IMEKO TC6 International Conference on Metrology and Digital Transformation
September 19 – September 21, 2022, Berlin, Germany

FUNDAMENTAL ASPECTS IN DATA ANALYSIS FOR SENSOR NETWORK METROLOGY

Sascha Eichstädt [a*], Anupam Prasad Vedurmudi[b], Maximilian Gruber[c], Daniel Hutzschenreuter[d]

[a] Department 9.4, PTB, Germany email address: sascha.eichstaedt@ptb.de
[b] Department 9.4, PTB, Germany email address: anupam.vedurmudi@ptb.de
[c] Department 9.4, PTB, Germany email address: maximilian.gruber@ptb.de
[d] Department 9.4, PTB, Germany email address: daniel.hutzschenreuter@ptb.de
*Corresponding author. E-mail address: sascha.eichstaedt@ptb.de

Abstract – Sensor networks underpin many developments in digital transformation, with applications ranging from regulated utility networks to low-cost Internet of Things (IoT). The metrological assessment of sensor networks necessitates a fundamental revision of calibration, uncertainty propagation and performance assessment and new approaches for information and data handling regarding the individual sensors and their interactions in the network to allow a systems metrology approach to be established. This contribution summarizes initial findings from three research projects and gives an outlook into future developments.

Keywords: sensor network, measurement uncertainty, Internet of Things, co-calibration

1. INTRODUCTION

Digital transformation includes the integration of digital technologies, such as software, communication and algorithms into products, processes, and services. A consequence of the revolution is that these technologies are in turn being used to generate completely new products, processes, and services.

Digital exchange of data and information is becoming the standard. Information is provided via cloud services in a machine-actionable way; digital infrastructures utilize information from calibration, self-diagnosis and other metadata communicated by individual measuring instruments; processes and services in the quality infrastructure are based on distributed databases and application programming interfaces (APIs).

Distributed measuring instruments and sensor networks are becoming more important than individual measuring instruments. Applications such as the Industrial Internet of Things (IIoT) and automated driving will belong to the first examples where the role of metrology is challenged. These challenges include methods for metrological traceable co-calibration and the metrological assessment of whole sensor networks in a systemic approach.

As algorithms and software become as important as the actual measurements, they will influence metrological traceability chains for measurands increasingly. In the digital age, artificial intelligence, sensor fusion and virtual measuring instruments will replace many of today's tools and principles. Their use will require a fundamental re-evaluation of the established methodologies for uncertainty evaluation and assessment of algorithms.

Quality of measurement data, trustworthiness of measurement results and reliability of measuring instruments are as important in the digital world as before. Hence, metrology plays an important role in the quality infrastructure in the digital age, too [1].

So far, metrology focuses on single measuring instruments and sensors. However, in a rapidly increasing number of applications networks of sensors are used to address measurement needs. Examples can be found in predictive maintenance of production machines in industry, urban air quality monitoring, and multi-modal human health assessment using wearables [2].

An important aspect that distinguishes sensor network applications from single sensor measurements is that rather than individual sensors, the combined information from all sensors is the main object of interest. For instance, the combination of microphone data and vibration measurement in predictive maintenance provides more insights into the actual status of the monitored machine than the individual measurements alone [3]. A consequence of the focus on combined sensor data rather than individual sensors is that the definition of the quantity of interest, the measurand, is not straightforward. Moreover, the assessment of quality and reliability of the system is more complex and challenging than the individual sensors alone, e.g., in terms of calibration result. Such an assessment also has to take into account potential sensor failure or networking issues.

Let us consider again the above example of predictive maintenance. The combination of different sensor data is carried out in a data-driven approach, i.e., using machine learning methods. The target is a classification of the remaining lifetime of the machine being monitored. Thus, the purpose of the measuring sensor network and data analysis is clear, but the outcome – expected remaining lifetime – is not a direct combination of the involved measured quantities based on physics. Since the combination of all sensor data is of interest instead of the individual sensor readings, an assessment of the measurement performance should consider the sensor network as a complex (often distributed) measuring instrument. The metrological treatment of such sensor networks thus requires a novel approach – called "systems metrology". This leads to the novel field of "sensor network metrology", which contains aspects such as, in-situ calibration and co-calibration, uncertainty evaluation for dynamic measurements and dynamically structured systems, semantic representation of metrological information, uncertainty-aware machine learning and explainable artificial intelligence applied to sensor networks.

This contribution introduces sensor network metrology aspects which were addressed in recent research projects. We outline the fundamental sensor network metrology aspects and discuss their combination into a coherent and consistent approach for a metrological treatment of sensor networks. Section 2 introduces general aspects of Internet of Things (IoT) type of sensor networks from the viewpoint of metrology. Section 3 addresses the digital representation of data and metrological information in sensor networks. Section 4 presents and discusses relevant uncertainty evaluation and propagation for processing of sensor network data. Section 5 discusses aspects related to the application of machine learning and artificial intelligence. Finally, Section 6 addresses the overall picture and gives an outlook to future developments.

2. METROLOGY AND THE INTERNET OF THINGS

In the concept of the Internet of Things (IoT) a set of physical devices communicate with each other via web technologies. With the rise of the IoT in Industry 4.0, Smart City, smart grids and more, the world of measurement is changing rapidly. As an example, the integration of measuring instruments in the IoT poses several specific requirements for the sensors itself, such as communicating via a digital interface, working reliably under a wide range of conditions and ideally, detecting and reporting adversarial conditions, as well as reporting on their health status upon request. These and other requirements have led to the development of so-called "smart sensors" [2]. These are measuring devices that contain some sort of pre-processing, which in turn poses new requirements for the calibration. Furthermore, measuring instruments which only provide pre-processed data usually don't fit well in today's calibration procedures and guidelines, because these assume access to the raw measurement data. A concrete challenge addressed in the project EMPIR Met4FoF was the dynamic calibration of a digital-output sensor using an external time stamping, e.g., based on GPS and a custom-built microcontroller (µC) board [4]. The same approach was then used to demonstrate the extension of a digital sensor such that it communicates not only raw measurement values, but also provides information about the measurement units, uncertainty, and calibration in a machine-readable way [4-6]. In this way, the most basic requirement of a metrological treatment of a sensor network can be met: the provision of measurement uncertainty and other metrological information for the individual sensors. In the concept developed in the Met4FoF project this information is provided by the "smart" sensor itself. However, other information architectures are possible, too. For instance, in the BMBF FAMOUS project a database approach combined with OPC-UA communication was considered instead. A similar approach is also discussed in the project BMWK GEMIMEG-II. More details are given in Sections 3 and 4.

The concept of sensors providing self-information upon request in a standardized way is also a fundamental element of OPC-UA, which is used mostly in industrial applications, but is increasingly adopted in other areas, too. For the metrological information communicated via OPC-UA to be machine-readable, it is necessary that the definition of a standard digital representation of units of measurement as well as commonly accepted data models for measured values are available. To this end, the D-SI data model developed in EMPIR SmartCom proposes an approach that is compatible with current guidelines and standards in metrology and calibration [5]. Other potential approaches for the digital representation of units of measurement and quantity kinds are UCUM or QUDT [7,8] – each optimized for different data usage approaches.

The concept of the IoT relies on a versatile and flexible combination of measuring instruments, the automated acquisition and processing of the measured data and the application of intelligent algorithms to derive conclusions or decisions. One consequence of this is that data analysis is typically carried out using data-driven machine learning. In contrast to mathematical models that rely on a physical understanding of the measured process, machine learning can be applied directly based on the sensors' output data. Thus, the need for calibration is not as obvious as for "traditional" measurements. However, calibrated measuring instruments in the IoT offer several benefits. For instance, calibrated sensors can serve as reference devices in the network to assess and improve data quality [9]; calibration of sensors enables the estimation of the measurand and thus, traceability [6], which itself is required to ensure the comparability of measurements between different sites and countries. Moreover, calibrated sensors improve the ability to explain the obtained output from the machine learning. That is, the calibration of a sensor enables direct interpretation based on the measurand whereas a non-calibrated sensors provides data streams which are only loosely related to the physical measured quantity. Moreover, the manufacturer's data sheet alone usually doesn't suffice as source of information to assess the type B uncertainty components. Hence, calibration plays an important role in IoT and provides benefits on all data processing layers and a way to quantify the trust one can have in the measurement system, see Fig. *1*.

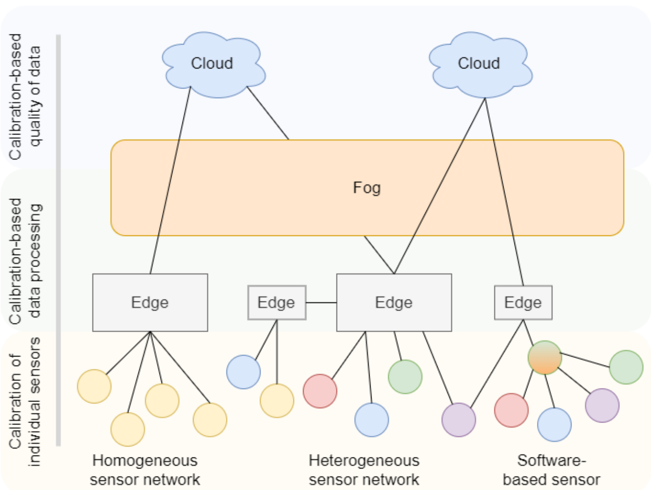

Fig. 1 Calibration information in the different layers of the IoT architecture.

3. DIGITAL REPRESENTATION OF DATA AND INFORMATION IN SENSOR NETWORKS

In the digital world, measurement data must also be readable and understandable by machines. This implies that the information about the measuring instruments, the units of measurement and other accompanying metadata must be available in a format that can be used by software or an algorithm. For instance, the software may need to verify that the unit of measurements of a given data set is consistent with previous entries of a data base.

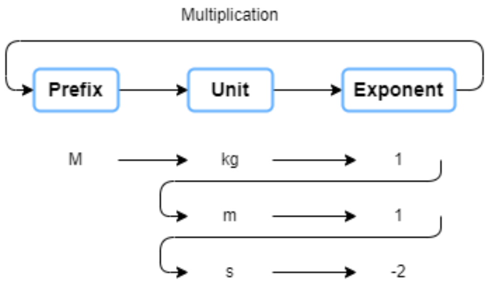

Fig. 2 Example for an algorithmic representation of units of measurement, which can be used by a software

The machine readability of data and measurement information is of particular importance in sensor network metrology. With hundreds of sensors measuring continuously, measurement data cannot be normalized and analysed manually, but requires a high level of automation. This, in turn, can only be achieved with machine-readable information. The machine readability begins with the description of the unit of measurement, for instance as shown in Fig. 2. Another important element is the information about the individual measuring instruments. For instance, a machine-readable digital calibration certificate [10] could be provided by the sensor itself, e.g., using OPC-UA, or from another source, e.g., an internal quality management platform. If needed, this information could be further extended by other data which could affect the quality of measurement data [11].

Machine-understandable representation of knowledge on information in sensor networks and its semantics can be modelled by means of (combinations of) ontologies. Several ontologies useful for sensor networks can be found within the semantic web community. These ontologies formalize the annotation of sensor data with spatial, temporal, and thematic metadata. Spatial metadata is particularly relevant for sensors distributed across a building or a country, or when mounted on a moving object like an automobile.

In the project FAMOUS a method to merge different kinds of metadata and ontologies along with the sensor measurement data was proposed [12]. The main idea of [12] is to split the self-description of a sensor into four aspects: (1) observation information, (2) general sensor description, (3) calibration information and (4) location information. Then, a sensor self-description can be achieved by combining existing ontologies that appropriately represent these aspects. In this way, one can build upon existing work and established principles and software tools. In the GEMIMEG-II project this development was extend to the integration of semantically described quality of data aspects [11].

4. MEASUREMENT UNCERTAINTY IN SENSOR NETWORK DATA PROCESSING

Sensor networks in IoT often contain various low-cost measuring instruments based on MEMS sensor technology. Consequently, sensor networks typically have a wide range of measurement data quality. Reliable data analysis for sensor networks thus requires taking data quality into account in a quantitative way.

An example of a fundamental quantity that can be considered as a quality metric is the measurement uncertainty. Other examples for parameters which may affect the quality of data in sensor networks are unstable network conditions, environmental interference, or malicious attacks. Moreover, in battery powered sensors, there is necessarily a trade-off between power-consumption and performance. Another common issue is drift, where sensor readings slowly deviate from the true value due to the degradation of the electronics [13]. One outcome from the GEMIMEG-II project is a framework for quality of data in sensor networks, expressed in terms of an ontology [11]. In a joint effort by the projects FAMOUS and EMPIR Met4FoF it was demonstrated how such an ontology could be utilized for an automated analysis in sensor network metrology [6,12]. Moreover, together with the project EMPIR SmartCom a sensor network data set was enriched with machine-readable metadata to demonstrate the metrological support of the FAIR principles [14].

Usually, the sensors used in IoT applications are measuring continuously irrespective of how the measured values are used. Thus, for a reliable quantification of data quality it must be ensured that the sensor behaviour is well known for a wide range of measurement situations. This includes situations where the measurand, i.e., the sensor input signal, changes rapidly over time. Thus, sensor properties such as effective bandwidth, internal analogue-to-digital conversion, time stamping reliability, resonance behaviour need to be considered.

Data analysis in the IoT typically relies on and greatly benefits from modern machine learning methods, because of the complexity of the sensor network and the amount of data acquired. Uncertainty evaluation for machine learning is an important topic and considered in several research activities. However, this is only possible when the uncertainty associated with the machine learning input values is available. Hence, uncertainty for data pre-processing must be addressed as the initial step for an uncertainty-aware machine learning for IoT.

Measurements in the IoT are usually time dependent, and often even dynamic. Examples are air quality monitoring, traffic surveillance, production control or mobile health measurements. Thus, signal processing methods are regularly applied for data pre-processing in IoT scenarios. For instance, the discrete Fourier transform is often applied to extract magnitude and phase values from a measurement of vibration, which are then used in a subsequent machine learning method as features. Other examples for pre-processing are synchronising the time axis of sensors; interpolation of sensor data to account for missing values or non-equidistant sampling; low-pass filtering to reduce noise or other unwanted high-frequency components in the measured data. Another reason for the application of data pre-processing is the reduction of data dimensionality. This may be necessary simply due to storage or data transfer bandwidth limitations

[3]. In the project EMPIR Met4FoF, the previously developed Python library *PyDynamic* [15] was extended to include the data pre-processing steps typically required in IoT. For each method, *PyDynamic* provides the propagation of uncertainties [15]. An important aspect in EMPIR Met4FoF and in FAMOUS was also the implementation of the methods in such a way that they can be applied online, i.e., during the measurement. For instance, the Discrete Wavelet Transform for uncertain input data was implemented using digital filters [16].

Another important aspect is the way of how the uncertainty propagation software is provided such that it is compatible with typical IoT architectures. In the project EMPIR Met4FoF, a so-called agent-based framework (ABF) was used [17]. In an ABF, data processing steps are encapsulated in software modules, called "agents". These agents can run on different locations in the network, if necessary, and allow for a very flexible demand-driven data analysis. For instance, one agent may acquire the data from a sensor, hands it over to an interpolation agent, which then provides it to a Fourier transform agent. With each agent taking care of the proper uncertainty treatment, very flexible data analysis pipelines for sensor network metrology can be realised. Usually there is an existing data analysis framework in place, which needs to be extended to include measurement uncertainty treatment. As a result, a web-service approach was used in the project FAMOUS instead of an ABF. That is, an uncertainty module was created to enrich existing sensor data streams with statements about the associated measurement uncertainty.

CONCLUSIONS AND OUTLOOK

Sensor network metrology combines several aspects from metrology, signal processing, semantics, IoT and web technologies. The treatment and metrological assessment of sensor networks, thus, needs to take these fields into account. Although sensor networks can be found in many applications, a rigorous sensor network metrology hasn't been established yet. Existing guidelines in metrology are typically focused on individual measuring instruments and quantities. The same holds true, by the way, for the organisation of metrology institutes and calibration laboratories. First metrology research efforts developed some basic elements required in sensor network metrology: dynamic calibration of digital sensors, cost-efficient calibration of MEMS sensors, digital representation of metrological metadata, evaluation and propagation of uncertainties, semantic modelling of sensor network information.

Future research needs to further develop these individual aspects and extend their integration into a consistent framework and toolset. Moreover, a systems metrology approach needs to be developed to assess sensor networks in a systemic way.

ACKNOWLEDGMENTS

Parts of this work was based on outcomes of the projects 17IND12 Met4FoF, 17IND02 SmartCom, BMBF FAMOUS and the BMWK GEMIMEG-II project. The 17IND12 Met4FoF and 17IND02 SmartCom projects have received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon2020 research and innovation programme.

The FAMOUS project received funding from the German Federal Ministry of Education and Research (BMBF). The GEMIMEG-II project received funding from the German Federal Ministry of Economy and Climate (BMWK).

REFERENCES

[1] S. Eichstädt, Metrology for the digital age, in De Gruyter Book Series on Measurement Science, Volume "Metrological Infrastructure". In press

[2] T. Schneider, N. Helwig, A. Schütze, Industrial condition monitoring with smart sensors using automated feature extraction and selection, in: Meas. Sci. and Technol., 29(4), 094002, 2018

[3] T. Dorst, T. Schneider, A. Schütze, S. Eichstädt, GUM2ALA–Uncertainty Propagation Algorithm for the Adaptive Linear Approximation According to the GUM, in: SMSI 2021

[4] B. Seeger and Th. Bruns, Primary calibration of mechanical sensors with digital output for dynamic applications, in:

[5] D. Hutzschenreuter et al., SmartCom Digital System of Units (D-SI) Guide for the use of the metadata-format used in metrology for the easy-to-use, safe, harmonised and unambiguous digital transfer of metrological data - Second Edition, 2020, DOI: 10.5281/zenodo.3816686

[6] S. Eichstädt, M. Gruber, A. P. Vedurmudi, B. Seeger, Th. Bruns and G. Kok, Toward Smart Traceability for Digital Sensors and the Industrial Internet of Things, in: Sensors, 21(6), 2021

[7] G. Schadow, C. J. McDonald, The unified code for units of measure, in: Regenstrief Institute and UCUM Organization: Indianapolis, IN, USA, 2009

[8] H. Rijgersberg, M. Van Assem and J. Top, Ontology of units of measure and related concepts, in: Semantic Web, 4(1), pp. 3-13, 2013

[9] G. Tancev, F. Grasso Toro, Sequential recalibration of wireless sensor networks with (stochastic) gradient descent and mobile references, in: Measurement: Sensors, vol. 18, 2018

[10] S. Hackel et al., The fundamental architecture of the DCC, in: Proceedings of the XXIII IMEKO World Congress 2021

[11] A. P. Vedurmudi, J. Neumann, M. Gruber and S. Eichstädt, Semantic Description of Quality of Data in Sensor Networks, in: Sensors, 21(9), 6462, 2021

[12] M. Gruber, S. Eichstädt, J. Neumann and A. Paschke, Semantic Information in Sensor Networks: How to Combine Existing Ontologies, Vocabularies and Data Schemes to Fit a Metrology Use Case, in: proceedings of IEEE International Workshop on Metrology for Industry 4.0 & IoT 2020

[13] K. Goebel and W. Yan, Correcting Sensor Drift and Intermittency Faults With Data Fusion and Automated Learning, in: IEEE Systems Journal, 2, 189–197, 2008

[14] T. Dorst, M. Gruber, A. P. Vedurmudi, Sensor data set of one electromechanical cylinder at ZeMA testbed (ZeMA DAQ and Smart-Up Unit), in: Zenodo, DOI: 10.5281/zenodo.5185952

[15] S. Eichstädt, C. Elster, I. M. Smith, T. J. Esward, Evaluation of dynamic measurement uncertainty – an open-source software package to bridge theory and practice, in: 6, 97-105, 2017, DOI: 10.5194/jsss-6-97-2017

[16] M. Gruber, T. Dorst, A. Schütze, S. Eichstädt and C. Elster, Discrete wavelet transform on uncertain data: Efficient online implementation for practical applications, in: Advanced Mathematical and Computational Tools in Metrology and Testing XII, pp. 249-261, 2022

[17] Eichstädt S. (2021), Metrology for the Factory of the Future. Research Outreach, Available at: https://researchoutreach.org/articles/metrology-for-the-factory-of-the-future/ (Accessed 2022/04/06)