MathMet2019

Portugal | Lisbon | LNEC
20-22 November, 2019

Book of abstracts

LABORATÓRIO NACIONAL DE ENGENHARIA CIVIL

Instituto Português da Qualidade

RELACRE
ASSOCIAÇÃO DE LABORATÓRIOS ACREDITADOS DE PORTUGAL

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin

Table of contentes

Chair Welcome Message

As chairman of the 2019 MATHMET International Workshop, it is an honour and a privilege to welcome you to Lisbon, Portugal. Lisbon is the city where in the XV century Portuguese sailors set out to give new worlds to the world, starting the first economic globalization. This wonderful cosmopolitan city, full of history, existing long before Roman times, was chosen to host the 2019 MATHMET International Workshop, to be held on November 20-22, 2019.

The event is jointly organized by IPQ – Portuguese Institute for Quality, LNEC – National Laboratory for Civil Engineering, RELACRE, Portuguese Network of Accredited Laboratories, MATHMET – The European Centre for Mathematics and Statistics in Metrology and PTB – Physikalisch-Technische Bundesanstalt. It has the support of the Portuguese Societies of Metrology, Mathematics, Statistics, Materials, Engineers and Chemistry.

The Symposium website is available at http://mathmet2019.lnec.pt/ where you will find useful details concerning the scientific programme, along with information about the accommodation, venue and social events. This website is the principal communication channel for the conference, so be sure to visit it often.

The Executive Committee, the Scientific Committee and the National Organizing Committee will do their best to offer you a Workshop at least as successful as the previous ones, held in Berlim, Germany, in 2016, 2014 and 2010. The aim is to provide a forum for applied mathematicians, statisticians, and metrologists to present and discuss contemporary methods and challenges in applications of mathematical models and statistical data analysis to measurement science, including uncertainty quantification, interlaboratory studies, medical and industrial imaging, atmospheric science and climatology, chemometrics, molecular biology, machine learning, dynamic measurements and big data. I do hope that such goals will contribute to the continued effort towards excellence in MATHMET events.

It is my special pleasure to host this MATHMET event. I hope that you will find it technically fulfilling and highly entertaining and that it will be an opportunity for useful interactions and communications with colleagues from all over the world, to network in sessions on theoretical and applied science, bringing together experts from related fields of knowledge. Beyond the technical and scientific aspects of the Symposium, I do hope that you will have the opportunity to enjoy Lisbon and its surroundings as well as the many cultural and recreational activities available in Portugal. A warm welcome to all of you coming to Lisbon, Portugal, for the MATHMET International Workshop.

Tony O'Hagan

Tony O'Hagan is Emeritus Professor of Statistics at the University of Sheffield, UK. His research in the methodology and applications of Bayesian Statistics is internationally recognised and has influenced practice in many other fields, including medicine, engineering, numerical analysis, health economics and environmental science. He has had some collaboration with metrologists, particularly with members of the GUM working party, since 2012.

A Solid Foundation for the Expression of Uncertainty in Measurement

Abstract: In any structure or enterprise, it is important to ensure that the foundations are solid. We can go on adding more fancy elements but if the foundations are weak then at some point cracks will appear.

Metrology is a fascinating field, the GUM is a remarkable document and some truly outstanding technical methods have been built on this foundation, but it is a mess. So this talk is about rebuilding the foundations and making them solid and fit for purpose. Any such rebuilding must ask fundamental questions. What do we mean by (the) measurement? What is the meaning of standard uncertainty, and is it fit for purpose?

My answers may be a little radical, so it is important to ask: are they realistic? The metrology community has already rejected change once because they didn't like, and didn't accept the need for, the practical consequences. But I will try to show that my proposals can be workable for testing labs, just as much as for NMIs.

Wendy Parker

Wendy Parker is Associate Professor of Philosophy and Co-Director of the Institute for Data Science (IDAS) at Durham University, UK. Her research examines the practices of contemporary science, especially meteorology and climate science, with a particular focus on the practice of computer simulation. Her work has been published in a variety of philosophical and scientific journals. She is currently Co-Editor-in-Chief of The British Journal for the Philosophy of Science.

Can we measure via computer simulation?

Abstract: In various fields, scientists now speak of “observing” or “measuring” the world via computer simulation. Some metrologists also now characterize measurement as an experimental or computational process of a particular sort. Does allowing that we can measure via computer simulation mark a radical shift in our understanding of measurement? Are there special conceptual or practical challenges associated with computational measurement? This talk will explore these and related questions, drawing on recent work in philosophy of science and examples from particular sciences.

Nicolas Fischer

Nicolas Fischer is a principal researcher in statistics and is head of the Data Science and Uncertainty department at LNE. He has more than ten years of experience as a statistician within LNE. Since 2010 he is responsible for the mathematical and statistical research program for metrology. His research mainly concerns the methods for evaluating the uncertainty of measurement and processing of interlaboratory comparisons. He provides numerous training courses in data analysis, sampling, quality control and evaluation of uncertainty. Since September 2014, he became a member of the Joint Committee for Guides in Metrology, Working group1 (JCGM-WG1) of the BIPM, international expert group, which produces and maintains the Guide to the expression of uncertainty in measurement. He is also involved in several scientific bodies (Société Française de Statistique, Institut de Maitrise des risques) as group leader within the french statisticians' community.

Characterization in size of aggregated nanoparticles measured by SEM: an illustration of deep generative models in metrology

Abstract: Recent advances in deep generative models based on convolutional neural networks (CNNs) are used to demonstrate the potential of these approaches for the estimation of particle size distribution on images of aggregated TiO2 particles obtained by Scanning Electron Microscopy (SEM).This very promising framework shall permit effective automation of SEM measurements analysis. Indeed, common image processing software bring the end-users with segmentation algorithms as well as measuring tools to estimate individual particle diameters. In the case of aggregated nanoparticles, most particles suffer missing contents and are not considered in the computations. In this work, we have used a recently developed method called "context encoders" to predict missing parts of the nanoparticles. The approach is tested against simulated and real dropped image regions.

Finally Consideration is made to evaluate the performance of the method based on both real and simulated particles using cross validation.

Clemens Elster

Clemens Elster is leading PTB's Working Group "Data Analysis and Measurement Uncertainty" since 2004. His research interests are mainly in the field of statistical data analysis with focus on Bayesian methods, and he has co-authored more than 100 papers in peer-reviewed journals.

Clemens Elster received his diploma with a thesis on developing a regularization method for solving an ill-posed inverse problem from the Faculty of Physics at the University of Freiburg, Germany, in 1990, and his PhD on designing methods for experimental design and the optimization of noisy functions in 1993 also from Freiburg University. He joined the Physikalisch-Technische Bundesanstalt (PTB), Braunschweig and Berlin, Germany, in 1994. Since 2004 he is leading the Working Group "Data Analysis and Measurement Uncertainty". In 2010 he also became a member of the JCGM Working Group on the Expression of Uncertainty in Measurement (GUM).

Application of Gaussian Markov random field priors for Bayesian spatial modeling

Abstract: Gaussian Markov Random Field (GMRF) priors are a popular tool in the Bayesian inference of spatially distributed parameters whose variation is expected to be smooth. Examples of applications comprise the analysis of functional magnetic resonance imaging or the inference of electron density of earth's upper atmosphere in the geosciences.

Applications involving GMRF priors are often high-dimensional which challenges the numerical calculation of the results of a Bayesian inference. This contribution reviews the concepts of GMRF priors and presents two examples of applications. The examples are high-dimensional and different approximation techniques are employed, including the use of Laplace-type approximations and approximate analytical expressions facilitating the sampling from the posterior.

Carola-Bibiane Schönlieb

Carola-Bibiane Schönlieb is Professor of Applied Mathematics at the Department of Applied Mathematics and Theoretical Physics, University of Cambridge. There, she is head of the Cambridge Image Analysis group, Director of the Cantab Capital Institute for Mathematics of Information, Director of the EPSRC Centre for Mathematical and Statistical Analysis of Multimodal Clinical Imaging, and a fellow of Jesus College Cambridge. Her current research interests focus on variational methods, partial differential equations and machine learning for inverse imaging problems. Her research has been acknowledged by scientific prizes, among them the LMS Whitehead Prize 2016, and by invitations to give plenary lectures at several renowned applied mathematics conference, among them the SIAM conference on Imaging Science in 2014, the SIAM annual meeting in 2017, the Applied Inverse Problems Conference in 2019 and the GAMM in 2020.

In her research she is interested in both the rigorous theoretical and computational analysis of the problems considered as well as their practical implementation and their use for real-world applications. She has active interdisciplinary collaborations with clinicians, biologists and physicists on biomedical imaging topics, chemical engineers and plant scientists on image sensing, as well as collaborations with artists and art conservators on digital art restoration.

From shallow to deep learning for inverse imaging problems: some recent approaches

Abstract: In this talk we discuss the idea of data-driven regularisers for inverse imaging problems. We are in particular interested in the combination of model-based and purely data-driven inversion approaches. In this context we will make a journey from “shallow” learning for computing optimal parameters for variational regularisation models by bilevel optimization to the investigation of different approaches that use deep neural networks for solving inverse imaging problems. The talk is furnished with application of these ideas to medical imaging, in particular computed tomography.

Peter Harris

Dr Peter Harris has worked at NPL since 1986 in the areas of mathematical and statistical modelling and experimental data analysis, uncertainty evaluation, algorithm design, numerical software development and software testing applied to a wide variety of measurement problems. He received a BSc in Mathematics from the University of Bath in 1986 and his PhD on spline approximation from the University of Brunel in 1991. He is a Principal Research Scientist in the Data Science Group at NPL, currently involved in projects on using sensor networks for environmental and climate monitoring, and on using data within the “factory of the future” and the “digitally-enabled supply chain” in support of advanced manufacturing.

Evaluating long-term trends in underwater noise in the Southern Ocean

Abstract: Underwater noise is classified as a form of pollution by international regulation, and there is increasing understanding of the effects such noise can have on the wellbeing of marine ecosystems. In this work we describe a method for performing long-term trend analysis of deep-ocean noise data measured by the hydro-acoustic monitoring stations of the Preparatory Commission for the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO). The analysis method uses a flexible discrete model that incorporates terms that capture seasonal variations in the data together with a moving-average statistical model to describe the serial correlation of residual deviations. The results show that statistically significant reductions in deep-ocean noise are observed at some of the monitoring stations. Strong seasonal variation in the recorded data is also observed, with a high degree of correlation with climatic factors such as sea surface temperature and Antarctic ice coverage. Some possible explanations for the observed behaviour are presented.

Blaza Toman

Blaza Toman is a member of the Statistical Engineering Division at the National Institute of Standards and Technology in the USA. Prior to this she taught statistics at Rutgers University and at George Washington University. She earned a PhD in statistics from Ohio State University and is interested mainly in Bayesian statistical methods as applied to uncertainty evaluation and design of experiments. Her recent publications include a new realization of the international system of units for organic chemical measurement, and a new procedure for interlaboratory studies and meta-analysis involving the novel concept of shades of dark uncertain

Design of a calibrated experiment for quantitation of a chemical composition

Abstract: Efficient experimental design is a critical aspect of practical scientific planning and measurement execution. Achieving fit-for-purpose measurement results using limited resources is a significant priority for laboratories. For new applications or procedures, specification of such an experimental design may not be straightforward or readily informed by comparable precedent. In this presentation we show how to optimally construct fit-for-purpose measurement schemes that achieve appropriate confidence. Specifically, we plan a two-stage experiment with a calibration phase followed by the measurement of an unknown. One such procedure is the determination of mass fraction of an analyte species *A* via LC-IDMS, using an isotopically-enriched internal standard *I*. An experimental design for the two phased procedure consists of the following quantities: the number of calibration standards n_I, the number of replications of the measurements for each calibration standard n_J, the set of nominal values of the standards $(\theta_1,\ldots,\theta_I)$, the number of samples of the unknown in the second experiment n_q, and the number of replicates per sample n_s. We will show how to select the experimental design $D= (n_I, n_J, (\theta_1,\ldots,\theta_I), n_q, n_s)$ which is locally optimum, and guarantees that the expected relative measurement uncertainty is at most p %. We will demonstrate, using our software App, the experimental design procedure on a specific example of measurement of mass fraction of 25(OH)D3 in serum.

Juris Meija

Juris Meija is a senior research officer at the National Research Council Canada whose research is aimed to improve the reliability of chemical measurements through development of certified reference materials and better understanding of the measurements themselves. His expertise lies in theoretical analytical chemistry, isotope ratio measurements, and data analysis. He serves as the Chair of the IUPAC Commission on Isotopic Abundances and Atomic Weights and is also IUPAC delegate to the Joint Committee for Guides in Metrology Working Group. He has been actively involved in many recent international activities such as the redefinition of the mole, naming of the new chemical elements, and revisions of the GUM.

Traceability in chemical measurements: the role of data analysis

Abstract: Countless chemical measurements are performed worldwide each day. While the results of a chemical measurement are determined by many crucial components such as the primary standards, choice of measurement methods, or the act of measurement, choice of the measurement model and its consequences is often less appreciated. This talk will address the importance to distinguish between the measurements and the measurement results. In this vein, chemical measurements cannot be performed without the recourse to data analysis, mathematical or statistical measurement models. These models, and their implementation, form an integral part of the measurement process and, much like the physical act of measurement, can lead to errors. This presentation will feature a variety of examples from traditional chemical analyses, including the titration, standard additions, and isotope dilution, showing that choices on how we interpret and model our measurements often have significant effects. The challenge is therefore for the analysts to explore the rich variety of modeling options and recognize that larger statistical toolkit can raise the bar for more reliable results.

Maria Antónia Turkman

Maria Antónia Amaral Turkman was, until 2013, full-time Professor in the Department of Statistics and Operations Research, Faculty of Sciences, University of Lisbon where she has taught courses on Bayesian Statistics and Computational Statistics, among many others. Though retired from the university, she is still an integrated member of its Centre of Statistics and Applications, where she held the position of scientific coordinator until 2017. Her research interests are Bayesian Statistics, Medical and Environmental Statistics, and Spatiotemporal Modelling, with most recent publications on computational methods in Bayesian statistics, with an emphasis on applications in health and forest fires. Her most recent book "Computational Bayesian Statistics: An Introduction", co-authored by Carlos Daniel Paulino and Peter Muller, was published in February 2019 by Cambridge University as the first text book on a series of the Institute of Mathematical Statistics (IMS) with International Society of Bayesian Statistics (ISBA). She was a founding member of the Portuguese Statistical Society (SPE), has served as vice president of the Society and is currently Chairman of the SPE General Assembly.

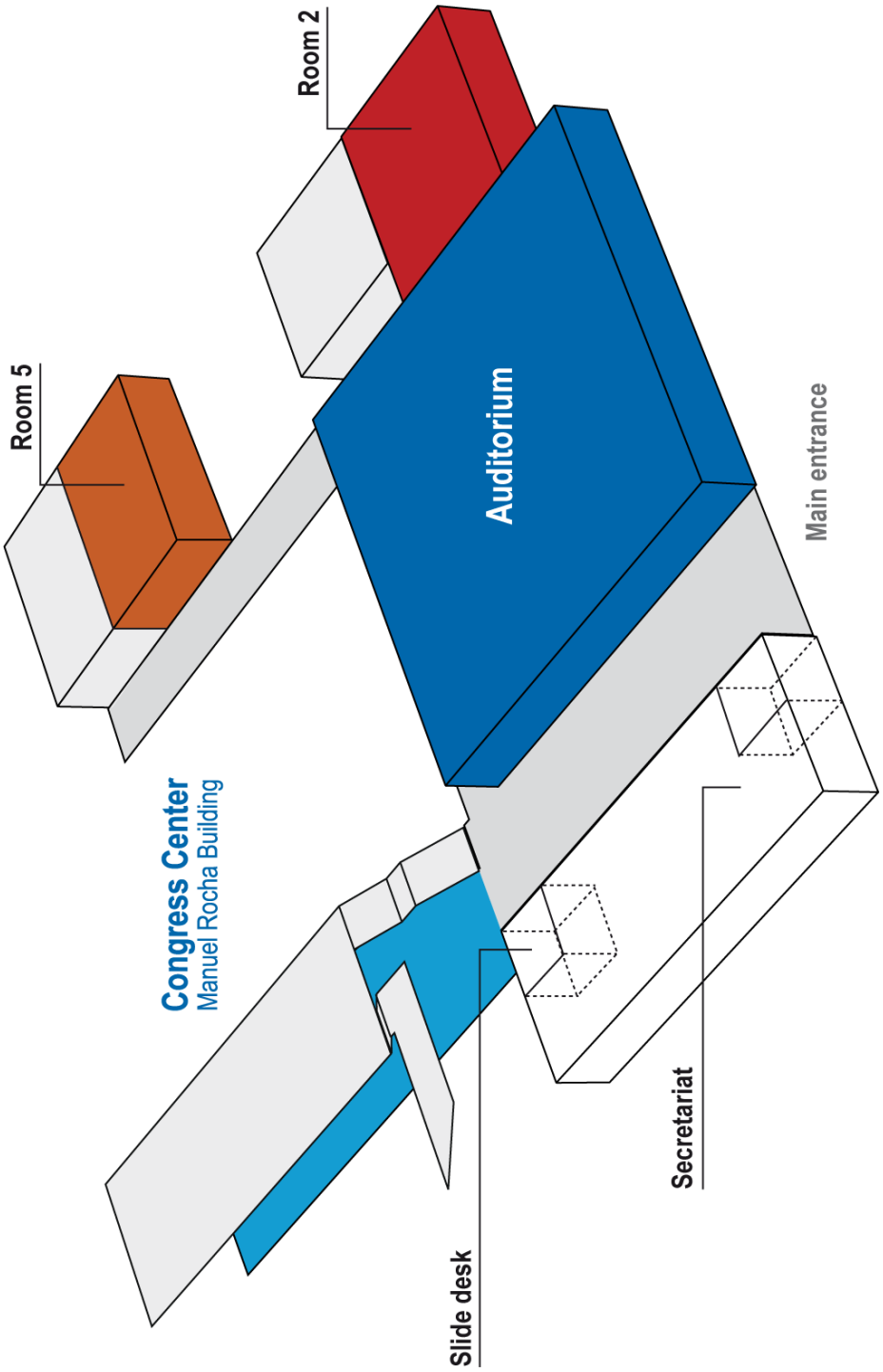
Data fusion/calibration methods to update simulated data based on the observed data: an application to wind speed data

Abstract: Extreme values of certain spatio-temporal processes, such as wind speeds, are the main cause of severe damage in property, from electricity distribution grid to road and agricultural infrastructures. Accurate assessment of causal relationships between environmental processes and their effects on risk indicators, are highly important in risk analysis, which in return depends on sound inferential methods as well as on good quality informative data. Often, information on the relevant environmental processes comes from monitoring networks, as well as from numerical-physical models (simulators) that typically solve a large set of partial differential equations, capturing the essence of the physical process under study.

Overall Schedule

	Tuesday // 19 November //	
18:30 – 20:00	Welcome Reception	
	Wednesday // 20 November //	**Chairperson**
8:45 – 9:15	Registration	
9:15 – 9:45	Opening Ceremony	
9:45 – 10:30	Invited Speaker: Tony O'Hagan	Maurice Cox
10:30 – 11:15	Invited Speaker: Wendy Parker	António Possolo
11:15 – 11:45	Coffee break	
11:45 – 13:15	Parallel session 1 (Measurement Uncertainty 1)	Tony O'Hagan
	Parallel session 2 (Machine Learning and Dynamic Measurements)	Sascha Eichstaedt
13:15 – 14:15	Lunch	
14:15 – 15:00	Invited Speaker: Nicolas Fischer	Markus Baer
15:00 – 15:45	Invited Speaker: Clemens Elster	Gertjan Kok
15:45 – 16:15	Coffee break	
16:15 – 18:00	Parallel session 3 (Modelling and Inverse Problems)	Clemens Elster
	Parallel session 4 (Measurement Uncertainty Training)	Katy Klauenberg
	Thursday // 21 November //	**Chairperson**
9:15 – 10:00	Invited Speaker: Carola Schönlieb	Francesca Pennecchi
10:00 – 10:45	Invited Speaker: Peter Harris	Álvaro Ribeiro
10:45 – 11:45	Coffee break + Poster session	
11:45 – 13:00	Parallel session 5 (Uncertainty Quantification for Computationally Expensive Models)	Peter Harris
	Parallel session 6 (Mathematics and Metrology in Medicine 1)	Carola Schonlieb
13:00 – 14:15	Lunch	
14:15 – 15:00	Invited Speaker: Blaza Toman	Olivier Pellegrino
15:00 – 15:45	Invited Speaker: Juris Meija	Alen Bosnjakovich
15:45 – 16:15	Coffee break	
16:15 – 17:35	Parallel session 7 (Statistical Methods for Interlaboratory Comparisons and Conformity Assessment)	Blaza Toman
	Parallel session 8 (Statistical Calibration and Regression Problems)	Juris Meija
19:30 – 23:30	Conference Dinner	
	Friday // 22 November //	**Chairperson**
9:30 – 11:10	Parallel session 9 (Mathematics and Metrology in Medicine 2)	Nicolas Fischer
	Parallel session 10 (Measurement Uncertainty 2)	Antónia Turkman
11:10 – 11:45	Coffee break	
11:45 – 12:30	Invited Speaker: Antónia Turkman	João A. Sousa
12:30 – 13:00	Invited Speaker: Emma Woolliams	João A. Sousa
13:00 – 13:15	Closing	
13:15 – 14:15	Lunch	
14:15 – 18:30	Special session on Introduction to Machine Learning for Metrology Apllications	
14:15 – 18:30	MATHMET meeting	

Congress Center Plan



Themes

- Measurement Uncertainty
- Statistical Calibration and Regression Problems
- Modeling and Inverse Problems
- Uncertainty Quantification for Computationally Expensive Model
- Statistical Methods for Interlaboratory Comparisons and Conformity Assessment
- Statistical Methods for Chemistry, Bioanalysis and Molecular Biology
- Machine Learning
- Dynamic Measurement
- Mathematics and Metrology in Medicine
- Measurement Uncertainty Training: A Survey and Developments
- Type A evaluations of measurement uncertainty

Tuesday // 19 November //

18:30 – 20:00 Welcome Reception

Wednesday // 20 November //

8:45 Registration

9:15 Opening Ceremony

9:45 INVITED SPEAKERS **AUDITORIUM**
Tony O'Hagan / Wendy Parker
Chairperson: Maurice Cox / António Possolo

11:15 Coffee break

11:45 PARALLEL SESSION 1 **ROOM 2**
Measurement Uncertainty 1
Chairperson: Tony O'Hagan

11:45 1.1 Asymmetrical uncertainties
ID1005 – Antonio Possolo

12:05 1.2 Evaluating the uncertainty in the measurement of nanoparticle size by means of SEM and DLS
ID1007 – Ignacio Lira

12:25 1.3 Redundant information in sensor networks and uncertainty quantification
ID1011 – Gertjan Kok

12:55 1.4 Uncertainty budget for gas mixtures preparation by dynamic dilution and subsequent use in the calibration of analytical instrumentation
ID1015 – Francesca Pennecchi

11:45 PARALLEL SESSION 2 **ROOM 5**
Machine Learning and Dynamic Measurements
Chairperson: Sascha Eichstaedt

11:45 2.1 Mathematical framework for metrology in the factory of the future
ID1034 – Sascha Eichstaedt

12:05 2.2 Metrology for virtual measuring systems: new competence centre: "VirtMet" at PTB
ID1033 – Sascha Eichstaedt

12:25 2.3 Influence of synchronization within a sensor system on machine learning results
ID1004 – Tanja Dorst

12:55 2.4 On the influence of inlet perturbations on the development of slugs in horizontal two-phase flow
ID1043 – Sonja Schmelter

13:15 Lunch

14:15 **INVITED SPEAKERS** **AUDITORIUM**

Nicolas Fischer / Clemens Elster

Chairperson: Markus Baer / Gertjan Kok

15:45 Coffee break

16:15 **PARALLEL SESSION 3** **ROOM 2**

Modelling and Inverse Problems

Chairperson: Clemens Elster

16:15 3.1 Recovery of smooth low-rank matrices using Bayesian inference
ID1009 – Gerd Wübbeler

16:35 3.2 Large scale inference with applications to environmental monitoring
ID1054 – Louis Sharrock

16:55 3.3 Development of a virtual flow meter
ID1051 – Andreas Weissenbrunner

17:15 3.4 Model-based determination of optical and geometrical properties of red blood cells from light scattering
ID1046 – Markus Bär

17:35 3.5 Targeted high-fidelity data to enrich surrogate models for uncertainty quantification in climate prediciton
ID1062 – Oliver Dunbar

16:15 **PARALLEL SESSION 4** **ROOM 5**

Measurement Uncertainty Training

Chairperson: Katy Klauenberg

16:15 4.1 Tutorial for a Bayesian evaluation of measurement uncertainty and its implementation
Séverine Demeyer

16:36 4.2 Measurement uncertainty training at PTB
Katy Klauenberg

16:46 4.3 Measurement uncertainty training at NPL
Peter Harris

16:55 4.4 Measurement uncertainty training at LNE
Michèle Désenfant

17:04 4.5 Measurement uncertainty training at METAS
Marc-Olivier André

17:13 4.6 Measurement uncertainty training: experiences at the INRIM and ideas for developing a dedicated international community
Francesa Pennecchi

17:22 4.7 Methodology of teaching the concept of measurement uncertainty
Anna Chunovkina

17:30 4.8 Discussion on future developments
All

Thursday // 21 November //

8:45 Registration

9:15 INVITED SPEAKERS **AUDITORIUM**

Carola Schönlieb / Peter Harris

Chairperson: Francesca Pennecchi / Álvaro Ribeiro

10:45 Coffee break + Poster session

POSTERS

Calculating coverage factors and coverage probabilities in study cases
ID1044 – Anna Chunovkina

Modelling of flow time series as an approach to compute its uncertainty
ID1036 – Maria Silva

Towards a new GUM – the suggested draft
ID1022 – Igor Zakharov

Indirect multiparameter measurements with correlated uncertainties
ID1021 – Jacek Puchalski

Estimation of the uncertainty in selected points of measured function from two control measurements
ID1035 - Jacek Puchalski

Testing software that implements least-squares fitting for nonlinear models
ID1024 – Ian Smith

Confidence regions for parameters in two-dimensional linear comparative calibration model
ID1013 – Gejza Wimmer

Mathematical model of the volume of no-OIML R111-1 standard weights
ID1059 – Omar-Jair Purata-Sifuentes

Musing on modelling in measurement science: from the perspective model to the descriptive model implementing the former in experiments
ID1058 – Franco Pavese

Efficient reliability analysis with model reduction techniques
ID1052 – Jörg Unger

Binary linear regression in dynamic force measurement and uncertainty estimation
ID1056 – Jun Yang

Research on analysis of amplitude-frequency response in dynamic pressure calibration with shock tube
ID1057 – Jun Yang

A statistical metrology approach to compare the quality of periodic cardiovascular waveforms
ID1055 – Janos Palhalmi

Implementation of GUM principles in Western Balkan countries
ID1047 – Alen Bosnjakovic

11:45 PARALLEL SESSION 5 **ROOM 5**

Uncertainty Quantification for Computationally Expensive Models

Chairperson: Peter Harris

11:45 5.1 Recurrent networks for parameter estimation in MV and LV grids
ID1008 – Natallia Makarava

12:05 5.2 Bayesian inversion for CD determination with uncertainties
ID1040 – Nando Farchmin

12:25 5.3 Global sensitivity analysis using polynomial chaos
ID1041 – Nando Farchmin

11:45 PARALLEL SESSION 6 **ROOM 2**
Mathematics and Metrology in Medicine 1
Chairperson: Carola Schonlieb

11:45 6.1 Quantitative imaging biomarkers: the need for Metrology
ID1025 – Nadia Smith

12:00 6.2 Measurement of drug-induced changes in cardiac contractility using blood pressure data
ID1020 – Philip Aston

12:15 6.3 Managing uncertainties in calculations involving normal tissue complication probability
ID1037 – Maurice Cox

13:00 Lunch

14:15 INVITED SPEAKERS **AUDITORIUM**
Blaza Toman / Juris Meija
Chairperson: Olivier Pellegrino / Alen Bosnjakovich

15:45 Coffee break

16:15 PARALLEL SESSION 7 **ROOM 2**
Statistical Methods for Interlaboratory Comparisons and Conformity Assessment
Chairperson: Blaza Toman

16:15 7.1 Shades of dark uncertainty and consensus value for the Newtonian constant of gravitation
ID1049 – Antonio Possolo

16:35 7.2 Efficient sampling plans for the EU measuring instruments directive
ID1026 – Cord Müller

16:55 7.3 CASoft: practical implementation of risk calculations in conformity assessment
ID1001 – Alexandre Allard

17:15 7.4 CASoft: an approach for the verification of software for conformity assessment
ID1023 – Ian Smith

16: 15 PARALLEL SESSION 8 **ROOM 5**
Statistical Calibration and Regression Problems
Chairperson: Juris Meija

16:15 8.1 Bayesian uncertainty analysis versus application of the GUM and its supplements for error-in-variables straight-line regression
ID1012 – Steffen Martens

16:35 8.2 Effective number of degrees of freedom and prior information
ID1031 – Alistair Forbes, Andrew Thompson

16:55 8.3 On inverse and direct prediction in polynomial comparative calibration
ID1014 – Viktor Witkovsky

17:15 8.4 Activities of EURADOS Working Group 6 "Computational Dosimetry"
ID1017 – Hans Rabus

19:30 Conference Dinner

Friday // 22 November //

9:00 Registration

9:30 PARALLEL SESSION 9 **ROOM 5**
Mathematics and Metrology in Medicine 2
Chairperson: Nicolas Fischer

9:30 9.1 A model for complex shape and motion pattern analysis in medical images
ID1042 – Noemie Debroux

9:50 9.2 Proper modelling of errors in B1-mapping for uncertainty quantification in electric properties tomography
ID1018 – Alessandro Arduino

10:10 9.3 Risk based assessment of the degree of severity of myocardial perfusion and the determination of an optimal decision rule
ID1039 – Kavya Jagan

10:30 9.4 Obtaining high accuracy measurements of brain stents from 2D X-Ray images
ID1029 – Philip Aston

10:50 9.5 Uncertainty related to flow modelling errors in medical perfusion imaging
ID1045 – Gertjan Kok

9:30 PARALLEL SESSION 10 **ROOM 2**
Measurement Uncertainty 2
Chairperson: Antónia Turkman

9:30 10.1 Bayesian sample size determination for Type A uncertainty evaluation
ID1010 – Jörg Martin

9:50 10.2 Assessing the uncertainty contribution of detection thresholds to the uncertainty of frequencies of radiation-induced and background DNA damage foci obtained by automatic scoring
ID1016 – Hans Rabus

10:10 10.3 Distribution detection and information loss in a measurement uncertainty network
ID1028 – Paul Duncan

10:30 10.4 Role of measurement uncertainty in the comparison of average areal rainfall methods and its impact on conformity assessment
ID1060 – Álvaro Ribeiro

10:50 10.5 Non-informative Bayesian Inference for Heterogeneity in a Generalized Marginal Random Effects Meta-Analysis
ID1030 – Olha Bodnar

11:10 Coffee break

11:45 INVITED SPEAKERS **AUDITORIUM**
Antónia Turkman / Emma Woolliams
Chairperson: João A. Sousa

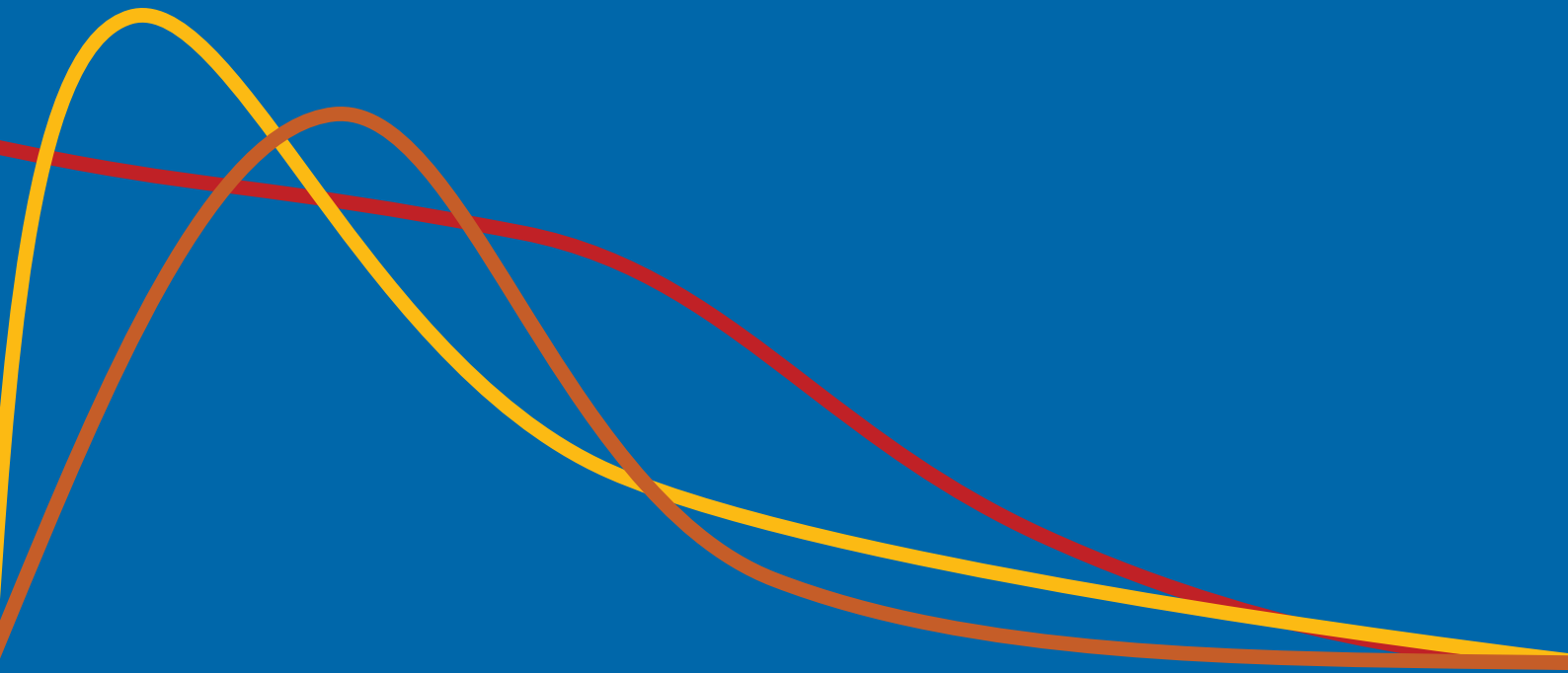
13:00 Closing

13:15 Lunch

14:15 **SPECIAL SESSION** on Introduction to Machine Learning for Metrology Apllications

14:15 MATHMET meeting

MathMet2019



Abstracts

MathMet2019

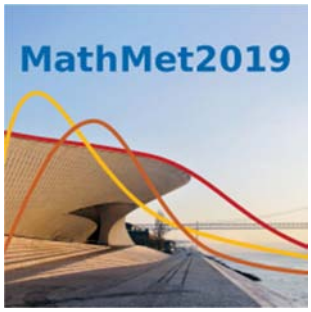
Abstracts

Measurement Uncertainty Training: A Survey and Developments

K. Klauenberg[1], M.-O. André[2], M. Désenfant[3], P. M. Harris[4], F. R. Pennecchi[5]

[1]Physikalisch-Technische Bundesanstalt, Abbestr. 2-12,10587 Berlin, Germany
[2] Federal Institute of Metrology METAS, Lindenweg 50, 3003 Berne-Wabern, Switzerland
[3]Laboratoire national de métrologie et d'essais, 1 rue Gaston Boissier 75724 Paris, France
[4]National Physical Laboratory, Hampton Rd, Teddington, TW11 OLW, UK
[5]Istituto Nazionale di Ricerca Metrologica – INRIM, Strada delle Cacce 91, 10135, Torino,Italy
E-Mail (corresponding author): Katy.Klauenberg@ptb.de

Evaluating the uncertainty of measurements is an inherent topic throughout metrology and wherever measurements are performed. Training on the evaluation of measurement uncertainty has a direct impact on how measurements and their uncertainty are understood at national metrology and designated institutes (NMIs / DIs), in the wider metrology community including calibration and testing laboratories, by legal, regulatory and standardization bodies, in industry as well as at universities and beyond. The quality of the training affects measurement capabilities in industry, metrological research and in society.
Therefore many, if not most, NMIs offer training courses on the evaluation of measurement uncertainty. These courses are usually based on the 'Guide to the expression of uncertainty in measurement' (*GUM*) and its supplements, but have developed independently and often differ with respect to the aim(s), the audience, content, duration, structure, the qualification of lecturers, integration into a training framework and possible constraints.
A dedicated session on measurement uncertainty training at MathMet2019 will offer a platform for overviewing different concepts at NMIs, for sharing and furthering good practice and for building a community. It will encourage and support NMIs and DIs without training courses to develop relevant expertise. During an open discussion common interests and their funding will be investigated.
In future, measurement uncertainty training could benefit from digital or E-learning solutions. An objective could be to develop harmonized, transferable and interconnected teaching or examination material. Video teasers could advertise measurement uncertainty courses, the *GUM* or adjacent topics. Frequently asked questions, problems or examples could be addressed and linked. Another development could aim at empowering the lecturers of courses, e.g. through exchange of expertise, didactical training, by establishing certification of courses or lecturers, or by reviewing measurement uncertainty software. More advanced courses e.g. on future *GUM* documents, on regression, Bayesian foundations or other active areas of research could be developed jointly.
An evolving community of lecturers creates a link between high-level research and every-day consumers of measurement uncertainty. The MATHMET centre and its research agenda could benefit from this link to potential stakeholders. Vice versa MATHMET's platform and network might provide an opportunity to support future activities of the community.
In a session on 'Measurement Uncertainty Training: A Survey and Developments' the authors of this proposal will briefly introduce their teaching concepts and/or ideas for developing measurement uncertainty training. The subsequent discussion will give room to debate concepts, propose ideas and shape objectives. All MathMet2019 attendees are encouraged to actively participate either in the discussion or by formally contributing to the overview part (through abstract submission to MathMet2019).



CASoft : Practical implementation of risk calculations in conformity assessment

A Allard[1], I Smith[2], P Harris[2], N Fischer[1], L Pendrill[3]

[1]Laboratoire National de métrologie et d'Essais, 1 rue Gaston Boissier, 75724 PARIS Cedex 15, Fance
[2]National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 OLW, United Kingdom
[3] RISE Research Institutes of Sweden AB, Eklandagatan 86, SE-412 61, Göteborg, Sweden
E-mail (corresponding author): alexandre.allard@lne.fr

The new version of standard ISO 17025:2017 has introduced a new requirement for calibration and testing laboratories, which consists of documenting the decision rule used for conformity assessment, taking into account the associated risk level. In particular, such a rule shall be communicated to, and agreed with, the customer. Indeed, due to measurement uncertainty, there is always a risk of taking a wrong decision: the risk of accepting a non-conforming item (called "consumer risk") and the risk of rejecting a conforming product (called "producer risk"). To this extent, JCGM106:2012 provides useful guidance on how to handle measurement uncertainty for evaluating risk in conformity assessment. The methodology includes the consideration of prior knowledge of the measurand. Such knowledge, combined with information provided from the measurement enables formulation of a posterior probability distribution that conveys all available information (pre and post measurement) about the measurand.

However, the practical application of the methodology is rather complex and requires high level probability calculations and the numerical evaluation of integrals. This complexity makes it difficult for practitioners to handle such calculations. EMPIR 17SIP05 "CASoft" aims at providing a free software tool for this purpose and at popularizing the implementation of the methodology for efficient decision-making in conformity assessment. The development of the software is based on the situations that are thought to arise most often for calibration and testing laboratories and industrialists. As an example, prior knowledge can be expected to be more accurate for the industrialist, whereas the calibration or testing laboratory may only have limited knowledge of the measurand before performing the measurement. Both situations will be handled with the developed software tool. Another key functionality is the determination of acceptance intervals, the interval of permissible measured quantity values, given a specified maximum level of risk. Indeed, the agreement between a calibration or testing laboratory and its customer (the industrialist for example) about the level of risk associated with the decision will result in the determination of a suitable acceptance interval.

The practical implementation of the calculations is illustrated through a case study that will involve both an industrialist and a calibration or testing laboratory. In particular, we propose to illustrate the calculation of specific or global risk on the one hand, and the determination of a suitable acceptance interval (agreed between both entities in the context of standard ISO 17025:2017) on the other hand.

Evaluating long-term trends in underwater noise in the Southern Ocean

P Harris, S Robinson, L Wang, V Livina and S-H Cheong

National Physical Laboratory, Teddington, United Kingdom
E-mail (corresponding author): peter.harris@npl.co.uk

Growth in international commerce and greater demands for energy have contributed to an increase in offshore human activity and the pollutants such activity produces. Underwater noise is classified as a form of pollution by international regulation, and there is increasing understanding of the effects such noise can have on the wellbeing of marine ecosystems. Consequently, the measurement of ambient sound levels in the deep ocean has been the subject of a number of recent studies, with particular interest in understanding long-term trends. In this work we describe a method for performing long-term trend analysis from measured deep-ocean noise data. Calculation of the uncertainty associated with the trend estimate is a fundamental requirement for statistical inference as it provides quantitative information about the quality of the estimate, and supports decisions made on the basis of the estimate.

Measurements of underwater ambient noise have been carried out since at least the 1960s. Most of the studies demonstrating an increase in the levels of low frequency sound in the deep-ocean have been undertaken in the Pacific Ocean. Such increasing trends have been partly attributed to increases in noise produced by shipping, but it is recognised that there is a variety of sound sources that contribute to the ambient sound field, both man-made and natural. The paucity of data over the last 50 years has meant that attempts to determine trends have often been based on very few data points and relied on simple statistical techniques such as least-squares straight-line fits. In more recent studies, covering the last 15 years, use has been made of much richer data sets where continuous monitoring has been undertaken. The measured data used here originate from the Southern Ocean spanning 15 years from 2003 to 2018. The data were obtained from the hydro-acoustic monitoring stations of the Preparatory Commission for the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO). The monitoring stations provide essentially continuous data at a sampling frequency of 250 Hz, leading to very large datasets, and coverage of acoustic frequencies up to 105 Hz.

The analysis method uses a flexible discrete model that incorporates terms that capture seasonal variations in the data together with a moving-average statistical model to describe the serial correlation of residual deviations. It is shown that using a (simple) straight-line model to describe the data leads to a model fit for which the residual deviations exhibit strong serial correlations, and not properly accounting for the serial correlation of the data-model differences can lead to uncertainties associated with trend estimates that are underestimated and unreliable. The trend analysis is applied to time series representing monthly and daily aggregated statistical levels for five frequency bands to obtain estimates for the change in sound pressure level (SPL) over the examined period with associated coverage intervals.

Evaluating long-term trends in underwater noise in the Southern Ocean

The analysis demonstrates that it is possible to determine statistically significant trends in deep-ocean noise data over periods exceeding a decade. The main features of the approach include (a) using a model that includes terms to represent explicitly seasonal behaviour, (b) applying a transformation to the data to improve the homogeneity of the distributions for the data-model differences, (c) using daily aggregation intervals derived from 1 minute SPL averages, and (d) applying a non-parametric approach to validate the uncertainties of the trend estimates that avoids the need to make an assumption about the distribution of those differences.

The results obtained (see Figure 1) show that statistically significant reductions in SPL are observed for all aggregated statistical levels for all the different frequency bands considered. Additionally, the relative differences between the various percentiles are found to be remarkably similar for all the frequency bands with higher percentiles following steeper trends than lower percentiles leading to a significant reduction of the dynamic range in the recorded noise. Also it is shown that trends in the data are dominated by the trends at low frequencies. Strong seasonal variation in the recorded data is also observed, with a high degree of correlation with climatic factors such as sea surface temperature and Antarctic ice coverage. Some possible explanations for the observed seasonal variations are presented.

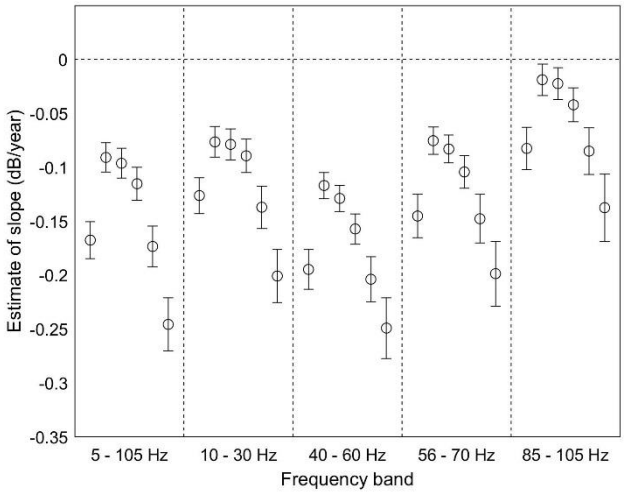


Figure 1: For daily aggregated data and each frequency band and each statistical level, estimates of the slope with associated 95% coverage intervals obtained by aggregating the results for the three hydrophones at the CTBTO station close to Cape Leeuwin, south west Australia. For each frequency band, results are presented for (left to right), average, P_1, P_{10}, P_{50}, P_{90}, and P_{99} statistical levels.

Influence of synchronization within a sensor system on machine learning results

T. Dorst[1], T. Schneider[2], S. Klein[3], S. Eichstädt[1], A. Schütze[2,3],

[1]Physikalisch-Technische Bundesanstalt, Abbestr. 2-12, 10587 Berlin, Germany
[2]Center for Mechatronics and Automation Technology (ZeMA), Eschberger Weg 46, 66121 Saarbruecken, Germany
[3]Lab for Measurement Technology, Department Systems Engineering, Saarland University, 66123 Saarbruecken, Germany
E-mail: tanja.dorst@ptb.de, t.schneider@zema.de, s.klein@zema.de, sascha.eichstaedt@ptb.de, schuetze@lmt.uni-saarland.de

Summary

Process sensor data allows not only the control of industrial processes, but also an assessment of plant condition to detect fault conditions and wear by means of machine learning [1]. To examine the influence of synchronization within a distributed sensor system, a testbed for condition monitoring, lifetime prediction and end-of-line tests of electro-mechanical cylinders is used. The sensors in the testbed are sampled between 10 kHz and 1 MHz. In this paper, time shifts between individual sensors of a maximum of 100 ms are considered and their influences on the performance of the remaining useful lifetime estimation for electromechanical cylinders is investigated. Therefore, a raw data set is manipulated by artificially simulated time shifts between the different sensors.

Motivation and results

Predictive maintenance, based on reliable condition monitoring, is a requirement for reducing repair costs and machine downtime and, therefore, increasing productivity. Therefore, an estimation of the remaining useful lifetime (RUL) of critical components is required, but this cannot be done directly. So, a testbed for electromechanical cylinders (EMCs) is used to perform RUL estimations. This testbed is equipped with 11 different sensors with different sampling rates between 10 kHz and 1 MHz, e.g. three acceleration sensors installed at the plain bearing, the ball bearing and the piston rod. In the testbed, an EMC operates cyclically against a pneumatic load of 7 kN until the cylinder fails due to increasing wear. Each working cycle, consisting of a forward and return stroke, lasts 2.8 s. The considered lifetime test consists of more than 600,000 working cycles. For the machine learning software toolbox (Fig. 1) [2], only the return stroke at constant speed with a duration of one second is used because of the constant pull against the pneumatic load in this period. This toolbox provides five different feature extraction and three different feature selection methods. To simulate synchronization problems, a random offset for the one second needed for the return stroke is set for each sensor. In this contribution the focus is on feature extraction methods, therefore Pearson correlation is always used for feature selection. The 10-fold cross-validation error increases for all feature extraction methods due to the simulated time shifts. The best RUL estimation is achieved with the raw data whereas the worst results from the maximum possible time shift of ± 100 ms. Both cases are shown in Fig. 2, where the cross validation error for the five feature extraction methods is plotted over the number of selected features. Other results indicate that only larger time shifts significantly influence the RUL estimation, resulting in constraints for the sensor synchronization.

Influence of synchronization within a sensor system on machine learning results

References

[1] A. Schütze, N. Helwig, and T. Schneider: Sensors 4.0 – smart sensors and measurement technology enable Industry 4.0, J. Sens. Sens. Syst., 7 (2018) 359–371

[2] T. Schneider, N. Helwig, and A. Schütze, “Industrial condition monitoring with smart sensors using automated feature extraction and selection,” Meas. Sci. Technol. 29 (2018) 094002 (15pp).

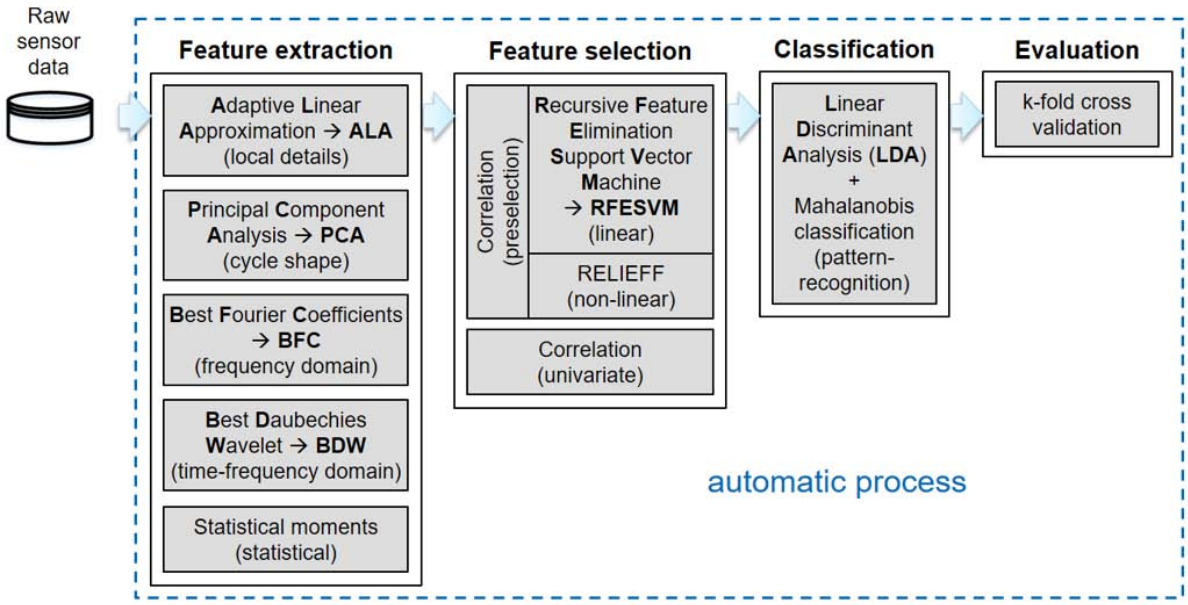


Figure 1: Scheme of the software toolbox

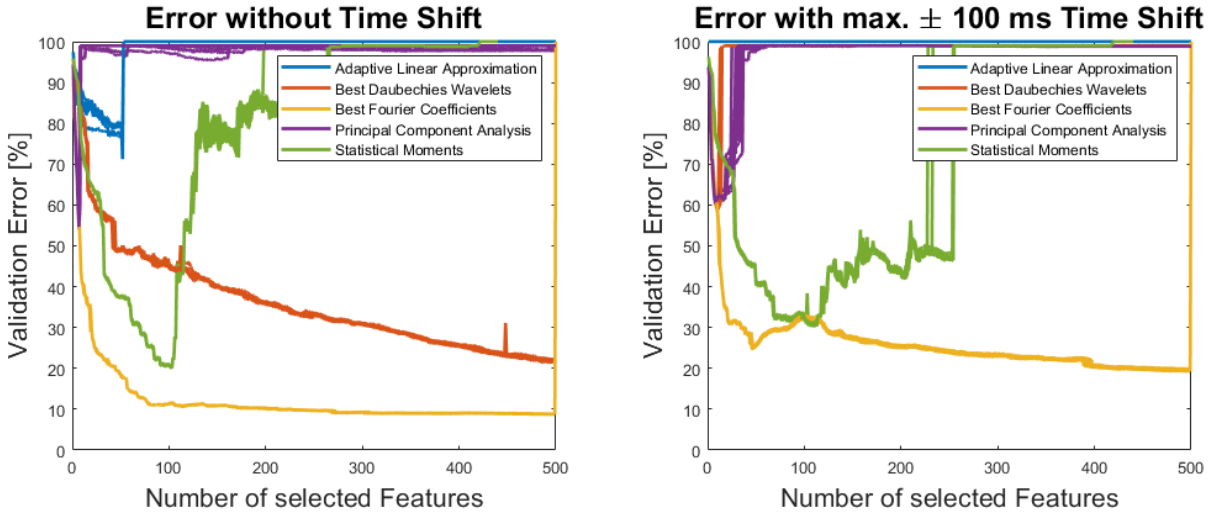


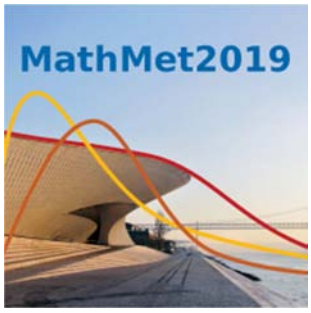
Figure 2: Comparison of the classification error without time shift (left) and with a time shift of maximum ± 100 ms per sensor (right) with a resolution of 1%. Pearson correlation with respect to used lifetime was used as a selector. In both cases, the smallest classification error is achieved with the Principal Component Analysis (8,72 % without time shift, 19,30 % with ± 100 ms time shift

Asymmetrical Uncertainties

A. Possolo[1], C. Merkatas[2], O. Bodnar[3]

[1]National Institute of Standards and Technology, Gaithersburg, MD, USA
[2] National Institute of Standards and Technology, Gaithersburg, MD, USA
[3] Mälardalen University, Västerås, Sweden
E-mail (corresponding author): antonio.possolo@nist.gov

In several disciplines, measurement results occasionally are expressed using coverage intervals that are asymmetric relative to the measured value. The conventional treatment of such results, when there is the need to propagate their uncertainties to derivative quantities, is to replace the asymmetric uncertainties by "symmetrized" versions thereof. We show that such simplification is unnecessary, illustrate how asymmetry may be modeled and recognized explicitly, and propagated using standard Monte Carlo methods. We present three distributions (Fechner, skew-normal, and generalized extreme value), among many available alternatives, that can be used as models for asymmetric uncertainties associated with scalar input quantities, in the context of the measurement model considered in the GUM. We provide an example where such uncertainties are propagated to the uncertainty of a ratio of mass fractions. We also show how a similar, model-based approach can be used in the context of data reductions from interlaboratory studies and other consensus building exercises where the reported uncertainties are expressed asymmetrically, illustrating the approach to obtain consensus estimates of the absorption cross-section of ozone, and of the distance to galaxy M83 in the Virgo cluster.



Tutorial for a Bayesian evaluation of measurement uncertainty and its implementation

S.Demeyer[1], C.Elster[2], N.Fischer[1]

[1]Laboratoire National de Métrologie et d'Essais, 29 avenue Roger Hennequin, 78197 Trappes, France
[2]Physikalisch-Technische Bundesanstalt, Abbestr.2-12, 10587 Berlin, Germany
contact author: severine.demeyer@lne.fr

The JRP 17NRM05 EMUE aims at providing comprehensive examples of measurement uncertainty (MU) evaluation as an extension of the Guide to the expression of uncertainty in measurement to cover problems arising in energy, environment and healthcare. This extension should provide methodologies to handle measurement problems where the GUM is not applicable (strong non linearities,…) i.e. when the law of propagation of uncertainty (LPU) is not valid. The GUM S1 still suffers from limitations for example when propagating distributions through complex systems.

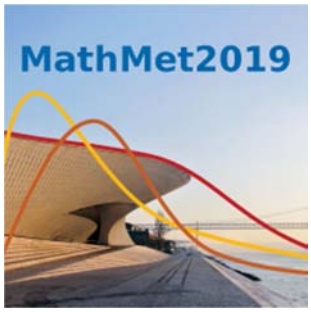
Complex situations require changing paradigm, moving from uncertainty/distribution propagation to statistical inference paradigm to allow better estimation of the quantity of interest. The latter consists in considering the measurand as a parameter controlling, among others, the distribution of observations/measurements and can be viewed as an alternative or a supplement to traditional GUM approaches.

The forthcoming new perspective of the GUM will highlight the use of Bayesian inference in measurement science following a large consensus in the metrology community and more generally in the statistical community to prefer Bayesian methods over classical least squares or maximum likelihood based methods for complex problems. Bayesian methods are famous for incorporating all kind of prior knowledge, as well as offering a flexible framework for modelling complex relationships such as inverse modelling and produce the joint posterior distributions of the parameters over the parameter space.

This tutorial addresses main operational aspects of Bayesian statistics in metrology such as building the likelihood of data from a measurement model, eliciting prior distributions from expert knowledge, deriving posterior distributions either analytically or based on MCMC (Markov Chain Monte Carlo) simulations. The comprehensive methodology is applied (with discussion) on the mass calibration example of GUM S1 and implemented in an R Shiny interface. Operational Bayesian MU evaluation procedure will be provided with the corresponding R code to implement it.

References :

BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement Joint Committee for Guides in Metrology, JCGM 100:2008



Tutorial for a Bayesian evaluation of measurement uncertainty and its implementation

BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 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, JCGM 101:2008

Elster C, *Bayesian uncertainty analysis compared with the application of the GUM and itssupplements*, Metrologia 51(2014), 159—166

Possolo A and Toman B, Assessment of measurement uncertainty via observation equations, Metrologia 44(2007) 464–75

Evaluating the uncertainty in the measurement of nanoparticle size by means of SEM and DLS

I. Lira[1], F. Plentz[2], D. Camarano[3], M. de Farias[4], M. Orlandi[5], G. Goldschmidt[6], D. Hotza[6]

[1]Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile
[2]Universidade Federal de Minas Gerais, 31270-901, Belo Horizonte, MG, Brazil
[3]Nuclear Technology Development Center, 30161-970, Belo Horizonte, MG, Brazil
[4]Nanotechnology National Laboratory, 13083-970, Campinas, SP, Brazil
[5]Universidade Estadual Paulista, 14800-900, Araraquara, SP, Brazil
[6]Universidade Federal de Santa Catarina, 88040-900, Florianópolis, SC, Brazil
E-mail (corresponding author): ilira@ing.puc.cl

Because of the wide variety of potential applications of nanomaterials in various fields (e.g. optics, catalysis, pharmaceutics, medicine or electronics), nanoparticles constitute currently an area of intense scientific research. For example, due to their extremely small size (less than 100 nm), these particles can diffuse in the body and accumulate in various organs and tissues, allowing for an improvement of drug delivery to specific targets, such as cancer cells. In this way, damage to healthy cells in the body is reduced and earlier detection of disease is possible. However, success in this and in many other applications depends largely on the ability to characterize nanomaterials, especially to establish their dimensions.

Scientists and technicians commonly use different procedures to measure nanoparticle size and size distribution reliably. A non-exhaustive list includes electron microscopy (EM), scanning probe microscopy (SPM), dynamic light scattering (DLS), small-angle X-ray scattering (SAXS), atomic force microscopy (AFM), nanoparticle tracking analysis (NTA), centrifugal liquid sedimentation (CLS), scanning mobility particle sizer (SMPS), light scattering particle counter (LPC), and others.

The majority of these techniques are either of the single particle type or of the ensemble type. In the former, images of individual particles are analysed; typical examples are EM and SPM. Such techniques are now well established and accurate calibration standards for them are available. However, they suffer from a relatively high cost, long analysis time and the requirement for highly trained personnel. In the ensemble type, many particles are analysed simultaneously; common examples are DLS and SAXS. These techniques do not measure the size of the particles directly, but rather the modelled effect of size on another property. They are generally cheaper and easier to use compared to single particle techniques and better suited for routine quality control applications [1].

In 2015, a project for the qualification and modernization of Brazilian technological research institutes was started. Signatories of this so-called Modernit Project were FINEP (funder of studies and projects) and FUNDEP (research development foundation); its focus was the construction of a research network in the area of nanotechnology. Participants were nine laboratories located throughout Brazil that are involved in the provision of nanotechnology services. The aim was to implement quality management systems in accordance with the

Evaluating the uncertainty in the measurement of nanoparticle size by means of SEM and DLS

ISO/IEC 17025 standard in order to attest the technical competence in performing the tests and measurements these laboratories perform.

Among other stipulations, this standard requires that:

- measuring equipment shall be calibrated and be capable of achieving the measurement uncertainty needed to provide a valid result;
- metrological traceability of measurement results shall be maintained by a documented unbroken chain of calibrations, each contributing to the measurement uncertainty, linking them to an appropriate reference;
- appropriate methods and procedures shall be used for laboratory activities and for the analysis of data, including the identification and quantification of all contributions to the measurement uncertainty.

Other documents (e.g. [2]) stress also the importance of evaluating measurement uncertainty, for it affects quality, costs, decisions, and risks. It was soon discovered, however, that the literature regarding uncertainties in nanometrology is rather scant. Scanning and transmission EM (SEM and TEM) and DLS are the two main techniques used by the laboratories involved in the Modernit Project to measure nanoparticle size. Yet, we found no guidance regarding EM at all, and only [1] provides some guidance regarding DLS.

In this talk, we shall discuss methods for evaluating the uncertainty in the measurement of nanoparticle size by means of SEM and DLS in accordance with the GUM uncertainty framework [3]. As mentioned above, these two techniques are very different, so the evaluation methods are different too. Thus, in SEM traceability is obtained through calibration with a certified reference standard such as Geller MicroAnalytical Laboratory's MRS-6. Instead, DLS does not require calibration in the common sense of the term, but the trueness of the instrument used should be validated through periodic measurement of certified reference materials such as gold or polystyrene nanoparticles [4]. A summary of our uncertainty evaluation proposals for these two techniques will be presented.

[1] Boyd, R. D. et al. (2015) "Measurement Good Practice Guide No.119; Good Practice Guide for the Determination of the Size and Size Distribution of Spherical Nanoparticle samples." Available in https://www.npl.co.uk/resources/gpgs.

[2] S. L. R. Ellison and A. Williams (Eds). Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement, Third edition. Available from www.eurachem.org, 2012.

[3] JCGM 100:2008 Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement.

[4] Varenne, F. et al. (2015) Colloids and Surfaces A: Physicochem. Eng. Aspects 486:124

Recurrent Neural Networks for parameter estimation in MV and LV grids

N. Makarava, S. Eichstädt
Physikalisch-Technische Bundesanstalt Institut Berlin, Abbestr. 2-12, 10587 Berlin, Germany
E-mail (corresponding author): natallia.makarava@ptb.de

State estimation in middle- (MV) and low-voltage (LV) electrical grids is a topic of active research since several decades. The basic challenge for this task is the lack of measurement data, because instrumentation of MV and LV grids is very sparse due to economic reasons. Many classical state estimation methods have been studied in the literature with varying efficiency and quality. In this contribution, we consider the application of artificial neural networks for state estimation, which have shown very good performance in other signal processing tasks. For instance, it was shown by many researches that artificial neural networks that are designed to simulate the biological neural systems, outperform many standard estimation methods. For the training of neural networks many algorithms are proposed in the literature. While some of them are based on the idea of optimizing the weights of the neural network to minimize a loss function (gradient descent), others are trying to optimize the network structure and hyperparameters (Bayesian approach) [1].

Classic gradient descent techniques can result in slow convergence and a nonlinear problem. Thus, the calculation of second-order derivatives can cause significant issues. Therefore, we approach the optimization of the weights through the Extended Kalman Filter method, where we interpret the weights as the states of a particular dynamic system [2]. Therefore, we consider the following system model for state estimation [3]:

$$\begin{cases} x(k+1) = \omega(k)\sigma\big(x(k)\big) + g(k) \\ \omega(k+1) = \omega(k) + e(k) \\ z(k) = C(k)x(k) + D(k)\,u(k), \end{cases}$$

where $x(k)$ is the state vector of nodal active and reactive powers, $\omega(k)$ are the weights, $z(k)$ is the vector of measured voltages, $u(k)$ is the input vector of measured nodal power values, $C(k)$ and $D(k)$ are time-dependent matrices of appropriate dimension, $e(k)$ and $g(k)$ are errors of known variance, and $\sigma(x) = 1/(1+\exp(-\beta x))$ is the sigmoid function with $\beta > 0$.

A previously proposed approach to the state estimation problem is the nodal load observer method [4], which aims to correct possibly incorrect pseudo-measurements of bus power and then determines the grid state based on reconstructed and corrected values of nodal power and voltage:

$$\begin{cases} x(k+1) = \alpha x(k) \\ z(k) = h(x(k), u(k)k). \end{cases}$$

Here, $h(.)$ is power-flow function and α is a number close to one for generating a simple dynamic behaviour of the states.

Recurrent Neural Networks for parameter estimation in MV and LV grids

The joint model of both approaches can also be considered:

$$\begin{cases} x(k+1) = \omega(k)\sigma\big(x(k)\big) + g(k) \\ \omega(k+1) = \omega(k) + e(k) \\ z(k) = h(x(k), u(k)k). \end{cases}$$

To illustrate the neural network approach for state estimation, we consider a grid model developed by a consortium of UK universities specifically for testing purposes. The grid presented in Figure 1 is a 11kV urban network fed from a 33kV supply point and has 12 buses, 11 branches with a generator at bus number 5 and the slack bus has number 0. This grid is a part of a 77 buses grid, which due to the radial topology can be examined independently. Data sets for this network typically represent the course of one day with measurements taken every 15 minutes, resulting in 96 data points. This small number of consecutive measurements is typical for such networks, but problematic for many online learning techniques due to the required speed of convergence. Measurements with this network have been simulated by inserting a time-dependent generator output and application of the method of optimal power flow to obtain matching measured values at all network buses.

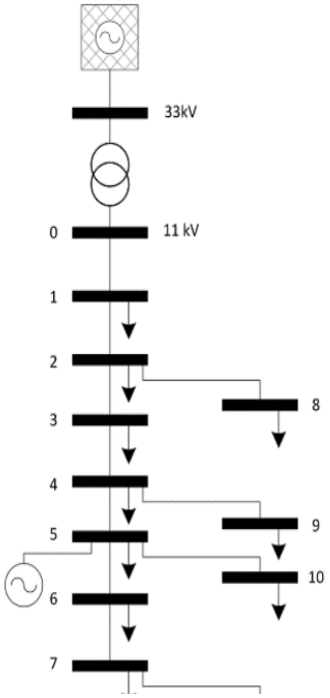


Figure 1. United Kingdom generic distribution system (UKGDS) grid.

We will compare the results with an extension of the nodal load observer method for considering the dynamical model based on autoregressive processes with online estimation learning technique [5].

References

1. Nabney, I. T. NETLAB. Algorithms for Pattern Recognition. 2002. London: Springer.
2. Haykin, S. Kalman Filtering and Neural Networks, Wiley–Interscience, 2001.
3. Alanis, Y., Sanchez, E. N., Loukianov, A. G., Perez, M. A. Real-Time Recurrent Neural State Estimation, IEEE Transactions on Neural Networks, vol. 22, no. 3, 2011.
4. Heins, W., Gewiss, G., Bohn, C., Beck, H.-P. State-space formulation for the nodal load observer for smart electrical grids with imperfect measurement infrastructure, XX IMEKO World Congress, 1-6, 2012.
5. Makarava, N., Lin, G., Eichstädt, S. Adaptive quasi-dynamic state estimation for MV and LV grids, submitted to EURASIP Journal on Advances in Signal Processing.

Recovery of smooth low-rank matrices using Bayesian inference

G. Wübbeler and C. Elster

Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany
E-mail (corresponding author): gerd.wuebbeler@ptb.de

The recovery of a matrix from a limited, and often small, number of entries is a common task arising in different fields such as recommender systems, gene-disease association, image inpainting, hyperspectral imaging, or magnetic resonance imaging. Corresponding matrix recovery schemes can be applied, e.g., when it is not possible to observe all matrix elements or when trying to accelerate a time-consuming measurement of high dimensional data.

Successful matrix recovery can often be achieved by a low-rank approximation of the sought matrix [1]. Many low-rank approaches utilize a convex optimization in terms of the nuclear norm given by the sum of the singular values. Alternatively, the low-rank matrix can be factorized by two smaller matrices. Low-rank matrix completion has also been suggested utilizing Bayesian inference by specifying priors enforcing a low-rank solution. The development of methods for low-rank matrix completion is a current topic of research [2].

We propose a hierarchical Bayesian inference for the recovery of a low-rank matrix from incomplete noisy observations. In addition to the low-rank, the sought matrix is assumed to exhibit smoothness along its rows and/or columns which can be expected in applications such as FTIR nano-spectroscopy [3] or magnetic resonance imaging. The proposed inference combines a recently suggested low-rank prior [4] with a Gaussian Markov random field (GMRF) prior accounting for smoothness [5]. The combination of these two priors again represents a GMRF prior. The parameters specifying the level of smoothness are modelled by hyperprior distributions which enables their determination from the observed data. An approximate Bayesian inference is achieved in terms of maximum *a-posteriori* estimates for which a sequential optimization scheme has been developed.

The performance of the proposed low-rank and smooth matrix recovery scheme is explored using synthetic data and real images. It is demonstrated that the additional smoothness assumption can lead to more stable solutions, enabling a successful matrix recovery even when only a small fraction of matrix elements is available.

Recovery of smooth low-rank matrices using Bayesian inference

[1] Candès, E. J., & Recht, B. (2009). Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6), 717.

[2] Chen, Y. and Chi, Y. (2018), Harnessing Structures in Big Data via Guaranteed Low-Rank Matrix Estimation: Recent Theory and Fast Algorithms via Convex and Nonconvex Optimization. *IEEE Signal Processing Magazine*, vol. 35, issue 4, pp. 14-31.

[3] Kästner, B., Schmähling, F., Hornemann, A., Ulrich, G., Hoehl, A., Kruskopf, M., Pierz, K., Raschke, M. B., Wübbeler, G. and Elster, C. (2018). Compressed sensing FTIR nano-spectroscopy and nanoimaging. *Optics express*, vol. 26, no. 14, pp. 18115-18124.

[4] Babacan, S. D., Luessi, M., Molina, R., & Katsaggelos, A. K. (2012). Sparse Bayesian methods for low-rank matrix estimation. *IEEE Transactions on Signal Processing*, 60(8), 3964-3977.

[5] Rue, H., & Held, L. (2005). Gaussian Markov random fields: theory and applications. Chapman and Hall/CRC.

Bayesian sample size determination for Type A uncertainty evaluation

Jörg Martin[1] and Clemens Elster[1]

[1]Physikalisch-Technische Bundesanstalt (PTB), Braunschweig and Berlin

E-mail (corresponding author): joerg.martin@ptb.de

The evaluation of measurement uncertainty for a small number of repeated measurements is a common task in metrology. The type A uncertainty evaluation of the GUM [1] attempts to assign an estimate of the scaled standard deviation of the underlying sampling distribution as the uncertainty. When the number of observations is small, this estimate is highly uncertain and does not provide a useful result from the perspective of classical statistics.

We consider instead a Bayesian Type A uncertainty analysis [2] for which the uncertainty is calculated as the standard deviation of a posterior distribution for the measurand. Bayesian inference [3] incorporates prior knowledge which can yield meaningful results also in the small sample case.

A Bayesian uncertainty analysis ought to be accompanied with a sensitivity analysis that reflects reasonable variations of the employed prior distributions. The number of required observations ought to be high enough such that the sensitivity analysis yields sufficiently stable results for the considered purpose, otherwise further measurements are needed.

The goal of this contribution is to illustrate such sensitivity analyses for a Bayesian type A uncertainty evaluation and to provide some means for assessing the adequacy of a chosen sample size. In addition, tools from Bayesian experimental design [4] are employed for the goal of sample size planning.

A Gaussian sampling distribution is considered. It is assumed that prior knowledge about the variance of the sampling distribution and vague prior knowledge about the measurand are available. This situation addresses the relevant case of a reliable, well-characterized measurement device that is applied in daily routine.

We start by illustrating the sensitivity of results for specific situations and different cases of small samples including $n = 1$ and $n = 2$. In a second step we apply Bayesian experimental design to derive means for determining a minimum sample size prior to sampling. Specifically, we determine the prior predictive distribution which is used to calculate the expected variation of the uncertainty reached. We then constrain the deviation to lie below a subjectively chosen threshold. This requirement is forced into a loss function and evaluated in a general framework. The derived expressions illustrate the tradeoff between the accuracy in the prior knowledge and the number of required observations needed to expect a stable uncertainty. Finally, we discuss the application of these expressions for sample size planning.

Bayesian sample size determination for Type A uncertainty evaluation

References

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, "Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement".*Joint Committeefor Guides in Metrology, JCGM 100:2008.*

[2] C. Elster, „Bayesian uncertainty analysis compared with the application of the GUM and its supplements,“ *Metrologia 51(4),* p. 159, 2014.

[3] C. Robert, The Bayesian choice: from decision-theoretic foundations to computational implementation., Springer Science & Business Media, 2007.

[4] K. Chaloner und I. Verdinelli, „Bayesian experimental design: A review,“ *Statistical Science,* pp. 273-304, 1995.

Redundant information in sensor networks and uncertainty quantification

G.J.P. Kok[1], P.M. Harris[2], Y. Luo[2],O. Panni[2] and T. Dorst[3]

[1]VSL B.V., Thijsseweg 11, 2629 JA Delft, the Netherlands
[2]National Physical Laboratory, Teddington, United Kingdom
[3]Physikalisch-Technische Bundesanstalt, Abbestr. 2-12, 10587 Berlin, Germany
E-mail (corresponding author): gkok@vsl.nl

The use of low cost sensor technology has been steadily increasing for quite some years now. One application of such technology is the “Factory of the Future”: production environments where a large network of sensors is used to monitor, control and optimise production. Very often there is some type redundant information present in the sensor network. Furthermore, whereas many applications exist where sensor data without uncertainty considerations are used to arrive at process information and control decisions, including uncertainty can improve the quality of the generated information and decisions, and give the user an understanding of the associated uncertainty. However, it is generally not clear how to incorporate uncertainty information into such applications, and research being performed in the EMPIR project “Metrology for the Factory of the Future” (Met4FoF) [1] is addressing this need. In this presentation the work performed for a few specific cases will be described, as well as how a more generally applicable framework is being developed.

The concept of ‘redundant information’ can be interpreted in different ways, as well as the concepts of ‘aggregated’ and ‘distributed’ measurement problems. A first step in this research was to define these concepts more precisely, allowing for various subcategories. These definitions can then be applied to various test environments or ‘testbeds’ that are part of the Met4FoF project.

In the STRATH radial forge testbed [2], metal parts are heated and formed by vibrating hammers, with approximately 60 sensor readings being recorded during the heating/forming process. Data has been collected for 81 forged parts, with final verification of the geometry undertaken via measurements using a Coordinate Measuring Machine. A physical understanding of the system provides some information about the usefulness of each of the sensors for a specific research question, whereas a statistical analysis (e.g. using Principal Component Analysis or cluster algorithms) can provide some more quantitative results. Sensitivity analysis by slightly mutating the data gives insight in the robustness and uncertainty of the methods. The main results will be presented. Special attention will be paid to the existence of redundant information.

In the ZEMA testbed [3,4] the goal is to predict the residual life time of electromechanical cylinders based on sensor data of various origin. In recent years ZEMA has developed an impressive toolset for analysing this data using machine learning techniques and training data [5]. By adding measurement noise, timing offsets and jitter, one can get an idea of the robustness of these, and of other methods and statistics, as well as an idea about the

uncertainty of the end result. Again, the most interesting results will be presented, highlighting (partially) redundant information.

The example results for the STRATH and ZEMA testbeds (complemented by altered and/ or simulated data) show a way of assessing the uncertainty of final results ('target information'), and how redundancy of information affects this. Future work will also address the SPEA testbed [3], where a machine is being developed for calibrating MEMS temperature sensors in large batches against a few reference sensors.

An outlook will be given on how the different methods and tools developed for different applications can be generalized so that they can be part of a mathematical framework that can be used for analysing industrial (and other) sensor networks including uncertainty quantification.

We gratefully acknowledge funding of this research by the EMPIR 17IND12 Met4FoF project. The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States.

[1] EMPIR project 17IND12 "Metrology for the Factory of the Future" (Met4FoF), www.met4fof.eu
[2] STRATH testbed description and dataset, https://zenodo.org/record/2573861#.XL_7BegzaUl
[3] ZEMA and SPEA testbed descriptions, https://www.ptb.de/empir2018/met4fof/information-communication/blog/
[4] ZEMA testbed dataset, https://zenodo.org/record/1326278#.XL_7ZegzaUk
[5] "Automatic Feature Extraction and Selection for Condition Monitoring and related Datasets", N. Helwig, T. Schneider, A. Schütze, IEEE I2MTC 2018

Bayesian uncertainty analysis versus application of the GUM and its supplements for errors-in-variables straight-line regression

S. Martens[1], K. Klauenberg[1], C. Elster[1]

[1]Physikalisch-Technische Bundesanstalt, Abbestraße 2–12, 10587 Berlin, Germany
E-mail: steffen.martens@ptb.de

Straight-line regression of experimental data is one of the most common problems in metrology and other fields of quantitative scientific work. Often, the independent variable, for instance, the quantity value provided by measurement standards in calibration processes or a reference procedure in method comparison studies, is assumed to be measured exactly, while the uncertainties of the dependent variable are used as weights for fitting the calibration curve. These presumptions are seldom rigorously true and errors-in-variables (EiV) regression models should be employed instead.

In this talk, we present a generic treatment of EiV straight-line regression. The focus is on the case of non-vanishing variance and covariance of each data point (x_i, y_i). The corresponding statistical model reads

$$x_i = X_i + d_i,\ y_i = Y_i + e_i,\ Y_i = \alpha + \beta X_i,\ i = 1, \ldots, n, \qquad (1)$$

where (X_i, Y_i) denote the latent values, each pair (d_i, e_i) is a realization of a bivariate normal random variable with expectation $(0,0)$ and known covariance matrix

$$\boldsymbol{U}_i = \begin{pmatrix} u^2(x_i) & \mathrm{Cov}(x_i, y_i) \\ \mathrm{Cov}(y_i, x_i) & u^2(y_i) \end{pmatrix}, \qquad (2)$$

and (α, β) are the regression parameters to be estimated. For this purpose, multiple standards[1,2] suggest the use of weighted total least-squares[3] (WTLS). Mostly, the underlying nonlinear optimization problem in WTLS can solely be solved numerically. To avoid the latter, simplifications like ordinary least-squares (OLS) regression, “effective-variance” OLS[4], and reduced major axis or limiting cases like orthogonal regression and Deming regression have been discussed[5] and applied[6].

EiV regression has been approached mostly from the frequentist perspective. In metrology, the associated uncertainties of the regression parameters are often determined by the law of

[1] ISO/TS 28037:2010. *Determination and use of straight-line calibration functions*, 2010. International Standards Organization, Geneva.
[2] ISO 6143:2006. *Gas analysis — Comparison methods for determining and checking the composition of calibration gas mixtures*, 2006. International Standards Organization, Geneva.
[3] D. York, *Can. J. Phys.* **44**, 1079 (1966); C. A. Cantrell, *Atmos. Chem. Phys.* **8**, 5477 (2008).
[4] M. Lybanon, *Am. J. Phys.* **52**, 22 (1984).
[5] D. York, *Can. J. Phys.* **44**, 1079 (1966); W. A. Fuller, *Measurement error models* (John Wiley & Sons, 2009).
[6] T. Isobe et al., *Astrophys. Journal* **364**, 104 (1990); D. I. Warton et al., *Biol. Rev.* **81**, 259 (2006).

Bayesian uncertainty analysis versus application of the GUM and its supplements for errors-in-variables straight-line regression

propagation of uncertainty (LPU)[1,7]. While the assessment of measurement uncertainties by LPU is consistent with the GUM uncertainty framework[8], the supplement 1 to the GUM[9] (GUM-S1) proposes the evaluation of uncertainty in terms of a probability density function. The latter expresses the knowledge about the value of the measurand, like the posterior distribution in Bayesian uncertainty analysis (BUA), without applying Bayes' theorem explicitly. Therefore, the resulting distribution for the measurand applying the GUM-S1 procedure may coincide with a posterior distribution resulting from BUA, but this is not true in general. For OLS with known variance, it has been shown that the GUM-S1 procedure yields a density identical to the Bayesian posterior for the usual non-informative prior, otherwise not[10]. Consequently, it is not clear when and if at all, the application of GUM-S1 results in a distribution for the regression parameters (α, β) in model (1) that could be reached by a BUA. Further, the estimates and their variances obtained according to the GUM may differ from the values derived by Bayesian uncertainty analysis. Despite the importance of EiV regression for metrology, neither the GUM nor its current supplements provide guidance and only few results are available on a Bayesian uncertainty analysis[11], in contrast to research areas like economics[12] and statistics[13].

Our aim is to investigate the relation between a Bayesian uncertainty analysis and the application of the GUM or GUM-S1 to WTLS and its simplifications. Similarities and differences in the approaches are illustrated. Because BUA requires a prior distribution to be assigned that expresses one's state of knowledge about all unknowns before the data are considered, prior elicitation is essential. We start by investigating suitable non-informative prior distributions and explore the relationship between the posterior distribution resulting from BUA and the probability density function obtained by GUM-S1. More importantly, we indicate in which circumstances the use of BUA with non-informative and informative priors has advantages; being quantified e.g. by the properties of the credible intervals or statistical measures for the regression parameters' estimate. Especially, we check the validity of the often blindly used OLS and provide guidance when OLS is expected to be insufficient.

---

[7] M. Krystek and M. Anton, *Meas. Sci. Technol.* **22**, 035101 (2011).
[8] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML. Evaluation of measurement data — Guide to the expression of uncertainty in measurement. Joint Committee for Guides in Metrology, JCGM 100:2008
[9] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. 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, JCGM 101:2008.
[10] C. Elster and B. Toman, *Metrologia* **48**, 233 (2011).
[11] I. Lira and D. Grientschnig, *Metrologia*, **54**, 133 (2017); A. M. H. van der Veen, *Metrologia*, 55, 670 (2018).
[12] A. Zellner, *An Introduction to Bayesian Inference Econometrics: Introduction to Bayesian Inference in Econometrics* (New York: Wiley-Interscience, 2010).
[13] D. Leonard, *Bayesian Anal.*, **6** 727 (2011); S. Muff et al., *J. Royal Stat. Soc. C* **64**, 231 (2015).

Confidence Regions for Parameters in Two-Dimensional Linear Comparative Calibration Model

Gejza Wimmer[1,2] and Viktor Witkovský[3]

[1] Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic
[2] Matej Bel University, Banská Bystrica, Slovak Republic
[3] Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Sovak Republic
E-mail: wimmer@mat.savba.sk

The contribution is concerned with two-dimensional comparative calibration, i.e. with situation when two-dimensional measurements are indicated by two measuring devices which are subject to normally distributed errors. The two dimensional comparative calibration model is introduced together with its replicated form. From statistical point of view the model is a linear regression model with nonlinear constraints on the parameters. After linearization it could be represented as a linear model of measurements with type II constraints [1]. An iterative algorithm for estimating the parameters of the linear calibration function is suggested. Also is solved the problem of deriving the approximate confidence region for the vector of calibration function parameters and also for linear functions of these parameters, based on using the method suggested by Kenward and Roger [2].

References

[1] Fišerová, E., Kubáček, L., Kunderová, P. (2007). *Linear Statistical Models: Regularity and Singularities*. Academia, Praha.
[2] Kenward,M.G., Roger, J.H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53, 983-997.

Supported by the Slovak Research and Development Agency, project APVV-15-0295, and by the Scientific Grant Agency VEGA of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences, project VEGA 2/0054/18.

On Inverse and Direct Prediction in Polynomial Comparative Calibration

Viktor Witkovský[1] and Gejza Wimmer[2,3]

[1] *Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Sovak Republic*
[2] *Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic*
[3] *Matej Bel University, Banská Bystrica, Slovak Republic*

E-mail: witkovsky@savba.sk

Calibration is an essential part of many measurement procedures. Linear and polynomial functions are the most frequently used calibration functions in comparative calibration, i.e. situation when measurements are indicated and compared by two measuring devices, both of which are subject to measurement errors and other uncertainties related to the measurement results. Parameters of the calibration function estimated from the calibration experiment are further used for predicting the value of the unobservable stimulus from future indication (measured response) received by using the calibrated measurement device, together with its associated measurement uncertainty.

The methods for proper determination and use of the calibration functions are broadly studied from the statistical and also from the metrological point of view. In metrology, the straight-line calibration and the polynomial calibration is well established in the ISO Technical Specifications [1-2], which are based on the GUM uncertainty framework [3], using the law of propagation of uncertainty. Here we shall assume that the calibration experiment provides direct measurements specified by state-of-knowledge distributions derived by using Type A and Type B methods of evaluation.

By using the Monte Carlo Method (MCM) for propagation of the distributions [4-5] it is possible to derive the state-of-knowledge distribution about the values of the calibration function parameters, and moreover, under specific conditions (as, e.g. monotonicity of the calibration function) also the associated state-of-knowledge distribution of the values attributed to the stimulus, given the state-of-knowledge distribution about the new indication received by the calibrated measurement device. However, in the classical setup based on using the inverse prediction, the MCM approach requires a root-finding solution in each simulation step, which can be computationally very demanding task. This raises a natural question about the possibility of using inverse relationship in modelling the calibration function, i.e., by changing the standard roles of the stimulus and the response in the considered calibration function. This leads to computationally simpler direct prediction.

In this contribution, we present a brief overview of results known for linear and polynomial comparative calibration, see [1-2] and [6], and investigate the effect of using inverse relationship in modelling the calibration function, i.e., by changing the standard roles of the stimulus and the response in the considered calibration function.

On Inverse and Direct Prediction in Polynomial Comparative Calibration

References

[1] ISO/TS 28037:2010. Determination and Use of Straight-Line Calibration Functions. International Standards Organization, Geneva, September 2010, 63 pp.
[2] ISO/TS 28038:2018. Determination and Use of Polynomial Calibration Functions. International Standards Organization, Geneva, December 2018, 52 pp.
[3] JCGM100:2008. Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement. JCGM – Joint Committee for Guides in Metrology.
[4] JCGM101:2008. Evaluation of Measurement Data – Supplement 1 to the Guide to the Expression of Uncertainty in Measurement – Propagation of Distributions Using a Monte Carlo Method. JCGM – Joint Committee for Guides in Metrology
[5] JCGM102:2011. Evaluation of Measurement Data – Supplement 2 to the Guide to the Expression of Uncertainty in Measurement – Extension To Any Number of Output Quantities. JCGM – Joint Committee for Guides in Metrology.
[6] Witkovský, V., Wimmer, G. (2018). Generalized polynomial comparative calibration: Parameter estimation and applications. In: S.Y. Yurish, (Ed.), Advances in Measurements and Instrumentation: Reviews, IFSA (International Frequency Sensor Association Publishing) Publishing, Barcelona, Spain, 2018, 15-52.

Supported by the Slovak Research and Development Agency, project APVV-15-0295, and by the Scientific Grant Agency VEGA of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences, projects VEGA 2/0054/18 and VEGA 2/0081/19.

Uncertainty budget for gas mixtures preparation by dynamic dilution and subsequent use in the calibration of analytical instrumentation

F. Pennecchi[1], F. Rolle[1], M. Sega[1], P. Spazzini[1], I. de Krom[2], A.M.H. van der Veen[2]
[1]Istituto Nazionale di Ricerca Metrologica - INRIM, Strada delle Cacce 91, Torino, Italy
[2]VSL, Thijsseweg 11, Delft, The Netherlands
E-mail (corresponding author): f.pennecchi@inrim.it

In the framework of the EMPIR 17NRM05 EMUE Project on "Advancing measurement uncertainty – comprehensive examples for key international standards" [1], INRIM and VSL are collaborating on the development of an example on the uncertainty evaluation concerning preparation of calibration gas mixtures of nitrogen oxides (NO_X) using dynamic dilution with mass flow controllers (MFCs) in accordance with ISO 6145-7:2018 [2].

Such a case study is meant to provide a response to the JCGM survey on the GUM [3] that highlighted the need for examples on correlation, arising in many applications across metrology and creating difficulties for many end-users. Indeed, correlations between the amount fractions of the analyte in the mixtures will be evaluated taking into account the effect introduced by the same parent mixture and the same MFCs used in the dilution.

The European Directive [4] prescribes the monitoring of NO_X by means of chemiluminescence as the reference method [5], which requires the use of proper reference gas mixtures for instrument calibration. To prepare such mixtures, dynamic dilution is a primary method considered as a valid alternative to the static gravimetric method: it allows preparing ready-to-use gas mixtures at low mass or amount fractions by diluting a standard mixture (parent mixture) with a proper diluent gas, thus avoiding stability problems related to diluted mixtures of reactive gases in high-pressure cylinders.

The foreseen operational flow is organized according to the following steps (for each of which a full uncertainty budget will be developed and provided as an input to the subsequent step):

1) Calibration of MFCs against a primary flow reference (Weighted Least-Squares regression will be adopted). The calibrated MFCs are then employed to dilute a static calibration gas mixture with a diluent gas to obtain reference gas mixtures having the analyte amount fraction in the range of interest (for environmental monitoring applications).
2) Starting from the classic model equation of the dynamic dilution, evaluation of the uncertainty associated with the amount fraction of the analyte within the obtained mixtures, by taking into account contributions arising from i) the flow of the parent mixture and that of the dilution gas, ii) the amount fraction of the analyte in the parent mixture and iii) the impurities of the analyte gas in the diluent gas.
3) Modelling and calculation of covariances between different levels of the analyte amount fraction obtained (in the different mixtures) by dilution using the same MFCs.
4) Use of the obtained NO_X reference mixtures to establish a response curve for a chemiluminescence analyser in the desired range of mass fractions (Weighted Total

Least-Squares regression will be applied, taking into account uncertainties of and covariances among the values of both the dependent and independent variables).

Some calibration cases will be presented together with corresponding validation studies, verifying the impact of uncertainties and covariances through the whole chain of measurements, up to the final estimates produced by the calibrated chemiluminescence analyser.

This project has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

References

[1] http://empir.npl.co.uk/emue/
[2] ISO 6145-7:2018 "Gas analysis – Preparation of calibration gas mixtures using dynamic methods – Part 7: Thermal mass-flow controllers"
[3] Willem Kool, JCGM-WG1-SC5-N12-15 "JCGM Survey (GUM) – Collated responses", 2012-07-19, https://www.bipm.org/wg/JCGM/JCGM-WG1/Allowed/sub-committee_5/WG1-SC5-N12-15_JCGM_GUM_Survey_Collated_responses.pdf
[4] Directive 2008/50/EC of the European Parliament and of the Council on ambient air quality and cleaner air for Europe
[5] EN 14211:2012 "Ambient air – Standard method for the measurement of the concentration of nitrogen dioxide and nitrogen monoxide by chemiluminescence"

Assessing the uncertainty contribution of detection thresholds to the uncertainty of frequencies of radiation-induced and background DNA damage foci obtained by automatic scoring

Ana Belchior[1], João Canhoto[2,3], Ulrich Giesen[3] and Hans Rabus[3]

[1] Centro de Ciências e Tecnologias Nucleares (C2TN), Instituto Superior Técnico, Universidade de Lisboa, E.N. 10 ao km 139.7, 2695-066 Bobadela LRS, Portugal
[2] Faculdade de Ciências da Universidade de Lisboa (FCUL), Campo Grande 016, 1749-016 Lisboa, Portugal
[3] Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany
E-mail (corresponding author): hans.rabus@ptb.de

The field of radiobiology studies the effects of ionizing radiation on living organisms. In radiobiological studies of early radiation effects to cells, immune-fluorescence assays are a frequently used technique for the detection of radiation-induced lesions to the double-helix deoxyribose nucleic acid (DNA) molecule. In these assays, the occurrence of DNA strand breaks is detected via fluorescence-tagged proteins. These molecules are involved in DNA damage repair and, hence, aggregate at the location of a DNA strand break, and these aggregations are observed in microscopy as fluorescent foci after stimulation of the cells with ultraviolet light.

A complication in the analysis of these assays is that foci also appear in cells that have not been exposed to ionizing radiation. The frequency of occurrence of such background foci depends on the protein used for detection and on the type of cell (cell line) investigated. Background foci may be artefacts or originate from DNA breaks that have occurred spontaneously or were induced by the handling of the cells (e.g. illumination of the cells for locating them). The contribution of background foci to the observed frequencies of foci in irradiated cells is generally assessed by studying additional cells samples that are treated in the same way as the irradiated cells as control groups. Cells that passed all steps of the treatment protocol except the irradiation are generally called sham-irradiated cells.

The present study builds on data produced during the BioQuaRT project [1], where human umbilical vein endothelial cells (HUVEC) were irradiated at the ion microbeam facility of PTB with alpha particles of 8 MeV, 10 MeV or 20 MeV energy or with protons of 3 MeV and a spatial targeting uncertainty accuracy of typically 1.6 µm (k=1). In each irradiation session all cells within a sample dish containing several thousands of cells were irradiated within typically 20 minutes. Each cell nucleus was targeted with 5 projectiles in a quincunx pattern that was chosen such as to assure an optimum between distinguishability of the foci forming at the loci of different tracks and minimizing the probability of projectiles passing outside the cell nucleus [2].

After the irradiation, the cells were fixed after about 30 minutes (corresponding to the time of maximum foci occurrence). Fluorescence microscope images of irradiated and sham-irradiated cells were taken later at the C2TN laboratories. In the present work, a second analysis of the microscopy images was performed using the software CellProfiler (CP) that was used for identifying of cell nuclei, for determining their cross-sectional area and

Assessing the uncertainty contribution of detection thresholds to the uncertainty of frequencies of radiation-induced and background DNA damage foci obtained by automatic scoring

orientation, and for counting the foci in each of them. The raw data obtained in this way were further processed applying filters to discriminate nuclei in division, incompletely imaged nuclei as well as objects that were erroneously identified by the software as foci. The filter criteria applied to the area of nuclei as well as to the size and intensity of the foci.

The frequency distributions of foci in irradiated cells and sham-irradiated cells were used to determine the parameters of model functions for the probability distributions of background foci and radiation-induced foci, respectively. The key parameter of these model functions is the probability p_f that a focus is observed at the location where a projectile track passed the cell nucleus. The other parameters are related to the experimental conditions at the ion microbeam and most of them can be determined independently. The parameter p_f is a measure of the biological effectiveness of the respective radiation quality (projectile type and energy) and may be considered as “the” measurement quantity in this type of assays.

In the data analysis of the sham-irradiated samples it was found that foci frequencies did not follow the expected Poisson distribution and this finding was independent on the choice of aforementioned thresholds. For the ensuing analysis of irradiated cell samples, the model parameter p_f was found to have values below 0.7, in agreement with values reported recently by Gonon et al. [2]. As the value of p_f was also found to be sensitive on the choice of thresholds, we exploited this sensitivity to determine the contribution of the thresholds to the uncertainty budget to be attributed to the value of p_f as determined by the experiment and analysis. Details will be reported at the conference.

[1] H. Rabus et al., EPJ Web of Conferences 77, 00021 (2014)
[2] G. Gonon et al., Radiation Research, accepted for publication 18.03.2019.

Activities of EURADOS Working Group 6 "Computational Dosimetry"

H. Rabus[1], J. M. Gómez-Ros[2], C. Villagrasa[3], J. Eakins[4], T. Vrba[5], V. Blideanu[6], M. Zankl[7], R. Tanner[4], L. Struelens[8], C. Domingo[9], G. Baiocco[10], B. Caccia[11], C. Huet[3]

[1] Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany
[2] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
[3] Institut de Radioprotection et de Sùreté Nucléaire (IRSN), Fontenay-aux-Roses, France
[4] Public Health England (PHE), Didcot, United Kingdom
[5] Czech Technical University in Prague (CTU), Prague, Czech Republic
[6] Commissariat à l'énergie atomique et aux énergies alternatives (CEA), Saclay, France
[7] Helmholtz Zentrum München German Research Center for Environmental Health (HMGU), Neuherberg, Germany
[8] Belgian Nuclear Research Center (SCK•CEN), Mol, Belgium
[9] Universitat Autonoma de Barcelona (UAB), Barcelona, Spain
[10] Physics Department, University of Pavia, Pavia, Italy
[11] Italian National Institute of Health (ISS), Rome, Italy
E-mail (corresponding author): hans.rabus@ptb.de

The European Radiation Dosimetry Group (EURADOS) e.V. is an association of more than 70 European institutions and 560 individual scientists as associate members. The mission of EURADOS is to promote the scientific understanding and technical development of dosimetry in the fields of radiation protection, radiobiology, and medical use of ionizing radiation (e.g. radiation therapy and diagnostic radiology) by stimulating collaboration between European research institutions. Currently, EURADOS has eight Working Groups (WGs) which organise scientific meetings, training activities, intercomparisons and benchmark exercises for promoting quality assurance [1].

WG 6 "Computational Dosimetry" has a cross-sectional role and promotes good practice in the application of computational methods for radiation dosimetry in radiation protection and the medical use of ionizing radiation. As computational methods are widely used in radiation protection and other areas of radiation dosimetry, e.g. in the design of experiments and in the interpretation of results, WG6 strongly engages in collaborations with the other WGs that are focused on subject areas rather than techniques. These collaborations cover a large range of current topics in radiation dosimetry including more fundamental studies of radiation effects in complex systems. In addition, WG 6 also performs scientific research and development as well as knowledge transfer activities, such as training courses.

Monte Carlo techniques, including the use of anthropomorphic and other numerical phantoms based on voxelized geometrical models, have a strong part in the activities pursued in WG6. However, other aspects and techniques, such as neutron spectra unfolding, play an important role as well. As a kind of hybrid activity between knowledge transfer and research, a number of intercomparison exercises have been carried out in the past where participants were invited to solve predefined computational problems with some freedom in choosing their methodology and approach [2-6].

Activities of EURADOS Working Group 6 "Computational Dosimetry"

In the past, the main purpose of such exercises was to provide information on the accuracy with which computational methods are applied and whether best practice is being followed. Within the still ongoing exercises on neutron spectra unfolding [5] and on assessing the uncertainty contribution of cross sections used as input parameters in track structure codes [6], the focus has changed towards assessing the uncertainty that can be achieved with these computational methods. Furthermore, the future strategy of WG 6 also includes an extension of the scope toward experimental benchmark activities and evaluation of cross sections and algorithms with the vision of establishing a gold standard for Monte Carlo methods used in medical and radiobiological applications.

[1] W. Rühm et al., The work programme of EURADOS on internal and external dosimetry *Annals of the ICRP* **47** 20-34 (2018)
[2] T. Vrba et al., EURADOS intercomparison exercise on MC modelling for the in-vivo monitoring of Am-241 in skull phantoms (Part I). R. Radiat. Phys. Chem. 104, 332-338 (2014)
[3] T. Vrba et al., EURADOS intercomparison exercise on MC modelling for the in-vivo monitoring of AM-241 in skull phantoms (Part II and III). Radiation Physics and Chemistry, 113, 59-71 (2015)
[4] B. Caccia et al., EURADOS intercomparison exercise on Monte Carlo modelling of a medical linear accelerator, Annali dell'Istituto Superiore di Sanità 53, 314-321 (2017)
[5] J. M. Gómez-Ros et al., International comparison exercise on neutron spectra unfolding in bonner spheres spectrometry: problem description and preliminary analysis, Radiation Protection Dosimetry 180 (1-4), 70–74 (2018)
[6] C. Villagrasa et al., Assessing the contribution of cross-sections to the uncertainty of Monte Carlo calculations in micro- and nanodosimetry, Radiation Protection Dosimetry ncy240 (2018), https://doi.org/10.1093/rpd/ncy240.

Proper modelling of errors in B1-mapping for uncertainty quantification in electric properties tomography

A. Arduino[1], O. Bottauscio[1], M. Chiampi[1], L. Zilberti[1]

[1]Istituto Nazionale di Ricerca Metrologica (INRiM), Strada delle Cacce 91, 10135 Torino, ITALY
E-mail (corresponding author): a.arduino@inrim.it

Electric properties tomography (EPT) is a novel quantitative imaging method based on magnetic resonance imaging (MRI) that promises to become, in the next future, a powerful non-invasive diagnostic technique for detection, characterisation and monitoring of pathologies, like breast cancer [1].
In order to provide a reliable map of the electric properties within the patient body, many methods have been proposed [2] that elaborates the MRI scanner transmit sensitivity, B_1^+, which is estimated by the scanner itself with a so-called B1-mapping technique [3]. Anyway, attempts to quantify the uncertainty propagation have been made only for few EPT strategies [4, 5], rarely accounting for non-Gaussian random errors [6].
The main difficulty in performing an accurate uncertainty quantification in any EPT strategy is the existence of different B1-mapping techniques, each one with a different random error distribution and non-systematic bias in the low signal regions [3]. To overcome this issue, a framework for proper modelling the random errors in any B1-mapping technique is here described.

In literature, there are papers in which the errors in some B1-mapping techniques for B_1^+ magnitude estimation are evaluated by applying the Monte Carlo method and by simulating the MRI acquisitions through the numerical solution of Bloch's equations [3, 6].
Following their example, in order to draw samples of error coherent with the employed B1-mapping technique without knowledge of the actual probability density function (usually not obtainable explicitly [3]), the simulated noise-free MRI images are corrupted in their real and imaginary parts with Gaussian noise of null mean and standard deviation σ, and then elaborated according to the B1-mapping technique to provide a B_1^+ estimate.
The noise-free MRI images can be obtained by an accurate simulation of Bloch's equations, or by an analytical approximation for simple MRI sequences. The standard deviation σ is assumed constant within the image and it is related to the peak signal-to-noise ratio (SNR) by $\mathrm{SNR} = \max(|I|) / \sigma$, where I is the complex-valued noise-free MRI image.

As an example of application of the described framework, the double-angle (DA) B1-mapping technique [3] is considered for B_1^+ magnitude estimation, whereas the transceive phase $\phi^{\pm}$ is directly obtained from the phase of a complex-valued MRI image. The estimation is performed for the transmit sensitivities of an 8-legs birdcage coil and an 8-channels TEM coil, both simulated in two-dimensions by the method of moments [7].
For simplicity, the MRI images are approximated, assuming a homogeneous proton density within the body and an ideally uniform receive sensitivity within the scanner bore, by $I(\boldsymbol{x}) = \sin(\alpha(\boldsymbol{x})) \exp(\mathrm{i}\, \phi^{\pm}(\boldsymbol{x}))$, where the flip angle $\alpha(\boldsymbol{x})$ is proportional to $|B_1^+|$ and such that its

Proper modelling of errors in B1-mapping for uncertainty quantification in electric properties tomography

maximum is equal to the nominal value α_{nom}. The DA technique acquires two images, I_1 and I_2, with α_{nom} equal to 60° and 120°, respectively. The flip angle is then estimated as $\alpha = \arccos(I_2 / (2\, I_1))$.
The results obtained with a peak SNR equal to 100 are reported in Fig. 1. Here, it is possible to recognise the drop in accuracy in the regions with low signal (typical for the channels of TEM coils), where, due to the presence of noise, the DA technique often cannot estimate the flip angle and introduces a significant bias in the recovery. Phase recovery is more robust, showing some sensible noise just in the regions of low signal (which are not present for the birdcage coil).

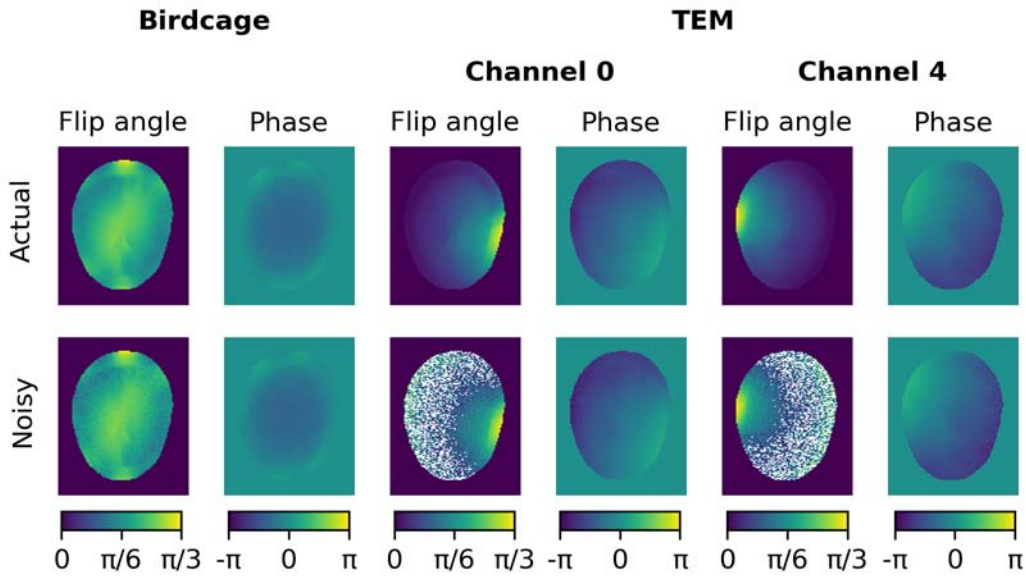


Fig. 1. Noise-free and measurements affected flip angles and B_1^+ phases for different coils.

The described framework for drawing realistic noisy B_1^+ maps can be used for uncertainty quantification in EPT based on the Monte Carlo method or on more advanced approaches. Some examples, comparing different B1-mapping techniques, will be shown in the talk.

[1] S.-Y. Kim, et al. “Correlation between conductivity and prognostic factors in invasive breast cancer using magnetic resonance electric properties tomography (MREPT).” Eur Radiol 2016; 26:2317-26
[2] J. Liu, et al. “Electrical properties tomography based on B1 maps in MRI: principles, applications, and challenges.” IEEE Trans Biomed Eng 2017; 64:2515-30
[3] D.J. Park, et al. “A statistical analysis of the Bloch–Siegert *B*1 mapping technique.” Phys Med Biol 2013; 58:5673-91
[4] S.K. Lee, et al. “Theoretical investigation of random noise-limited signal-to-noise ratio in MR-based electrical properties tomography.” IEEE Trans Med Imag 2015; 34:5100304
[5] A. Arduino, et al. “Monte Carlo method for uncertainty propagation in magnetic resonance-based electric properties tomography.” IEEE Trans Magn 2017; 53:2220-32
[6] S. Gavazzi, et al. “Accuracy and precision of electrical permittivity mapping at 3 T: the impact of three B1+ mapping techniques.” Magn Reson Med 2019; 81:3628-42
[7] A. Arduino, et al. “Magnetic resonance-based imaging of human electric properties with phaseless contrast source inversion.” Inverse Problems 2018; 34:084002

Methodology of teaching the concept of measurement uncertainty

A. Chunovkina[1], V.Sulaberidze[2]

[1] *The D.I.Mendeleyev Institute for Metrology, Moskivsky pr.19, St.Petersburg, Russia*

[2]*Saint Petersburg State University of Aerospace Instrumentation 67, Bolshaya Morskaya str., St. Petersburg, Russia ,*

E-mail : A.G.Chunovkina@vniim.ru

In the Russian Federation, there are several levels in training qualified specialists: three stages of higher education (bachelor, master and specialty, postgraduate) and the level of professional advance training and / or retraining specialists in the field of metrology. Specialists of such a kind work in testing and calibration laboratories and metrological services. The objectives of training, acquired qualifications, knowledge, skills and abilities at all levels correspond to requirements of professional standard "Specialist in Metrology", educational standards: “Standardization and Metrology”, “Metrology, standardization and certification” for bachelors and masters. Postgraduate education is organised according to the requirements of the passports of scientific specialties: ”Instruments and methods of measurement” and “Metrology and metrological assurance”.

The programs of training the concept of measurement uncertainty significantly differ for the listed groups of students. The differences are mainly determined by the a priori knowledge of students in the field of data processing and accuracy assessment, on the one hand, and expectations - the tasks of the students, on the other hand.

We would like to discuss the professional development of specialists, or their retraining. For employees of testing, calibration laboratories and metrological services of enterprises and organizations in the learning process, the following sections/topics are considered:

1. **The notion of accuracy, trueness and precision of measurements. Basic terms in the field of measurement accuracy expression**

The main idea of this section is to form a holistic view of how to express the accuracy of measurements, to “connect” the a priori knowledge of students with the stated concept of measurement uncertainty. In teaching the national and international normative documents as well as basic publications are used (JCGM Guidelines, ISO, OIML documents and others).

2. **Sources of uncertainty and how to calculate uncertainty based on available information**.

In each course of lectures, this section is mostly focused on the area of measurements, which is the field of students’ activity. The main issues are related to the formulation of the measurement model and the use of available information on the accuracy of the measuring instruments used, measurement methods, etc. An important aspect of the lecture courses for specialists of this group is the need for references to the relevant regulatory documents that confirm one or another assessment of uncertainty.

An effective approach in mastering the material is the consideration of examples of listeners, answers on questions from listeners, as well as the consideration of the most common errors in calculating uncertainties. As forms of testing, the tasks of developing

projects for measurement methods, calibration methods, and also answers on questions are used.

3. **Use of measurement uncertainty in conformity assessment, the concept of target uncertainty**

This section of the lectures is mainly based on the following two documents JCGM 106 and OIML G19. The greatest interest and difficulties arise from the questions of applying calibration results when checking the compliance with the maximum permissible error, calculating the instrumental component of measurement uncertainty, calculating correction for systematic bias and some others.

4. **Confirmation of measurement uncertainty**

This section of lectures is often initiated by the wishes of listeners who attend a course of lectures with the aim to prepare their laboratory for accreditation in accordance with ISO / IEC 17025 General requirements for testing and calibration laboratories. In the lectures the procedures for intra-laboratory and inter-laboratory control / confirmation of measurement accuracy are considered. As inter-laboratory control procedures the inter-laboratory comparisons, models and methods for processing experimental data are considered.

The content of postgraduate programs is largely determined by the level of knowledge of students in the field of mathematics, probability theory and mathematical statistics. In the course given a great attention is focused more on theoretical issues of measurement modelling, theoretical-probabilistic methods of calculating uncertainty, uncertainties (pdf) propagation, using software to calculate uncertainties, as well as on comparing probabilistic methods with other approaches to treating uncertain data.

Measurement of Drug-Induced Changes in Cardiac Contractility Using Blood Pressure Data

P.J. Aston[1,2], E. Bonet-Luz[2], M. Nandi[3]

[1]Data Science Group, NPL, Hampton Road, Teddington TW11 0LW, UK
[2]Department of Mathematics, University of Surrey, Guildford GU2 7XH, UK
[3]School of Cancer and Pharmaceutical Science, King's College London, 150 Stamford Street, London SE1 9NH, UK

E-mail (corresponding author): philip.aston@npl.co.uk

All new drug candidates have to be tested for their effect on the heart. Drugs can have the side effect of changing cardiac contractility, which is the strength of contraction of each heartbeat. Such effects are a major cause of drug attrition.

The gold standard for measuring drug-induced changes in contractility in animal studies uses the maximum slope of the pressure in the left ventricle of the heart. The problem with collecting this left ventricular pressure (LVP) signal is that it requires the insertion of a probe directly into the heart. The (continuous) arterial blood pressure (BP) signal is much easier to collect from a peripheral artery. The problem that we consider is whether the changes in contractility shown from the LVP signal can be consistently reproduced using the BP signal.

We use data that was collected from dogs in a HESI-sponsored study that involved different doses of four drugs at six different laboratories [1]. The obvious first choice for monitoring changes in contractility is to consider the maximum slope of the BP data, mimicking the approach used for LVP. However, in many cases this gives a poor correlation with LVP data over a 24-hour post-dose period.

An alternative approach uses the novel attractor reconstruction method that has recently been developed for analysing physiological data [2,3]. In this case, the first derivative of the BP signal is derived which has a large positive spike corresponding to the upstroke of each cycle, which is due to the contraction of the heart. By using Takens' delay coordinates, using the methodology described in [2,3], an attractor with threefold rotation symmetry is generated from this derivative signal. The attractor has three long arms which relate to the large spikes in the data (see Fig. 1). Various candidate quantities associated with the arms of the attractor are derived and traced out over time. The correlation between these and the maximum slope of the LVP signal over a 24-hour post-dose period is

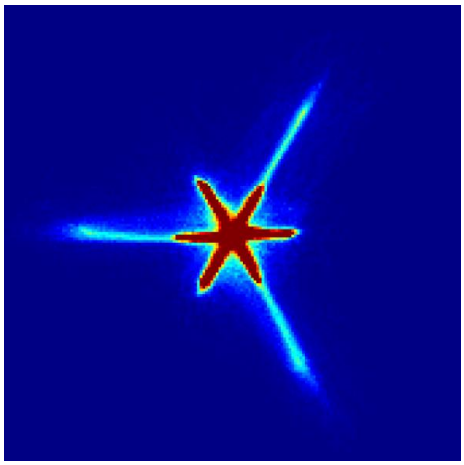


Figure 1: An attractor generated from dog BP data.

obtained and several of these measures show excellent correlations for all the drug doses and for different drugs.

This approach shows that the attractor reconstruction method can be used to derive quantities from BP signals that demonstrate the same changes in cardiac contractility as are observed in the LVP signals and hence provide an alternative method for detecting changes in drug-induced contractility of the heart.

References

[1] B.D. Guth *et al.* The evaluation of drug-induced changes in cardiac inotropy in dogs: Results from a HESI-sponsored consortium. *J Pharmacol Toxicol Methods* 75, 70-90, 2015.

[2] P.J. Aston, M.I. Christie, Y.H. Huang and M. Nandi. Beyond HRV: attractor reconstruction using the entire cardiovascular waveform data for novel feature extraction. *Phys. Meas.* 39, 024001, 2018.

[3] M. Nandi, J. Venton and P.J. Aston. A novel method to quantify arterial pulse waveform morphology: attractor reconstruction for physiologists and clinicians. *Phys. Meas.* 39, 104008, 2018.

Indirect multiparameter measurements with correlated uncertainties

Z.L. Warsza[1], J. Puchalski[2],

[1] Industrial Research Institute of Automation and Measurements (PIAP), Warszawa, Poland
E-mail: (corresponding author) zlw1936@gmail.com
[2] Central Office of Measures (GUM), Warszawa Poland, E-mail: jacek.puchalski@gum.gov.pl

The paper presents an upgraded version of the vector method of evaluation of multiparameter measurement uncertainties stated in the Supplement 2 to GUM guide. This was done on the example of two-parameter jointed measurements. It consists the correlation of individual components of the type A and/or type B uncertainties of input measurands. The general formulas for the covariance matrix, final uncertainties and correlation coefficient were determined. The 3D graph shows the correlation coefficients of the output quantities as a function of the type B contributions in the uncertainty of two input quantities. It has been demonstrated that the inclusion of correlations of uncertainty components makes the uncertainty evaluations more reliable and accurate.

The GUM Guide on the expression of uncertainty in measurement is based on the determination of the type A and type B uncertainty components, designated as u_A and u_B, respectively [1]. The standard deviation of the $p(x)$ distribution of the single measurand values, i.e. its uncertainty u, is the geometric sum of statistically independent component uncertainties u_A and u_B:

$$u = \sqrt{u_A^2 + u_B^2} \qquad (1)$$

After the GUM guide creation, in many items of literature discussed how in the one parameter measurements determine in practice the type B uncertainty in different environments, including those caused by correlated random quantities. But in literature on multivariable measurements we have not found an analysis of impact of the correlation of uncertainty components type A and/or type B of two different input quantities. Proposal, how to do that is given below. This task is explained in Fig 1. It is to determine the $\boldsymbol{U_X}$ covariance matrix considering correlations (Fig. 2) of uncertainty components type A or/and type B of pairs of input quantities of multimeasurand $\boldsymbol{X}$ by the $\boldsymbol{U}_{AB}$ matrix.

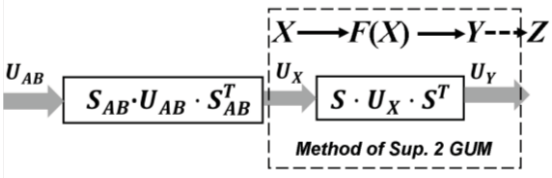


Fig. 1. The extended vector method of propagation uncertainty for multiparameter measurements: $\boldsymbol{X}$ and $\boldsymbol{Y}$ - input and output multivariable measurands; $\boldsymbol{U_X}, \boldsymbol{U_Y}$ – their covariance matrixes, $\boldsymbol{U_{AB}}$ – input matrix with correlated uncertainty components of type A or/and B of $\boldsymbol{X}$ quantities.

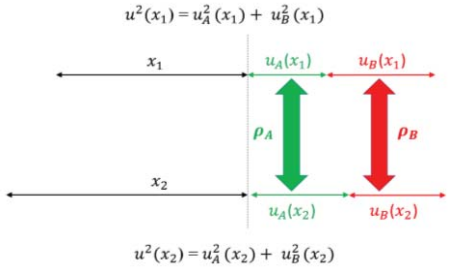


Fig 2. Relations of Type A and/or Type B correlated uncertainties of the 2D measurand

Main Equations for 2D measurand X

$$\boldsymbol{U_X} = \boldsymbol{S_{AB}}\boldsymbol{U_{AB}}\boldsymbol{S_{AB}}^{\mathrm{T}} = \begin{bmatrix} u_{x1}^2 & \rho_{x1,2}u_{x1}u_{x2} \\ \rho_{x1,2}u_{x1}u_{x2} & u_{x2}^2 \end{bmatrix} \qquad (2)$$

Indirect multiparameter measurements with correlated uncertainties

Where:

$$\boldsymbol{U_{AB}} = \begin{bmatrix} u_{1A}^2 & \rho_A u_{1A} u_{2A} & 0 & 0 \\ \rho_A u_{1A} u_{2A} & u_{2A}^2 & 0 & 0 \\ 0 & 0 & u_{1B}^2 & \rho_B u_{1B} u_{2B} \\ 0 & 0 & \rho_B u_{1B} u_{2B} & u_{2B}^2 \end{bmatrix} \quad (2a)$$

$$\boldsymbol{S_{AB}} = \begin{bmatrix} \frac{\partial u_1}{\partial u_A} & 0 & \frac{\partial u_2}{\partial u_A} & 0 \\ 0 & \frac{\partial u_1}{\partial u_B} & 0 & \frac{\partial u_2}{\partial u_B} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad (2b)$$

Covariance matrix of correlated u_1, u_2:

$$\boldsymbol{U_X} = \begin{bmatrix} u_{1A}^2 + u_{1B}^2 & \rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B} \\ \rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B} & u_{2A}^2 + u_{2B}^2 \end{bmatrix} \quad (3)$$

Uncertainties

$$u_{x1}^2 = u_{1A}^2 + u_{1B}^2, \quad u_{x2}^2 = u_{2A}^2 + u_{2B}^2 \quad (3a, b)$$

Correlation coefficient

$$\rho_{x1,2} = \frac{\rho_A u_{1A} u_{2A} + \rho_B u_{1B} u_{2B}}{\sqrt{u_{1A}^2 + u_{1B}^2}\ \sqrt{u_{2A}^2 + u_{2B}^2}} \quad (3c)$$

Continuation: by the GUM Supplement 2

$$\boldsymbol{Y}=\boldsymbol{F}(\boldsymbol{X}) \quad (4a) \qquad \boldsymbol{U_Y} = \boldsymbol{S}\boldsymbol{U_X}\boldsymbol{S}^{\mathrm{T}}. \quad (4b)$$

and for the function $\boldsymbol{Z}$ of $\boldsymbol{Y}$

$$\boldsymbol{Z}=\boldsymbol{G}(\boldsymbol{Y}) \quad (5a) \qquad \boldsymbol{U_Z} = \boldsymbol{S_G}\boldsymbol{U_Y}\boldsymbol{S_G}^{\mathrm{T}} \quad (5b)$$

Designations for ratios of components and standard values of uncertainty:

$$0 \le k_{1\mathrm{B}} \equiv \frac{u_{1B}}{u_{x1}} \le 1, \qquad 0 \le k_{2\mathrm{B}} \equiv \frac{u_{2\mathrm{B}}}{u_{x2}} \le 1.$$

Then: $\frac{u_{1A}}{u_{x1}} \equiv \sqrt{1-k_{1B}^2}$, $\frac{u_{2\mathrm{A}}}{u_{x2}} \equiv \sqrt{1-k_{2B}^2}$,

and a simpler pattern of (3c) is obtained

$$\rho_{x1,2} = \rho_A\sqrt{1-k_{1B}^2}\sqrt{1-k_{2B}^2} + \rho_B k_{1\mathrm{B}}\, k_{2\mathrm{B}} \quad (6)$$

For both quantities measured, the coefficient $\rho_{x1,2}$ in (6) depends only on the uncertainty ratio $k_{i\mathrm{B}}$, as:

$$\left(\frac{u_{iB}}{u_{xi}}\right)^2 + \left(\frac{u_{iA}}{u_{xi}}\right)^2 = k_{iB}^2 + 1 - k_{iB}^2 = 1.$$

In a case, when $\rho_A = 0$, $\rho_B = 1$, $k_{1B}^2 \approx k_{1B}^2 \approx \frac{1}{2}$, it is: u_{x1}=u_{x2}=$\sqrt{2}u_B$, and $\rho_{y1,2} = \frac{1}{2}$.

The formula (6) is used for Fig 3 diagrams.

It is the output coefficient function 3D for three pairs of correlation coefficients ρ_A, ρ_B of the relative uncertainties type B, i.e.

$$\rho_{x1,2} = f(k_{1B}^2, k_{2B}^2) \quad (7)$$

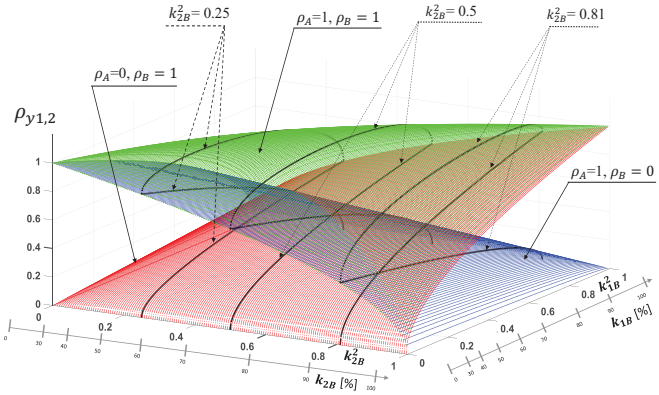


Fig. 3. Relations of correlation coefficient $\rho_{y1,2} = f(k_{1B}, k_{1B})$ of measurand $\boldsymbol{X}$ as 3D charts for the 3 correlation coefficient pairs of its uncertainty components type A or B: $\rho_A = 0, \rho_B = 1$; $\rho_A = 1, \rho_B = 1$; $\rho_A = 1, \rho_B = 0$.

Shorted conclusions:

- The function (6) is invariant to changes of value of $u_{1A}$to u_{1B}, u_{2A} to u_{2B} and ρ_A to ρ_B.
- The maximum correlation coefficient $\rho_{y1,2} = 1$ is achieved for the quantities of fully correlated components of type A and type B, ρ_A=1, $\rho_B = 1$, when $k_{1B} = k_{2B}$. This leads to the condition $\frac{u_{1B}}{u_{1A}} = \frac{u_{2B}}{u_{2A}}$.
- For the values of $k_{1B}^2 \ll \left(1 + \frac{\rho_B^2 k_{2B}^2}{\rho_A^2(1-k_{2B}^2)}\right)^{-1}$ a strong negative correlation for the curves $\rho_A < 0$ is observed.
- If both input values are measured in equal or similar conditions, their uncertainties B-type may be correlated. For example, for the sum of the input quantities and the positive correlation coefficient, the uncertainty will be greater than for the geometric summation of uncertainties of A and B types according to GUM, and for the difference of quantities - smaller.
- Conclusions about uncertainty can be generalized for multivariable measurands.

[1] Evaluation of measurement data – Supplement 2 to the “Guide to the expression of uncertainty in measurement” – Extension to any number of output quantities. JCGM 102:2011.

Towards a new GUM – the suggested draft

I.P. Zakharov[1,2], P.I. Neyezhmakov[1], O.A. Botsiura[2]

[1]National Scientific Centre "Institute of Metrology", 42, Myronosytska str., 61002, Kharkiv, Ukraine
[2] Kharkiv National University of Radioelectronics, 14, Nauka ave., 61166, Kharkiv, Ukraine
E-mail (I.P. Zakharov): newzip@ukr.net

Introduction

In 2014 the WG-1 of the JCGM developed the Committee Draft (CD) of the new (revised) GUM, which based on the Bayesian approach [1]. The basic algorithm for evaluating and reporting uncertainty in CD contain 5 steps:

1. Modelling the measurement;
2. Evaluating input quantities, standard uncertainties and covariances;
3. Evaluating the measurand and standard uncertainty;
4. Determining a coverage interval for the measurand;
5. Reporting and recording measurement result.

The steps 1, 2 and 5 we leave unchanged and describe in this paper our propositions for improvement the steps 3 and 4.

Evaluating the measurand and standard uncertainty

The CD concentrates on measurands that can be described by a linear model or a model that can safely be linearized for the purpose of providing a best estimate of the measurand and the associated standard uncertainty. In this case the best estimate y of Y is given by evaluating expression

$$Y = f(X_1, X_2, ..., X_N)$$

at the best estimates x_i of the input quantity X_i:

$$y = f(x_1, x_2, ..., x_N).$$

Clause 10.3.2 of CD recommended confirmed such estimate y for non-linear measurement models applying the Monte Carlo method of GUM-S1 [2].

For this confirmation we proposed evaluate the bias of the y using the formula:

$$\Delta_y = -\frac{1}{2}\sum_{i=1}^{N} c_{ii} u_i^2 ,$$

where u_i is the standard uncertainty associated with x_i and c_{ii} is the second partial derivative of Y with respect to X_i evaluated at $X_1 = x_1, ..., X_N = x_N$ and compare it with y [3]. If this bias will be significant, it can be used as additive correction to y. This approach, in contrast to CD, is suitable in the presence of uncertainties in the input quantities, estimated by both statistical and non-statistical methods. For facilitation to decide this task we propose used method of finite increments [3].

For independent input quantities $X_1, X_2, ..., X_N$, in clause 10.4.2 of CD standard uncertainty of measurand $u(y)$ is obtained by combining the standard uncertainties $u_1, u_2, ..., u_N$ associated with the best estimates of those quantities in the following way:

$$u^2(y) = \sum_{i=1}^{N} c_i^2 u_i^2 ,$$

where c_i is sensitivity coefficients.

This equation is right for independent input quantities only for linear model. For confirmation on linearity we proposed evaluate the bias of the $u^2(y)$ using the formula [4]:

$$\Delta_{u^2} = \frac{1}{4}\sum_{i=1}^{N} c_{ii}^2(\eta_i + 2)u_i^4 + \sum_{i=2}^{N}\sum_{j=1}^{i-1} c_{ij}^2 u_i^2 u_j^2 ,$$

where c_{ij} is the mixed partial derivative the second order of Y with respect to X_j, X_i evaluated at $X_1 = x_1, ..., X_N = x_N$; η_i is the kurtosis of the *i*-th input quantity. If this bias will be significant, it can be used as additive correction to $u^2(y)$.

Determining a coverage interval for the measurand

The main disadvantage of the CD is independence of estimates of expanded uncertainty from the distributions of input quantities. Conservative coverage factors, derived in CD from Chebyshev and Gaussian inequalities, give overstated estimates of the expanded uncertainty. To eliminate the disadvantages we proposed to apply the kurtosis method [5].

In this case expanded uncertainty $U_{0.95}$ given by the product of a coverage factor $k(\eta)$ and the standard uncertainty $u(y)$ associated with y:

$$U_{0.95} = k(\eta) \cdot u(y).$$

The coverage factor $k(\eta)$ for $p = 0,95$ is calculated by the formula [5]:

$$k_{0.95} = \begin{cases} 0,1085\eta^3 + 0,1\eta + 1,96, \text{ at } \eta < 0; \\ 1,96, \text{ at } \eta \geq 0, \end{cases}$$

where η is the kurtosis of the measurand that found by the formula [6]:

$$\eta = \left(\sum_{i=1}^{N} \eta_i c_i^4 u_i^4 \right) \Big/ u^4(y).$$

It is shown that the deviation of the estimates of the expanded uncertainty, obtained by the kurtosis method from the estimates obtained using the Monte Carlo method, does not exceed ± 2.5% for the number of repeated measurements of input values more than 5.

References

[1] Bich, W., Cox, M., Michotte, C. (2016). Towards a new GUM – an update. Metrologia, 53, S149–S159.

[2] JCGM 101:2008 Evaluation of measurement data − Supplement 1 to the "Guide to the expression of uncertainty in measurement" − Propagation of distributions using a Monte Carlo method.

[3] Zakharov, I., Neyezhmakov, P., Botsiura, O. (2018) Reduction of the measurand estimate bias for nonlinear model equation. *Journal of Physics: Conf. Series* 1065, 212002.

[4] Zakharov, I., Botsiura O. (2018) Measurement Uncertainty Evaluation in Nonlinear Model Equations. Metrology and Metrology Assurance 2018: Proceedings of 28-th International Scientific Symposium, September 10-13, 2018, Sozopol, Bulgaria, p. 35-38.

[5] Zakharov I. Botsiura O. (2019) Calculation of expanded uncertainty using the kurtosis method when implementing a Bayesian approach. calculation of expanded. *Measurement techniques.* Volume 62, Issue 4.

[6] P. Täubert (1987). Abschätzung der Genauigkeit von Meßergebnissen. VEB Verlag Technik, Berlin.

CASoft: An approach for the verification of software for conformity assessment

I. M. Smith[1], P. M. Harris[1], J. Nguyen[1], A. Allard[2], N. Fischer[2], L. Pendrill[3]

[1]National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 OLW, United Kingdom
[2]Laboratoire National de métrologie et d'Essais, 1 rue Gaston Boissier, 75724 PARIS Cedex 15, France
[3] RISE Research Institutes of Sweden AB, Eklandagatan 86, SE-412 61, Göteborg, Sweden
E-mail (corresponding author): ian.smith@npl.co.uk

The European Metrology Programme for Innovation and Research (EMPIR) Support for Impact Project "Software to maximize end user uptake of conformity assessment with measurement uncertainty" (short name "CASoft") is concerned with enabling end users to undertake calculations for risk-based decision-making in conformity assessment. This objective is being achieved through the development and provision of a software tool to implement the relevant calculations. The software is being developed using the MATLAB App Designer and will be made freely available as a Windows executable.

It is imperative that the software undergoes appropriate testing prior to its release. A key aspect of testing is verification, i.e., ensuring that calculations have been implemented correctly. The European Metrology Research Programme (EMRP) Joint Research Project "Traceability for computationally-intensive metrology" (short name "TraCIM") [1] developed a framework for the verification of mathematical software. The framework is based on three key aspects: computational aims, reference pairs, and an information and communications technology (ICT) infrastructure.

The availability of specifications of computational aims, i.e., complete and unambiguous descriptions of the calculations to be undertaken, is fundamental to the verification of implementations of those calculations. The document [2], prepared by the Joint Committee for Guides in Metrology (JCGM) and an accompanying document to the Guide to the Expression of Uncertainty in Measurement (GUM) [3], provides details of the calculations to be undertaken by the CASoft software. For each calculation, a specification of the computational aim is developed that lists the inputs to and outputs of the calculation, and describes the mathematical calculations to be undertaken.

Having specified a computational aim, software is then developed to generate reference pairs. A reference pair comprises reference data (inputs) and reference results (outputs). Software under test is applied to reference data to obtain test results. The test results are then compared with corresponding reference results. The process is repeated for many reference pairs to allow an overall assessment of the performance of the software under test to be made.

The ICT infrastructure, referred to as the "TraCIM system", allows verification of mathematical software to be undertaken via the internet. For each calculation, a verification service is developed. The software developer, using a TraCIM client, is able to request

reference data from the TraCIM server. The test results obtained upon processing of the reference data by the software under test are then sent to the TraCIM server which undertakes comparison of corresponding test and reference results, and provides a software evaluation report summarising the performance of the test software.

This paper describes the application of the TraCIM system to the verification of the software developed within the CASoft project. Details of the three main aspects of the verification framework, including the approach to comparing test and reference results, are discussed.

References

[1] A. B. Forbes, I. M. Smith, F. Härtig and K. Wendt. Overview of EMRP Joint Research Project NEW06 “Traceability for Computationally Intensive Metrology”. Advanced Mathematical and Computational Tools in Metrology and Testing X (2015).
[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of measurement data – The role of measurement uncertainty in conformity assessment. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 106: 2012.
[3] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML. Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 100: 2008.

Testing software that implements least-squares fitting for nonlinear models

I. M. Smith[1], J. Nguyen[1]

[1]National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 OLW, United Kingdom
E-mail (corresponding author): ian.smith@npl.co.uk

In metrology, it is frequently required to solve least-squares fitting problems that involve nonlinear models. Solution of such problems generally requires the use of an iterative approach such as the Gauss-Newton algorithm or Levenberg-Marquardt algorithm, and involves the evaluation, or estimation, of partial derivatives of the model function with respect to the model parameters.

Many software packages are available that claim to solve least-squares fitting problems. The user may be allowed to either select a nonlinear model from a list of commonly-encountered models or specify a 'user-defined' model, e.g., as a textual expression. In general, the user is required to select options (e.g., the algorithm to implement, the method of estimating partial derivatives) or provide numerical information (e.g., tolerance values for convergence criteria, values of initial estimates of the model parameters) that can significantly influence the accuracy of the results returned by the software. In many cases, there is little guidance provided to the user on the consequences of options selected prior to processing data. Often, guidance that is provided is almost meaningless. For example, while software documentation may state that the quality of results returned by the software is dependent on the initial estimates of the model parameters provided by the user, how does such a statement help the user? Onus is therefore put on the user to ensure that the software is fit for purpose.

This paper considers a specific nonlinear model, viz., the sum of an exponential decay function and a constant, for which least-squares fitting to data can be implemented in several commercially-available packages. Similarities and differences between the options and numerical information that the user is expected to select and provide when using those packages are discussed. A simple measure that allows the performance of software packages to be assessed is introduced, and results that demonstrate the performance of the software packages are presented.

Quantitative Imaging Biomarkers: the need for Metrology

Nadia A. S. Smith[1], Matt G. Hall[1, 2]

[1]National Physical Laboratory, Hampton Road, Teddington, United Kingdom
[2] UCL Great Ormond Street Institute of Child Health, University College London, London, United Kingdom
E-mail (corresponding author): nadia.smith@npl.co.uk

The National Institute of Health (NIH) defines a biomarker as "…a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention" [1]. A reliable biomarker can be almost anything: the result of a genetic test, a biochemical response, the presence or absence of a certain molecule, or an image-based signature of some important structural or physiological change. The concept of a biomarker changes greatly from field to field, but one concept unifies them all: biomarkers are measured. As such, biomarkers have an associated uncertainty, and the devices used to measure them require calibration against some agreed-upon standard or procedure.

When considered from this perspective, the importance of both reproducibility and traceability of biomarkers is very clear. Biomarkers are the result of measurements and are used to indicate the presence or progress of some biomedical object of inquiry. To be useful they must be reproducible so that clinicians and researchers can be confident that their trials are based on usable data, and in turn patients can be confident that the results of tests and the clinical conclusions those tests support are as reliable as possible. Additionally, to be confident that biomarkers are reliable and deployable in different contexts, the values of biomarkers must be traceable to primary standards. Unfortunately, these are issues which are frequently overlooked in the development of novel biomarkers.

Quantitative imaging biomarkers (QIBs) are of particular interest in this paper. These are biomarkers measured using quantitative imaging, which is defined as "the extraction of quantifiable features from medical images for the assessment of normal [findings] or the severity, degree of change, or status of a disease, injury, or chronic condition relative to normal [findings]" [2].The medical imaging community has been very active in developing QIBs, but few of them have been rigorously evaluated for their technical and clinical performance. This omission is possibly due to an inconsistent use of terminology and methods related to evaluation of the technical performance of these markers [3]. Standardisation and metrology can play an important role to improve this issue by helping to describe accurately and consistently what physical phenomenon is being measured, its relation to disease progress or outcome, and the measurement uncertainty associated with it.

One important method for validating measurements and providing traceability is through the use of well-characterised test objects or phantoms. National Measurement Institutes such as NIST, NPL, PTB and LNE, are already working on aspects of validation with the

development and testing of numerical and physical imaging phantoms. A phantom provides traceable, verifiable ground-truth values against which quantitative measurements can be compared. They can be used to validate image processing pipelines resulting in a better understanding of the uncertainty and bias in that process or even mimic pathological changes or disease states.

In this paper, several case studies of validation of physiological measurements using phantoms will be presented, including measurements of perfusion in the myocardium and the brain and other quantitative MR parameters. The data analysis tools will be discussed, paying particular attention to the application of well-known metrological techniques within a relatively new application area for the field of quantitative imaging. We will also describe the current state-of-the-art in imaging biomarker traceability and make recommendations for future developments.

References

[1] NIH Biomarkers Def. WG, Clin Pharmacol Therap. 69:89-95 (2001).

[2] Quantitative Imaging Biomarkers Alliance. http://rsna.org/QIBA/ Accessed April 27, 2019.

[3] Sullivan et al., Metrology Standards for Quantitative Imaging Biomarkers, Radiology: Volume 277: Number 3—December 2015.

Efficient sampling plans for the EU measuring instruments directive (MID)

C.A. Müller[1], K. Klauenberg[2], C. Elster[2]

[1]*Deutsche Akademie für Metrologie, Bayerisches Landesamt für Maß und Gewicht,83435 Bad Reichenhall, Germany*
[2]*Physikalisch-Technische Bundesanstalt, Abbestr. 2-12, 10587 Berlin, Germany*
E-mail (corresponding author): cord.mueller@lmg.bayern.de

The European Measuring Instruments Directive (MID) [1] harmonizes legal metrology within the EU. It provides requirements for different types of measuring instruments, encompassing utility meters for water, gas, electricity and heat, as well as automatic weighing instruments, various material measures, taximeters, exhaust gas analysers, and many more. The MID's conformity assessment modules F and F1, for product verification by a notified body, provide the option to either test every instrument or to proceed by "statistical verification". For the latter, modules F and F1 require that attribute sampling plans ensure

"(a) a level of quality corresponding to a probability of acceptance of 95 %, with a non-conformity of less than 1 %;

(b) a limit quality corresponding to a probability of acceptance of 5 %, with a non-conformity of less than 7 %."

How should these legal requirements be translated into mathematical relations, so that sampling plans can be derived? In fact, the requirements are ambiguous: Firstly, what precisely do the clauses "with a non-conformity of p%" mean? Secondly, an exact "probability of acceptance of (9)5 %" is almost never met with discrete sampling plans and finite-sized lots.

The WELMEC guide 8.10 [2] recommends ISO 2859-1 [3] table 1 and 2 A, but admits: "There is no standard covering the MID conditions to the full extent".[1] Concerning the interpretation of (a) and (b), it states: "The OC curves have to be on the left hand side of the points" $(P_{AQL},p_{AQL}) = (95\%,1\%)$ and $(P_{LQ},p_{LQ}) = (5\%,7\%)$, referring to the operating characteristic (OC) curve, i.e. the probability of acceptance P as a function of the quality level p. Together with the OC's monotonicity, this implies the two mathematical conditions

(a*) $p = p_{AQL} = 1\% \Rightarrow P(p) < P_{AQL} = 95\%$, and

(b*) $p \geq p_{LQ} = 7\% \Rightarrow P(p) < P_{LQ} = 5\%$

for a sampling plan to be admissible. Recent work [4] derived from this interpretation a set of optimized sampling plans, which are being applied already by the notified body 0104 of Bavaria, Germany.

[1] In fact, sampling plans of ISO 2859-1 are ill-suited here since, firstly, they apply to processes, as opposed to single lots that are the object of the MID modules F and F1 (see 12.6.1 in [3]), and, secondly, they do not contain the pairs of values $(P_{AQL},p_{AQL}) = (95\%,1\%)$ and $(P_{LQ},p_{LQ}) = (5\%, 7\%)$ defined by (a) and (b).

Efficient sampling plans for the EU measuring instruments directive (MID)

However, conditions (a*) and (b*) do not fit into the paradigm of statistical hypothesis tests (as described, e.g., in [5] in the context of legal metrology), often employed for developing sample plans in practice. Therefore, we propose to consider an alternative interpretation based on statistical hypothesis tests that results in the conditions

(a') $p = p_{AQL} = 1\% \Rightarrow P(p) \geq P_{AQL} = 95\%$, (the null-hypothesis), and

(b') $p \geq p_{LQ} = 7\% \Rightarrow P(p) < P_{LQ} = 5\%$ (the one-sided alternative hypothesis).

Note that conditions (b') and (b*) coincide, whereas conditions (a') and (a*) are inverted such that now both producer's and consumer's risks are bounded from above (somewhat paradoxically, (a*) bounds the producer's risk $1 - P(p_{AQL})$ from below - see Fig. 1). This interpretation is consistent with ISO 2859-1 (see def. 3.1.20), which devises sampling plans to secure a high probability of acceptance for quality levels $p \leq p_{AQL}$.

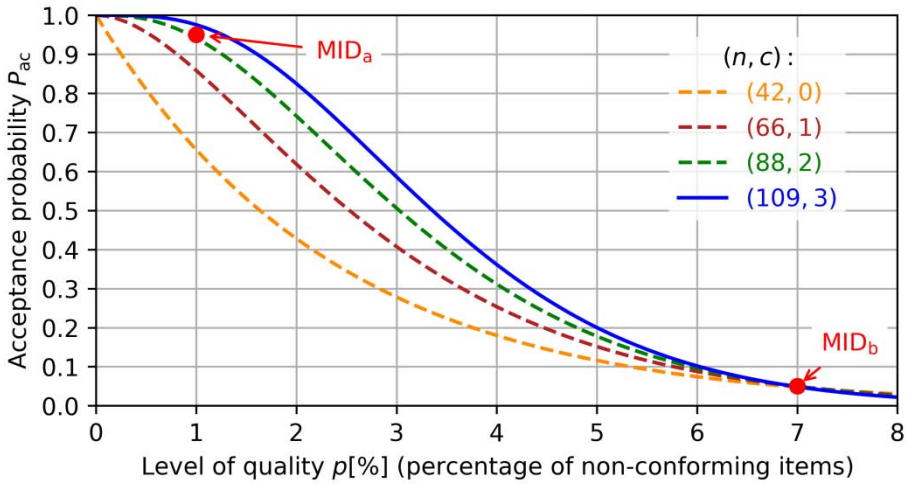


Figure 1: Operating characteristics of optimized sample plans satisfying criteria (a*) and (b*) (dashed lines, from [4]) as opposed to the alternative criteria (a') and (b') (solid line, this work). (n,c) denote sample size and acceptance number, respectively. Curves are computed using the binomial model valid for very large lots.

For this work, we optimize sampling plans under the constraints (a') and (b') while considering both large lots (using a binomial model) and finite-size lots (using a hypergeometric model). These plans cover the MID conditions to the full extent. For the two different interpretations, we discuss and illustrate the general implications, contrast the resulting sampling plans and conclude by comparing advantages and disadvantages of each approach.

[1] Directive 2014/32/EU of the European Parliament and of the Council of 26 February 2014, OJ L96/149 (2014)
[2] WELMEC 8.10 Guide for generating sampling plans for statistical verification according to Annex F and F1 of MID 2014/32/EU, (2018)
[3] ISO/TC 69/SC 5, "Sampling procedures for inspection by attributes - Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection", ISO 2859-1:1999(E)
[4] C. A. Müller, J. Appl. Stat., DOI: 10.1080/02664763.2019.1588235, (2019)
[5] K. Klauenberg and C. Elster, Metrologia 54,pp. 59-68, 2017.

Distribution Detection and Information Loss in a Measurement Uncertainty Network

P. M. Duncan[1], D. S. Whittaker[1]

[1]National Physical Laboratory - Scotland, Technology & Innovation Centre, 99 George Street, G1 1RD, Glasgow, Scotland
E-mail: paul.duncan@npl.co.uk, david.whittaker@npl.co.uk

The development of a measurement uncertainty network comprising an information flow along the supply chain from its inception to an application endpoint is proposed and discussed. This network will contain a traceable set of internally consistent measurement models connected through common quantities. Various approaches to calculating measurement uncertainty can be taken at each stage and, crucially, how to represent this uncertainty as a transferrable input to the next stage in the chain.

Conventional approaches include calculation of the first and second moments of the distribution or Monte Carlo sampling. Thereby obtaining, non-parametrically, both the mean and standard deviation of the data concerned or the parameters defining some supplied or assumed representative distribution. This layered method, however, where the first and second distribution moments at each stage are calculated before being passed on to the next is undermined by aggregated information loss from layer to layer. This information loss can affect downstream decision making. Parametric sampling is also dependent on the veracity of the supplied sample distribution, and can be computationally expensive.

To address this concern, we propose a method for quantifying the information loss and its potential compounded effects on a network. Focusing on the determination of a probability distribution from sample data and how that information is propagated, we utilise the Kolgomorov-Smirnov test to determine the best-fitting standard probability distribution function for each dataset. Then we attempt to measure the information lost through assuming this distribution by employing methods such as the Kullback-Leibler divergence to determine entropic loss. Understanding this information loss could help minimise and avoid decisions influenced by misleading data representation.

Obtaining High Accuracy Measurements of Brain Stents From 2D X-Ray Images

P.J. Aston[1], J. Talbott[1], N. McCormick[1]
K. Spranger[2], L. Asner[2], M. Tsakalakis[2]

[1]NPL, Hampton Road, Teddington TW11 0LW, UK
[2]Oxford Heartbeat Ltd, Health Foundry, Canterbury House,
1 Royal Street, London SE1 7LL, UK

E-mail (corresponding author): philip.aston@npl.co.uk

Oxford Heartbeat is an early-stage start-up who have developed software that helps surgeons plan and rehearse minimally invasive stent placements inside blood vessels in the brain. Once the stent has been deployed, two x-ray scans are taken from different angles from which the location of the deployed stent is required.

A framework is established taking account of the geometry of the x-ray device. Three triangulation methods are described for deriving a point in 3D from a point identified in each of the two 2D images using homogeneous image coordinates. Two of the methods involve solving linear systems of equations while the third requires the singular vector of a matrix associated with the smallest singular value.

We then perform a Monte Carlo uncertainty analysis for each of the three methods in which the image coordinates are regarded as normally distributed random variables. This results in a point cloud of reconstructed points in 3D. An ellipsoid containing 95% of the reconstructed points is constructed using the method described in Supplement 2 to the *Guide to the expression of uncertainty in measurement* [1]. The volume of the ellipsoid is taken as an overall measure of uncertainty and it is found that this is very similar for the three different reconstruction methods used. The reconstructed points in 3D are then projected onto the centreline of the stent and a one-dimensional distribution along the centreline is also derived, together with a 95% coverage interval.

An analytic solution of two of the methods can be obtained and so the uncertainty analysis is repeated for these methods using the uncertainty framework which requires the covariance and sensitivity matrices [1]. This approach gives accurate results without the requirement for the many iterations involved in the Monte Carlo method.

References

[1] Evaluation of measurement data - Supplement 2 to the “Guide to the expression of uncertainty in measurement” - Extension to any number of output quantities. JCGM 102:2011.

Non-Informative Bayesian Inference for Heterogeneity in a Generalized Marginal Random Effects Meta-Analysis

O. Bodnar[1]

Mälardalen University, Västerås, Sweden
E-mail (corresponding author): olha.bodnar@mdh.se

This presentation describes an objective Bayesian inference procedure for the heterogeneity parameter in a generalized marginal random effects model. Models of this kind are widely used in meta-analysis and in inter-laboratory comparisons. Assuming that the data can be adequately modeled using elliptically contoured distributions, we derive a reference prior for the model parameters, and produce an analytical expression of the corresponding posterior. We also state necessary conditions for the resulting posterior to be proper and for its first two moments to be finite. The general theoretical results are illustrated for three well-known families of elliptically contoured distributions: normal, Student's t, and Laplace.

Effective number of degrees of freedom and prior information

A B Forbes

Data Science Group, National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, UK

E-mail: alistair.forbes@npl.co.uk

This paper is concerned with the approximation of data using linear models for which there is prior information associated with some or all of the model parameters ***a.*** This situation arises, for example, in Tikhonov regularisation, Bayesian inference and approximation with Gaussian process models. In the approximation process, the m-vector of data $\boldsymbol{y}$ is approximated by a linear function H$\boldsymbol{y}$ of the data, where H is an m x m matrix. The effective number of degrees of freedom [1] associated with the model is given by the sum of the eigenvalues of H. For standard linear regression, the matrix H is a projection and has n eigenvalues equal to 1 and all others zero, where n is the number of parameters in the model. Thus, the effective number of degrees of freedom is equal to the number of free parameters in the model.

Incorporating prior information reduces the effective number of degrees of freedom since the ability of the model to approximate the data vector $\boldsymbol{y}$ is constrained by the prior information. We give a general approach for providing bounds on the effect number of degrees of freedom for models with prior information on the parameters and illustrate the approach on common problems. In particular, we show how the effective number of degrees of freedom depends on spatial or temporal correlation lengths associated with Gaussian processes [2]. The correlation lengths are seen to be tuning parameters used to match the model degrees of freedom to the (generally unknown) number of degrees of freedom associated with the system giving rise to the data.

[1] Hastie et al.: Elements of Statistical Learning, Springer, New York, 2011
[2] Rasmussen and Williams: Gaussian Process for Machine Learning, MIT Press, Cambridge, 2006.

Metrology for virtual measuring systems: new competence centre “VirtMet” at PTB

S. Eichstädt[1], C. Elster[1], I. Fortmeier[1], D. Heißelmann[1], M. Anton[1], T. Kretz[1], S. Schmelter[1], M. Stavridis[1], A. Weissenbrunner[1], T. Wiedenhöfer[1]

[1] Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany

E-mail (corresponding author): sascha.eichstaedt@ptb.de

In the course of the digitalisation in science, economy and society, the importance of simulations and in silico experiments is increasing rapidly. In many areas, so-called "virtual measurements" as simulations based on physical-mathematical modelling and statistical methods are now in everyday use. For example, simulations serve to gain a better understanding of the real experiment, to plan new experiments or to evaluate existing ones. Moreover, simulations are increasingly being used as an essential part of measurement, usually as part of an inverse problem.
In this development, the task of metrology is to secure confidence in simulation results if they are to be used in the same way as real measurements. Concrete existing examples at the Physikalisch-Technische Bundesanstalt (PTB) include the Tilted-Wave Interferometer (TWI) or the Virtual Coordinate Measuring Machine (VCMM). In a national workshop "Metrology for Virtual Measuring Instruments" organised by PTB in March 2018, the following overarching questions and cross-sectional tasks were identified for these and other application examples:

1. How to ensure trust in simulation results?
2. How can virtual and real measurements be compared?
3. What standards are required for interfaces, metadata and data formats?
4. How can virtual experiments for complex measurement systems with large amounts of data be handled using machine learning methods?

Addressing these issues requires continuous and intensive interdisciplinary cooperation. For this reason, PTB launched in 2019 a new competence centre "Metrology for Virtual Measuring Instruments" (VirtMet), in which the existing expertise is bundled, and the interdisciplinary exchange is continuously promoted. In addition, the competence centre will further strengthen the exchange and cooperation with external partners in this area with regular workshops.

Based on concrete research questions, the higher-level questions will be dealt with in cross-sectoral projects. By embedding the projects in the competence centre, there will be an intensive and regular exchange between all participants in order to exploit synergy effects and pursue a joint strategy. The corresponding projects within VirtMet are

A) Simulations for medical imaging methods using X-rays
B) Transfer of the VCMM concept to other areas and for use in digital twins
C) TWI as example for hand-in-hand calibration of real and virtual experiments
D) Development of a virtual flow meter

Metrology for virtual measuring systems: new competence centre “VirtMet” at PTB

Project A) aims at the development of so-called "model observers" for image quality assessment using X-rays and the characterisation of their uncertainty. These developments are supported by using virtual experiments for the assessment and validation of statistical procedures.
Project B) focuses on the dissemination of the established VCMM concept to other metrological areas within PTB and industry through the development of a universal modular system for virtual measurement processes (VMP).
In Project C) the central objective is the determination of uncertainty, which is to be determined by combining the virtual experiment with the real measurement setup. In addition, application of machine learning methods for the simulation part will be addressed.
Project D) addresses the use of computational fluid dynamics (CFD) simulations to simulate different flow configurations and to analyse their effects on the measuring instruments. With the development of a virtual measuring system, flow meters can already be optimised on the computer and correction factors can be adopted to different flow conditions. This work is complemented by the development of a virtual flow meter for water meters as part of the ongoing EMPIR project METROWAMET, which is to enable the estimation of the influence of water quality on the measurement quality of water meters

This contribution will present these projects, discuss how they are addressing the overarching metrological questions for virtual measuring systems, and outline future developments at PTB in this field.

Mathematical framework for metrology in the factory of the future

S. Eichstädt[1], P. Harris[2], G. Kok[3], B. Ludwig[1], B. X. Yong[4], N. Fischer[5]

[1]Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany
[2]National Physical Laboratory, Teddington, United Kingdom
[3]VSL, Netherlands
[4]Univsersity of Cambridge, United Kingdom
[5]Laboratoire national de métrologie et d'essais, Paris, France
E-mail (corresponding author): sascha.eichstaedt@ptb.de

The EMPIR project "Metrology for the Factory of the Future" (Met4FoF) [1] aims for a holistic implementation of metrological principles to the flow of data in an interconnected manufacturing scenario. The factory of the future is characterised by the use of digital sensors and networks of such sensors providing a flow of data for purposes of automated monitoring, prediction and control. The mathematical framework developed in this project covers various aspects of modelling, data quality and measurement uncertainty applied to sensor networks relevant for the factory of the future. In addition, issues originating from the network itself, such as the treatment of redundant information and network design are addressed. This contribution outlines the tasks in the work package "Metrological infrastructure for aggregation of data from industrial sensor networks".

In general, a mathematical framework is much more than a collection of mathematical methods, publications and items of software. It provides a coherent approach to addressing problems relevant to a specific application or scientific area. For instance, in the EMPIR project 14SIP08 "Dynamic" [2], a mathematical framework for the analysis of dynamic measurements has been developed based on mathematical methods and software developed in the EMRP project IND09 "Traceable Dynamic Measurement of Mechanical Quantities" [3]. This framework consists of the open source software toolbox PyDynamic [4] with introductory examples for various use cases as well as dedicated publications to describe its use in dynamic measurements. In a similar way, Met4FoF aims at a framework for addressing problems relevant to industrial sensor networks. Therefore, each task of the Met4FoF project work package "Metrological infrastructure for aggregation of data from industrial sensor networks" covers a specific type of measurement or measurement issue, see Figure 1. In a particular sensor network, not all aspects covered in these tasks may arise. However, the methods together form a rich toolset for the treatment of measurement uncertainty in sensor networks.

In order to provide a usable implementation of this toolset, one task of the project is to develop a versatile software implementation using a so-called "agent-based framework" (ABF). In such an ABF, each sensor is represented by a software module called a "sensor agent". This module contains all relevant information about the sensor as well as its specific methods and properties. For instance, a calibrated pressure sensor can be represented by a "sensor agent" that contains the complete calibration data as well as methods to calculate a measured value (for pressure) and its associated uncertainty at any instant. This

calculation may also include sophisticated methods, such as filtering or deconvolution. The aim is to make it as simple as possible to apply the mathematical toolset developed in the project.

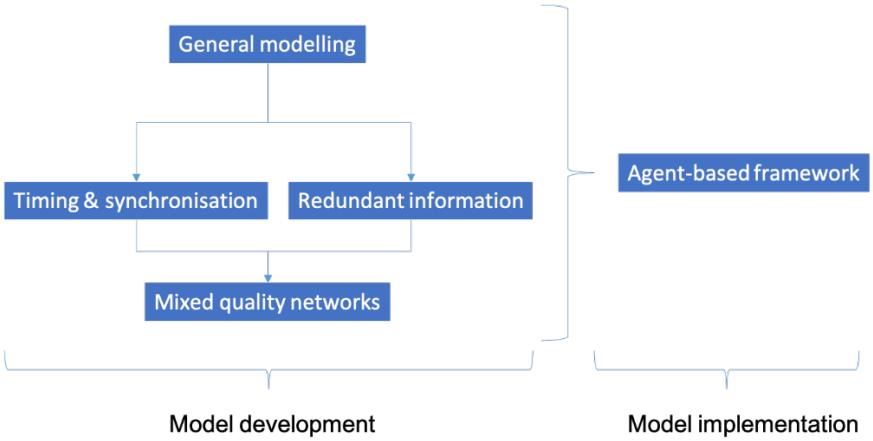


Figure 1 Organisation of the work package „Metrological infrastructure for aggregation of data from industrial sensor networks“ into its various technical tasks.

In this contribution we present the aims and developments from the individual model development tasks and their integration into the ABF for the model implementation. In particular, we focus on the derivation of mathematical models for estimating the quantity of interest (or the process output) and associating an uncertainty with that estimate based on analysing sensor network data. Two measurement scenarios are considered: distributed measurements and aggregated measurements. In this project “distributed” refers to the case in which a number of sensors are deployed within a space (such as a test cell, laboratory or workshop) from which the quantity of interest is estimated as a function of spatial position (and possibly time). The “aggregated” measurement scenario refers to the requirement to assign a single aggregated value of the quantity of interest at a defined point in space and time using a collection of sensors that are providing data that are distributed in space and time.

For the mathematical developments, information about sensor types, calibration data and measurement capabilities are obtained from the work package “Calibration framework”. Realistic measurement data from actual sensor networks is obtained from three testbeds considered in the project [5-7]. For some issues to be addressed by the mathematical framework this data will be taken as the basis for simulations of sensor network data. For the testbed owners the mathematical framework will provide a way to analyse different measurement scenarios as well as to investigate the implementation and inclusion of additional or different sensors. As a starting point for the analysis, previous work by the testbed owners will be used, see, for instance [8].

Mathematical framework for metrology in the factory of the future

Acknowledgement
We gratefully acknowledge funding of this research by the EMPIR 17IND12 Met4FoF project. The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States.

References

[1] EMPIR project 17IND12 "Metrology for the Factory of the Future" (Met4FoF), https://www.ptb.de/empir2018/met4fof/, last visited 30th April 2019
[2] EMPIR project 14SIP08 "Standards and software to maximise end user uptake of NMI calibrations of dynamic force, torque and pressure sensors", http://www.mathmet.org/projects/14SIP08/, last visited 30th April 2019
[3] EMRP project IND09 "Traceable Dynamic Measurement of Mechanical Quantities", https://www.ptb.de/emrp/ind09.html, last visited 30th April 2019
[4] S. Eichstädt, C. Elster, I. M. Smith and T. J. Esward "Evaluation of dynamic measurement uncertainty – an open-source software package to bridge theory and practice" J. Sens. Sens. Syst., 6, 97-105, 2017, DOI: 10.5194/jsss-6-97-2017
[5] STRATH testbed description and dataset, DOI: 10.5281/zenodo.2573860
[6] ZEMA testbed dataset, DOI: 10.5281/zenodo.1326277
[7] ZEMA and SPEA testbed descriptions, https://www.ptb.de/empir2018/met4fof/information-communication/blog/
[8] N. Helwig, T. Schneider, A. Schütze "Automatic Feature Extraction and Selection for Condition Monitoring and related Datasets", IEEE I2MTC 2018

Estimation of the uncertainty in selected points of measured function from two control measurements

Z. L. Warsza[1], J. Puchalski[2]

[1] Industrial Research Institute of Automation and Measurements (PIAP), Warszawa, Poland
E-mail: (corresponding author) zlw1936@gmail.com
[2] Central Metrology Office (GUM), Warszawa Poland, E-mail: jacek.puchalski@gum.gov.pl

This paper discusses two methods for estimating the uncertainty of any selected values of the function which describes a characteristic of tested device, substances or engineering process properties. This estimation is based on performing measurements at two control points. The method I, deterministic one, estimates these uncertainties using a linear approximation based on values of maximum permissible measurement errors at control points and can be used for correlation factor equal to 1. The method II relies on the statistics. Values of points of the tested function, their uncertainties and correlation are estimated as a linear combination of measurement results in two control points. Matrix approach of the uncertainties propagation in indirect multivariate measurements was used. Method I is the boundary case of method II when the correlation coefficient is equal to 1. The absolute and relative uncertainties of interpolated values of characteristic curve and of their linear or nonlinear functions can be properly estimated. It is an extension of the scope of the Supplement 2 to the GUM Guide and can be useful in many metrology applications.

This paper considered is the estimation of values and uncertainties for selected points of the curve modeled with known linear or nonlinear function $y = f(x)$, which are not measured directly. The possibility of estimation based on measurements in two control points were analyzed. The estimated uncertainties depend on the location of tested and control points This issue is not be found yet in the literature.

Basing on uncertainties and correlation coefficients obtained from measurements at control points, it should be performed:

- assessment of the uncertainty and correlation coefficients for points of ested function which are not measured directly, and type A and B uncertainties in these points if they are known at control points,
- determination of confidence bands of this function, such as the permissible maximum error of the digital voltmeter and its linear to voltage dependence, and
- computing correlation coefficients of the estimated values if their values and uncertainties can be processed together.

Method I – deterministic

In method I, the uncertainty u_B (1b) is estimated directly from the linear function (1a) of maximum permissible error of instrument $\Delta_x|_{max}$, i.e.:

$$|\Delta_x| \leq \Delta_x|_{max} = \Delta_{x_0}|_{max} + (x - x_0)\varepsilon_S|_{max}$$
$$\sigma_x \leq \sigma_x|_{max} = \sigma_{x0}|_{max} + (x - x_0)\delta_x|_{max} \quad (1a, b)$$

Where: $\sigma_x|_{max} \equiv u_B = \Delta_x|_{max}/\sqrt{3}$;
$\delta_x|_{max}$ – max. permissible relative error, and $\varepsilon_S|_{max}$ - type B uncertainty of the difference $(x - x_0)$.

Method I is very simple and can be used for estimation of uncertainty u_B in the full range x-x_0. However, it cannot consider the statistical nature of uncertainties of controlled and tested values x_i. Also the correlation coefficients between any two

Estimation of the uncertainty in selected points of measured function from two control measurements

estimated values and uncertainties cannot be determined and are assumed equal 1. For the uncertainty of sum and difference of two values x with the correlation coefficient equal to +1, the component uncertainties should be added algebraically. For the sum, they are larger, and for the difference - smaller than for the geometric summation according GUM.

Method II.

Method II relies on a statistical description of accuracy using uncertainty. Uncertainty evaluation of the elements $\boldsymbol{Y}$ is divided into two stages. In the first, based on measurement results of two controlled values x_1, x_2, a linear scale of values $\boldsymbol{X_c}=\boldsymbol{F_c}(\boldsymbol{X})$ is created for the considered range of x_{max}–x_0 and estimates their uncertainties. Any x_c value is the linear combination described by formula

$$x_c = x_1 + k(x_2 - x_1) \qquad (2)$$

where k means relative locations of the point x_c. In the interval $\langle x_1, x_2 \rangle$ is $0 \le k \le 1$.

The dependence of uncertainty σ_c on k of point x_c is described by the formula:

$$\sigma_c = \sqrt{(1-k)^2\sigma_{x1}^2 + k^2\,\sigma_{x2}^2 + 2k(1-k)\rho_{x1,2}\sigma_{x1}\sigma_{x2}}$$

when no correlation $\rho_{x1,2} = 0$

$$\sigma_c|_{\rho_{x1,2}=0} = \sqrt{(1-k)^2\sigma_{x1}^2 + k^2\,\sigma_{x2}^2}$$

for $\rho_{x1,2} = \pm 1$ the formula becomes linear

$$\sigma_c|_{\rho_{x1,2}=1} = |(1-k)\sigma_{x1} \pm k\,\sigma_{x2}|$$

After normalizing σ_c and σ_{x2} to σ_{x1}, $\varepsilon \equiv \frac{\sigma_{x1}}{\sigma_{x2}}$

$$\sigma_{nc} = \frac{\sigma_c}{\sigma_{x1}} = \sqrt{\varepsilon^2(1-k)^2 + k^2 + 2k(1-k)\varepsilon\,\rho_{x1,2}}$$

Fig 1 shows 3D functions $\sigma_{nc} = f(k,\ \rho_{x1,2})$ for three values of $\varepsilon = \frac{\sigma_{x1}}{\sigma_{x2}} = 4/5; 2/3; 1/2$.

Normalized σ_{nc} uncertainties of value x_c as function of its related location k for $\rho_{x1,2}$ = 1; $\rho_{x1,2}$ = 0 (methods I and II) and ratios of uncertainties $\varepsilon = 1; 2/3$ are on Fig.2.

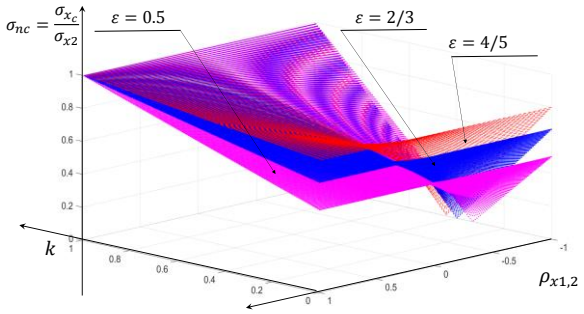


Fig. 1.

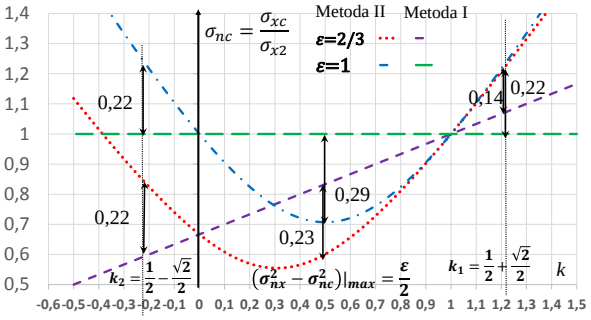


Fig. 2.

Inside the interval $\langle x_1, x_2 \rangle$ the uncertainty reaches the minimum at related location

$$k_{min}(\sigma_{nxc} = min) = \frac{\varepsilon(\varepsilon - \rho_{x1,2})}{1 + \varepsilon^2 - 2\varepsilon\,\rho_{x1,2}}$$

The minimum value of uncertainty is

$$\sigma_{n\,min} = \frac{\varepsilon\sqrt{1-\rho_{x1,2}^2}}{\sqrt{\varepsilon^2 + 1 - 2\rho_{x1,2}\varepsilon}}$$

Uncertainties outside the interval $\langle x_1, x_2 \rangle$ and differences of methods are studied.

In the second stage, from the selected elements x_c of $\boldsymbol{X_c}$, values and uncertainties of the vector $\boldsymbol{Y}$ elements are determined according to the linear or non-linear function $\boldsymbol{Y}=\boldsymbol{F_y}(\boldsymbol{X_c})$. It is based on the vector method for determining uncertainties of multivariate indirect measurements, as is presented in Supplement 2 of GUM Guide.

General conclusion: Method II is based on results of two control measurements with

the required accuracy, is a complementary to the regression methods needs much less work and can be used also on line.

Modelling of flow time series as an approach to compute its uncertainty

M. A. Silva[1,2], C. Amado[1], D. Loureiro[2], A. S. Ribeiro[2]

[1]CEMAT, Instituto Superior Técnico, Av. Rovisco Pais 1,1049-001 Lisbon, Portugal
[2]National Laboratory for Civil Engineering, Avenida do Brasil, 101, 1700-066 Lisbon, Portugal
E-mail (corresponding author): maria.jose.silva@tecnico.ulisboa.pt; mjsilva@lnec.pt

---

The flow measurement is essential for better knowledge about the flow that is abstracted or pumped to drinking water systems, transferred between water utilities or is monitored at the entrance of the subsystems. In the case of water that is transferred and billed between water utilities, flow measurements are a crucial factor for the confidence level between the different agents and the economic sustainability of the service. When flow measurements are used for network monitoring and control of water losses, data quality is fundamental to ensure that results from well-known approaches (e.g., water balance, minimum night flow analysis) are reliable.

The study of the factors that influence the flow measurement uncertainty is a topic insufficiently studied both nationally and internationally. Nevertheless, some factors should be accounted for according to some studies. The elbows in the neighbour of the flowmeter, the orientation of the flow and its nominal diameter are examples of factors influencing the measurement uncertainty. However, most of these studies are done in a laboratory, considering ideal conditions. For a water utility, the goal is to know the uncertainty *in-situ,* considering the existing characteristics of the environment and the installation.

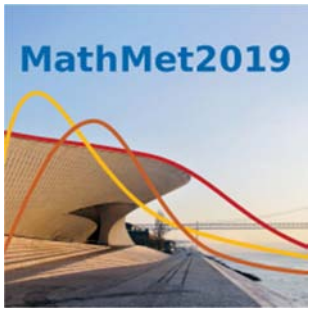
Although there are methods to calibrate the flowmeter *in-situ* through the installation of a different type of flowmeter in parallel, a few water utilities apply this technique. However, the water utilities have the data recorded by the flowmeters. Therefore, the goal of this study is to use the flow data available to measure the uncertainty without installing a different type of flowmeter in parallel or removing the flowmeter from its location to a laboratory calibration.

There are several sources of uncertainty. However, in this study, only the model uncertainty and the parameter uncertainty will be considered. For time series, the problem is that the seasonal effects and the autocorrelation must be taken into account on the uncertainty estimation.

In distribution networks where domestic consumption is predominant, flow time series have weekly and daily seasonality due to water consumption habits. In general, working days have similar patterns among themselves and Saturdays and Sundays have different patterns. Therefore, in this work, three different daily patterns will be considered: working days, Saturdays and Sundays. Considering each pattern individually, it can be modelled considering a proper flow time series model and estimating its parameters. Consequently, the uncertainty of each time series model and parameters can be obtained. Depending on

the type of model used to estimate the pattern, the law of propagation of uncertainty, Monte Carlo simulations or bootstrap can be explored to estimate the uncertainty.

The results allow a better understanding of the drinking water systems by the water utilities. The usefulness of the results is various. Based on them, the water utilities can prioritise the network sectors for operational and management actions, including an indication of which flowmeters should be calibrated. Furthermore, better estimates for some water balance input variables can be obtained, which increases the reliability of the water losses performance indicators calculated based on the water balance. The business relation between water utilities also become more confident in terms of the measurement of the transacted water and, consequently, in terms of the billing.



Managing uncertainties in calculations involving normal tissue complication probability (NTCP)

Maurice Cox[1]
National Physical Laboratory, maurice.cox@npl.co.uk, UK

Absorbed dose in normal tissue is an unwanted consequence of the administration of a particular pharmaceutical to a cancer patient. One of the grand challenges of treatment planning is the quantitative correlation of absorbed dose with clinical outcome. Any difficulties in obtaining such a relationship are exacerbated by the limited data available and the coarseness of such data on the 'outcome' or 'complication' scale. Normal tissue complication probability (NTCP) is a parameter used in radiation therapy for estimating the risk of harmful side effects being induced. An NTCP curve gives the relationship between absorbed dose to normal tissue and complication probability. This study investigates the effects of absorbed dose uncertainty on NTCP as specified by the NTCP curve. Evaluation of NTCP uncertainty has the potential to indicate acceptable levels of uncertainty for absorbed dose for molecular radiotherapy. There are two major sources of uncertainty, namely the data on which the NTCP curve is based and the model used for the NTCP curve. Since there is no widely accepted theoretical basis for the relationship between NTCP and absorbed dose, an empirical function with appropriate properties is chosen. Two-parameter models in common use for the NTCP curve are the Niemierko model (logistic) and the Lyman-KutcherBurman model (normal distribution function). Of the order of 10 % standard uncertainty is often claimed in an estimate of absorbed dose. It is relevant to understand how such an uncertainty propagates to an uncertainty in NTCP. The same standard uncertainty associated with absorbed dose propagates to a much greater standard uncertainty in NTCP in the steeper parts of the curve. However, the left-hand tail, where the gradient is smaller, is important since it is where we want most patients to be, since it corresponds to a relatively small complication probability. Application of GUM principles requires valid linearization of the model about an estimate of absorbed dose, and the assumption of underlying normality in providing a coverage interval. Examination of the results obtained for representative data, in particular infeasible confidence intervals, indicates that the GUM assumptions must be questioned. Application of the Monte Carlo guide in the GUM suite of documents, assuming probability distributions for the input quantities, gives results that are feasible and differ considerably from those of the GUM in terms of the estimate of complication probability, the associated standard uncertainty and coverage. Realistic examples are given to illustrate the above points and conclusions are drawn in a clinical context.

Measurement uncertainty training: experiences at the INRIM and ideas for developing a dedicated international community

F. Pennecchi[1]
[1]*Istituto Nazionale di Ricerca Metrologica - INRIM, Strada delle Cacce 91, Torino, Italy*
E-mail (corresponding author): f.pennecchi@inrim.it

As a National Metrology Institute (NMI), INRIM aims at the *knowledge transfer* as one of its main missions. Within such framework, training on measurement uncertainty (MU) plays a fundamental role. INRIM holds the chairmanship of the BIPM JCGM-WG1 (Working Group on the Expression of Uncertainty in Measurement) and it is represented into several scientific and standardization committees dedicated to MU evaluation, such as the ISO/TC 69/SC 6/WG 7 on Statistical methods to support measurement uncertainty evaluation, the Eurachem/CITAC Measurement Uncertainty and Traceability Working Group and the ENBIS Special Interest Group on Measurement Uncertainty, just to mention some. The Italian NMI has participated and it is still involved into several MU-related European Metrology Projects and Networks, such as the EMRP NEW04 Project on Novel mathematical and statistical approaches to uncertainty evaluation, the EMPIR EMUE Project on Examples of Measurement Uncertainty Evaluation and the EURAMET EMN MATHMET. Given the above-mentioned commitment to MU evaluation, INRIM has always put effort into fostering MU training for different kinds of audience and applications and at a different degree of detail.

At the INRIM, general courses have been organized in the past for the internal staff, providing basic knowledge about the GUM and related Supplements. Many courses are regularly or occasionally provided to industries and test and calibration laboratories in their field of interest. Typically, these courses concern the specific measurement field or are conceived as a general course on metrology: in either case, a certain time span of the lessons is invariably dedicated to MU evaluation. Within such courses for industries and laboratories, the Excel software is the typical tool used to show how to perform calculations concerning the law of propagation of uncertainty (LPU) and corresponding uncertainty budget, the Monte Carlo Method (MCM) for propagation of probability distributions and the least-squares regression. The metrological areas covered by this kind of courses span from thermodynamic to electrical, from mass-related to meteorological measurements and so on.

There are also a number of institutional courses or seminars which have been given to Associations, Bodies and Schools (such as the International School of Physics "Enrico Fermi", the PhD School "Italo Gorini" and the INMETRO Advanced School on Evaluation of Uncertainty in Measurement), concerning research, quality, accreditation, legal metrology and education, as well as to other research institutes and non-European NMIs asking for a support in the MU field. A very important effort is also provided by INRIM personnel to the training at the academic level, both for undergraduate students (giving courses at the University of Pavia and Torino, for example) and for PhD students. In particular, it is worthwhile mentioning that INRIM coordinates, with the Politecnico di Torino, a doctorate course on Metrology, in which a specific course on MU evaluation was regularly hold in the

past, whereas at the moment it is scheduled as the “Uncertainty of measurements” module within the “Fundaments of metrology” course. Courses held at the university level typically allow to treat uncertainty topics in more details, also presenting to the students the available softwares and tools dedicated to the MU evaluation (see, e.g., https://en.wikipedia.org/wiki/List_of_uncertainty_propagation_software) and/or showing them how to develop their own software facility.

Despite all the above-mentioned experience in the MU training framework, those activities were rarely overall coordinated. An exchange of views and materials, not only at a national institute level but also at an international level, could certainly serve as a base for creating a platform of experts/teachers updated and active on MU topics. The choice of didactical materials, tools and softwares and the degree of in-depth analysis need to be specific to the target audience (academia, industry, laboratory, …) and to the aim of the course (general knowledge, particular application, adherence to normative standards, …). These thoughts and other considerations will be given in the presentation.

I would like to warmly thank all the INRIM colleagues who kindly shared with me their teaching experience and materials.

Risk based assessment of the degree of severity of myocardial perfusion and the determination of an optimal decision rule

K. Jagan[1], P. M. Harris[1], N. Smith[1], A. Fenwick[1], Jarmo Teuho[2], Jussi Schultz[2], Reetta Siekkinen[2], Juhani Knuuti[2], Antti Saraste[2,3] and Mika Teräs[4,5]

[1]National Physical Laboratory, Hampton Road, Teddington, United Kingdom
[2]Turku PET Centre, Turku University Hospital and Turku University (TUCH), Finland
[3]Heart Centre, Turku University Hospital, Finland
[4]Department of Biomedicine, University of Turku, Turku, Finland
[5]Department of Medical Physics, Division of Medical Imaging, Turku University Hospital, Turku, Finland
E-mail (corresponding author): kavya.jagan@npl.co.uk

Cardiovascular disease (CVD) is a leading cause of death in Europe and costs the European economy approximately €210 billion each year [1]. Over the last two decades, several clinical landmark studies have shown that accurate measurement of heart muscle blood supply (perfusion) could serve as a gatekeeper for treating the right patients [2]. Perfusion is essential to the functioning of the heart and is an early marker of the so-called ischemic cascade that leads to non-reversible tissue damage and thus chronic heart disease. Accurate quantification of perfusion is currently only possible through invasive measurements and as an alternative, different medical imaging techniques (modalities) have been developed to measure perfusion non-invasively, namely Positron Emission Tomography (PET), Computed Tomography (CT), Single Photon Emission Computed Tomography (SPECT) and Magnetic Resonance Imaging (MRI).

The European Metrology Programme for Innovation and Research (EMPIR) project 15HLT05 "*Metrology for multi-modality imaging of impaired tissue perfusion"* aims to address metrology issues for the health sector by developing a physical phantom for quantitative imaging of blood perfusion of heart muscles that can be imaged in a range of modalities, and by developing new data analysis methods that will make blood perfusion imaging traceable and quantitative [3]. As part of this project, a classification framework of patients as ischemic/non-ischemic has been developed which reflects the uncertainty in the classification process and in the myocardial perfusion (blood flow) values.

This paper describes a risk based [4] decision making framework on the degree of severity of myocardial perfusion in patients based on a clinically predefined decision rule used by clinicians at the Turku PET centre. A method based on Receiver Operating Characteristic (ROC) curves [5] to determine the optimised decision rule for PET studies is also presented.

The methods are illustrated on clinical data from patients participating in a cardiac PET perfusion study at the Turku PET centre. The PET data is processed to obtain perfusion maps [6, and references within] which are the inputs to the classification methods developed.

Risk based assessment of the degree of severity of myocardial perfusion and the determination of an optimal decision rule

By combining patient data and expert insights, this classification tool could (i) help less experienced clinicians make better decisions regarding patient health, (ii) serve as a starting point for further investigation, (iii) be used as a screening to categorize patients in order of severity, so that the most severe cases can be prioritized on the clinical list.

References

[1] Wilkins E et al. *European Cardiovascular Disease Statistics 2017*. European Heart Network, Brussels.

[2] Shaw L et al. Optimal medical therapy with or without percutaneous coronary intervention to reduce ischemic burden: results from the Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation (COURAGE) trial nuclear substudy. *Circulation* 2008, 117(10), 1283–1291.

[3] https://www.ptb.de/empir2018/perfusimaging/home/. Last accessed: 27/04/2019.

[4] BIPM, IEC, ILAC IFCC, IUPAC ISO, and OIML IUPAP. *Evaluation of measurement data—the role of measurement uncertainty in conformity assessment. JCGM (2012);* 106:1-06.

[5] Fawcett T. An introduction to ROC analysis. *Pattern recognition letters*. 2006 Jun 1; 27(8):861-74.

[6] http://www.turkupetcentre.net/petanalysis/quantification.html. Last accessed: 27/04/2019.

Bayesian inversion for CD determination with uncertainties.

S. Heidenreich[1], N. Farchmin[1], H. Gross[1] and M. Bär[1]

[1]Physikalisch-Technische Bundesanstalt – Braunschweig and Berlin, Berlin, Germany
E-mail: sebastian.heidenreich@ptb.de

Optical scattering techniques are fast indirect and non-destructive methods for the determination of profile parameters at the nanoscale. Profile parameters are obtained from diffracted light intensities by solving an inverse problem. To comply with the upcoming need for improved accuracy and precision and thus for the reduction of uncertainties, a finite element method based computationally expansive forward models have to be used. In the talk, we present a Bayesian inversion to reconstruct six geometry parameters from light diffracted intensities. We determine the related posterior distribution by a Markov Chain Monte Carlo Metropolis random walk method. To cope with the computational effort, we approximate the forward model by a polynomial chaos based surrogate model. Finally, we discuss the strengths and limitations of the approach presented.

S. Heidenreich, H. Gross and M. Bär, Bayesian approach to determine critical dimensions from scatterometric measurements. Metrologia, 55(6), 201, 2018.

Global sensitivity analysis using polynomial chaos.

N. Farchmin[1], M. Bär[1], S. Heidenreich[1]

[1]Physikalisch-Technische Bundesanstalt – Braunschweig and Berlin, Berlin, Germany
E-mail: nando.farchmin@ptb.de

Optical scattering techniques are fast non-destructive methods for the determination of critical dimensions (CDs) at the nanoscale. CDs are obtained as profile parameters from diffracted light intensities by solving an inverse problem. In many applications, a large number of such parameters have to be estimated which leads to enormous computational costs. In order to reduce the number of parameters as well as for experimental design a sensitivity analysis is quite useful. In this talk, we present a general and efficient method for a global sensitivity analysis based on a polynomial expansion of stochastic variables of interest. In particular, we first demonstrate the functionality of the method by applying it to the Sobol function, where the sensitivity parameters are analytically known. Secondly, we apply the method to a line grid of a photomask. Here the global sensitivity of diffracted light is determined by changes in the line geometry. Light diffraction is simulated by solving the Helmholtz equation for electromagnetic waves by a FEM solver. Finally, we discuss the accuracy and efficiency of the method.

A model for complex shape and motion pattern analysis in medical images

N. Debroux[1], J. Aston[2], F. Bonardi[3], A. Forbes[4], C. Le Guyader[5], M. Romanchikova[4], C. Schönlieb[6]

[1]Department of Applied Mathematics and Theoretical Physics,University of Cambridge and National Physical Laboratory, Cambridge, U.K.
[2]Statslab, University of Cambridge, Cambridge, U.K.
[3] IBISC, Université d'Évry, Évry, France
[4]National Physical Laboratory, Teddington, U.K.
[5]Laboratoire de Mathématiques, Normandie Univ, Institut National des Sciences Appliquées de Rouen, Rouen, France.
[6]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, U.K.
E-mail (corresponding author): nd448@cam.ac.uk

Context : In medical image analysis, constructing a statistically meaningful representative of a set of images of the same object, is a critical task for practitioners to estimate variability inside a population, and to characterise and understand how structural changes have an impact on health. This involves identifying the main shapes in each image of the set, a process called segmentation resulting in a structural image representing a complex shape, and mapping this group of images to an unknown mean structural image, a task called registration for atlas generation. Using the so-created mean complex shape and together with the mappings to the individual images, further statistical analysis of the image dataset is possible.

Motivation : We present a joint segmentation, registration and atlas generation model that reduces error propagation and takes advantage of mutual beneficial influence of the registration and the segmentation tasks. We then perform a Principal Component Analysis on the registration mappings to get the main modes of variations in terms of geometric transformations and identify motion patterns for different anatomical sites and different patient populations.

Methodology : From a set of initial images $\left(T_i\right)_{i=1}^{M}$, we seek to simultaneously extract a segmentation of each image $\left(\theta_{T_i}\right)_{i=1}^{M}$, to generate a structural and statistical representative of this set called θ_R , as well as to estimate the deformation maps $\left(\varphi_i\right)_{i=1}^{M}$ aligning each segmentation to this structural mean. In a nonlinear elasticity setting, we propose to view the shapes to be matched as homogeneous, isotropic and hyperelastic materials and more precisely as Ogden materials. This nonlinear-elasticity-based regulariser prescribing the physical nature of the deformation is then complemented by novel hard constraints on the L^{∞} norm of the Jacobian and its inverse to ensure topology preservation. Segmentation is based on the Potts model which allows for a partition into more than two regions. The shapes are assumed to have homogeneous intensity values and the resulted segmentations are piecewise constant images representing complex shapes. The connection between these two tasks is ensured by a dissimilarity measure that aims at

A model for complex shape and motion pattern analysis in medical images

aligning the corresponding regions of interest. The overall functional to be minimised is given hereafter (Equation 1). We obtain an approximate solution by an alternating optimisation scheme.

$$\inf F\left(\theta_R,\left(\theta_{T_i},\varphi_i\right)_{i=1}^{M}\right)=\frac{1}{M}\underbrace{\sum_{i=1}^{M}\frac{\gamma_T}{2}\sum_{l=1}^{N}\mathrm{TV}\left(\theta_{T_i,l}\right)+\sum_{l=1}^{N}\int_{\Omega}\theta_{T_i,l}\left(c_{T_i,l}-T_i\right)^2 dx}_{\text{Template segmentation}}+\underbrace{\frac{\lambda}{2}\sum_{l=1}^{N}\mathrm{TV}\left(\theta_{T_i,l}\circ\varphi_i-\theta_{R,l}\right)}_{\text{homogeneous region pairing}}$$

$$+\underbrace{\frac{\gamma_R}{2}\sum_{l=1}^{N}\mathrm{TV}\left(\theta_{R,l}\right)+\int_{\Omega}\sum_{l=1}^{N}\theta_{R,l}\left(c_{R,l}-T_i\circ\varphi_i\right)^2 dx}_{\text{structural nature of the atlas}}+\underbrace{\int_{\Omega}W_O\left(\nabla\varphi_i\right)dx+1_{\{\|.\|_{L^\infty(\Omega)}\leq\alpha\}}\left(\nabla\varphi_i\right)+1_{\{\|.\|_{L^\infty(\Omega)}\leq\beta\}}\left(\left(\nabla\varphi_i\right)^{-1}\right)}_{\text{regularisation of the deformations}},$$

$$W_0(x)=\begin{cases}a_1\|x\|^4+a_2(\det x-1)^2+\dfrac{a_3}{(\det x)^{10}} & \text{if } \det x>0\\ +\infty & \text{otherwise}\end{cases}.$$

Equation 1 : Proposed optimisation model.

Then, in a D^m spline setting, we propose to approximate the deformation maps in a linear space and perform a Principal Component Analysis to determine the main modes of variations in terms of geometric distortions.

Results : We test our method implemented in C on a Intel core i7 with 2.60GHz and 8GB RAM, and on a CINE MRI sequence representing a cardiac cycle. We show in the following image (Illustration 1) the obtained two first modes of variation that are consistent with the anatomical dynamic of the heart : while the first mode encodes the dilation/contraction of the right ventricular chamber in the transverse direction, the second mode conveys the vertical stretching of the left ventricular chamber. The modes of variation are visualised by showing $\theta_R\circ(Id+c(v_1,v_2))$ with c varying from -5 to 5 and where (v_1,v_2) are the resulting displacements from the PCA.

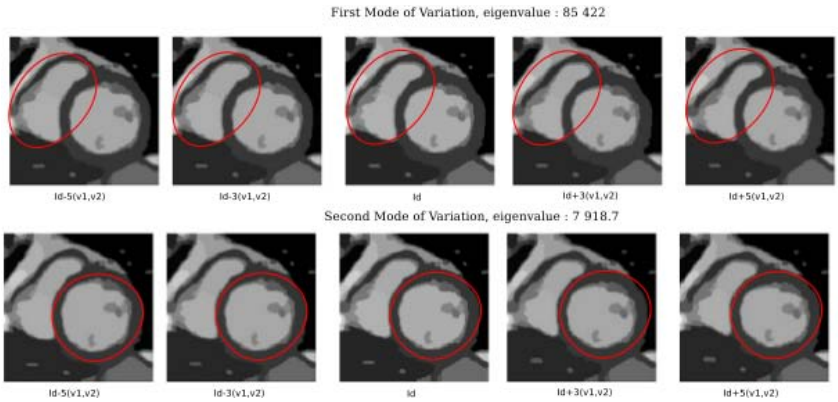


Illustration 1 : Modes of variation for a cardiac CINE MRI sequence. The first row represents the first mode of variation while the second row corresponds to the second mode with red circles pointing out the moving region.
Computation time : first step 49min, second step 3min 27sec.

Conclusion : Our joint approach exploits the synergy between segmentation and registration to generate an anatomically consistent complex shape with sharp edges and physically meaningful deformations maps. The PCA on the approximated deformations enables to identify motion patterns for different regions of the heart and show promising potential for clinical practice to classify motion patterns for different anatomical sites and different patient populations.

On the influence of inlet perturbations on the development of slugs in horizontal two-phase flow

S. Schmelter[1], M. Olbrich[1,2], E. Schmeyer[1], M. Bär[1]

[1]Physikalisch-Technische Bundesanstalt (PTB), Abbestr. 2-12, 10587 Berlin, Germany
[2]Institute of Fluid Dynamics and Technical Acoustics, Technische Universität Berlin, Müller-Breslau-Str. 8, 10623 Berlin, Germany
E-mail (corresponding author): sonja.schmelter@ptb.de

Horizontal two-phase flow is of great importance in many industrial applications, for example in the nuclear, chemical, and petroleum industries. When two-phase liquid and gas flow occurs inside a horizontal pipe, the phases can be distributed in diverse ways resulting in different flow patterns [1]. Slug flow is one of the most commonly observed patterns in two-phase flow and it is of special interest due to its intermittent behavior. It can cause vibrations in the pipe system, which may lead to damage [2]. Therefore, it is necessary to understand the development of slugs.

In this contribution, we investigate the influence of different inlet perturbations on the development of slugs further downstream in a horizontal pipe by computational fluid dynamics (CFD). Two different approaches are considered. In the first approach, a sinusoidal agitation of the interface between the two phases is prescribed on the inlet as was first proposed in [2]. In the second approach, the radial and angular components of the inlet velocity vector field are randomly perturbed. The numerical simulations show that both approaches enhance the development of slugs in the pipe. However, since the frequency of the sinusoidal agitation is reflected in the frequency of slug occurrence, this approach is only feasible if the slug frequency is known in advance, for example from corresponding experiments.

Fig. 1 shows simulation results for an air-water slug flow test case, where the interface between the phases was agitated with a frequency of 1 Hz. The pictures show the gas volume fraction in a longitudinal section through the middle of the pipe at different time points. In the top picture, one can see the sinusoidal structure of the interface between the phases, which was prescribed as initial condition. In the first seconds of the simulation, this sinusoidal structure decreases in amplitude until it is almost levelled out after ca. 2 seconds. After about 2.4 seconds, the first slug occurs. This slug grows continuously while moving through the pipe. For this and other test cases, relevant parameters (like slug length or velocity) are compared with experimental data for validation. Furthermore, the influence of different agitation frequencies and perturbation amplitudes on the development of slugs are studied.

[1] H. Pineda-Pérez, T. Kim, E. Pereya, and R. Ratkovich, CFD modelling of air and highly viscous liquid two-phase slug flow in horizontal pipes, *Chem. Eng. Res. Des.* **136**:638-653, 2018.

On the influence of inlet perturbations on the development of slugs in horizontal two-phase flow

[2] T. Frank, Numerical simulation of slug flow regime for an air-water two-phase flow in horizontal pipes, in: *The 11th International Topical Meeting on Nuclear Reactor Thermal-Hydraulics (NURETH-11)*, Avignon, France, 2005.

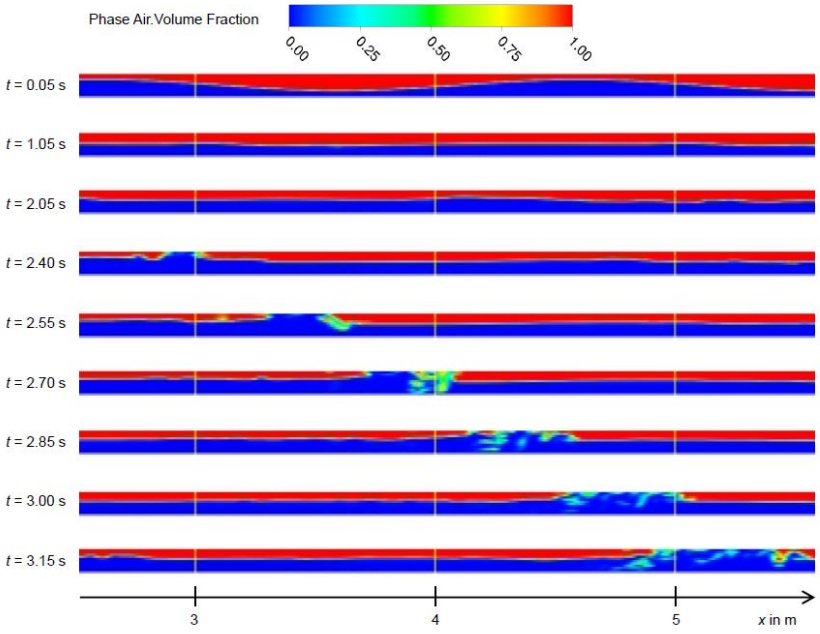


Figure 1: Development of slug flow for an air-water test case with sinusoidal perturbation of the prescribed liquid level at the inlet.

Calculating coverage factors and coverage probabilities in study cases.

A. Chunovkina[1], A.Stepanov[1]
[1] *The D.I.Mendeleyev Institute for Metrology, Moskivsky pr.19, St.Petersburg, Russia*
E-mail : A.G.Chunovkina@vniim.ru

The widespread use of GUM is due to the simplicity of the estimates given in it and their correctness in the case of linear measurement models, which are quite common in practice. However, a significant drawback of GUM is its internal inconsistency. Therefore, when developing guidelines for calculating uncertainty in routine tasks, it seems important to keep the simplicity of GUM and overcome its inconsistency by setting out the calculation of uncertainty on the basis of the Bayesian approach. This can be achieved by considering study cases of calculating measurement uncertainty, taking into account the fact that differences in uncertainty estimates of the order of 5% in routine tasks can usually be considered insignificant.
In this article the detailed consideration is given to study cases that are characterized by the following factors:

1. There is a linear measurement model $X = Y - B$, where X is the measurand, Y is the indication of the measuring instrument, B is the bias of the measuring instrument indication. Note, that a large number of linear measurement models are reduced to this model, when B is treated as the sum of biases caused by various systematic effects.
2. There are repeated observations of the measuring instrument $y_1, \dots, y_n, n \geq 4$, , the standard uncertainty, u_A, corresponding to Y, is calculated according to type A. The normal and uniform pdf for the repeated observations are considered in the paper . Accordingly, for the normal pdf the following expression is valid:

$$u_A^2 = \frac{n-1}{n-3}\frac{S^2}{n} \quad , \text{ where } S \text{ is the sample variance} \qquad (1).$$

and for a uniform pdf:

$$u_A^2 = \frac{r^2}{2(n-2)(n-3)} \text{ , where } r \text{ is the range of the series } y_1, \dots, y_n, n \geq 4 \qquad (2).$$

3. The standard uncertainty u_B, associated with the bias B is calculated according to type B. If B is the sum of the corrections for various systematic effects, u_B, is calculated as the square root of the sum of the squares of the uncertainties corresponding to each systematic effect. For modeling the B, a family of exponential probability distributions was chosen, which includes, as particular cases, the normal and uniform (as a limit case) distributions [1,2] :

$$p(b) = \frac{\alpha}{2\lambda u_B \Gamma(1/\alpha)} exp\left\{-\left|\frac{b}{\lambda u_B}\right|^{\alpha}\right\}, \lambda = \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}. \qquad (3)$$

Calculating coverage factors and coverage probabilities in study cases.

In the article, using the Bayesian approach, the probability distribution for the measurand *X* is obtained and the coverage factors (for probabilities 0.9 and 0.95) are calculated for different values of α and *n*. The coverage factors are obtained depending on the parameter $= \frac{u_B}{S}\sqrt{n}$ (in the case of a normal distribution of repeated observations) and $= \frac{u_B}{r}\sqrt{n}$ (in the case of a uniform distribution of repeated observations).
An analysis of the calculation results shows that for the cases of the considered coverage factors, *k,* for the probability of 0.95 vary within the interval 1.65-2.15, which is more than 25%. At the coverage probability of 0.99, these differences will reach 60% already. However, for the coverage probability of 0.90, the range of variation of the coverage factors is significantly less, 1.56 - 1.65, which does not exceed 7%.
Thus, if the accuracy of estimating a coverage factor less than 5% is considered acceptable, then for a coverage probability of 0.90, it is possible to recommend a single coverage factor for the considered pdf equal to 1.6. Note, that the result obtained is valid if formulae (1) - (2) are used for the standard uncertainty calculated by type A.

At $\rightarrow 0$, the following asymptotes take place $k_p = \sqrt{\frac{n-1}{n-3}}t_p(n-1)$: where $t_p(n-1)$ is the quantile of the Sudent distribution (in the case of the normal pdf of repeated observations) and $k_p = \sqrt{\frac{(n-2)(n-3)}{2}} \times \sqrt[n-1]{\frac{1}{1-p}}$ (in the case of a uniform pdf of repeated observations).

In practice, for example, in calibration certificates a coverage factor equal to 2 is often used, specifying that it corresponds to a coverage probability of about 0.95. In this paper, within the considered models, the coverage probabilities are obtained for the coverage factors of 1.6, 1.8, 2.0, 2.2, and 2.4. It is shown that for coverage factor 2, the corresponding coverage probabilities vary from 0.94 to 0.99 (in the case of a normal pdf of repeated observations) and from 0.92 to 1.0 (in the case of a uniform pdf of repeated observations). Note that for the coverage factor of 1.6, the probabilities of coverage vary from 0.89 to 0.94, regardless of the pdf of repeated observations.

[1] Stepanov, A. V. , Chunovkina, A. G. and Burmistrova, N. A. Calculation of coverage intervals: some study cases. Advanced mathematical and computational tools in metrology and testing X Advances in mathematics for applied sciences (Volume 86), World Scientific Publishing Co, Singapore, 2015, 429 pp, ISBN: 978-981-4678-61-2
[2] Chunovkina, A., Stepanov, A. Calculation of coverage intervals for repeated measurements (Bayesian inference), Journal of Physics: Conference Series 1065(21), 212009, 2018

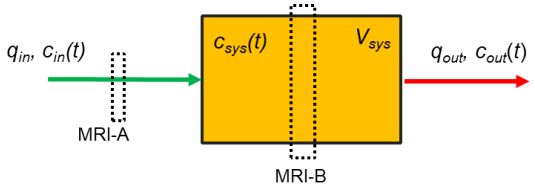
Uncertainty related to flow modelling errors in medical perfusion imaging

G. Kok[1] and X. Milidonis[2]

[1]VSL B.V., Thijsseweg 11, 2629 JA, Delft, Netherlands,
[2]School of Biomedical Engineering and Imaging Sciences, King's College London, London, United Kingdom
E-mail (corresponding author): gkok@vsl.nl

In the medical sector, various imaging methodologies or *modalities* (e.g. MRI, PET, CT) are used to assess the health of various parts of the bodies of patients. One such investigation is the normalized blood flow or *perfusion* of the heart muscle. The perfusion rate (of human tissue) is defined as the (blood) flow rate normalized by the mass of the volume of interest, and usually expressed in mL/min/g. A decreased perfusion of the heart muscle or *myocardium* is an imaging biomarker for increased risk for coronary artery disease and can cause chest pain. However, there is no physical flow standard for the assessment and validation of myocardial perfusion imaging methodologies, resulting in a large proportion of medical diagnoses being inaccurate and highly dependent on the scanner type, software used and the clinical operator. In the EMPIR 15HLT05 PerfusImaging project [1], basing on an earlier work [2], a physical standard ("phantom" in medical imaging terminology) simulating myocardial perfusion has recently been developed [3]. Part of this project focuses on the data analysis of the measurement images, and the assessment of the uncertainty of quantitative perfusion imaging. In this paper one particular source of uncertainty is analysed. Note that in daily practice, perfusion values are rather assessed on a qualitative scale than on a quantitative scale, and that a shift to quantitative assessment of perfusion values requires amongst others an analysis of their uncertainty.

A standard approach being used in perfusion imaging is based on the model as shown in Figure 1 for which the convolution problem of Equation (1) plays a central role (notation is explained in the caption of Figure 1).



$$c_{\mathrm{sys}}(t) = (c_{\mathrm{in}} * R_f)(t). \qquad (1)$$

Figure 1: Standard system view used in perfusion quantification with MRI. The system volume V_{sys} is fed by a constant flow rate q_{in} with a time dependent concentration of contrast agent (CA) $c_{\mathrm{in}}(t)$, the same flow rate leaves the system (q_{out}) with a CA concentration $c_{\mathrm{out}}(t)$, and the average system CA concentration is $c_{\mathrm{sys}}(t)$. Using MRI the time dependent CA concentration is measured at the positions MRI-A and MRI-B.

It is assumed that the functions $c_{\mathrm{sys}}(t)$ and $c_{\mathrm{in}}(t)$ can be measured with a sufficient quality. However, in many situations it is the CA concentration $c_{\mathrm{out}}(t)$ at the exit of a compartment that is being measured, rather than the average concentration $c_{\mathrm{sys}}(t)$ of the complete compartment. Alternative equations involving additional assumptions can be formulated to solve this measurement problem. These will be presented as well as the results of the

Uncertainty related to flow modelling errors in medical perfusion imaging

different methods involving both simulated data (using different models) and measurement data.

Another issue with the standard method is that as a base assumption it treats various image pixels or image regions as completely independent. Sometimes some purely statistical approaches are applied to smoothen the image data or the calculated perfusion image, or to reduce the effect of measurement noise when solving the convolution problem, but this does not take into account existing physical relationships between pixels or image regions in terms of a mathematical model of the flow physics. In this presentation it will be shown how physical interaction between phantom parts can be modelled, and how the various analysis approaches perform, again involving both simulated data (using different models) and measurement data. In Figure 2 the main ideas of this alternative system view are shown.

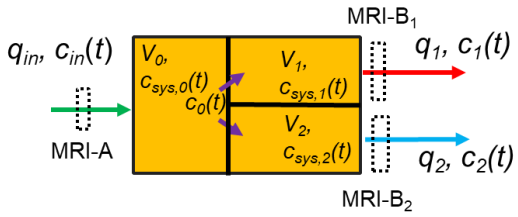


Figure 2: Alternative system view for perfusion quantification with MRI. The compartment volumes are denoted by V_i, compartment average concentrations as function of time are denoted by $c_{\mathrm{sys},i}(t)$, compartment outflow concentrations are denoted by $c_i(t)$ and outflow flow rates by q_i for $i = 0, 1, 2$.

The advantage of the proposed method is that it may give more accurate results when all model parameters are known, and that it can give an idea about (one part of) the model uncertainty of the standard approach used in perfusion imaging. Its main drawback is that various additional parameters and assumptions are needed, which are difficult to obtain and verify for measurements involving real patients.

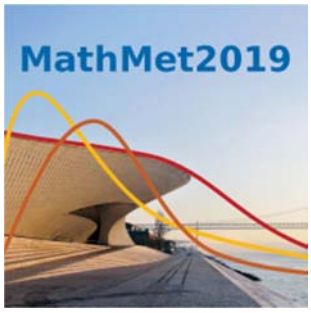
We gratefully acknowledge funding of this research by the EMPIR 15HLT05 PerfusImaging project. The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States.

References

[1] EMPIR project 15HLT05 PerfusImaging, *Metrology for multi-modality imaging of impaired tissue perfusion*, https://www.ptb.de/empir2018/perfusimaging/home/

[2] Chiribiri et al., "Perfusion Phantom: An Efficient and Reproducible Method to Simulate Myocardial First-Pass Perfusion Measurements With Cardiovascular Magnetic Resonance", *Magnetic Resonance in Medicine*, 2012, DOI 10.1002/mrm.24299

[3] Milidonis et al., *Dynamic contrast-enhanced and phase contrast MRI of a novel 3D printed cardiac phantom mimicking transmural myocardial perfusion gradients*, ISMRM conference, Paris 2018.



Model-based determination of optical and geometrical properties of red blood cells from light scattering

Jonas Gienger, Hermann Gross, Nando Farchmin, Sebastian Heidenreich, Jörg Neukammer, and Markus Bär

Physikalisch-Technische Bundesanstalt (PTB), Abbestr. 2 – 12, 10587 Berlin, Germany
e-mail: markus.baer@ptb.de

Light scattering off single cells is widely applied for flow cytometric differentiation of cells. The knowledge of optical properties of biological cells is essential to interpret their interaction with light and to derive morphological information and parameters associated with cell function like the oxygen transport capacity of human red blood cells (RBCs). First, we present a method to determine the dependence between the refractive index (RI) of human RBCs and their intracellular hemoglobin (Hb) concentration from spectral extinction measurements of a cell suspension. The procedure is based on the analysis of the corresponding ensemble averaged extinction cross section. Our approach yields the RI increment with Hb concentration of intact, reversibly sphered, oxygenated RBCs over a wide wavelength range from 290 nm to 1100 nm from macroscopic measurements [1]. Second, we analyze data from a flow cytometer built to simultaneously observe forward light scatter of RBCs for orthogonal laser beams. Oftentimes, bimodal distributions are observed in the distributions of the forward scattering cross section (FSC) of RBCs in their native state. Then, simulations of the light scattering by single RBCs were performed using the discrete dipole approximation for a range of sizes, orientations and Hb concentrations to obtain FSC distributions from RBC ensembles. If an elongated shape model is employed that accounts for the stretching of the cell by hydrodynamic forces, the features of the strongly bimodal measured frequency distributions are qualitatively reproduced by the simulation. Elongation ratios significantly greater than 1 in the range of 1.5 to 2.5 yield the best agreement between experiments and simulated data, indicating a deformation of the RBCs by hydrodynamic gradients. This observation provides hope for a quantitative determination of the RBC geometry or elastic properties from light scattering measurements. Preliminary results on the solution of this inverse problem of optical flow cytometry will be presented.

[1] J. Gienger, K. Smuda, R. Müller, M. Bär and J. Neukammer. Refractive index of human red blood cells between 290 nm and 1100 nm determined by optical extinction measurements. Scientific Reports 9, 4623 (2019).

Implementation of GUM principles in Western Balkan countries

A. Bosnjakovic[1], V. Karahodzic[2], E. Jugo[3], M. Causevic[4], Z. Dzemic[5]

[1,2,3,4] Institute of metrology of Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina
E-mail: alen.bosnjakovic@met.gov.ba, vedran.karahodzic@met.gov.ba, ehlimana.jugo@met.gov.ba, merima.causevic@met.gov.ba, zijad.dzemic@met.gov.ba

---

In year 2018 European Association of Metrology Institutes (EURAMET) started with the activities for establishment of European Metrological Networks (EMNs), including EMN: Mathematics and Statistics - MATHMET. Institute of Metrology of Bosnia and Herzegovina (IMBIH) is one of the MATHMET's active members, focused on achieving network's objectives. Measurement uncertainty by itself is a complex area which requires broad spectrum of knowledge and skills for its understanding and implementation. When it comes to improvement of implementation of measurement uncertainty it is necessary to understand knowledge and need of different sectors that deal with measurement uncertainty. For that purpose Institute of Metrology of Bosnia and Herzegovina has started a survey with aim of discovering different approaches for providing training courses on implementation of Guide to the expression of uncertainty in measurement (GUM) principles among Western Balkan National Metrology Institutes and their National Accreditation Bodies. The overall goal of this survey is to obtain information on capacities for the implementation of GUM and its Supplements, as well as on the knowledge transfer to stakeholders in these countries.

After visiting several National Metrology Institutes during the past years, it was concluded that not all of them use same methodologies when it comes to applying of GUM. Some of the NMIs also organize training courses related to implementation of measurement uncertainty in their laboratories and these trainings differ from country to country. This fact led to an idea for making a survey to discover how the GUM is implemented and how the training courses on GUM implementation are developed in Western Balkan countries including Bosnia and Herzegovina.

The survey has been already started and it will consist of two phases. First phase includes sending the questionnaire to several NMIs and Accreditation Bodies, which will give the information on the GUM implementation and training courses related to measurement uncertainty. The second phase will imply analysis of the research results and presenting the findings.

The results of this research will provide overall framework on what are the needs of National Metrology Institutes and their National Accreditation Bodies when it comes to GUM implementation training courses. Besides, the results will create a basis for measurement uncertainty trainings at NMIs - from large NMIs with long-standing, well-developed courses via smaller NMIs/DIs with an occasional or specific short course to NMIs who wish to offer courses. The results of these findings will be presented at MATHMET2019 conference.

Data fusion/calibration methods to update simulated data based on the observed data: an application to wind speed data.

Amaral Turkman, M. A.[1], Pereira, S.[1], Pereira, P.[2], Turkman, K. F.[1], de Zea Bermudez, P.[1], and de Carvalho, M.[3]

[1]Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, Lisboa, Portugal
[2]Centro de Estatística e Aplicações, EST Setúbal, Instituto Politécnico de Setúbal, Setúbal, Portugal
[3] School of Mathematics, University of Edinburgh, United Kingdom and Centro de Estatística e Aplicações, Universidade de Lisboa, Portugal
E-mail (corresponding author): maturkman@fc.ul.pt

Abstract: Extreme values of certain spatio-temporal processes, such as wind speeds, are the main cause of severe damage in property, from electricity distribution grid to road and agricultural infrastructures. Accurate assessment of causal relationships between environmental processes and their effects on risk indicators, are highly important in risk analysis, which in return depends on sound inferential methods as well as on good quality informative data. Often, information on the relevant environmental processes comes from monitoring networks, as well as from numerical-physical models (simulators) that typically solve a large set of partial differential equations, capturing the essence of the physical process under study. In general, monitoring networks are formed by a sparse set of stations, whose instrumentation are vulnerable to disruptions, resulting in data sets with many missing observations, whereas, simulated data from numerical simulators typically supply average yield of the process in grid cells of pre-specified dimensions, often at high resolutions, spanning large spatial domains, with no missing observations. However, simulated data typically mismatch and misalign observed data, therefore calibrating it and bringing it in line with observed data may supply modellers with more reliable and richer sources of data. Data assimilation methods, namely combining data from multiple sources, are well known in environmental studies. There is a very rich statistical literature on data fusion with the objective of enriching the information for inference. These statistical methods are often based on Bayesian Hierarchical methods for space time data and are constructed around the idea of relating the monitoring station data and the simulated data using spatial linear models with spatially varying coefficients. Since these relations involve data measured at different spatial resolutions, the models often are called downscaler models. The principal objective of these downscaler models is to relate observations measured at different space resolutions using spatial linear models. However, as a by-product, these models can be used for calibrating one set of data as a function of the other.

In this talk we will address several aspects of the calibration problem of interest namely, calibrating extreme values of simulated data based on observed data—using different methods and models. Although, our ultimate goal is the development of statistical methods for data fusion and calibration that can extrapolate beyond the range of observed data, into the tails of a distribution, we will first address calibration methods for the whole range of data. The data available for this particular study corresponds to simulated wind speeds from a simulator (The WRF model, version 3.1.1) at a regular grid of 81ksq grid cell size, obtained at 10 minutes interval from 2006-2013; however only daily maximum wind speed will be used.

Data fusion/calibration methods to update simulated data based on the observed data: an application to wind speed data.

Observed daily maximum wind speed is also available during the same period of time, from 117 stations in Portugal mainland, but missing observations reach to 90% in some stations. Only around one third of the stations have less than 30% missing observations. There is an additional challenge: Simulated and observed daily-maxima wind speeds, particularly at some stations and at the right tail, do not match well. Our objective is to explore several methods to model the relationship between simulated and observed wind speeds at observation sites, so to extrapolate this relationship in space at grid cell or county level resolution. In other words, more than imputing missing observations, we want to use simulated wind speeds for risk assessment, after being calibrated, i.e., brought in line with observed wind speeds.

After giving a brief description of standard data fusion/calibration methods to update simulated data based on the observed data, we will describe in detail two specific data fusion/calibration methods and show how our wind speed data can be calibrated using these methods. We also briefly explain how calibration can be extended specifically to data coming from the tails of simulated and observed data, using asymptotic models and methods suggested by extreme value theory.

Keywords: Data fusion; Calibration; Bayesian hierarchical modelling; Spatial extremes.

Acknowledgments:The authors acknowledge the financial support received by Fundação para a Ciência e Tecnologia, Portugal, through the projects PTDC/MAT-STA/28649/2017 and UID/MAT/00006/2019.

Shades of Dark Uncertainty and Consensus Value for the Newtonian Constant of Gravitation

C. Merkatas[1], B. Toman[1], A. Possolo[1], S. Schlamminger[1]

[1]National Institute of Standards and Technology, Gaithersburg, MD, USA
E-mail (corresponding author): antonio.possolo@nist.gov

The Newtonian constant of gravitation, *G*, stands out in the landscape of the most common fundamental constants owing to its surprisingly large relative uncertainty, which is attributable mostly to the dispersion of the values measured for it by different methods and in different experiments, each of which may have rather small relative uncertainty.

This study focuses on a set of measurements of *G* comprising results published very recently as well as older results, some of which have been corrected since the original publication. This set is inconsistent, in the sense that the dispersion of the measured values is significantly larger than what their reported uncertainties suggest that it should be. Furthermore, there is a loosely defined group of measured values that lie fairly close to a consensus value that may reasonably be derived from all the measurement results, and then there are one or more groups with measured values farther away from the consensus value, some appreciably higher, others lower.

This same general pattern is often observed in many other interlaboratory studies and meta-analyses. In the conventional treatments of such data, the mutual inconsistency is addressed by inflating the reported uncertainties, either multiplicatively, or by the addition of "random effects", both reflecting the presence of *dark uncertainty*. The former approach is often used by CODATA and by the Particle Data Group, and the latter is common in medical meta-analysis and in metrology. However, both achieve consistency ignoring how the measured values are arranged relative to the consensus value, and measured values close to the consensus value often tend to be penalized excessively, by such "extra" uncertainty.

We propose a new procedure for consensus building that models the results using latent clusters with different shades of dark uncertainty, which assigns a customized amount of dark uncertainty to each measured value, as a mixture of those shades, and does so taking into account both the placement of the measured values relative to the consensus value, and the reported uncertainties. We demonstrate this procedure by deriving a new estimate for *G*, as a consensus value $G = 6.67408 \times 10^{-11}$ m^3 kg^{-1} s^{-2}, with $u(G) = 0.00024 \times 10^{-11}$ m^3 kg^{-1} s^{-2}.

Bayesian inference to evaluate predictions with calibration function

C Yardin[1], M Sancandi[2]

[1]Laboratoire National de métrologie et d'Essais, 1 rue Gaston Boissier, Paris, France
[2]Commissariat à l'énergie atomique et aux énergies alternatives, CEA/CESTA, Bordeaux, France
E-mail (corresponding author): catherine.yardin@lne.fr

Calibration of a measurement device is processing in two steps: in a first step, a calibration function is estimated; in a second step, the function is used to predict a new value. A lot of articles deal with the estimation of the function and some of them compare different approaches to evaluate its associated uncertainty (GUM, GUM_S1 and Bayesian inference) [1,2].

On the contrary, the second step devoted to predicting a new value is still too rarely processed. The ISO 28037 standard [3] is a complete document which proposes a GLS (generalised least squares) estimate of the function and the calculation of direct or inverse forecasts. In this standard, the uncertainties associated to parameters and forecasts are evaluated using the GUM method. It would be appropriate to have a similar document both for Monte Carlo and Bayesian methods (although it is not impossible to mix two methods, for example estimate the function with Bayesian inference and evaluate the forecast with GUM approach).

Recently, Bayesian inference for calibration is deeper investigated [4,5,6]. Last year, we have proposed a bayesian model where uncertainties are associated both to x and y variables, and the posterior distribution of the calibration function estimated with numerical methods [7].
To conduct the prediction study, we have identified cases where the forecasts can be computed with Bayesian inference. In these cases, the new value used as an entry in the calibration function is an observation i-e a new measurement (x or y). In the other cases, the new value is not an observation - it is a target or a simulated value tested in the analysis of the device. These cases are processed in [4,5].

Because the first cases are more innovative (for calculation) and illustrate the common use of a measuring instrument, we focused our study on them. Bayesian inference applied to the forecasts was conducted using two methods of calculation. The first, called sequential mode, is a direct application of the calibration procedure: the posterior pdf of the calibration function parameters and a prior associated to the forecast are used to deduce a posterior pdf. This posterior pdf defines the forecast estimates. The second method, called global mode, allows to evaluate both the function parameters and the forecast in a one-step procedure. Compared to the first step of calibration, the calibration function could be updated.

Bayesian inference to evaluate predictions with calibration function

We tested both methods on the example of a scale whose calibration function is a straight line. The results depend on the assumptions made about the variables, observations and parameters.

References

[1] C. Elster and B. Toman. Bayesian uncertainty analysis for a regression model versus application of GUM Supplement 1 to the least-squares estimate. Metrologia 48,233-240, 2011.
[2] C. Elster and al., A Guide to Bayesian Inference for Regression Problems.Deliverable of EMRP project NEW04 Novel mathematical and statistical approaches to uncertainty evaluation, 2015.
[3] ISO/TS 28037 Determination and use of straight-line calibration function
[4] G J P Kok, A M H van der Veen,, P M Harris, I Smith and C Elster, Bayesian analysis of a flow meter calibration problem, Metrologia 52, 2015
[5] K. Klauenberg, G. Wübbeler, B. Mickan, P. Harris, and C. Elster. A tutorial on Bayesian Normal linear regression, Metrologia 52,878-892, 2015.
[6] I Lira and D Grientschnig, Error-in-variables models in calibration, Metrologia, 54, S133-S145, 2017.
[7] C. Yardin, S. Amar, N. Fischer, M. Sancandi and M. Keller, Bayesian estimation of a polynomial calibration function associated to a flow meter, Advanced Mathematical and Computational Tools in Metrology and Testing XI, Décembre 2018, pp 417 – 426

Development of a Virtual Flow Meter

A. Weissenbrunner[1], S. Schmelter[1]

[1]Physikalisch-Technische Bundesanstalt (PTB), Abbestr. 2-12, 10587 Berlin
E-mail (corresponding author): andreas.weissenbrunner@ptb.de

Flow meters are sensitive to the flow conditions present at its application. These are most often not the same as at their calibration, which are performed downstream of a long straight piping leading in an ideal symmetric flow profile. Elbows, T-Junctions and other fittings cause asymmetric flow profiles with swirls, which leads to errors of the installed flow meter. To assess these errors a virtual flow meter is in development at the Physikalisch-Technische Bundesanstalt (PTB) within the competence center "Metrology for Virtual Measuring Instruments" (VirtMet).

The virtual flow meter is used to estimate uncertainties of ultrasonic and electromagnetic flow meters as well as a Laser Doppler Velocimeter (LDV). Therefore, a model of the measurement principle is applied to the flow conditions downstream different elbow combinations simulated with Computational Fluid Dynamics (CFD). It is demonstrated that simulations with standard Reynolds Averaged Navier-Stokes (RANS) turbulence models predict enhanced uncertainties, caused by simplified assumptions of the turbulence model. In particular, the azimuthal orientation of the flow profile cannot be predicted, compare Figure 1. To reduce the uncertainties of the flow simulation, and yet of the virtual flow meter, a higher order transient turbulence model Detached Eddy Simulation (DES) is used.

With this approach a virtual flow meter can not only be used to estimate the uncertainty of the real flow meter at its installation position, but also to correct its measurement value. To overcome the enormous computational costs of the DES model, a reduced order surrogate model is developed, which permits an accurate prediction of the flow field within low computation time.

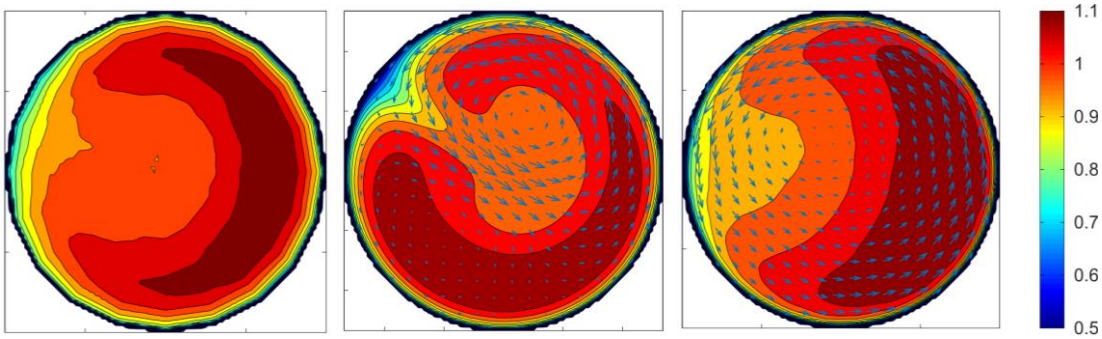


Figure 1: Axial velocity profile 6 diameters downstream of a combination of six elbows, measured with an LDV system (left), simulated with a RANS (middle) and a DES (right) turbulence model. The arrows represent the cross-flow components.

Efficient reliability analysis with model reduction techniques

J.F. Unger[1], A. Robens-Radermacher[1], F. Held[1], I. Coelho Lima[1]

[1]BAM, Federal Institute for Materials Research and Testing, Unter den Eichen 87, 12205 Berlin, Germany
E-mail (corresponding author): joerg.unger@bam.de

One of the most important goals in civil engineering is to guaranty the safety of constructions. National standards prescribe a required failure probability in the order of 10^{-6} (e.g. DIN EN 199:2010-12). The estimation of these failure probabilities is the key point of structural reliability analysis. Generally, it is not possible to compute the failure probability analytically. Therefore, simulation-based methods as well as methods based on surrogate modelling or response surface methods have been developed. Nevertheless, these methods still require a few thousand evaluations of the structure, usually with finite element (FE) simulations, making reliability analysis computationally expensive for relevant applications.
The aim of this contribution is to increase the efficiency of structural reliability analysis by using the advantages of model reduction techniques. Model reduction is a popular concept to decrease the computational effort of complex numerical simulations while maintaining a reasonable accuracy. Coupling a reduced model with an efficient variance reducing sampling algorithm significantly reduces the computational cost of the reliability analysis without a relevant loss of accuracy.

Introduction

The probability of failure P_f is prescribed in standards to guaranty the safety of the construction. P_f is the integral over the failure domain F over the n -dimensional joint distribution function f_X of all given random parameters $\boldsymbol{X} = [X_1, X_2, \dots, X_n]$ (e.g. material parameters).
The failure domain is implicitly given by the limit state function $g(\boldsymbol{X})$ separating the space into a safe $g(\boldsymbol{X}) > \boldsymbol{0}$ and a failure $g(\boldsymbol{X}) < \boldsymbol{0}$ domain. Usually, g compares the resistance R of the structure and the corresponding current stress S: $g(\boldsymbol{X}) = 1 - S(\boldsymbol{X})/R(\boldsymbol{X})$.
Generally, it is not possible to solve the probability integral analytically. Therefore, simulation-based methods such as Monte Carlo, importance sampling [1] or subset simulation [2], to name a few, have been developed. Nevertheless, these methods still require a few thousand numerical simulations of the problem, making reliability analyses computationally expensive.
In other fields, a common way to reduce the computational effort of numerical simulations is the application of model reduction techniques. The proper generalized decomposition method (PGD) [3] is one recently developed model reduction approach. The main idea is to compute a so-called abacus, which computes all solutions of a problem depending on many parameters directly. The approach is based on the assumption that the solution field (in our case the displacements $\boldsymbol{u}$) can be represented in a separated form

$$\boldsymbol{u}(\boldsymbol{\kappa}) \approx \boldsymbol{u}^{N_{PGD}} = \sum_{i=1}^{N_{PGD}} \prod_{k=1}^{d} F_k^i(\kappa_k)$$

consisting of a number N_{PGD} of PGD modes F_k^i each depending only on a single coordinate κ_k. The PGD coordinates include all possible coordinates, e.g. material or load parameters. This way, a multidimensional problem is split into several low-dimensional problems. The modes are computed in an iterative way using a progressive PGD solver with fixed-point iteration [4, 5].
The goal of this contribution is to couple a PGD model with an efficient sampling strategy to reduce the computational cost of structural reliability analyses.

Efficient reliability analysis

The efficient reliability analysis is based on the coupling of a PGD reduced model for the structure with an importance sampling technique. A PGD model is generated for the structural problem using the space $\boldsymbol{x}$ and the random parameters $\boldsymbol{X}$ as PGD coordinates. The integral related to the computation of the failure probability is approximated using importance sampling by means of

$$P_f \approx \frac{1}{N}\sum_{k=1}^{N} I(\boldsymbol{v}^k)\frac{f_X(\boldsymbol{v}^k)}{h_X(\boldsymbol{v}^k)}$$

using the importance sampling density h_X the indicator function $I(\boldsymbol{v}^k) = 1\ if\ g(\boldsymbol{v}^k) < 0\ otherwise\ I(\boldsymbol{v}^k) = 0$ and N samples $\boldsymbol{v}^k$ in $\boldsymbol{X}$. h_X is chosen as f_X with a shifted mean (in the best case it is the design point). This shifted mean is computed using an approach similar to subset simulations [2, 6]. The idea is to decompose the problem into a sequence of sub-problems with a higher failure probability than the expected one of the original problem. In an adaptive process, the failure events F_i with corresponding limit states g_i are defined in such a way that the corresponding probabilities $P_{f_{g_i}}$ are close to a prescribed threshold p_0. The process ends if the original limit state g_i is reached. The shifted mean is then computed by means of the average of the samples lying close to the limit state. The evaluation of the limit state function at each sample $\boldsymbol{v}^k$ is done by evaluating the PGD solution which is computationally very efficient. The approach is demonstrated for several numerical examples demonstrating the accuracy as well as the speed-up.

Conclusion

An efficient reliability analysis method was presented. A high reduction in the computational time of standard reliability analysis was achieved by embedding a PGD model reduction in a variance reducing sampling technique. On the other hand, the same accuracy level of the estimated failure probabilities can be reached as using a full FE model. The extension of the proposed method to higher dimensions, nonlinearity as well as error estimation will be done in future works.

Literature

[1] S. Engelund and R. Rackwitz, Struct. Saf. 12, 255 – 276 (1993).
[2] S.-K. Au and J. L. Beck, Prob. Eng. Mech. 16, 263 – 277 (2001).
[3] F. Chinesta, P. Ladeveéze and E. Cueto, Arch. Comp. Meth. Eng. 18, 295 – 404 (2011).
[4] S. Zlotnik, P. Díez, D. Gonzalez, E. Cueto and A. Huerta, Adv. Model. Simul. Eng. Sci. 2, 2213–7467 (2015).
[5] A. Falcó, N. Montés, F. Chinesta, L. Hilario and M.C. Mora, J. Comput. Appl. Math. 330, 1093 – 1107 (2018).
[6] I. Papaioannou, W. Betz, K. Zwirglmaier and D. Straub, Prob. Eng. Mech. 41, 89 – 103 (2015).

Characterization in size of aggregated nanoparticles measured by SEM: an illustration of deep generative models in metrology

N.Fischer, L.Coquelin,

LNE, 29 avenue Roger Hennequin, 78197 Trappes, France
nicolas.fischer@lne.fr

Recent advances in deep generative models based on convolutional neural networks (CNNs) are used to demonstrate the potential of these approaches for the estimation of particle size distribution on images of aggregated TiO_2 particles obtained by Scanning Electron Microscopy (SEM).This very promising framework shall permit effective automation of SEM measurements analysis.
Indeed, common image processing software bring the end-users with segmentation algorithms as well as measuring tools to estimate individual particle diameters. In the case of aggregated nanoparticles, most particles suffer missing contents and are not considered in the computations.
In this work, we have used a recently developed method called ”context encoders” to predict missing parts of the nanoparticles. The approach is tested against simulated and real dropped image regions.
Finally Consideration is made to evaluate the performance of the method based on both real and simulated particles using cross validation.

Large Scale Inference with Applications to Environmental Monitoring

L. Sharrock[1], N. Kantas[1]

[1]Department of Mathematics, Imperial College London, London, UK
E-mail (L. Sharrock): louis.sharrock16@imperial.ac.uk

Ambient air pollution poses a significant threat to global health and has been associated with a vast range of adverse health effects, including cardiovascular and respiratory diseases in addition to various cancers [1,2]. Fine particulate matter (PM2.5) in particular has been established as a key driver of global health issues, with an estimated 3 million deaths in 2014 being attributable to PM2.5 [3]. In this context, the accurate and timely monitoring and prediction of pollutant exposure is of clear relevance.

Environmental phenomena such as air pollution, even if defined continuously over a region and in time, can typically only be monitored and measured at a limited number of spatial locations and time points. Furthermore, they are often highly complex in so far as their dependence structure across time and space is non-trivial, non-separable or non-stationary. It is generally also the case that the number of spatial locations at which inference is desired is large, that the data are obtained with significant observational uncertainty, and that there are missing observations at numerous spatial and temporal locations.

These factors motivate the use of partially observed space-time varying stochastic dynamical models. These are both temporally dynamic *and* spatially descriptive; and provide a natural framework by which to account for uncertainty in both the data, and our knowledge of the underlying physical process.

Such models are now relatively well understood. However, their practical implementation is often hampered by the need for expensive matrix factorisations, an operation whose computational complexity increases in cubic order with the number of spatial and temporal locations. This is commonly termed the *big n problem*. In environmental monitoring applications, where the advancement of sensor and storage technologies, the advent of personal air pollution sensors, and the reduced cost of data acquisition have rendered very high-dimensional data sets increasingly common, this problem is of growing concern.

In this work, we consider modelling spatio-temporal environmental phenomena via a linear Gaussian state space model. Following [4], we assume that the process of interest evolves according to a physically motivated *stochastic partial differential equation* (SPDE) - namely, the stochastic advection diffusion equation. The parameters in this SPDE may be interpreted as explicitly modelling physical phenomena such as advection and diffusion, which occur in many natural processes.

We assume that this process is hidden, but generates noisy observations according to a linear *observation equation*. The parameters in this equation may be interpreted as the

variance(s) and bias(es) of sensors used to measure the latent process defined by the SPDE.

We propose two likelihood-based methods for the estimation of the unknown parameters in this model. Significantly, a truly *online* parameter estimation method is achieved via the Recursive Maximum Likelihood (RML) approach. In this case, parameter estimates are sequentially updated as soon as new observations become available, without the need to revisit the past. This approach relies on gradient methods to maximise the average log-likelihood of the unknown parameters. In particular, parameter estimates are recursively updated in the direction of the gradient of the conditional log-likelihood, for which we have obtained an explicit analytical expression.

On the basis of numerical simulations, we establish the ability of both offline and online likelihood methods to accurately estimate *static* parameters in the SPDE model. We further establish the ability of the online likelihood method to accurately estimate *dynamic* parameters in the SPDE model: that is, to track changes in these parameters over time. Notably, this includes the ability to track instantaneous changes in the parameters at certain distinct time instances, or change-point detection.

Extending the contributions of [4], we also demonstrate that, under certain assumptions, these methods can be implemented efficiently in the spectral space at a total computational cost of O(TN log(N)), where T and N denote the number of points in time and space respectively.

The proposed methodology is then applied to estimate the parameters of a dynamical spatio-temporal model for the daily maximum nitrogen dioxide (NO2) concentrations in Greater London for the five year period 2007-2011. While there remains a need for further refinement, preliminary results indicate that this model outperforms the simulation based Air Quality Unified Model, as well as some other simple statistical models used for benchmarking.

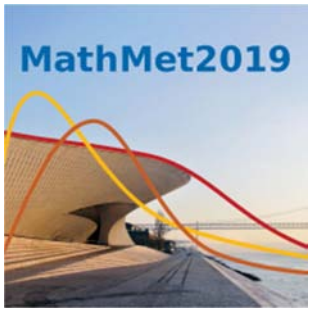
---

[1] R.D.Brook,S.Rajagopalan,C.A.Pope,J.R.Brook,A.Bhatnagar,A.V.Diez-Roux,F.Holguin,Y.Hong,R.V. Luepker, M. A. Mittleman, A. Peters, D. Siscovick, S. C. Smith, L. Whitsel, and J. D. Kaufman. Particulate Matter Air Pollution and Cardiovascular Disease. *Circulation*, 121(21) : 2331–2378, June 2010.

[2] G. Hoek, R. M. Krishnan, R. Beelen, A. Peters, B. Ostro, B. Brunekreef, and J. D. Kaufman. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ. Health*, 12(1) : 43, May 2013.

[3] World Health Organisation. Ambient Air Pollution: a Global Assessment of Exposure and Burden of Disease. Technical report, World Health Organisation, Geneva, 2016.

[4] F. Sigrist, H. R. Kunsch, and W. A. Stahel. Stochastic partial differential equation based modelling of large spacetime data sets. J. Roy. Stat. Soc.: Series B (Statistical Methodology), 77(1) : 3–33, 2015.



A Statistical Metrology Approach to Compare the Quality of Electrocardiogram Waveforms

Janos Palhalmi[1,2]
[1]*DataSenseLabs Ltd. Budapest, Hungary,*
[2]*EastWestData Ltd. Budapest, Hungary*
E-mail (corresponding author): janos.palhalmi@datasenselabs.net

---

Introduction

Biometric metrology is becoming increasingly important as wearable application specific biosensors are capable of generating accurate raw signals representing certain vital states.

A comparative statistical approach has been worked out to answer questions arising from a health informatics and bioinformatics perspective.

In this specific work a high-quality prototype MEMS (micro electromechanical system) biosensor solution and a gold standard physiological signal measurement system (BioPac Systems Inc.) will be compared to analyse the deep data level similarities or differences between the electrocardiogram (ECG) raw biomedical signals generated by the two individual systems.

Data collection and pre-processing

Two minutes long parallel recordings (sampling frequency = 500 Hz) will be carried out with both systems on 8 healthy human subjects. Both systems have been controlled by the manufacturer's user-interface. For the test, configuration settings and filter settings were optimized to achieve the highest signal quality.

Results and Conclusions

A wavelet coherence-based method was worked out to compare the relevant frequency bands and two different correlation-based methods were developed to compare the wave-to-wave stability and similarity of the two examined waveforms.

Magnitude squared coherence values will be compared in all the explored frequency bands (0.5-40 Hz). Correlation coefficients and p values will be systematically analysed in the ECG wave-to-wave comparisons.

From a biosensor test measurement perspective, we will conclude whether the quality of the ECG signal recorded by the prototype MEMS system reaches the same performance level as the gold standard system.

From a statistical metrology perspective, the uncertainty levels of the above described comparative method will be discussed based on the purpose to satisfy questions arising from the relevant health informatics application areas.

Binary linear regression in dynamic force measurement and uncertainty estimation

Jun Yang[1], Yinbao Cheng[2], Li Zhang[1], Yuan Liu[1], Xiao Yin[1]

[1] Changcheng Institute of Metrology& Measurement (CIMM), huanshan village, wenquan town, haidian district, Beijing100095, China
[2]School of Instrument Science and Opto&electronics Engineering, Beihang University, Beijing 100083,China
E-mail (corresponding author): yangjun@cimml.com.cn

Aiming at the problem of effective mass measurement of the force sensor in the impulse force calibration with drop mass by laser interferometry, a binary linear regression based on the least squares method is proposed. The effective mass and the sensor's amplitude sensitivity are directly obtained based on the measured waveforms. Based on the Monte Carlo method, the measurement uncertainty of the effective mass and the amplitude sensitivity of the sensor are estimated. It is compared with the repeatability of multiple measurements to verify the validity of the Monte Carlo method.

1. Introduction

Impulse calibration is one important calibration method for the piezoelectric sensor's amplitude sensitivity. In the drop-mass impulse force calibration by laser interferometry, the traceable measurement of dynamic force requires correction of the effective inertial force of the force sensor and its fixture.

The calibration device after considering the effective inertial force of the sensor is shown in Figure 1. The dynamic force measurement model at this time is:

$$f = M(a+g) + m(a_m + g) \qquad (1)$$

Where f is the force measured by the sensor. a is the dropping hammer' acceleration which is by the laser vibrometer. g is the local gravity acceleration. a_m is the equivalent mass of the force sensor and its fixture, which is measured by one accelerometer.

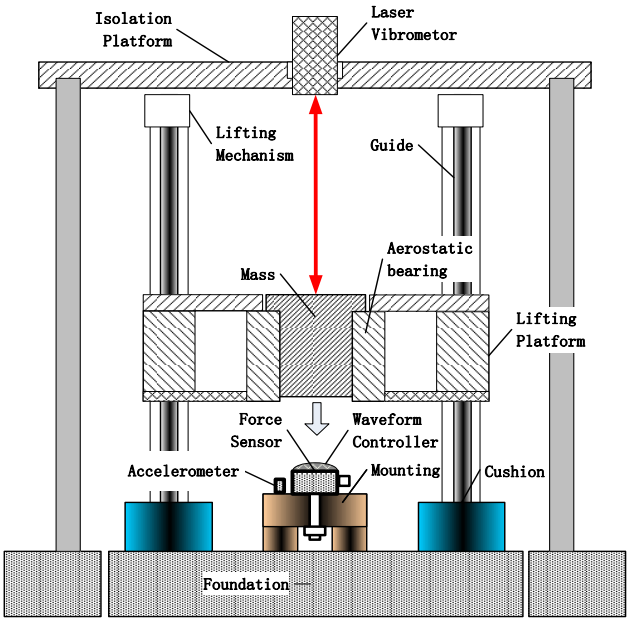


Figure 1. Structure of the impulse force calibration device by laser interferometry

However, the analysis of the effective mass of the force sensor and its contribution to the measurement uncertainty of the amplitude sensitivity is a problem.

2. Measurement method

Binary linear regression in dynamic force measurement and uncertainty estimation

One possible method for obtaining the equivalent mass of the sensor is variable quality experiments with the sinusoidal force calibration device. Aiming at the insufficient driving ability of the existing sinusoidal force calibration device and the cumbersome operation of the method, a binary linear regression method based on the measurement curves and the least squares method is designed to calculate the equivalent mass and the sensitivity of the sensor.

U is the sensor's output. According to the formula (1), the definition of the amplitude sensitivity and the three measured curves, which include $a[i]$, $U[i]$ and $a_m[i]$ (i=0~n-1), a binary linear equation is obtain:

$$y[i] = b_1 x_1[i] + b_2 x_2[i] + b_0 \quad (2)$$

$y[i] = Ma[i]$ is taken as the dependent variable, $x_1[i] = U[i]$ and $x_2[i] = a_m[i]$ are the independent variables. $b_1 = \frac{1}{s_d}$ and $b_2 = m$ are parameters to be calculated.

Binary linear regression base on the least square is used to obtain b_1、b_2. This eliminates the need for additional special measurements of *m* to get the sensitivity needed for calibration.

3. Uncertainty estimation method

It is difficult to estimate the binary linear regression's uncertainty especially considering the uncertainty of x1, x2 and y.

The Monte Carlo method is used to estimate the measurement uncertainty of s_d and m. The number of Monte Carlo trials is *N*. Storage requirements at least:

$$N + 3 * n^N$$

4. Experiments and analysis

In six repetitive experiments with a peak force of 117kN, the standard deviation of the measurement of m is 0.0496kg. The m measurement uncertainty obtained by the Monte Carlo method is 0.3383 kg (k=2), which is much larger than the measurement standard deviation of m, indicating that the result obtained by the method is available. The measurement uncertainty of s_d is 0.0197 pC/kN (k=2), and the relative measurement uncertainty is 1.1% (k=2).

by

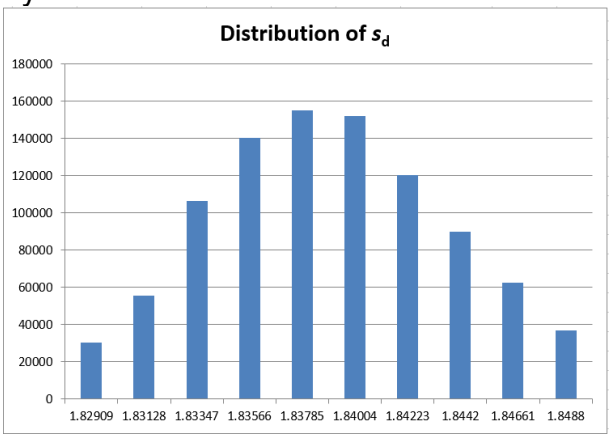


Figure 2. distribution of s_d by Monte Carlo method

5. Conclusions

A novel binary linear regression method realizes the fast measurement and calculation of the effective mass and sensitivity of the sensor in impulse force calibration, and the measurement uncertainty analysis of binary linear regression and calibration result is completed by the Monte Carlo method.

Research on Analysis Method of Amplitude-Frequency Response in Dynamic Pressure Calibration with Shock Tube

Li Zhang, Yahui Wu, Jun Yang

Changcheng Institute of Metrology& Measurement (CIMM), huanshan village, wenquan town, haidian district, Beijing 100095, China
E-mail (corresponding author): zhangli@cimm.com.cn

The shock tube can generate a step pressure with a rise time of about 10ns. It is widely used in the dynamic performance calibration of the pressure test system. The frequency response can be used to analyze the dynamic performance of the test system from the frequency domain. This paper introduces the common step signals, and analyzes the problems that the discrete Fourier transform will deal with when signal sequences do not satisfy the absolute summable conditions. The solution is proposed and verified by the experiments.

1. Introduction

The shock tube can generate a pressure step with a very short rise time and is widely used in the field of dynamic pressure and dynamic force calibration. Limited by the pressure test system, it has not been possible to measure the true pressure process of the reflected shock, and the dynamic performance of the test system can be obtained from the frequency domain by amplitude-frequency response analysis.

The frequency response of the system is determined by the external excitation and the characteristics of the system itself. When the input signal is a step signal or a shock signal, the zero state response of the system is called the step response and the impulse response. The step response is often used to determine the dynamic performance of the system because the step signal is generally the most critical operating condition for the system. If the dynamic performance of the step response of the system meets the requirements, the dynamic performance of the system will be satisfactory under the influence of other inputs.

Shock tubes are often used as excitation sources to calibrate the dynamic performance of pressure test systems. Since the shock tube can generate a step pressure with a very short rise time, which approximates the ideal step signal, the output signal of the pressure test system approximates the step response. For this step response analysis, the dynamic characteristics of the pressure test system can be studied.

2. Analysis methods and the problems

The main analysis methods of amplitude-frequency response include discrete Fourier transform and differential analysis.

2.1 Frequency response analysis based on Fourier transform

The input signal x(n) and the output signal y(n) of the system are separately subjected to discrete Fourier transform to obtain the spectrum X(s) and Y(s), corresponding to the input and output signals, by dividing the spectrum of the output signal by the spectrum of the input signal, the amplitude-frequency characteristics of the system can be obtained, i.e

$$A(s) = \left| \frac{Y(s)}{X(s)} \right|$$

2.2 Differential analysis

The differential response of the system's step response is the impact response of the system. Therefore, when the input signal of the system is a step signal, the output of the system is the step response y(n) of the system, and the step response y(n) is differentiated. After that, the impact response curve of the system is obtained, and then the impulse response curve is subjected to discrete Fourier transform to obtain the amplitude-frequency response of the system.

2.3 The problems

Theoretically, two methods for processing the same set of experimental data should obtain the similar amplitude-frequency characteristics. However, the experimental results show that the

Research on Analysis Method of Amplitude-Frequency Response in Dynamic Pressure Calibration with Shock Tube

amplitude-frequency characteristics obtained by the differential analysis method are relatively coherent(Fig1(b)), and the amplitude-frequency response obtained by the Fourier transform analysis method has many jumps(Fig1(a)).

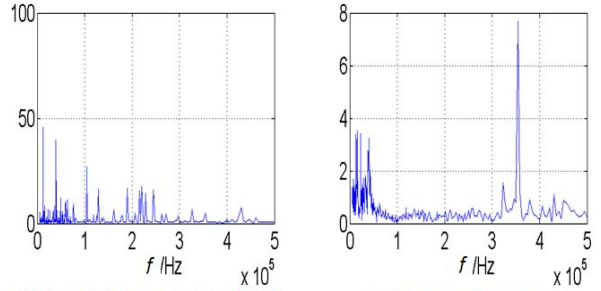


Figure 1 different amplitude-frequency characteristics for the same signal.

3. Solution

Two methods to analyze the same data will get different spectrum, which is related to the requirement of absolute sum of the signals by the discrete Fourier transform. If the aperiodic and non-harmonic signals such as step signals are processed, non-convergence problem will occur.

When the discrete Fourier transform of the finite-length signal is actually processed, it is equivalent to obtaining a reciprocal signal by adding zeros at the end of the signal, and then performing discrete Fourier transform. Therefore, when the frequency response of the system is analyzed by the Fourier transform analysis method, the input of the system is not a step signal, but a square wave signal. The output signal only contains the response of the pressure test system to the rising edge of the input signal, and does not include the response of the system to the falling edge caused by the zero padding. This results the difference between the spectrum analysis and differential analysis method. Assuming that the pressure test system is a causal system, the response of the rising and falling edges of the square wave is the same, and the original signal is extended, that is, for the signal x(n) of length N, the signal is constructed:

$$x'(n)=\begin{cases} x(n) & n \le N \\ x(N)-x(n-N) & n > N \end{cases}$$

Using the same sensor, multiple sets of experiments were carried out under 100k low-pass filtering and unfiltered conditions. The experimental results were analyzed by frequency response analysis based on Fourier transform, differential analysis, and extended frequency response analysis based on Fourier transform to analyze the amplitude-frequency response of the pressure test system. The results as shown in Fig. 2 and 3 are obtained.

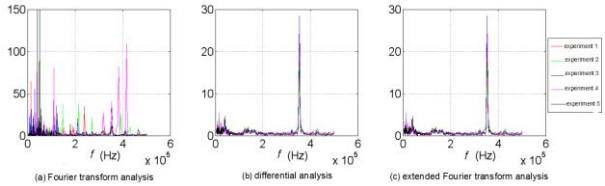


Figure 2. Spectrogram obtained with experimental data from low-pass filtering

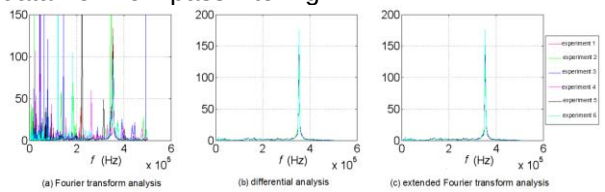


Figure 3 Spectrogram obtained from unfiltered experimental data

From the results of Fig. 2 and Fig. 3, the glitch point of the suspected noise is commonly found in the spectrum obtained by the frequency response based on Fourier transform, and the spectrum obtained by the frequency response analysis after the signal is extended is analyzed. There is no such problem. Therefore, it is basically possible to determine the extension of the output signal of the pressure test system, which can effectively solve the influence of the discrete Fourier transform on the spectrum of the test system in processing the finite-length signal.

Musing on modeling in measurement science: from the prescriptive model to the descriptive model implementing the former in experiments

F. Pavese

[1]IMEKO, Torino, Italy
E-mail (corresponding author): frpavese@gmail.com

In measurement science the object of a measurement is called “measurand”. It is defined in VIM:2012 clause 2.3 as “quantity intended to be measured” (but in GUM:1995, B.2.9 it is defined instead according to VIM:1993 clause 2.6 “particular quantity subject to measurement”. In VIM:2012, NOTE 1 to clause 2.3 specifies “the specification of a measurand requires … description [i.e., a model] of the state of the phenomenon, body, or substance …”, and NOTE 3 adds “the measurement, including the measuring system and the conditions under which the measurement is carried out, might change the phenomenon, body, or substance such that the quantity being measured may differ from the measurand as defined. In this case, adequate correction is necessary” (emphasis added).
The presentation will start from observing that the concept (model) of any particular measurand has to be shared by the relevant Community, because the same measurand is supposed to be the object of replicated measurements that must be comparable with each other, i.e., that the measurand has to be recognised as a quantity having a current recognisable common meaning in the community. In the dialect of the science philosophers, this means that it should be “projected into a social framework”.
Also in the scientific frame this means that the measurand model must be one of the “prescriptive” type, meaning “giving directions or injunctions”—which does not always mean “physical model”, or “measurement model” in VIM:2012 (clause 2.48), or “mathematical model” in GUM:1995 (clauses 3.1.6, 3.4.2, 4.1).
It will be shown that it follows the fact that the design of an experiment (DoE) must start from this initial conceptual model of the measurand, “socially shared”, not from building up the descriptive model (often called “experimental model”) of the measuring system.
In addition, the latter modeling is necessarily specific to each measurement arrangement; i.e. based on the specific solutions that are chosen in order to implement the prescription in each experiment.
The initial conceptual model, being independent on any specific experimental implementation, is clearly a highly idealised one (in fact is often called the “ideal condition” of the measurand). It does not even allow appreciating the experimental difficulties and compromise (which, in turn, are graduated depending on the target uncertainty). The latter arise from three categories of sources indicated in the VIM:2012:
1) the phenomenon, body, or substance;
2) the measuring system;
3) the conditions under which the measurement is carried out.

A corresponding descriptive model must indicate the specific measurement conditions, often called “physical” or “experimental” model—here it will not necessarily correspond to any of them—where most often the conditions do not correspond to the “ideal” one.

Musing on modeling in measurement science: from the prescriptive model to the descriptive model implementing the former in experiments

The paper discusses the non-simple roadmap bringing from the prescriptive model to the descriptive one, and will show that this process brings to the explicit need to define a “reference” state for each influence quantities and to an alternative classification of the measurement uncertainties (errors), with respect to the traditional one—random and systematic.
The presentation will conclude with a proposal for a different logical scheme with respect to the one reported by the concept diagram in VIM:2012 of the “measured quantity value”.

Mathematical model of the volume of no-OIML R111-1 standard weights

O. J. Purata-Sifuentes[1], A. A. Guillén-Capetillo[2], A. Espinal-Jiménez[3]

[1, 3]Universidad de Guanajuato, Lascurain de Retana 5, Guanajuato, México
[2]CIATEC, Omega 201, León, México
E-mail (corresponding author): opurata@ugto.mx

During mass measurements in the air, the volume or density of the body being measured must be known, to be able to determine the corresponding correction due to air buoyancy. Most of the standardized weights that are calibrated in the Mexican market have geometry as described in recommendation R-111 of the International Organization of Legal Metrology. However, there are also many requests in the Mexican market for the calibration of weights whose geometry does not follow the one illustrated in recommendation R-111.

Although in the literature various methods are reported for the determination of the volume of weights, undoubtedly the hydrostatic weighing is the most accurate, but also the most expensive to perform. The determination of the volume of weights by means of its geometric characterization is a much more economical alternative, although also less exact.

A mathematical model for the calculation of the volume of standard weights by geometric measurement was deducted. The model is suitable for weights with geometry according to ASTM E617 which is different from the specified and published in the OIML R 111-1 recommendation. The proposed model involves the possibility of making geometric measurements without surface contact of the weights, for example, with an optical comparator, which gives the additional advantage of eliminating the risk of scratching the weights. An assessment of the current model against hydrostatic weighing allows highlighting the relevance and higher accuracy of the proposed mathematical model, which makes possible to calculate the density of standard weights through just geometric measurements.

Role of measurement uncertainty in the comparison of average areal rainfall methods and its impact on conformity assessment

A. S. Ribeiro[1], M. C. Almeida[1], M. G. Cox[2], J. A. Sousa[3], L. Martins[1], D. Loureiro[1], R. Brito[1], M. Silva[1], A. C. Soares[1]

[1]LNEC – National Laboratory for Civil Engineering, Avenida do Brasil, 101, 1700-066 Lisbon, Portugal
[2]National Physical Laboratory, Teddington TW11 0LW, UK
[3]Portuguese Institute for Quality, Rua António Gião, 2, 2829-518, Caparica, Portugal
E-mail (corresponding author): asribeiro@lnec.pt

1. INTRODUCTION

The growing awareness of the impact of climatic change and the United Nations' sustainable development goals (SDG) [1] show the need to have reliable measurements of quantities (like precipitation and rainfall) to support management of water resources. Measurement of precipitation (vertical depth of collected water equivalent to the water that would be collected at a level surface area) and of rainfall (in terms of liquid water droplets condensed from atmospheric water vapour that form precipitation, measured by the depth of water accumulated on the horizontal projection of Earth without losses) are real needs.

2. AVERAGE AREAL RAINFALL METHODS

Rainfall can be measured by different types of techniques and rain gauges (e.g., tipping bucket, weighing, optic) by collecting the quantity of liquid precipitation over a defined period of time [2]. Measurements provide information to meteorologists, hydrologists and scientists, allowing computation for predictive models to be developed.

Rainfall measurements obtained at single stations allow the average areal rainfall using different methods to be calculated: *arithmetic mean method*; *Thiessen polygons method*; and *isohyetal method*. These methods use distinctive interpretations of the physical quantity in relation to the geometric context. For this reason, uncertainty of the common output plays a relevant role, namely, for the comparison of the accuracy of the methods.

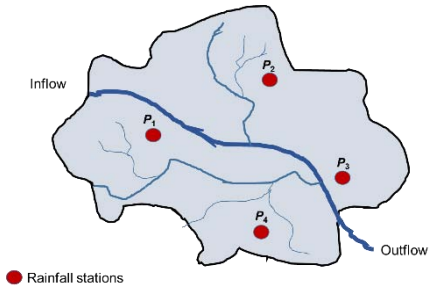


Figure 1 – Position of four rainfall stations in a basin

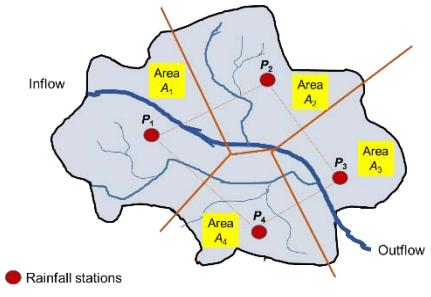


Figure 2 – Approach based on the Thiessen polygons' method

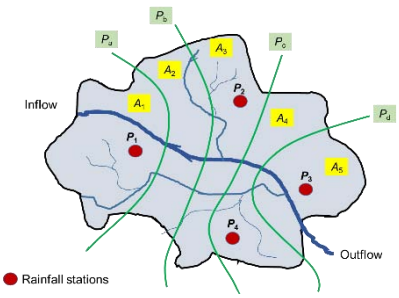


Figure 3 – Approach based on the isohyetal method

The methods consider that there are *m* measurement stations able to obtain rainfall values P_i $(i = 1, \ldots, m)$ distributed across a basin (see Figures 1 to 3). The arithmetic mean method evaluates the average without establishing a relation between the position of the stations and the geometry of the area of observation. The average areal rainfall is

Role of measurement uncertainty in the comparison of average areal rainfall methods and its impact on conformity assessment

$$P_{\mathrm{av}} = \frac{1}{m}\sum_{i=1}^{m} P_i = \frac{1}{4}(P_1 + P_2 + P_3 + P_4). \quad (1)$$

The method of Thiessen polygons evaluates the average areal rainfall as

$$P_{\mathrm{av,tp}} = \sum_{i=1}^{m} w_i \cdot P_i = \sum_{i=1}^{m} \frac{A_i}{A} \cdot P_i, \quad (2)$$

using a weighted approach to the arithmetic mean, with the weights w_i defined by the relative area of the polygons obtained using a geometric approach (see Fig. 2), A_i the area of the polygon related to the station i and A the total area of observation. A change of value of the rainfall measurement does not affect the areas considered, as the geometry is fixed.
The isohyetal method (Fig. 3) also applies weights, now determined by the contour map areas. In this case, a change in the rainfall values reshapes the contour areas and the overall estimate of the areal rainfall average, $P_{\mathrm{av,isoh}}$, will change accordingly:

$$P_{\mathrm{av,isoh}} = \sum_{i=1}^{m} w_i \cdot P_i = \frac{1}{A}\left[A_1 \cdot P_{\mathrm{a}} + A_2 \cdot \left(\frac{P_{\mathrm{a}}+P_{\mathrm{b}}}{2}\right) + A_3 \cdot \left(\frac{P_{\mathrm{b}}+P_{\mathrm{c}}}{2}\right) + A_4 \cdot P_{\mathrm{d}}\right]. \quad (3)$$

3. MEASUREMENT UNCERTAINTY EVALUATION AND FINAL REMARKS

Measurement uncertainty evaluation for the methods mentioned allows a comparison of the performance and compliance regarding the requirements of applications (from simple monitoring of environmental quantities to meteorological forecast) to be made. Considering the functional relations presented, uncertainty propagation using the original GUM [3] and the Monte Carlo approach of GUM Supplement 1 [4] allows a discussion of the results taking into consideration the different interpretation of the methods with regards to the influence given to the geometric analysis for the average areal rainfall.

4. ACKNOWLEGEMENTS

The authors would like to acknowledge the European Metrology Programme for Innovation and Research (EMPIR) developed by EURAMET, integrated in Horizon 2020, the EU Framework Programme for Research and Innovation, for the funding of project 17NRM05 EMUE (Examples of Measurement Uncertainty Evaluation), Advancing measurement uncertainty – comprehensive examples for key international standards.

5. REFERENCES

[1] Transforming our World: The 2030 Agenda for Sustainable Development. A/RES/70/1. United Nations. Following the United Nations Sustainable Development Meeting 2015, New York, 25 – 27 Sept. 2015.
[2] WMO–No.168: 2008. Guide to Hydrological Practices. Vol. 1. Hydrology – from measurement to hydrological information. 6th Ed., World Meteorological Organization, Geneva, Switzerland.
[3] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML. Evaluation of measurement data – Guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology. JCGM 100:2008.
[4] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML. 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. JCGM 101:2008.

Application of Gaussian Markov random field priors for Bayesian spatial modeling

C. Elster, J. Lehnert, S. Metzner and G. Wübbeler

Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany
E-mail (corresponding author): clemens.elster@ptb.de

Gaussian Markov Random Field (GMRF) priors [1] are a popular tool in the Bayesian inference of spatially distributed parameters whose variation is expected to be smooth. Examples of applications comprise the analysis of functional magnetic resonance imaging [2] or the inference of electron density of earth's upper atmosphere in the geosciences [3]. Applications involving GMRF priors are often high-dimensional which challenges the numerical calculation of the results of a Bayesian inference. This contribution reviews the concepts of GMRF priors and presents two examples of applications. The examples are high-dimensional and different approximation techniques are employed, including the use of Laplace-type approximations and approximate analytical expressions facilitating the sampling from the posterior.

The first example presents the Bayesian inference for magnetic resonance fingerprinting (MRF) in human brain imaging. MRF is a recent technique for quantitative magnetic resonance imaging allowing for short acquisition times [4]. From a statistical point of view, MRF constitutes a nonlinear large-scale regression problem. A hierarchical Bayesian model is applied including GMRF priors to model the smoothness of tissue properties in the brain. The theoretical properties of the posterior are explored and a Laplace approximation is applied for the numerical calculation of the results [5].

In the second example application of GMRF priors is demonstrated for the task of inferring myocardial perfusion from cardiovascular magnetic resonance (CMR). While this non-invasive technique is qualitatively applied in clinical routine [6], quantitative CMR is challenging and represents a topic of current research. The use of GMRF priors in a hierarchical Bayesian approach significantly stabilizes the quantification of perfusion at a high spatial resolution. In using approximate, analytical calculations for marginal and conditional posteriors, sampling can be achieved through applying standard MCMC in low dimensions combined with sampling from a high-dimensional truncated Gaussian. The latter can be realized efficiently using a Gibbs sampler.

[1] H. Rue and L. Held. Gaussian Markov random fields: theory and applications. CRC press, 2005.

[2] A. Brezger, L. Fahrmeir, and A. Hennerfeind. Adaptive Gaussian Markov random fields with applications in human brain mapping. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 56, no. 3, pp. 327-345, 2007.

[3] J. Norberg, L. Roininen, J. Vierinen, O. Amm, D. McKay-Bukowski, and M. Lehtinen. Ionospheric tomography in Bayesian framework with Gaussian Markov random field priors. *Radio Science*, vol. 50, no. 2, pp. 138-152, 2015.

[4] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. L. Sunshine, J. L. Duerk, and M. A. Griswold. Magnetic resonance fingerprinting. *Nature*, vol. 495, no. 7440, p. 187-193, 2013.

[5] S. Metzner, G. Wübbeler, and C. Elster. Approximate large-scale Bayesian spatial modeling with application to quantitative magnetic resonance imaging. *AStA Advances in Statistical Analysis*, online first, 2018.

[6] J. R. Panting, P. D. Gatehouse, G.-Z. Yang, F. Grothues, D. N. Firmin, P. Collins, and D. J. Pennell. Abnormal subendocardial perfusion in cardiac syndrome x detected by cardiovascular magnetic resonance imaging. *New England Journal of Medicine*, vol. 346, no. 25, pp. 1948-1953, 2002.

Targeted high-fidelity data to enrich surrogate models for uncertainty quantification in climate prediciton

O. Dunbar[1], A. Stuart[1], T. Schneider[1]

[1]California Institute of Technology, Pasadena, California, USA
E-mail (corresponding author): odunbar@caltech.edu

Climate prediction features large uncertainties due to the dependence of large-scale phenomena on fluctuations on small length scales that cannot be resolved globally. Parameterized closures are provided to approximate small scale behaviour, and local (and therefore feasible) high-resolution simulations can give information about these closure parameters. This raises the question, where should we place these high-resolution simulations, and how might we use this high-resolution data?

We consider a closure model for moist convection within an idealized aquaplanet general circulation model (GCM). We use a Bayesian design framework for inverse problems. We locate optimal latitudes based on information content of sampled posterior distribution for parameters at different designs. At these optimal latitudes, we perform a local high-resolution simulation and incorporate the different resolution data sets by training a multi-fidelity Gaussian process emulator. The Gaussian process is used as a surrogate in a sampling technique (MCMC), to obtain enriched posterior distributions for the parameters.

Institutional Support

EURAMET

EMRP
European Metrology Research Programme
Programme of EURAMET

The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union

eurolab aisbl
European Federation of National Associations of Measurement, Testing and Analytical Laboratories

spm
SOCIEDADE PORTUGUESA DE MATEMÁTICA

SPMet
Sociedade Portuguesa de Metrologia

SOCIEDADE PORTUGUESA DE QUÍMICA

http://mathmet2019.lnec.pt/