-k_{12})u_1(t) - k_{21}u_2(t), \quad (9)$$

$$x_4(t) = u_4(t+1) - (1-k_{43})u_4(t) - k_{34}u_3(t). \quad (10)$$

It is a backward calculation in which the values in the previous time are obtained from the data of the later time. In the cases that we do not know the exact values of heat transfer coefficients, we have to determine them by searching the condition conforming physical laws. As the conformity used for the judgment, we used the fact that any heat values must be nonnegative. If the heat amount of each storage $u_1(t)$, $u_2(t)$, $u_3(t)$ and $u_4(t)$ or heat generation $x_1(t)$ becomes negative through the calculation, it indicates that the simulated combination of parameters is not feasible.

Heat flow between nodes in the backward calculation can be easily determined because it can be performed by one-way calculation as shown in Fig. 3 unlike the case of one heat source described in [8]. The broken line in the figure represents the link transmitting heat contrary to the flow of the actual heat. The solid line represents the link where the reverse flow can be processed with the forward flow in the equation (7)-(10). Although there are six heat transfer coefficients, they have the following relationship with the ratios of heat capacities. Then, the three transfer coefficients can be determined by other parameters,

$$\begin{aligned} k_{21} &= k_{12}C_1 / C_2, \\ k_{32} &= k_{23}C_2 / C_3, \\ k_{43} &= k_{34}C_3 / C_4. \end{aligned} \quad (11)$$

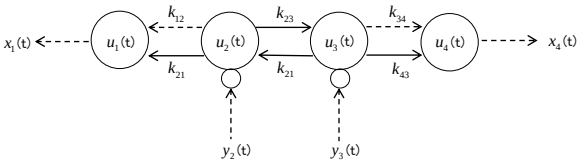


Fig. 3. Backward heat simulation model.

III. SIMULATOR IMPLLEMENTATION

Our simulator was implemented by using Node-RED [18] (Fig. 4). The coefficients of the backward calculation formula, the heat amount in each storage, the heat transfer, and the temperature values are all processed as range signals [15]. There is no concern of the simulation case leakage which is unavoidable by discrete simulations. The ranges of the signals are determined by dividing all possible range by the parameter ndiv. For example, in the case that k_{12} can be any value between 0 to 1 and if we set ndiv as 100, the range for k_{12} is tried from 0 to 0.01, 0.01 to 0.02, and so on.

Since increasing the ndiv takes calculation time proportional to $ndiv^2$ for the simulation, we adopted a reducing method which doubles ndiv stepwise. The range that is shown not feasible in the previous step is not calculated in the subsequent steps by using an array of conformance results. To allow some errors in calculation

for testing non-negative condition, the error allowance is introduced.

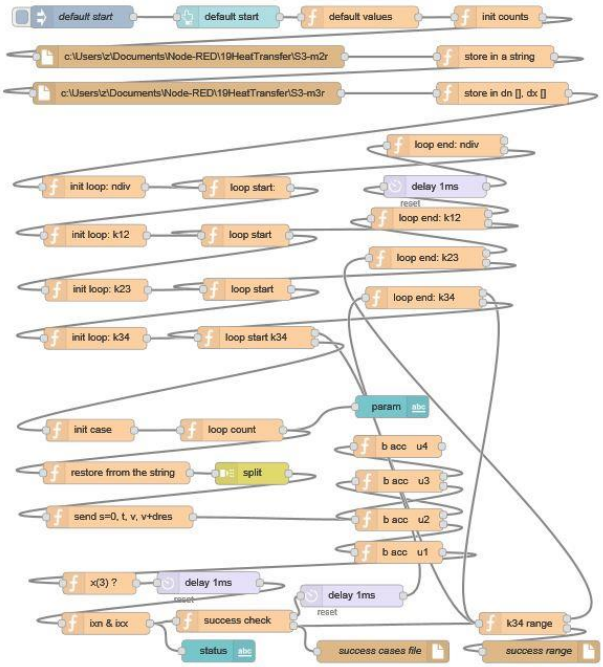


Fig.4. Backward simulation program using node-red.

IV. TEST DATA AND ITS RESULT

The time sequence of x_1 (1 only at t = 6-9, otherwise 0) and x_4 (-0.01 only at t = 21-60, otherwise 0) is shown in Fig. 5. The temperature y_2 and y_3 were determined by equations (1)-(6). The result obtained by setting $k_{12} = 0.3$, $k_{23} = 0.1$, $k_{34} = 0.05$, $C_1 = 1$, $C_2 = 4$, $C_3 = 8$ and $C_4 = 40$ is shown in Fig.6.

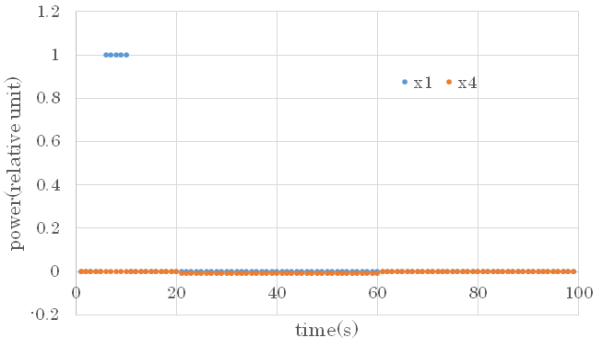


Fig.5. Simulated heat input(x1 and x_2).

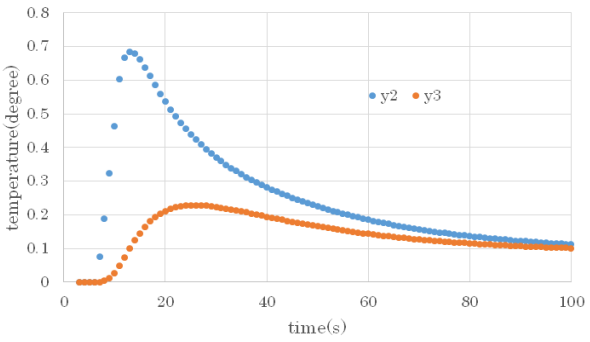


Fig.6. Simulated temperature change (y_2 and y_3).

To confirm that the reverse calculation can be correctly performed, the backward calculation was performed for the correct set of k_{12}, k_{23}, k_{34}, C_1, C_2, C_3 and C_4, and shown in Fg7 and Fg8. The result of x_4 is perfect, while the result of x_1 has some distortion. The larger the value such as k_{12}, the larger the distortion.

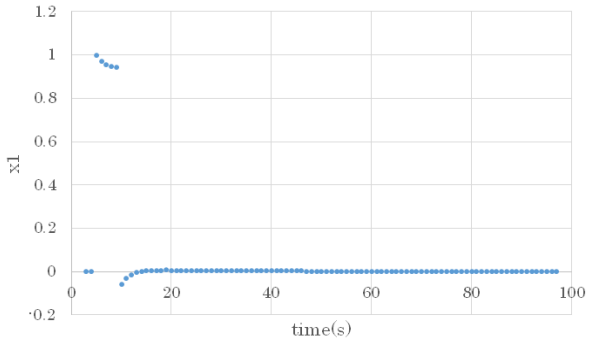


Fig.7. Estimated heat input x_1.

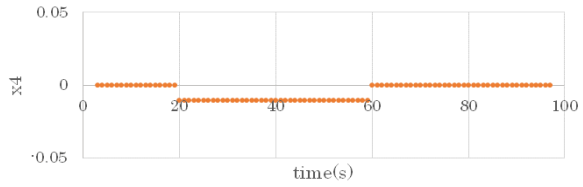


Fig.8. Estimated heat input x_4.

In our simulations, parameters such as k_{12} are searched assuming that C_1, C_2, C_3 and C_4 are known. The entire range is divided into ndiv divisions for k_{12}, k_{23} and k_{34}, and it is mapped whether the conforming data is obtained or not for each combination of parameters. The result maps are shown in Fig. 9 (k_{12} vs. k_{23}) and Fig. 10 (k_{23} vs. k_{34}). The central upper left edge in Fig. 9 indicates the correct value of k_{12}, and the lower right corner in Fig. 10 indicates the correct values of k_{23} and k_{34}.

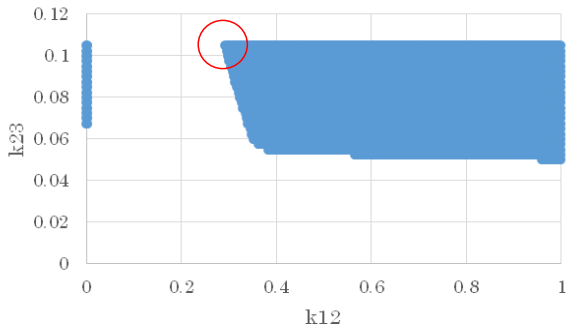


Fig.9. Feasible parameter map for ndiv=400.

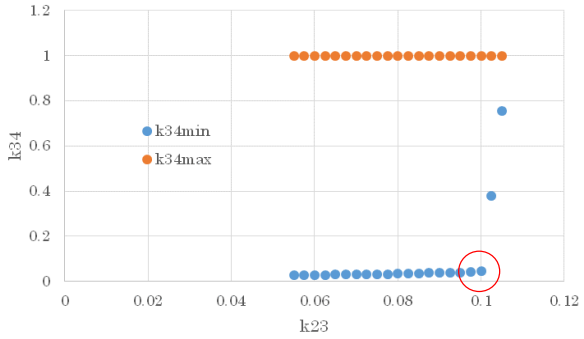


Fig.10. Feasible parameter map for ndiv=400 and k_{12}=0.5.

The reason why k_{12}, k_{23} and k_{34} can be determined from Fig. 8 and Fig. 9 is as follows. At first, the descending part after the peak of y_3 curve in Fig. 6 does not fall more rapidly than that. For example, if x_1 has additional heat input at a time later than the simulation setting, the temperature at each point in the transfer route will rise, but the fall after that will be limited by the outflow parameter k_{34}. On the other hand, even if the descending rate is high, the descent can be changed arbitrarily slow by adding the x_1 input during the descent. Therefore, the descending rate appearing in the graph at times when there is no x_1 input will mainly represent the minimum value of k_{34}. Similarly, the descent part after the peak of y_2 curve in Fig. 6 represents the minimum value of k_{23}.

If k_{34} is tried for all values in the possible range between 0.0 and 1.0, allowing any possible time change of x_1, the feasible value of k_{34} fitted to the curve must be equal to or more than the value corresponding to the descending rate represented by the y_3 curve. Of course, we have to run simulations with sufficient accuracy. However, the values of k_{23} and k_{43} also affect the descent of y_3. Since k_{43} is determined by k_{34} as in the equation (11), only k_{23} must be considered. If k_{23} increases, which means inflow increase to u_3, the outflow parameter k_{34} must be increased to match the increase to keep the descent rate of y_3 unchanged. Therefore, if a value that fits the temperature measurement curve in two dimensions of parameters k_{23} and k_{34} is obtained as shown in Fig. 10, k_{34} will be close to the correct value (0.05) when k_{23} is small. When k_{23} is lager then some value, feasible k_{23} will go up and the corner of the feasible range appear in the lower right of the figure. This corner corresponds the point where the range of feasible k_{34} is unduly reduced by k_{23} value, which means k_{23} beyond the corner is not valid.

On the other hand, the rising of the graph y_2 in Fig. 6 mainly depends on k_{12}, which determines the inflow to u_2. The rise can be made steeper by arbitrarily setting of x_1 input, but the slowest rise with no x_1 input is determined by k_{12}. For the same reason as k_{34}, the minimum value of the curve matching k_{12} will represent the actual value of k_{12}. Since the rise of y_2 also becomes slow as the outflow coefficient k_{23} becomes large, the range of k_{12} and k_{23} that fit in the two dimensional map is determined as shown in Fig. 9. The corner at the upper center represents the minimum value of k_{12} in the feasible range of k_{23}. Then, the actual value of k_{12} can be determined as 0.3 (correct value). Also at the corner, k_{23} is 0.1 and we find that it is appropriate to estimate k_{34} at the lower right corner of Fig.10.

In the application to real measurements, the heat capacities C_1, C_2, C_3, and C_4 have to be determined by the simulation, it is necessary to comprehensively test the conformity for the capacities other than the transfer coefficients. It needs further reduction of calculation time to perform such simulations.

Although the simulation tool Node-RED is a convenient simulation tool, it takes long computation time because of interpreting JavaScript language. It took about 30 minutes for calculation of ndiv = 100 in the case of 100 measured data points. However, the processing time can be largely reduced by concentrating only in the boundary of the conforming area in the Fig. 9 and Fig. 10, more specifically concentrating only in the effective area for parameter determination.

V. REAL MEASUREMENT DATA AND ISSUES

Fig. 11 shows the results of transient temperature change caused by an infrared heater (800W) activated for three minutes and measured by two temperature sensors in a small room. The sensors were placed at two locations in the room (near and far from the heater). Because of some issues we are still in the way to estimate the model parameters k_{12}, k_{23} and k_{34}.

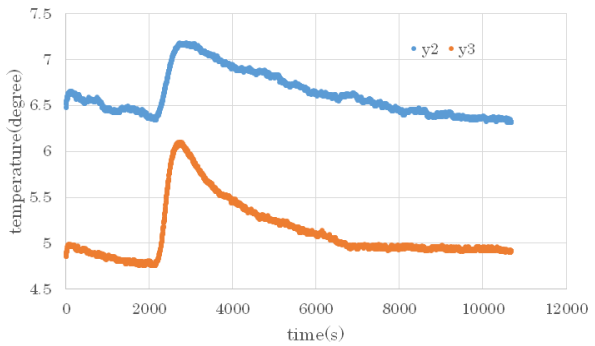


Fig.11 An example of a real temperature change.

There are noise and continuous temperature drift, which is clearly shown before the rising of temperature changes. It is expected that the drift is a contribution of negative heat source corresponding to the external environment and can be used to estimate the negative heat value.

The influence of noise should be suppressed to get a precise estimation. First, with regard to the quantization noise, we convert the measured values to value ranges, of which range width is half the quantization step. With regard to the noise above the quantization step, there is a method of correcting the value of each time point by checking the model match described in [16]. However the calculation time tend to be huge, we are considering to use optimal search method such as used in ANN processings.

Furthermore, the heat capacity ratios C_1/C_2, C_2/C_3, and C_3/C_4 are supposed to be known in the previous sections. However, for practical applications, it is necessary to treat them as unknown. Also, the range of these ratios may not be limited. Therefore, we should search the six parameters of k_{12}, k_{21}, k_{23}, k_{32}, k_{34} and k_{43} all in the range of 0 to 1, basically exhaustively. Since it is a six-dimensional grid search, the calculation time will be huge again. As a little relief, one of the six dimensions can be searched at only the end points of feasible ranges, so that the search is substantially five-dimensional. As an example in the case of k_{34}, Fig. 10 shows that our task is to find only the minimum and the maximum of feasible k_{34} values. To handle five-dimensional search area, we are doubling the parameter ndiv stepwise to get more precise results. In the course, if we find nonconformity to the model in some search area, we can omit the area in the further search. Currently, it takes several days for calculation in the case of ndiv 32 for 200 time point measurement data. We are looking for further time savings.

VI. CONCLUSIONS AND OUTLOOK

It was shown that the time-series estimation of internal conditions and external factors by backward simulation can be effectively done for multiple measured time series. If the number of model parameters to be determined increases, the processing time will be extended. However, we can shorten the processing time by using high performance algorithm and hardwares, and we expect to apply the method effectively to real measured data.

REFERENCES

[1] **J. V. Beck, B. Blackwell and C.R.S. Clair Jr**., Inverse Heat Conduction Ill-posed Problems, John Wiley & Sons, 1985, ISBN 0-471-08319-4.

[2] **Y. Jarny and D. Maillet**, Linear Inverse Heat Conduction Problem – Two Basic Examples, http://www.sft.asso.fr/Local/sft/dir/user-3775/documents/actes/Metti5_School/Lectures%26Tutorials-Texts/Text-L10-Jarny.pdf.

[3] **T.B.Bako and T.Daboczi**, Improved-Speed Parameter Tuning of Deconvolution Algorithm, IEEE Trans. Instrum. Meas., vol.65, no.1, pp.1568-1576, 2016.

[4] **A.Muqaibal, A,Safaai-Jazi, B.Woerner and S.Raid**, UWB Channel Impulse Response Characterization Using Deconvolution Techniques, Proc. MWSCAS, 2002.

[5] **Kenji Oguni**, Inverse problem and instrumentation, Ohmsha, Tokyo, 2011, ISBN 978-4-274-06829-4.

[6] **C.D. Twigg and D.L. James**, Back Steps in Rigid Body Simulation, ACM trans. Graph., vol. 27, no.3, article 25, 2008.

[7] **Ibrahim Mohamed Elshafiey**, Neural network approach for solving inverse problems, https://lib.dr.iastate.edu/rtd/16966, 1991.

[8] **Housen Li, Johannes Schwab, Stephan Antholzer, Markus Haltmeier**, NETT: Solving Inverse Problems with Deep Neural Networks, arXiv preprint arXiv:1803.00092, 2018.

[9] **Liqiang Zhang, Luoxing Li and HuiJu BiwuZhu**, Inverse identification of interfacial heat transfer coefficient between the casting and metal mold using neural network, Energy Conversion and Management, vol. 51, Issue 10, pp.1898-1904, 2010.

[10] **S. Deng and Y. Hwang**, Applying neural networks to the solution of forward, and inverse heat conduction problems, International Journal of Heat and Mass Transfer, 49, pp. 4732–4750, 2006.

[11] **M. Heidari and M. Garshasbi,** Using Artificial Neural Networks in Solving Heat Conduction Problems, International Journal of Operations Research Vol. 12, No. 1, pp.016-020 , 2015.

[12] **Obed Cort´es, Gustavo Urquiza, Marco A. Cruz J. Alfredo Hern´andez**, Artificial Neural Networks for Inverse Heat Transfer Problems, Electronics, Robotics and Automotive Mechanics, 2007.

[13] **Sharath Kumar, Harsha Kumar, N. Gnanasekaran**, A NEURAL NETWORK BASED METHOD FOR ESTIMATION OF HEAT GENERATION FROM A TEFLON CYLINDER, Frontiers in Heat and Mass Transfer (FHMT), 7, 15, 2016.

[14] **S. Ghosh, D.K. Pratihar, B. Maiti and P.K. Das**, Inverse estimation of location of internal heat source in conduction, Inverse Problems in Science and Engineering, Taylor & Francis , 2011.

[15] **Yukio Hiranaka and Toshihiro Taketa**, DESIGNING, BACKWARD RANGE SIMULATOR FOR SYSTEM DIAGNOSES, Proc. XX IMEKO World Congress, 2012.

[16] **Yukio Hiranaka, Shinichi Miura, Toshihiro Taketa**, Hybrid backward simulator for determining causal heater state with resolution improvement of measured temperature

data through model conformation, ACTA IMEKO, Vol 6, No 1, pp.13-19, 2017.

[17] **Yukio Hiranaka, Shinichi Miura and Toshihiro Taketa**, Generic Causal Determination of thermal model and heat input by using backward simulator, Proc. XXII IMEKO World Congress, OR-134, 2018.9.6.

[18] **Node-RED**, https://nodered.org/

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Challenges of dimensional quantification in CCTV inspection in drain and sewer systems

L. L. Martins[1], A. S. Ribeiro[1], M. C. Almeida[1]

[1]*LNEC – National Laboratory for Civil Engineering, Lisbon, Portugal, lfmartins@lnec.pt, +351 21 844 3000*

***Abstract* – This paper addresses the metrological quality of dimensional measurements based on images obtained from CCTV inspections in drain and sewer systems. In this type of indirect visual inspection, a significant number of absolute and relative dimensional quantities can be quantified, contributing to the characterization of the observations and, consequently, to the analysis of performance of drain and sewer systems outside buildings.**

Unfavourable environmental factors and conditions within the drain or sewer components affect estimation of the quantities of interest and the quality of the recorded images (lighting, lack of reference points, geometric irregularities and subjective assessments, among others). Quantification improvement of the dimensional quantities is a key objective to achieve a better value from these inspections. This study contributes to improve the quality of the dimensional measurements by defining experimental procedures, which can be applied to the optical systems used in CCTV inspection.

The paper describes the European normative framework for these inspection activities and proposes approaches aiming at increasing confidence in the dimensional measurements, based on the metrological characterization of the optical systems used, as well as their integration in a traceability chain. Results presented and discussed include the evaluation of measurement uncertainty.

Keywords – CCTV inspection, drain system, sewer system, dimensional measurement.

I. INTRODUCTION

Drain and sewer systems are integrated in urban areas, often large and complex. These networks are composed by multiple types of components, such as drains and sewers, manholes, gullies, combined sewer overflows, storage and retention tanks, pumping stations, among others. These systems operate essentially under gravity to convey wastewater and stormwater to a treatment works or receiving water.

In the last decades, extensive work has been carried out by European standardization committees [1-3], aiming at the harmonization and improvement of strategic, policy and operational activities concerning drain and sewer systems. The established framework ambition is to contribute to fulfilment of strategic objectives of these services (including protection of public and occupational health and safety, environmental protection and sustainability). Set principles and functional requirements detailed in the standards are related to three main lines of action: (i) investigation and assessment; (ii) design and construction; (iii) management and control.

The investigation and assessment of drain and sewer systems outside buildings [2] aims to establish an overview of their condition and performance and it can be defined as a sequential process, which includes: (i) purpose; (ii) scope; (iii) review of existing information; (iv) inventory update; (v) hydraulic, environmental, structural and operational investigations. The results of these investigations allow determining the hydraulic performance, environmental impact, structural condition and operational deficiencies, supporting the comparison with the established performance requirements from which non-conformities can be identified and management plans can be updated.

The investigations are carried out using several sources of information, including external or internal inspection activities for the detection and characterization of anomalies which can negatively affect the performance of the drain or sewer system. Close circuit television (CCTV) inspection is a largely used visual inspection technique for non-man entry components. The use of a remotely controlled CCTV camera is generally motivated by: (i) safety issues, avoiding direct visual inspection by persons inside the drain and sewer systems; (ii) faster and economic advantages, when compared with quantitative methods such as laser scanning, ground piercing radar, sonar, infrared thermography and other available techniques [4].

This technique allows recording videos and individual images of the anomalies in drains, sewers, manholes and inspection chambers. A standard European coding system has been developed [3], in line with existing national

systems in member countries, facilitating common approaches and circulations of services within EU. Data on observations includes: (i) type (e.g. fissure, deformation, displaced joints, defective connections, roots, infiltration, settled deposits, attached deposits and other obstacles, subsidence, defects in manholes and inspection chambers, mechanical damage or chemical attack); (ii) characterisation; (iii) location (longitudinal and radial); (iv) quantification.

A significant number of absolute and relative dimensional quantities can be quantified, based on the recorded videos and images. Since these values are determining the results of the investigations and prioritization for correction actions [5], efforts towards reduction of systematic deviations and measurement uncertainties are extremely relevant. In parallel with technological developments, on-going research efforts include automated image processing techniques [6] aiming at the reduction of human intervention in image analysis (which is time-consuming, prone to human error and to subjectivity) and the increase of image quality e.g. by noise reduction [7]. However, discussion from a metrological perspective on the measurement accuracy was not found to be sufficiently developed.

In the following sections, this paper presents the problem of accurate dimensional measurements in CCTV inspections and proposes several approaches aiming to reduce subjectivity in the image dimensional analysis and increase measurement reproducibly, by the metrological characterization of the applied optical systems and their integration in a traceability chain.

II. THE PROBLEM OF ACCURATE DIMENSIONAL MEASUREMENTS IN CCTV INSPECTIONS

According the EN 13508-2:2003+A1:2011 standard [3], a number of quantities should be recorded as to fully characterise the observation identified in the inspection.

In the case of absolute dimensions measurements (e.g. in millimetres or in degrees), record of the quantities that can be required include:

- width of fissures, cracks, connections, channels and sections;
- height of connections and channels;
- length of breaks, collapsed regions, intruding connections and sections;
- maximum dimension of obstacles;
- thickness of deposits;
- depth of sediments, anomalies in walls;
- longitudinal and radial displacement of joints;
- level difference between coverage and surface;
- curvature;
- angular displacement at joints.

In addition, relative quantities that can be required (expressed in percentage) include [3]:

- dimensional reduction;
- level relative to the diameter or vertical dimension;
- reduction in effective cross section area;
- intrusion length relative to the diameter or vertical dimension.

Examples of images obtained from a CCTV camera (e.g. Figure 1) in drain and sewer systems inspections are shown in Figures 2 and 3, illustrating some anomalies and the corresponding quantification parameters.

Fig. 1. Remotely controlled CCTV camera used in drain and sewer systems inspections [8].

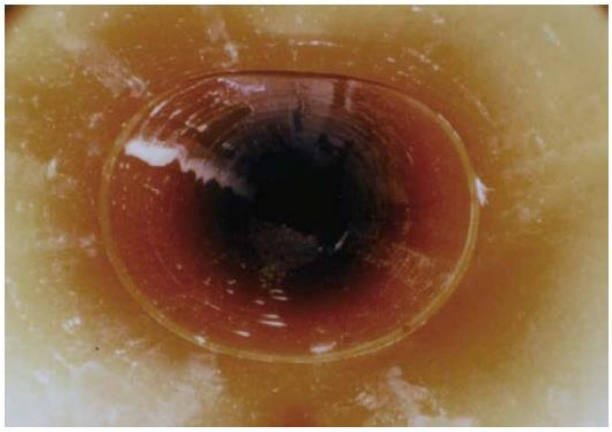


Fig. 2. Inspection image showing dimensional reduction by deformation effect [8].

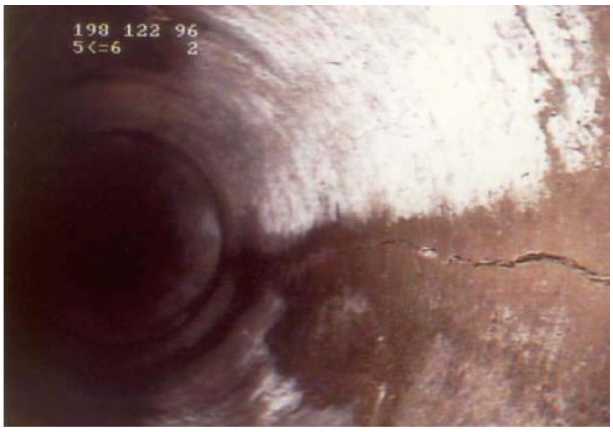


Fig. 3. Inspection image showing fissure [8].

As a rule, the longitudinal location of observations in this type of inspection, is measured by the CCTV system and recorded with a recommended resolution equal to 100 mm and the operational speed should be in the range of 0.1 m/s and 0.2 m/s to ensure adequate detection of observations.

In order to improve the accuracy of the measurements, special attention is given to the CCTV camera position inside the sewers or drains, thus reducing geometrical projection deviations. A tolerance up to 10% is defined for the CCTV camera vertical position (1/2 and 2/3 of the drain/sewer height, for a circular/regular or egg shape, respectively).

Lighting tests should be carried out before each inspection, using image quality testing charts such as the Marconi Resolution Chart no. 1, allowing in situ qualitative evaluation of the recorded images, namely, the radiometric and spatial resolution as well as the geometrical distortion.

Although benefits arise from these operational practices, limitations on non-existence of reference points inside the component under inspection makes any dimensional measurement vulnerable to subjective image analysis. In the following section, approaches aiming at improving measurement accuracy are proposed as well as the comparability between measurements undertaken by different operators or in different moments in time.

III. PROPOSED APPROACHES

A first approach, aiming at increasing the confidence in the dimensional measurements carried out, is to ensure the traceability of the metrological characterization of the optical system – the CCTV camera – used in drain or sewer inspection. This characterization includes both a geometrical component and a radiometric component, allowing a rigorous comparison between different CCTV cameras available in commercial solutions and their corresponding suitability to the measurement environment.

The main objective of the geometrical characterization is the quantification of CCTV camera intrinsic parameters such as the focal distance, the principal point coordinates and the lens distortion coefficients, which are input quantities for the determination of accurate dimensions in the field-of-view. This task can be achieved in a laboratorial setup, using traceable reference dimensional patterns and applying known algorithms such as the DLT – Direct Linear Transform [9], the Tsai [10] and the Zhang method [11].

The radiometric characterization can be carried out following the EMVA standard guidelines [12], allowing to determine the CCTV camera sensitivity, linearity, noise, dark current, spatial non-uniformity and defect pixels.

In addition to the CCTV camera metrological characterization, the measurement model itself must be defined and used in the dimensional measurement uncertainty evaluation, following the GUM framework [13-14].

If both the intrinsic (focal distance, principal point coordinates and lens distortion coefficients) and extrinsic (camera position and orientation in the world coordinate system) parameters of the camera are known, the perspective camera model [15] can be used to determine the coordinates related to points of interest inside the sewer or drain. Assuming, in a first approach to this problem, that the lens distortion does not have a significant impact on the dimensional accuracy when compared with other uncertainty components, the perspective camera model assumes a straight line between the interest point in world (X, Y, Z) and in the image (x, y) coordinate systems, defining the so-called collinearity equations

$$X = X_0 + (Z - Z_0) \cdot \frac{r_{11} \cdot (x - x_0) + r_{12} \cdot (y - y_0) - r_{13} \cdot f}{r_{31} \cdot (x - x_0) + r_{32} \cdot (y - y_0) - r_{33} \cdot f}, \quad (1)$$

$$Y = Y_0 + (Z - Z_0) \cdot \frac{r_{21} \cdot (x - x_0) + r_{22} \cdot (y - y_0) - r_{23} \cdot f}{r_{31} \cdot (x - x_0) + r_{32} \cdot (y - y_0) - r_{33} \cdot f}, \quad (2)$$

where: (i) intrinsic parameters – f is the focal distance and (x_0, y_0) are the principal point image coordinates; (ii) extrinsic parameters – (X_0, Y_0, Z_0) are the camera's coordinates in the world and r_{ij} are the rotation matrix elements defined by the camera's orientation angles Ω, Φ, K, given by

$$r_{11} = \cos\Phi \cdot \cos K, \quad (3)$$

$$r_{12} = -\cos\Phi \cdot \sin K, \quad (4)$$

$$r_{13} = \sin\Phi, \quad (5)$$

$$r_{21} = \cos\Omega \cdot \sin K + \sin\Omega \cdot \sin\Phi \cdot \cos K, \quad (6)$$

$$r_{22} = \cos\Omega \cdot \cos K - \sin\Omega \cdot \sin\Phi \cdot \sin K, \quad (7)$$

$$r_{23} = -\sin\Omega \cdot \cos\Phi, \quad (8)$$

$$r_{31} = \sin\Omega \cdot \sin K - \cos\Omega \cdot \sin\Phi \cdot \cos K, \quad (9)$$

$$r_{32} = \sin\Omega \cdot \cos K + \cos\Omega \cdot \sin\Phi \cdot \sin K, \quad (10)$$

$$r_{33} = \cos\Omega \cdot \cos\Phi. \quad (11)$$

If the camera's intrinsic and extrinsic parameters are not available, a less rigorous approach can be followed, assuming a parallel geometrical relation between the image plane and cross-section plane in the drain or sewer and using an orthographic projection camera model [15] to define a scale coefficient, SC, between real dimension (in millimetres) and image dimension (in pixels). In the

observed field-of-view, the only available dimensional reference is related to the height of the drain or sewer, *H*. Therefore, the dimensional quantification of a certain observation must be supported by the determination of the scale coefficient relative an average cross-section plane in the drain or sewer which includes the quantity to be measured and where its height is also visible, as shown in Figure 4.

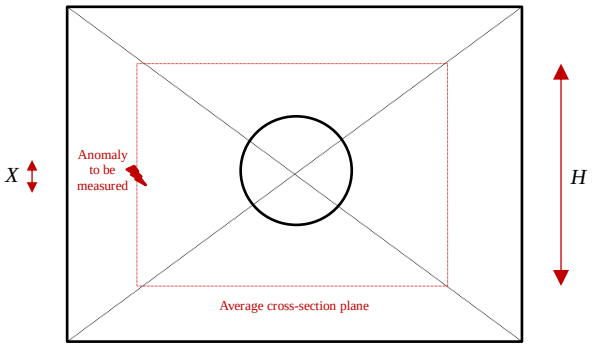


Fig. 4. Schematic representation of the measurement procedure in CCTV camera images.

In this approach, the image dimensions (in pixels) related to the drain or sewer height in the average cross-section plane, p_H, and to the observed anomaly, p_X, are obtained by image processing. Therefore, the scale coefficient, expressed in millimeters per pixel, is obtained by

$$SC = \frac{H}{p_H}, \tag{12}$$

while the dimensional measurement of the observation, *X*, (in millimeters) is given by

$$X = SC \cdot p_X. \tag{13}$$

Since these are linear mathematical models, the application of the Uncertainty Propagation Law [13], results in the standard measurement uncertainty expressions for the scale coefficient, $u(SC)$, and for the dimensional measurement of the anomaly $u(X)$, respectively,

$$u(SC) = \sqrt{\frac{1}{p_H^2} \cdot u^2(H) + \frac{H^2}{p_H^4} \cdot u^2(p_H)}, \tag{14}$$

$$u(X) = \sqrt{p_X^2 \cdot u^2(SC) + SC^2 \cdot u^2(p_X)}, \tag{15}$$

where $u(H)$ is the standard uncertainty of the height of the drain or sewer and $u(p_H)$ and $u(p_X)$ are the standard uncertainties of the image dimensions of p_H and p_X.

IV. RESULTS AND DISCUSSION

This section shows the evaluation of the dimensional measurement uncertainty related to the adoption of the perspective camera model or the orthographic projection camera model. The quantified uncertainty components and estimates are merely illustrative, reflecting values which are expected to have in a sewer or drain CCTV inspection scenario. Therefore, in a real case scenario, the presented probabilistic formulation and quantification must be confirmed and updated if required.

Due to the non-linear and complex mathematical models related to perspective camera model (expressions 1 to 11), the Monte Carlo method [14] was applied in the calculation of the dimensional measurement uncertainty. The computational simulation algorithm was developed in Matlab®, using the Mersenne-Twister pseudo-random number generator [16] and validated computational routines. 10^6 trials were used to obtain convergent solutions.

In the simulated observation scenario, the following estimates of the input quantities were considered: (i) camera location – the camera is placed near the centre of the drain or sewer cross-section (X_0=Y_0=0.10 m) and in a longitudinal position Z_0=1,00 m from the origin (the initial inspection position, for example); (ii) camera orientation – modern remotely controlled CCTV inspection systems often allow the camera rotation in three axis; in this case, the camera is considered to be orientated towards the centre of the drain or sewer cross-section, with reduced rotation angles (Ω=0.02 rad; Φ= 0.03 rad; K=0.01 rad); (iii) location of the observation plane – the point of interest is considered to be 50 mm in front of the camera location (Z=1.05 m); (iv) camera's intrinsic parameters, obtained from previous laboratorial characterization – focal distance, f=0.0010 m; the principal point is considered to be in the image centre (x_0=y_0=0); (v) image coordinates – the interest point is located in the following image coordinates: x=0.000 15 m and y=0.000 225 m.

Table 1 presents the adopted probabilistic formulation for the mentioned input quantities.

Table 1. Probabilistic formulation of the input quantities.

Uncertainty component	**Probability distribution**	**Standard uncertainty**
Camera location $u(X_0)$, $u(Y_0)$, $u(Z_0)$	Uniform	0.005 m / √3
Camera orientation $u(\Omega)$, $u(\Phi)$, $u(K)$	Uniform	0.01 rad / √3
Focal distance $u(f)$	Uniform	0.000 05 m / √3
Principal point coordinates $u(x_0)$, $u(y_0)$	Gaussian	$7.5 \cdot 10^{-7}$ m *
Location of the observation plane, $u(Z)$	Uniform	0.005 m / √3
Image coordinates $u(x)$, $u(y)$	Gaussian	$1.5 \cdot 10^{-4}$ m, $2.1 \cdot 10^{-4}$ m *

* Considering a ¼ squared pixel with 3 µm linear dimension.

Table 2 shows the obtained simulation results in terms of estimates, 95% expanded uncertainties and computational accuracy.

Table 2. Simulation results for X and Y.

Quantity	Estimate	95% expanded uncertainty	Computational accuracy
X	94.1 mm	5.0 mm	0.014 mm
Y	87.7 mm	5.5 mm	0.019 mm

Figures 5 and 6 correspond to the output normalized probability density functions related to *X* and *Y*, respectively, showing a trapezoidal shape as the result of the adopted input probabilistic formulation (uniform and Gaussian distributions).

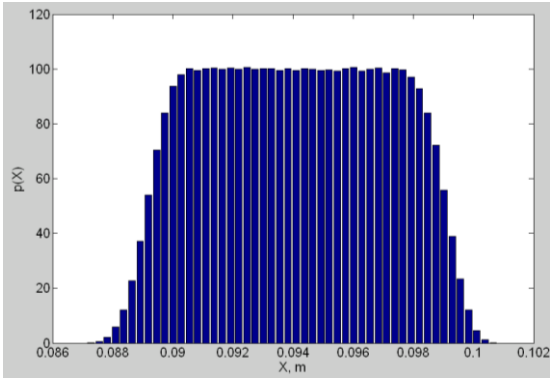


Fig. 5. Normalized probability density function for X

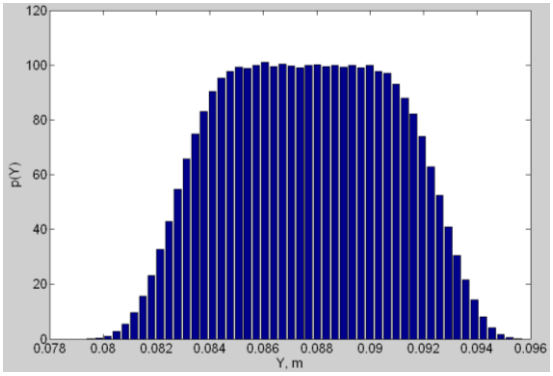


Fig. 6. Normalized probability density function for Y

A sensitivity analysis was performed with the develop simulation algorithm, revealing that the major uncertainty contributions correspond to the camera's location coordinates in the world coordinate system (X_0, Y_0, Z_0), as well as the *Z* longitudinal location of the observation plane (cross-section of the drain or sewer).

In order to evaluate their impact on the dimensional measurement accuracy, additional simulations were performed, considering extreme accuracy levels of 1 mm and 100 mm related to the above mentioned input quantities. The results are shown in Table 3.

Table 3. Impact of the camera and observation plane locations in the dimensional measurement accuracy

X_0, Y_0, Z_0 and Z accuracy levels	$U_{95\%}(X)$	$U_{95\%}(Y)$
1 mm	0.9 mm	1.1 mm
10 mm	5.0 mm	5.5 mm
100 mm	49 mm	54 mm

For the case of the orthographic projection camera model, Table 4 shows the input estimates used in the performed measurement uncertainty evaluation, while Table 5 refers to the probabilistic formulation of the corresponding standard measurement uncertainties.

Table 4. Measurement estimates of the input, intermediate and output quantities.

H	p_H	*SC*	p_X	*X*
/ mm	/ pixel	/ mm·pixel^{-1}	/ pixel	/ mm
300	350	0.857	30	25.7

Table 5. Probabilistic formulation of the measurement uncertainties of the input, intermediate and output quantities.

Uncertainty component	Probability distribution	Standard uncertainty	Relative standard uncertainty
$u(X)$	Gaussian	From 0.33 mm up to 4.3 mm	1% - 17%
$u(p_X)$	Gaussian	From 0.25 pixel up to 5 pixels	1% -17%
$u(SC)$	Gaussian	From 0.008 mm·pixel^{-1} up to 0.015 mm·pixel^{-1}	1% - 1.7%
$u(H)$	Uniform	5 mm / $\sqrt{3}$ = 2.9 mm	1%
$u(p_H)$	Gaussian	From 0.25 pixel up to 5 pixels	0.07% -1.4%

In the adopted formulation, the standard measurement uncertainty related to image dimensions was varied between 0.25 pixel and 5 pixels in order to show the impact of the image quality (related to lighting) and of measurement reproducibility (related to the manual selection of image points by different operators). The results show that, for a dimensional estimate of 25.7 mm, the relative standard uncertainty can vary between 1% and 17%.

V. CONCLUSIONS

This paper illustrated the main contributions of a metrological perspective to the problem of the accuracy of dimensional measurements in CCTV inspections. The proposed approaches constitute a first step towards the improvement of measurement accuracy in this field of inspection.

Measurement uncertainty analysis tools were developed and are now available to be used for the proposed approaches and experimental input information

regarding estimates and measurement uncertainties. However, the results obtained so far in this study, already indicate a high impact of the accuracy related to the camera and plane locations in the perspective camera model. In the orthographic projection camera model, the results show the strong influence of the image quality in the accuracy of the dimensional measurements.

These calculation tools can also be used to quantify accuracy operational improvements, namely, due to changes in field lighting, image analysis, CCTV camera selection and metrological characterization.

Future work will be focused on determining the uncertainty component related to the adoption of the orthographic projection camera model, which is considered an approximation of the perspective camera model. In this last camera model, the impact of the lens distortion should also be investigated.

REFERENCES

[1] **EN 752:2017**: *Drain and sewer systems outside buildings. Sewer system management*, CEN - European Committee for Standardization, 2017.

[2] **EN 13508-1:2012**: *Investigation and assessment of drain and sewer systems outside buildings – Part 1: General requirements*, CEN - European Committee for Standardization, 2012.

[3] **EN 13508-2:2003+A1:2011**: *Investigation and assessment of drain and sewer systems outside buildings – Part 2: Visual inspection coding system*, CEN - European Committee for Standardization, 2011.

[4] **Makar, J.M.**: Diagnostic techniques for sewer systems, *J. Infrastruct. Syst*, 5, 1999, pp. 69-78.

[5] **Fenner, R.A.**: Approaches to sewer maintenance: a review, *Urban Water*, 2, 2000, 343-356.

[6] **Su, T.C.; Yang, M.D.**: Application of morphological segmentation to leaking defect detection in sewer pipelines, *Sensors*, 14, 2014, pp. 8686-8704.

[7] **Su, T.C.; Yang, M.D.; Pan, N.F.; Yang, Y.F.**: Systematic image quality assessment for sewer inspection, *Expert Syst. Appl.*, 38, 2011, pp. 1766-1776.

[8] **Henley, P.**: Sewer condition – classification and training course, *WRC*, 2017.

[9] **Heikkilä, J.; Silvén, O.**: A four-step camera calibration procedure with implicit image correction, *Conference on Computer Vision and Pattern Recognition – Proceedings CVPR 1997*, IEEE Computer Society, 1997.

[10] **Tsai, R.**: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE Journal of Robotics and Automation*, RA-3, 4, 1987.

[11] **Zhang, Z.**: A flexible new technique for camera calibration, *IEEE Transctions on Pattern Analysis and Machine Intelligence*, 22, 11, 2000.

[12] **EMVA 1288:2010**: *EMVA Standard 1288. Standard for Characterization of Image Sensors and Cameras*, European Machine Vision Association, release 3.0, 2010.

[13] **GUM**: *Guide to the expression of Uncertainty in Measurements*, International Organization for Standardization (ISO), 1st edition, Genève (Switzerland), 1993, ISBN 92-67-10188-9.

[14] **JCGM 101**: *Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method*, Joint Committee for Guides in Metrology, 1st edition, 2008.

[15] **Hartley, R.; Zisserman, A.**: *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd edition, New York (USA), 2003, ISBN 978-0521-54051-3.

[16] **Matsumoto, M.; Nishimura, T.**: Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator, *ACM Transactions on Modeling and Computer Simulation*, 8, 1, 1998, 3-30.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"
Berlin, Germany, September 3-4, 2019

A consensus ranking based proposal for combining data in adjustment of the fundamental physical constant values

Sergey V. Muravyov, Liudmila I. Khudonogova

Division for Automation & Robotics, National Research Tomsk Polytechnic University, Tomsk, Russian Federation, email: [muravyov, likhud]@tpu.ru

***Abstract* – To ensure the traceability of testing and diagnostic equipment it is necessary to provide a chain of comparisons connecting the equipment with primary standards of the SI units based on fundamental physical constants. The values of the constants are regularly determined by an adjustment procedure which requires consistency of input data and an assumed statistical model. In this paper, it is proposed to apply the developed interval fusion with preference aggregation (IF&PA) method for combining data and determining a consensus value of a fundamental constant. Due to its high robustness, accuracy and reliability confirmed by the numerical experimental results, the IF&PA does not require a consistency check and works well without using any statistical assumptions. Usage of the IF&PA is demonstrated by example of processing the simulated interval data and real values of the Planck constant used in the adjustment in 2006 and 2017. The outcome comparison with the estimates obtained by other methods, including the procedures based on Birge ratio, modified Birge ratio, random effects model and fixed effects model, is carried out.**

Keywords – Fundamental Constant Adjustment, Consensus Estimate, Ranking, Data Fusion, Preference Aggregation.

I. INTRODUCTION

According to international requirements [1], all measuring instruments (including those intended for testing and diagnostics) have to be calibrated before putting into operation; they also should be recalibrated within prescribed calibration period. The programme for calibration is supposed to ensure that measurement results are traceable to the SI units. Traceability is established by means of an unbroken chain of comparisons that connects measuring instruments with primary standards, which are agreed representations of the SI units based on fundamental physical constants. Thus, the accuracy of any measuring, testing or diagnostic procedure directly depends on the correctness of determination of the constant values.

A self-consistent set of internationally recommended values of the basic fundamental constants is periodically provided by the Committee on Data for Science and Technology (CODATA) [2]. These values of the constants along with their quoted uncertainties are determined by a special procedure called *adjustment*. The adjustment is the processing of numerical set of measured (or calculated) values, obtained from experiment or theory by different methods and provided from different laboratories, to estimate the best *consensus value* of a constant. Every four years, the CODATA Task Group on Fundamental Constants carries out a *least squares adjustment* (LSA) of the fundamental constants taking into account all relevant data from diverse researches [3].

The obligatory condition for the adjustment procedure is a *consistency* of input data. It is an often situation where the data provided by different laboratories are inconsistent, i.e. discrepancy between the input values are larger than it is expected taking into consideration the quoted uncertainties. Such inconsistency is caused by some unknown sources of uncertainty, which should be identified, but in practice it is hard to implement. To deal with data inconsistencies, the Task Group has to omit some values from the adjustment set or increase the initial uncertainties of the values by an expansion factor.

More to the point, the determination of a consensus value of a constant on the basis of an inconsistent input data set requires additional assumptions. Thus, existing techniques used for the adjustment, including LSA, suppose that the input data are normally distributed, independent random variables with the same mean μ and exactly known variances u_k^2, but this hypothesis is not always valid [4].

This paper is the authors' first attempt to apply an alternative approach for combining the data and determining a consensus value of a fundamental constant. The approach is based on our interval fusion with preference aggregation (IF&PA) method [5], which is found to be highly robust to successfully deal with

inconsistent data. The IF&PA doesn't require any consistency check as well as any assumptions about the distribution law and independency of data, and uses all the available values of a constant for the adjustment procedure.

II. COMBINING DATA IN ADJUSTMENT OF THE FUNDUMENTAL CONSTANTS

In this Section, the procedure for combining data used in adjustment of the fundamental constants is considered.

Let $x_1, \ldots, x_m$ be a set of measured (or calculated) values of the same fundamental constant, along with corresponding standard uncertainties $u_1, \ldots, u_m$. The aim is to combine these data for obtaining a single estimate x^*, and to determine the uncertainty u^* associated with x^*.

The estimate x^* is typically calculated as a weighted mean

$$x^* = \sum_{k=1}^{m} u_k^{-2} x_k \Big/ \sum_{k=1}^{m} u_k^{-2} \qquad (1)$$

with corresponding uncertainty

$$u^* = \left(1 \Big/ \sum_{k=1}^{m} u_k^{-2} \right)^{1/2}, \qquad (2)$$

where m is a number of values of a constant.

Decision about the consistency of the data set is made by conducting the chi-square test and computing the statistics:

$$\chi^2 = \sum_{k=1}^{m} (x^* - x_k)^2 / u_k^2, \qquad (3)$$

where the χ^2 is a value of a chi-square random variable with $v = m - 1$ degrees of freedom.

Among procedures used for combining inconsistent data one of the most popular is a procedure based on the Birge ratio [6]. It is applied by the CODATA Task Group for the adjustment. The procedure assumes that values x_k are normally distributed with mean μ and variance $c^2 u_k^2$, $k = 1, \ldots, m$. Then the standard uncertainty of the x^*, is expanded and estimated as $u_B^* = cu^*$.

To evaluate c, the Birge ratio, which is a measure of the data consistency, is calculated with v degrees of freedom:

$$R_B = \sqrt{\chi^2 / v} = c. \qquad (4)$$

In case of combining data of the same constant, the value $(\chi^2 / m - 1)$ is an unbiased estimator of c^2, hence, the standard uncertainty of x^* is computed as follows:

$$u_B^* = u^* \sqrt{\chi^2 / m - 1}. \qquad (5)$$

In literature, there are alternative procedures for combining data in the adjustment of the constants, for example, procedures based on the random effects model [7, 8] or fixed effects model [7, 9]. This paper is not focused on their detailed description, but exploits the results obtained by these procedures in [7] for comparison with our IF&PA method.

III. INTERVAL FUSION WITH PREFERENCE AGGREGATION

This Section presents the interval fusion with preference aggregation (IF&PA) method for combining inconsistent data in adjustment of the fundamental constants.

Every value x_k of a fundamental constant along with its uncertainty u_k can be represented as an *interval* I_k on the real line. We consider a set of initial intervals I_k, $k = 1, \ldots, m$, as input data for the IF&PA method.

By *interval data fusion* we understand a procedure of determination of a resulting interval $[x^* \pm u^*]$ which is consistent with the maximal number of initial intervals I_k, $k = 1, \ldots, m$, and with maximum likelihood contains a point x^* (fusion result) that can be considered as a representative of all I_k with an uncertainty u^*.

The fusion result x^* is supposed to be selected from finite set of points related to initial intervals. To form this finite set we introduce a concept of a range of actual values (RAV) $A = \{a_1, a_2, \ldots, a_n\}$ consisting of ordered discrete values a_i, $i = 1, \ldots, n$.

The RAV is composed by union of all intervals I_k with the lower bound $a_1 = \min\{l_k \mid k = 1, ..., m\}$ and the upper bound $a_n = \max\{u_k \mid k = 1, ..., m\}$, where l_k and u_k are the lower and upper bounds of some k-th interval correspondently. Next, we partition this union into $n - 1$ equal subintervals of length h to obtain elements $a_2, a_3, \ldots, a_{n-1}$. Number n, or cardinality of the RAV partition, is calculated by (6):

$$n = \lceil (a_n - a_1) / 0.31\sigma \rceil + 1, \qquad (6)$$

where σ is a standard deviation value computed before RAV partition [5]. After partitioning, the length h of interval is defined as $h = |a_i - a_{i-1}|$, $i = 2, \ldots, n$. Thus, as a result of RAV partitioning it is obtained a set $A = \{a_1 < a_2 < \ldots < a_i < \ldots < a_n\}$ of fully ordered discrete values a_i, $i = 1, \ldots, n$, which are then ranked in a specific way to introduce every particular interval.

On the next stage, the initial intervals I_k are represented by rankings $\lambda_k = (a_1 \succ a_2 \succ \ldots \sim a_n)$ over a set of n discrete values $A = \{a_1, a_2, \ldots, a_n\}$ *belonging to* these intervals. Ranking λ_k induced by interval I_k, or *inranking*, is composed according to conditions (7) for $i, j = 1, \ldots, n$:

$$\begin{aligned} &\text{(i) } a_i \in I_k \wedge a_j \notin I_k \Rightarrow a_i \succ a_j; \\ &\text{(ii) } a_i, a_j \in I_k \vee a_i, a_j \notin I_k \Rightarrow a_i \sim a_j; \\ &\text{(iii) } a_i \notin I_k \wedge a_j \in I_k \Rightarrow a_i \prec a_j; \\ &\text{(iv) } a_i, a_j \in I_k \text{ are neighbors } \Rightarrow j \equiv i + 1 \end{aligned} \qquad (7)$$

A set of m inrankings $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ called *preference profile* is formed for initial intervals I_k. For Λ we determine a single preference relation (or *consensus ranking*) β which represents the best compromise between the initial inrankings. To do this, we apply recursive branch and bounds algorithm implementing the Kemeny rule [10]. It finds such a linear order β of the alternatives of A that the distance $D(\beta, \Lambda)$ (defined in terms of the number of pairwise inconsistencies between

inrankings) from β to the inrankings of the initial profile $\Lambda(m, n)$ is minimal for all possible strict orders ρ, that is

$$\beta = \arg\min \sum_{i<j} p_{ij} , \quad (8)$$

where $p_{ij} = \sum_{k=1}^{m} [1 - \mathrm{sgn}(a_i^k, a_j^k)]$ is an element of the $(n \times n)$ profile matrix $[p_{ij}] = P$, rows and columns of which are labeled by the alternatives' numbers; $\mathrm{sgn}(a_i, a_j)$ is a function that reveals the sign (or direction) of the pair $(a_i, a_j) \in \lambda$. The recursive algorithm determines multiple consensus rankings $\beta_1, \beta_2, \ldots, \beta_N$ which are transformed into a single final consensus ranking β_{fin} by means of a convolution rule [11].

The convolution rule is formulated as follows. Let $B(N, n) = \{\beta_1, \beta_2, \ldots, \beta_N\}$, $B \subset \Pi_n$, be a set of consensus rankings for a profile $\Lambda(m, n) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ given over some set of alternatives $A = \{a_1, a_2, \ldots, a_n\}$. Let *rank* r_i^t be a position of an alternative a_i in the consensus ranking $\beta_t \in B$, $t = 1, \ldots, N$. Then, for all $i < j$, $i, j = 1, \ldots, n$, we have

$$\sum_{t=1}^{N} r_i^t < \sum_{t=1}^{N} r_j^t \Rightarrow a_i \succ a_j \text{, and } \sum_{t=1}^{N} r_i^t = \sum_{t=1}^{N} r_j^t \Rightarrow a_i \sim a_j, \quad (9)$$

where both strict order relation $\succ$ and tolerance relation $\sim$ are in the single final consensus ranking β_{fin}.

The main stages of the IF&PA method are as follows.

1. Forming the RAV from initial intervals I_k, $k = 1, \ldots, m$ and partitioning it into $n - 1$ subintervals in accordance with (6) to obtain a set $A = \{a_1, a_2, \ldots, a_n\}$.
2. Composing preference profile $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ consisting from inrankings in accordance with (7).
3. Determining a fusion result x^* as the best alternative in consensus ranking β_{fin} by (8-9).
4. Determining an uncertainty u^* of the fusion result x^* by the formula (10)

$$u^* = \min\left(\max_{k=1,\ldots,m'} \{l_k \le x^*\}, \min_{k=1,\ldots,m'} \{u_k \ge x^*\} \right), \quad (10)$$

where m' is a number of intervals from I_k which include the value x^*.

IV. RESULTS AND DISCUSSIONS

This Section presents the results of processing of interval data by the IF&PA as well as the comparison with the results obtained by other methods. The processed data are both simulated random values and real Planck constant values from the CODATA set.

A. Processing data of numerical simulation

To verify the proposed IF&PA method for combining the data we developed special software IntFusion in the NI LabVIEW environment. The software generates synthetic pseudo-random interval data, which represent measured values of the Planck constant with relative uncertainties as the pairs $\langle x_k, u_k \rangle$, $k = 1, \ldots, m$, using Monte-Carlo method [12].

In experiments, the CODATA recommended value 6.626 070 150(69) × 10^{-34} J s was used as a nominal (true) value x_{nom} of the measured quantity. The number m of intervals was set to be 9, and there were conducted 100 individual experiments for both normal and uniform distributions of interval data.

Since the nominal value is preset in simulation one can judge the procedure performance by the deviation ξ of the obtained value x^* from the nominal value x_{nom}:

$$\xi = |x_{\mathrm{nom}} - x^*|. \quad (11)$$

The deviation ξ allows to estimate a set of three characteristics of the procedure: robustness, accuracy, and reliability. We treat *robustness* as an independence of the procedure outcome x^* on a particular law of interval data distribution. Under *accuracy* we understand a closeness of the outcome x^* to the nominal value x_{nom}. *Reliability* is considered as a degree of belief to the obtained outcomes x^*.

The deviation ξ can be graphically demonstrated for the visual comparison of robustness, accuracy, and reliability of the procedures. The outcomes x^* of 100 experiments is presented as a curve $\xi(v)$, where values ξ are sorted in ascending order and v is a number of a particular experiment. On the graph, robustness is demonstrated by the discrepancy between the curves $\xi(v)$ for normal and uniform distributions; the smaller the discrepancy the higher the robustness of a procedure. Accuracy is determined by the distance between the curve $\xi(v)$ and x-axis.

As a measure of reliability the probabilities $P(\xi \le \xi_b)$ were used that the value ξ does not exceed some fixed *boundary deviation* ξ_b. For instance, if $\xi \le 0.5$ in 95 experiments out of the 100, then the probability $P(\xi \le 0.5) = 0.95$.

The generated interval data were processed by the procedure IF&PA and by the procedure based on Birge ratio. The resulting estimate x^*, its relative uncertainty $u_r^* = u^*/x^*$ and deviations ξ were determined. An example of the simulated values of x_k (which represent Planck constant values h_k) and $u_r = u_k / x_k$ for one arbitrary selected experiment are presented in Table 1.

Table 1. Simulated measured values of the Planck constant with corresponding relative uncertainties.

Number k of interval	$x_k = h_k$, J s	u_r
1	6.626 070 138 × 10^{-34}	3.2×10^{-9}
2	6.626 070 163 × 10^{-34}	4.1×10^{-9}
3	6.626 070 152 × 10^{-34}	3.3×10^{-10}
4	6.626 070 144 × 10^{-34}	3.0×10^{-10}
5	6.626 070 144 × 10^{-34}	1.0×10^{-9}
6	6.626 070 155 × 10^{-34}	3.3×10^{-9}
7	6.626 070 143 ×10^{-34}	4.6×10^{-9}
8	6.626 070 157 ×10^{-34}	1.1×10^{-9}
9	6.626 070 174 ×10^{-34}	1.4×10^{-9}

After the processing, for the particular example the following results were obtained:

- for the IF&PA method – the resulting estimate $x^* = 6.626\ 070\ 152 \times 10^{-34}$ J s, the relative standard uncertainty $u_r^* = 2.34 \times 10^{-10}$, the deviation $\xi = 8.58 \times 10^{-10}$ J s;
- for the Birge ratio procedure – the resulting estimate $x^* = 6.626\ 070\ 148 \times 10^{-34}$ J s, the relative standard uncertainty $u_r^* = 2.09 \times 10^{-10}$, the deviation $\xi = 2.30 \times 10^{-9}$ J s.

The simulated interval data as well as the obtained outcomes of two procedures are shown in Fig 1. The dashed line on the graph indicates the nominal value.

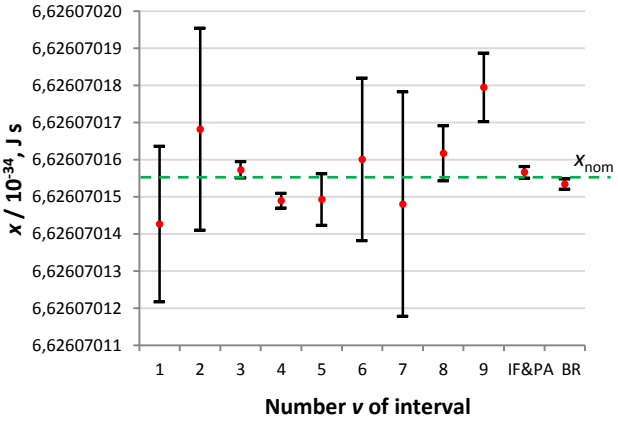


Fig. 1. An example of simulated interval data and the outcomes of the IF&PA and the Birge ratio procedures.

For the given example, the relative uncertainties for both procedures are almost the same, while resulting estimate obtained by the IF&PA is closer to the nominal value.

To evaluate the characteristics of procedure performance, the curves $\xi(v)$ were constructed (Fig. 2). One can see from Fig. 2 that the average discrepancy between normal and uniform distributions (curves 3 and 4), in case of the IF&PA, is about 0.60×10^{-7}, whereas for the Birge ratio procedure (curves 1 and 2) the discrepancy is three times larger – about 1.99×10^{-7}. These results denote the noticeably higher robustness of the IF&PA against the Birge ratio procedure.

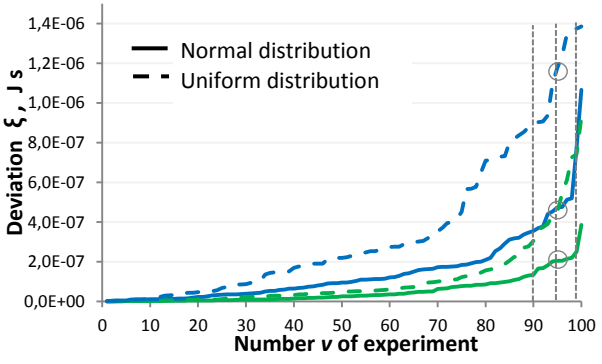


Fig. 2. Deviations ξ obtained by the IF&PA and the Birge ratio procedures for uniform and normal distributions of data.

It is evident from Fig. 2 that the curves $\xi(v)$ for the IF&PA are closer to the x-axis than those obtained by the Birge ratio procedure for both normal and uniform distribution. The average distances to the x-axis are 0.52×10^{-7} and 1.48×10^{-7} (case of normal distribution), and 1.12×10^{-7} and 3.47×10^{-7} (case of uniform distribution), for the IF&PA and the Birge ratio procedure correspondently. Therefore, the IF&PA procedure have demonstrated higher (approximately 3 times) accuracy than the Birge ratio procedure.

The calculated values of boundary deviations ξ_b for probabilities $P = 0.90$, $P = 0.95$ and $P = 0.99$ are shown in Table 2. They are also illustrated in Fig. 2 as cross points of the curves $\xi(v)$ with dashed perpendiculars to the x-axis.

Table 2. Boundary deviations ξ_b for probabilities $P = 0.90$, $P = 0.95$ and $P = 0.99$.

Probability	Boundary deviation ξ_b, 10^{-7}·J s			
	IF&PA		Birge ratio	
	Normal	Uniform	Normal	Uniform
0.90	1.33	3.04	3.54	8.80
0.95	2.04	4.50	4.73	11.77
0.99	2.50	7.42	7.73	13.73

It follows from the data given in Table 2 that for all considered probabilities P the smallest values of boundary deviations ξ_b were provided by the IF&PA procedure.

Thus, the IF&PA procedure has shown improved performance characteristics, namely robustness, accuracy, and reliability, in comparison with the Birge ratio procedure.

Fig. 3 demonstrates the obtained values of the the relative standard uncertainty u_r^* sorted in ascending order for 100 numerical experiments. As can be seen, the obtained uncertainties are of the same order of magnitude and in general commensurate for two procedures considered in the experiment.

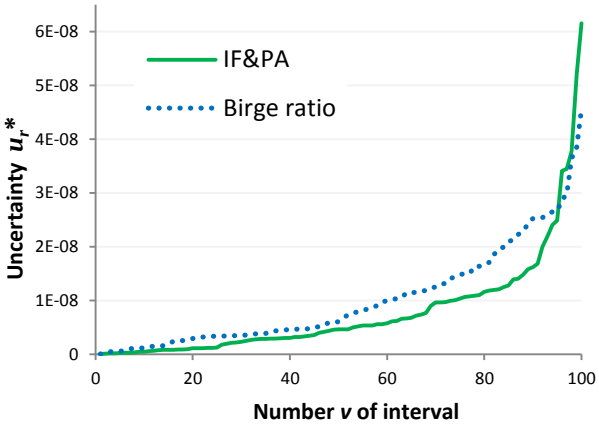


Fig. 3. Relative uncertainties u_r obtained by the IF&PA and the Birge ratio procedures.

B. Processing data of the CODATA adjustment 2006

For the sake of comparison simplicity, we analyzed data set from [7], which represents the values of the Planck constant h (Table 3). The set includes eight values used for the CODATA adjustment of 2006, and four values obtained after that adjustment. Twelve measured values of the Planck constant h with corresponding relative uncertainties u_r are also shown in Fig.4.

The resulting estimate of the Planck constant obtained by the IF&PA is given in Table 4 and graphically demonstrated in Fig.4.

Table 3. Measured values of the Planck constant with relative uncertainties as in [7].

Identification	h, J s	u_r
NPL-79	$6.626\,0729 \times 10^{-34}$	1.0×10^{-6}
NIST-80	$6.626\,0657 \times 10^{-34}$	1.3×10^{-6}
NMI-89	$6.626\,0684 \times 10^{-34}$	5.4×10^{-7}
NPL-90	$6.626\,0682 \times 10^{-34}$	2.0×10^{-7}
PTB-91	$6.626\,0670 \times 10^{-34}$	6.3×10^{-7}
NIM-95	$6.626\,071\times 10^{-34}$	1.6×10^{-6}
NIST-98	$6.626\,068\,91 \times10^{-34}$	8.7×10^{-8}
NIST-07	$6.626\,068\,91 \times10^{-34}$	3.6×10^{-8}
METAS-11	$6.626\,0691 \times10^{-34}$	2.9×10^{-7}
NPL-12	$6.626\,0712 \times10^{-34}$	2.0×10^{-7}
NRC-12	$6.626\,070\,63 \times10^{-34}$	6.5×10^{-8}
Avogadro-11	$6.626\,070\,09 \times10^{-34}$	3.0×10^{-8}

In [7], the following combining data procedures were considered based on: (1) Birge ratio, (2) modified Birge ratio, (3) random effects model and (4) fixed effects model. The outcomes obtained by means of these procedures are presented in Table 4.

Table 4. Resulting estimates of the Planck constant for data from Table 3.

Method	h^*, J s	u_r^*
Birge ratio	$6.626\,069\,67 \times 10^{-34}$	3.13×10^{-8}
Modified Birge ratio	$6.626\,069\,67 \times 10^{-34}$	3.46×10^{-8}
Random effects	$6.626\,069\,60 \times 10^{-34}$	6.68×10^{-8}
Fixed effects	$6.626\,069\,60 \times 10^{-34}$	9.21×10^{-8}
IF&PA	$6.626\,069\,34 \times 10^{-34}$	1.43×10^{-8}

As can be seen, the resulting estimate h^* obtained by the IF&PA noticeably differs from the others. It is expectable since the procedures (1)–(4) estimate the value h^* by calculating a weighted mean using formula (1) directly or with some modifications. This yields almost similar values of the estimate for all procedures (1)–(4).

On the other hand, Fig.4 demonstrates that the IF&PA provides the best compromise value for the largest subset of consistent intervals without preliminary elimination of any initial interval. Moreover, the IF&PA has allowed to have the resulting estimate of the Planck constant with a lower standard uncertainty u_r^*.

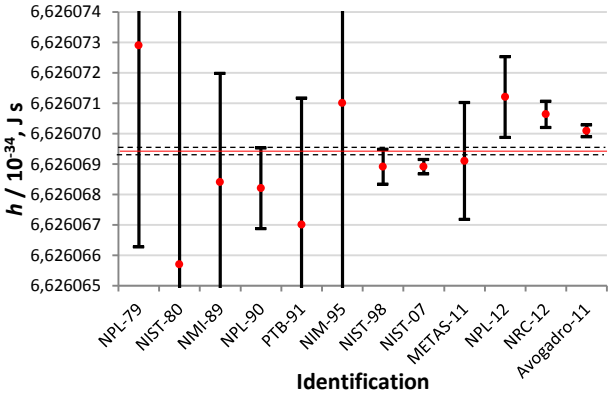


Fig. 4. Measured data of the Planck constant as in [7]. The IF&PA resulting estimate is demonstrated by the red line with the uncertainty boundaries denoted by dashed lines

C. Processing data of the CODATA special adjustment

The IF&PA method was applied to processing the latest values of the Planck constant h (Table 5) used for the CODATA special adjustment in 2017 carried out for revision of International System of Units [13, 14].

Table 5. Measured values of the Planck constant with relative uncertainties as in [13].

Identification	h, J s	u_r
NRC-17	$6.626\,070\,133(60) \times 10^{-34}$	9.1×10^{-9}
IAC-17	$6.626\,070\,405(77) \times 10^{-34}$	1.2×10^{-8}
NIST-17	$6.626\,069\,934(88) \times 10^{-34}$	1.3×10^{-8}
IAC-15	$6.626\,070\,22(13) \times 10^{-34}$	2.0×10^{-8}
NMIJ-17	$6.626\,070\,13(16) \times 10^{-34}$	2.4×10^{-8}
IAC-11	$6.626\,069\,94(20) \times 10^{-34}$	3.0×10^{-8}
LNE-17	$6.626\,070\,40(38) \times 10^{-34}$	5.7×10^{-8}
NIST-15	$6.626\,069\,36(38) \times 10^{-34}$	5.7×10^{-8}
NIST-98	$6.626\,068\,91(58) \times 10^{-34}$	8.7×10^{-8}

The results of processing data from Table 5 by the IF&PA are illustrated in Fig.5. The resulting estimates h^* of the Planck constant along with the relative standard uncertainties u_r^* obtained by the IF&PA and by the Birge ratio procedure (i.e. CODATA 2017 recommended value) are shown in Table 6.

Table 6. Resulting estimates of the Planck constant for data from Table 5.

Method	h^*, J s	u_r^*
Birge ratio	$6.626\,070\,150\,(69) \times 10^{-34}$	1.0×10^{-8}
IF&PA	$6.626\,070\,170\,(90) \times 10^{-34}$	3.5×10^{-9}

One can see that the IF&PA estimate is quite close to the one recommended in [13], but has a much lower uncertainty. It should be mentioned that the IF&PA exploited all nine available values of the Planck constant to find a result, while the Birge ratio procedure did not include the value of NIST-98 in the adjustment.

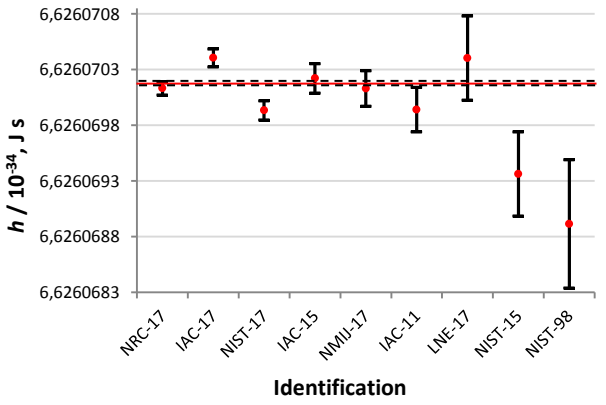


Fig. 5. Measured data of the Planck constant as in [13]. The IF&PA resulting estimate is demonstrated by the red line with the uncertainty boundaries denoted by dashed lines

V. CONCLUSIONS AND OUTLOOK

We proposed an alternative approach for combining inconsistent data in adjustment of the fundamental constants which we called the interval fusion with preference aggregation (IF&PA) method. In contrast with other procedures, it does not need a preliminary consistency check. As demonstrated by the results of numerical experimental investigations, the IF&PA shows high robustness, accuracy, reliability, and works well without using any statistical assumptions.

The results of processing real values of the Planck constant have shown that the resulting estimates provided by the IF&PA method differ from estimates obtained by Birge ratio procedure, modified Birge ratio procedure, procedures based on the random effects model and on the fixed effects model. This is caused by the fact that the IF&PA does not assign weights to the data or eliminate inconsistent values, i.e. it combines all available input data. However, we may consider the outcomes of the IF&PA method as rather reliable, as the experimental investigation results demonstrated the ability of the IF&PA to provide accurate (close to the true value) resulting estimate.

In case of processing real data, the standard uncertainty determined by the IF&PA is lower in comparison with other procedures. But we should note that the IF&PA method supposes to find the "best case" value of uncertainty. In practice it could be too optimistic, and to overcome this difficulty the dependence of the uncertainty on the RAV partition should be further investigated. Another line of development of the IF&PA concerns taking into account possible correlation among data.

VI. ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation, project # 18-19-00203, in part of the interval data fusion by preference aggregation approach development, and by the Russian Ministry of Education and Science, basic part of the state task "Science", project 2.5760.2017/8.9, in part of numerical experimental investigations.

REFERENCES

[1] **ISO/IEC 17025:2017** General requirements for the competence of testing and calibration laboratories, 2017.

[2] **Mohr, P.J., Newell, D.B., and Taylor, B.N.** CODATA recommended values of the fundamental physical constants: 2014, *Reviews of Modern Physics*, Vol. 88, 2016, 035009.

[3] **Mohr, P.J., and Taylor, B.N.** CODATA recommended values of the fundamental physical constants: 1998, *Reviews of Modern Physics*, Vol. 72, No. 2, 2000, pp. 351-495.

[4] **Mana, G., Massa E., and Predescu M.** Model selection in the average of inconsistent data: an analysis of the measured Planck-constant values, *Metrologia*, Vol. 49, 2012, pp. 492-500.

[5] **Muravyov, S.V., Khudonogova, L.I, and Emelyanova, E.Y.** Interval data fusion with preference aggregation, *Measurement*, Vol. 116, 2018, pp. 621-630.

[6] **Birge, R.T.** Probable Values of the General Physical Constants, *Reviews of Modern Physics*, Vol. 1, Iss. 1, 1929, pp. 1-73.

[7] **Toman, B., Fischer, J., and Elster C.** Alternative analyses of measurements of the Planck constant, *Metrologia*, Vol. 49, 2012, pp. 567-571.

[8] **Bodnar, O., Elster C., Fischer, J., Possolo, A., and Toman, B.** Evaluation of uncertainty in the adjustment of fundamental constants, *Metrologia*, Vol.53, 2016, pp. 46-54.

[9] **Mana, G.** Model uncertainty and reference value of the Planck constant, *Measurement*, Vol. 94, 2016, pp. 26-30.

[10] **Muravyov, S.V.** Ordinal measurement, preference aggregation and interlaboratory comparisons, *Measurement*, Vol. 46, No. 8, 2013, pp. 2927-2935.

[11] **Muravyov, S.V., Baranov, P.F., Emelyanova, E.Y.** How to transform all multiple solutions of the Kemeny Ranking Problem into a single solution, *Proceeding of Joint IMEKO TC1-TC7-TC13-TC18 Symposium 2019 "The future glimmers long before it comes to be"*, Saint Petersburg, Russia, July 2-5, 2019.

[12] **ISO/IEC Guide 98-3:2008/Suppl 1:2008** (JCGM/WG1/101) Propagation of distributions using a Monte Carlo method.

[13] **Mohr, P.J., Newell, D.B., Taylor, B.N., and Tiesinga, E.** Data and analysis for the CODATA 2017 special fundamental constants adjustment, *Metrologia*, Vol. 55, 2018, pp. 125-146.

[14] **Newell, D.B. *et al*** The CODATA 2017 values of *h*, *e*, *k*, and N_A for the revision of the SI, *Metrologia*, Vol. 55, 2018, pp. L13- L16.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Improving the Planning Quality in Production Planning and Control with Machine Learning

Thomas Ryback[1], Lukas, Lingitz[1], Alexander Gaal[1], Viola Gallina[1*], Dávid Gyulai[2]

[1]*Fraunhofer Austria Research GmbH,*
Theresianumgasse 27, A-1040, Vienna, Austria,
[*]*viola.gallina@fraunhofer.at, +43 676 888 616 46*
[2] *Centre of Excellence in Production Informatics and Control (EPIC), Institute for Computer Science and Control (SZTAKI), Hungarian Academy of Sciences (MTA),*
Kende *S*tr. 13-17, H-1111, Budapest, Hungary

Abstract – There are always deviations between production planning and subsequent execution. These deviations are caused by uncertainties, e.g. inaccurate or insufficient planning data (e.g. data quality and availability), inappropriate planning and control systems or unforeseeable events. Production planners therefore use buffers in the form of inventories or extended transitional periods to create possibilities for implementing corrective measures in production control. Buffers, however, lead to increased coordination and control effort and to negative effects, e.g. on inventory, throughput time and capacity utilization. Furthermore, it was found that the reliability of the production plans and thus the planning quality (PQ) can drop down to 25% in the first three days after plan creation [1]. Potentials for more accurate planning remains largely unexploited. **The objective of this paper is to investigate the possibilities to increase planning quality. Two approaches are presented, focusing on reducing gaps between master data and predicted data used during the production planning process.**

Keywords – ***production planning,*** *planning quality, master data, prediction, machine learning.*

I. INTRODUCTION

Industry 4.0, cyber-physical production systems (CPPS) [2,3] and the Industrial Internet of Things have a significant influence on production planning and control (PPC). Deploying CPPS raises several challenges for industries addressed in [4], in particular with regard to extraction of knowledge from heterogeneous data sources, interoperations with production information systems as well as changeability, adaptability and re-configurability in production management. Compared to traditional production planning based on a static knowledge base, smart factories enable a collection of real time information and share from and between products, machines, processes and operations [5]. The application and exchange of data by the elements of a smart factory leads to an automated and decentralized production, which is an essential characteristic of Industry 4.0 [6,7]. However, there is a need to study how the different solutions – enabled by digitalization – can support PPC and contribute to an increased corporate competiveness [8].
PQ is a commonly used term in PPC, but evidently, it is not clearly defined. A new definition for planning quality will be proposed by the authors and it is investigated how this suggested planning quality can be used in production planning systems.

II. RELATED RESULTS IN THE LITERATURE

Literature in PPC approaches the phrase PQ from different angles, considering several influencing factors. The first aspect focuses on classical logistical targets of PPC such as flow times, due dates, setup costs or product features. Since accurate and high-quality planning data is one of the most important parts of a good production plan, master data management is an important aspect in our discussion as well[9]. These first two approaches do not deal with (planned or unforeseeable) changes or uncertainties of the production environment, which are typical elements of the paradigm shift towards Industry 4.0 [10–12]. Therefore, robustness and resilience consider such disturbances. A production plan is robust if its performance is guaranteed – even when facing events not known at the time of planning [13,14]. Resilience is, in contrast, the ability of a

system to cope with changes of all kinds [15]. Efficient ways of dealing with uncertainty are either applying stochastic, fuzzy models or using adaptive and cooperative approaches [16]. The overall objective of PPC is the creation of reliable production plans, so as their realization on the shop floor should be close to – or ideally the same as – the production plan as originally planned. The deviation between planning and reality on the shop floor increases up to 75% after just 3 days in medium sized mechanical engineering enterprises [1]. It is desirable to have more reliable production plans. Measuring quality of the prediction, as an alternative, may reveal potentials for bridging the gap between planned and actual figures. Besides, the success of a good production plan depends on the process of decision-making itself. In the era of Industry 4.0, the automation of decision-making processes and the level and way of human engagement are essential topics [17,18] as well. It can be concluded that there is no exact definition of PQ. In the paper a novel industry-oriented concept for measuring, evaluating and improving the PQ is going to be developed.

A general truth is that data does not bring any added value on its own, but in practice, domain-specific knowledge and algorithms are needed to extract useful information from heterogeneous and scalable data sources [19]. Simple statistical analysis is often not sufficient as it is time-consuming and oftentimes do not lead to the desired results. Hence, automated data extraction and analytics methods are needed. Together with the rise of data science as one of the most emerging research and application fields today, machine learning (ML) has gained increasingly high attention in the recent past.

At the very beginning of the development of ML, the vast majority of papers were published in journals related to the topic area of Computer Science. However, with increasing demands to computational capabilities and big data analytics, the area is growing, with far-reaching applications in diverse disciplines. Nowadays, many different disciplines use ML algorithms, as was shown in 2012 [20]. The first ML applications in management science can be found e.g. in finance or marketing [21]. In 2009, Choudhary et al. identified that emerging application of ML in PPC has not been systematically explored [22]. However, in the following years, several research papers were published in production management focusing on applying ML for advanced planning and scheduling [23], quality improvement, process monitoring and defect analysis [24]. Yet, researchers did not intensively focus on (sub-)topics relevant for PPC – such as flow time prediction, lot cycle time prediction or lead time prediction, and thus the improvement potentials are not completely identified and used up.

The results of our literature survey determine that the current trend in PPC is to employ ML-based simulation and optimization algorithms. Furthermore, it can be recognized that the focus of purpose in most analysed publications is either on production scheduling (47%) or other applications (33%), while prediction of planning relevant times is rarely (20%) focused as shown in [19].

III. PLANNING QUALITY – DEFINITION AND APPLICATION POSSIBILITIES

Until now, there has been no uniform mathematical definition and no uniform understanding of the term PQ in the scientific literature within the framework of PPC. Therefore, we will briefly explain and define what is meant by the term PQ.

PQ is foreseen as a key indicator for the planner to assess the reliability of the production plan in the planning phase (*time t*) and to continuously improve operational reliability of production plans in forthcoming phases (*time t+n*). In principle, planning quality is high when:

Ideally no deviation, but a deviation in an at least an acceptable range (which can vary depending on different industrial context), between predicted times and actual times (e.g. lead time, set up time, operation time) exists. Reliable dynamic-based prediction models are required to continuously reduce the deviation. Ideally no deviation between planning times and the predicted times exists. A novel data-driven approach is required to generate production plans when times are functions of features, and to evaluate and minimize the gap between planning time and predicted time. The PQI (Planning Quality Index) shall be formulated mathematically in the course of further research to be able to use it as a KPI and for objective decision making.

The application of ML to increase PQ is the main innovation of this paper. The method in which high quality production plans are created lead to the development and evaluation of two possible, but quite different approaches.

Evolutionary Approach:

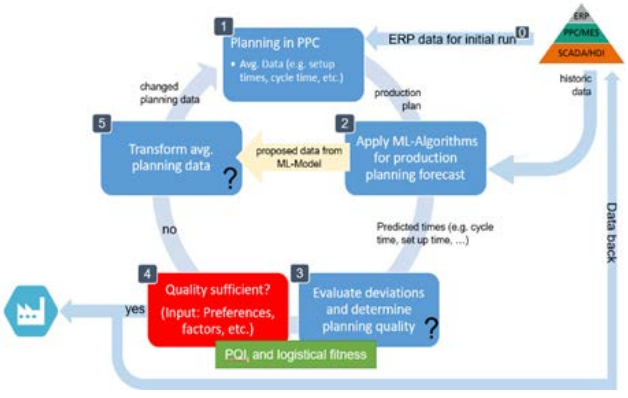


Figure 1. Evolutionary approach

The overall approach consists of six steps, numbered from 0 to 5 while step 1 to 5 run in ongoing loops. It is assumed that in every iteration, the planning quality is increased. This iterative process aims at continuous improvement is the reason why the consortium named the approach

"evolutionary". The planning starts with the generation of an initial production plan, whereby the planning itself is carried out by the conventional planning system, e.g., ERP or MES of the company, and the planning data is coming from the master data of the underlying system (steps 0 and 1). The production plan is the input for the second step. In step 2, ML models, which have been trained by utilizing historical data from the company prior, are employed to predict different time slices such as cycle time, setup time, operation time, etc. In the next step, the deviations between the planned times and the forecasted times, as well as their impact will be evaluated. The results are summarized in a dashboard, showing the logistical fitness as proposed by Lödding et al. [25], and the newly developed PQ index. Depending on the impact of the deviation, the quality will give feedback about the reliability of the production plan. In the fourth step, the planner will decide whether the production plan meets the preferences and the objectives of the company or not. If not, it will be possible to change the planning related master data by using the support system that is capable of recommending the planner some alternative proposals to change the data. The overall decision, and therefore, the responsibility stays with the human (i.e. human-in-the-loop model) [18]). After the refinement and adjustment of the planning data, the planning can be restarted and the cycle starts again.
If the planner is convinced of the key performance indicators (KPIs) in step 4, he/she can activate the plan and the planning related master data that was used to create the production plan, will be transferred to planning systems. The innovative characters of the proposed evolutionary approach can be specified as follows:

- Providing an evaluation method for the assessment of the production plans in Step 3.
- Creating a dashboard to comprehensibly provide feedback to the planner through visualizing the main KPIs. Establishing a recommendation method for proposing planning data that are more suitable.

Function-based Approach:

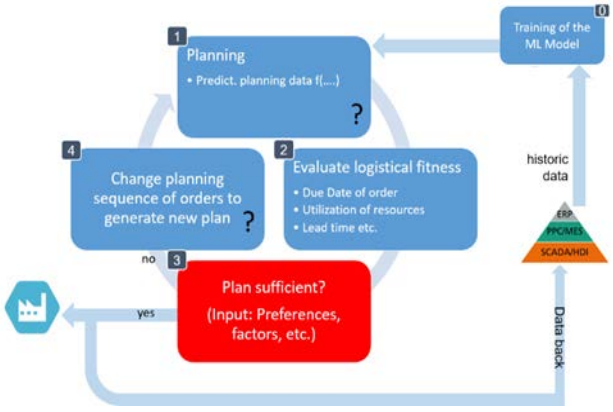


Figure 2. Function-based approach

The second approach differs, as the planning cannot be carried out by the standard/conventional planning system. Whereas in the evolutionary approach the quality of the production plan continuously increases, in the function based-approach, the PQ is always set to high, while the other logistical fitness of the plan will be optimized iteratively as shown in Figure 2.
Step 0 is necessary to generate the prediction models and therefore be done regularly. Based upon these ML-models of the planning related data, a novel planning method generates production plans. The novelty of the planning models lays within the usage of the dynamic set of functions (i.e. an extendable vector of possible functions) for the planning related data and not a single value or a set of values. The general approach looks as follow: orders will be subsequently scheduled, always taking the current information of the production plan as information for the prediction models of the planning related data. Some features for the prediction will not be available at the time of planning, and will therefore these dynamic elements will be predicted with different models. As there is no comparable approach in the field of research, this will be the highest overall innovation within this approach.
After planning the evaluation of the logistical fitness follows. The logistical goal weighting of the company (e.g., avg. lead-time of 6 weeks, timeliness of 87% of all orders, etc.) models the benchmark, a production plan must reach. Since the PQ is already high, only these KPIs must be checked. In Step 3, the planner again decides, whether the plan is acceptable or not. If not, he/she will try another planning sequence. As a support function, the system offers several strategies (based on the order size, due date, costumer priorities etc.) Additionally, the planner can do manual changes. If the sequence is defined he/she starts the planning again. These steps will be repeated until the plan is satisfactory. If so, the plan can be transferred/released to the enterprise resource system (ERP) / manufacturing execution system (MES). In sum, the innovative characters of the proposed function-based approach can be specified as follows:

- Providing a novel planning method for dynamic times.
- Creating a recommendation method to propose changes in the planning sequence.

IV. RESULTS AND DISCUSSIONS

The objective of both approaches is to increase the planning quality in the sense of smaller deviations between planning and subsequent execution. The key is seen in incorporation of uncertainty in the planning phase by carrying out production planning on the basis of dynamic instead of static time values. Only with the use of dynamic time values (e.g. standard times for machining and set-up) the interdependencies of different influencing factors, that occur naturally can be depicted close to reality. These influencing factors can be well known, but most planning system do not offer the possibility to consider them in

planning or scheduling activities. In the worst case these influencing factors remain unknown or at least undetected. An example for the former case could be fluctuations in the machining and set-up times per week and/or shift. This is commonly caused by different skill levels of employees or the actual machine or tools that executes the job. Even though these effects are well known, most companies do not have the resources to levy and document all the effects so that a system would be able to process the information. Still this information is available in the past that and ML is able to quantify the impact to some extent.

In order to be able to represent these dynamics as realistically as possible using an ML algorithm, it is essential to ensure high data quality. In both approaches, the prediction provided by the ML algorithm can be as reliable and valid for the actual production system as the underlying historical data source that is used. For further discussion we want to focus on different data sources for MES data. We distinguish between Plant Data Acquisition (PDA) and Machine Data Acquisition (MDA) data. MDA data can generally probably considered to be more reliable compared to PDA data as the actual machine status is automatically captured. E.g. status like "no spindle rotation" and "no malfunction" indicate that the process has obviously been completed and a time feedback via MDA is done correctly. In the same case with a feedback via PDA there can be either a time delay in the feedback of the machine operator or feedback is missing at all.

Therefore, in the case of PDA feedback, higher fluctuations and less quality training data are to be expected in comparison to MDA data. However, this does not mean that the approach is only applicable for MDA generated data or data that is generated with a simulation model of the production system. Longer observation periods and therefore a larger data source help to increase the reliability. In conclusion, it can stated that the use of reality reflecting data is essential.

To insure that the machine statuses are mapped in production planning, it is furthermore important to consider them as input features of the ML algorithm. Depending on the input data of the machine statuses, the underlying causes should be questioned instead of a direct link between the change of the data and a change of the production plan. Example: Under the assumption that the machining times of the machine increase over time, a variety of causes can be assumed, such as a change in the condition of the machine tool or the production machine (e.g. blunt tool, slow feed). In this case, it seems reasonable to invest resources to adjust the machine condition instead of generally increasing the machining times. It is important to provide the production planner with a decision support system in order to present decision options and their effects on the planning system. It seems conceivable to constantly increase the processing times due to a lack of investment in old machines or to restore the original machine condition. In the second case, the actual machining times should be significantly reduced. A major advantage of the function-oriented approach is the not required time-consuming master data maintenance, but instead the approach generates benefits through the dynamic adjustments of the relevant values. However, it should be noted that the logic of the ML algorithm can only create a production plan on the basis of the current data situation. Missing or defective data usually leads to a reduced PQ. Even if the condition of the machine worsens, immediate repair of the machines and the associated shortening of the processing times is generally not to be expected nor is it appropriate. Permanent monitoring and visualization of the machine status, can alternatively be used to initiate measures to reduce production time. As an example, a maintenance or change of the production parameters shall be mentioned. As an ideal solution, the integration of a logic into the ML-algorithm can be considered, which makes it possible to examine data from different data sources, machines, etc.. The algorithm can assist in interpretation and decision making, or if required to perform appropriate weighting of individual data and states. By weighting, the effects of individual influences on production planning can be adapted and implicit knowledge of planning can be represented to further increase PQ.

Evolutionary Approach:

Since this approach uses the existing production planning system, there are fewer interventions in the current, existing system. This means that the existing systems do not become obsolete and can still be used to generate production plans. The evolutionary approach is meant to be an additional decision support system (DSS) to the actual planning system. Within this additional DSS a comparison is made between the planned times and the predicted times (cycle time, set-up time, operating time, etc.), the probability that the planned and the predicted time is calculated in feedback to the planner is given in form of various KPIs and the PQ index. Furthermore the planner gets feedback why a certain time prediction is different from the planned time. Based on the chosen ML method "driver" for predicted times can be identified and give a note to planner. Since the system is an additional system, the implementation is expected to be easier and quicker. Furthermore, the acceptance by the planners is expected to be higher as the actual planning is still done by the existing system. Transparency and sovereignity of the planner about decision-making is an important factor for the acceptance of employees and proves to be a significant advantage of this approach.

Function-based Approach:

As the production plan is created directly using the time prediction models and these models get updated frequently, the quality is independent from the stored master data. Furthermore, the "dynamic" models are

always updated automatically. The authors expect this approach to need a smaller number iteration cycles. The reason for this lays in the fact, that the system always uses the planning data, that most likely depicts the later execution. The iterations are needed to meet the objectives for the logistical KPIs. Therefore the number of iterations are expected to be similar to the current number of iterations that are needed today for creating a proper plan. However there are several open research questions that need to be answered. Within the function-based approach the planning algorithm pics one order after the other. This means, that the prediction only considers orders that have already been planned. However, subsequent orders have an impact on the features that are used for the prediction e.g. the WIP when, an order arrives at the same workstation and the prior planned order is not finished. It is therefore crucial to derive correlations from planned orders and unplanned orders in the planning process. According to the authors, further research is needed to define a way in which the necessary features for the time prediction can be determined.

The proposed approach is designed to replace the current planning system with a new planning algorithm. The planner has the option of weighting the KPIs (on time delivery, lead-time, stock, etc.) and thus manually adapting the priorities for production planning. Since the PQ index is always high, it is important to check the corresponding KPIs and only implement the production plan when the KPIs deliver sufficient results. A further advantage of this approach is that, in contrast to the Evolutionary Approach, no adjustments of master data have to be carried out, but a collection of the feedback data is necessary to train the ML algorithm.

Evaluation of the approaches:

For the functionality and correct use of both approaches to increase planning quality, it is also important when which approach is applied. Therefore, in addition to further developing the approaches, attention should be paid to creating an evaluation method for the two approaches. This should clearly demonstrate the advantages of each approach as well as the complexity of implementation and ensure that the right approach is always used depending on the application, the industry and other influencing factors in order to achieve optimal results.

Implementation in ERP/MES:

In order to ensure that both approaches are fully functional and can be used separately from the implemented ERP or MES solution, they must made accessible platform-independently. Different variants of ERP and MES systems in the industry do not represent restrictions for the application of the new approaches. Therefore the transfer of historical data for the training of the ML model is a general solution and must guarantee a data transfer at any time, although the training data can consist of individual characteristics depending on the application case. Even if the accuracy of the production plan prediction increases with the amount of used training data, it is important not to select a disproportionately large amount of data. The horizon of the data also has an important influence. The use of long-term data ensures that the results are based on a long history, while the use of short-term data is suitable for representing outliers and random events. The storage of data duplicates in MES and ERP has to be prevented as well.

In each planning run, only the delta from old and new data is to be transferred, which guarantees speed advantages in the data transfer and keeps the transfer duration low. In fact, data transfers are not only possible for order confirmations, but also for partial confirmations via MES or ERP. According to the current status, however, only historical data of a completed order (order confirmation) is used to forecast production plans. Since the time horizon of the data used can influence the result of the prediction, it is also possible to differentiate between two variants of time-based use. In variant 1, the time horizon extends over the entirety of all training data, with all previous training data having an unknown influence on the result. In variant 2, the time horizon can be individually selected according to the principle of ongoing planning, which has advantages, for example, for the more accurate reproduction of short-term events. It is assumed that old data is less valid and can lead to distortions. In the further course of a research project, the advantages and disadvantages of different time horizons on the quality of production planning will be investigated.

An essential basis for a successful introduction and use of these approaches in ERP or MES is an analysis and optimization of the existing process flows. A large part of the benefits that can be realized by the integration can hardly be evaluated in advance using quantitative criteria. The improvement of the internal production planning is an essential benefit of the presented solution, but its valuable support in the entire order processing is difficult to present. The benefit assessment in particular is considered problematic, since only parts of the achievable benefits can be quantitatively assessed in advance, for instance adherence to delivery dates or error avoidance. Another element that needs to be considered are the basic hardware requirements at the factory level that the constant use of ML algorithms in planning entails. A quantification of the required computing power is evaluated in the course of a research project.

V. CONCLUSIONS AND OUTLOOK

During the work, a mixture of the two approaches was discussed. In the first place an evaluation method for the two approaches should be developed. The function-based approach appears to be smarter due to the usage of a new planning method based on historical data creating a production plan immediately, instead of creating a

production plan with master data and then comparing it with forecast data. However, both approaches have to be examined more closely in the further proceeding of the research before an exact assessment can take place.
In fact, production plans are currently created from ERP or MES data. The acceptance level of employees in the planning department depends strongly on the recognisability of deviations between the predicted and the classic (ERP/MES) production plan. This is a clear advantage of the Evolutionary Approach. The confidence of the planner in a production plan, which was created by an unknown planning logic, could initially be low and thus lead to a restraint on the implementation. Therefore, it becomes clear that no matter which approach is chosen, the transparency of the approach must be consistent and comprehensible for the employee. Only if the two approaches are accepted and trusted a correct implementation can be achieved and the planning quality in production planning increased what leads to an optimization of the logistic target values.

REFERENCES

[1] **Schuh, G. (ed.):** Ergebnisbericht des BMF-Verbundprojektes PROSense. Hochauflösende Produktionssteuerung auf Basis kybernetischer Unterstützungssysteme und intelligenter Sensorik. 2015, pp. 7-8.

[2] **Ansari, F., Khobreh, M., Seidenberg, U. and Sihn W.**: A Problem-Solving Ontology for Hu-man-Centered Cyber Physical Production Systems. *CIRP Journal of Manufacturing Science and Technology*, Elsevier, 22C, 2018, pp. 91–106.

[3] **Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W. and Ueda, K.**:. Cyber-physical systems in manufacturing. *CIRP Annals-Manufacturing Technology*, Vol. 65, 2016,pp. 621–41.

[4] **Ansari, F.**: Cyber-Physical Systems. Bericht für Ergebnispapier "Forschung, Entwicklung & Innovation in der Industrie 4.0", *Verein Industrie 4.0 Österreich*, 2018, pp. 26–28

[5] **Bauernhansl, T., Hompel, M. and Vogel-Heuser, B.**: Industrie 4.0 in Produktion, *Automatisierung und Logistik. Anwendung, Technologien, Migration*. 2014

[6] **Yongkui, L. and Xun, X.**: Industry 4.0 and Cloud Manufacturing: A Comparative Analysis. *Journal of Manufacturing Science and Engineering*, No. 139. 2016

[7] **Sanders, A., Elangeswaran, C. and Wulfsberg, J.**: Industry 4.0 implies lean manufacturing. Research activities in industry 4.0 function as enablers for lean manufacturing. *JIEM*, Vol. 9, No. 3, 2016, p. 811.

[8] **Kamble, S. S., Gunasekaran, A. and Gawankar, S. A.**: Sustainable Industry 4.0 framework. A systematic literature review identifying the current trends and future perspectives. *Process Safety and Environmental Protection*, Vol. 117, 2018, pp. 408–25.

[9] **Hees, A. and Reinhart, G.**: Approach for Production Planning in Reconfigurable Manufacturing Systems. *Procedia CIRP*, Vol. 33, 2015, pp. 70–75.

[10] **Schuh, G., Potente, T., Fuchs, S. and Hausberg, C.**: Methodology for the Assessment of Changeability of Production Systems Based on ERP Data. *Procedia CIRP*, Vol. 3, 2012, pp. 412-417.

[11] **Nyhuis, P., Heinen, T., Rimpau, C., Abele, E. and Wörn, A.**: Wandlungsfähige Produktionssysteme. Theoretischer Hintergrund zur Wandlungsfähigkeit von Produktionssystemen. Werkstattstechnik online, Vol. 98, 2008, pp. 85–91.

[12] **Wiendahl, H.-P., ElMaraghy, H. A., Nyhuis, P., Zäh, M., Wiendahl, H.-H., Duffie, N. and Brieke, M.**: Changeable Manufacturing - Classification, Design and Operation. *Annals of the CIRP*, 56/2, 2007, pp. 783–809.

[13] **Tolio, T., Urgo, M. and Váncza, J.**: Robust production control against propagation of disruptions. *CIRP Annals*, Vol. 60, No. 1, 2011, pp. 489–92.

[14] **Gyulai, D., Pfeiffer, A. and Monostori, L.**: Robust production planning and control for multi-stage systems with flexible final assembly lines. *International Journal of Production Research*, Vol. 55, No. 13, 2017, pp. 3657–73.

[15] **Bergmann, U. and Heinicke, M.**: Resilience of Productions Systems by Adapting Temporal or Spatial Organization. *Procedia CIRP*, Vol. 57, 2016, pp. 183–88.

[16] **Gyulai, D., Kádár, B. and Monosotori, L.**: Robust production planning and capacity control for flexible assembly lines. *IFAC-PapersOnLine*, Vol. 48, No. 3, 2015, pp. 2312–17.

[17] **D'Addona, D. M., Bracco, F., Bettoni, A., Nishino, N., Carpanzano, E. and Bruzzone, A. A.**: Adaptive automation and human factors in manufacturing. An experimental assessment for a cognitive approach. *CIRP Annals*, Vol. 67, No. 1, 2018, pp. 455–58.

[18] **Ansari, F., Hold, P. and Sihn, W.**: Human-Centered Cyber Physical Production System: How Does Industry 4.0 Impact on Decision-Making Tasks? *IEEE Technology and Engineer-ing Management Society Conference*, 27 June-1 July 2018.

[19] **Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W. and Monostori, L.**: Lead time prediction using machine learning algorithms. A case study by a semiconductor manufacturer. *Procedia CIRP*, Vol. 72, 2018, pp. 1051–56.

[20] **Domingos, P.**: A few useful things to know about machine learning. Commun. ACM, Vol. 55, No. 10, 2012, p. 78.

[21] **Rainer, C.**: Data Mining as Technique to Generate Planning Rules for Manufacturing Control in a Complex Production System. K. Windt (ed.), Robust Manufacturing Control, *Springer Berlin Heidelberg (Lecture Notes in Production Engineering)*, 2013, pp. 203–214.

[22] **Choudhary, A. K., Harding, J. A. and Tiwari, M. K.**:. Data mining in manufacturing. A review based on the kind of knowledge. *J Intell Manuf*, Vol. 20, No. 5, 2009, pp. 501–21.

[23] **Csáji, B. and Monosotori, L.**: Value Function Based Reinforcement Learning in Changing Markovian Environments. *Journal of Machine Learning Research (JMLR), MIT Press and Microtome Publishing*, Vol. 9, 2008, pp. 1679–709.

[24] **Cheng, Y., Chen, K., Sun, H., Zhang, Y. and Tao, F.**: Data and knowledge mining with big data towards smart production. Journal of Industrial Information Integration, Vol. 9, 2018, pp. 1–13.

[25] **Lödding, H.**: Handbook of Manufacgturing Control. Fundamentals, Description, Configuration. 2013

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Remaining Useful Life Estimation of Industrial Circuit Breakers by Data-Driven Prognostic Algorithms Based on Statistical Similarity and Copula Correlation

Michy Alice[1], Dejan Pejovski[2], Loredana Cristaldi[3]

[1,2,3]*Department of Electronics, Information and Bioengineering (DEIB),Politecnico di Milano, Milan, Italy*
[1]michy.alice@mail.polimi.it
[2]dejan.pejovski@mail.polimi.it
[3]loredana.cristaldi@polimi.it

***Abstract* – Predicting the future behaviour of an item or a complex system based on its past history is the aim of data-driven algorithms. In our paper, we present two algorithms for predicting the Remaining Useful Life (RUL) of industrial circuit breakers (CB) which make use of on-site collected data related to CB's health condition. In the first algorithm, a sub-fleet of CBs is identified by applying the two-sample Kolmogorov-Smirnov Test which relies on statistical similarity between the observations. Once chosen the sub-fleet, the algorithm attempts to exploit correlations between the variation of health condition and sampling time using copulas. The second algorithm models the correlation structure between the time at which a certain degradation level occurs and the item's End of Life (EOL). Both algorithms are used to estimate the item's Remaining Useful Life through the Monte Carlo method. The use of copulas attempts to exploit also the information on the correlation structure in the data in order to obtain a higher accuracy in the estimation.**

Keywords - Remaining Useful Life, Copula correlation, Data-driven prognostics, Health condition variation.

I. INTRODUCTION

In the recent years, condition monitoring (CM) has become available and cost effective for large sets of products; this allowed development of data-driven algorithms which are able to predict the health condition (HC) and the Remaining Useful Life (RUL) of a product. Data-driven algorithms are based on the collected run-to-failure times and do not deal with failure mechanisms, which makes them suitable for analysing systems with complex physical relations between the components. RUL of an item or a system, at a given time instant is defined as the remaining time interval in which it is able to fulfil its required function. Predicting the future behaviour of the product (in terms of HC and RUL) based on the ability to learn from its past history and from the past behaviour of similar products is an essential objective when aiming to reduce maintenance costs and increase the system availability [1].

No health condition is defined as "the extent of degradation or deviation from an expected normal behaviour" [2]. In the case of CB, HC refers to a component profile based on specific parameters which are monitored: degradation of the switching contacts, leakage of the interrupting chamber, SF_6 gas density, temperature of the interrupting chamber, etc.

Relevant scientific papers have been dealing with different aspects of the RUL estimation and maintenance decision making for critical components of the system. Using the fleet concept, in [3] simulation models are presented for maximizing the fleet reliability compared to the target reliability in order to optimize maintenance activities. The importance of fleet size is discussed in [4]: the paper deals with large scale problems reducing them to single items and calculating their reliability individually. An essential need for uncertainty analysis when estimating the RUL is described in [5]. Previous versions of the first algorithm developed in our paper are presented in [6]: the condition monitoring data of a fleet of a product is analysed by statistical methods to extract the usage and degradation profile of the product. This profile is then represented by statistical distributions used later to predict the behaviour of the component.

This paper is structured as follows: in Section 2 a methodology for identifying a suitable sub-fleet is briefly

described. In Section 3 two algorithms for RUL prediction are explained and illustrated on an industrial CBs dataset. Finally, in Section 4 the results are reported and analysed.

II. BASIC CONSIDERATIONS

As proposed in the relevant literature [1], a subset (i.e., sub-fleet) is selected among a given set of products (fleet) that show higher similarity in terms of observed degradation in time, with respect to the item whose RUL estimation is required. The sub-fleet identification is based on a statistical test for grouping those products which present a statistical distribution of their degradation rate similar to the target product. The two-sample Kolmogorov-Smirnov Test (KST) is used in order to decide whether the two samples are drawn from the same continuous statistical distribution or not, i.e., if they belong to the same sub-fleet [1]. The KST uses the maximum absolute difference between the distribution functions of the samples. In general, the test makes use of each individual data point in the samples, independently of their direction and ordering [7]. The confidence level α determines the selectivity of the test.

The information about the past usage of the product is reported as a time series of HC from the initial value of 100% up to 0%. The sampling time and variation of the HC can be calculated as the difference between two subsequent points of the monitored values. Degradation rate is the ratio between the HC variation and sampling time, for i=1, 2,…, n, where n is the number of monitored values [8]. These definitions can be initially used for each CB in the fleet and then combined to obtain vector representations for the whole fleet.

$$\Delta t_i = t_i - t_{i-1}, \tag{1}$$

$$\Delta HC_i = HC_i - HC_{i-1}, \tag{2}$$

$$d_i = \frac{\Delta HC_i}{\Delta t_i}. \tag{3}$$

The correlation structure between sampling time and health condition variation is determined by estimating the underlying copula. A copula is a function which joins (or couples) multivariate distribution functions to their one-dimensional marginal distribution functions, i.e., contains information about the correlation structure between random variables and, in general, can also capture nonlinear relationships [9]. Mathematically speaking, let us assume that F_x and F_y are the cumulative distribution functions of the random variables X and Y. Their joint distribution can be written as [10]:

$$\begin{aligned} &F(x,y) = P(X < x, Y < y) = \\ &P(X < F_x^{-1}(u), Y < F_y^{-1}(v)) \\ &= F(F_x^{-1}(u), F_y^{-1}(v)) = C(u,v) \end{aligned}, \tag{4}$$

where $F^{-1}_x(u) = x$, $F^{-1}_y(v) = y$ and u and v belong to the unit square. In eq. (4), $F_x(x)$ and $F_y(y)$ are the marginal distribution functions of random variables X and Y and C(·) denotes the copula function. The interesting aspect about copulas is that they allow to model separately the marginal distributions and the correlation structure.

III. DESCRIPTION OF THE METHOD

In this paper two algorithms are proposed in order to predict the RUL: the first one is based on KST for sub-fleet identification and copula correlation modelling in order to predict the HC vs time curve, while the second one completely relies on copula modelling in order to estimate the RUL. For both algorithms, the accuracy of the prediction is calculated as the ratio between the sum of correct predictions over the total number of CBs tested as proposed in [1].

A. Algorithm 1

Once chosen the sub-fleet (for which we have a complete or partial HC profile), the prognostic algorithm 1 consists of two main phases:

A). Knowledge extraction from the condition monitoring data of the product as described in section 2: past usage information is extracted by taking for each product the distribution of the sampling time and the distribution of the health condition variation in the related condition monitoring data.

B). Knowledge exploitation: prediction of the future HC profile over time and extracting a confidence interval for the test product RUL. This is done through the steps shown in Fig. 1 (left), assuming that the correlation structure between the random samples of ΔHC and Δt is captured by the copula. The product is not able to fulfil its required function for HC = 0, when it reaches its estimated End of Life (EOF). The algorithm is run for every item in the reference fleet, repeating it a significantly large number of times so that Monte Carlo method can be applied to obtain a 5% confidence interval for the RUL.

Maximum Likelihood Estimation (MLE) is utilized to set the parameters of the copula which fits best the data [11]. In the baseline scenario the independence copula is considered, and the same results are obtained as in the relevant papers [1]. Then other copulas types are used and based on the log-likelihood values, the most suitable one is selected. The final aim is to observe RUL variations obtained by taking into account the dependence between sample time and health condition variation.

The algorithm performance is estimated for a given

observed degradation level (the % of collected data for the test product with respect to its actual lifetime) and different values of α (which defines the desired level of selectivity of the KST) [1].

B. Algorithm 2

Another approach to solve the same problem is proposed in the following section. A set of CM data related to *N*=90 products is considered again. By inspecting the data, a significant pairwise correlation is found between t_i when i = 25, 50, 75 (t_i is the time needed for the product to reach a degradation level equal to i expressed in % [12]) and t_f (the EOL). For instance, the time needed to reach HC=100-25=75% appears to be significantly correlated with the time needed to reach the EOL. This is shown in the scatter plots in Fig. 2 where the correlation is calculated using Spearman's rho (a non-parametric measure of correlation between variables [13]).

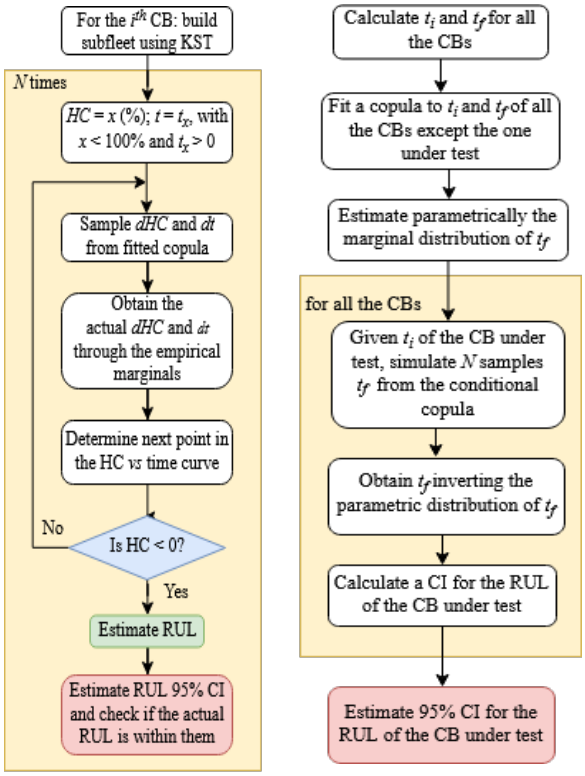


Fig. 1. Flowcharts of algorithm 1 (left) and algorithm 2 (right)

Since from Fig. 2 it becomes clear that some kind of relationship exists, the aim is to predict the RUL (defined as the difference $t_f - t_i$) of a specific CB at a given t_i, by using the information of all the other CBs in the dataset. The algorithm 2 consists of the steps shown in Fig. 1 (right) and it runs for all the *N* CBs in the dataset.

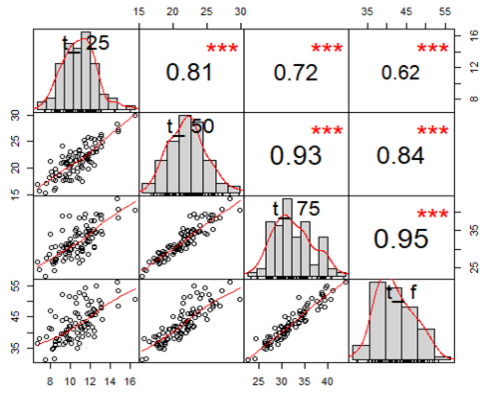


Fig. 2. Pairwise comparison between t_i and t_f

IV. RESULTS AND DISCUSSIONS

In the baseline scenario a naturally assumed correlation is discussed between the sampling time and the variation of health condition. A reasonable expectation is the longer the sampling time, the higher the variation of HC. Since this case is not taken into account in the previous similar work [1], at first it is examined through the analysis of correlation between the sampling time and HC variation from the available dataset. The data, however, indicates that there is almost no correlation pattern between these two variables both when considering a single item (Fig. 3 left), and when considering the entire fleet (Fig. 3 right): indeed the Spearman's rho for the entire fleet is ρ= - 0.025, which indicates that almost no correlation exists. However, at 1% significance level, a bivariate asymptotic independence test based on Kendall's τ (which counts the number of different pairs between two ordered sets and gives the symmetric difference distance [14]) still suggests some possible dependence. Based on the log-likelihood values, a 180° rotated Clayton [10] copula has been selected for modelling the correlation structure.

As it is shown in Fig. 4, some level of correlation and highly visible asymmetry in the upper part exist between the degradation rate and the sampling time.

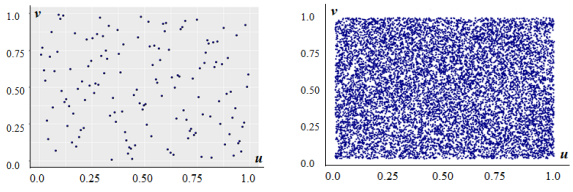


Fig. 3. Pseudo observations of sampling time (u) vs. HC variation (v) for a CB (left) and for the entire fleet (right)

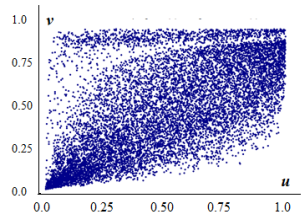


Fig. 4. Pseudo observations of sampling time (u) vs. degradation rate (v) for the fleet

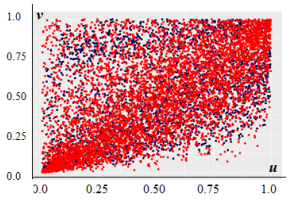


Fig. 5. Pseudo observations of sampling time (u) vs. HC variation (v) for the fleet: red-simulated observations, blue-real observations

The degradation rate and its relationship with the sampling time are considered to be inherent in the model, since they are mathematically linked through Eq. 3. However, the correlation between them is taken into account in a second implementation of algorithm 1 by sampling d (instead of ΔHC) and Δt from a fitted copula (a 180° rotated Tawn [14]): the results are shown in Fig. 5 (in red the simulated data, while in red the observed one) and the algorithm performance is estimated for different levels of α. In both the implementations of algorithm 1 very similar results have been obtained so only the results of the first one are reported in Fig. 6. As it can be seen, the alpha value does not seem to have a notable impact while a significant accuracy (>90%) is reached for observed degradation greater than 50%.

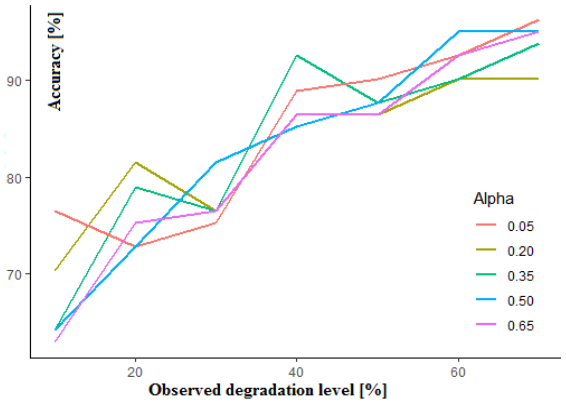


Fig. 6. Algorithm 1 performance N_f= 81

For what it concerns the second algorithm, depending on the (t_i, t_f) pair, Frank and Gaussian copulas [9] are used in the analysis. Independence test at 5% significance level is done prior to fitting each copula through MLE. As it can be seen in Fig. 7, algorithm 2 shows a very high accuracy (>90%) even at low observed degradation level (20%).

A significant advantage of algorithm 1 is the fact that it can be applied to products for which the entire HC time series is not available. On the other hand, algorithm 2 shows higher accuracy even for very low observed degradation levels but requires the complete HC time series for at least some of the products.

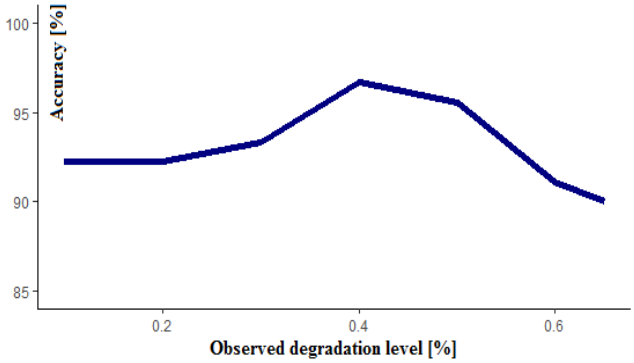


Fig. 7. Algorithm 2 performance N = 90

Another point of comparison between the two models can be the width of their confidence interval (CI). In Fig. 8 and 9 are plotted the CIs of prediction as a percentage of the true RUL value for each algorithm respectively.

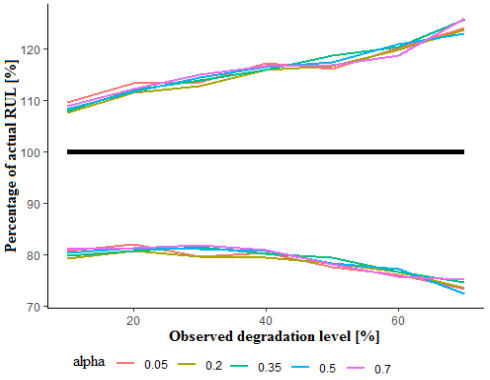


Fig. 8. Actual RUL in case of algorithm 1

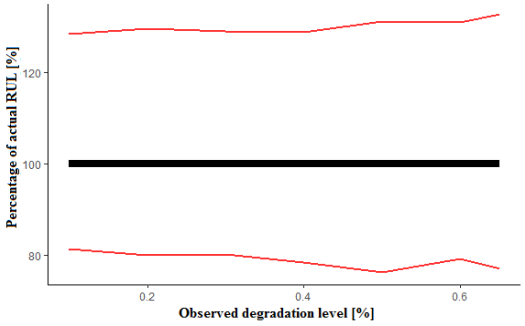


Fig. 9. Actual RUL in case of algorithm 2

The following observations can be made for the first algorithm:

- Regardless of the selectivity value α, the same trend for the confidence intervals is observed;
- As the degradation level increases, the confidence intervals (as a percentage of the actual RUL) increase as well. The overall increase is about 20% (10% on both sides of the CI). This can be explained by the fact that the values to be predicted become smaller and therefore the error weighs more.

For algorithm 2 instead, it can be noted that:

- The confidence intervals do not vary significantly with the observed degradation level;
- The uncertainty on the prediction is, however, larger if compared with the confidence intervals of algorithm 1.

As far as computational complexity is concerned, generally speaking the two algorithms do not pose any particular problems, at least when dealing with limited data such as in this case. Indeed, the most demanding operation is the fitting of the copula. However, this operation is done only once, given the information available on the CBs and then it is updated only when new information become available. To predict the CI of the RUL for a CB, algorithm 2 requires extracting N samples only, while algorithm 1 requires repeating N times the simulation of the entire life of the CB. This means that if the life of a CB is described by m samples, then algorithm 1 needs to sample $m \cdot N$ samples, i.e. it needs m times the number of samples of algorithm 2 to achieve the same purpose.

Overall, a possibility for combined usage of both algorithms seems to exist: algorithm 1 could be used on new products for which historical data is limited by exploiting the fleet and sub-fleet concepts. Algorithm 2 instead can be used as soon as some historical data is collected. Another possible use would be to combine these algorithms through ensemble learning [16] or choose which algorithm to use based on the observed degradation level.

V. CONCLUSIONS AND OUTLOOK

The novelty in the proposed models is the attempt to exploit all the information enclosed in the product HC profile by including the correlation structure in the models. Taking into account the entire HC profiles instead of RUL values only, the models obtain more detailed information, such as Probability of Failure within a predetermined time interval. Considering the appropriate dependences by using the copula approach allows for further exploiting the information contained in the data and improving the performances of the prediction algorithms, particularly evident in the second algorithm. In case of algorithm 1, another advantage is that the sub-fleet is not strictly required to include only products with a known RUL, but also products characterized by a partial HC profile knowledge. A further usage could be explored by combining both the algorithms depending on the data available.

REFERENCES

[1] **Leone, G.; Cristaldi, L.; Turrin, S**.: A data-driven prognostic approach based on statistical similarity: An application to industrial circuit breakers, *Measurement*, No. 108, 2017, pp. 163-170.

[2] **Pecht, M.G**.: *Prognostics and Health Management*, Electronics, John Wiley & Sons, 2008.

[3] **Schneider, K.; Cassady, C. R.:** Fleet Performance Under Selective Maintenance, *Reliability and Maintainability, 2004 Annual Symposium –RAMS*, 2004.

[4] **Yanagi, S.**: An Iteration Method for Reliability Evaluation of a Fleet System, *Journal of Operations Research*, Vol. 43, No. 9, 1992, pp. 885-896.

[5] **Wouters, P.A. A. F.; Schijndel V. A.; Wetzer, J. M.:** Remaining Lifetime Modeling for Power Transformers: Individual Assets and Fleets, *Electrical Insulation Magazine*, IEEE Vol: 27, pp. 45-51.

[6] **Turrin, S.; Subbiah, S.; Leone, G.; Cristaldi, L.:** An Algorithm for Data-Driven Prognostics Based on Statistical Analysis of Condition Monitoring Data on a Fleet Level, *2015 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, Pisa, Italy, 2015.

[7] **Li, M.; Vitanyi, P. M. B.:** Two decades of applied Kolmogorov complexity: in memoriam Andrei Nikolaevich Kolmogorov 1903-87, *Structure in Complexity Theory Third Annual Conference, Washington*, 1988, pp. 80-101.

[8] **Turrin, S.; Subbiah, S.; Leone, G.; Cristaldi, L**.: An algorithm for data-driven prognostics based on statistical analysis of condition monitoring data on a fleet level, *International Instrumentation and Measurement Technology Conference*, 2015, pp. 629-634.

[9] **Peng. W.; Zhang X.; Huang, H-Z.**; A failure rate interaction model for two-component systems based on copula function, *Risk and Reliability*, Vol. 230(3), 2016, pp. 278-284.

[10] **Nelsen, B. R**.; *An Introduction to Copulas*, 2nd Ed., Springer, 2006.

[11] **Jia, X.; Cui, L**.: Reliability research of k-out-of-n: Supply chain system based on copula, *Communications in Statistics – Theory and Methods*, 41:21, 2012, pp. 4023-4033.

[12] **Xi, Z.; Jing, R.; Wang, P.; Hu, C.**; A copula-based sampling method for data driven prognostics, *Reliability Engineering and System Safety*, 132, 2014, pp. 72-82.

[13] **Rebekic, A.; Loncarric, Z.; Petrovic, S.; Maric, S.:** Pearson's or Sperman's correlation coefficient – which to use?, *Poljoprivreda*, 21(2), 2015, 47-54.

[14] **Herve, A**.: The Kendall rank correlation coefficient, *Encyclopedia of Measurement and Statistics*, 2007.

[15] **Eschenburg P**.; *Properties of Extreme-Value Copulas*, PhD thesis, Technical Universituy of Munchen, 2013.

[16] Zhang C., Ma Y. (Eds.), *Ensemble Machine Learning*: *Methods and Applications*. Springer, 2012.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Quality Improvement of Milling Processes Using Machine Learning-Algorithms

Maik Frye[1], Robert H. Schmitt[2]

[1]Fraunhofer Institute for Production Technology IPT,
Steinbachstraße 17, Aachen 52074, Germany,
maik.frye@ipt.fraunhofer.de, +49 241 8904-538
[2]Laboratory for Machine Tools WZL RWTH Aachen University,
Cluster Production Engineering 3A 540, Aachen 52074, Germany,
robert.schmitt@rwth-aachen.de, +49 241 80-20283

***Abstract* – The increasing digitalization and industrial efforts towards artificial intelligence foster the use of Machine Learning (ML)-algorithms in the production environment. Within production, different application areas and use-cases arise for the usage of ML. In this paper, we focus on the implementation of ML-algorithms for a milling process where critical process conditions are predicted. Based on the predicted process conditions, the machining parameters can be adjusted in advance to avoid critical conditions of the process. The avoidance of critical process conditions increases the quality of the products, since quality characteristics such as surface roughness or dimensional deviations can be influenced. To ensure the transferability of the results to other applications, we follow a methodical approach. The results of the ML-models are discussed critically and further steps are derived in order to use ML-models successfully in the future.**

Keywords –Quality Improvement, Predictive Process Control, Machine Learning, Artificial Intelligence, Data Preprocessing, Artificial Neural Networks, Random Forest, Gradient Boosting.

I. INTRODUCTION

In recent years, the application of data-driven models (DDM) strongly increased in the production environment. Main reasons for this trend are higher computing power, easier data acquisition, simpler application of DDM and higher reliability of DDM [1]. Currently, 30 % of German companies use ML for process optimization [2]. According to the study, the application of ML in production will double by 2023. This trend will be further intensified due to the "Artificial Intelligence (AI)" research year proclaimed by the German government, which opens up a wide range of applications areas for ML in production [3]. Application areas range from Predictive Maintenance (PdM) to optimization of routing and scheduling to Predictive Process Control (PPC). Even process as well as product design can be influenced by applying ML-algorithms. In this paper, we focus on PPC and, more specifically, on a milling process of high precision and highly stressed components used in aerospace industry. A major challenge during milling pose the emergence of critical process conditions such as vibrations, high forces and accelerations, which have a negative impact on the component surface quality and, thus, on the product quality. Since the machining parameters such as axis data and spindle speed influence critical process conditions, data is acquired, which is analysed in order to gain insights of the process and, in a second step, to predict critical process conditions by applying ML-algorithms.

The accurate prediction of critical process conditions can be used to adjust machining parameters in advance so that critical process situations do not even occur. Based on the ML-results, machining parameters can be adjusted. The machining parameter adaptation requires a further model, which in practice are heuristics. In this work, the focus is not on heuristics, but rather lies on the evaluation of suitable ML-algorithms that are able to predict critical process parameters, in this case the magnitude of the amplitude of vibrations, accurately.

This work differs from previous IMEKO-publications, which have already focused on ML. So far, results have mainly been published which concentrate on the entire production or individual components [4–6]. In addition, further IMEKO-publications from the past years are related to other fields of application such as electrical engineering [7, 8]. In the present case, reference is made to a single milling process and thus to mechanical engineering. Moreover, this publication can be assigned to the research topic "Data Analytics, artificial intelligence techniques and machine learning for testing, diagnostics and inspection" (T5).

II. RELATED RESULTS IN THE LITERATURE

Related results in the literature refer mainly to the prediction of product and process quality [9]. An overview of existing use-cases shows the great potential of ML algorithms in production. [10, 11]

Krauß et al. (2019) focused on the prediction of product quality based on a process chain consisting of multiple processes. For each process, a classification and regression tree (CART)-algorithm was trained in order to classify whether the corresponding product would be in or out of specification at the end of the process chain. The CART-algorithm was evaluated based on the Matthews correlation coefficient and achieved a performance of MCC = 0.70. [12]

Kim (2018, p. 556) showed that most ML-studies are focusing on the prediction of tool wear, surface roughness as well as process parameter in milling processes. In this context, the most common ML-algorithms are artificial neural networks (ANN), support vector machines (SVM) followed by decision trees. [11]

Krishnakumar et al. (2015) conducted experiments in order to classify the cutting tool condition in milling of titanium alloys. For this purpose, the ML-algorithms decision tree and artificial neural network (ANN) are applied. The results reveal almost identical performance of the decision tree and ANN. Both algorithms predicted the tool condition of 87 samples with a classification accuracy of about ACC = 90 % [13]. Recently, Saadallah et al. (2018) applied ML-algorithms for time series data to predict the stability of a milling process. In addition, simulation data was used to train an ensemble of deep learning (DL)-algorithms to realise a live recommendation system. [9]

Regarding milling and surface finish predictions, Venkata Rao et al. (2014) used a feed forward neural network to predict tool wear, amplitude of work piece vibration and workpiece vibration [14]. Wu et al. (2018) applied random forests to predict tool wear in a milling process for a condition monitoring system. Statistics of cutting forces, vibrations and acoustic emissions where fed to a random forest (RF) of 10,000 trees. [15]

It can be concluded that ML is already used in the production area, even if its use is limited to prototypes and pilots. In the development of pilots, however, the focus is on the optimization of production processes such as milling. However, the data basis is not sufficient to achieve good generalisation results in the vast majority. Moreover, the performance of ML-algorithms are yet not evaluated on the milling of high precision and highly stressed components used in the aerospace industry.

III. DESCRIPTION OF THE METHODOLOGY

As in many cases within production, the methodology utilised in this work is inspired by cross-industry standard process for data mining (CRISP-DM)-framework [16]. Figure 1 shows the consecutive stages of the derived methodology. Exemplarily, the methodology is displayed only with two ML-algorithms. In the first step, data of the milling process is acquired. The next stage comprises the data preprocessing, in which the data is cleaned, synchronized, augmented and features are selected.

Subsequently, ML-algorithms are selected, which includes the derivation of requirements from the data set characteristics and requirements of the milling process. After implementation, the selected ML-algorithms are validated. Finally, all results are evaluated and compared. In the following, each step of the methodology is presented in detail.

A. Process Understanding

During milling, two separate data sets are acquired with different sample frequencies but the same time stamp. The first data set comprises the machine data including the coordinates of the 5-axis milling machine as well as the spindle speed, feed and torque, while the second data set has two attributes consisting of the sound pressure level with corresponding time stamps.

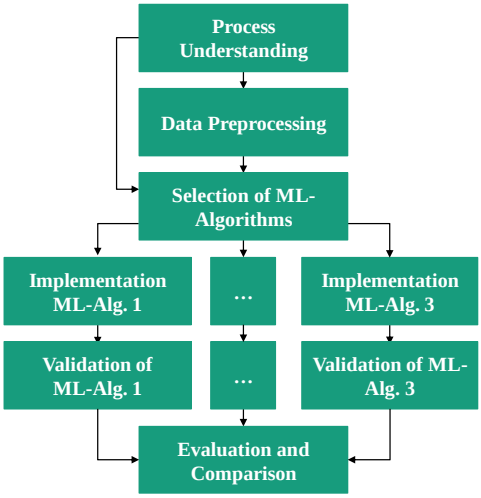


Figure 1. Methodology of ML-implementation

Table 1 summarizes the characteristics of the data set. The number of observations in the machine data set is less than 1 % of the amount of observation of the microphone data. A microphone is mounted within the machine in order to record the sound pressure level (SPL). By recording the pressure level over time, a Fast Fourier Transformation (FFT) can be applied to obtain the amplitude and frequency of the vibration.

B. Data Preprocessing

In a first step, both data sets are integrated into one data base. Because of different sample frequencies, the data sets need to be synchronized. In practice, data synchronization pose a major challenge when synchronizing different data

sources. In addition to the different frequency, a different staring time of data acquisition can also occur. The machine data is assigned to the vibration data with the corresponding time stamp. Since it is assumed that the machine data change linearly, missing values are augmented by using the imputation technique linear regression.

Before selecting appropriate ML-algorithms, features of the data sets are engineered. Firstly, the feed of the spindle is removed since the values remain constant at 100 % all the time. Next, the vibration data is modified through mapping of all data points to the absolute value.

In addition, the signal of the SPL was further engineered. The SPL of the current and the past two instances were used as an input of the ML-algorithm. In addition, the first derivative and second derivative of the SPL of the current and past instance were calculated to determine the change and slope of the signal. Outliers in the data are automatically removed and all features are transformed through standardisation. Since the ML-output serves as input for the heuristic, which triggers the adjustment of process parameter during machining, a prediction horizon H needs to be calculated. The prediction horizon includes the calculation time of the ML-algorithm, the calculation time of the heuristic, the signal processing of the machine as well as the actual physical change of the parameter in the machine. Based on a comprehensive literature research, the prediction horizon H is set as H = 184.255 ms.

Table 1. Data set characteristics.

	Machine data set	Microphone data set
No. of observations	12,875	21,720,837
No. of attributes	8	2
Milling Time [s]	424	424

C. ML-Algorithm Selection

Since all the data is labelled and the goal is to predict the magnitude of amplitudes, we applied supervised learning-algorithms with the regression learning task. The ML-algorithms are selected on the basis of the requirements of the milling process, the properties of the data set, external requirements such as computing time and the successful ML-application in similar use-cases. These approach leads to the selection of RF, gradient boosting and Long Short-Term Memory (LSTM) neural networks.

RF is a tree-based algorithm with many single predictors, whose separate predictions are aggregated to one superior prediction [17]. A RF utilises a group of decision trees that are built separately while including random operations in training. Therefore, Bootstrap Aggregating (bagging) is performed to build the ensemble learners [18]. By applying bagging, training data sets are selected randomly for the respective decision tree. Each bootstrap comprises the same amount of instances such as the original data set [19]. The decision trees are trained using the Classification and Regression Tree (CART)-algorithm. When training is finished, prediction of decision trees are aggregated. [20]

Gradient boosting is also a tree-based algorithm. First, one decision tree is trained based on the training data set. Afterwards, a second decision tree is trained to predict the residuals of the first one. Subsequently, a third decision tree is trained to predict the residuals of the second decision tree. This iterative process can be extended arbitrarily. In the end, all prediction are summed up. [21]

LSTM neural networks are a special form of the Recurrent Neural Network (RNN). RNN are capable of modelling specially sequential data through the inclusion of connections that are pointing backwards to neurons of previous layers or to the input of the same neuron. Thus, timely dependencies are incorporated in the architecture [17]. An issue with RNNs are exploding and, more likely, vanishing gradients [22]. In case of exploding gradients, weights between hidden layers may become very big during backpropagation, which let the ML-algorithm diverge. Through limiting the norm of gradients, the exploding gradients can be mitigated [22]. This technique is called gradient clipping. When gradients approximate to the norm zero, gradients are vanishing, which lead to the fact that the early layers in the network will not contribute to learning, since the gradient tend to become smaller and vanish while propagating backwards during training. [23] As a solution of vanishing gradient problem in RNN, Hochreiter & Schmidhuber suggested the LSTM cell. Instead of a neuron with an activation function such as rectified linear unit (ReLU), a LSTM cell consists of several activation functions and signal flows. [24]

D. ML-Algorithm Implementation

ML-algorithms can be applied using either open-source or proprietary platforms. Generally, proprietary platforms like Matlab are useful for fast model development, however, licences can lead to high operational costs. On the other hand, there are open source platforms such as Python or R, which are not only free of charge but also plays a dominant role in recent ML-developments. Thus, Python including the existing libraries such as scikit-learn for RF as well as gradient boosting and TensorFlow for LSTM neural networks are used. The computing operations were conducted applying a computing cluster that consists of 64 cores and an internal memory of 256 Gigabyte.

Two RF are trained with 20 decision trees and 40 decision trees respectively. The decision trees are fully grown and not pruned. Moreover, two gradient boosting algorithms are trained. The number of predictors are 50 and 100. Studies by Hastie et al. (2017) lead to the setting

of the learning rate of $\eta = 0.1$ and the number of leafs $n_{leaf} = 8$ [25]. Based on the findings of Friedman (2002), a random sub sample size of 40 % is applied for training the predictors. [26]

Two LSTM neural networks are implemented, consisting of two hidden layers, 20 LSTM-cells per layer, a batch size of batch = 32 and the learning rate of $\eta = 0.001$ [27]. In both cases, the gradient optimizer Adam is used. The neural networks are trained over 50 epochs. According to Srivastava et al. (2014), one neural network is regularized based with a dropout rate of $p = 0.5$ as well as the max-norm regularization with a norm boundary for weight vectors of $C = 4$. [28]

E. Validation and Evaluation

The validation of the ML-algorithms are conducted based on forward chaining cross validation. The data set is split into six different folds as indicated in the literature and common in comparable use-cases [29]. The folds of the data set, which are located after the test fold, are ignored in order to allow realistic training scenarios. Since the learning task is a regression, the performance of the ML-algorithms are evaluated based on the root mean square error (RMSE). In the end, results are compared.

IV. RESULTS AND DISCUSSIONS

Table 2 shows all ML-algorithms results ranked by the validation error. The regularized LSTM-neural network has the lowest validation error of RMSE = 0.2920. There are differences in training and validation error. Contrary to what has been proven in the literature, validation errors are not always higher than training errors [30]. For instance, gradient boosting and LSTM neural networks have lower validation than training errors. RFs achieve the lowest training error while having the highest validation error.

Table 2. Results of Ml-algorithms.

	RMSE (validation)	RMSE (training)
LSTM neural net (incl. regularization)	0.2920	0.3890
Gradient boosting (100 predictors)	0.3089	0.4628
Gradient boosting (50 predictors)	0.3144	0.4765
LSTM neural net (excl. regularization)	0.3248	0.3800
Random forest (40 predictors)	0.3269	0.1478
Random forest (20 predictors)	0.3329	0.1549

Although RFs performed worst, peaks in the time series were approximated the best (see Figure 2). Figure 2 shows the comparison between actual amplitudes and the amplitudes predicted by the 40-tree RF exemplarily for all ML-algorithms. In case of RFs, the predictions on the validation set reflect a high validation error and poor generalisation of the trained algorithm. The RF underestimated the real values, which originates from averaging the single predictors. The differences to the real values are much higher for predictions on the validation set. According to Géron (2017, p. 127), a low training error together with a high validation error indicates overfitting [17]. Since the majority of data is actually background noise, the algorithm learns primarily to predict the low amplitudes (93.64 % of the data) during training. RFs of fully-grown trees overfit on noisy data sets.

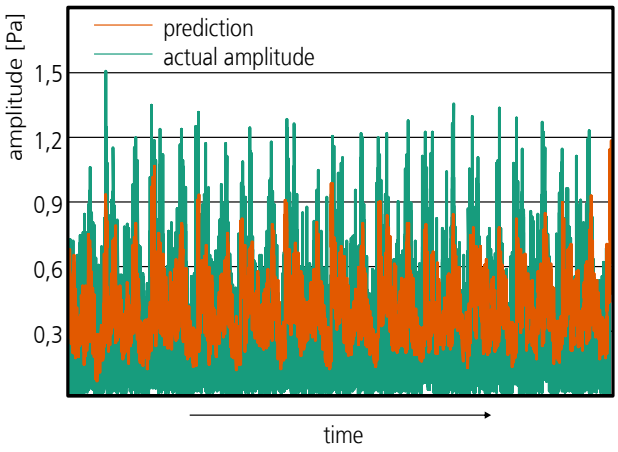


Figure 2: Validation results of RF with 40 predictors

For every ML-algorithm, the learning curves are calculated in order to identify when the ML-algorithm converge. Figure 3 shows the learning curve of the RF with 40 predictors. It can be seen that the training error continues decreasing but tends to start converging up from around 15 decision trees. The same principle can be taken from the validation error, which stays on a level of around RMSE = 0.33 Pa after approximately ten decision trees. In Table 2, it was shown that the RMSE-values for both training and validation of RF with 20 and 40 predictors do not differ significantly through the early convergence. Adding more predictors does not contribute new information to the model.

In addition, the relative importance of the feature are recorded. The features of the machine axes y and a as well as the second derivative of the amplitude of the sound pressure level deliver most information for the RF. Additional time series features contribute around 53 % to the prediction. The training time was $t_{train} = 28.50$ h by utilising 30 processor cores.

The results of the gradient boosting algorithms are characterized by higher training than validation errors (see Table 2), which is not expected from literature [30]. The feature importance and the predictions reveal that both algorithms only consider axis data. The reason is the

limited selection of features, since the training of the predictors is stopped after three splits. Furthermore, the overall vibration is decreasing over time. Based on the chosen validation approach and the fact that the algorithms predicts noise of the data, the algorithm achieve better results during validation than training. The same principle can be observed at LSTM neural networks. Here, the algorithm with the lowest validation error was already reached after the first epoch of training, which means that the models start to overfit directly.

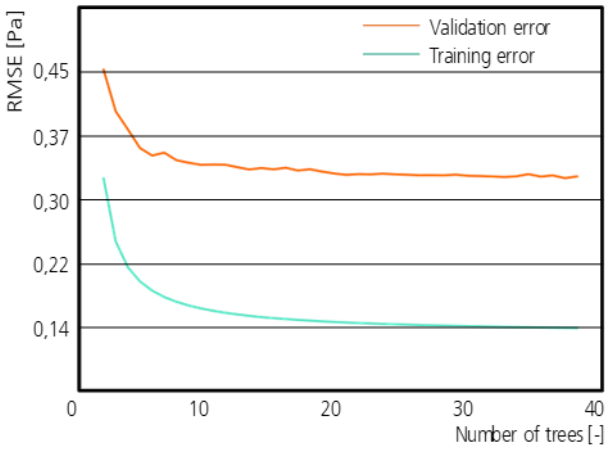


Figure 3: Learning curve of RF with 40 predictors

Based on the results, fields of action can be derived. In the first step, the data quality needs to be enhanced through the elimination of noise. In following works, different noise filter methods will be tested in order to reduce the noise in a proper manner. Three different ensemble and vote-based methods will be implemented such as Ensemble Filter [31], the Cross-Validated Committees Filter [32] and the Iterative-Partitioning Filter [33].

In order to deploy an automated system that controls process parameters, new data has to be prepared automatically and standardised to feed the ML-models. When feeding the ML-model with new data automatically, the data preprocessing need to be automated and a data pipeline needs to be created in order to stream data. Thus, further research need to ensure that the process of streaming and preprocessing data for ML-application for the given use-case is stable.

The focus of further research regarding ML-models lays in methods to adjust hyperparameters. Recently published hyperparameter methods such as BOHB seem to outperform basic approaches such as random search in terms of accuracy and advanced approaches like Bayesian optimization in terms of computational efforts [34]. Applying these methods will lead to higher accuracy of ML-models.

In addition, research has to determine how to change process parameters depending on the outcome of the forecast. Based on the predictions of the ML-model, heuristics can be used to adjust process parameters in order to stay within tolerance after the end of the process.

V. CONCLUSIONS AND OUTLOOK

Critical components in the aerospace industry have to meet high requirements regarding product quality characteristics. In order to improve product quality, critical process conditions during milling processes must be avoided. Parameter optimisation during the milling process serves as an opportunity for avoiding critical process conditions. Following a methodological approach, the initially acquired data from different data sources needed to be preprocessed. For this reason, data synchronization was performed and most relevant features were engineered. Based on the derived requirements, three suitable ML-algorithms were investigated to forecast vibrations during milling. The ML-algorithms were the decision tree-based algorithms RF and gradient boosting as well as LSTM neural network. By applying forward chaining cross validation and by calculating the RMSE-errors, the ML-algorithms were validated and evaluated.

The results show that in the course of this work, a RF approximates high amplitudes better than gradient boosting and LSTM neural networks. The reasons can be traced back to the selected hyperparameters of the chosen algorithms. Aside from the detection of peaks, the regularized LSTM achieved the lowest validation error.

Further improvements can be achieved by investigating ways of handling noisy data, searching for optimal hyperparameter settings, establishing standardised and automated data preprocessing methods and developing models for process parameter adjustments based on predictions of the ML-algorithm. In the context of production processes, in which especially critical components of aeronautical engineering are manufactured and in which ML-algorithms are applied at the same time, certification aspects have to be considered. For this reason, interest groups are already dealing with this topic in order to determine guidelines for certification. This can lead to a successful deployment of ML-algorithms to improve the quality in milling processes.

VI. ACKNOWLEDGMENTS

The European Union under H2020-Widespread 2016/2017 "Teaming for Excellence" sponsors this project to promote further development of data analytics and machine learning with the granted number 739592. Research Executive Agency (REA) supervises the development of the project under the power delegated by the European Commission.

VII Literatur

[1] **Driscoll, M.:** Building data startups: Fast, big, and focused. Low costs and cloud tools are empowering new data startups, 2011

[2] **Geissbauer, R.; Schrauf, S.; Berttram, P.; Cheraghi, F.:**

Digital Factories 2020. Shaping the future of manufacturing, https://www.pwc.de/de/digitale-transformation/digital-factories-2020-shaping-the-future-of-manufacturing.pdf., 2017

[3] **Bundesministerium für Bildung und Forschung**: Wissenschaft im Dialog: Wissenschaftsjahr 2019 - Künstliche Intelligenz, https://www.wissenschaftsjahr.de/2019/.

[4] **Viharos, Z. J.; Csanaki, J.; Nacsa, J.; Edelényi, M.; Péntek, C.; Balázs Kis, K.; Fodor, A.; Csempesz, J.:** Production trend identification and forecast for shop-floor business intelligence. ACTA IMEKO 5, No. 4, 2016, pp. 300-306.

[5] **Capriglione, D.; Carratù, M.; Sommella, P. et al.:** ANN-based IFD in Motorcycle Rear Suspension., https://www.imeko.org/publications/tc10-2017/IMEKO-TC10-2017-002.pdf., 2017.

[6] **Presenti Campagnoni, V.; Ierace, S.; Floreani, F.:** A diagnostic tool for Condition-Based Maintenance of circuit breaker, https://www.imeko.org/publications/tc10-2016/IMEKO-TC10-2016-026.pdf, 2016.

[7] **Polok, B.; Bilski, P.:** Optimization of the neural RBF classifier for the diagnostics of electronic circuit, https://www.imeko.org/publications/tc10-2017/IMEKO-TC10-2017-020.pdf., 2017.

[8] **Bilski, P.:** Unsupervised learning-based hierarchical diagnostics of analog circuit, https://www.imeko.org/publications/tc10-2017/IMEKO-TC10-2017-016.pdf., 2017.

[9] **Saadallah, A., Finkeldey, F., Morik, K., Wiederkehr, P.:** Stability prediction in milling processes using a simulation-based machine learning approach. In: Procedia CIRP, 2018, pp. 1493–1498.

[10] **Köksal, G., Batmaz, I. and Testik, M. C.:** A review of data mining applications for quality improvement in manufacturing industry. In: Expert Systems with Applications, 2011, pp. 13448–13467.

[11] **Kim, D.-H., Kim, T. J. Y., Wang, X., Kim, M., Quan, Y.-J., Oh, J. W., Min, S.-H.:** Smart Machining Process Using Machine Learning: A Review and Perspective on Machining Industry. In: International Journal of Precision Engineering and Manufacturing-Green Technology, 2018, pp. 555–568.

[12] **Krauß; J., Döhler Beck, G. T., Frye, M., Schmitt, R. H.:** Selection and Application of Machine Learning-Algorithms in Production Quality. In: Niggemann, O.; Kühnert, C.; Beyerer, J. (Hrsg.): Machine Learning for Cyber Physical Systems, Springer-Verlag, 2019, pp. 46–57.

[13] **Krishnakumar, P., Rameshkumar, K., Ramachandran, K. I.**: Tool Wear Condition Prediction Using Vibration Signals in High Speed Machining (HSM) of Titanium (Ti-6Al-4V) Alloy. In: Procedia Computer Science, 2015, pp. 270–275.

[14] **Venkata K., Murthy, B., Mohan N.:** Prediction of cutting tool wear, surface roughness and vibration of work piece in boring of aisi 316 steel with artificial neural network. In: Measurement, 2014, pp. 63-70.

[15] **Wu, D., Jennings, C., Terpenny, J., Kumara, S. and Gao, R. X.:** Cloud-based parallel machine learning for tool wear prediction. In: Manufacturing Science and Engineering 2018.

[16] **Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R.:** CRISP-DM. Step-by-step data mining guide, 1999.

[17] **Géron, A.:** Hands-on machine learning with Scikit-Learn and TensorFlow. Concepts, tools, and techniques to build intelligent systems. Sebastopol: O'Reilly Media, 2017.

[18] **Breiman, L.:** Random Forests, In: Machine Learning, Volume 45, 2001, pp. 5–32.

[19] **Pham, D. T.; Afify, A. A.:** Machine-learning techniques and their applications in manufacturing. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 219, No. 5, 2005, pp. 395–412.

[20] **Han, J., Kamber, M., Pei, J.:** Data mining: Concepts and techniques. The Morgan Kaufmann series in data management systems. Amsterdam: Elsevier/Morgan Kaufmann, 2012.

[21] **Friedman, J. H.:** Greedy function approximation: A gradient boosting machine. In: The Annals of Statistics, pp. 1189–1232, 2001

[22] **Goodfellow, I., Bengio, Y. and Courville, A.:** Deep learning. London: MIT Press 2016.

[23] **Pascanu, R.; Mikolov, T.; Bengio, Y.:** On the difficulty of training Recurrent Neural Networks, 2012.

[24] **Hochreiter, S.; Schmidhuber, J.:** Long Short-Term Memory. Neural Computation 9, No. 8, 1997, pp. 1735–1780.

[25] **Hastie, T., Tibshirani, R., Friedman, J. H.:** The elements of statistical learning: Data mining, inference, and prediction. In: Springer series in statistics. New York: Springer 2017.

[26] **Friedman, J. H.:** Stochastic gradient boosting. In: Computational Statistics & Data Analysis - Nonlinear methods and data mining, 2002, pp. 37–378.

[27] **Greff, K., Srivastava, K., Koutnik, J., Steunebrink, R., Schmidhuber, J.:** LSTM: A Search Space Odyssey. In: IEEE transactions on neural networks and learning systems, 2017, pp. 2222–2232.

[28] **Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.:** Dropout a simple way to prevent neural networks from overfitting. In: Journal of Machine Learning Research, 2014, pp. 1929–1958.

[29] **Bergmeir, C.; Costantini, M.; Benítez, J. M.:** On the usefulness of cross-validation for directional forecast evaluation. Computational Statistics & Data Analysis 76, 2014, pp. 132–143.

[30] **Witten, I. H. and Frank, E.:** Data mining: Practical machine learning tools and techniques. Amsterdam and Boston: Morgan Kaufman, 2005.

[31] **Batista, G. E. A. P. A.; Prati, R. C.; Monard, M. C.:** A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD Explorations Newsletter 6 (2004), No. 1, p. 20.

[32] **Kibler, D.; Aha, D. W.:** Learning Representative Exemplars of Concepts: An Initial Case Study. In: Proceedings of the Fourth International Workshop on MACHINE LEARNING, Elsevier, 1987, pp. 24–30.

[33] **Devijver, P. A.:** On the editing rate of the Multiedit algorithm. Pattern Recognition Letters 4, 1986, pp. 9–12.

[34] **Falkner, S.; Klein, A.; Hutter, F.:** BOHB: Robust and Efficient Hyperparameter Optimization at Scale 2018, http://arxiv.org/pdf/1807.01774v1, 2018.

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Integration of Automated Structure Mechanic Analyses into Production Process Simulation

M.B.Schmitz[1], R.Stark[2]

[1]*Fraunhofer IPK, Berlin, Germany, michael.schmitz@ipk.fraunhofer.de, +49 30 39006 125*
[2]*Fraunhofer IPK, Berlin, Germany, rainer.stark@ipk.fraunhofer.de, +49 30 39006 243*

***Abstract* – To produce individual variants of a given product, each variant has to be assessed with respect to their mechanical integrity. This is usually done in the course of product design by assessing all required variants. However, during the lifetime of a product several changes of the existing variants occur or new variants will be created. In this paper a method is described how to integrate and automate the assessment of the mechanical integrity of the product variants directly into the production process by exploiting modern computer methods and automated set-up of state-of-the-art CAE methods. We show this on a demonstrator of a smart factory, that produces disks which can be customized to a large extend. The disks undergo a milling process, which alters the structural properties of the disks. Generally, it cannot be guaranteed upfront that a variant that can be produced is also strong enough to withstand the loads in the application. Therefore, the integrated FEM analyses are computing and assessing the load cases deduced from the use cases of the disks and the loads that occur during milling and handling in the factory. The assessment of the results is done automatically and if positive the production process is started.**

***Keywords* – Finite Element Analysis, In-situ assessment, production process.**

I. INTRODUCTION

In the product creation process, design, verification and validation of the base product and its variants are usually objectives of the R&D department. The development process is often based on product requirements and takes specifications of all required variants into account. Here, we describe a change in the process that allows for higher development speed since only the basic variants of a product need to be designed upfront. All other variants, as fa as their mechanical integrity is of concern, will be created on demand as part of the production process. The new process will be implemented into the production process of the demonstrator of a smart factory at Fraunhofer IPK in Berlin, that produces customized disks (so-called coasters). In this process a freely designable pattern will be engraved onto the raw disk. For the purpose of this study we assume that not all possible producible patterns are mechanically strong enough to be acceptable for a customer viable product. Consequently, after checking for manufacturability i.e. after creating the control file for the milling process a Finite Element analysis of the new disk geometry will be performed to assure the mechanical integrity of this particular disk variant.

II. STATE-OF-THE-ART

Today's product design processes are often following a sequential approach. Starting from top-level design criteria the product architecture and requirements are defined and the design process will be conducted. The V-Model (Figure 1) illustrates this process with the necessary testing and verification steps.

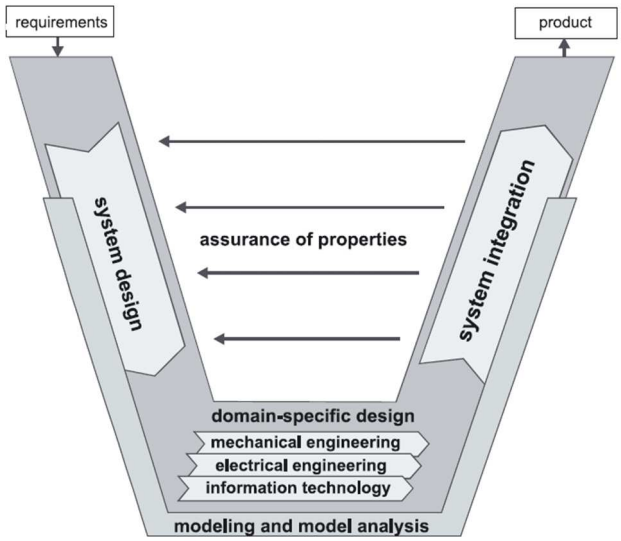


Figure 1: The V-Model according to guideline VDI 2206 [1]

Today, agile methods foster continuous development in smaller steps with integrated testing cycles and enable continuous release cycles of ever changing products [7]. With increasing duration of the overall product life cycle more and more variants will be created in shorter time cycles. Nevertheless, often the technical validity of new variants of an existing product will be done by dedicated teams, sometimes in smaller organisation by the R&D team. This is usually argued by dedicated skills or tools, that have been used during the initial design process. Consequently, the variant or product life cycle management still relies on well distinguished disciplines that act on well-defined boundary conditions e.g. product development requires input from product management or production planning requires input from development.

Stark et. al. [2] described concepts to solve deficits in product development processes for mechatronic products by using data models. In this study, we integrate the structural mechanic assessment of a new variant immediately as part of the production process in a smart factory. Consequently, the variants will be defined and assessed automatically as part of the production process. This becomes possible by modern computer power and sufficient automatization and standardisation of the underlying computational methods, thereby fostering faster turnaround time to create variants.

To assess production, *OEE* (overall equipment efficiency) is often used [9] despite its major drawback of being too general. *OEE* is defined as $OEE = AV \cdot PR \cdot QR$ with AV availability (in %), PR productivity (in %) and QR quality rate (in %). Nevertheless, it provides a first indication of the performance of a given production line [8]. Today, plant and process simulations do not include aspects like in-situ structural mechanics assessment of products during production or the assessment of new possible variants, as addressed in this study. However, including this capability into the production process simulation should affect the *OEE* immediately by increasing all three factors of the *OEE*. The quality rate can be increased by reducing waste due to producing parts that do not meet the required use-cases or by exposing variants to excessive loads in production. The performance rate can be increased by scheduling the production such, that with altered tools only those variants will be produced that can bear these. And the availability can be increased by reducing commissioning time due to proper scheduling, if for the set-up of the production line is still sufficient for some of the variants.

III. Description Of The Method

For the purpose of this study, the SIEMENS NX tool suite is used to integrate structural mechanics calculation and the assessment of the results into the production process of the demonstrator of a Smart Factory at Fraunhofer IPK [3], [4]. The factory and its digital twin represent a one-piece-flow fabrication line, that produces engraved disks (Figure 2). This allows for a broad variation from mass production down to lot size 1. The pattern can be freely customized via a user interface at the factory. The design will then be engraved by a milling process on the disk(s).

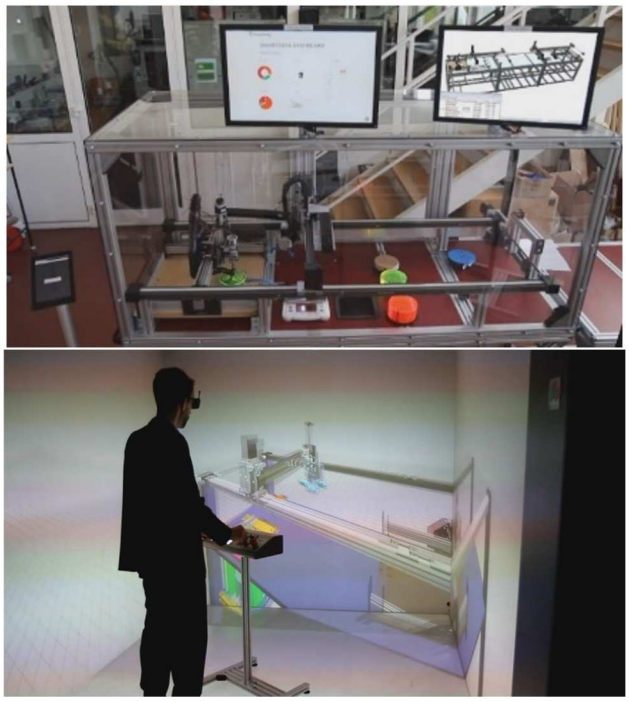


Figure 2: Smart Factory (top) at Fraunhofer IPK and visualisation of its digital twin (bottom).

After defining the design of the disks, the CAD variant is created automatically and the G-code for the milling process is computed. In case this step is successfully completed, the FEM model will be created form the CAD file. Here, it is important to use the same best-practises as they were applied during the design process of the base products in order to ensure reliable and repeatable results. In addition, all changes made to the FEM pre-processing in the future e.g. by exchanging FEM software or using different finite element types or mesh strategies, must be verified and fed back to the design process. Ref. [6] gives good suggestions how to develop and verify such best-practise methods for engineering applications.

The disks are exposed to three types of load cases, illustrated in Figure 3.

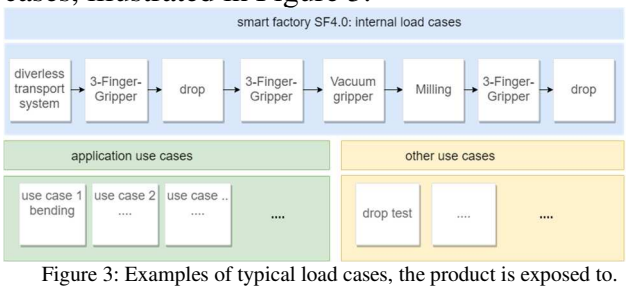


Figure 3: Examples of typical load cases, the product is exposed to.

Firstly, the loads applied during the use of the disk for its intended purpose. Since this demonstration case is fairly simple, the load case is of equal complexity: the use case is defined as a bending load as it may occur, when a disk is bended on a table as it is typical for such coasters (Figure 4). Such use-cases are often defined upfront, the product is verified against it and often the requalification of the product in production also reflect these use cases. If a different use case is desired, the validation cycle must be conducted again. With automated and validated FEM it becomes possible to assess new load cases simply by changing the FEM pre-processing as part of the production line set-up. Figure 4 shows a simulation set-up of a virtual test rig, that may be the digital counterpart of a real physical test rig. The test object is supported by the lower left clamp and a round support underneath the disk. The load is applied via a moving bar on the top right with specified force. The material of the provisions is steel.

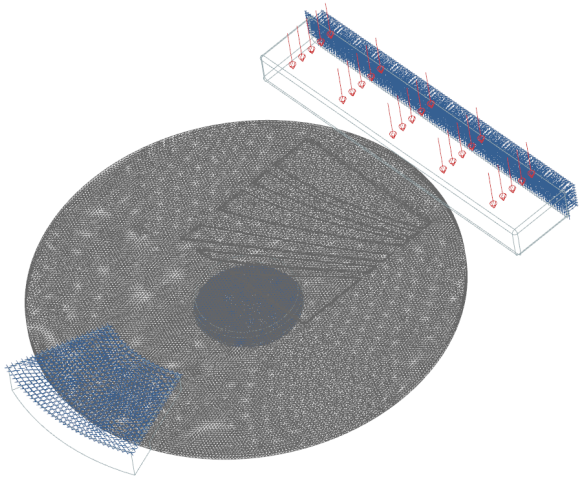


Figure 4: Mesh and applied loads of a disk variant with a typical pattern to be engraved. Load applied by a moving bar and disk supported on top and bottom by a fixed sector provision and a round support, respectively.

The second type of loads applied to the product are those loads that occur during the production process itself. In the Smart Factory, the disks are exposed to the following loads: the disks are transported by a driverless system to their storage place. A single blank disk is picked-up by a mechanical 3-finger gripper and dropped onto an electronic scale. In case the disk mass is

within the correct range it is picked-up again and transported to the milling process. Here, it is fixed by a vacuum device during milling thereafter picked up again and finally dropped to its end station.

These load types are different in character and their magnitude may change over lifetime. E.g. the gripper force or the suction pressure of the fixture may change due to wear. Since these factors are measurable by sensors, the actual values can be used to perform the structural mechanic calculation. The order of load cases to which the disks are exposed is of importance as well. This becomes important when changing the order of the process steps itself or when exchanging e.g. one fixture mechanisms. Consequently, the FEM assessment procedure needs to be adopt accordingly.

The third type are loads that may occur by external events like dropping a package during transportation. For the purpose of this study, we exclude these load cases because they can be treated similar to the first type of load cases.

It is desirable to have fast response time of the results the FEM calculations. This allows for just-in-time production of lot-size 1 variants. However, it is thought that under realistic conditions variable quantities are more important. Therefore, a proper planning of the production, based on respective process simulations allow for some calculation and response time of this load calculations. Nevertheless, fast turnaround time of the computation process can be get by leveraging cloud-based computation. In this study, the actual FEM calculations were performed outside of the smart factory using the network of the Fraunhofer Simulation Alliance.

As described above, the automated structural assessment procedure reflects (external) use cases and covers internal loads created by the production process itself. Since the parameters of latter can always be deduced from the actual status of the factory (e.g. from sensor data) the calculation and assessment can be done with a reduced amount of safety factors. Every set of calculation becomes unique and should be added to the data set of the digital twin of the product [5].

The data set for the structural mechanic assessment is therefore unique for each variant and if desired for every individual part. The results of the FEM may be either positive, meaning the design passes the structural mechanic criteria and the factory starts producing the disks or negative. In this case the user wants to use the results to modify the design of reject it depending on the particular results. Therefore, the FEM data are made available to be used more detailed investigation.

A negative assessment does not necessarily mean that this variant will not be produced. In case the internal loads due to the production process are higher than those of the actual use cases, the factory setup may be changed directly in order to avoid pre-damages on the product. Alternatively, a load reduction on the disks may be possible simply by exploiting the natural degradation of the factory components towards e.g. lower clamping forces by using smart planning processes. In these scenarios, more robust variants will be produced earlier to the more delicate variants. A sufficient process simulation should then include the results and methods of the FEM assessments as well.

In this study we focus on one internal load cases and the particular use case illustrated in Figure 4.

IV. Results And Discussions

For this study, the gripping of the disk after milling and the bending (Figure 4) have been selected as internal load case and as application case, respectively. Also, it is assumed that the calculation set-up (i.e. grid size and type, software, etc.) is taken firstly from best practice from development and product qualification. Here, we used PPM as disk material and ABS as gripper material. The disk mesh has 234000 elements. Siemens NX has been used to perform the Finite Element computation and to control the entire automated calculation.

In the first step, the geometry of the variant is processed and handed over to the preprocessor of

NX. The inner load case is then set up and computed. In this study, the three-finger ABS gripper is modelled as three blocks that each apply a specified force of 43.33N in radial direction, the PMMA disk. (Fig. 5). The values of the radial forces are taken from the gripper specifications.

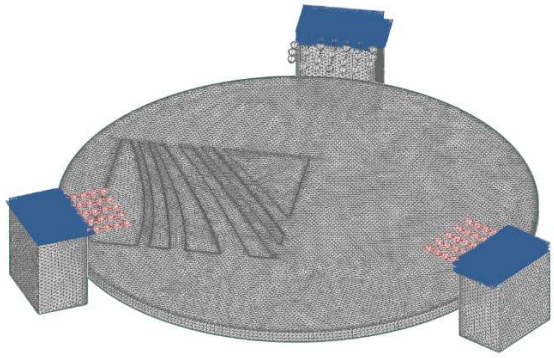


Figure 5: Disk and three-finger gripper as modelled in the FE-calculation. Gripper fingers apply force radially to the disk.

The rule-based automatic pre-processing results in an input file for the actual FEM solver. This input file can be sent to any computer that has a solver license and it powerful enough to perform the computation. A convenient way is to sent it to a cloud service provider to perform the computation and transfer just the acceptance criteria back to the factory. This minimized network load and circumvents the need for hosting an own software.

As assessment criteria the maximum stresses in the all components involved, i.e. the three gripper fingers and the disk itself are used. This allows for assessing the disk for the processing in the factory and also the actual status of the gripper since e.g. material ageing models can be included to the FE-model of the gripper in order to allow for life expectancy assessment and alike of this plant component. Clearly, the entire plant could be simulated by including more and more components and their behavior to the Finite-Element model. In a later study it is planned to elaborate more on this topic to assess efforts vs. benefit of such attempts.

At the moment, the assessment is based on comparing the computed maximum stress with the maximum allowable value for each component involved in the set-up. E.g. in case the maximum calculated stress value of the disk exposed to the loads of the external use-case exceeds the allowable value (e.g. obtained in experiments during the product development or simply taken from a material database) this variant will be rejected as not suited and no production will be granted.

Also, if the computation of the internal load cases yields negative results for one or more factory components, e.g. the gripper is not able to carry the disk due to its material shortcomings, production for this particular variant will not be started, however the negative feedback will be given to schedule maintenance. Figure 6 shows the resulting von Mises stresses of the configuration. The maximum stress value of about 3.1 MPa occur at the gripper finger close to the engravement, however they are very well below of the critical limits of either the disk (73 MPa) or the gripper (35 MPa).

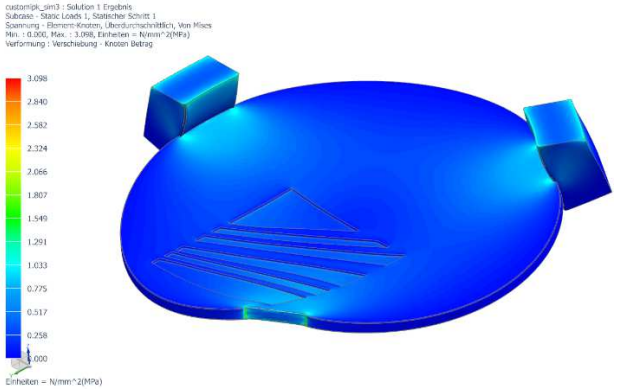


Figure 6: Von Mises stresses of the clamped disk under load. Note: third gripper finger not show in this post-processing picture and displacement five times enhanced for illustration.

Figure 7 shows the results of the application use case as it occurs in the virtual test rig set-up discussed in Figure 4. The maximum stress value is computed as 37.6 MPa, which is of the order of magnitude of the said max. limit for this disk material. The control script automatically picks the maximum value and compares it with the allowable value and grants or rejects production. Here, production will be granted. In case of rejection, the results can be studies further and serve as input for a feedback-to-design process for future product developments. The same is true for changing load cases either due to new

demands on the product yielding new test procedures or due to new production methods yielding to new load cases during production of the disks.

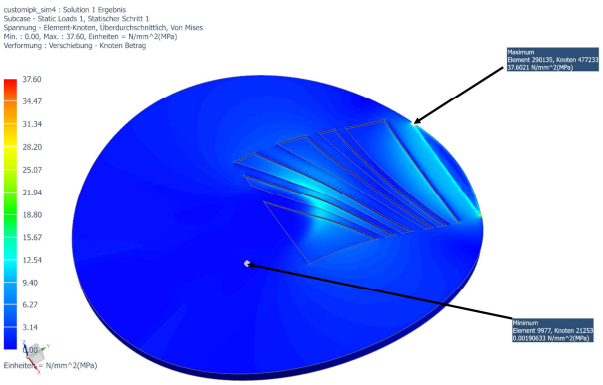


Figure 7: Von Mises stresses of the bended disk under application loads. Note: the provisions of the virtual test-rig are not shown and the displacement is five time enhanced for illustration.

Both categories can be assessed beforehand and the learnings can be fed back to design of the product and/or the production plant.

V. Conclusions And Outlook

The automated set-up of FEM calculations for product variants allows an assessment of the mechanical integrity during the variant production. In order to do so, the best-practice methods for the FEM calculation have to be applied to the variant calculation, too. The pre- and post-processing is set-up, controlled and assessed automatically while the calculation itself may be performed externally in e.g. cloud environment. After the positive assessment of both use case and internal machine loads onto the product, the production is started.

A negative assessment does not necessarily mean that this variant will not be produced. In case the internal loads due to the production process are higher than those of the actual use cases, the factory set-up may be changed in order to avoid pre-damages on the product.

In the next steps, the numerical results and the FEM setup will be part of the digital twin of the individual product. Furthermore, an individual treatment of the product and machine deviations will be taken into account. Also, possible next steps are the enhance the speed of the computation further.

In the future we will use randomly created customization pattern in addition to historic results to train a neural network and use those answers instead of direct FEM calculations.

VI. Acknowledgments

The authors would like to thank Lennert Pannitz for conducting the FEM analyses and integrating it into the workflow of the smart factory.

References

[1] **VDI Verein Deutscher Ingenieure e.V., VDI 2206:** Entwicklungsmethodik für mechatronische Systeme, 2004.

[2] **Stark, R., Beier, G., Figge, A., Wöhler, T.:** Cross-Domain Dependency Modelling - How to achieve consistent System Models with Tool Support. In: Proceedings EuSEC 2010 - European Systems Engineering Conference (CD-ROM), Stockholm, Sweden, 23.-26.05.2010

[3] **Damerau, T.; Vorsatz, T.**: Demozelle Smarte Fabrik 4.0, *FUTUR Ausgabe 2, 2016.*

[4] **ProSTEP iViP Association**: Production Lifecycle Information Management; White Paper, 2019

[5] **Stark, R., Damerau, T.:** Digital Twin. In: The International Academy for Production Engineering, Chatti S., Laperrière L., Reinhart G., Tolio, T. (eds) CIRP Encyclopedia of Production Engineering. Springer, Berlin, Heidelberg, in print

[6] **Hellen, T. K.; Becker, A.A.**: Finite Element Analysis for Engineers - A Primer, NAFEMS, Hardback 438 pages, Dec. 2013

[7] **Scaled Agile, Inc.:** SAFe 4.6 Introduction, White Paper, November 2018

[8] **Fuchs, F.:** Key Performance Indicators in Automated Production, Integrationsaspekte der Simulation: Technik, Organisation und Personal; Gert Zülch & Patricia Stock (Hrsg.); Karlsruhe, KIT Scientific Publishing 2010

[9] **Koschnitzke, T.**: Kontinuierliche Verbesserung mit Total Productive Management. Ziele und Kennzahlen für Verbesserungsprogramme in der Produktion. Hamburg: Diplomica Verlag GmbH, 2008

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

A Novel Condition Monitoring Methodology Based on Neural Network of Pump-Turbines with Extended Operating Range

Weiqiang Zhao[1], Eduard Egusquiza[1], Carme Valero[1], Mònica Egusquiza[1], David Valentín[1], Alexandre, Presas[1]

Center for Industrial Diagnostics and Fluid Dynamics (CDIF), Polytechnic University of Catalonia (UPC), Barcelona., Spain, weiqiang.zhao@upc.edu, +34 657080743

***Abstract* – Due to the entrance of new renewable energies, water-storage energy has to be regulated more frequently to keep the stability of power grid. Consequently, pump-turbines have to work under off-design conditions more than before, which will cause more damage and decrease their useful life. Advanced monitoring methodologies that can balance the degradation of machine and revenues to the power plant has been required. To develop an innovative condition monitoring approach, vibration data was collected from different components of a pump-turbine which is running in an extended operating range. The consequences of operating range extension on the vibration of the pump-turbine have been studied by analysing the vibration signatures. The changing rule of the vibration behavior of the machine with the operating parameters has been obtained. An artificial neural network based model has been applied to build an autoregressive normal behavior model. The results indicated that the normal behavior model based on multi-layer neural net has the ability to predict the vibration characteristics of the machine in different operating conditions. This monitoring method can be adapted to the similar type of hydraulic turbine units.**

Keywords – Condition monitoring, Pump-turbine, Neural networks, Normal behaviour models.

I. INTRODUCTION

Due to their environment friendly and sustainable characteristics, new renewable energies (NREs) such as wind power, solar power, tidal power and so on are more and more popular. With the scenario of the entrance of NREs, a kind of energy system that can response to the demand of the grid instantly is required. A reversible pump-turbine (RPT) is a high-performance turbine which can switch itself between turbine mode and pump mode by reversing the rotation direction of the runner in a short time[1]. At present, pump-storage is the only system that can store huge amounts of energy[2].

Comparing with the other types of turbines such as Francis and Kaplan turbine, a RPT has a higher rotation speed and suffers higher pressure pulsation due to its special design characteristics. In addition, due to the new demand of energy grid brought by the massive entrance of NREs, RPTs have to increase their start/stop cycles dramatically as well as operate at extended operating condition rather than the best efficiency points (BEPs). In an extreme off-design region, the components of the machine are subjected to strong excitation forces from turbulence, rotor-stator interaction (RSI), vortex, etc. The excessive turbulence and pressure pulsation will generate great vibration and stress on the machine and cause damage[3–5]. When the synchronous precession frequency coincides with one natural frequency of the hydraulic circuit, a hydraulic resonance is induced. The hydraulic resonance will cause hydraulic instability and power swing, endangering the stability of the power grid[6–8]. Lots of cases of damage have been reported due to this new scenario[9,10]. On the other hand, the increased regulation capacity is always translated into a large increase of revenues. Therefore, it's urgent and necessary to develop advanced monitoring methodology for the maintenance on the machine and maximizing revenues to the power plant.

Some artificial intelligent (AI) techniques have been applied on condition monitoring and diagnosis. M. Schlechtingen has applied neural network (NN) on the condition monitoring of wind turbine and detection of the fault[11]. M. Garcia has built an intelligent system based on neural network for estimating the health condition of wind turbine units[12]. AI techniques have also been tried to apply on the condition monitoring and fault diagnosis on hydraulic machines. G.Succi has used NNs to distinguish different flow rates and mechanical conditions of helicopter pump based on experimental data[13]. R. Saeed has used NNs and adaptive neuro-fuzzy inference systems (ANFISs) to predict the operating conditions and crack length of a Francis turbine based on numerical simulation results[14]. Artificial neural works (ANNs)

have been widely applied on condition monitoring on different types of machines while it has rarely been applied on the monitoring and diagnosis on PTs, let along the monitoring methodology based on on-site measurement.
In this paper, vibration signals were collected from a prototype pump-turbine. Based on these vibration data, a normal behaviour model is built by means of artificial neural networks due to their ability to model dynamic non-linear industrial processes. The model is able to predict on-line the normal behavior (or reference behavior) expected for each pump-turbine component according to its current operating condition. Based on the normal behaviour model, a predictive operating as well as maintenance scheme can be performed rather than only according to a general guideline. The advanced monitoring system of PTs will guarantee their long life and large benefits.

II. SYSTEM DESCRIPTION

The researched pump turbine unit is a vertical shaft machine with a maximum capacity of 85 MW. It has 7 blades in the runner and 16 vanes in the distributor. The rotating speed is 500 rpm.

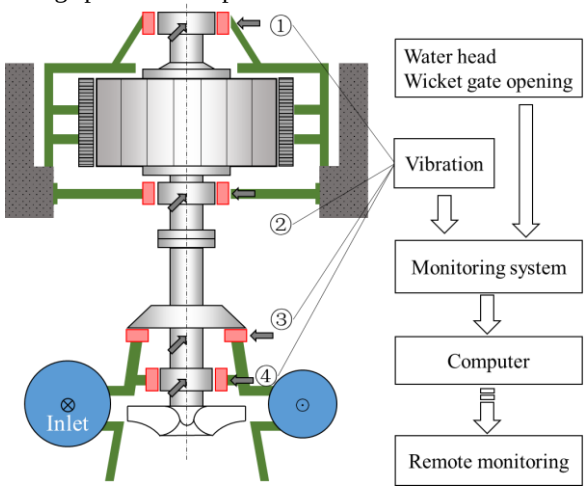


Figure 1 Sketch map of the pump turbine and the monitoring system.

As shown in Figure 1, the accelerometers are mounted on upper generator bearing, lower generator bearing and the turbine bearing. All of the signals generated from the accelerometers as well as parameters related to operation conditions including head level and guide vane opening (GVO) are stored by the monitoring system installed in the unit. Via internet connection, it's possible for a computer that has also installed the monitoring software to remote control or diagnosis.
Over 30 days, 614 operation conditions with different working parameters were acquired and the PT operation and vibration data is used as the research database. During this period of time, the head level varies up and down between 310 m and 340 m while GVO between 20% and 100% (Figure 2). Except head, it is important for us to notice that the output in turbine mode is changing continually in order to match the demand of the electric grid. Being different from the continually changed head level, GVO change dramatically and frequently in turbine mode in order to adjust the output in time. Whereas in pump mode, the GVO degree changes little (from 65.4° minimum to 70.2° maximum) comparing with water head.

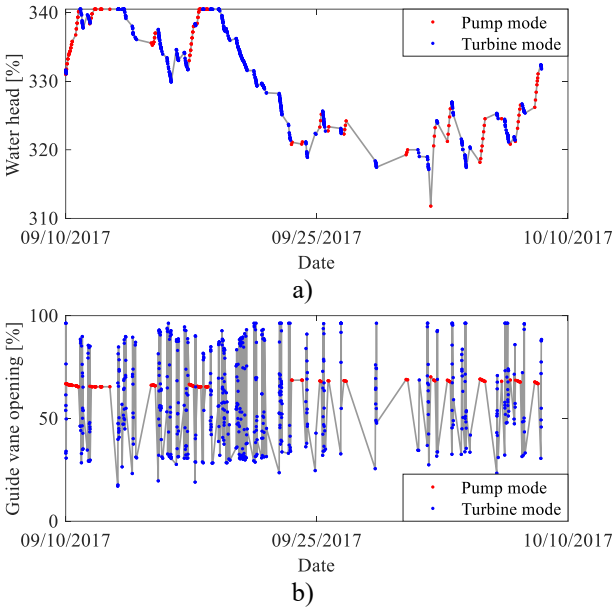


Figure 2 Variation of head level in pumping operation (denoted in red) and in turbine operation (denoted in blue).

III. VIBRATION ANALYSIS

It is well known that the behavior of the structure depends on both the excitation forces and the structure response. Once a defection or damage appears on one of the components of the machine, the excitation forces or/and structure response will change, which may induce the change on vibration. Further inspection can be performed when an abnormal vibration value is detected from the monitoring system. However, the traditional monitoring is only based on the revolution of the vibration level with time, without considering the operating conditions in which they are working. For a PT which is running in an extended range, the type and intensity of excitation forces change dramatically. The structure response (natural frequencies, modal shapes, damping) might even also change with the boundary condition. All of these can be transferred into vibration in the machine behavior. Therefore, it is so hard to predict the machine behavior under off-design operating conditions purely by numerical simulation or model test that it is necessary to perform measurement on a prototype machine to get the real data.
Within deep-part load or over load conditions, the vibration level may reach alarm or even damage level whereas without any damage happens on it. Moreover, the monitoring on vibration overall level is only applied for

protection. The damage or defection happens on one part of the machine without having any great influence on the overall level can only be detected from the spectrum. Therefore, it its necessary to build a monitoring system including not only the overall level but also the details of the frequency components for detecting the faults hidden behind the overall level.

Basically, the excitations of hydraulic turbine units are mainly come from hydraulic, mechanical and electromagnetic origin. Mechanical origin excitations are generated by any system with masses in rotation around an axis and are widely existing in rotation machines. In Figure 3 the peak in f_n=8.33Hz is caused by the unbalance of the rotation components.

For PTs, the most important hydraulic excitation is the interaction between the rotating runner blades and the stationary guide vanes, which is known as RSI. The harmonic frequencies excited by RSI are determined as:

$$f_b = n \cdot z_b \cdot f_f \tag{1}$$

Where n is the order of harmonic, z_b is the number of blades, f_f is the rotating frequency. So that in our case f_b=1×7×500/60=58.33Hz and its twice harmonic $2f_b$=1×7×500/60=116.67Hz. Being one of the most important excitation force in PTs, it may lead to fatigue damage in the runner. In Figure 3 the peaks with f_b and $2f_b$ are caused by RSI (a side band of $2f_b$ with a frequency of 116.67+8.33=125Hz has been noticed and it will not influence the amplitude of $2f_b$).

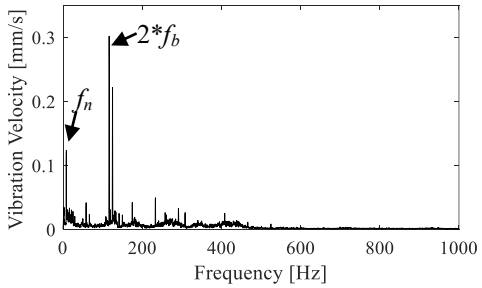


Figure 3 Spectral vibration signature in turbine bearing

Once a mechanical defect or damage appears on any component of PTs, frequency components may have different change in amplitude respectively. It is a complicated work for choosing appropriate indicators from the spectrum for monitoring. In our research, overall level (A_{OL}) and amplitude of f_n (A_{fb}) and twice f_b (A_{2fb}) are chosen as examples for illustrating the feasibility of vibration indicators on building the normal behavior models of PTs.

IV. CONDITION MONITORING WITH ARTIFICIAL NEURAL NETWORKS

Artificial neural network (ANN) is one of the most popular interdisciplinary science that has a large development in recent years. ANN is inspired by biological systems with a large number of neurons collectively performance tasks that even the largest computer has not been able to match. They are parallel computational models comprised of densely interconnected process units, which are called neurons. The networks' learning results are restored within the inter-neuron connection in the form of synaptic weights[15]. Synaptic weights allow the networks to adaptive themselves by "training" instead of "programming" in solving problems, which is the most important feature of ANNs. With this feature, ANNs are very appealing in the cases where one has incomplete or little understanding of the problems to be solved while the existed data can be used for training, such as taxonomic problem, pattern recognition, trend analysis and prediction, etc. Neural Networks Method (NNM) has been widely applied in biology, medicine, finance, industrial engineering and so on. Comparing with linear regression model, NN based model can represent better the vibration signal of PT because the vibration signal contains some nonlinear relationships that linear regression or other regression methods cannot fit so well.

The structure of a neurone is shown in Figure 4. One neuron is a basic information-processing element of an NN. Generally, it is consisted of a set of synaptic weights, an adder and an activation function (or "transfer function")[16].

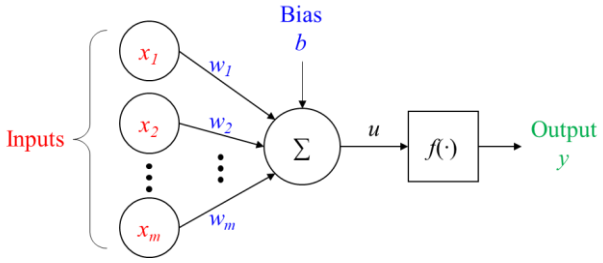


Figure 4 Schematic diagram of one neuron

The neuron depicted in Figure 4 can be described in mathematical terms as the pair of equations:

$$u = \sum_{j=1}^{m} (w_m \cdot x_m + b) \tag{2}$$

and

$$y = f(u) \tag{3}$$

Where $x_1, x_2, \ldots, x_m$ are the input values; $w_1, w_2, \ldots, w_m$ are the respective synaptic weights of the inputs; b is the bias and $f(\cdot)$ is the activation function. Activation functions are usually sigmoid function like tang and exp function which normalize the output of each neuron into a finite range [-1,1] or [0 1]. The normalization helps increase the convergence speed. On the other hand, sigmoid function builds a nonlinear relationship between the input and output which enables the network to approach any nonlinear function. Another characteristic of the sigmoid functions is they are differentiable, which is essential for a BP-NN to correct their synaptic weights by computing the partial derivatives of the error to each weight.

A. *Input and output patterns pre-processing*

a. Validity check. A simple way for validity check is a data range check. Meanwhile, the other conditions like the consistence between inputs and between inputs and outputs are also need to be checked. In this paper, the extreme outliners and data that has unexpected gradient have been removed. 521 sets of input data have passed the validity check. Especially, for checking the accuracy of the trained network, 20% of the samples have been selected from the whole data set and are used as validation data set. The left 80% samples are used as training set.

b. Data scaling. Data scaling is essential for the pre-processing of the input data because the sigmoid functions are always used as transfer functions and the output is most sensitive to variations of the input values[17]. Each group of input data is scaled within minus one to one with the following equation:

$$X = \frac{2x - x_{max} - x_{min}}{x_{max} - x_{min}} \tag{4}$$

Where x is the observed value of each sample, x_{max} and x_{min} are the maximum and minimum value of each signal respectively, X is the normalized value of observed value x.

B. *Building NNs*

a. Network topology selection. Backpropagation neural network (BP-NN), which is a typical kind of feedforward neural network (FNN), is a combination of back propagation (BP) algorithm and Multi-Layer Perceptron (MLP) FNN. Gradients are calculated and are used to fix the error of each synaptic weights. The typical structure of BP-NN model applied in this research consists one input layer, one output layer and one or more so-called hidden layers (because these layers do not express themselves to the external environment). Each neuron (or perceptron, cells, units, nodes, processing elements) at subsequence layer is connected by arc (weight) with neurons at prior layer, while neurons in the same layer are independent between each other[18].

b. Numbers of perceptron and layers. It has been proven that a feed forward network with one hidden layer can represent any continuous function. However, there is no generally accepted theory to determine how many hidden neurons are needed in each hidden layer to approximate any given function. It is recommended in [19] that the optimum number of neurons should be found by performing at least 10 runs where only the number of neurons is changed. In our case a single-layer feedforward network with 12 perceptron in the hidden layer is applied, as shown in Figure 5.

According to Section III, A_{OL}, A_{fn} and A_{2fb} are chosen as the examples of the monitoring indicators. The auto-regression normal behavior model is mainly focus on these 3 parameters.

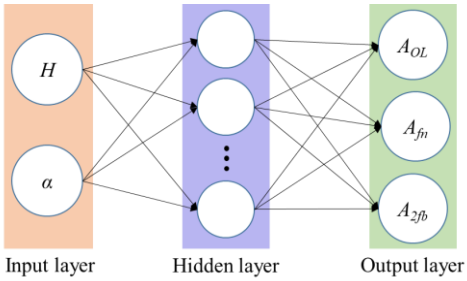


Figure 5 Feed-forward network architecture (Input: head level and GVO; output: head, guide vane opening and vibration overall level; Numbers of hidden layers: 12)

c. Network transfer/activation function. In order to avoid generating extreme values by the neurons, transfer functions are applied in the neurons to limit the permissible amplitude range of the output signal to some finite value[16]. The transfer function, or activation function, is usually a strictly increasing function that exhibits a graceful balance between linear and nonlinear behavior. Sigmoid functions, which have characteristic "S"-shaped curves, are by far the most common form of activation function used in the construction of neural networks. In our case, the log-sigmoid function, which transfer the layer's input into a value between 0 and 1, is chosen as the transfer function:

$$f(x) = \frac{1}{1 + exp^{-x}} \tag{5}$$

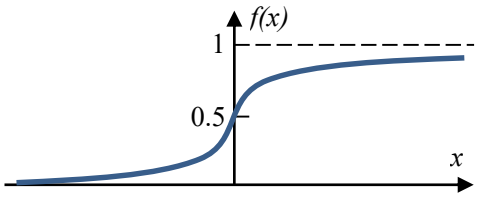


Figure 6 Log-sigmoid function in perceptron

d. Training method. Levenberg-Marquardt backpropagation function is used as the training method in our case. LM algorithm appears to be the fastest method for training moderate-sized feedforward neural networks (up to several hundred weights), although it does require more memory than other algorithms. It has added a damping factor λ into the Gauss–Newton algorithm:

$$x_{k+1} = x_k + [J^T J + \lambda I]^{-1} J^T e \tag{6}$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, **e** is a vector of network errors and I is an identity matrix. The advantage of lm algorithm is the scalar λ can be adjusted: when it is large, this algorithm is more like a gradient descent method. While with the proceed of the iteration, λ decreases, which makes it closer to Gauss–Newton algorithm which has a higher accuracy on approximation[20].

V. RESULT AND DISCUSSION

The network was obtained after 2000 epochs training. Beside the existing operating parameters, it is able to respond to any new operating condition within the input range. Given the number of input is only 2 parameters, the variation of each output can be represented in a 3-dimention figure and its contour plot as shown in Figure 7. The PT nomal behavior model representing the variation of overall level A_{OL} versus operating parameters is shown in Figure 7 a) while its sub-models represent the variation of A_{fn} and A_{2fb} are shown in Figure 7 b) and c).

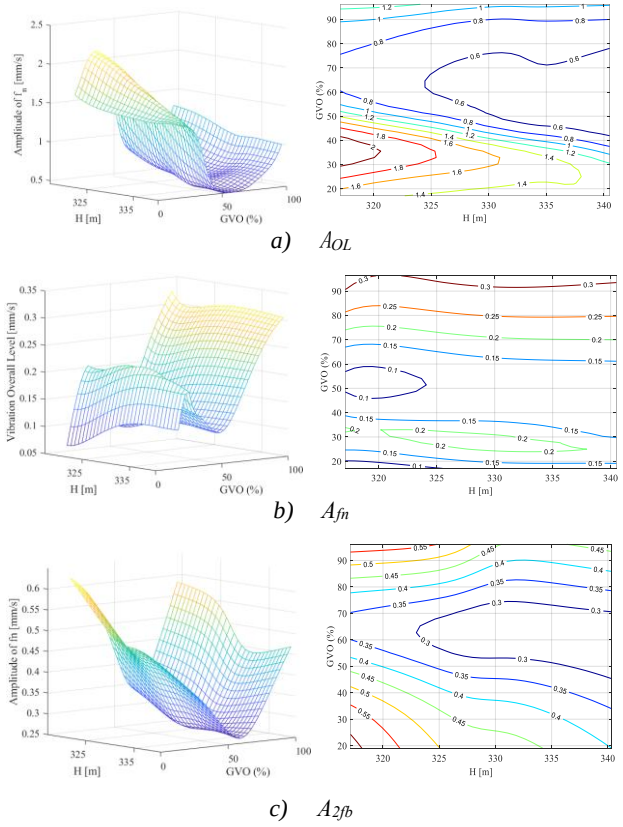


Figure 7 Vibration amplitude of bearing turbine changing with operating condition (left) and their contour lines plots (right)

The coefficient of determination R^2 is one of the most widely used goodness-of-fit metrics. It represents the proportion of observed variation that can be explained by the simple linear regression model.

$$R^2 = 1 - \frac{\sum[(\hat{y} - y)^2]}{\sum[(y - \bar{y})^2]} \tag{7}$$

Where $\hat{y}$ is the predicted value from the regression model. y is the observed value and $\bar{y}$ is the mean value of y. Whereas the shortage of R^2 is its value always rises with the increase of the number of the input variations. In this paper, the adjusted coefficient of determination R^2_{adj}, which has considered the number of inputs, is used for checking the precision of the trained model in predicting the the vibration characteristics of the operating conditions in validating set.

$$R^2_{adj} = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1}\right] \tag{8}$$

Where n is the number of the data samples; k is the number of variables in the regression model.
The validating result are listed in Table 1. R^2_{adj} of each indicator is near to 1, which proved that the networks provid fairly good results.

Table 1 multiple determination R^2 of each indicator

Indicator	*Overall level*	f_n	$2*f_b$
R^2_{adj}	0.9426	0.9560	0.7793

On one hand, normal behavior models provide us a reference behaviour expected for each component, according to its current working conditions. The conventional alarm limit is a constant value defined by the same standard. False alarms could be easily happen when the machine is working far away from the BEP. In an extended range, the increase of the overall level might not mean fault but the increase of the excitation forces or the change on the dynamic response.With the normal behavior model, the alarm threshold can be optimized according to the reference vibration levels. Misinformation in fault diagnosis would be decreased.
Comparing with normal behavior model, the sub-models give different vibration ranges and response to the change of operating parameters. The sub-models are useful for improving monitoring ability since they correspond to specific vibration frequencies caused by different damages. When the overall vibration level is within the normal range according to the normal range, sub-models alarm may be triggered by a great change in A_{fn} or the other sub-models. On the other hand, the normal behavior models are helpful for estimating the remaining useful life (RUL) of the machine by providing the power plant operators information on the distribution of each type of vibration under different operating conditions. For instance, according to Figure 7, in high head condition the machine is more prone to suffer machinery origin damage like bearing damage, etc. since A_{fn} increases with the rise of head while A_{2fb} decrease. According to the normal behaviour models, operators in the power plant can make a plan on the operating condition in which the machine works to control the excitation forces applied on the machine. Some specific impulse or strains can be avoided and the overall RUL of the machine would be extended. According to the operation plan based on normal behavior model, the turbine unit can run in the off-design conditions for a certain time, when there is a demand from the power grid. This reasonable extension will be transferred into a great amount of revenue for the plant without destroying the machine.

VI. CONCLUSIONS AND OUTLOOK

In this paper, vibration signals over one-month monitoring have been obtained from the turbine bearing of a pump-turbine which is working in an extended range. The main excitation forces and vibration of the machine have been analysed.

A single-layer neural network has been designed for building normal behaviour models. Head level and GVO were used as input patterns and vibration overall level A_{OL}, A_{fn} and A_{2fb} are chosen as the monitoring parameters. The prediction result proved that the network provided fairly good results.

Comparing with the conventional monitoring system, the NNs based normal behaviour models have the following advantages:

1) With the normal behaviour models the possibility of monitoring the signal is widely decoupled from the operational mode. Misinformation in fault diagnosis would be decreased.
2) Diagnosis accuracy would be increased since the sub-models focus on specific damage in the machine;
3) Predictive operating plan can be drawn up according to the normal 'damage model' and the requirement from the power grid. the maximum revenue can be achieved within the RUL of the machine.

In the follow-up studies. The effectiveness of damage behaviour models based on more condition indicators (e.g. vortex rope, natural frequency bands) from different components (generator bearings, thrust bearings) of the hydro turbine unit by means of more AI techniques will be researched.

REFERENCES

[1] **A，J， Wood.；B，F， Wollenberg**., Power generation, operation, and control, John Wiley & Sons, 2012.

[2] **E，Egusquiza.；C，Valero.；D，Valentin.；A，Presas.；C，G，Rodriguez**., Condition monitoring of pump-turbines. New challenges, *Measurement*, Vol. 67., (2015) 151–163.

[3] **P，Dörfler.；M，Sick.；A，Coutu**., Flow-induced pulsation and vibration in hydroelectric machinery: engineer's guidebook for planning, design and troubleshooting, Springer Science & Business Media, 2012.

[4] **M，Nishi.；S，Matsunaga.；M，Okamoto.；K，Takatsu**., Wall pressure measurements as a diagnosis of draft tube surge, in: Proc. 15th IAHR Symp. Hydraul. Mach. Cavitation, Belgrade, Serbia, 1990.

[5] **F，Avellan**., Flow investigation in a Francis draft tube: the FLINDT project, in: Proc. 20th IAHR Symp. Hydraul. Mach. Syst., Charlotte (NC), 2000.

[6] **S，Alligné；P，Maruzewski.；T，Dinh.；B，Wang.；A，Fedorov.；J，Iosfin.；F，Avellan**., Prediction of a Francis turbine prototype full load instability from investigations on the reduced scale model, in: IOP Conf. Ser. Earth Environ. Sci., IOP Publishing, 2010: p. 12025.

[7] **E，Egusquiza.；D，Valentín.；A，Presas.；C，Valero**., Overview of the experimental tests in prototype, in: J. Phys. Conf. Ser., IOP Publishing, 2017: p. 12037.

[8] **D，Valentin.；A，Presas.；M，Egusquiza Montagut.；C，Valero.；E，Egusquiza**., Behavior of Francis turbines at part load. Field assessment in prototype: Effects on power swing, 2019.

[9] **E，Egusquiza.；C，Valero.；X，Huang.；E，Jou.；A，Guardo.；C，Rodriguez**., Failure investigation of a large pump-turbine runner, *Eng. Fail. Anal.*, Vol. 23., (2012) 27–34.

[10] **E，Egusquiza.；C，Valero.；A，Estévez.；A，Guardo.；M，Coussirat**., Failures due to ingested bodies in hydraulic turbines, *Eng. Fail. Anal.*, Vol. 18., (2011) 464–473.

[11] **M，Schlechtingen.；I，F，Santos**., Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection, *Mech. Syst. Signal Process.*, Vol. 25., (2011) 1849–1875.

[12] **M，Schlechtingen.；I，F，Santos.；S，Achiche**., Using data-mining approaches for wind turbine power curve monitoring: a comparative study, *IEEE Trans. Sustain. Energy*, Vol. 4., (2013) 671–679.

[13] **G，P，Succi.；H，Chin**., Helicopter Hydraulic Pump Condition Monitoring Using Neural Net Analysis of the Vibration Signature, SAE Technical Paper, 1996.

[14] **R，A，Saeed.；A，N，Galybin.；V，Popov**., 3D fluid–structure modelling and vibration analysis for fault diagnosis of Francis turbine using multiple ANN and multiple ANFIS, *Mech. Syst. Signal Process.*, Vol. 34., (2013) 259–276.

[15] **A，Choubey.；D，K，Sehgal.；N，Tandon**., Finite element analysis of vessels to study changes in natural frequencies due to cracks, *Int. J. Press. Vessel. Pip.*, Vol. 83., (2006) 181–187.

[16] **S，S，Haykin.；K，Elektroingenieur**., Neural networks and learning machines, Pearson Upper Saddle River, 2009.

[17] **M，Y，Rafiq.；G，Bugmann.；D，J，Easterbrook**., Neural network design for engineering applications, *Comput. Struct.*, Vol. 79., (2001) 1541–1552.

[18] **J，Schmidhuber**., Deep learning in neural networks: An overview, *Neural Networks*, Vol. 61., (2015) 85–117.

[19] **L，Tarassenko**., Guide to neural computing applications, Elsevier, 1998.

[20] **K，Madsen.；H，B，Nielsen.；O，Tingleff**., Methods for non-linear least squares problems, (1999).

16th IMEKO TC10 Conference
"Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety
Berlin, Germany, on September 3-4, 2019

Application of the fusion of regression machines for the analog circuit state identification

Piotr Bilski[1]

[1]*Institue of Radioelectronics and Multimedia Technology, Warsaw University of Technology, Warsaw, Poland, pbilski@ire.pw.edu.pl, +48 601056969*

***Abstract* – The paper presents the application of the combined group of regression algorithms to identify state of the analog circuit. Implementing the fusion of regression machines is aimed at obtaining high accuracy of the diagnosed system's state, especially compared to the single parameter identification algorithm. The large number of simple methods (such as linear regression techniques) is expected to give the high accuracy without the need of time consuming and complex optimization of the selected approach (such as Support Vector Machines – SVM). The approach consists in preparing the ensemble architecture, selecting computational methods, optimizing features extracted from the diagnosed system and testing the approach. The tests were conducted to evaluate efficiency of various fusion architectures, determine their accuracy for different sets of features and confront them against the single optimized regression algorithm. The time analysis verified the ability of the approach to use the framework in the online mode. Obtained results show the potential of the proposed framework for the accurate identification of analog system parameters, which can be used to analyse other types of systems.**

Keywords – regression, diagnostics of analog systems, parameter identification.

I. INTRODUCTION

The modern diagnostics of analog systems faces multiple problems, which are solved using various approaches, among which the most prominent belong to the Artificial Intelligence (AI) domain. One of the pressing issues is the accurate Parameter Identification (PI) in the system [1], which allows for determining its state during the long-term exploitation. Initially in the nominal state, the system degrades with time and its parameters change, starting from small deviations from the original values, further going beyond the tolerance regions. In the complex industrial systems and sophisticated devices it is important to learn in advance that they do not (or will not in the nearest future) operate exactly as they should. This way faulty elements may be replaced before the object becomes permanently damaged. Systems critical for society (such as power plants [2]) are constantly monitored, as their parameters must not deviate beyond the nominal state.

The PI is difficult because of multiple factors and phenomena. These include the noise present in the observed System Under Test (SUT) responses, tolerance regions, making the definition of the nominal state "fuzzy", existence of ambiguity groups [3], etc. Also, in most systems there is the non-linear relation between the parameters' values and the observed diagnostic features. These pose a challenge for the applied algorithm, usually leading to a long and mundane process of adjusting parameters. In many applications linear regression (approximation) approaches provide the acceptable accuracy. They can be the alternative to more complex methods, such as SVM [4].

The paper presents the fusion (or ensemble) architecture of AI-originated regression machines determining the state of the electronic circuit. Application of multiple algorithms estimating its parameters is confronted against the single, more sophisticated approach, i.e. .SVM regressor. It is assumed that voting mechanism allows multiple approximators to obtain the greater accuracy. The tested circuit contains discrete elements, which change over time and influence its work regime. Though integrated circuits are currently majority and in the case of leaving the nominal state they can be easily disposed of, in some applications, such as expensive audio amplifiers it is still justified to replace the damaged elements instead of exchanging it into the new device.

The paper structure is as follows. In Section II the parameter identification task is introduced. Section III discusses the proposed fusion architecture with the characteristics of the usable algorithms. In Section IV details of the prepared solution (such as voting mechanism) are considered. Section V presents the SUT, while in Section VI conducted tests are presented. Section VII contains conclusions and future prospects.

II. AI-BASED MULTIPLE PARAMETER IDENTIFICATION PROBLEM

Every SUT, disregarding its technical nature

(electronic circuits, electrical machines, automation elements, etc.), should operate according to the design principles, defining the nominal state. It depends on the configuration of SUT parameters $\boldsymbol{p}=\{p_1, \ldots, p_m\}$ (such as capacitances, resistances' values, amplification coefficient, etc.), which determine produced output (response) signals ***y(t)*** as reaction to excitations ***x(t)***.

$$\boldsymbol{y}(t) = f(\boldsymbol{x}(t), \boldsymbol{p}) \tag{1}$$

The task is to construct the function *g()*, which will be able to map information represented by SUT responses into the approximated set of parameters $\hat{\boldsymbol{p}}$, which should be as close to the actual set $\boldsymbol{p}$ as possible (usually in the least squares sense).

$$\hat{\boldsymbol{p}} = g\big(\boldsymbol{y}(t), \boldsymbol{x}(t)\big) \tag{2}$$

Unfortunately, the function *f()* (1) is not easily reversible, therefore based on the observation of responses it is difficult to determine exact values of parameters, even if excitation patterns ***x(t)*** are known. The relation (1) is usually described by the computer model, therefore it is easy to establish the approximations $\hat{\boldsymbol{p}}$ heuristically, by generating multiple simulations for varying configurations of SUT parameters. This allows for creating the diagnostic module, representing the function (2), i.e. the function *g()*.

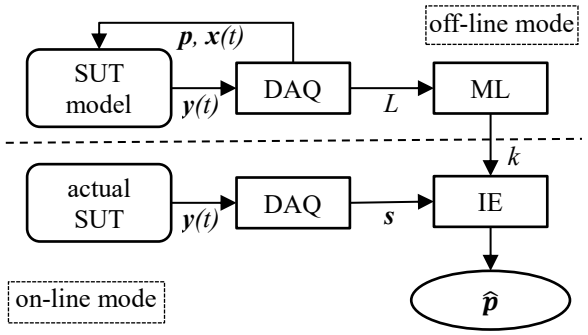


Fig. 1. Architecture of the PI module

Currently the most popular are data-driven approaches. They enable defining the PI task based on the algorithm extracting knowledge from available sets, for instance through the machine learning process [5]. The sets are first created during the simulations of different SUT states. The diagnostic features (symptoms) are extracted from the output signals to form the symptoms' vectors ***s***, based on which knowledge *k* about the actual SUT state can be induced. The AI-based identification implements the scheme in Fig. 1, operating in the off-line and on-line mode. In the former the DAQ module initiates model simulations, acquires response signals and extracts symptoms from them. The Machine Learning (ML) module operates on the learning set *L* to generate knowledge *k* connecting observed symptoms and predicted parameters' values. It is the core of the Inference Engine (IE), which (working in the on-line mode) will process the vector of symptoms extracted from the actual SUT and estimate values of parameters $\hat{\boldsymbol{p}}$.

Contrary to the popular fault detection and location scheme, where the system must be classified as working correctly or damaged (described by the discrete category), in the presented PI task the expected outcome is the predicted real-number value of the parameter p_i. According to Fig. 1, it can be described as follows:

$$\{k(L), \boldsymbol{s}\} \rightarrow \mathbb{R}^m \tag{3}$$

Data sets used in AI-based identification consist of vectors of diagnostic symptoms $\boldsymbol{s}_i$ (further also called examples) coming from model simulations for the specific configurations of parameters $\boldsymbol{p}$, which values describe each example. This way it is possible to perform supervised learning, discovering relation between symptoms and SUT parameters. The scheme requires two data sets of the same form (4), used for training and testing (*L* and *T*, respectively) the regression machine. From *L* knowledge *k* is extracted, while *T* is used to test generalization abilities of the regression method. Both sets shoudl then contain different, exclusive examples.

$$L \equiv T = \begin{bmatrix} \boldsymbol{s}_1 & \boldsymbol{p}_1 \\ \vdots & \vdots \\ \boldsymbol{s}_n & \boldsymbol{p}_n \end{bmatrix} = \begin{bmatrix} s_{11} & \cdots & s_{1l} & p_{11} & \cdots & p_{1m} \\ & \ddots & & & \ddots & \\ s_{n1} & \cdots & s_{nl} & p_{n1} & \cdots & p_{nm} \end{bmatrix} \tag{4}$$

Quality of generated knowledge is measured using one of the possible determination scores. In the presented research the measure (5) is used [6], where p_{ij} is the actual value of the *i-th* SUT parameter for the *j-th* example from the set, $\hat{p}_{ij}$ is its estimated value and $\bar{p}_{ij}$ is the average of the actual values. The maximum possible score is equal to 1 (when the predicted and actual parameters' values are identical) with the unlimited minimum going towards -∞. The score is calculated for each parameter separately. This way it is possible to optimize the regression machines regarding all parameters for the whole available set, by maximizing value of (5). In practice, only the range (0.5; 1) is useful (i.e. close enough to the actual values) PI.

$$acc(p_i) = 1 - \frac{\sum_{j=1}^{n}(p_{ij} - \hat{p}_{ij})^2}{\sum_{j=1}^{n}(p_{ij} - \bar{p}_{ij})^2} \tag{5}$$

The score (5) is used in the presented research to test the regression quality on *L* and *T*. Result for the former is used to configure the ensemble, while the latter shows the performance of the designed approach. Vectors of symptoms prepared for both sets should cover the most typical SUT states: nominal and the most probable deviations. It is assumed that the framework detects signle faults, i.e. only one parameter is beyond nominal, while all other are within their tollerance margins. Though the

proposed scheme is able to identify values of all parameters indopendently, its efficiency for the multiple faults prediction must be evalutaed separately.

In the following sections the architecture of the classifiers' fusion is introduced. Next, the implementation of the ensemble as the software module is presented. To create the architecture, Python language was used, including the sci-kit package [7], which is the standard library of the AI-based algorithms.

III. THE PROPOSED FUSION OF REGRESSION ALGORITMS

To solve the PI problem, the single regression method is usually applied (such as RBF Artificial Neural Networks or SVM). These approaches consider non-linear relations between the parameters' values and observed symptoms, additionally SVM is able to cope with the data inconsistency. The disadvantage of these approaches is the need to tune them to obtain the best results. Alternatively, simpler, linear regression methods may be used here. Usually not as effective as their non-linear counterparts, they can be exploited in the more accurate way by combining them into the single regression machine.

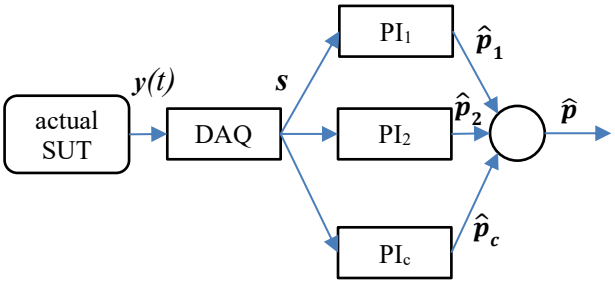


Fig. 2. Ensemble of regression machines

A. Architecture overview

Each algorithm in the ensemble is trained separately on the same set *L* (4), but knowledge *k* extracted from it depends on the particular method. Therefore it is expected that regression results will be different for each algorithm and the combined outcome can be better than of any other separate approach. The scheme in Fig. 2 is similar to the classifier ensemble, used in diagnostics before [8]. The difference lies in the way of calculating the output value, strongly related to the applied voting mechanism (which is investigated further).

The software architecture is aimed at training all approaches subsequently in the off-line mode on the previously generated set *L*. The multiple PI is performed by all algorithms working in parallel, the overall result is calculated after their job is complete.

The linear regression implements the following relation between the *i-th* SUT parameter and analysed symptoms [7]:

$$p_i = \sum_{j=1}^{l} s_j \cdot w_j + w_0 \qquad (6)$$

where weights w_j are adjusted during the training process. Each algorithm introduces modifications to (6), leading to different PI results.

B. Considered regression algorithms

Among multiple algorithms the following were selected, based on their work regime and potential usefulness for the task. The simpler ones were preferred, to check their applicability for complex data processing:

- Least Squares (LS) regression, which implements directly equation (6) and provides the simple approximation of independent features.
- Ridge (R) regression, which modifies (7) by adding penalty for too large coefficients, imposing as simple approximation as possible.
- Bayesian Ridge (BR) regression includes the regularization coefficients to the approximation. The produces value of ***y(t)*** is treated as the random variable with normal distribution.
- Orthogonal Matching Pursuit (OMP) is the implementation of the linear regression by defining the specific number of non-zero coefficients w_i.
- Lasso Least Angle Regression with Information Criterion selection (LLIC) is the combined approach of high-dimensional data processing with sparse coefficients in (6) and the model selection using the Akaike or Bayesian criterion.
- Automatic Relevance Determination (ARD) is the variant of the Bayesian Ridge regression with the sparse coefficients w_i.
- Theil-Sen (TS) estimation is the non-parametric method insensitive to outliers, with the result curve calculated as the median among data points.

Other possible algorithms, including Passive Aggressive estimator or Elastic Net were initially tested and discarded due to their low scores. Each algorithm has multiple copies, responsible for their individual SUT parameter. The *i-th* PI machine in Fig. 2 consists of *m* separate approximators, generating together the vector of values $\hat{\boldsymbol{p}}$. If needed, the fusion may be supplemented by other algorithms.

The presented approach was confronted with the more sophisticated approximation algorithm, i.e. SVM regressor. It is a well-established approach, especially for processing uncertain data. Its disadvantage is the time-consuming process of selecting optimal parameters of the kernel. The aim of the experiment was to check if the ensemble can outperform the SVM-based approximator.

IV. ARCHITECTURE DETAILS

During the ensemble construction, it is crucial to determine the set of algorithms and their significance during the overall result calculation. Initially, all available

methods were tested on the provided learning data sets, then *c* the best ones were selected for the fusion, which was finally evaluated on the testing sets. The comparison was made for each parameter separately. It turned out that there is the subset of regression methods being more applicable to the PI task than others, but it is not possible to point out the single one dominating all of them. This justifies using the ensemble rather than the single method. In this section issues related with the architecture construction are presented. For each parameter the ensemble is constructed only of algorithms with the score on *L* high enough.

A. Voting mechanism

To generate the single output out of the intermediate values produced by different approximators, they must be combined in a way reflecting significance of each algorithm. The proposed relation for the *i-th* parameter estimation $\hat{p}_i$ is as follows:

$$\hat{p}_i = \frac{1}{c} \cdot \sum_{j=1}^{c} \hat{p}_{ij} \cdot \alpha_i \qquad (7)$$

where $\boldsymbol{\alpha}=\{\alpha_1, \ldots, \alpha_c\}$ is the vector of weighting coefficients for subsequent regression methods, calculated based on the performance on *L*. This way the most accurate approaches influence the most the estimation result. The weights are calculated as follows, where $acc_j(p_i)$ is the score of the *j-th* algorithm of the *i-th* estimated SUT parameter:

$$\alpha_j = \frac{\frac{1}{1-acc_j(p_i)}}{\sum_{i=1}^{c}\frac{1}{1-acc_j(p_i)}}, \quad \sum_{j=1}^{c}\alpha_j = 1 \qquad (8)$$

All weights are scaled to represent values between *0* and 1 and they are proportional to the accuracy (5). To avoid influencing the overall result by the regressor with the low score on the training set, the weight if calculated provided that $acc_j(p_i)$ is greater than the threshold θ, which was set to *0.45*. When the score is lower, the algorithm does not take part in voting (its weight is zero).

B. Features significance

The estimation accuracy depends on the symptoms available to the algorithm. Without the prior knowledge about the features importance, all should be used. Their large number may significantly degrade regression. Therefore elimination of redundant and not-important features was proposed. Besides the correlation analysis (to find dependent variables), there is the need to determine the set of symptoms representing changes in the specific parameter. This way it could be easier to decompose PI task for different parameters into independent routines, based on different sets of symptoms. Due to the existence ambiguity groups, such a separation may not be always possible. The algorithm is presented in Fig. 3.

```
features_selection(L)
    foreach p_j
        L(p_j) ← ∅
        foreach s_i
            if change(s_i, p_j) == TRUE
                L(p_j) ← L(p_j) ∪ s_i
```

Fig. 3. Features selection algorithm

For each SUT parameter the subset of features from the learning set *L* is identified, depending on whether their change determines the change of the corresponding symptom's value. The features' values should be beyond the 10% margin from their values in the nominal state (the **change**() function). In result the regression machine estimating the selected parameter operates on the individual set of symptoms (represented by $L(p_j)$). The approximators' scores for the original and reduced sets were compared to verify which gives better results.

V. EXPERIMENTAL TEST STAND

The proposed architecture was tested on the 5th order lowpass analog filter, presented in Fig. 4. Its structure contains three operational amplifiers, five capacitors and resistances. The latter two groups of elements were subject to PI task, based on the symptoms extracted from the SUT responses. These are output signals generated after providing the sinusoidal excitation (1V amplitude and 9kHz frequency, i.e. close to the 10kHz cut-off frequency of the circuit) to the SUT input (node No. 1). Responses were recorded at accessible nodes No. 2, 3, 5, 6, 8 and 9. From output patterns the first three maxima, minima, zero crossings and their corresponding time instants were extracted as symptoms. This way each symptoms vector contains 54 features, which can further be minimized.

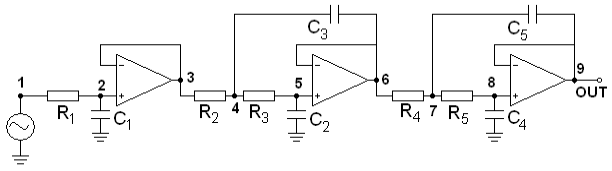


Fig. 4. Fifth order lowpass filter

The nominal values of identified elements are as follows: $R_1=R_2=R_3=R_4=R_5=1k\Omega$, $C_1=16nF$, $C_2=19nF$, $C_3=13nF$, $C_4=51nF$ and $C_5=49nF$. The model of the SUT was implemented in the Simulink environment. Experiments consisted in setting the selected parameter out of tollerance margins, while leaving all remaining ones within the nominal state (affected by the random variable representing tollerance dispersion and the uniform noise). Elements' changes ranged from 10% to 90% of the nominal value (in both directions, i.e. above and below it).

VI. EXPERIMENTAL RESULTS

The experiments were divided into a couple of stages, described in the subsequent sections. Firstly, all described algorithms were tested separately to check their usefulness to estimate SUT parameters. Next, algorithms selected to ensemble were combined to form the PI module and confronted against SVM regression. Finally, the features' selection algorithm was implemented to increase the regression accuracy if possible.

Four pairs of data sets (i.e. *L* and *T*) were used for tests. The first two (L_1, T_1 and L_2, T_2, respectively) included 70 and 180 examples, with actual parameters' values not affected by tolerances. The remaining two pairs (L_3, T_3 and L_4, T_4), also containing 70 and 180 examples respectively, included simulations affected by elements' tolerances and noise. The smaller sets covered 7 simulations for each diagnosed element, while the more extensive tests included 18 changes of each parameter's value.

A. Accuracy of selected regression algorithms

Algorithms introduced in Section III.B were tested separately on full vectors of symptoms. Results on the set L_1 for each parameter are in Table 1, while for the set T_1 are in Table 2. The bold font indicates the highest score, while italics show, which algorithm's accuracy allows for including it in the ensemble. The "0.00" values indicate the approximator was unable to converge on the learning set, therefore it will not take part in voting. In most cases the LS regression is the best on the learning set, but as Table 2 shows, knowledge extracted in this case is not well generalized, which justifies application of the fusion. Although the analysis of L_1 is dominated by two methods: LS and TSR, knowledge verification on T_1 shows more algorithms involved in the accurate regression: also BR, LLIC have the highest scores for some parameters. Table 2 also shows, which SUT elements are difficult to identify: for R_1, C_1 and C_2 it is hard to get scores high enough to make them usable.

Table 1. Regression scores on the set L_1.

	LS	R	BR	OMP	LLIC	ARD	TSR
R_1	**0.95**	0.41	*0.49*	*0.50*	*0.50*	*0.50*	0.09
R_2	*0.91*	0.18	*0.67*	*0.70*	*0.80*	*0.68*	**0.98**
R_3	**0.97**	0.10	*0.73*	*0.72*	*0.75*	*0.70*	0.89
R_4	**1.00**	0.26	*0.73*	*0.66*	*0.71*	*0.74*	0.99
R_5	**1.00**	0.08	0.00	*0.64*	*0.65*	0.38	0.97
C_1	**0.95**	0.41	0.48	0.00	0.00	0.10	-0.87
C_2	*0.92*	0.04	0.10	0.00	0.00	0.00	**0.98**
C_3	*0.90*	*0.52*	*0.65*	0.00	0.00	0.01	**0.97**
C_4	**1.00**	0.07	0.11	*0.55*	0.00	0.00	0.99
C_5	**1.00**	0.27	0.40	*0.53*	0.00	0.00	0.96

Results on other data sets are similar, still showing advantage of LS and TSR over other approaches, but after introducing tolerances and noise the maximum scores are dispersed among other methods. Symptoms affected by these phenomena are more difficult to process, therefore regression accuracy for T_3 and T_4 is lower (on average by *0.1*) than for T_1 and T_2. Increasing the number of examples in the set did not improve scores. Accuracy on the testing set is always lower on the learning one, which suggests the loss of generalization during the training.

Table 2. Regression scores on the set T_1.

	LS	R	BR	OMP	LLIC	ARD	TSR
R_1	-0.09	0.41	0.49	0.50	**0.51**	0.50	-2.20
R_2	-1.21	0.18	0.68	0.68	**0.82**	0.68	0.42
R_3	-1.23	0.10	0.73	0.72	**0.74**	0.69	0.29
R_4	0.93	0.25	0.74	0.64	0.71	0.75	**0.94**
R_5	**0.90**	0.08	0.00	0.61	0.66	0.35	0.83
C_1	-0.01	0.41	**0.47**	0.00	0.00	0.10	-1.59
C_2	-1.24	0.04	**0.12**	0.00	0.00	0.00	-0.02
C_3	-1.16	0.51	**0.64**	0.00	0.00	0.01	0.44
C_4	0.92	0.06	0.10	0.53	0.00	0.00	**0.93**
C_5	**0.92**	0.31	0.44	0.50	0.00	0.00	0.77

B. Efficiency of the ensemble

Based on the information from algorithms used individually, the ensemble was constructed with the voting mechanism defined by (7) and (8). In Table 3 results for the test T_3 are presented. In the subsequent columns, scores of the algorithms (with least one the most accurate estimation) treated separately, ensemble (ENS) and the SVM regressor (SVR) outcomes are shown. The latter was designed to use the RBF kernel with the optimal width $\sigma=0.001$. Again, bold font indicates the best estimation. In all cases the fusion is more accurate than SVM, which requires complex and time-consuming tuning (covering selection the best kernel and adjusting its parameters). The overall performance shows that although not the best for all parameters, the fusion of regressors is the most reliable approach, maintaining high score for all parameters.

Analysis of the ensemble performance has also shown that in some cases eliminating the LS from the fusion leads to increasing the score, as this method fails to properly generalize knowledge about values of R_1, C_1 and C_2.

Table 3. Regression scores on the set T_3, including ensemble and reference method.

	LS	BR	LLIC	ENS	SVR
R_1	**0,63**	0,49	0,50	0,59	0,34
R_2	0,81	0,69	0,66	**0,82**	0,27
R_3	0,77	0,66	0,71	**0,78**	0,65
R_4	0,35	0,70	**0,74**	0,45	0,38
R_5	0,56	**0,79**	0,69	0,70	-0,17
C_1	**0,68**	0,50	0,00	0,55	-0,02
C_2	0,82	0,23	0,00	**0,83**	0,63
C_3	**0,77**	0,69	0,00	0,75	0,48
C_4	**0,81**	0,13	0,00	0,80	0,23
C_5	0,37	0,55	0,00	**0,59**	-0,11

The actual accuracy of regression is presented in Fig. 5. Here variability of the selected parameter, i.e. R_5 for the set T_4 (180 examples with tolerances and noise) with its estimated values obtained by the LS regression and fusion of all methods from Section III.B are shown. Both outcomes are close to the actual values, especially around the nominal range, eliminating the threat of the false alarm.

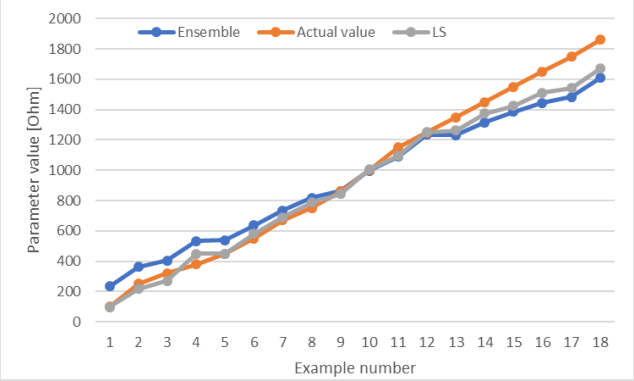


Fig. 5. Comparison between the regression accuracy of the Least Squares and ensemble regression

C. Influence of the feature selection algorithm

Experiments presented so far were conducted in the full data sets, containing all 54 symptoms, extracted from the SUT responses at accessible nodes. Introduction of the algorithm from Section IV.B was aimed at decreasing the amount of data processed by the algorithms, and hopefully maximizing the approximation scores. For each SUT element a different subset of symptoms was returned, causing different conditions for training the regression algorithms. In Table 4 results for processing the set T_1 by all algorithms are presented, including their ensemble. Outcomes for the symptoms' reduction method are inconclusive, as in some cases (C_2, C_3, C_4) the performance of both individual algorithms and ensemble has increased, but in other the accuracy dropped (especially for R_5, which is caused by the elimination of too many symptoms).

Table 4. Regression scores on the set T_1 for the reduced sets of symptoms.

	LS	R	BR	OMP	LLIC	TSR	ENS
R_1	-0,04	0,19	0,26	**0,49**	0,48	0,00	0,32
R_2	0,41	0,12	0,41	0,23	**0,41**	0,00	0,38
R_3	**0,68**	0,00	0,00	0,45	0,56	0,00	0,64
R_4	**0,33**	0,11	0,30	0,30	0,32	0,00	0,28
R_5	0,15	0,05	0,10	**0,16**	0,15	0,00	0,12
C_1	**0,75**	0,39	0,49	0,00	0,00	-31,28	0,47
C_2	**0,70**	0,02	0,10	0,00	0,00	0,13	0,36
C_3	**0,85**	0,32	0,45	0,00	0,00	0,33	0,57
C_4	0,92	0,06	0,10	0,53	0,00	0,91	**0,93**
C_5	0,92	0,31	0,44	0,50	0,00	0,85	**0,93**

VII. CONCLUSIONS

The proposed ensemble of regression machines has proven the usefulness during the estimation of parameters of the electronic analog circuit. Although it is a complex system with multiple parameters to identify simultaneously, most parameters can be evaluated with the acceptable accuracy despite the noisy conditions and influence of elements' tolerances. In most cases simple linear dependencies between the symptoms and actual SUT parameter values are accurate enough to use simple methods, not requiring the extensive tuning.

The ensemble of regressors proven to be useful for the PI, as the combined operation of multiple algorithms gives in the global term better accuracy than any method used individually. The criterion of selecting the approach to the fusion based on its performance on the learning set is effective, but in some cases knowledge is not well generalized, which can be supressed by other approaches.

The features selection algorithm requires modifications and extensions, as its usefulness is currently limited. Accuracy for some parameters is improved, while for other – degraded. The most probable solution for this problem is the introduction of additional criteria for leaving or discarding the specific feature. Also, additional tests of ensemble with additional methods, not tested here (such as CART), should be performed in the future.

REFERENCES

[1] **Moradi, M.; Naraghi, M.; Nikoobin A.**: Parameter identification of nonlinear systems using indirect solution of optimal control problem, *Proceeding of the 3rd RSI International Conference on Robotics and Mechatronics*, Oct. 7-9 2015, Tehran, Iran, DOI: 10.1109/ICRoM.2015.7367797.

[2] **Lemma T.A.**: A Hybrid Approach for Power Plant Fault Diagnostics, *Studies in Computational Intelligence*, Springer, 2018,.

[3] **Stenbakken G.N.; Souders T.M.;. Stewart G.W.**: Ambiguity groups and testability, *IEEE Transactions on Instumentation and. Measurement*, Vol. 38., No. 5., 1989, pp. 941-947, DOI: 10.1109/19.39034.

[4] **Ravikumar B.; Thukaram D.; Khincha H.P.**: Application of support vector machines for fault diagnosis in power transmission system, *IET Generation Transmission & Distribution*, Vol. 2., No. 1., 2008, pp. 119-130, DOI: 10.1049/iet-gtd:20070071.

[5] **Mitchell T.**: Machine learning, *McGraw-Hill*, 1997.

[6] **Devore J. L.**: Probability and Statistics for Engineering and the Sciences (8th ed.), Boston, MA: Cengage Learning, 2011, ISBN 978-0-538-73352-6.

[7] **Pedregosa F.; et al.**: Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, Vol. 12., 2011 pp. 2825-2830.

[8] **Bilski P.**: Analysis of the classifier fusion efficiency in the diagnostics of the accelerometer, *Measurement*, Vol. 67., 2015, pp. 116–125.