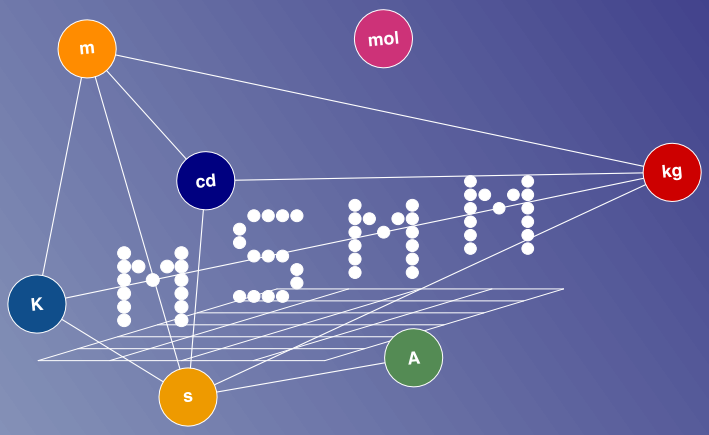


ENBIS and EMN Mathmet
Joint Workshop

Mathematical and Statistical Methods for Metrology

30-31 May 2023

http://www.msmm2023.polito.it/

INRiM
Istituto Nazionale di Ricerca Metrologica
Strada delle Cacce, 91 – Torino, Italy

Co-chairs

Francesca Pennecchi (INRiM)

Antonio Pievatolo (IMATI – CNR)

Scientific committee

João Alves e Sousa (IPQ)
Markus Baer (PTB)
Rossella Berni (Università di Firenze)
Enrico Bibbona (Politecnico di Torino)
Walter Bich (INRIM)
Alen Bosnjakovic (IMBIH)
Oriano Bottauscio (INRIM)
Maurice Cox (NPL)
Severine Demeyer (LNE)
Gianfranco Durin (INRIM)
Stephen Ellison (LGC)
Clemens Elster (PTB)
Nicolas Fischer (LNE)
Bernard Francq (GSK)
Alistair Forbes (NPL)
Fiorenzo Franceschini (Politecnico di Torino)
Maurizio Galetto (Politecnico di Torino)
Gianfranco Genta (Politecnico di Torino)
Rainer Göb (University of Wuerzburg)
Sebastian Heidenreich (PTB)
Katy Klauenberg (PTB)
Gertjan Kok (VSL)
Alessandra Manzin (INRIM)
Francesca Pennecchi, co-chair (INRIM)
Antonio Pievatolo, co-chair (IMATI – CNR)
Jacek Puchalski (GUM)
Amalia Vanacore (Università di Napoli)
Adriaan van der Veen (VSL)
Grazia Vicario (ex-Politecnico di Torino)
Luca Zilberti (INRIM)

Local organizing committee

Enrico BIBBONA (Politecnico di Torino)
Antonella BIANCHI (Politecnico di Torino)
Francesca COLLINI (Politecnico di Torino)
Eleonora CREVACORE (Politecnico di Torino)
Silvia CAVALLERO (INRIM)
Graziano COPPA (INRIM)
Elisabetta MELLI (INRIM)
Francesca DURBIANO (INRIM)

The Programme

(with hyperlink to the abstracts)

ENBIS - Mathmet

MSMM 2023

Mathematical and Statistical Methods for Metrology

PROGRAMME
30-31 MAY 2023

30 MAY - FIRST DAY

TIME (CEST)*	
08.00 - 09.00	**Registration at the INRiM** Strada delle Cacce 91, 10135 TORINO
09.00 - 09.15	**INRiM and Politecnico di Torino welcome** Room: Conference Hall **Francesca Pennecchi** (MSMM 2023 co-chair, Istituto Nazionale di Ricerca Metrologica, IT), **Pietro Asinari** (Scientific Director of the Istituto Nazionale di Ricerca Metrologica, IT) and **Enrico Bibbona** (DISMA, Politecnico di Torino, IT)
09.15 - 09.30	**ENBIS and Mathmet welcome** Room: Conference Hall **Antonio Pievatolo** (MSMM 2023 co-chair, ENBIS Past President, IMATI-CNR) and **Alistair Forbes** (ENBIS SIG on Measurement Uncertainty co-chair, National Physical Laboratory, UK), **Francesca Pennecchi** (MSMM 2023 co-chair, EMN Mathmet Chair) and **Nicolas Fischer** (EMN Mathmet Vice-Chair, Laboratoire national de metrologie et d'essais, FR)
09.30 - 10.20	**Invited speaker - Lorenzo Tamellini**, Institute of Applied Mathematics and Information Technologies "Enrico Magenes" (CNR-IMATI Pavia, IT) *A multi-fidelity method for uncertainty quantification in engineering problems* - **Chair: Antonio Pievatolo** Room: Conference Hall
10.20 - 10.40	**COFFEE BREAK**
	MATHMET Project - Chair: Nicolas Fischer Room: Conference Hall
10.40 - 11.05	**ID 122** *The strategic research agenda of the European Metrology Network Mathmet* **Sebastian Heidenreich** - Physikalisch-Technische Bundesanstalt, DE
11.05 - 11.30	**ID 126** *Mathmet Quality Assurance Tools for data, software, and guidelines* **Keith Lines** - National Physical Laboratory, UK
	Applications of Machine Learning Methods for Solving Inverse Problems **Chair: Philipp Benner** Room: Conference Hall
11.30 - 11.50	**ID 53** *Invertible neural networks for estimating electron densities from X-ray scattering measurements* **Philipp Benner** - Bundesanstalt für Materialforschung und -prüfung, DE
11.50 - 12.10	**ID 123** *Reconstructions of nano-geometries from grazing incidence X-Ray fluorescence measurements using neural networks* **Sebastian Heidenreich** - Physikalisch-Technische Bundesanstalt, DE
12.10 - 12.30	**ID 70** *Determining radius and refractive index of nanoparticles using machine learning* **Federica Gugole** - Nationaal Metrologisch Instituut, NL
12.30 - 12.50	**ID 76** *Predicting Equivalent Electrical Circuits from Electrochemical Impedance Spectroscopy (EIS) Data with Convolutional Neural Networks and Global Optimization* **Alexander Kister** - Bundesanstalt für Materialforschung und -prüfung, DE
12.50 - 14.00	**LUNCH BREAK - EXPO ROOM**

*ORAL PRESENTATION - all time listed are CEST (UTC+2)

PARALLEL SESSIONS

ROOM: CONFERENCE HALL

Design and optimisation methods
Chair: Enrico Bibbona

14.00 - 14.20
ID 58 *Bridging the Gap between Design and Metrology using Statistical Tolerance Analysis*
Mattia Maltauro - Department of Management and Engineering, University of Padova, IT

14.20 - 14.40
ID 18 *Simulated Annealing for Covariate-Adaptive Designs*
Marco Novelli - Department of Statistical Sciences, University of Bologna, IT

14.40 - 15.00
ID 81 *A procedure for optimal designs and modeling in technological processes: a case-study on freight trains*
Nedka D. Nikiforova - Department of Statistics Computer Science Applications "G. Parenti", University of Florence, IT

15.00 - 15.20
ID 55 *Improving cancer diagnosis times by optimising resource allocation*
Elizabeth A. Cooke - National Physical Laboratory, UK

Industrial applications
Chair: Gianfranco Genta

15.20 - 15.40
ID 96 *Subjective vs objective assembly complexity assessment: a comparative study in a Human Robot Collaboration framework*
Elisa Verna - DIGEP, Politecnico di Torino, IT

15.40 - 16.00
ID 155 *The thermal dynamics of a brake pad, and the estimation of its thermal parameters*
Francesca Collini - DISMA, Politecnico di Torino, IT

ROOM: SEMINAR ROOM

Conformity assessment
Chair: Francesca Pennecchi

14.00 - 14.20
ID 26 *Advanced methods for assessment of chemical compositions of multicomponent substances or materials and their categorical property values*
Ilya Kuselman - Independent Consultant on Metrology, IL

14.20 - 14.40
ID 97 *How do asymmetric measurement distributions affect risks in conformity assessment?*
Stephen L R Ellison - LGC limited, UK

Longitudinal data and time series
Chair: Marco Coïsson

14.40 - 15.00
ID 38 *Modeling Lifetime Drift of Discrete Electrical Parameters for Automotive Semiconductors*
Lukas Sommeregger - Infineon Technologies Austria AG, AT

15.00 - 15.20
ID 125 *Forecasting oxygen content in seawater*
Gianfranco Durin - Istituto Nazionale di Ricerca Metrologica, IT

Reference data
Chair: Alistar Forbes

15.20 - 15.40
ID 128 *Towards Reference Point Cloud Generation for Data Fusion in Dimensional Metrology*
Louis-Ferdinand Lafon - Laboratoire Commun de métrologie et d'essais, FR

15.40 - 16.00
ID 132 *Reference data for Electrical Resistance Tomography*
Alessandro Cultrera - Istituto Nazionale di Ricerca Metrologica, IT

16.00 - 16.30 **COFFEE BREAK & POSTER SESSION**
ROOM: EXPO ROOM

ROOM: CONFERENCE HALL

Uncertainty and regression problems
Chair: Walter Bich

16.30 - 16.50
ID 88 *On the dB-to-linear conversion*
Luca Callegaro - Istituto Nazionale di Ricerca Metrologica, IT

16.50 - 17.10
ID 121 *Data smoothing and its application to the evaluation of the measurement uncertainty in a humidity standard*
Rezvaneh Nobakht - Istituto Nazionale di Ricerca Metrologica, IT

17.10 - 17.30
ID 103 *Callendar Van Dusen interpolation by means of Piecewise Constrained Least Squares with nullspace method – an update*
Graziano Coppa - Istituto Nazionale di Ricerca Metrologica, IT

ROOM: SEMINAR ROOM

Methods for dosimetry
Chair: Stephen Ellison

16.30 - 16.50
ID 50 *Meta-analysis of dosimetry audits*
Ellie L. Smyth - National Physical Laboratory, UK

16.50 - 17.10
ID 56 *Sensitivity Analysis for Gamma Index Calculations in Dosimetry Audits for Advanced Radiotherapy*
Nadia Smith - National Physical Laboratory, UK

17.10 - 17.50
Methods for Electric Properties Tomography
Chair: Oriano Bottauscio

17.10 - 17.30
ID 68 *Repeatability and Reproducibility Uncertainty Assessment in Magnetic Resonance-based Electric Properties Tomography of a Homogeneous Phantom*
Alessandro Arduino - Istituto Nazionale di Ricerca Metrologica, IT

17.30 - 17.50
ID 84 *Electric Properties Tomography via Green's Integral Identity*
Luca Zilberti - Istituto Nazionale di Ricerca Metrologica, IT

Social dinner
Venue: Kipling Restaurant & Wines
Via Giuseppe Mazzini, 10 - 10123 Torino
19.30 - 22.00

31 MAY - SECOND DAY

TIME (CEST)*

08.30 - 09.00 — **Registration at the INRiM**
Strada delle Cacce 91, 10135 TORINO

09.00 - 09.10 — **Welcome to day 2**
Room: Conference Hall

Francesca Pennecchi (MSMM 2023 co-chair, Istituto Nazionale di Ricerca Metrologica, IT)

09.10 - 10.00 — **Invited speaker - Botond Tibor Szabó,** Bocconi University, Department of Decision Sciences (Milano, IT)
On the theoretical understanding of Bayesian methods in complex models
Chair: Francesca Pennecchi
Room: Conference Hall

Methods for Deep Learning
Chair: Sebastian Heidenreich
Room: Conference Hall

10.00 - 10.20 — **ID 112** *GUM-compliant uncertainty propagation for deep neural networks*
Björn Ludwig - Physikalisch-Technische Bundesanstalt, DE

10.20 - 10.40 — **ID 93** *Efficient learning of the copula distribution using WGANs*
Jorg Martin - Physikalisch-Technische Bundesanstalt, DE

10.40 - 11.20 — **COFFEE BREAK & POSTER SESSION**
ROOM: EXPO ROOM

PARALLEL SESSIONS

ROOM: CONFERENCE HALL

iMet-MRI Project
Chair: Nadia Smith

11.20 - 11.40 — **ID 65** *T2 or not T2? A new tool for consistent processing of qMRI parameters*
Jack D. Clarke - National Physical Laboratory, UK

11.40 - 12.00 — **ID 111** *An efficient way to generate synthetic spin echo signals by the extended phase graph*
Asante Ntata - National Physical Laboratory, UK

12.00 - 12.20 — **ID 129** *Simulation of acquisition process in Magnetic Resonance Imaging to support standardization*
Riccardo Ferrero - Istituto Nazionale di Ricerca Metrologica, IT

Bayesian methods
Chair: Gianfranco Durin

12.20 - 12.40 — **ID 108** *Separation of effects associated with measurement data*
Alistair Forbes - National Physical Laboratory, UK

12.40 - 13.00 — **ID 80** *Investigation of a Bayesian approach for the calibration of large batches of sensors*
Andrea Prato - Istituto Nazionale di Ricerca Metrologica, IT

ROOM: SEMINAR ROOM

ViDit Project
Chair: Sonja Schmelter

11.20 - 11.40 — **ID 71** *Challenges related with Virtual Experiments in Metrology*
Gertjan Kok - VSL, NL

11.40 - 12.00 — **ID 113** *Trustworthy virtual experiments and digital twins (ViDiT) – Uncertainty evaluation for Digital Twins*
Giacomo Maculotti - DIGEP, Politecnico di Torino, IT

12.00 - 12.20 — **ID 130** *Monte Carlo simulations for uncertainty estimation of error separation techniques*
Saint-Clair T. Toguem - Laboratoire national de metrologie et d'essais, FR

Modelling for engineering applications
Chair: Maurizio Galetto

12.20 - 12.40 — **ID 42** *Multilayer Delamination Model*
Kirill Ivanov - Infineon Technologies Austria AG, AT

13.00 - 14.00 — **LUNCH BREAK - EXPO ROOM**

*ORAL PRESENTATION - all time listed are CEST (UTC+2)

ROOM: CONFERENCE HALL

QUIERO Project
Chair: Luca Zilberti

14.00 - 14.20
ID 106 *Physiological variability in brain electric conductivity: correcting the effect of the age for the detection of pathological alterations*
Sebastien Marmin - Laboratoire national de métrologie et d'essais, FR

14.20 - 14.40
ID 114 *Combining experimental design with digital twin and phantom experiments to optimise data acquisition for magnetic resonance fingerprinting (MRF)*
Stephen L.R. Ellison - LGC Limited, UK

14.40 - 15.00
ID 127 *Myocardial Fibrosis Segmentation from MRF Images*
Aleksander Sadikov - Faculty of Computer and Information Science, University of Ljubljana, SI

ROOM: SEMINAR ROOM

RaCHy Project
Chair: Alessandra Manzin

14.00 - 14.20
ID 107 *A machine learning approach for the estimation of magnetic nanoparticles specific loss power*
Riccardo Ferrero - Istituto Nazionale di Ricerca Metrologica, IT

14.20 - 14.40
ID 120 *In silico experiments to investigate the heating efficiency of magnetic nanoparticles in hyperthermia preclinical tests*
Marta Vicentini - Istituto Nazionale di Ricerca Metrologica, IT

14.40 - 15.00
ID 124 *Thermo-acoustic simulation in ultrasound hyperthermia applications*
Silvia Pozzi - National Center for Radiation Protection and Computational Physics, Italian National Institute of Health, IT

PLENARY SESSION

ROOM: CONFERENCE HALL

Explainable Deep Learning
Chair: Gertjan Kok

15.00 - 15.20
ID 87 *Explainable deep learning inference to decode decision-making processes from multidimensional patterns of neural activities*
Andrea Ciardiello - "Sapienza" University of Rome, IT

15.20 - 15.40
ID 94 *From "Wich" to "Why": Interpretation map for Explainable Deep Learning based on Influence methods*
Andrea Ciardiello - "Sapienza" University of Rome, IT

15.40 - 16.00
CONCLUSIONS
Conference Hall

ONE LAST COFFEE
ROOM: EXPO ROOM

30/05
16.00 - 16.30
POSTER SESSION
Room: Expo Room

31/05
10.40 - 11.20
POSTER SESSION
Room: Expo Room

ID 28
Black-Box Uncertainty Estimation of Machine Learning Models
Georgi Tancev - Federal Institute of Metrology METAS, CH

ID 45
Quantitative analysis and processing of surfaces and profiles from profilometry images
Andrea Giura - Istituto Nazionale di Ricerca Metrologica, IT

ID 51
Ensuring the validity of measurement results through the use of triangulation rules
Iulian Mihai - Istituto Nazionale di Ricerca Metrologica, IT

ID 101
PyES - an open source software for the computation of in solution and precipitation equilibria
Lorenzo Castellino - University of Turin, IT

ID 105
Obsidian sourcing by combining SEM images and machine learning
Marco Coïsson - Istituto Nazionale di Ricerca Metrologica, IT

ID 115
ViDiT project "Trustworthy virtual experiments and digital twins"
Sonja Schmelter - Physikalisch-Technische Bundesanstalt, DE

ID 156
Employing machine learning models to enhance the prediction of cocrystals formation
Eugenio Alladio - University of Turin, IT

A multi-fidelity method for uncertainty quantification in engineering problems

Lorenzo Tamellini

Institute of Applied Mathematics and Information Technologies “Enrico Magenes”
CNR-IMATI, Pavia, Italy

Computer simulations, which are nowadays a fundamental tool in every field of science and engineering, need to be fed with parameters such as physical coefficients, initial states, geometries, etc. This information is however often plagued by uncertainty: values might be e.g. known only up to measurement errors, or be intrinsically random quantities (such as winds or rainfalls). Uncertainty Quantification (UQ) is a research field devoted to dealing efficiently with uncertainty in computations. UQ techniques typically require running simulations for several (carefully chosen) values of the uncertain input parameters (modeled as random variables/fields), and computing statistics of the outputs of the simulations (mean, variance, higher order moments, pdf, failure probabilities), to provide decision-makers with quantitative information about the reliability of the predictions. Since each simulation run typically requires solving one or more Partial Differential Equations (PDE), which can be a very expensive operation, it is easy to see how these techniques can quickly become very computationally demanding.

In recent years, multi-fidelity approaches have been devised to lessen the computational burden: these techniques explore the bulk of the variability of the outputs of the simulation by means of low-fidelity/low-cost solvers of the underlying PDEs, and then correct the results by running a limited number of high-fidelity/high-cost solvers. They also provide the user a so-called "surrogate-model" of the system response, that can be used to approximate the outputs of the system without actually running any further simulation.

In this talk we illustrate a multi-fidelity method (the so-called multi-index stochastic collocation method) and its application to a couple of engineering problems. If time allows, we will also briefly touch the issue of coming up with good probability distributions for the uncertain parameters, e.g. by Bayesian inversion techniques.

References

C. Piazzola, L. Tamellini, R. Pellegrini, R. Broglia, A. Serani, and M. Diez. Comparing Multi-Index Stochastic Collocation and Multi-Fidelity Stochastic Radial Basis Functions for Forward Uncertainty Quantification of Ship Resistance. Engineering with Computers, 2022

J. Beck, L. Tamellini, and R. Tempone. IGA-based Multi-Index Stochastic Collocation for random PDEs on arbitrary domains. Computer Methods in Applied Mechanics and Engineering, 2019

On the theoretical understanding of Bayesian methods in complex models

Botond Tibor Szabó

Bocconi University, Department of Decision Sciences
Milano, Italy

Bayesian methods are becoming increasingly popular in various fields of sciences. They offer a principled way to incorporate expert knowledge into the statistical model and provide built in uncertainty quantification (i.e quantifying the remaining uncertainty in the statistical procedure). To deal with the ever increasing amount of available information and increasingly more complex models new approximation methods (e.g. parallel computing, variational approximations) are being developed to speed up the computations and reduce the memory requirement. I will demonstrate the wide applicability of Bayesian methods over several concrete examples ranging from epidemiology through astronomy.

However, Bayesian methods are subjective by nature (by the choice of the prior) and inaccurate use can result in misleading interpretation and contradictory conclusions.
To better understand their behaviour and achieve acceptance by a wider scientific community their frequentist properties have to be understood. The main focus of research is on understanding whether Bayesian methods can recover the underlying parameters of interest (with an optimal rate) as the sample size increases (called Bayesian consistency), and whether Bayesian uncertainty quantification provides reliable confidence statement (i.e. frequentist coverage). I will briefly discuss the state of the art literature in this field and introduce some standard techniques achieving such guarantees.

References

[1] Ghoshal, Ghosh, van der Vaart (2000) Convergence rates of posterior distributions. Annals of Statistics
[2] Szabo, van der Vaart, van Zanten (2015) Frequentist coverage of adaptive nonparametric Bayesian credible sets. Annals of Statistics (Discussion paper)
[3] Szabo, van Zanten (2019) An asymptotic analysis of distributed nonparametric methods. Journal of Machine Learning Research.

Simulated Annealing for Covariate-Adaptive Designs

Marco Novelli, Alessandro Baldi Antognini and Maroussa Zagoraiou

Key words: Balancing Covariates, Loss of information, Treatment Comparison

Abstract

In comparative experiments, balancing the covariates across experimental groups is a fundamental requirement to perform reliable inference about the treatment effects [Rosenberger and Sverdlov(2008)]. In this regard, several procedures have been suggested in the literature. However, most of the proposed methods are applicable only with categorical factors, while quantitative variables are generally either ignored or discretized, thus affecting the inferential accuracy. The few exceptions (see for example [Atkinson(1982)]) show performances that are strongly related to both the correctness of the model specification and its complexity [Baldi Antognini et al.(2023)]. In the last decade, thanks to the recent advances in the biomarkers-based personalized medicine, it has become increasingly common to include several covariates and their interactions in the analysis [Karczewski and Snyder(2018)]. Nevertheless, an efficient Covariate-Adaptive (CA) procedure able to deal with mixed covariates profile with potentially complex interaction structure is still missing.

In this work we discuss a new class of CA designs based on the Simulated Annealing (SA) algorithm recently proposed in [Baldi Antognini et al.(2023)]. Originally suggested by [Metropolis et al.(1953)] in the field of statistical mechanics, SA is a stochastic local search algorithm which has been employed to approximate

Marco Novelli
Department of Statistical Sciences, University of Bologna, e-mail: m.novelli@unibo.it

Alessandro Baldi Antognini
Department of Statistical Sciences, University of Bologna, e-mail: a.baldi@unibo.it

Maroussa Zagoraiou
Department of Statistical Sciences, University of Bologna, e-mail: maroussa.zagoraiou@unibo.it

global optimization solutions for large search spaces. It comes from the physical process of the annealing of metals by gradual cooling: at high temperatures, the particles are rather free to move, leaving the structure subject to substantial changes, while as the temperature gradually decreases, the probability that a particle will move decreases accordingly, until the system reaches a steady state.

Here, the SA algorithm is exploited to control covariate imbalance and a new CA procedure, the Simulated Annealing Design (SADe), is presented. Such a design is very flexible since:

- it can deal with continuous and/or categorical variables,
- it allows the adoption of any specific measure of covariate imbalance,
- it can be applied to both fixed (i.e., non-sequential) experiments, where all the covariate information is available before the trial begins, and sequential ones in which statistical units enter the trial sequentially (where at each step, only a partial information about the covariates is available).

Due to the nature of the SA algorithm, SADe is intrinsically randomized and completely unpredictable, thus avoiding any possible selection bias. Moreover, it turns out to be particularly effective also in the case of small sample sizes and a large number of covariates. The finite sample properties of the SADe are further explored through an extensive simulation study, exhibiting a remarkable improvement in the ability to balance the experimental groups as well as in the inferential accuracy with respect to all the other procedures proposed in the literature.

References

[Atkinson(1982)] Atkinson AC (1982) Optimum biased coin designs for sequential clinical trials with prognostic factors. Biometrika 69, 61–67.
[Baldi Antognini et al.(2023)] Baldi Antognini A, Novelli M and Zagoraiou M (2023) Simulated Annealing for Balancing Covariates. Statistics in Medicine, to appear DOI:10.1002/sim.9672.
[Karczewski and Snyder(2018)] Karczewski KJ and Snyder MP(2018) Integrative omics for health and disease. Nature Reviews Genetics 19.5, 299.
[Metropolis et al.(1953)] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH and Teller E(1953) Equation of state calculations by fast computing machines. The journal of chemical physics 21.6, 1087–1092 .
[Rosenberger and Sverdlov(2008)] Rosenberger WF and Sverdlov O(2008) Handling covariates in the design of clinical trials. Statistical Science, 23, 404–419.

Advanced methods for assessment of chemical compositions of multicomponent substances or materials and their categorical property values

Ilya Kuselman[1], Francesca R. Pennecchi[2], Tamar Gadrich[3] and D. Brynn Hibbert[4]

Key words: Chemical composition, Conformity assessment, False decisions, Categorical property values, Comparison

1. Conformity Assessment of a Chemical Composition

A Bayesian methodology for the evaluation of risks of false decisions on conformity of a multicomponent substance, material or object, due to measurement uncertainty was developed and published as the IUPAC/CITAC Guide [Kuselman et al.(2021a)]. In continuation of this development, another IUPAC/CITAC Guide, taking into account a mass balance constraint of the object composition [Kuselman et. al.(2019)], is now under preparation for publication.

The mass balance constraint means, according to the law of conservation of mass, that the sum of the actual ('true') values of the component contents under conformity assessment is equal to 100 % (or 1 when expressed as fractions). At the same time, the sum of measured contents can differ from 100 % (or 1) because of measurement uncertainty. As a consequence, the actual component contents are intrinsically correlated. This correlation influences the evaluation of risks of false decisions in addition to effects of possible contributions of metrological, native or technological correlations of the data. The Guide will be helpful for a range of applications in analytical chemistry and metrology.

Different scenarios of risks of false decisions due to measurement uncertainty at the mass balance constraint are considered in the following examples of conformity assessment of chemical compositions of:

- a platinum-rhodium alloy produced by a manufacturer in a two-year period [Pennecchi et. al. (2020)];
- a batch of potassium iodate as a candidate reference material of given purity [Pennecchi et. al.(2021a)];
- sausage "Braunschweigskaya" from two manufacturers produced over three years [Pennecchi et. al.(2021b)];
- synthetic air prepared by NMIs as 'zero or balance gas' in calibration mixtures for Key Comparisons, and also by an industrial manufacturer for medicinal purposes [Pennecchi et. al.(2022)].

2. Assessment of Categorical Property Values

[1] Independent Consultant on Metrology, Israel, ilya.kuselman@bezeqint.net

[2] Istituto Nazionale di Ricerca Metrologica (INRIM), Italy, f.pennecchi@inrim.it

[3]Braude College of Engineering, Karmiel, Israel, tamarg@braude.ac.il

[4] School of Chemistry, UNSW Sydney, Australia, b.hibbert@unsw.edu.au

Application of a newly-developed decomposition of total variation in two-way categorical analysis of variation of nominal values (CATANOVA) [Gadrich et al.(2020)], and two-way analysis of variation of ordinal values (ORDANOVA) [Gadrich & Marmor(2021)] for interlaboratory or similar comparisons of examination/test results, has been targeted in the next IUPAC project [Kuselman et al.(2021b)]. Two-way CATANOVA has been applied for interlaboratory examination of weld imperfections [Gadrich et al.(2020)]. A case study of expert responses of 45 ecological laboratories to intensity of odor and taste of drinking water was an example of the application of ORDANOVA [Gadrich et al.(2022a)]. Another example was a study of sensory responses of experts to the quality properties of samples of market-purchased sausage "Moscowskaya" manufactured by 16 producers. Dependence of the probabilities of classification of five quality properties to one of the categories as a function of the chemical composition of the sausage was assessed using multinomial ordered logistic regression [Gadrich et al.(2022b)].

References

1 Gadrich T., Kuselman, I., Andrić, I.: Macroscopic examination of welds - Interlaboratory comparison of nominal data. SN Appl. Sci. (2020) 10.1007/s42452-020-03907-4

2 Gadrich T. and Marmor Y.N.: Two-way ORDANOVA - analyzing ordinal variation in a cross-balanced design. J. Stat. Plan. Inference (2021) doi: 10.1016/j.jspi.2021.04.005

3 Gadrich, T., Kuselman, I., Pennecchi F. R., Semenova, A.A., Chew, P.S., Naidenko, V.N.: Interlaboratory comparison of the intensity of drinking water odor and taste by two-way ordinal analysis of variation. J. Water Health (2022a) doi: 10.2166/wh.2022.060

4 Gadrich, T., Kuselman, I., Pennecchi F.R., Hibbert, D.B., Semenova, A.A., Chew, P.S.: Ordinal analysis of sensory responses in combination with multinomial ordered logistic regression vs. chemical composition. J. Food Qual. (2022b) doi: 10.1155/2022/4181460

5 Kuselman, I., Pennecchi, F.R., Hibbert, D.B., Di Rocco, A., Botha, A.: Influence of mass balance constraint on uncertainty of test results of a substance or material and risks in its conformity assessment. IUPAC Project (2019) https://iupac.org/project/2019-012-1-500

6 Kuselman, I., Pennecchi, F.R., da Silva, R.J.N.B., Hibbert, D.B.: IUPAC/CITAC Guide: Evaluation of risks of false decisions in conformity assessment of a multicomponent material or object (IUPAC Technical Report). Pure Appl. Chem. (2021a) doi: 10.1515/pac-2019-0906

7 Kuselman, I., Gadrich, T., Pennecchi, F.R., Hibbert, D.B., Semenova, A.A., Chew, P.S.: Harmonization of approaches to comparison of qualitative and related property values of a substance or material. IUPAC Project (2021b) https://iupac.org/project/2021-017-2-500/

8 Pennecchi, F.R., Di Rocco, A., Kuselman, I., Hibbert, D.B., Sega, M.: Correlation of test results and influence of a mass balance constraint on risks in conformity assessment of a substance or material. Measurement (2020) doi: 10.1016/j.measurement.2020.107947

9 Pennecchi, F.R., Kuselman, I., Di Rocco, A., Hibbert, D.B., Sobina, A., Sobina, E.: Specific risks of false decisions in conformity assessment of a substance or material with a mass balance constraint. Measurement (2021a) doi: 10.1016/j.measurement.2020.108662

10 Pennecchi, F.R., Kuselman, I., Di Rocco, A., Hibbert, D.B., Semenova, A.A.: Risks in a sausage conformity assessment due to measurement uncertainty, correlation and mass balance constraint. Food Control (2021b) doi: 10.1016/j.foodcont.2021.107949

11 Pennecchi, F.R., Kuselman, I., Hibbert, D.B., Sega, M., Rolle, F., Altshul V.: Fit-for-purpose risks in conformity assessment of a substance or material – A case study of synthetic air. Measurement (2022) doi: 10.1016/j.measurement.2021.110542

Black-Box Uncertainty Estimation of Machine Learning Models

Georgi Tancev

Key words: bootstrap, machine learning, Monte Carlo, uncertainty propagation

1 Introduction

Estimating the uncertainty of measurements is a routine task in measurement science. The expanded uncertainty U of a model $Y = m_\theta(X) + \epsilon$ (with variables X and Y) typically contains three terms categorized into aleatoric and epistemic uncertainty: the uncertainty of the input/output realizations x and y, as well as the model parameters θ [5].

$$U = 2 \times [\underbrace{(\frac{\partial m_\theta}{\partial x} u_x)^2 + \sigma_\epsilon^2}_{\text{aleatoric}} + \underbrace{(\frac{\partial m_\theta}{\partial \theta} u_\theta)^2}_{\text{epistemic}}]^{\frac{1}{2}} \tag{1}$$

The increasing use of machine learning models such as neural networks [6], random forests [2], or Gaussian processes [3] may require dedicated methods for uncertainty propagation. For parametric models, Eq. 1 can be computed analytically. If derivatives are not available (e.g., in nearest neighbors), Monte Carlo simulations [1] can be performed instead. While such random trials allow to estimate the effect uncertain inputs or outputs, they do not necessarily take into account the uncertainty (i.e., variability) of the model. This work proposes bootstrapping [4] combined with Monte Carlo simulations as a black-box (i.e., plug-in) method to obtain an empirical estimate of the uncertainty distribution of the model output (Fig. 1).

Georgi Tancev
METAS, Lindenweg 50, 3003 Bern, e-mail: georgi.tancev@metas.ch

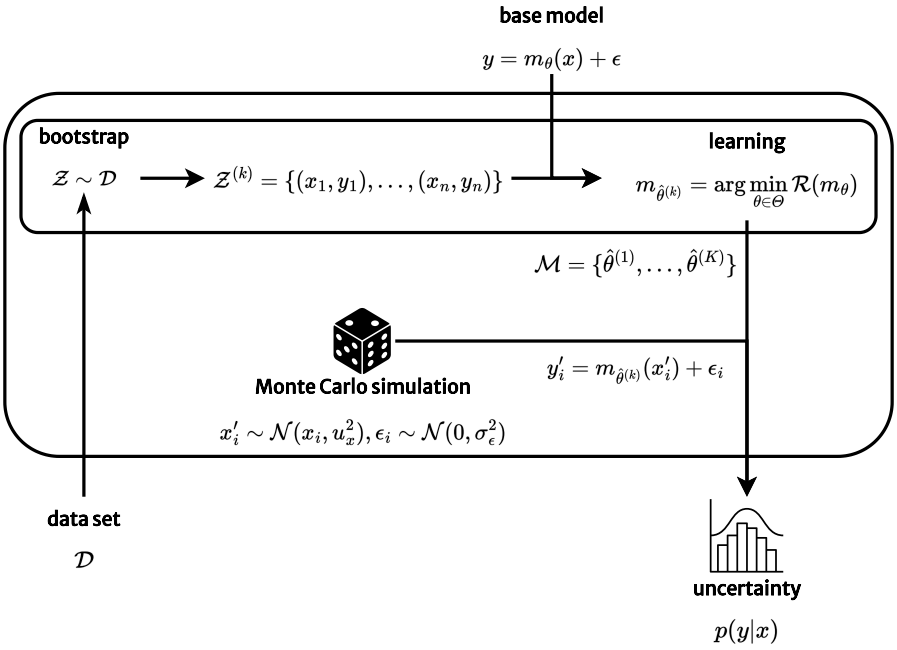


Fig. 1: Visualization of the approach.

References

[1] Daniel R. Albert. "Monte Carlo Uncertainty Propagation with the NIST Uncertainty Machine". In: *Journal of Chemical Education* 97.5 (2020), pp. 1491–1494. ISSN: 19381328. DOI: 10.1021/acs.jchemed.0c00096.

[2] Leo Breiman. "Random Forests". In: *Machine Learning* 45 (2001), pp. 5–32. DOI: 10.1023/A:1010933404324.

[3] Andrew Gelman et al. *Bayesian Data Analysis*. 3rd ed. CRC Press, 2013, p. 677. ISBN: 978-1439840955.

[4] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. 2nd ed. Springer, 2009.

[5] International Organization for Standardization. *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. Tech. rep. Geneva: International Organization for Standardization, 2008. URL: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6.

[6] Yann Lecun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". In: *Nature* 521.7553 (2015), pp. 436–444. ISSN: 14764687. DOI: 10.1038/nature14539.

Modeling Lifetime Drift of Discrete Electrical Parameters for Automotive Semiconductors

Lukas Sommeregger and Horst Lewitschnig

Key words: accelerated stress test, discrete data, lifetime drift, measurement error, statistical model.

Abstract

Autonomous driving technology presents new challenges in the field of semiconductor manufacturing for the automotive industry. Longer-than-before usage times and longer life cycles lead to increased requirements in guaranteeing quality targets for the whole lifetime of the chips.

To simulate these lifetime behaviors, accelerated stress tests are used which simulate the lifetime of the device under harsher-than-usual conditions. A sample of devices is stressed, electrical parameters are measured, then the device is stressed again and so on until the desired lifetime is simulated. The measured electrical parameters are retrieved as longitudinal data. As defined in their data sheets, electrical parameters have to stay within their specified limits over lifetime. The drift of these parameters is an indication of degradation of semiconductor devices. Statistical models are needed to assess the lifetime drift and guarantee the quality targets, e.g., 1 or 10 ppm (one part per million) allowed maximum exceedance of the limits.

Based on previous works, [1] and [2], we present a method to adapt a model for continuous parameters to the case of discretized continuous parameters. This typically takes place in the case of limited tester resolution or analog-to-digital con-

Horst Lewitschnig
Infineon Technologies Austria AG, Siemensstraße 2, 9500 Villach e-mail: Horst.Lewitschnig@infineon.com

Lukas Sommeregger
Infineon Technologies Austria AG, Siemensstraße 2, 9500 Villach e-mail: Lukas.Sommeregger@infineon.com

version of signals. It introduces additional errors in the data. The idea is based on a correction comparable to the correction of the gauge repeatability and reproducibility (GR&R) error. Furthermore, data are corrected for tester offsets.

The case of a combination of measurement uncertainty and discretization error is simulated and discussed.

Furthermore, we give an example of how to use the model to efficiently calculate tighter-than-usual guard bands used at production testing to guarantee the quality of shipped parts to customers. The issue is formulated as an optimization problem and solved under assumptions of linear changes in the data at time points between readouts.

Acknowledgements The work has been performed in the project ArchitectECA2030 under grant agreement No 877539. The project is co-funded by grants from Germany, Netherlands, Czech Republic, Austria, Norway and Electronic Component Systems for European Leadership Joint Undertaking (ECSEL JU). All ArchitectECA2030 related communication reflects only the author's view and ECSEL JU and the Commission are not responsible for any use that may be made of the information it contains.

References

1. V. Hofer, Johannes Leitner, Horst Lewitschnig, Thomas Nowak, "Determination of tolerance limits for the reliability of semiconductor devices using longitudinal data," Quality and Reliability Engineering International., vol. 33, pp. 2673-2683, 2017.
2. H. Lewitschnig and L. Sommeregger., "Quality Control of Lifetime Drift Effects," Microelectronics Reliability, Volume 139, December 2022, 114776

Multilayer Delamination Model

Horst Lewitschnig, Kirill Ivanov, and Renato Podvratnik

Key words: beta distribution, data simulation, delamination, multivariate distributions, parameter estimation.

1 Abstract

There has been a significant demand for stacked-die technologies during the past few years. Delamination is one of the typical failure modes of stacked dies. In Lin et al (2003), authors define delamination as a separation of layers within a molded part. Mold compound, imide, die attach, plasma cleaning, or design factors can attribute to delamination.

Delamination is usually modelled by the beta distribution because it can be up to one hundred percent, and the beta distribution is a bounded continuous probability distribution defined on the interval from zero to one. In order to describe delamination at several levels at the same time, a mathematical model is needed. In this case, a multivariate beta distribution model comes into use.

One of the most intuitive ways to construct a multivariate beta distribution model is through the relationship between beta and gamma functions. In Olkin et al (2003), authors were the first to use this relationship in order to construct a bivariate beta

Horst Lewitschnig
Infineon Technologies Austria AG, Siemensstraße 2, 9500 Villach e-mail: Horst.Lewitschnig@infineon.com

Kirill Ivanov
Infineon Technologies Austria AG, Siemensstraße 2, 9500 Villach e-mail: Ivanov.Kirill@infineon.com,

Renato Podvratnik
Infineon Technologies Austria AG, Siemensstraße 2, 9500 Villach e-mail: Renato.Podvratnik@infineon.com

distribution. This construction is extendable up to n dimensions by adding independent gamma random variables with common shape parameter (Arnold and Tony (2011)). Scalability and flexibility are the advantages of this construction, while large number of parameters and the absence of density equations for some parameter combinations are the disadvantages.

Another way to construct a multivariate beta distribution is by using the Gaussian copula. Straightforward parameter estimation and the possibility to specify the desired correlation matrix are the advantages of this construction, while only linear dependencies and additional degrees of freedom in choosing the copula are the disadvantages.

The last construction we investigate was proposed in Ferrari and Cribari-Neto (2004), where authors used the beta regression to build a multivariate beta distribution model. High flexibility and non-linear dependencies are the advantages of this model, while scalability and the proper choice of a link function are the challenges.

We simulate the data using the proposed models, where we vary initial model parameters, sample size, and the number of dimensions. Then, we simulate the model parameters based on the simulated data. In real measurements, delamination is always measured with an uncertainty. This, in turn, affects the accuracy of the estimated parameters. We reflect this measurement uncertainty in the simulated data as well.

We assess the fit of the models using distance measures such as Kolmogorov-Smirnov statistic, Kullbach-Leibler divergence, and Hellinger distance. The results are based solely on the simulated data.

Acknowledgements The work has been performed in the project ArchitectECA2030 under grant agreement No 877539. The project is co-funded by grants from Germany, Netherlands, Czech Republic, Austria, Norway and Electronic Component Systems for European Leadership Joint Undertaking (ECSEL JU). All ArchitectECA2030 related communication reflects only the author's view and ECSEL JU and the Commission are not responsible for any use that may be made of the information it contains.

References

Arnold B, Tony HK (2011) Flexible bivariate beta distributions. Journal of Multivariate Analysis 102(8):1194–1202, DOI https://doi.org/10.1016/j.jmva.2011.04.001

Ferrari S, Cribari-Neto F (2004) Beta regression for modelling rates and proportions. Journal of Applied Statistics 31(7):799–815, DOI 10.1080/0266476042000214501

Lin TY, Xiong ZP, Yao YF, Tok L, Yu ZY, Njoman B, Chua KH, Y MY (2003) Failure analysis of full delamination on the stacked die leaded packages. Journal of Electronic Packaging 125(3):392–399, DOI 10.1115/1.1602704

Olkin I, Liu R, Stuff K, SA K (2003) A bivariate beta distribution. Statistics & Probability Letters 62(4):407–412

Quantitative analysis and processing of surfaces and profiles from profilometry images

A. Giura[1*], M. Zucco[1], A. Balsamo[1], L. Ribotta[1]

[1] Division of Applied metrology and Engineering - Istituto Nazionale di Ricerca Metrologica (INRiM), Strada delle Cacce 91, 10135, Torino, Italy

a.giura@inrim.it

Key words: Profilometry, texture, roughness, parameters, Python, topography, levelling, resampling, filtering

1. Introduction

Surface metrology is concerned with inspecting morphological parameters of a surfaces or profiles, by using contact or non-contact profilometers.

The following abstract describes the development of a software in Python environment that implements various processing methods on images from optical and stylus profilometers. In particular, the program focusses on image pre-processing and determination of dimensional parameters for 2D areas and 1D profiles.

It is worth mentioning that many open and closed source programs are already distributed, but they do not provide a sufficient automatization in the image processing, often requiring the user to repeat the same steps for each image to obtain the expected results.

The program has been initially developed within the framework of the EMPIR 20IND07 TracOptic project [1] for the processing of a batch of topographies on RS-M and RS-N linear step samples, in order to compensate for the lack of automation for the calculation of height parameters. The developed program [2] is designed to be modular and scalable for expanding the processing capabilities.

1.1 Surface processing

Surface measurements are usually obtained with optical profilometers or microscopes, and the resulting topographies must be processed to extract the parameters of interest.

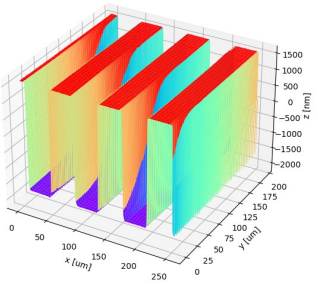


Figure 1: Topography

For the pre-processing of the images, three different levelling methods have been implemented which allow to correct any tilt of the topography: least square plane levelling, three points plane levelling and bounded least square plane levelling. The program also includes a method for resampling the image and for extracting cross-sectional profiles of the topography.

1.2 Profile processing

Profile measurements are usually obtained with stylus profilometers, confocal point sensors or by extracting a cross-sectional profile from a surface.

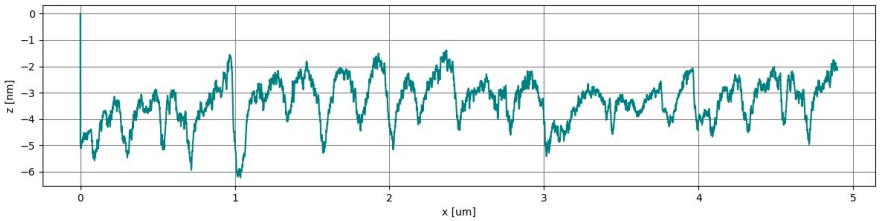


Figure 2: Profile

In order to pre-process the profile, least square line levelling and bounded least square line levelling were implemented as for the topography. An additional histogram levelling method has been developed to allow the tilt correction of profiles with a large feature on a flat baseline.

The program provides Gaussian filters according to the ISO 16610-21:2011 written standard [3] for the extraction of the morphological roughness parameters, and the erosion morphological filter according to the new ISO 21920-2:2021 written standard [4].

2. Conclusions

The methods previously described were validated using MountainsMap 7.4 [5] commercial software.
Regarding profile roughness parameters, an agreement within 0,05% for Ra, 0,2% for Rq, 0,5% for Rsk and Rku, below 1 ‰ for Rt, is achieved on different types of profile from stylus profilometers measurements.

Future developments will focus on uncertainty estimation of morphological parameters both using GUM and Monte Carlo methods. Moreover, the capabilities of the software will be expanded to allow the extraction of more complex surface features.

Acknowledgements

The 20IND07 TracOptic 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] https://www.ptb.de/empir2021/tracoptic/home/
[2] https://github.com/andeledea/pluopen
[3] EN ISO 16610-21:2012 Geometrical product specifications (GPS) - Filtration - Part 21: Linear profile filters: Gaussian filters (ISO 16610-21:2011)
[4] EN ISO 21920-2:2022 Geometrical product specifications (GPS) - Surface texture: Profile - Part 2: Terms, definitions and surface texture parameters (ISO 21920-2:2021, Corrected version 2022-06)
[5] https://www.digitalsurf.com/software-solutions/profilometry/

Meta-analysis of dosimetry audits

E. L. Smyth[1], S. A. Thomas[1], E. A. Cooke[1], M. Hussein[1], C. H. Clark[1,2] and N. A. S. Smith[1]

Key words: dosimetry, radiotherapy, meta-analysis, summary statistics

In the current digital age, data from measurements are collected routinely for a wide range of applications. Curating these data in structured databases can greatly enhance data access, searchability and reuse through data queries. Storing digital information in standardised and searchable formats is crucial for measurement science and the metrology community. Additionally, standardised digital records enable meta-analysis of the data which may reveal unknown patterns, clusters or correlations improving our understanding of the measurements and their application domains. This is critical in complex areas such as in healthcare, and in particular in those areas related to clinical decision making, such as in radiotherapy planning.

Here we curate data from a range of advanced radiotherapy dosimetry audits conducted at hospitals across the UK using test objects developed at the National Physical Laboratory. The audits contain several predictions and measurements, as well as information regarding the type of equipment used at the hospital. We use this curated data to perform meta-analysis. Specifically, we look at data from lung Stereotactic Ablative Body Radiotherapy (SABR) dosimetry audits [Distefano(2017)] to assess the positional and dosimetric accuracy of a SABR treatment delivery.

We perform meta-analysis [Hunter(2004)] of the audit data using both traditional statistical methods and data-driven methods. Fig. 1 shows a visualisation of the predicted and measured doses which, along with summary statistics, allow hospitals and auditors to be able to easily visualise and compare audit data from different sites and using different equipment. The predicted value is calculated by each respective hospital using their radiotherapy treatment planning system, which is specialised software that simulates the radiotherapy dose deposition. Highlighting individual treatment centres (orange circles, Fig. 1) allows us to visualise the spread of data from a given centre and detect points which are outside the expected confidence intervals. These results may be used to provide insight to the hospitals by giving a quantitative comparison between them and other sites using similar equipment. In addition, these visualisations are useful for auditors as they can highlight outliers and help to identify previously unknown patterns or dependencies in the data that need to be investigated further.

We also use data-driven machine learning techniques to search for hidden patterns within the audit data and use feature importance to determine the key factors influencing the outcome of radiotherapy audits. Fig. 2 shows results from a random forest model [Cutler(2011)] used to predict the measured dosage based on features such as the equipment used, temperature and pressure in the chamber, and the energy at which the dose was delivered. The model is able to predict the

[1]National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, UK
[2]University College London NHS Foundation Trust, London, NW1 2PG, UK
E-mail (corresponding author): ellie.smyth@npl.co.uk

measured dosage with a RMS error of 0.654, although we caution that this is using a small sample size. The results from this type of analysis can give evidence that certain conditions of the audit may be influencing the outcome of the audit, allowing auditors and hospitals greater insight when planning future treatments.

The results from our meta-analysis and visualisation will inform future work as the curated database grows and expands to additional audit types.

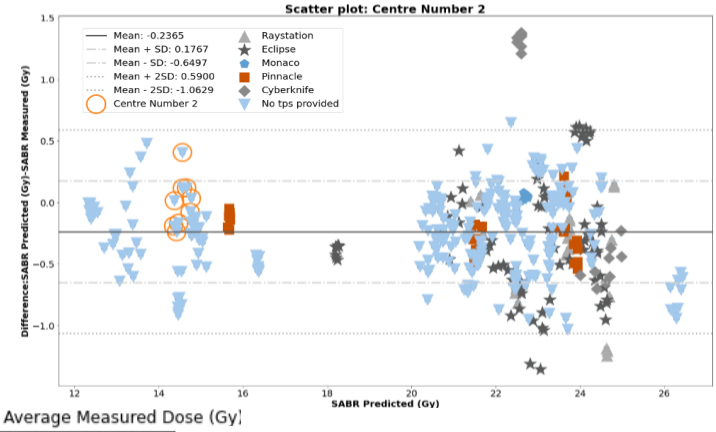


Figure 1: *Scatter plot of the measured and predicted doses (Gy). Each point is a measurement, and the orange circles represent a specific centre/hospital. The different symbols show different equipment used for treatment planning.*

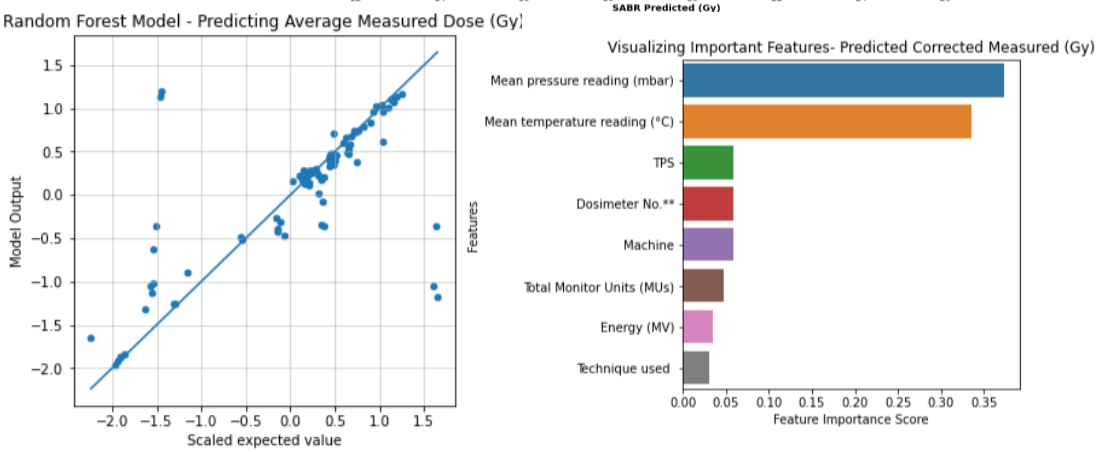


Figure 2: *Results from a random forest model predicting measured dosage for lung SABR audits. Left: Scatter plot showing the model predictions. Each point is an individual result, and the line shows a 1:1 correlation for comparison. The model is able to predict the measured dosage with a RMS error = 0.654. Right: Bar plot showing the feature importance for the model. Higher scores indicate that the model is more dependent on those features when predicting the output values.*

References

[Cutler(2011)]A. Cutler, D. Cutler, J. Stevens, *Random Forests*, Machine Learning, **45,** 157-176, (2011)

[Distefano(2017)]G. Distefano, J. Lee, J. Shakardokht, C. Gouldstone, C. Baker, H. Mayles, C. Clark, Catharine, *A national dosimetry audit for stereotactic ablative radiotherapy in lung*. Radiotherapy and Oncology, **122**, (2017)

[Hunter(2004)],J. Hunter, F. Schmidt, *Methods of Meta-Analysis Corrected Error and Bias in Research Findings*, Educational Researcher, **20**, (2004)

Ensuring the validity of measurement results through the use of triangulation rules

Iulian Mihai and Marcus Vinícius Viegas Pinto

Key words: Uncertainty, Triangulation rules, Validity

1 Introduction

Measurement results are critical in various industries, including healthcare, aerospace, and manufacturing. Inaccurate measurements can lead to severe consequences, such as faulty medical diagnoses, airplane crashes, and defective products. Therefore, it is essential to ensure the validity of measurement results to maintain the integrity and reliability of measurements.

The ISO/IEC 17025 standard provides guidelines for laboratories to ensure the validity of measurement results. This standard specifies requirements for the competence, impartiality, and consistent operation of laboratories [ISO/CASCO(2017)]. It outlines the procedures for testing, calibration, and sampling that laboratories must follow to produce reliable measurement results.

To comply with the ISO/IEC 17025 standard, laboratories must demonstrate the validity of their measurement results. They can achieve this by using the triangulation rules, which involves using multiple methods or instruments to measure the same quantity. By comparing the results of multiple measurements, laboratories can identify any discrepancies.

Iulian Mihai
Istituto Nacionale de Ricerca Metrologica - INRiM, Strada delle Cacce, 91 10135 Torino, Italy, e-mail: i.mihai@inrim.it

Marcus Vinícius Viegas Pinto
Instituto Nacional de Metrologia, Qualidade e Tecnologia - Inmetro , Av. Nossa Senhora das Graças, 50 - Vila Operária - Duque de Caxias - Rio de Janeiro (RJ), Brazil, e-mail: mvviegas@inmetro.gov.br

2 The triangulation rules

The proposed triangulation tool consists of performing the calibration of a certain equipment using different traceability chains related to each other. The proposed tool is indicated, for example, for measurements involving ratios, such as the calibration of high standard resistors by means of a Dual Source High Resistance Ratio Bridge.

The triangulation tool can help laboratories guarantee the validity of their measurement results by providing a systematic approach to calibration. The tool can also help laboratories identify and correct measurement errors and improve the accuracy and precision of their measurements.

To demonstrate the robustness of the presented tool, several tests were conducted to evaluate its performance. The tests involved simulations using Monte Carlo method. The results showed that the triangulation tool was suitable for detecting inconsistencies in the measurement and therefore for ensuring the validity of the results.

This tool can be used as a statistical technique to compare results obtained by different methods, different standards, or a combination thereof, and is used in conjunction with monitoring results. The methods used to guarantee the validity of the results, such as retesting or recalibration of retained items, use of calibrated alternative instrumentation to provide traceable results, intralaboratory comparisons, replicated tests or calibrations, using the same methods or different methods, among others, should have the application of statistical techniques for analysis whenever practicable [ISO/CASCO(2017)].

An application example of the triangulation tool includes the calibration of high value resistors [Mihai(2022b), Mihai(2022a)], where the limited availability of measurement methods prevents a comparison using different methods. In this case, the tool provides a possibility to increase confidence in the functioning of the measurement system even if it is not possible to use different methods. Next, the theoretical background and real examples carried out in the laboratory and tests, even as simulations are shown to demonstrate the use of the proposed tool.

3 Examples and performed tests

When determining a resistance value using different ratios, one can use resistors of different resistance values and compare them all. In this way, there are different traceability chains related to each other. You can then compare three resistors of different values $A = 1$, $B = 1$, and $C = 10$, where the ratios between each pair would be $R_{ab} = 1$, $R_{ac} = 0.1$ and $R_{cb} = 10$, and $R_{ab} = R_{cb} \times R_{ac}$. Each ratio has its associated uncertainty. In the case where only one pair is analyzed, if the values are incompatible, that is, with the difference greater than the limit, there is no indication of which of the two resistors could have been the origin of the problem. When analyzing the different pairs, one can indicate which resistor has a problem or which

measurement needs to be revised, which helps the user in detecting discrepancies, and provides an indication of the validity of the measurement results.

4 Conclusion

In conclusion, ensuring the validity of measurement results is critical for maintaining the integrity and reliability of measurements. Compliance with the ISO/IEC 17025 standard is essential for laboratories to achieve this goal. The presented triangulation tool is a reliable and robust approach to testing and calibration that can help laboratories guarantee the validity of their measurement results.

References

[ISO/CASCO(2017)] ISO/CASCO (2017) General requirements for the competence of testing and calibration laboratories. Standard ISO/IEC 17025:2017, International Organization for Standardization, Geneva, CH, URL https://www.iso.org/standard/66912.html

[Mihai(2022a)] Mihai I (2022a) Analyses for equivalent ratio model in measurements of high standard resistance bridges. Tech. rep., INRIM - Istituto Nazionale di Ricerca Metrologica, DOI 10.13140/RG.2.2.33636.65923, URL https://rgdoi.net/10.13140/RG.2.2.33636.65923

[Mihai(2022b)] Mihai I (2022b) Metrological triangulation rules in ratio measurements of high standard resistance bridges. Tech. rep., INRIM - Istituto Nazionale di Ricerca Metrologica, DOI 10.13140/RG.2.2.12240.58888, URL https://rgdoi.net/10.13140/RG.2.2.12240.58888

Invertible neural networks for estimating electron densities from X-ray scattering measurements

Philipp Benner and Sofya Laskina and Brian Pauw

Key words: Inverse Problems, Invertible Neural Networks, SAXS

1 Abstract

Continuing progress in the field of X-ray scattering empowers scientists with new possibilities to capture the 3D electron density of materials and molecules. Although first methods appeared almost a century ago, recovering the density structure of a sample is still challenging. Scattering techniques measure the intensity of scattered waves, for which a very accurate mathematical description exists based on the Fourier transform of the electron density. However, the resulting measurements capture not all information about the sample. Firstly, instead of a 3D measurement, only a 2D image is measured. In addition, also the phase information of the scattered waves is lost. The latter is known as the ”phase problem” and poses a serious obstacle when trying to recover the 3D electron density. We tackle this problem using invertible neural networks trained on theoretical electron densities and simulated Small Angle X-ray scattering (SAXS) measurements. We implemented a computationally highly efficient software library for simulating SAXS measurements, also called the forward model. Based on this implementation, we were able to create

Philipp Benner
Bundesanstalt für Materialforschung und -prüfung (BAM), Unter den Eichen 87, 12205 Berlin, e-mail: philipp.benner@bam.de

Sofya Laskina
Bundesanstalt für Materialforschung und -prüfung (BAM), Unter den Eichen 87, 12205 Berlin, e-mail: sofya.laskina@bam.de

Brian Pauw
Bundesanstalt für Materialforschung und -prüfung (BAM), Unter den Eichen 87, 12205 Berlin, e-mail: brian.pauw@bam.de

a library of theoretical electron densities and their corresponding SAXS measurements. The library contains electron densities composed of different shapes, including spheres and cylinders. We used this library to develop an invertible neural network that is able to recover certain parameters of the electron densities. Despite the large loss of information when applying the forward model, our method is able to reliably identify size and shape parameters. This shows that invertible neural networks have large potential to interpret the measurements of SAXS experiments. Ideally, we hope that this work is a first step towards a fully automated measurement and analysis workflow.

Improving cancer diagnosis times by optimising resource allocation

Elizabeth A. Cooke[1], Nadia A. S. Smith[1,2], Spencer A. Thomas[1], Carolyn Ruston[1], Sukhraj Hothi[3] and Derralynn Hughes[3]

Key words: cancer diagnosis, particle swarm optimisation, discrete event simulation, resource allocation

The UK's National Health Service (NHS) constitution sets out minimum standards for rights of access of patients to NHS services. The 'Faster Diagnosis Standard' (FDS) states that 75 % of patients should be told whether they have a diagnosis of cancer or not within 28 days of an urgent general practitioner (GP) referral [NHS(2019)]. Timely diagnosis and treatment lead to improved outcomes for cancer patients, however, compliance with these standards has recently been challenged, particularly in the context of operational pressures and resource constraints relating to the COVID-19 pandemic. In order to minimise diagnostic delays, we have proposed addressing this problem by treating it as a resource optimisation problem, aiming to minimise the number of patients who breach the FDS [Cooke(2022)].

Using the historical patient numbers for the Royal Free London (RFL) NHS Trust and the recommended pathways for different cancer types, Fig. 1 shows a visualisation of all cancer diagnosis pathways. The width of each line is proportional to the number of patients referred for each type of cancer. Fig. 1 gives a clearer picture of where bottlenecks – i.e., investigations required by larger numbers of patients – might be, aiding the RFL in planning resourcing.

We created a resource allocation model to simulate different scenarios of patients flowing through the diagnosis pathways over the course of six months. The model was validated against actual performance at the RFL, using historical data. The model was then optimised using a particle swarm optimisation algorithm to see where capacity might need to be increased to improve patient flow through the system and to reduce the number of patients breaching the FDS. Fig. 2 shows the output optimised resource capacities needed for the RFL to meet their FDS targets.

Previous work has addressed patient diagnosis time in single cancer pathways [England(2021)]. Here, we examined all cancer diagnosis pathways as a complete system. Fig. 2 shows that changing resource in one investigation impacts resourcing needed in other areas. This suggests that studies which look at individual areas may miss the wider context, and not necessarily improve overall patient diagnosis times. Looking at the system as a whole, we have identified areas for improvement which can be used for future resourcing and will have system-wide impact.

[1] National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, UK
[2] Royal Surrey NHS Foundation Trust, Guildford, GU2 7XX, UK
[3] Royal Free London NHS Foundation Trust, Pond St, London, NW3 2QG, UK
e-mail (corresponding author): elizabeth.cooke@npl.co.uk

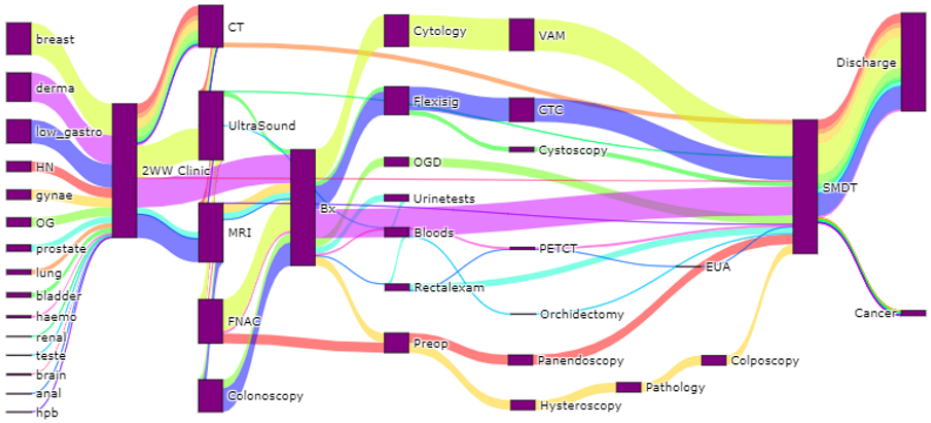


Figure 1: *Sankey flow diagram depicting the various routes through cancer diagnosis pathways taken by patients. Purple blocks show investigations in the pathways; their height represents the number of patients requiring that investigation. Coloured ribbons show routes taken by patients with different suspected cancer types; the width of the lines represents the number of patients.*

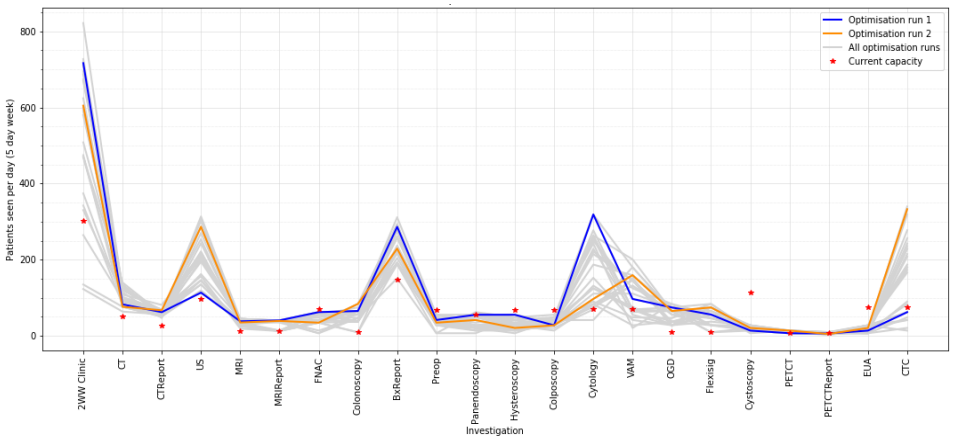


Figure 2: *Results from 25 optimisation runs showing the capacity of patients needing to be seen for each investigation in a working week. Each grey line represents a solution which results in the RFL meeting its target for patient diagnosis. Highlighted are two runs to show the spread in results and the importance of optimising all investigations simultaneously. By lowering capacity in one investigation, capacity must be increased elsewhere to meet targets. The red stars show the current working capacity at the RFL.*

References

[Cooke(2022)] Cooke, Smith, Thomas, Ruston, Hothi, and Hughes, *An integrated discrete event simulation and particle swarm optimisation model for optimising efficiency of cancer diagnosis pathways*, Healthcare Analytics, **2**, (2022)

[England(2021)] England, Harper, Crosby, Gartner, Arruda, Foley and Williamson, "*Examining the diagnostic pathway for lung cancer patients in Wales using discrete event simulation*," Translational Lung Cancer Research, **10**, no. 3, (2021)

[NHS(2019)] NHS England, [Online]. Available: https://www.england.nhs.uk/cancer/faster-diagnosis/

Sensitivity Analysis for Gamma Index Calculations in Dosimetry Audits for Advanced Radiotherapy

Padmini Krishnadas[1,*], Nadia Smith[1], Mohammed Hussein[1]

Key words: Sensitivity analysis, dosimetry, Radiotherapy, Gamma index

Measurement pipelines are becoming increasingly complex with technological advancements, making it harder to establish sources of variability in measurements and understand if they stem from true differences in the measurands or in the measurement pipelines themselves. Radiotherapy dosimetry audits [Hussein (2017)] are carried out by NPL to independently validate radiotherapy technique implementations in the clinic. The audits make use of anthropomorphic tissue-mimicking phantoms embedded with passive 2D detectors. By sending the phantom through the typical patient treatment pathway it is possible to metrologically verify the agreement between the radiation dose deposited in the phantom and that expected by the clinic. The gamma index calculation is a widely accepted metric used for the comparison of the measured and predicted dose. Various steps in the analysis approach can introduce variation in the output that requires quantification to get accurate results. The in-house Versatile Independent Gamma Analysis (VIGO) software [Hussein (2017)] was designed at NPL to provide standardised gamma index computation for complex radiotherapy audits. In this work, we perform a sensitivity analysis to better understand the sensitivity of gamma index calculations in VIGO with respect to factors that each vary in a defined interval.

We planned an experimental design and employed an Analysis of Variance (ANOVA) to perform sensitivity analyses that calculates the *Type III Sum of Squares* [Raphael (2021)] for factors involved in gamma index calculation. This indicates the relative sensitivity of measurements to changes in the pipeline factors in the experiment. The factors included were the calibrated shape and size, reference shapes and sizes of regions of interest, the calibrated offset dose and position along the axis. Figure 1 is a pie chart illustrating the Pearson coefficients for main interactions of the factors that had significant influence over the output measurements (> 0.999 %) for a global gamma index measured with 5 % as the dose difference criterion and using 2 mm as the distance difference criterion, calculated in an area enclosed by 50 % of the prescription isodose. It was found that the factor introducing most variation was the calibrated shape. We split the data based on this factor and subject the resulting data frames to sensitivity analyses. Figure 2 is a comparison bar plot representing

[1] National Physical Laboratory, Hampton Rd, Teddington TW11 0LW,
* corresponding author: padmini.krishnadas@npl.co.uk

Pearson coefficients for main interactions after the dataset was split based on the calibrated shape, with measurements obtained under the same constraints. We observed the variation caused by higher order interactions was significantly larger when using a circle calibration as opposed to a square calibration and that the calibrated shape and the calibrated size of the region of interest significantly contribute to variations in the output measurements.

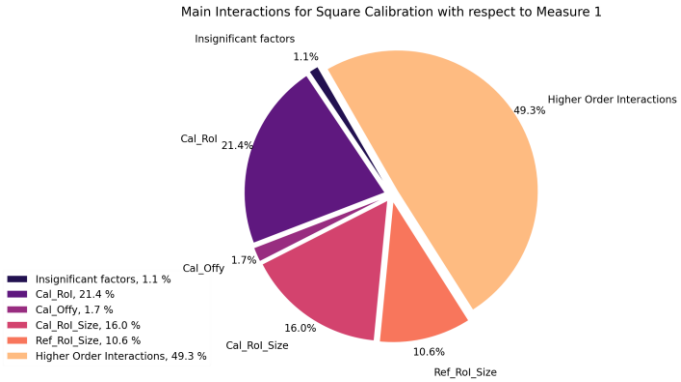


Figure 1: Pie chart representing Pearson Coefficients of the factors when using parameter 5 % / 2 mm to calculate global gamma index.

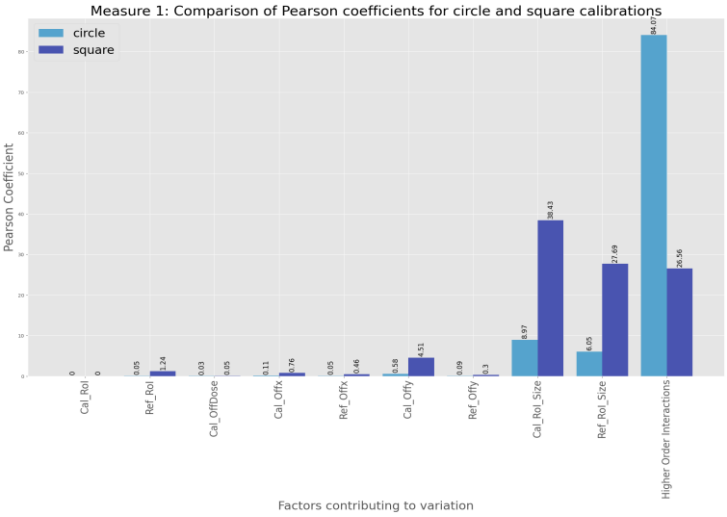


Figure 2: Comparison bar plots representing Pearson Coefficients of the factors when using parameter 5 % / 2 mm to calculate global gamma index with the calibrated shape of region of interest set to square and circle respectively.

References

1. Hussein, Mohammad, C. H. Clark, and Andrew Nisbet. "Challenges in calculation of the gamma index in radiotherapy–Towards good practice." Physica Medica 36 (2017): 1-11.
2. Raphael Vallat. Pingouin ANOVA documentation. Pingouin statistics https://pingouin-stats.org/generated/pingouin.anova.html (2021)

Bridging the Gap between Design and Metrology using Statistical Tolerance Analysis

Mattia Maltauro[1], Roberto Meneghello[1] and Gianmaria Concheri[2]

Key words: ISO GPS, geometric inspection, custom reports, design for metrology

1. Introduction

Design and metrology departments are not always properly linked in the industrial context. In the conceptual framework of design for metrology [Morse (2019)], three items can be defined: design for metrology, design using metrology (metrology feedback), and metrology for manufacturing. This contribution falls within the idea of design using metrology. For designers is often difficult to understand the functional impact of metrological inspection, also because metrological reports are created without functionality in mind. In the framework depicted in [Maltauro, Meneghello, and Concheri (2023)] the design department provides the quality control department with the functional geometric specification, and the latter defines and performs an inspection based on this specification, giving back the measurement. To ease the designers' understanding of the measurement report, we propose a framework in which, through statistical tolerance stack-up analysis, the functional impact of each measurand can be determined. This information is used to create a custom inspection report showing only the critical quantities, from a functional point of view, and in a function-based order.

2. Implementations

A statistical tolerance analysis can be performed based on the functional geometric specification defined by the design department. For the present work the software

[1] Department of Management and Engineering, University of Padova, Stradella San Nicola 3, 36100 Vicenza Italy, e-mail: mattia.maltauro@phd.unipd.it & roberto.meneghello@unipd.it

[2] Department of Civil, Environmental and Architectural Engineering, University of Padova Laboratory of Design Methods and Tools in Industrial Engineering, Via Venezia 1, 35131 Padova Italy, e-mail: gianmaria.concheri@unipd.it

CETOL6sigma from Sigmetrix is used. The statistical core of this software is explained in [Cox (1986)]. Each functional requirement, at the assembly level, is translated into a critical dimension, i.e., a gap to be controlled. The critical dimension is statistically computed based on the admissible variability defined by the tolerances at the components level. The results are the statistical distributions of the critical dimension and the contribution factors showing the impact of each geometric tolerance on functionality: the higher the contribution the more the deviation of the geometric feature will degrade the performance. The contribution is defined as the impact of each tolerance on the variability (standard deviation) of the functional output.
A part, pertaining to the assembly, can be inspected based on the functional geometric specification obtaining the actual deviations for the features controlled by the tolerances. Parts from a simple assembly are 3D scanned with ATOS 5 and inspected with GOM Inspect. The inspection is exported in xml format.
A correlation core is programmed in phyton reading the CETOL source file and the GOM Inspect output file. Congruent features in the two files are paired together and the information regarding the contribution, tolerances and deviations is collected.
Based on a specific functional requirement it is possible to generate specific customized measurement reports where the information is displayed in order of contribution. Further filtering is even possible, as displaying only out-of-tolerance measurands, or measurands with a contribution higher that a predefined percentage. A specific report is generated per each part and functional requirement combination.

3. Conclusions

The proposed framework tries to link together two departments whose communication may be difficult. The bonding is given by statistical tolerance analysis that can assign each tolerance, and therefore measurand, a contribution to the quality of the assembly. This information is used to filter and rank the measurand that are displayed in the inspection report resulting in a clean and effective measurement report that can be easily understood by designers.

References

1. Cox, N. D. 1986. Volume 11: How to Perform Statistical Tolerance Analysis. Milwaukee, WI: American Society for Quality Control.
2. Maltauro, Mattia, Roberto Meneghello, and Gianmaria Concheri. 2023. Tolerance Specifications Management in Industrial Product Design Cycle. Preprint. doi: 110.21203/rs.3.rs-2556637/v1.
3. Morse, Edward. 2019. “Design for Metrology – a New Idea?” Procedia CIRP 84:165–68. doi: 10.1016/j.procir.2019.04.240.

T2 or not T2? A new tool for consistent processing of qMRI parameters.

Jack D. Clarke[1], Elizabeth A. Cooke[1], Nadia A.S. Smith[1], the iMet-MRI consortium[2] and Matt G. Hall[1]

Key words: quantitative MRI, Standardization, T2 Parameter

Magnetic Resonance Imaging (MRI) is an important medical imaging modality. Despite its utility, there are issues relating to consistency of MR images – specifically, comparing images from different scanners or at different times is problematic [Seiberlich 2020]. Traditional clinical MRI employs only relative contrast – one tissue region may be brighter or darker than another, but the contrast is not intrinsically meaningful or quantitative. This means the comparison of different images of different subjects is not straightforward. To counter this, quantitative MRI (qMRI) techniques can be used, which define physical and quantitative properties of the tissue. Quantitative measurements are far more reproducible than relative contrast and enable the performance of different scanners to be analysed and benchmarked.

In this work, we apply qMRI techniques to sets of phantom MRI data by creating qMRI Consistent Processing (QCP) tools. The work has been done as part of a project which aims to develop test objects, procedures, analysis tools, and best practice guidance for various qMRI techniques [iMet-MRI 2023]. The tools provide a homogeneous qMRI measurand estimation process, by applying the same equations and methods to any set of input images. Here, we outline the development and testing of our new QCP tool used to measure T2 (transverse relaxation time) parameters from MRI data. T2 is a standard qMRI measurand, which is different to clinical T2-weighted imaging because we find the T2 value for each voxel.

Our Python tool is designed to extract T2 DICOM MRI data and a semi-automatic region detection algorithm is employed to find the regions of interest (ROIs) at the first echo time point. The T2 value for each pixel in each ROI is estimated using a fitting method from [Raya 2009] taking into account the noise distribution of the measured signal intensity (Eqn. 1).

$$S = \sqrt{\frac{\pi}{2}C^2}\left((1+2\alpha)I_0(\alpha) + 2\alpha\, I_1(\alpha)\right), \qquad \alpha = \left(\frac{S_0}{2C}\, e^{\frac{-t_e}{T_2}}\right)^2 \qquad (1)$$

Key: S: image pixel intensity (in arbitrary units); t_e: echo time (time since initial excitation pulse); T_2: T_2 decay constant; S_0: Initial pixel magnitude; C: noise parameter for Rician model (equivalent to std dev); I_0 and I_1: exponentially scaled modified Bessel functions of order 0 and 1, resp. [Jackson 2022].

The overall T2 value for each ROI is then calculated as the mean of the pixel T2 values from each ROI. An example of this is shown in Figure 1.

[1] Jack D. Clarke, Elizabeth A. Cooke, Nadia A. S. Smith, Matt G. Hall
National Physical Laboratory – Hampton Rd, Teddington TW11 0LW
Email: jack.clarke@npl.co.uk
[2] iMet-MRI consortium
Improved Metrology for quantitative MRI http://empir.npl.co.uk/imet-mri/

To test the QCP tool, we used data from four test centres to calculate and compare T2 parameters. These centres also each used distinct T2 calculating algorithms on their respective datasets, which we include for comparison with our own fitting results. All T2 values are shown in Table 1.

The alignment between the QCP calculated results and the site calculated results provides evidence that the QCP algorithm is a suitable tool for aiding the standardisation of qMRI parameters. Future work will extend this tool to other qMRI techniques, such as T1 mapping (longitudinal relaxation time) and proton density fat fraction. The standardisation of qMRI calculations has significant potential advantages for the qMRI field, including in analysis and comparison of studies, clinical trials, and the use of Artificial Intelligence in healthcare.

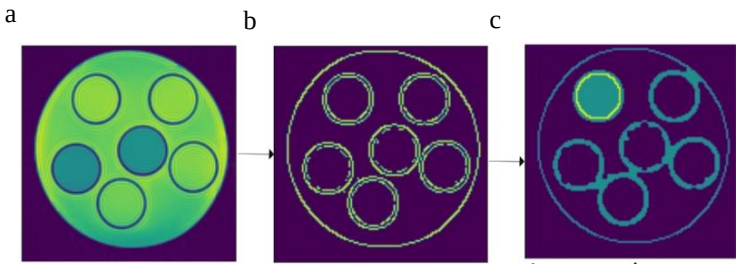


Figure 1 *:* Demonstration of the QCP algorithm with example data and T2 parameter fitting. First, MRI images are loaded from the directory (a). Then, the bounded regions are found (b). Then T2 parameters are calculated for each pixel in each ROI (c). Then, the mean of those T2 values is calculated per ROI and stored for comparison.

Table 1: Comparison between centre calculated T2 parameters and QCP calculated T2 parameters on each site's respective dataset. Numbers following ± are an expanded uncertainty corresponding to a coverage probability of 95%. The expected T2 value for these vials was 387 ms.

Site	Site T2 (ms)	QCP T2 (ms)
Azienda Ospedaliero Universitaria Careggi (AOUC), Italy	392	397.2±10.6
Belfast Health and Social Care Trust Hospitals (HSC), Ireland	505	516.2±22.5
University College London (UCL), UK	232	228.9±9.4
University Hospitals Bristol and Weston NHS Foundation Trust (UHBW), UK	435.1	430.4±22.8

References

1. iMet-MRI Publishable Summary, 2023
2. Jackson, L. *et al.* hazen (Version 1.1.0) 2022
3. Raya J *et al*. T_2 measurement in articular cartilage: Impact of the fitting method on accuracy and precision at low SNR. Magnetic Resonance in Medicine 2010; 2009
4. Seiberlich N *et al*. Quantitative Magnetic Resonance Imaging. Advances in Magnetic Resonance Technology and Applications 1 **xxxvii** 2020

Repeatability and Reproducibility Uncertainty Assessment in Magnetic Resonance-based Electric Properties Tomography of a Homogeneous Phantom

Alessandro Arduino[1], Francesca Pennecchi[1], Ulrich Katscher[2], Maurice Cox[3] and Luca Zilberti[1]

Key words: electric properties tomography, magnetic resonance imaging, phantoms, quantitative imaging, uncertainty

Magnetic resonance-based electric properties tomography (MR-EPT) is a quantitative imaging technique that non-invasively estimates the electric property values inside a human body [Leijsen (2021)]. Some early clinical applications of MR-EPT have been presented in the literature and most of them rely on the MR-EPT implementation called Helmholtz-EPT, in particular its phase-based approximation [Shin (2015), Kim (2016), Tha (2018, 2021)]. Helmholtz-EPT suffers a large sensitivity to input noise and the presence of systematic errors at tissue boundaries. Recognizing this issue, this contribution presents a repeatability and reproducibility uncertainty assessment in the electric conductivity of a homogeneous phantom estimated by phase-based Helmholtz-EPT.

A homogeneous cylindrical phantom of a known solution of NaCl in distilled water was acquired with a 3 T Ingenia TX scanner (Philips Healthcare, Best, The Netherlands) with a body coil in transmission and a 32-channel head coil in reception using a steady-state free precession (SSFP) sequence with nominal flip-angle of 30° and isotropic resolution of 2 mm. The measured phase of the acquired complex-valued image is a good estimate of the transceive phase from which the electric conductivity is estimated. Repeatability conditions are obtained by acquiring 25 images with the phantom centered at the scanner isocenter. Reproducibility conditions are obtained with 8 additional images acquired by moving the phantom at

[1]Istituto Nazionale di Ricerca Metrologica, 10135 Torino, ITALY, e-mail: a.arduino@inrim.it
[2]Philips Research Laboratories, 22335 Hamburg, GERMANY
[3]National Physical Laboratory, Teddington TW11 0LW, UK

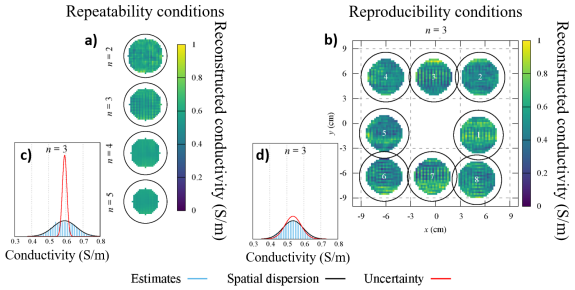


Fig. 1 Estimated maps under repeatability (a) and reproducibility (b) conditions, with kernels of various sizes n. Statistical distribution of the estimated conductivity under repeatability (c) and reproducibility (d) conditions, compared with the spatial dispersion and the median uncertainty.

a distance of about 6 cm from the isocenter between the scans (see Fig. 1b). The covariance matrices of the means of the phase and conductivity maps are evaluated according to the James–Stein shrinkage estimator [Schäfer (2005)] and the law of propagation of uncertainty. The EPTlib implementation of phase-based Helmholtz-EPT [Arduino (2021)] and the code for uncertainty propagation published in [Arduino (2020)] were used.

The conductivity maps estimated under repeatability conditions show spatial noise, corresponding to a spatial dispersion of the estimated conductivity, larger than the median standard uncertainty associated with each voxel. Hence, averaging multiple repeated acquisitions does not correct for this, apparently random, spatial noise. On the other hand, the conductivity maps estimated under reproducibility conditions show that the repositioning of the phantom changes the distribution of the repeatable noise, making it random from a practical point of view. Indeed, we observed that the spatial dispersion of the conductivity values estimated under reproducibility conditions is comparable with the median standard uncertainty associated with each voxel (see Fig. 1).

Acknowledgements The results here presented have been developed in the framework of the EMPIR Project 18HLT05 QUIERO. This project 18HLT05 QUIERO 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

Arduino A (2021) Appl Sci 11:3237
Arduino A, et al (2020) URL https://zenodo.org/record/4248879, [Dataset]
Kim SY, et al (2016) Eur Radiol 26:2317–2326
Leijsen R, et al (2021) Diagnostics 11:176
Schäfer J, Strimmer K (2005) Stat Appl Genet Mol Biol 4
Shin J, et al (2015) J Magn Reson Imaging 42:371–378
Tha KK, et al (2018) Eur Radiol 28:348–355
Tha KK, et al (2021) J Magn Reson Imaging 54:1689–1691

Determining radius and refractive index of nanoparticles using machine learning

Federica Gugole[1], Martine Kuiper[1,2] and Richard Koops[1]

Key words: machine learning, inverse problem, validation, nanoparticle characterization, Mie theory

1. Introduction

Extracellular vesicles (EVs) are nanoparticles released by the cells and are present in all body fluids. EVs have the potential to be a biomarker to distinguish between healthy and diseased individuals. EV measurements are generally performed using flow cytometry, which measures the light scattering of particles. Mie theory [Mie (1908)] determines how the particle will scatter light (i.e., the scattering pattern) using given properties of the particle such as its radius and refractive index. In a measurement setup the challenge is to solve the inverse problem, i.e., given the scattering pattern determine the radius and refractive index of the particle that generated it. Here we present preliminary results and future steps of an approach to solve the inverse problem using convolutional neural network (CNN).

2. Solving the inverse problem using CNN

Machine learning algorithms are becoming popular in many application areas thanks to the large availability of data and their fast computation times when making predictions for new input data. In particular, CNN is the state of the art for image processing without any competitor in terms of performance. CNNs find application in many areas such as medical imaging [Kim(2022)] and autonomous driving [Sonata(2020)]. Recently CNNs have been applied to characterize nanoparticles using images obtained with optical microscopy [Midtvedt(2021)]. Here we follow a similar approach for a flow cytometer setup.

2.1 Training of the CNN on synthetic data

[1] Federica Gugole, Martine Kuiper, Richard Koops
VSL, Thijsseweg 11, 2629 JA Delft, The Netherlands, e-mail: fgugole@vsl.nl, mkuiper@vsl.nl, rkoops@vsl.nl
[2] Martine Kuiper
Laboratory Experimental Clinical Chemistry / Vesicle Observation Center / Biomedical Engineering and Physics, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands, e-mail: m.w.kuiper@amsterdamumc.nl

To train the CNN we first generate synthetic data using numerical simulations of the forward problem in Mie theory (i.e., simulate the scattering pattern given the radius and refractive index). Artificial noise is added to the generated images to simulate the noise present in the measurement setup. For a good performance of the CNN on experimental data, it is key to ensure that the data used in training are representative of the experimental data.

2.2 Validation of the CNN

The CNN will be validated using reference particles. Scattering patterns of reference particles obtained in the flow cytometer will be used to assess the similarity of the synthetic data used for training with the experimental data, and to evaluate the accuracy of the CNN predictions.

2.3 Uncertainty estimate associated to the CNN prediction

In order to use the CNN predictions as part of a traceable measurement procedure a comprehensive measurement uncertainty should be calculated. Ensemble and Bayesian methods [Gawlikowski (2021)] are considered for the uncertainty evaluation.

Acknowledgment

Part of this work has been developed within the 18HLT01 MetVes II project. This project (18HLT01 MetVes II) 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. Gawlikowski J., Tassi C. R. N., Ali M., Lee J., Humt M., Feng J., Kruspe A., Triebel R., Jung P., Roscher R., Shahzad M., Yang W., Bamler R., Zhu X. X., *A survey of uncertainty in Deep Neural Networks*, arXiv, (2021)
2. Kim H. E., Cosa-Linan A., Santhanam N., Jannesari M., Maros M. E., Ganslandt T.: *Transfer learning for medical image classification: a literature review*, BMC Medical Imaging, **22**, (2022)
3. Midtvedt B., Olsén E., Eklund F., Höök F., Adiels C. B., Volpe G., Midtvedt D.: *Fast and Accurate Nanoparticle Characterization Using Deep-Learning-Enhanced Off-Axis Holography*, ACS Nano, **15**, 2240-2250 (2021)
4. Mie G.: *Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen*, Annalen der Physik, **330** (3), 377-445 (1908)
5. Sonata I., Heryadi Y., Lukas L., Wibowo A.: *Autonomous car using CNN deep learning algorithm*, Journal of Physics: Conference Series, **1869**, (2020)

Challenges related with Virtual Experiments in Metrology

Gertjan Kok[1], Marcel van Dijk[1], Federica Gugole[1], Gerd Wübbeler[2], Manuel Marschall[2], Clemens Elster[2]

Key words: virtual experiments, uncertainty evaluation, validation, coordinate measurement machine, coherent Fourier scatterometry

1. Introduction

Virtual experiments (VEs) in metrology are simulations of measurements in which potential error sources are carefully addressed. They usually express the measurement data as function of the measurand, in opposite of what the GUM [GUM(2008)] requires as a measurement model. VEs can give a greater understanding of the measurement, and potentially help with the calculation of the best estimate of the measurand and its uncertainty, cf. [Wübbeler(2022)] for the case of linear models.

Recently, the European research project 'Trustworthy virtual experiments and digital twins' (ViDiT) [ViDiT(2023)] has been granted. In this project several challenges related to VEs will be investigated. In this contribution some of these challenges will be presented on the basis of the practical applications of a coordinate measurement machine (CMM) [Kok(2022)] and a coherent Fourier scatterometer (CFS) [Kumar(2014)] and their virtual counterparts (VCMM, VCFS).

2. Uncertainty evaluation

In the case of a CMM often the measured coordinates are plugged into the VCMM and a Monte Carlo method is employed. From a simulation point of view, this procedure results in adding noise to measurement data that already reflects random effects associated with the measurement system, which is statistically incorrect. More importantly, if the model is strongly non-linear, the calculated uncertainty can strongly vary with the input data, as shown in [Kok(2022)]. As an alternative, a Bayesian approach can be used.

[1] Gertjan Kok, Marcel van Dijk, Federica Gugole
VSL, Thijsseweg 11, 2629 JA Delft, the Netherlands, e-mail: gkok@vsl.nl, mvdijk@vsl.nl, fgugole@vsl.nl

[2] Gerd Wübbeler, Manuel Marschall, Clemens Elster
PTB, Braunschweig and Berlin, Germany, e-mail: gerd.wuebbeler@ptb.de, manuel.marschall@ptb.de, clemens.elster@ptb.de

In CFS the measurement data consists of an image and the geometrical parameters of the measured sample need to be reconstructed by the inversion of (a simplified version of) the VE. The fit result is never perfect, as the measurement contains measurement noise and as the VE, and a fortiori a simplified model, is not a completely faithful representation of the physical reality. Inversion can be done with respect to different norms or by taking a Bayesian approach, and this also affects the calculated uncertainty.

In this talk we will give more details on the challenges encountered for these applications, and present the planned work in ViDiT with the goal of defining an uncertainty evaluation method that is in line with the GUM [GUM(2008)].

3. Validation

For VCMMs there exists a validation method that consists of calibrating all individual error sources of the real CMM, plugging the resulting relevant parameters as uncertainty sources in the VCMM, and performing validation measurements on reference artefacts. This is very labour intensive. In ViDiT we will investigate if alternative approaches, e.g., based on repeated measurements, can alleviate this burdensome calibration. We will also study the existing validation approaches in the light of the GUM.

For CFS an important aspect of validation is the definition of an analysis method based on a simplified version of the VE that allows inversion within reasonable time. We will show various possible simplifications. In ViDiT guidance will be developed for such type of problems.

Acknowledgment

The project (22DIT01 ViDiT) has received funding from the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

References

1. JCGM. Guide to the expression of uncertainty in measurement. Sèvres, France: BIPM, September 2008. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, JCGM 100:2008.
2. G. Wübbeler, M. Marschall, K. Kniel, D. Heißelmann, F. Härtig and C. Elster. GUM-Compliant Uncertainty Evaluation Using Virtual Experiments. Metrology, 2(1), 114-127, 2022
3. Partnership project 'Trustworthy virtual experiments and digital twins' (22DIT01 ViDiT), 2023–'26
4. Kok, G.; Wübbeler, G.; Elster, C. Uncertainty calculation using virtual experiments: a case study for a virtual scatterometer. Presentation at MathMet international conference, Paris, November 2022.
5. Nitish Kumar, et. al., "Reconstruction of sub wavelength features and nano positioning of gratings using coherent Fourier scatterometry ," Opt. Express 22, 24678 24688 (2014)

Predicting Equivalent Electrical Circuits from Electrochemical Impedance Spectroscopy (EIS) Data with Convolutional Neural Networks and Global Optimization

Bashar Albakri[1], Marco Favaro[2], Alexander Kister[1]

[1] Federal Institute for Materials Research and Testing (BAM), S.3 eScience
[2] Helmholtz-Zentrum Berlin, Institute for Solar Fuels

Key words: Equivalent Electrical Circuits, Electrochemical Impedance Spectroscopy, Neural Networks, Global Optimization

Abstract

Researchers study the properties of polymeric membranes - an important building block of batteries – by first measuring its impedance spectra and then determining a so called equivalent electronic circuit (EEC), that is an electronic circuit that approximately reproduces these measurements. Determining EEC is not completely automatized since it includes the search for the topology of the electronic circuit. This search is an inverse problem, since the formulas for determining the impedance spectra given the circuit topology and parameters are known while there are no formulas for the invers direction. To solve this invers problem, the original problem is split into two parts. The first part is to determine the topology of the circuit and the second part is to determine the parameters of the circuit. For the first part a Convolutional Neural Network Classifier is trained on a simulated data set, where the simulator is an implementation of the formulas for determining the impedance spectra given the complete description of the electronic circuit. After the topology is determined, the parameters of the electronic circuit are found by minimizing the error between the observed spectra and the spectra corresponding to these parameters. This minimization is a global optimization.

Investigation of a Bayesian approach for the calibration of large batches of sensors

Andrea Prato[1], Francesca Pennecchi[2], Gianfranco Genta[3] and Alessandro Schiavi[4]

Key words: Bayesian statistics, calibration, large production batch, MEMS

1. Introduction

Low-cost sensors and in particular MEMS (Micro-Electro-Mechanical Systems) devices are widely used in many applications, including consumer electronics, healthcare, automotive, and industrial automation. Their production, which is typically in the order of millions per week in a single factory, involves the calibration of these devices on a large scale which can be costly and time-consuming. This is a significant challenge for manufacturers who need to guarantee the required traceability to SI. To address these challenges, as also requested by the Consultative Committee for Acoustics, Ultrasound, and Vibration of BIPM [1], a solution can be found in the use of statistical process control techniques [2]. Based on probabilistic models that take into account prior knowledge and uncertainties, Bayesian approaches can be very useful for "statistically calibrating" large batches of sensors. In this paper, a Bayesian method is investigated and proposed in this respect: it implies experimentally calibrating a sample of sensors from an unknown batch of larger dimension and using them to estimate the true number of reliable sensors in the whole batch, process that can be considered as a "statistical calibration" of the batch. The advantage of this statistical approach is that it allows for the incorporation of the prior knowledge coming from the calibration of a golden batch representative of the whole production process. This method allows for the calibration of large batch sensors by reducing the number of actual experimental calibrations.

[1] Andrea Prato
INRiM, Strada delle Cacce 91, 10135 Torino, e-mail: a.prato@inrim.it
[2] Francesca Pennecchi
INRiM, Strada delle Cacce 91, 10135 Torino, e-mail: f.pennecchi@inrim.it
[3] Gianfranco Genta
Politecnico di Torino, DIGEP, Corso Duca degli Abruzzi 24, 10129 Torino, e-mail: gianfranco.genta@polito.it
[4] Alessandro Schiavi
INRiM, Strada delle Cacce 91, 10135 Torino, e-mail: a.schiavi@inrim.it

2. The proposed method in brief

The first step of the proposed method consists in the usual calibration in the laboratory of a golden batch of N sensors, representative of the production process: each sensor is provided with an estimate of its sensitivity S and the associated uncertainty. Using a mixture distribution model to account for the calibration uncertainties of the individual sensors, the overall distribution of the sensitivities of the golden batch is modelled. By setting a bearable probability (p_{def}) of finding out-of-tolerance sensors in the golden batch, the mixture distribution is used to find the minimum and maximum sensitivity limits, S_{min} and S_{max} respectively, encompassing the 1-p_{def} fraction of acceptable sensors. Hence, the expected number of out-of-tolerance sensors in the golden batch is $C_{gold} = p_{def}\ N$. For other/future unknown batches of the same kind of sensors, only a sample of $n < N$ devices from each batch is required to be experimentally calibrated, hence reducing time and cost efforts. The calibrated devices are checked whether their sensitivity is within or outside the out-of-tolerance limits: k indicates that number. The statistical calibration of the whole batch of N sensors is then based on the following Bayesian model. Assuming that the expected number of out-of-tolerance sensors in the unknown batch is equal to that in the golden batch, a binomial prior distribution $P_{prior}(C_{unknown}; N, p_{def})$ is used to model the number of $C_{unknown}$ defective sensors in the unknown batch. This prior (assuming that experimental calibration and manufacturing process do not change from batch to batch) models the state of knowledge of a typical batch from that production. Multiplying the prior by a likelihood function $P_{likelihood}(k; n, C_{unknown}, N)$ defined as a hypergeometric distribution with n sensors, k of which are defective, drawn from the unknown batch of N sensors, $C_{unknown}$ of which are defective, yields an un-normalized posterior $P_{post,un}(C_{unknown}; k, n, N, p_{def})$ for $C_{unknown}$ out-of-tolerance sensors in the unknown batch. Normalizing this distribution leads to a probability mass function for the number $C_{unknown}$ of out-of-tolerance pieces in the unknown batch, provided that k out-of-tolerance devices were found in the small calibrated sub-batch. At this point, it is possible to define different metrics based on the posterior cumulative probability function that can provide an assessment of the reliability and quality of the whole unknown batch.

References

1. BIPM - Consultative Committee for Acoustics, Ultrasound, and Vibration (CCAUV), Strategy plan 2019 to 2029, (2019)
2. Prato A., Mazzoleni F., Pennecchi F. R., Genta G., Galetto M., Schiavi A.: Towards large-scale calibrations: a statistical analysis on 100 digital 3-axis MEMS accelerometers. IEEE International Workshop on Metrology for Industry 4.0 & IoT, 578--582 (2021)

A procedure for optimal designs and modeling in technological processes: a case-study on freight trains

Rossella Berni[1], Luciano Cantone[2], Alessandro Magrini[3], Nedka D. Nikiforova[4]

Key words: genetic algorithm, D-optimality criterion, random effects, hierarchical optimal design

1. Introduction

Complex engineering and technological processes typically generate data with a non-trivial hierarchical structure which make the experimental planning challenging. However, when preliminary data are available, the relevance of the information gained through a preliminary data analysis may be verified through the building of a pilot design. Therefore, the pilot design results allow to verify the validity of the experimental planning, and to help to define the most suitable statistical model. This point is particularly important when considering the optimal design framework, given the model-dependent nature of optimal designs. In this talk we propose a full procedure for the design and analysis of experiments in complex engineering and technological processes; the talk is entirely based on a recently published paper of Berni et al. (2022). The proposal is applied to a real case-study aiming to optimize the payload distribution of freight trains, thus making it possible to minimize the in-train forces to avoid freight trains derailment and/or disruption. The suggested procedure also allows to achieve the best train configuration for minimizing the in-train forces.

2. Brief description of the full procedure

The proposed full procedure consists of four main steps, as also detailed in Berni et al. (2022). First, we consider (Step #1): (i) the statistical analysis of preliminary data, and (ii) the building of a pilot design, both of which allow for efficiently

[1] Rossella Berni
Department of Statistics Computer Science Applications "G. Parenti", University of Florence, e-mail: rossella.berni@unifi.it
[2] Luciano Cantone
Department of Engineering for Enterprise "Mario Lucertini", University of Rome "Tor Vergata", e-mail: luciano.cantone@uniroma2.it
[3] Alessandro Magrini
Department of Statistics Computer Science Applications "G. Parenti", University of Florence, e-mail: alessandro.magrini@unifi.it
[4] Nedka D. Nikiforova
Department of Statistics Computer Science Applications "G. Parenti", University of Florence, e-mail: n.nikiforova@unifi.it

planning and validating the experimental design. Step #2 consists of performing the experimental runs of the pilot design, and following, the estimation of suitable statistical models is checked through diagnostic measures; this step#2 makes it possible to establish the most suitable statistical model for building the optimal design. In Step #3 we consider the D-optimality design criterion, and a genetic algorithm (GA), opportunely adapted, for obtaining a hierarchical optimal design. The proposed full procedure includes the experimental planning at several hierarchical levels for obtaining the hierarchical D-optimal design. In doing so, one of the main key-points is the presence of several types of experimental factors; more precisely, we deal with nested, branching and shared factors, as well as with a new type of factor, which we call composite-form-factor. Lastly, Step #4 relates to the achievement of best configurations for the problem under study according to the target values of several response variables.

3. Main results for the case-study

According to the full procedure outlined (Section 2), we begin with the analysis of existing data containing information about a large number of trains, used for building the pilot design. Once the pilot design runs are performed, suitable linear mixed-effect models are estimated and checked with diagnostic measures. The satisfactory results obtained confirm the validity of the experimental planning. More specifically, we consider the following three hierarchical levels: i) level-1 related to the entire train, ii) level-2 related to two consecutive wagons, and iii) level-3 related to each individual wagon. Following, we build the D-optimal hierarchical design through a GA; the true values of compressive and tensile forces are simulated through the TrainDy software [Cantone (2011)]. In our setting, we select the GA by specifically considering the non-trivial hierarchical structure, and the types of experimental factors defined (e.g. the composite-form factor); for further details refer to Berni et al. (2022). Two linear mixed-effect models are estimated, one for each response variable (e.g., compressive forces at 10m, tensile forces at 2m); the model results are good enough. The final step #4 corresponds to finding the best train configuration, by considering the experimental factors and variables involved, and allowing to avoid the risk of derailment and disruption. To this end, the two response variables are minimized taking the roles of both forces into account, also considering the target values. The minimization of the defined objective function achieves good performances for the global scenarios, also considering the simultaneous minimization of both responses. The results are very satisfactory and confirm that our full procedure represents a valid method to be successfully applied for solving similar technological problems.

References

1. Berni, R, Cantone, L., Magrini, A., Nikiforova, N.D.: Hierarchical optimal designs and modeling for engineering: a case-study in the rail sector. Appl. Stoch. Models Bus. Ind. **38**, 1061--1078 (2022).
2. Cantone, L.: TrainDy: the new Union Internationale des Chemins de Fer software for freight train interoperability. J. Rail Rapid Transit. **225**, 57--70 (2011).

Electric Properties Tomography via Green's Integral Identity

Luca Zilberti[1], Oriano Bottauscio[1], Umberto Zanovello[1] and Alessandro Arduino[1]

Key words: electric properties tomography (EPT), magnetic resonance imaging (MRI), quantitative imaging, Green identity.

Electric Properties Tomography (EPT) is a quantitative imaging technique performed starting from data acquired during a Magnetic Resonance Imaging (MRI) exam [Leijsen (2021)]. EPT aims at measuring, in each voxel/pixel of the tomographic image, the electric conductivity and, when possible, also the dielectric permittivity of the scanned biological tissues, to be used as objective biomarkers.

In recent years, a plethora of EPT methods were proposed, but the creation of reliable EPT images is still challenging. Typical problems in EPT are represented by the amplification of the input measurement noise and the presence of image artifacts at the boundary between different tissues. To overcome these issues, here we propose an EPT approach based on Green's integral identity. The starting point is the complex Helmholtz equation describing the propagation of the B^+ component of the magnetic flux density (which rotates with angular frequency ω) through a body with conductivity σ and permittivity ε (that may change from point to point):

$$\nabla^2 B^+ = j\omega\mu_0\left(\sigma + j\omega\varepsilon\right)B^+ \tag{1}$$

where j is the imaginary unit and μ_0 is the magnetic permeability of vacuum.

If the right-hand side of (1) were a known term, we would have a Poisson equation, whose solution, exploiting Green's integral identity, would take the form:

$$B^+\left(P\right) = \oint_{\partial\Omega}\left[\Psi\frac{dB^+}{dn} - B^+\frac{d\Psi}{dn}\right]ds + j\omega\mu_0\int_\Omega\left(\sigma + j\omega\varepsilon\right)B^+\Psi dv \tag{2}$$

where P is a generic point within a region Ω surrounded by a surface $\partial\Omega$ (oriented according to an outward normal direction n) and Ψ is the Green function for 3D elliptic problems [Morse (1953)].

[1] Istituto Nazionale di Ricerca Metrologica, 10135 Torino, ITALY, e-mail: l.zilberti@inrim.it

During MRI, the spatial distribution of B^+ can be sampled in a regular grid of N_V voxels. If we assume that each i-th voxel is homogeneous, (2) can be rewritten as

$$B^+(P) = \oint_{\partial\Omega} \left[\Psi \frac{dB^+}{dn} - B^+ \frac{d\Psi}{dn} \right] ds + j\omega\mu_0 \sum_{i=1}^{N_V} (\sigma + j\omega\varepsilon)_i \int_{\Omega_i} B^+ \Psi dv \qquad (3)$$

In a general case, (3) is a complex equation involving N_V complex unknowns (the complex conductivities $\sigma + j\omega\varepsilon$ of the N_V voxels). By placing the reference point P at the barycentre of each voxel in turn, a system of N_V independent equations is obtained, from which the conductivity and permittivity can be calculated.

The proposed formulation has been specifically conceived to model heterogeneous regions (including discontinuities in the spatial distribution of the target parameters), avoiding image artifacts at the boundary between different tissues. Moreover, the use of the integrals in the right-hand side of (3) introduces a spontaneous compensation of the noise that affects the B^+ values used in the integrals themselves. Thus, the size of the region Ω in which the EPT formulation is applied must be chosen as a trade-off between the computational burden (the larger N_V, the bigger the algebraic system to solve) and the need for noise mitigation.

A residual issue resides in the noise that affects the reference value $B^+(P)$ in each equation of the system. To make the procedure more robust with respect to it, the unknowns in (3) can be grouped based on the tissue they belong to. This can be done exploiting the contrast image produced by the traditional MRI exam, which allows differentiating the different tissues in the scanned region. In this case, the number of unknowns becomes lower than N_V and the application of (3) produces a rectangular system of equations, which can be solved in the least-square sense.

If the magnitude of B^+ is quite homogeneous (a common condition in clinical MRI scanners), it can be shown that a phase-based version of the proposed formulation holds, which allows calculating the conductivity σ from the knowledge of the phase of B^+ only, hence simplifying the acquisition stage.

At the conference, the performances of the proposed method will be illustrated and compared to those of other state-of-the-art EPT techniques.

Acknowledgements. The results here presented have been developed in the framework of the EMPIR Project 18HLT05 QUIERO. This project 18HLT05 QUIERO 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. Leijsen R., Brink W., van den Berg C.A.T., Webb A., Remis R.: Electrical properties tomography: a methodological review. Diagnostics **11** (2021).
2. Morse P. and Feshbach H.: Methods of Theoretical Physics. McGraw-Hill, New York (1953).

Explainable deep learning inference to decode decision-making processes from multidimensional patterns of neural activities

Andrea Ciardiello, Antonio Roberto Buonfiglio, Giampiero Bardella, Maurizio Mattia, Pierpaolo Pani, Stefano Ferraina and Guido Gigante

Key words: Deep learning, Explainable AI, Neuronal dynamics, Decision making, Behavioral prediction

Connecting effective behavioural models to high-dimensional experimental data is one of the most important challenges in neuroscience. In this work, we show how a link can be inferred between intracortical neural recordings performed during a simple decision task and the ramping variable postulated by a threshold decision model exploiting an artificial neural network (ANN).

We recorded the multi-unit activity (MUA) from a 96-channel array in the dorsal premotor cortex of two monkeys performing a countermanding reaching task that requires, in a subset of trials, to cancel the planned movement before its onset [2]. We trained a WaveNet-inspired and a multilayer perceptron causal ANN architecture to map this complex data to an accumulation process as derived from theoretical models.

Andrea Ciardiello
"Sapienza" University of Rome, Rome, Italy, e-mail: andrea.ciardiello@uniroma1.it

Antonio Roberto Buonfiglio
Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy, e-mail: antonio.buonfiglio2@unibo.it

Giampiero Bardella
"Sapienza" University of Rome, Rome, Italy, e-mail: giampiero.bardella@uniroma1.it

Maurizio Mattia
Istituto Superiore di Sanità, Rome, Italy, e-mail: maurizio.mattia@iss.it

Pierpaolo Pani
"Sapienza" University of Rome, Rome, Italy, e-mail: pierpaolo.pani@uniroma1.it

Stefano Ferraina
"Sapienza" University of Rome, Rome, Italy, e-mail: stefano.ferraina@uniroma1.it

Guido Gigante
Istituto Superiore di Sanità, Rome, Italy, e-mail: guido.gigante@iss.it

Our results show that neural recordings can be, to a large extent, mapped at the single trial level to a ramping process, whose angular coefficient is tied to the inverse of the movement reaction time (RT). From the predicted ramp it is possible to perform an early estimation of the RT and the outcome of a stop signal presented in a subset of trials with a good level of accuracy. Moreover, the network generalises nicely when tested on sessions having large temporal gaps with the ones used for training (generalisation through time).

We then applied explainability techniques (xAI) to our network to extract insights into the information hidden in the input data [4]. By combining already established xAI algorithms based on Gradient methods and a newly proposed xAI method, that we call "functional explainability", we show how, by perturbing the output function in principled ways, we obtain different spatiotemporal patterns of "saliency" in the input. Notably, our results suggest that the information used to build the ramps (so the decision process and consequently the RT) can be found very early after, if not before, the instruction to initiate the movement (go signal)

We also employed methods of training influence (TracIn) to find the training examples most relevant for a given prediction [3]. The results confirm that the performance of the network on a given test trial is positively influenced by training examples recorded weeks and even months before or after it. Furthermore, when the network is trained on data from two monkeys, TracIn highlights a substantial inter-subject influence, thus hinting at a partial shared representation of the hypothesised ramping process at the neuronal level.

Influence functions are known to sometimes be biased by "global explanations" [1] that are training examples that have a large impact on the training procedure without being directly linkable to any relevant property of other elements present in the dataset. We found that this holds also for our methods, where a few trials are reported as highly influential for the vast majority of the tested data.

Such training examples are usually distinguished by a high "self influence", meaning that they are very important for learning themselves. This is typically seen in noisy, affected by artefacts, mislabeled, or otherwise unique elements of the dataset.

The correct identification of these data points is twofold useful. First, it is possible to clean the dataset, removing the inconsistent data and reducing the noise in the training set. This may lead to better performances during successive retraining of the network. Secondly, these examples may be linked to anomalous recordings of the signal or to a deviation in the behaviour of the subject. This information can be used by neurophysiologists to control or improve their experimental setting or to find new interesting structures in the data.

References

1. E. Barshan, M.-E. Brunet, and G. K. Dziugaite. Relatif: Identifying explanatory training examples via relative influence. In *International Conference on Artificial Intelligence and Statistics*,

2020.
2. G. Mirabella, P. Pani, and S. Ferraina. Neural correlates of cognitive control of reaching movements in the dorsal premotor cortex of rhesus monkeys. *Journal of Neurophysiology*, 106(3):1454–1466, 2011.
3. G. Pruthi, F. Liu, M. Sundararajan, and S. Kale. Estimating training data influence by tracing gradient descent, 2020.
4. G. Ras, N. Xie, M. Van Gerven, and D. Doran. Explainable deep learning: A field guide for the uninitiated. *Journal of Artificial Intelligence Research*, 73:329–397, 2022.

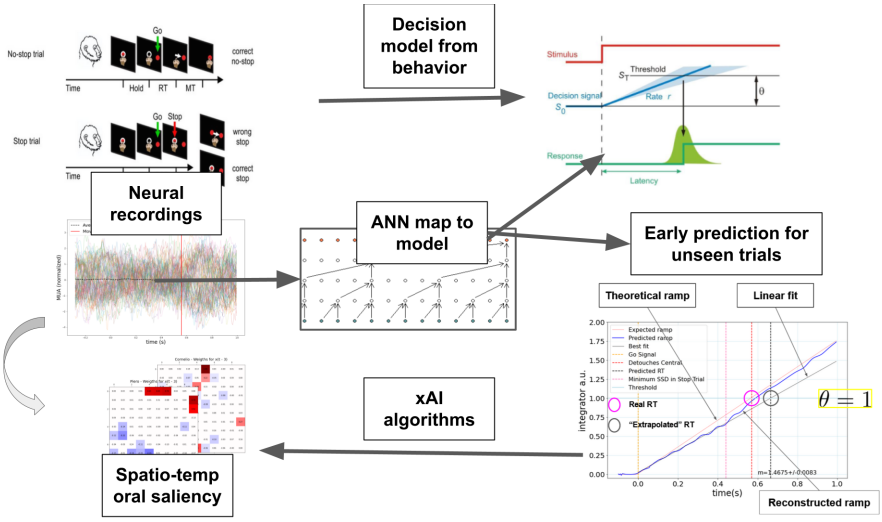


Fig. 1 Scheme of the workflow: Combined registration of subject behaviour and neural recordings in a trial. Subject behaviour is used to extract the parameter of the theoretical model. An ANN is trained to map the recordings to the theoretical model. During inference, it is possible to find the model parameters of unseen trials and from that predict the subject behaviour. Explainability algorithms applied to the ANN prediction can be used to discover new structures in the data.

On the dB-to-linear conversion

Luca Callegaro and Francesca Pennecchi and Walter Bich

Key words: Uncertainty, nonlinearity, estimate, decibel

1 Expression of quantities in decibel

When dealing with a *power-related quantity*, P, one is often interested in the ratio P_{dB} of its value to a reference value P_0 rather than to its SI value. The quantity P_{dB} is expressed in decibel (dB) as the decimal logarithmic ratio of P with respect to P_0:

$$P_{\mathrm{dB}} = 10 \log_{10} \frac{P}{P_0} \tag{1}$$

$$P = P_0 10^{P_{\mathrm{dB}}}. \tag{2}$$

Equations (1) and (2) can be considered as measurement models, whose measurands P_{dB} and P are function of the input quantities P and P_{dB}, respectively (P_0 is a constant).

2 Estimate and uncertainty

The JCGM 100 [1] approach to uncertainty evaluation is based essentially on Gauss' law of propagation of variances, normally implying the linearisation of the measurement model around the estimates of the input quantities. It is acknowledged that non-linearities sufficiently strong in neighborhoods of the input estimates comparable with the magnitudes of the associated standard uncertainties might affect

Luca Callegaro, Francesca Pennecchi, Walter Bich
INRIM - Istituto Nazionale di Ricerca Metrologica, Torino, Italy, e-mail: l.callegaro@inrim.it

the evaluated uncertainty. What is not acknowledged is that non-linearities (as well as asymmetrically distributed input values) *can also bias the measurand estimate*. JCGM 101 [2] automatically takes these effects into account by assigning state-of-knowledge probability distributions for the input quantities and propagating them through the measurement model to produce a probability distribution for the possible values of the measurand. JCGM 100, on the contrary, does not explicitly discuss this topic, and only considers a correction to the measurand estimate in an example given in an Annex (see [1], F.2.4.4). Thus, in general, both the estimates and the standard uncertainties provided by [1] differ from those of [2], the latter document being considered as the gold standard.

The strong non-linearity of models (1) and (2) suggests that the JCGM 100 approach might be inadequate.

Indeed, in GUM-6 [3] it is written "... it is not recommended to perform uncertainty calculations using logarithmic quantities in, for example, decibel. Logarithmic quantities should be converted to the equivalent linear format before performing uncertainty calculations" (JCGM-GUM6).

When P_{dB} has a normal distibution, P has a lognormal one. This assumption appears reasonable when considering the structure of a typical acoustic or radio-frequency power measurement. The signal is given by a number of devices in cascade (transducers, amplifiers, filters, detectors ...), each one characterised by a signal gain with an uncertainty. When the signal is expressed in dB, all device gains sum up with the original signal power; in the linear domain, instead, gains are multiplied. It is therefore reasonable that for dB quantities the additive central limit theorem can be applied, and for linear quantities the multiplicative one, resulting in normal and log-normal distributions of the quantity value respectively.

The talk will discuss exact conversion formulae for the distribution parameters, the mean and the standard deviation of power quantities, from the linear to dB representation to the linear one and vice versa, highlighting the bias introduced by the JCGM 100 approach.

References

[1] 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, 2008.

[2] 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, 2008.

[3] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Guide to the expression of uncertainty in measurement — Part 6: Developing and using measurement models. Joint Committee for Guides in Metrology, JCGM GUM-6:2020, 2020.

Efficient learning of the copula distribution using WGANs

Jörg Martin and Clemens Elster

Key words: Deep neural networks, Copula, Correlation, Generative models

1 Introduction

A key task in statistics is to describe the dependency structure between random variables. However, such descriptions are often influenced by non-linear transformations of the considered variates. For instance, consider the chemical concentrations c_1, c_2 of two substances. To find the dependency between both concentrations, given some measurements of both variates, one might compute their Pearson correlation. However, it could be argued that the log-concentration $\log c_1, \log c_2$ is a more natural quantity, as it often tends to be normally distributed. In general, there is no obvious relation between the Pearson correlation of the concentrations and the correlation of the log-concentrations which then poses the question which of the correlations is more the "correct" description of dependency.

The copula distribution is a powerful statistical tool to overcome such issues. By separating the joint distribution of random variables from their marginals it allows for a description of the dependence structure that is freed of arbitrary or misleading scalings and transformations. The key idea behind copulas is based on Sklar's theorem [Sklar(1959)] that states that the cumulative distribution function (cdf) of a multivariate random variable can be written as a composition of the cdfs of the marginals with the cdf C of a multivariate random variable with uniform marginals. C is known as the copula and the according distribution as the *copula distribution*. In

Jörg Martin
PTB, Abbestraße 2, 10587 Berlin e-mail: joerg.martin@ptb.de

Clemens Elster
Abbestraße 2, 10587 Berlin e-mail: clemens.elster@ptb.de

this talk we present a way to sample from the copula distribution using a Wasserstein Generative Adversarial Network (WGAN), a generative model from deep learning.

2 Method and results

Only for a few distributions the copula distribution is explicitly known [Nelsen(2017)]. When modelling the copula of a multivariate distribution often additional assumptions have to be made about the underlying structure [Ling(2022)]. Using a generative model, such as a WGAN [Gulrajani(2017)], to learn the distribution in a non-parametric fashion avoids such restrictions. However, we observe that the naive learning of such a model on the copula distribution leads to an unstable behavior. In fact, for a long part of the training time the samples of the trained WGAN only partially cover the support the underlying copula distribution. To fix this issue we propose a new model, which we call *CopWGAN*, whose loss function is enhanced by the addition of a regularization term

$$\mathscr{L}_{\text{CopWGAN}} = \mathscr{L}_{\text{WGAN}} + \frac{1}{n}\sum_{i=1}^{n} W_1(\Gamma^i; \text{Unif}([0,1]))\,, \tag{1}$$

that measures the Wasserstein-1-distance between the marginals Γ^i of the learned WGAN distribution and the uniform distribution. This term substantially enhances the training performance and leads to a good description of the dependency structure by our trained model, cf. Figure 1. We study our method along various multivarariate datasets and use different metrics to compare our results to competing approaches. We observed throughout all datasets that our CopWGAN approach performed comparable or better than other methods with respect to the considered metrics.

This talk is based on the preprint [Martin(2023)], that is currently under review.

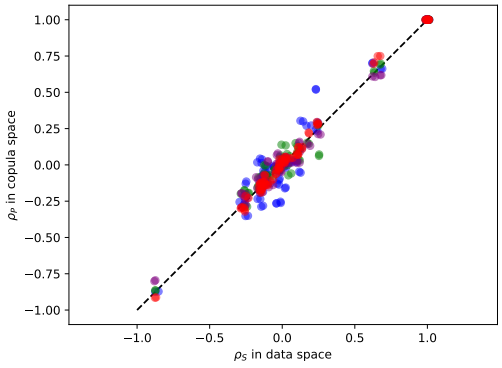


Fig. 1 Correlations of pairs of variates from the california housing dataset [Pace(1997)] in copula space for different methods according to the data (x-axis) and trained generative models (y-axis). Different generative models are depicted in different colors. Our method (CopWGAN) is depicted in red. The diagonal is shown as a dashed line.

3 References

[Martin(2023)] Efficient learning of the copula distributions using WGANs. Martin, J. and Elster, C. Preprint - currently under review.
[Nelsen(2017)] An introduction to copulas. Nelsen Roger, B. Springer Science & Business Media, 2017.
[Ling(2020)] Deep archimedean copulas. Ling, C. K. et al. Advances in Neural Information Processing Systems, 2020.
[Sklar(1959)] Fonctions de repartition an dimensions et leurs marges. Sklar, M. Publ. inst. statist. univ. Paris, 1959.
[Gulrajani(2017)] Improved training of wasserstein gans. Gulrajani, I. et al. Advances in neural information processing systems, 2017.
[Pace(1997)] Sparse spatial autoregressions. Pace, R. K. and Barry, R. Statistics & Probability Letters, 1997.

From “Wich” to “Why”: Interpretation map for Explainable Deep Learning based on Influence methods

Tommaso Torda, Andrea Ciardiello, Simona Gargiulo, Greta Grillo, Cecilia Voena and Stefano Giagu

Key words: Deep Learning, Explainable AI, Interpretable xAI, Segmentation, Medical Imaging

Abstract

In recent years, thanks to improved computational power and the collection of a multitude of data, AI has become a fundamental tool in both basic research and industry. Despite this very rapid development, AI, in particular deep neural networks, remains black boxes that are difficult to explain. This is a major limitation in the application of these algorithms in sensitive fields such as clinical diagnosis, where the robustness, transparency and reliability of the algorithm are indispensable for its use. In tasks involving the study of images, especially their classification, a multitude of xAI methods have been developed. However, in the medical field, saliency

Tommaso Torda
“Sapienza” University of Rome, Rome, Italy, e-mail: tommaso.torda@uniroma1.it

Andrea Ciardiello
“Sapienza” University of Rome, Rome, Italy, e-mail: andrea.ciardiello@uniroma1.it

Simona Gargiulo
“Sapienza” University of Rome, Rome, Italy, e-mail: simona.gargiulo@uniroma1.it

Greta Grillo
“Sapienza” University of Rome, Rome, Italy, e-mail: greta.grillo02@gmail.com

Cecilia Voena
“Sapienza” University of Rome, Rome, Italy, e-mail: cecilia.voena@roma1.infn.it

Stefano Giagu
“Sapienza” University of Rome, Rome, Italy, e-mail: stefano.giagu@uniroma1.it

map techniques without quantitative evaluation of the goodness of explanation are often used [Singh et al., 2020].

A very interesting class of xAI techniques is based on approximating the influence a train example has on the predictions made by the model. Theoretically, this influence should be calculated as the difference between the score of network trained without the example of interest, and the score of the network trained on the entire dataset. While from a conceptual point of view, influence is a simple thing to understand, from a computational point of view it quickly becomes unmanageable and necessarily requires approximations. Often, such approximations use the Hessian of the loss function with respect to the network parameters, but this is not always easy to calculate. Indeed, we rarely find ourselves at the end of a training cycle in an absolute minimum of the landscape, but often in a local minimum, which does not ensure that the Hessian matrix is defined as positive. There are methods where the approximation of the influence only involves the first derivatives of the loss function, one of the most promising being known as TracIn [Pruthi et al., 2020]. This method was originally presented on classification tasks but is fully generalisable to different tasks and architectures.

The aim of this work is to implement this xAI technique in a specific clinical problem, the segmentation of tumour brains in MRI, and also to provide information regarding the interpretation and robustness of the algorithm. Influencing methods are very powerful tools and seemingly easy to interpret by expert and non-expert users. However, when the task becomes difficult to solve, it can become complicated to interpret the results of the explanation. In the case of segmentation, this problem is even more pronounced as the influence is an average over all pixels of the predicted image, thus it is the collective behaviour of many classifications viewed in unison.

Algorithms such as TracIn answer the question "what is the train example that positively or negatively influenced the specific prediction?", but they do not give any information as to "why" that particular example influenced the prediction, i.e. on the basis of which features the choice is made. Our proposal is to calculate the influence only on areas that the network predicts with less certainty. To do this we need to calibrate [Guo et al., 2017] the neural network in such a way that we can interpret the output of the neural network as the probability of the prediction, in this way we reduce the number of features related to the influence metric. Then we would like to find a similarity metric to answer the question "which train example, had a positive influence on the prediction made by the model and why?". In fact, we expect that examples that have a positive influence also have features in common with the prediction. However, the individual features used by the neural network to complete the task are usually of little interest to the end user, as these are difficult for humans to interpret. The idea is therefore to use a modality importance technique [Jin et al., 2021] to extract which acquisition channel played a more important role in prediction. Based on this, a similarity metric between train examples and model predictions is constructed. This algorithm then needs to prove robust against different network architectures, so it is tested on different models to ensure that the explanation does not depend on them.

References

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberge. On calibration of modern neural networks. *PLMR*, 2017.

Weina Jin, Xiaoxiao Li, and Ghassan Hamarneh. One map does not fit all: Evaluating saliency map explanation on multi-modal medical images. *ICML*, 2021.

Garima Pruthi, Frederick Liu, Mukund Sundararajan, and Satyen Kale. Estimating training data influence by tracing gradient descent. *NeurIPS 2020*, 2020.

Amitojdeep Singh, Sourya Sengupta, and Vasudevan Lakshminarayanan. Explainable deep learning models in medical image analysis. *MDPI*, 2020.

Subjective vs objective assembly complexity assessment: a comparative study in a Human-Robot Collaboration framework

Elisa Verna[1] and Stefano Puttero[2] and Gianfranco Genta[3] and Maurizio Galetto[4]

Key words: manufacturing complexity, perceived complexity, assembly, quality

Extended Abstract

The impact of manufacturing complexity on company performance can be significant, affecting productivity, efficiency, affordability, and quality if not managed correctly. Assessing and managing manufacturing complexity is a multifaceted task that involves both objective and subjective features, such as product complexity, assembly sequence, operator factors, and operation/management strategies. This study proposes a structured methodology to assess the perceived complexity of human-robot collaboration assembly processes. The methodology is based on 16 assembly complexity criteria and a multi-expert decision-making method for evaluation. Operators assign importance scores and agreement levels to each criterion using a five-level ordinal scale, and the linguistic data is processed using the Multi-Expert Multi-Criteria Decision Making (ME-MCDM) method [Yager(1993)]. This approach combines linguistic information provided for non-equally important criteria using maximum, minimum, and negation operators to obtain an overall synthetic linguistic value of perceived complexity using fuzzy logic. The proposed approach provides an assessment of perceived complexity at both individual and overall levels, aggregating all individual complexity assessments by the operator Ordered Weighted Average (OWA) [Yager(1993); Filev(1994)].

[1] Elisa Verna
DIGEP, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: elisa.verna@polito.it
[2] Stefano Puttero
DIGEP, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: stefano.puttero@polito.it
[3] Gianfranco Genta
DIGEP, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: gianfranco.genta@polito.it
[4] Maurizio Galetto
DIGEP, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: maurizio.galetto@polito.it

The proposed approach for assessing perceived complexity of assembly is compared with a purely objective assessment method, firstly proposed by Sinha et al. [Sinha(2012)]. This model was validated in various studies, and its effectiveness in quantifying the complexity of industrial products was demonstrated [Verna(2022)]. It is based on the molecular orbital theory and is applied to the engineering domain to analyse the complexity of cyber-physical systems. The model represents a cyber-physical system as several connected components where each component can be thought of as an atom, and the interfaces between them as inter-atomic interactions or chemical bonds. The complexity of the assembly is defined as the combination of three complexity components: handling complexity (C_1), connections complexity (C_2), and topological complexity (C_3), as follows $C = C_1 + C_2 \cdot C_3$. This objective model, based on structural characteristics of the assembly process, was used as a reference model for the subjective complexity model.

The comparison between subjective and objective assessment of complexity was performed in a real-world production environment, using a human-robot collaboration process for manufacturing custom electronic boards with different levels of complexity. The results showed a significant correlation between individual perceived complexity and objective complexity, indicating that the proposed perceived complexity model can be linked to the objective model. As structural complexity increases, higher levels of individual perceived complexity become more likely, but the variability in perceived complexity varies with structural complexity. These findings suggest that individual operator ability and cognitive factors, such as training, knowledge, and cultural and organisational factors, play a role in perceived complexity and require further investigation. The study also suggests that using perceived complexity to assess assembly complexity is suitable for low- and medium-complexity products, but not for high-complexity products, where objective complexity models may be more appropriate, since after a certain point operators do not distinguish between different levels of complexity.

The proposed methodology and data analysis approach offer a new perspective on assessing perceived complexity, relying solely on synthesis operators and statistical tools suitable for categorical data. Engineers can use the study's results to minimise perceived complexity and ensure alignment between perceived and objective complexity.

References

1. Filev, D., Yager R.R.: Essentials of Fuzzy Modeling and Control. Sigart Bulletin **6**(4), 22–23 (1994).
2. Sinha, K., de Weck O.L., Onishi M., Maurer M., Kirner K., Lindemann U.: Structural Complexity Metric for Engineered Complex Systems and Its Application. In: Gain Competitive Advantage by Managing Complexity: Proceedings of the 14th International DSM Conference Kyoto, Japan, pp. 181–94 (2012).
3. Verna, E., Genta G., Galetto M., Franceschini F.: Defect Prediction for Assembled Products: A Novel Model Based on the Structural Complexity Paradigm. International Journal of Advanced Manufacturing Technology **120**(5–6), 3405–3426 (2022).
4. Yager, R.R.: Non-Numeric Multi-Criteria Multi-Person Decision Making. Group Decision and Negotiation **2**(1), 81–93 (1993).

How do asymmetric measurement distributions affect risks in conformity assessment?

Stephen L R Ellison[1]

Key words: conformity assessment, consumer risk, producer risk, distribution asymmetry

1. Introduction

Product conformity assessment is vital for a wide range of manufacturing, health and regulatory activities. For example, the nutritional composition of foods, levels of environmental contaminants in surface waters, the dosage of active ingredients in pharmaceuticals, the levels of biological indicators in clinical patients, thermal yield of fuels, and the geometric characteristics of engineered parts, all rely on measurement results falling within defined tolerances. As with all measured values, the results have associated uncertainties, which must be taken into account by accredited testing laboratories [1]. A number of guides and standards have been developed to assist laboratories in setting suitable decision criteria that take due account of measurement uncertainty (see, for example, references [2-6]).

An important consideration in setting decision rules is the control of so-called producer and consumer risks – respectively, the probabilities of falsely rejecting or falsely accepting a test item. Typical guidance provides for low risks to both parties. Detailed guidance on the calculation of producer and consumer risks has been published by JCGM [7], and simpler explanations have been included in guidance for testing laboratories [2,5].

Conformity assessment documentation often assumes normal, or at least symmetric, distributions for both the production process and the measurement process. In addition, measurement uncertainty is often assumed to be close to constant across the range of interest, an assumption that leads to simple symmetric 'guard bands' for decision rules [2]. These are often reasonably safe assumptions where product or other tolerances are narrow relative to the expected mean value and where measurement uncertainties are small. Often, however, and particularly for

[1] LGC limited, Queens Road, Teddington, Middlesex TW11 0LY, UK
e-mail: s.ellison@lgcgroup.com

natural processes or contaminant emission monitoring, the process can generate material with a wide range of possible values, and measurement or sampling challenges can lead to comparatively large uncertainties near control limits [8]. This can require the use of asymmetric uncertainty intervals [5, 8] and corresponding adjustments to decision rules. It then becomes more challenging to provide clear guidance for laboratories on the choice and implementation of decision rules.

2. Allowing for asymmetry in conformity assessment guidance

This study examines the effect on producer and consumer risks of selected departures from symmetry in the distribution of both process output and of measurement results. The study uses a combination of numerical methods, including numerical integration and simulation. The challenges of comparing and generalising results across different distributional assumptions and scenarios will be discussed, and results for a number of scenarios will be presented, together with some considerations for future guidance.

References

1. ISO/IEC 17025:2017, General requirements for the competence of testing and calibration laboratories. ISO, Geneva, 2017.
2. ILAC-G8:09/2019 Guidelines on Decision Rules and Statements of Conformity, International Laboratory Accreditation Cooperation, Guidance, 2019. [Online]. Available: https://ilac.org/.
3. ASME, B89.7.3.1-2001, Guidelines for Decision Rules: Considering Measurement Uncertainty in Determining Conformance to Specifications. ASME, 2001
4. ISO 14253-1:2017, Geometrical product specifications (GPS) – Inspection by measurement of workpieces and measuring equipment – Part 1: Decision rules for verifying conformity or nonconformity with specification. ISO, Geneva, 2017
5. A. Williams and B. Magnusson (eds.) Eurachem/CITAC Guide: Use of uncertainty information in compliance assessment (2nd ed. 2021). ISBN 978-0-948926-38-9. Available from www.eurachem.org.
6. EUROLAB Technical Report No.1/2017, Decision rules applied to conformity assessment. Eurolab, 2017. Available from www.eurolab.org
7. BIPM, IEC, IFCC, et al., Evaluation of measurement data – The role of measurement uncertainty in conformity assessment, JCGM 106:2012. BIPM, 2012.
8. I. Kuselman, S. Shpitzer, F. Pennecchi, and C. Burns. Investigating out-of-specification test results of mass concentration of total suspended particulates in air based on metrological concepts - a case study. Air Qual Atmos Health, 5:269–276, 2012. DOI: 10.1007/s11869-010-0103-6
9. M H Ramsey, S L R Ellison and P Rostron (eds.) Eurachem/EUROLAB/ CITAC/Nordtest/AMC Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches. Second Edition, Eurachem (2019). ISBN (978-0-948926-35-8). Available from www.eurachem.org.

PyES - an open source software for the computation of in solution and precipitation equilibria

Lorenzo Castellino and Eugenio Alladio and Silvia Berto and Demetrio Milea

Key words: Software, Open-Source, Speciation, Analytical Chemistry

1 Science and Computers

Computer tools have increasingly become a necessity in the world of modern research. From simple spreadsheets applications, such as Microsoft Excel, to more elaborate instruments for solving complex computation tasks scientists has to interface with a computer that assists him in their tasks. Unfortunately, more often then not these tools are not readily available or are scarcely supported becoming "abandonware". From a survey conducted among the participants of the Network for Equilibria and Chemical Thermodynamics Advanced Research, COST ACTION - NECTAR CA18202(NEC, 2020), a European network of researchers dedicated to the study of thermodynamic equilibria, it has been noted how scarce and inadequate the available software to support the study of stability constants is.

It was therefore decided to re-write ES4, a terminal-based computer program originally written by Prof. Silvio Sammartano from the Università degli Studi di Messina and his co-workers in the last century using the BASIC programming language (Arena et al, 1979) (Maggiore et al, 1976) (Daniele et al, 1984) (Casale et al,

Lorenzo Castellino
University of Turin, Via P. Giuria 7, 10125, Torino, Italy e-mail: lorenzo.castellino@unito.it

Eugenio Alladio
University of Turin, Via P. Giuria 7, 10125, Torino, Italy e-mail: eugenio.alladio@unito.it

Silvia Berto
University of Turin, Via P. Giuria 7, 10125, Torino, Italy e-mail: silvia.berto@unito.it

Demetrio Milea
Università degli Studi di Messina - CHIBIOFARAM, Viale Ferdinando Stagno d'Alcontres 31, 98166, Messina, Italy e-mail: dmilea@unime.it

1989) (De Stefano et al, 1993), created with the aim of solving chemical equilibria in solutions.

2 PyES

Building on its predecessor we developed PyES a new, open-source, practical, graphical and multi-platform Python application. Our implementation has reached feature parity with the original ES4, allowing the simulation of titrations and the computation of species distribution of system of a theoretically infinite number of chemical species. All the stability constants used in the computation can be dynamically adjusted to correct for the effect of the ionic strength through an expanded Debye-Hückel equation.

In addition to the revised logic for these calculations, this new version allows the inclusion of precipitable species in the system, empowering the user to investigate more complex systems. Finally, given the uncertainties for the input data, these are propagated in the calculation, allowing to report the results with their respective fiduciary range.

Various tests were conducted considering various real chemical systems, comparing results of our program with both ES4 and other analogous software to assess the correctness of our code.

References

(2020) NECTAR COST. https://www.cost-nectar.eu/index.html

Arena G, Rizzarelli E, Sammartano S, Rigano C (1979) A non-linear least-squares approach to the refinement of all parameters involved in acid—base titrations. Talanta 26(1):1–14, DOI 10.1016/0039-9140(79)80146-0

Casale A, Destefano C, Sammartano S, Daniele P (1989) Ionic-strength dependence of formation constants—XII A model for the effect of background on the protonation constants of amines and amino-acids. Talanta 36(9):903–907, DOI 10.1016/0039-9140(89)80028-1

Daniele PG, Ostacoli G, Rigano C, Sammartano S (1984) Ionic strength dependence of formation constants. Part 4. Potentiometric study of the system Cu2+-Ni2+-citrate. Transition Metal Chemistry 9(10):385–390, DOI 10.1007/BF00637025

De Stefano C, Mineo P, Rigano C, Sammartano S (1993) IONIC STRENGTH DEPENDENCE OF FORMATION CONSTANTS. XVII: THE CALCULATION OFEQUILIBRIUM CONCENTRATIONS AND FORMATION CONSTANTS. Annali di chimica 83(5-6):243–277

Maggiore R, Musumeci S, Sammartano S (1976) Computer calculation of distribution diagrams. Talanta 23(1):43–44, DOI 10.1016/0039-9140(76)80007-0

Callendar Van Dusen interpolation by means of Piecewise Constrained Least Squares with nullspace method – an update

Graziano Coppa[1]

Key words: temperature metrology, Callendar-Van Dusen, spline

1. Introduction and scope

The Callendar-Van Dusen (CVD) equation (Van Dusen, 1925) has been used for decades in the interpolation of platinum resistance thermometers (PRTs) calibration points. This empirical equation, which represents the extension to $t < 0$ ℃ of the original Callendar equation (Callendar, 1899), can be written in several forms, one of the most used being

$$R = R0\ (1 + A\ t + B\ t^2 - 100\ C\ t^3 + C\ t^4), \quad (1)$$

where *R* is the resistance of the PRT at temperature *t*, *R0* is the resistance of the PRT at 0 °C, and *A*, *B* and *C* are coefficients of the interpolation. For temperatures above 0 °C, the equation simplifies as the original Callendar equation:

$$R = R0\ (1 + A\ t + B\ t^2). \quad (2)$$

There are currently no prescriptions on the mathematical method of interpolation for PRT calibration points, therefore there are a number of sub-optimal methods in use. Some of those are geared towards the optimization of subranges, while the well-known Ordinary (or Weighted) Least Squares, while providing the best overall regression in terms of minimization of the loss function, can run into numerical issues.

This work inquires into a novel method of interpolation that minimizes the squared sum of fit residuals along all the calibration range, therefore representing a means of uncertainty reduction in an overall calibration budget.

2. Method

[1] Graziano Coppa
INRiM, Strada delle Cacce, 91, Torino, e-mail: g.coppa@inrim.it

The interpolation system presented in this work makes use of Piecewise Constrained Least Squares (PCLS) and consists in the familiar least squares method to which a system of constraints is applied.

Basically, the whole calibration range is interpolated by a spline, made up by the two forms of the CVD equations (Eqn. 1 for $t < 0$ ℃ and Eqn. 2 for $t > 0$ ℃). The constraints set is composed by the continuity conditions of Eqns 1 and 2, as well as their first derivative, at $t = 0$ ℃. Other constraints are given by the equality conditions of coefficients of Eqns 1 and 2.

The problem can be written as the minimization of the objective function:

$$\|Xb - y\|^2$$

constrained by

$$Cb = d,$$

where *b* is the unknown vector of CVD coefficients, dimension *n*, of which only 3 are linearly independent; *X* is the Vandermonde matrix of the observations (Shores, 2018), dimensions $m \times n$, where *m* is the number of calibration points; $y=R_i/R0$ is the observations vector (resistance measurements), dimension *m; C* is the constraints matrix, dimension $p \times n$, with *p* representing the number of constraints*; d* is the response vector of constraints, dimension *p*.

3. Conclusions

The theoretical construction of the system, based on the null-space method (Benzi *et al.*, 2005), will be shown, as well as its practical implementation (Excel and R functions). Its computational advantages over the direct method will be shown, as well as the formal expression of the solution in case of heteroscedastic errors on the random variable (with variance-covariance matrix). Also, a comparison with other methods of interpolation already in use will be provided using case studies.

This work extends and completes the work presented at MSMM 2021.

References

1. Benzi M, Golubt GH, Liesen J. 2005. Numerical solution of saddle point problems. *Acta Numerica*, 14: 1–137. https://doi.org/10.1017/S0962492904000212.
2. Callendar HL. 1899. XIII. Notes on platinum thermometry . *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 47(285): 191–222. https://doi.org/10.1080/14786449908621251.
3. Shores TS. 2018. *Applied Linear Algebra and Matrix Analysis*. *Springer*. Springer International Publishing: Cham.
4. Van Dusen M V. 1925. Platinum-resistance thermometry at low temperatures. *Journal of the American Chemical Society*, 47(2): 326–332. https://doi.org/10.1021/ja01679a007.

Obsidian sourcing by combining SEM images and machine learning

Marco Coïsson, Elena S. Olivetti, Patrizio Ansalone, Luca Martino, Enzo Ferrara

Key words: Machine learning, classification

Provenance studies of obsidian archaeological artefacts allow to understand the links among ancient populations, their trade routes and their socio-economic context. In the Mediterranean region, the number of geological sources of obsidian is limited, and although chemical analysis remains the reference technique for sourcing [Orange(2017)], other techniques have been investigated for screening purposes. Among these, the examination in various samples of granulometry and morphology of iron-based microphenocrysts embedded within the vitreous matrix allowed to validate their provenance according to studies of their associated magnetic properties [Ferrara(2019)]. Building upon these bases, we aim at exploiting all the microcrystalline phases typically present in obsidian rock, and visible in scanning electron microscopy (SEM), as a fingerprint to assess the provenance of geological samples [Burton and Krinsley(1987)]. The identification of the origin of a sample is a labelling process suited for supervised machine learning (ML) methods. In this work, we exploit them for classifying the sources of obsidian samples coming from different Mediterranean islands: Lipari, Sardinia, Pantelleria, Palmarola and Milos. To this goal, we collected SEM backscattered electron images and Energy Dispersive X-ray Spectroscopy (EDX) data of geological samples. In typical SEM images (Figure 1), features like microphenocrysts of different size, shape and contrast are present. Although not specific, their average size, density, number, etc., can be extracted and fed to a ML algorithm that will find relationships among them uncatchable by the human eye. The challenging objective is to identify the provenance of an obsidian sample within a limited set of possible sources using only SEM images, an easily accessible and minimally invasive technique, by comparison with a reference database of pictures of known origin. However, in addition to morphological features extracted from SEM images, relative intensities of EDX peaks of selected chemical elements can also be used, at the expense of a reasonable increase in anal-

INRIM, strada delle Cacce 91, 10135 Torino (TO), Italy, e-mail: m.coisson@inrim.it

ysis time, to enrich the set of parameters exploited by a suitable ML algorithm (e.g. neural network, random forest), to properly classify the provenance of the obsidian samples of the Mediterranean region. A critical analysis of the best parameters and methods is proposed. The validation of the approach is studied with samples not used to train the ML model. The comparison between the algorithms performance on SEM pictures only or on SEM and EDX data and the scalability of this approach to larger datasets and to a more comprehensive list of obsidian sources will be discussed. Perspectives of application of the method to archaeological artefacts will be explored by analysing its robustness with a set of pictures of fragments raw surfaces.

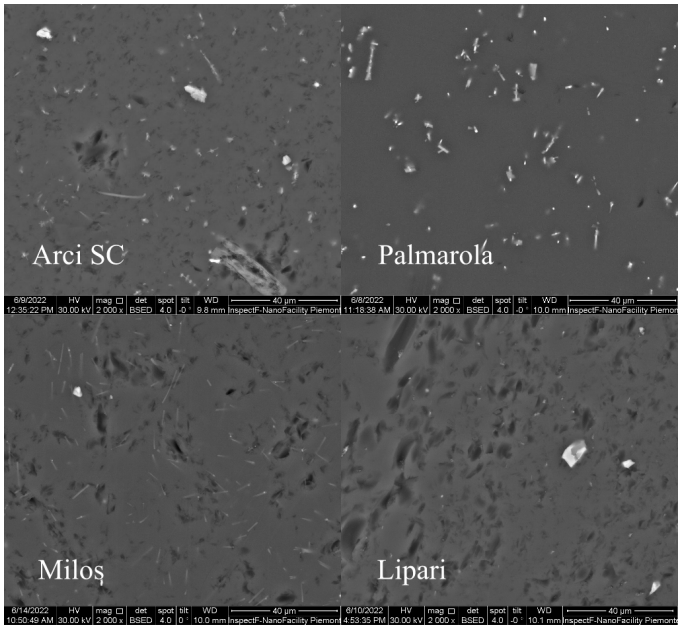


Fig. 1 SEM backscattered electron images of obsidian fragments of different Mediterranean sources. The magnification is the same for all the pictures.

References

[Burton and Krinsley(1987)] Burton J, Krinsley D (1987) Obsidian provenance determination by back-scattered electron imaging. Nature (6113):585–587

[Ferrara(2019)] Ferrara ea E (2019) Investigating the provenance of italian archaeological obsidian tools based on their magnetic properties. Archaeological and Anthropological Sciences 11(7):3329–3341

[Orange(2017)] Orange Mea (2017) On sourcing obsidian assemblages from the mediterranean area: Analytical strategies for their exhaustive geochemical characterisation. Journal of Archaeological Science: Reports 12:834–844

Physiological variability in brain electric conductivity: correcting the effect of the age for the detection of pathological alterations

Sébastien Marmin and Alessandro Arduino and Matteo Cencini and Marta Lancione and Laura Biagi and Michela Tosetti and Luca Zilberti

Key words: Electric properties tomography, Quantitative imaging, Magnetic resonance imaging, white matter disorders

Overview

The aim of this work is to analyze the physiological variability of electric conductivity in the brain measured through electric properties tomography (EPT) using Magnetic Resonance Imaging (MRI). Electric conductivity maps of 27 healthy volunteers and 45 patients with pathologies of the white matter (WM) were estimated based on water content [Michel et al. (2017)] measured using MR Fingerprinting [Cencini et al. (2022)] (Fig. 1A).

For each subject, tissue class segmentation was performed to obtain a white matter mask (Fig. 1B) and the median conductivity within this region was measured. A non-linear mixed effect model was used to identify sources of variability. The inter-subject standard deviation of the median WM conductivity was estimated to 11 mS/m. The results in Fig 1C showed strong dependence of WM conductivity with age.

We reported higher WM conductivity in patients with respect to control subjects. To distinguish pathological changes from physiological variability, we established a threshold value of 22.5 mS/m more than the age-dependent average of observed

Sébastien Marmin
Laboratoire national de métrologie et d'essais, 29 avenue Hennequin, 78190 Trappes, France, e-mail: sebastien.marmin@lne.fr

Alessandro Arduino and Luca Zilberti
Istituto Nazionale di Ricerca Metrologica, Strada delle Cacce, 91 10135 Torino, Italy e-mail: a.arduino@inrim.it, l.zilberti@inrim.it

Matteo Cencini
National Institute for Nuclear Physics (INFN), Largo Bruno Pontecorvo 3, 56127 Pisa, Italy e-mail: matteo.cencini@pi.infn.it

Marta Lancione, Laura Biagi and Michela Tosetti
Fondazione IRCCS Stella Maris, Viale del Tirreno, 331, 56128 Pisa, Italy e-mail: marta.lancione@fsm.unipi.it, laura.biagi@fsm.unipi.it, michela.tosetti@fsm.unipi.it

in the control group (Fig. 1D). These findings are important for the development of biomarkers and personalized medicine using EPT, and demonstrate the potential of metrology in this field.

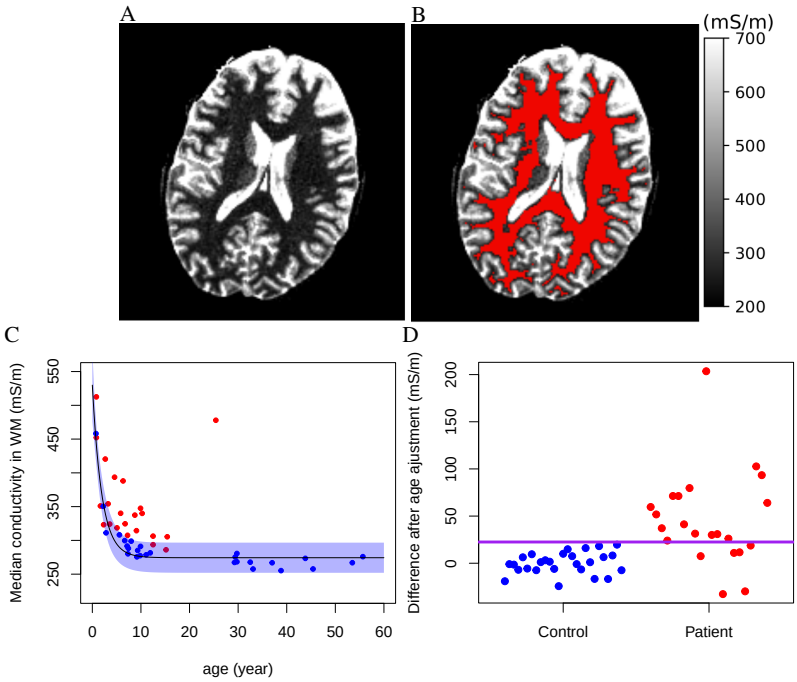


Fig. 1 A. Conductivity map on a transverse section for one subject. B. Segmentation of the white matter. C. Median conductivity in WM for each subject with respect to age (black line); The blue area is the prediction envelope at level 95%. D. For each subject, difference with the age-dependent average calculated from the control group (threshold in purple).

Acknowledgements This work was developed within the QUIERO project. 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. This study was also partially supported by grant RC and the 5x1000voluntary contributions to IRCCS Fondazione Stella Maris, funded by the Italian Ministry of Health.

References

Cencini, Matteo et al. (2022). "Fast high-resolution Electric Properties Mapping using three-dimensional MR Fingerprinting based water fraction estimation (MRF-EPT)". In: *Proceedings of the Joint Annual Meeting ISMRM-ESMRMB, London*.

Michel, Eric, Daniel Hernandez, and Soo Yeol Lee (2017). “Electrical conductivity and permittivity maps of brain tissues derived from water content based on T1-weighted acquisition”. In: *Magnetic Resonance in Medicine* 77.3, pp. 1094–1103. DOI: 10.1002/mrm.26193.

A machine learning approach for the estimation of magnetic nanoparticles specific loss power

Riccardo Ferrero[1] Marta Vicentini[1] and Alessandra Manzin[1]

Key words: Magnetic Hyperthermia, Magnetic nanoparticles, Micromagnetism, Machine Learning, Numerical models.

1. Introduction

In the last decade, magnetic nanoparticles (MNPs) made of iron oxides have been intensively studied for potential application in cancer therapies based on magnetic hyperthermia [Vassallo (2023)]. One of the quantification indexes typically used to express the capability of MNPs to release heat is the specific loss power (SLP), which is the power dissipated per unit mass of magnetic material. The SLP strongly depends on the MNP hysteresis losses, which can be estimated by numerically solving the Landau-Lifshitz-Gilbert (LLG) equation. The MNP hysteresis losses depend on a plethora of parameters, including magnetic properties, shape, size and state of aggregation of the MNPs, as well as the frequency and amplitude of the exciting magnetic field. However, the numerical integration of the LLG equation is computationally very intensive and analysing all the possible combination of MNPs properties and magnetic field conditions is a daunting task.

Machine learning (ML) techniques are a powerful tool in the analysis of large amount of data and very well suited to develop regression and predictive models depending on a high number of parameters. Leveraging the large dataset of micromagnetic simulation results, obtained by numerically solving the LLG equation for various types of MNPs under different experimental conditions, we present a comparison of the predictive performance of several regression models and ML techniques in the evaluation of MNP hysteresis losses and SLP.

2. Methodology and results

The initial dataset is obtained using the micromagnetic numerical solver we developed [Ferrero (2021)]. The magnetization dynamics of each MNP, assumed to be in the single-domain state, is modelled by the LLG equation, with the inclusion of thermal effects, using the Langevin dynamics.

[1] Riccardo Ferrero e-mail: r.ferrero@inrim.it · Marta Vicentini e-mail: m.vicentini@inrim.it ·Alessandra Manzin, e-mail: a.manzin@inrim.it
INRiM, Strada delle Cacce 61, Torino, Italy

Using the numerical model we analyse the magnetic hysteresis behaviour of spherical MNPs made of different materials, and with a diameter ranging between 10 nm and 40 nm. The applied magnetic field frequency varies between 100 kHz and 1 MHz, while the field amplitude varies between 5 kA/m and 50 kA/m for a total of nearly 800 calculated hysteresis loops. The hysteresis loops are characterized by three fundamental parameters, i.e. the area of the loop for the derivation of the SLP, the coercive field and the magnetic remanence.

The inputs for the predictive models are the magnetic properties and size of the MNPs, and the magnetic field amplitude and frequency.

Two ML approaches from the SkLearn library [Pedregosa (2011)] were tested: a Random Forest (RF) Regression algorithm and a shallow Neural Network (NN) with few layers. In Fig. 1 we report the preliminary analysis performed on the two models on randomly chosen test dataset, not used in the training process, which amounts to 10% of the total original data.

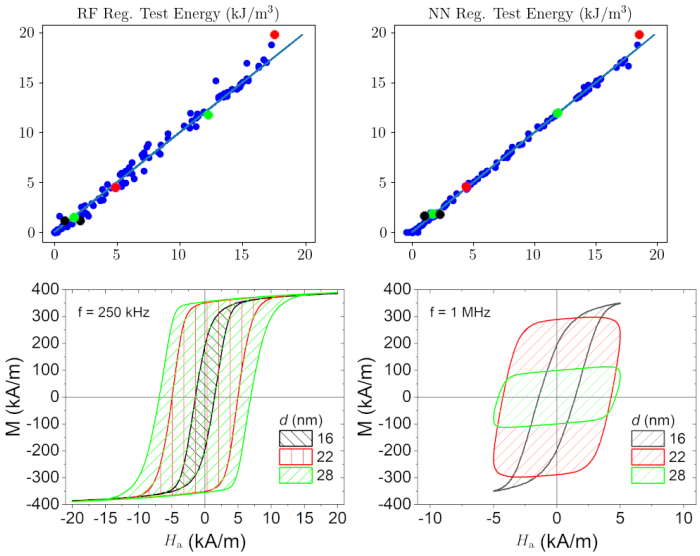


Figure 1: Regression analysis on a subset of the simulated dataset with two different regression methods: a Random Forest (RF) and a 3 layer Neural Network (NN). Below the computed magnetic hysteresis loops for a combination of different parameters to illustrate the variety of hysteresis loops shape.

References

1. Ferrero, R., Manzin, A.: Adaptive geometric integration applied to a 3D micromagnetic solver, Jour. Magn. Magn. Mat., 518, 167409, (2021).
2. Pedregosa , F. et al.: Scikit-learn: Machine Learning in Python. JMLR **12**, 2825-2830, (2011).
3. Schaffer, S., Mauser; N. J., Schrefl, T., Suess, D., Exl, L.: Machine Learning Methods for the Prediction of Micromagnetic Magnetization Dynamics, IEEE Trans. Magn. **58**, 2, 1-6. (2022).
4. Vassallo, M. et al.: Improvement of Hyperthermia Properties of Iron Oxide Nanoparticles by Surface Coating. ACS Omega, **8**, 2, 2143–2154, (2023).

Separation of effects associated with measurement data

Alistair Forbes

Key words: Bayesian inference, multiple effects model

1 Abstract

The Guide to the Expression of Uncertainty in Measurement (GUM) gives, through the construction of uncertainty budgets and the law of propagation of uncertainty (LPU), a way of adding up uncertainty contributions from various sources in order to assign an uncertainty associated with a measurand. In this paper we are interested in the reverse process, given (redundant) measurement data, what can be said about the effects that contributed to the measurement data. We consider a response model of the form

$$\eta = C\alpha + \sum_{k=1}^{K} \delta_k, \quad \delta_k \in \mathrm{N}(\mathbf{0}, V_k),$$

where the response η depends on α, often the parameters of primary interest, and random effects δ_k. Given an observed response $\mathbf{y} \in \eta$, we wish to make inferences about α but also δ_k. A common situation has two effects sources, $K = 2$, where δ_1 models some non-ideal or random behaviour of the system under study, and δ_2 models effects associated with the measurement system. Assuming a noninformative prior $p(\alpha) \propto 1$ for α, the posterior distribution for α is given by

$$\alpha \sim \mathrm{N}(\mathbf{a}, V_{\mathbf{a}}), \quad \mathbf{a} = (C^\top V^{-1} C)^{-1} C^\top V^{-1} \mathbf{y}, \quad V_{\mathbf{a}} = (C^\top V^{-1} C)^{-1}, \quad V = \sum_{k=1}^{K} V_k.$$

Alistair Forbes
National Physical Laboratory, Hampton Road, Teddington, UK, TW11 0LW e-mail: alistair.forbes@npl.co.uk

Setting $\mathbf{r} = \mathbf{y} - C\mathbf{a}$, it is possible to derive Gaussian posterior distributions for δ_k involving straightforward expressions involving $\mathbf{r}$, V_k and V, $k = 1, \ldots, K$.

We are also interested in the case where the variance matrices V_k are only partially characterised, e.g., $V_k|\sigma_k^2 = \sigma_k^2 V_{k,0}$ where $\phi_k = 1/\sigma_k^2$ has a gamma prior distribution. We explore the use of the generalised singular value decomposition and other approaches in deriving computationally efficient algorithms for providing information about the posterior distribution $p(\boldsymbol{\phi}|\mathbf{y})$. The approaches will be illustrated in applications in coordinate metrology and microflow measurement.

An efficient way to generate synthetic spin echo signals by the extended phase graph

Zeinab. Al-Siddiqui[1], Nadia. A. S. Smith[1] and Matt. G. Hall[1]

Key words: quantitative MRI, extended phase graphs, RF pulses, dephasing, uncertainty propagation

Magnetic resonance imaging (MRI) is a non-invasive imaging modality used by clinicians to diagnose and monitor a variety of medical conditions. MRI utilizes both a powerful magnetic field and radiofrequency pulses to generate detailed images of the internal structures of the body. Furthermore, T2-mapping is a quantitative MRI (qMRI) technique used to calculate the transverse relaxation times (T2) of different tissues and display them for each voxel on a parametric map. Although this technique provides reliable quantitative information, there is little work on the uncertainty of measurements in MRI and their quantitative impact. Here, we aim to simulate signals from MRI pulse sequences, specifically the spin echo signal using the extended phase graph (EPG) principle.

EPG [Henning (1991); Henning (1991); Weigel (2015)] is commonly used for simulating signals obtained from MRI pulse sequences. This technique applies a Fourier approach to solving the Bloch equations [Weigel (2015)]. The EPG method provides efficient modelling of MR sequences by isolating different events that lead to echo formation and modelling each one as a matrix containing the physics of excitation and evolution. The events, including the effect of gradients, RF pulses, relaxation and dephasing are characterised by matrix operations; therefore, efficiently quantifying the intensities of the echo signal. This approach allows us to build a modular flexible framework, whereby uncertainty propagation can then be effectively performed in a flexible, efficient, and extensible manner.

By applying the EPG concept from [Weigel (2015)], we expressed each component (effect of gradients, RF pulses, relaxation, and dephasing) using matrix notation in python. We then simulated both noiseless (figure 1a) and noisy (figure 1b) spin echo signal data in python. The aim is to fit the T2-weighted signals to extract T2 values utilizing the qMRI Standardisation Processing (QSP) algorithms (Clarke et al., T2 or not T2? A new tool for consistent processing of qMRI parameters. Abstract submitted to MSMM (2023)) to aid in the standardisation of qMRI parameters.

The EPG method is both an efficient technique in modelling a range of MR signals and enabling flexibility in uncertainty propagation. This concept could prove to be very useful in qMRI and can potentially be applied for the use of QSP.

[1] National Physical Laboratory, Hampton Rd, Teddington, TW11 0LW, UK
e-mail: zeinab.al-siddiqui@npl.co.uk, nadia.smith@npl.co.uk, matt.hall@npl.co.uk

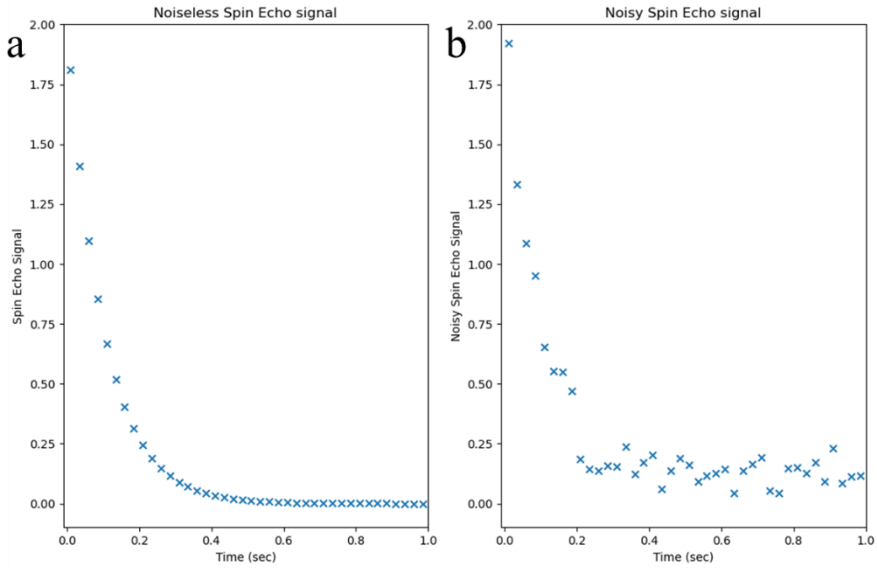


Figure 1: Simulated Spin Echo Signal using the EPG concept. (a) A clean or noiseless representation of the spin echo signal and (b) a noisy spin echo signal, where a 2D gaussian distribution was added to both real and imaginary components of the spin echo vector.

References

1. Hennig J. Echoes—how to generate, recognize, use or avoid them in MR-imaging sequences. I: Fundamental and not so fundamental properties of spin echoes. Concepts Magn. Reason. **3**:125–143 (1991).
2. Hennig J. Echoes—how to generate, recognize, use or avoid them in MR-imaging sequences. II: Echoes in imaging sequences. Concepts Magn. Reason. 3:179–192 (1991).
3. Weigel M. Extended phase graphs: dephasing, RF pulses, and echoes—pure and simple. J. Magn. Reson. Imaging. 41:266–295 (2015).

GUM-compliant uncertainty propagation for deep neural networks

Björn Ludwig, Barbara Jung and Sascha Eichstädt

Key words: guide to the expression of uncertainty in measurement, measurement uncertainty, uncertainty propagation, deep neural networks, software

1 Introduction

Advances in algorithms and computing capacities continue to increase the popularity of neural networks. In particular, their use in safety-critical applications (e.g. aviation, transport, telecommunications) poses new challenges with regard to the consideration of uncertainties in their input data. If these consist of sensor measurements, metrology provides a standard reference for dealing with such uncertainties in the internationally harmonised and recognised "Guide to the expression of uncertainty in measurements" (GUM). In this work, the key results of a solid mathematical framework established at Physikalisch-Technische Bundesanstalt for applying the principles of the GUM to simple neural networks, so-called multilayer perceptrons, is presented. The results are applied to sample network architectures equipped with a choice of three distinct activation functions, of which one was developed specifically to meet the needs of metrological use cases. Neural network robustness verification served as a theoretical, metrological use case to apply the results. FAIR software is made available to enable emergent research (Ludwig, 2023).

Björn Ludwig, e-mail: bjoern.ludwig@ptb.de
Barbara Jung, e-mail: barbara.jung@ptb.de
Sascha Eichstädt, e-mail: sascha.eichstaedt@ptb.de
Physikalisch-Technische Bundesanstalt, Abbestr. 2–12, 10587 Berlin

2 Law of propagation of uncertainty for deep neural networks

One of the key result of the presented work is the deduction of a law of propagation of uncertainty for deep neural networks (DNN) based on Supplement 2 of the Guide to the expression of uncertainty in measurement (GUM S2) (JCGM, 2011):

Theorem 1 (law of propagation of uncertainty for deep neural networks). *Let f be a DNN in the shape of an already trained, fully-connected, feed-forward multi-layer perceptron (MLP), with completely certain parameters, $\ell \in \mathbb{N}$ the number of layers in addition to the input layer, $n^{(0)}, \ldots, n^{(\ell)} \in \mathbb{N}$ the number of neurons per layer, $x^{(0)} = \left(x_j^{(0)}\right)_{j=1}^{n^{(0)}} \in \mathbb{R}^{n^{(0)}}$ the input vector comprised of the best estimates of the input quantities X_j, $\mathbf{U}_{x^{(0)}}$ the associated covariance matrix and $W^{(i)}, i = 1, \ldots, \ell$ the i-th layer's weight matrix. Further let each activation function $h_k^{(i)} \colon \mathbb{R} \to \mathbb{R}, i = 1, \ldots, \ell, k = 1, \ldots, n^{(i)}$ be continuously differentiable on $\mathbb{R}$. Then the law of propagation of uncertainty according to GUM S2 (JCGM, 2011, paragraph 6.2.1.3) can be applied to propagate the covariance matrix $\mathbf{U}_{x^{(0)}}$ associated with the inputs to compute the covariance matrix $\mathbf{U}_{x^{(\ell)}}$ associated with the outputs $x^{(\ell)}$ as*

$$\mathbf{U}_{x^{(\ell)}} = \Big(\prod_{i=0}^{\ell-1} H^{(\ell-i)} W^{(\ell-i)}\Big) \mathbf{U}_{x^{(0)}} \Big(\prod_{i=1}^{\ell} {W^{(i)}}^T H^{(i)}\Big), \qquad \text{(LPU for DNN)}$$

with the diagonal matrices $H^{(i)}$

$$H^{(i)} \coloneqq \operatorname{diag}\left(\frac{\mathrm{d}h_k^{(i)}}{\mathrm{d}z}(z_k^{(i)})\right)_{k=1,\ldots,n^{(i)}}, \quad i = 1, \ldots, \ell, \tag{1}$$

where diag *denotes the* diag *operator for vectors.*

References

JCGM. (2011). Evaluation of measurement data – Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Extension to any number of output quantities. https://www.bipm.org/documents/20126/2071204/JCGM_102_2011_E.pdf/6a3281aa-1397-d703-d7a1-a8d58c9bf2a5?version=1.4&download=true

Ludwig, B. (2023). Pytorch_gum_uncertainty_propagation. https://github.com/BjoernLudwigPTB/pytorch_gum_uncertainty_propagation

Trustworthy virtual experiments and digital twins (ViDiT) – Uncertainty evaluation for Digital Twins

Giacomo Maculotti[1] and Gianfranco Genta[2] and Maurizio Galetto[3]

Key words: Digital Twin, Uncertainty, Traceability, Nanoindentation, Quality

Extended Abstract

Digital Twins (DTs) are key enabling technologies to achieve and realise European strategic policies devoted to sustainability and digitalisation within Industry 4.0 and the European Green Deal [He(2021)]. DTs are simulation models that accurately replicate physical systems in a virtual environment and include dynamic updates of the virtual model according to the observed state of its real counterpart to achieve a physical control of the latter. They consist of two parts, a Physical to Virtual (P2V) connection modelling the considered system, and a Virtual to Physical (V2P) connection, which implements prevention and control strategies to achieve the target accuracy in the physical system [Jones(2020)]. DTs find application in several strategic manufacturing fields [de Ketelaere(2022)] but, within manufacturing processes, the DT's creation often neglects quality control measurements [Xin(2022)]. By neglecting measurements for quality control, their traceability, and the related uncertainty, the overall DT's results, at process level, lack traceability and they cannot be associated with a confidence level in the prediction [Zheng(2020)]. Very few examples of DTs measurement related are reported in the literature

[1] Giacomo Maculotti
Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: giacomo.maculotti@polito.it
[2] Gianfranco Genta
Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: gianfranco.genta@polito.it
[3] Maurizio Galetto
Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: maurizio.galetto@polito.it

[Jones(2020)], and even less literature discusses DT uncertainty evaluation [Kochunas(2021)].

DTs require complex modelling which often leads to data driven approaches rather than explicit analytical models to comply with big data management and real time control. This challenges GUM application for uncertainty evaluation and rather calls for Bayesian and non-parametric approaches, often requiring a task-based evaluation. Furthermore, currently available methods for DT's uncertainty evaluation often neglect coupling with the different parts of the DT, especially the links to the closed-loop feedback control and the V2P connection. This is particularly critical because it introduces a closed-loop feedback and time-dependent control. Therefore, the precision and the accuracy of the prediction are a time and system state dependent function, which shall be evaluated and managed accordingly. Bayesian approaches with iterative update of the distribution of the DT's output variables will allow a rigorous management of such coupling effect while allowing for the prior distribution estimation by non-parametric approaches, e.g. bootstrap sampling, to better condition the problem and achieve faster convergence. A rigorous definition of DT's metrological characteristics is unavailable. Both accuracy and precision shall be defined catering for the V2P closed-loop feedback control.

ViDiT project, funded by the European Partnership on Metrology, will progress beyond the state of the art by developing methods, suitable to cope with the above mentioned challenges, for uncertainty quantification for DTs representing complex measurement processes and mechanisms for four different applications (nanoindentation, NanoCyl, 3D robotic measurement, electrical measurement). Specifically, a case study for nanoindentation will be developed at Politecnico di Torino and Università degli Studi di Padova. In nanoindentation applications, the wear state of the indentation tip is a critical factor, and its shape needs to be continuously controlled to update the calibration or to know when it needs to be disposed of. Similarly, the thermal flux between the sample and the indenter is a critical factor biasing the results if the test is not performed a conditioned environment. Moreover, nanoindentation is highly sensitive to vibrations. A trustworthy DT of this process needs to be established to allow the status of the indentation tip to be monitored along with the correction of real time thermal and vibration effects.

References

1. de Ketelaere, B., Smeets, B., Verboven, P., Nicolai, B., Saeys, W.: Digital Twins in quality engineering. Qual. Eng. 34:404-408 (2022)
2. He, B., Bai, K.-J.: Digital twin-based sustainable intelligent manufacturing: a review. Adv. in Manuf. 9(1):1-21 (2021).
3. Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B.: Characterising the Digital Twin: A systematic literature review. CIRP J. Manuf. Sci. Technol. 29:36-52 (2020)
4. Xin, Y., Chen, Y., Li, W., Wu, F.: Refined simulation method for computer-aided process planning on Digital-Twin technology. Micromachines 13:620 (2022)
5. Zheng, Y., Wang, S., Li, Q., Li, B.: Fringe projection profilometry by conducting deep learning from its digital twin. Opt. Exp. 28:36568 (2020)
6. Kochunas, B., Huan, X.: Digital Twins concepts with uncertainty for nuclear power applications. Energies. 14:4235 (2021)

Combining experimental design with digital twin and phantom experiments to optimise data acquisition for magnetic resonance fingerprinting (MRF)

Stephen L.R. Ellison[1], Christoph Kolbitsch[2] , Constance G.F. Gatefait[2]

Key words: experimental design, magnetic resonance imaging, magnetic resonance fingerprinting, digital twin

1. Introduction

Magnetic Resonance Imaging (MRI) is an important medical imaging technique frequently used in cardiology. Most current medical MRI applications provide qualitative images, relying on the contrast between healthy and pathological tissue for diagnosis. More recently, however, it has become possible to provide more quantitative information on (bio)-physical parameters, which can then be compared to reference values for healthy and pathological tissue or for different tissues.

One important quantitative technique is Magnetic Resonance Fingerprinting (MRF) [1]. MRF uses variation in acquisition parameters in order to create unique temporal signal curves, called "fingerprints", which can be matched to a predetermined dictionary of signal evolution patterns for various tissue types. This allows quantitative determination of parameters such as longitudinal (T1) and transverse (T2) relaxation times, at the voxel level, for different tissues [1-3]. These parameters can provide useful diagnostic information for a range of human diseases.

MRF is a complex series of events, dependent on the choice of a number of key MR acquisition parameters, including pulse and acquisition timings and Rf excitation ('flip angles'). Recent work within the Quiero project [4] has focused on improvement of MRF through uncertainty quantification [5], improved data analysis [6] and optimisation of MRF acquisition parameters for more accurate determination of relaxation times in cardiac MRF [7].

[1] LGC Limited, Teddington, Middlesex, UK
[2] Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany

2. Application of experimental design

In this presentation, we describe the practical application of experimental design approaches, in conjunction with digital twin and phantom experiments, to choose optimal acquisition parameter combinations for accurate relaxation time determination.

The experimental design approach used factorial and fractional factorial screening designs to identify the most significant effects on relaxation time accuracy, and to choose between important qualitative features of the acquisition schemes used. Factors studied included flip angle amplitude, lobe size, the presence of T2-preparation pulses, repetition rates, number of pulses and pulse durations. The designs were chosen to allow characterisation of main effects and two-way interactions. The designs were applied to a series of simulated experiments using digital twins, for which exact relaxation parameters were known, and the results were compared to experiments on a physical phantom with previously characterised relaxation times. Results were assessed using linear modelling for raw coefficients, with Cohen's f used to characterise standardised effect size. Following identification of qualitative settings, a central composite design was used to assess the effect of pulse flip angle and lobe duration on the accuracy of T1 and T2. Lobe size and pulse duration (flip angle) were chosen for optimisation on the basis of effect size. Results were assessed using response surface modelling to locate optimal settings.

Practical considerations discussed will include the choice of parameters for study, including the use of simplified descriptors for the many pulse angles and timings in each acquisition scheme, together with accommodation for operational factors that prevented independent analysis of some parameters.

3. Results and conclusions

Based on standardised effect sizes, T2-preparation pulses were found to influence the accuracy of T2. T2 accuracy and precision were found to improve with larger flip angle amplitudes, for both simulations and phantom experiments. The effects were also verified qualitatively in in-vivo scans [7]. Accuracy and precision of T1 appeared robust to different design parameters, with improved values for faster flip angle variation. The use of formal experimental design approaches assisted in identifying key acquisition parameters, exploring the design space, and in identifying optimal settings.

Acknowledgement

The results presented here have been developed in the framework of the EMPIR Project 18HLT05 QUIERO. Project 18HLT05 QUIERO 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. Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, et al. Magnetic resonance fingerprinting. Nature 2013 495(7440):187–92. https://doi.org/10.1038/nature11971
2. Bipin Mehta B, Coppo S, Frances McGivney D, Ian Hamilton J, Chen Y, Jiang Y, et al. Magnetic resonance fingerprinting: a technical review. Magn Reson Med 2019;81:25–46. https://doi.org/10.1002/mrm.27403.
3. Assländer J. A perspective on MR fingerprinting. J Magn Reson Imaging 2021;53(3):676–85. https://doi.org/10.1002/jmri.27134.
4. L Zilberti, Publishable Summary for 18HLT05 QUIERO - Quantitative MR-based imaging of physical biomarkers. 2023. https://quiero-project.eu/, accessed 20 March 2023.
5. S. Metzner, G. Wübbeler, S. Flassbeck, C. Gatefait, C. Kolbitsch, C. Elster, Bayesian uncertainty quantification for magnetic resonance fingerprinting, Physics in Medicine and Biology, 2021, https://doi.org/10.1088/1361-6560/abeae7
6. S. Metzner, G. Wubbeler, C. Kolbitsch, C. Elster, A comparison of two data analysis approaches for quantitative magnetic resonance imaging, Measurement Science and Technology, 2022, https://doi.org/10.1088/1361-6501/ac5fff.
7. Constance G.F. Gatefait, Stephen L.R. Ellison, Stephen Nyangoma, Sebastian Schmitter, Christoph Kolbitsch.. Optimisation of data acquisition towards continuous cardiac Magnetic Resonance Fingerprinting applications. Physica Medica 2023 105 102514. https://doi.org/10.1016/j.ejmp.2022.102514.

ViDiT project ”Trustworthy virtual experiments and digital twins”

Sonja Schmelter

Key words: Virtual experiment, digital twin, uncertainty evaluation, validation

1 Introduction

Virtual experiments (VEs) and digital twins (DTs) are key enabling technologies to achieve and realise European strategic policies devoted to sustainability and digitalisation within the complex framework of Industry 4.0 and the European Green Deal [He(2021)]. VEs and DTs are both simulation models that accurately replicate physical systems and their characteristics in a virtual environment. DTs also include dynamic updates of the virtual model according to the observed state of its real counterpart. Hence, they consist of two parts, a Physical-to-Virtual connection which models the considered system, and a Virtual-to-Physical connection, which implements control strategies to achieve target results in the physical system [Jones(2020)].

In [Orji(2018)], several examples are stated which explain how virtual metrology could be used to enhance the usefulness of real metrology, e.g. by predicting outcomes for parameters that cannot be measured non-destructively. This use of VEs and DTs in metrology will require uncertainty evaluation methods to be developed. Furthermore, frameworks for the validation of VEs/DTs need to be constructed to ensure traceability for these virtual models. The European Partnership on Metrology project ”Trustworthy virtual experiments and digital twins” (ViDiT) addresses these challenges. The methods developed in this project will ensure trust in the results generated by the VEs/DTs. Hence, they are a premise for further digitalisation in metrology.

2 The ViDiT project

The aim of the ViDiT project is to develop methods and tools that enable ensuring the reliability and trustworthiness of VEs and DTs in metrology. The project has a duration of three years. The consortium consists of eight National Metrology Institutes (NMIs), five universities, two research centres, and six industrial partners bringing together metrological expertise, scientific knowledge, and experience in industrial applications.

The project is divided into four technical work packages (WPs), see Fig. 1. The objectives of WP 1 and WP 2 are to develop uncertainty evaluation methods for VEs and DTs, respectively. In this context, current state-of-the-art uncertainty evaluation methods will be used, such as Bayesian or Monte Carlo approaches or as documented in the GUM [GUM(2018)]. In WP 1, these methods will be applied to the

Sonja Schmelter
Physikalisch-Technische Bundesanstalt (PTB), e-mail: sonja.schmelter@ptb.de

following three metrological applications: coordinate measuring machine (CMM), tilted wave interferometer (TWI), and flow measurement. In WP 2, the following four complex measurement use cases will be considered: 3D robotic measurement, nanoindentation, NanoCyl, and electrical measurements. Since all the considered DTs include dynamic effects, the models have to be updated according to sensor data. WP 3 develops validation approaches for all seven applications mentioned above. Here, statistical procedures will be used for the assessment of differences between calibrated standards and corresponding data from their virtual counterpart. Furthermore, surrogate models will be constructed for one VE application (flow measurement) and one DT application (3D robotic measurement). The surrogate models will be validated against traceable measurement data, and improved as necessary. After having developed the methods for uncertainty evaluation and validation in WPs 1-3, they will be applied to twelve industrial case studies in WP 4. This will demonstrate the practical applicability and transferability of the developed methods in real industrial environments.

Acknowledgements The project (22DIT01 ViDiT) has received funding from the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

References

B. He and K.-J. Bai. Digital twin-based sustainable intelligent manufacturing: a review. *Advances in Manufacturing*, 9:1–21, 2021.

Joint Committee for Guides in Metrology. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM)*. International Bureau of Weights and Measures (BIPM), Sèvres, France, 2008.

D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks. Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, 29:36–52, 2020.

N. Orji, Y. Obeng, C. Beitia, S. Mashiro, and J. Moyne. Virtual Metrology White Paper - INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS (IRDS), 2018.

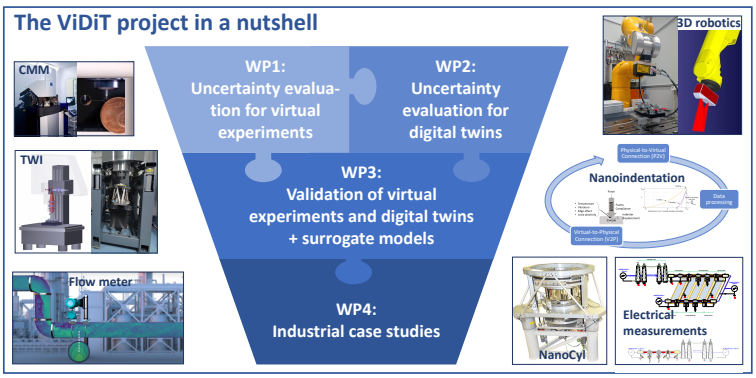


Fig. 1 Illustration of the four technical work packages of the ViDiT project, as well as the seven different metrological applications considered throughout the project.

In silico experiments to investigate the heating efficiency of magnetic nanoparticles in hyperthermia preclinical tests

Marta Vicentini[1], Riccardo Ferrero[2] and Alessandra Manzin[3]

Key words: Magnetic hyperthermia, Magnetic nanoparticles, Bio-heat transfer model, *In silico* tests, Computational anatomical models, Finite element method, Regression methods.

1. Introduction

Magnetic hyperthermia is a cancer treatment that involves the use of magnetic nanoparticles (MNPs) to selectively heat diseased tissues. The MNPs are injected in the tumour region and then activated through the application of an alternating current (AC) magnetic field, with a frequency in the range from 50 kHz to 1 MHz [Perigo(2015)]. MNP activation causes the local release of heat, which leads to an increase in the temperature within the tumour, promoting the damage of cancer cells. A crucial aspect when planning magnetic hyperthermia preclinical tests, commonly conducted on murine models (mice, rats) [Rodrigues(2020)], is the monitoring of the temperature increase, considering that the target range of 40-45 °C should be ideally reached in the entire tumour region for a sufficiently long time (from 20–30 min up to one hour).

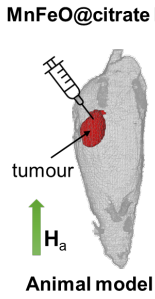
In this framework, we develop a physic-based modelling approach that combines *in silico* models with regression methods to simulate the thermal effects due to AC magnetic field exposure and MNP heat release during magnetic hyperthermia preclinical tests. Several types of MNPs are tested, under different magnetic field conditions and for variable MNP dose administered to the tumour region.

1.1 Methods

The thermal effects due to MNP activation and magnetic field exposure are investigated by evaluating the maximum temperature T_{max} and the average temperature T_{avg} within the tumour region. To this aim, we determine the temperature distribution within the animal body by solving Pennes' bioheat equation, taking into account possible eddy current effects due to field exposure by means of a low-

[1]Marta Vicentini, INRIM, 10135 Torino, Italy, e-mail: m.vicentini@inrim.it
[2]Riccardo Ferrero, INRIM, 10135 Torino, Italy, e-mail: r.ferrero@inrim.it
[3]Alessandra Manzin, INRIM, 10135 Torino, Italy, e-mail: a.manzin@inrim.it

frequency electromagnetic field solver [Manzin(2021), Vicentini(2022)]. Different regression models (e.g., linear, support vector machine, and Gaussian process regression) are compared to predict T_{max} and T_{avg}, and thus enlarge the amount of synthetic data for experimental conditions that are not numerically simulated.



Tmax		
Regression model	RSME (validation)	Max discrepancy % (test)
Stepwise linear	0.042	0.147
Gaussian Process	**0.021**	**0.084**
Neural Network	0.868	2.568
Quadratic SVM	0.927	3.164

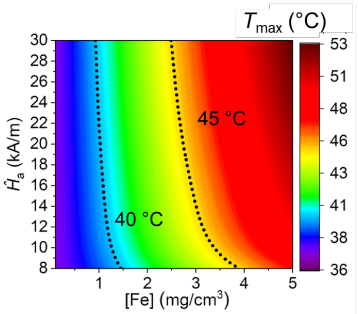


Figure 1: *The table reports the comparison between four regression models for the evaluation of the maximum temperature (T_{max}) within the tumour region of a 30 g mouse model (schematized on the left) for manganese ferrite (MnFeO) NPs with citrate coating [Soetaert(2017)], activated with a uniform 150 kHz magnetic field. The diagram on the right shows the results obtained for T_{max}, as a function of magnetic material concentration [Fe] and magnetic field peak amplitude $\hat{H}a$, predicted with the Gaussian process regression method. The black dashed lines define the therapeutic range of 40-45 °C.*

1.2 Results

We perform a comprehensive analysis, significantly expanded thanks to regression models, to compare the heating efficiency of different types of MNPs [Soetaert(2017)] as a function of AC magnetic field parameters (frequency and peak amplitude) and MNP dose. The investigation is carried out on two murine models, a 30 g mouse and a 500 g rat, under the assumption that MNPs are uniformly distributed within the tumour region.

As an example, Figure 1 reports the results obtained for T_{max}, calculated for the mouse model when treated with manganese ferrite (MnFeO) NPs with citrate coating.

References

1. Périgo, E.A., et al.: Fundamentals and advances in magnetic hyperthermia. Appl. Phys. Rev. **2(4)**, 041302 (2015)
2. Rodrigues, H.F., Capistrano, G., Bakuzis, A.F.: In vivo magnetic nanoparticle hyperthermia: a review on preclinical studies, low-field nano-heaters, noninvasive thermometry and computer simulations for treatment planning. Int. J. Hyperthermia **37**, 76--99 (2020)
3. Manzin, A., Ferrero, F., Vicentini, M.: From micromagnetic to in silico modeling of magnetic nanodisks for hyperthermia applications. Adv. Theory Simul. **4(5)**, 2100013 (2021)
4. Vicentini, M., Vassallo, M., Ferrero, F., Androulakis, I., Manzin, A.: *In silico* evaluation of adverse eddy current effects in preclinical tests of magnetic hyperthermia. Comput. Methods Programs Biomed. **223**, 106975 (2022)
5. F. Soetaert, F., Kandala, S.K., Bakuzis, A., Ivkov, R.: Experimental estimation and analysis of variance of the measured loss power of magnetic nanoparticles. Sci. Rep. 7, 6661 (2017)

Data smoothing and its application to the evaluation of the measurement uncertainty in a humidity standard

R. Nobakht, R. Cuccaro, V. Fernicola

Abstract In the production of ultra-high purity industrial gases used in advanced semiconductor fabrication, measuring water vapor contamination at the parts per billion level is crucial. INRiM developed a primary trace humidity standard able to producing a stream of moist gas with a frost-point temperature between -100 °C and -25 °C (or, equivalently, with an amount fraction from 15 ppb to 600 ppm) with an expanded uncertainty of better than 0.5 °C at the lowest temperature. The study evaluates different smoothing techniques like moving average, Gaussian-weighted moving average, and Savitzky-Golay, and their impact on measurement standard deviation and output uncertainty.

Key words: uncertainty analysis, data smoothing, humidity standard, calibration facilities

1 Theory

Humidity generators use gas stream (N_2, Ar, H_2, air) passing over water or ice at known temperature T_{sat}. Gas is saturated with water vapor, expressed as mole fraction x using Eq. 1, where $e(T_{sat})$ is the saturation water vapor pressure and $f(T_{sat},P_{sat})$ corrects for non-ideal behavior of water vapor gas mixture at saturator temperature T_{sat} and pressure P_{sat}. Empirical equations calculate $e(T_{sat})$ and $f(T_{sat},P_{sat})$ [Sonntag(1990), Greenspan(1976)].

R. Nobakht
Istituto Nazionale di Ricerca Metrologica, and Politecnico di Torino, Italy, e-mail: rezvaneh.nobakht@polito.it

R. Cuccaro
Istituto Nazionale di Ricerca Metrologica e-mail: R. Cuccaro@inrim.it

V. Fernicola
Istituto Nazionale di Ricerca Metrologica e-mail: V. Fernicola@inrim.it

$$x = \frac{e(T_{\mathrm{sat}})f(T_{\mathrm{sat}}, P_{\mathrm{sat}})}{P_{\mathrm{sat}}} \tag{1}$$

2 Method

In order to assess humidity standard generators, a detailed uncertainty analysis is needed. The combined uncertainty of the amount fraction, $u_{\mathrm{c}}(x)$, can be estimated by propagating input parameter uncertainties weighted by their sensitivity coefficients. Empirical formulations' standard uncertainties and the standard uncertainties of temperature and pressure, $u(T_{\mathrm{sat}})$ and $u(P_{\mathrm{sat}})$ respectively, are used to estimate this uncertainty. This calculation can be performed using equation 2: [CIPM(1999), BIPM(2008), Lovell-Smith(2000), Lovell-Smith(2009)].

$$u_{\mathrm{c}}(x) = \sqrt{\left(u(T_{\mathrm{sat}})\frac{\partial x}{\partial T_{\mathrm{sat}}}\right)^2 + \left(u(P_{\mathrm{sat}})\frac{\partial x}{\partial P_{\mathrm{sat}}}\right)^2 + \left(u(e)\frac{\partial x}{\partial e}\right)^2 + \left(u(f)\frac{\partial x}{\partial f}\right)^2} \tag{2}$$

In this work, statistical filters were applied to enhance the precision and reliability of calibration and measurement capabilities by minimizing noise and improving data distribution [Orfanidis(1996)]. The use of data smoothing techniques reduced the standard deviation of pressure data by a factor of 4.5 and brought it closer to a normal distribution. The study also revealed that an optimal filter can decrease the standard deviation of temperature measurements as low as 1 mK, which represents the ultimate threshold of temperature refinement. These findings underscore the importance of statistical filters in enhancing the performance of the humidity measurements.

Acknowledgements This work has been carried out within the project "PROMETH2O - Metrology for trace water in ultra-pure process gases". The 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 (Grant number: EMPIR 20IND06 PROMETH2O).

References

[BIPM(2008)] BIPM (2008) Guide to the expression of the uncertainty in measurement
[CIPM(1999)] CIPM (1999) Mutual recognition arrangement. (revised 2003, accepted July 2008)
[Greenspan(1976)] Greenspan LH (1976) A simple approximation for calculating dew point. National Bureau of Standards
[Lovell-Smith(2000)] Lovell-Smith (2000) Uncertainty analysis for humidity generators. Measurement Science and Technology 11(6):783–788
[Lovell-Smith(2009)] Lovell-Smith (2009) Calibration of dew point hygrometers. Metrologia 46:607–615
[Orfanidis(1996)] Orfanidis SJ (1996) Introduction to Signal Processing. Prentice Hall
[Sonntag(1990)] Sonntag D (1990) Important new values of physical constants of 1986. Zeitschrift für Meteorologie 70(5):340–344

The strategic research agenda of the European Metrology Network Mathmet

Sebastian Heidenreich, Markus Bär, Clemens Elster and Oliver Henze[1], Keith Lines, Susan Rhodes, Maurice Cox, Louise Wright and Peter Harris[2], Nicolas Fischer, Diane Gumuchian and Sebastien Marmin[3], Stephen Ellison[4], Adriaan van der Veen and Gertjan Kok[5], Luca Zilberti, Alessandra Manzin and Francesca Pennecchi[6], Alen Bosnjakovic[7], Joao Sousa, Carlos Pires and Olivier Pellegrino[8], and Kirstin Vogel[9]

Key words: strategic research agenda, artificial intelligence, virtual metrology, uncertainty, data analysis

In 2014, several metrology institutes initiated the foundation of a European Centre for Mathematics and Statistics in Metrology. Three years later, the Memorandum of Understanding was signed by seven members. The aim of the centre was to aggregate, strengthen, focus and disseminate best practice in the field of mathematics and statistics in metrology. After EURAMET decided to launch metrology networks, the European Centre for Mathematics and Statistics was re-established in 2019 as Mathmet, the European Metrology Network (EMN) for mathematics and statistics in metrology. In the same year, a four-year Joint Network Project (JNP 18NET05) was launched, pursuing the following key objectives:

i. establish a stakeholder consultation process,
ii. develop a strategic research agenda, and
iii. develop quality assurance tools (data, software, guidelines).

[1]Sebastian Heidenreich, Markus Bär, Clemens Elster and Oliver Henze - PTB, Abbestrasse 2-12, 10587 Berlin, Germany, e-mail: sebastian.heidenreich@ptb.de
[2]Keith Lines, Susan Rhodes, Maurice Cox, Louise Wright and Peter Harris - NPL, Hampton Road, Teddington TW11 0LW, UK, e-mail: peter.harris@npl.co.uk
[3]Nicolas Fischer, Diane Gumuchian and Sebastien Marmin - LNE, 1 rue Gaston Boissier, 75724 Paris Cedex 15, France, e-mail: Nicolas.Fischer@lne.fr
[4]Stephen Ellison LGC Limited, Queens Road, Teddington, TW11 0LY, UK, e-mail: S.Ellison@LGCGroup.com
[5]Adriaan van der Veen and Gertjan Kok VSL B.V., Thijsseweg 11, 2629 JA, Delft, the Netherlands, e-mail: GKok@vsl.nl
[6]Luca Zilberti, Alessandra Manzin and Francesca Pennecchi
INRIM, Strada delle Cacce 91, 10135 Turin, Italy, e-mail: f.pennecchi@inrim.it
[7]Alen Bosnjakovic IMBIH**,** Augusta Brauna 2, 71000 Sarajevo, e-mail: alen.bosnjakovic@met.gov.ba
[8]Joao Sousa, Carlos Pires and Olivier Pellegrino
IPQ, Rua António Gião, 2, 2829-513 Caparica, Portugal, e-mail: jasousa@ipq.pt
[9]Kirstin Vogel BAM, Unter den Eichen 87, 12205 Berlin, Germany, e-mail: kristin.vogel@bam.de

This talk will present Mathmet's Strategic Research Agenda (SRA) as an output of the JNP. The development of the SRA was based on stakeholder consultations and the involvement of all EMN Mathmet members.

The SRA identified cutting-edge developments in society for which mathematics and statistics will play an essential role:

i. machine learning and artificial intelligence (AI), and
ii. computational modelling and virtual metrology (VM).

Neither topic belongs to the traditional research topics in the field of mathematics and statistics in metrology but are highly relevant in the dynamic development of digital transformation, with expected impact on a plethora of applications, in the health, energy, environment and industry sectors.

Based on further stakeholder consultations and the needs of other EMNs, a classic (third) topic was included in the SRA:

iii. data analysis and quantification of uncertainty (UQ).

The talk will discuss the metrology requirements and challenges related to AI, VM and UQ and present a corresponding roadmap.

Acknowledgement

The JNP 18NET05 Mathmet has received funding from the European Union's Horizon 2020 research and innovation programme.

Reconstructions of nano-geometries from grazing incidence X-Ray fluorescence measurements using neural networks

Sebastian Heidenreich, Victor Soltwisch, Nando Farchmin[1], Paul Hagemann[2],

Key words: Bayesian inversion, Invertible neural networks, Scatterometry

Optical scattering techniques are frequently used for the characterization of periodic nanostructures on surfaces in the semiconductor industry [1,2]. As a non-destructive technique, grazing incidence X-ray fluorescence (GIXRF) is of particular interest for many industrial applications. Mathematically, the reconstruction of nanostructures, i.e., of their geometrical parameters, can be rephrased in an inverse problem. Given grazing incidence X-ray fluorescence measurements, we want to recover the distribution of the geometry parameters of the grating. To account for measurement errors, we use Bayesian inversion.

The standard approach to recover the distribution of the parameters are Markov Chain Monte Carlo (MCMC) based algorithms [3]. In this talk, we present invertible neural networks (INNs) [4,5] within the general concept of transport maps [6]. This means that we sample from a reference distribution and seek a diffeomorphic transport map, or more precisely its approximation by an INN, which maps this reference distribution to the posterior of interest. This approach has some advantages over standard MCMC–based methods:

i) Given a transport map, which is computed in an offline step, the generation of independent posterior samples essentially reduces to sampling the freely chosen reference distribution. Additionally, observations indicate that learning the transport map requires less time than generating a sufficient amount of independent samples via MCMC.

ii) Although the transport map is conditioned on a specific measurement, it can serve as a good initial guess for the transport related to similar measurements or as a prior in related inversion problems. Hence the effort to find a transport for different runs within the same experiment reduces drastically. An even more sophisticated way of using a pretrained diffeomorphism has been recently suggested in [7].

[1]Sebastian Heidenreich, Nano Farchmin, Victor Soltwisch - PTB, Abbestrasse 2-12, 10587 Berlin, Germany, e-mail: sebastian.heidenreich@ptb.de

[2]Paul Hagemann – Institute of Mathematics, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany

Having trained the INN, we compare its ability to recover the posterior distribution with the established MCMC method for fluorescence experiments.

In the talk, we introduce INNs with an appropriate loss function to sample posterior distributions of inverse problems [7]. Furthermore, we describe the forward model in GIXRF and finally, we compare the performance of the INN with MCMC posterior sampling.

1. M.-A. Henn, H. Gross, S. Heidenreich, F. Scholze, C. Elster, and M. Ba ̈r. Improved reconstruction of critical dimensions in extreme ultraviolet scatterometry by mod- eling systematic errors. Measurement Science and Technology, 25(4):044003, 2014.

2. C. Mack. Fundamental principles of optical lithography: the science of microfabrication. John Wiley & Sons, 2008.

3. C. Andrieu, N. de Freitas, A. Doucet, and M. I. Jordan. An introduction to MCMC for machine learning. Machine Learning, 50:5–43, 2003.

4. L. Dinh, J. Sohl-Dickstein, and S. Bengio. Density estimation using real NVP. In 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings, 2017.

5. J. Kruse, G. Detommaso, R. Scheichl, and U. Ko ̈the. HINT: Hierarchical invertible neural transport for density estimation and Bayesian inference. ArXiv preprint arXiv:1905.10687, 2020.

6. Y. Marzouk, T. Moselhy, M. Parno, and A. Spantini. Sampling via measure trans- port: An introduction. Handbook of Uncertainty Quantification, page 1–41, 2016.

7. A. Siahkoohi, G. Rizzuti, M. Louboutin, P. A. Witte, and F. J. Herrmann. Precon- ditioned training of normalizing flows for variational inference in inverse problems. ArXiv preprint arXiv:2101.03709, 2021.

8. A. Anderle, N. Farchmin, P. Hagemann, S. Heidenreich. V. Soltwisch, G. Steidl. Invertible neural networks versus MCMC for posterior reconstruction in grazing incidence X-ray fluorescence. Scale Space and Variational Methods in Computer Vision: 8th International Conference, SSVM 2021, Virtual Event, May 16--20, 2021, Proceedings 528-339.

Thermo-acoustic simulation in ultrasound hyperthermia applications

Silvia Pozzi[1] and Barbara Caccia[2] and Stefano Valentini[3] and Raphaela Baesso[4] and Aoife Ivory[5] and Piero Miloro[6]

Key words: ultrasound hyperthermia, in silico models, ultrasound simulation, thermal simulation, computational physics.

Abstract

Numerous studies show the sensitization effect of a local increase in body temperature (hyperthermia) when applied in combination with radiation therapy [Datta(2015), Mallory(2016), Peeken(2017)].

One of the challenges of hyperthermia is delivering the appropriate amount of heat to a defined anatomical district, thus it is necessary to have an *in-silico* tool to better plan and evaluate the spatial and temporal temperature profiles.

The aim of the project RaCHy (Radiotherapy Coupled with Hyperthermia) is to provide a reliable metrology framework for the evaluation of a class of

[1] Silvia Pozzi
National Center for Radiation Protection and Computational Physics, Italian National Institute of Health, Rome, Italy, e-mail: silvia.pozzi@iss.it
[2] Barbara Caccia
National Center for Radiation Protection and Computational Physics, Italian National Institute of Health, Rome, Italy, e-mail: barbara.caccia@iss.it
[3] Stefano Valentini
National Center for Radiation Protection and Computational Physics, Italian National Institute of Health, Rome, Italy, e-mail: stefano.valentini@iss.it
[4] Raphaela Baesso
Ultrasound and Underwater Acoustics, National Physical Laboratory, Teddington, UK, e-mail: raphaela.baesso@npl.co.uk
[5] Aoife Ivory
Ultrasound and Underwater Acoustics, National Physical Laboratory, Teddington, UK, e-mail: aoife.ivory@npl.co.uk
[6] Piero Miloro
Ultrasound and Underwater Acoustics, National Physical Laboratory, Teddington, UK, e-mail: piero.miloro@npl.co.uk

radiation-based therapies coupled with hyperthermia induced by therapeutic ultrasound, conventional electromagnetic radiation, magnetic nano particles. This contribution will describe the work conducted within the RaCHy project to develop a hyperthermia treatment planning tool for hyperthermia induced by therapeutic ultrasound.

The entire simulation process and related critical issues will be described. The process consists of an initial simulation phase in which the propagation of the ultrasonic wave is described and the deposited heat is estimated. This is performed using the k-Wave toolbox [Treeby(2010)] with the holographic technique, in order to best reproduce the acoustic field generated by the HIFU transducer. The evaluation of the temperature increase is carried out from the heat deposited by the ultrasound waves by numerically solving the Pennes' equation [Pennes(1948)] using the Alternating Direction Implicit Method.

The results obtained for an *ad-hoc* phantom realized within the project will be illustrated.

References

1. Datta, N.R., Gomez Ordonez, S., Gaipl, U.S., Paulides, M.M., Crezee, H., Gellermann, J., Marder, D., Puric, E., Bodis, S.: Local hyperthermia combined with radiotherapy and-/or chemotherapy: Recent advances and promises for the future. Cancer Treatment Reviews. **41**(9), 742--753 (2015)
2. Mallory, M., Gogineni, E., Jones, G.C., Greer, L., Simone, C.B. 2nd.: Therapeutic hyperthermia: the old, the new, and the upcoming. Crit Rev Oncol Hematol. **97**, 56--64 (2016)
3. Peeken, J.C., Vaupel, P., Combs, S.E.: Integrating Hyperthermia into Modern Radiation Oncology: What Evidence Is Necessary? Front Oncol. **7,** 132 (2017)
4. Pennes, H.H.: Analysis of tissue and arterial blood temperatures in the resting human forearm. J Appl Physiol. **1**(2), 93-122 (1948)
5. Treeby, B.E., Cox, B.T.: k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields. J Biomed Opt. **15**(2) (2010)

Forecasting oxygen content in seawater

Gianfranco Durin, Marco Coïsson, Francesca Pennecchi, Andrea Bordone, Tiziana Ciuffardi, Giancarlo Raiteri and Chiara Petrioli

High resolution monitoring stations and big data production are two of the occurring challenges in ocean studies. To validate satellite data, make predictions on fast seawater changes and provide early warnings for maritime activities, networks of sensors producing high quality-data in real time are becoming a priority. The Internet of Underwater Things (IoUT) technology of the Green Star project at Smart Bay Santa Teresa in Lerici, Italy [SmartBay], aims to create a sensor network producing high resolution and quality data for marine ecosystem monitoring. A W-SENSE wireless sensors network provides various parameters, such as temperature, oxygen, conductivity etc. High-cost multi parametric probes such as CTD (Conductivity, Temperature and Depth) devices are also employed for IoUT sensor calibration [Raiteri(2018)]. In this paper, we investigate the daily variability of the seawater oxygen with the aim to link its temporal variations with a few environmental parameters such as solar irradiation, air and water temperature. The final goal is to forecast the oxygen content on the base of weather forecast only, once time series of a few years are acquired and tested with statistical and machine learning methods.
Since Dec. 2021, data have been acquired every 30 minutes, (see Figure - May 18-27, 2022). Here we consider the oxygen percentage (red curve), being 100% the thermodinamic value calculated at the measured water temperature and salinity. The oxygen content actually increases due to the phythoplankton production under the solar radiation. In addition, various algae progressively grow over the sensor (an effect named fouling), increasing the oxygen daily variability. To limit the fouling, the sensors are cleaned every 8-10 days, when are compared with the CTD sensor for calibration (red dot in Figure). Even with periodic cleaning, the fouling increases

Gianfranco Durin, Marco Coïsson, Francesca Pennecchi
INRIM, strada delle Cacce 91, Torino, Italy, e-mail: g.durin@inrim.it

Andrea Bordone, Tiziana Ciuffardi, Giancarlo Raiteri
Marine Environment Research Centre of S. Teresa - Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) - Lerici, Italy,

Chiara Petrioli
W•SENSE S.r.l., Rome, Italy & University of Rome La Sapienza, Rome, Italy,

with time, especially in spring-summer due to the higher temperatures. Our first goal is then to properly decompose the signal into a daily component (the *Seasonal* one in time series analysis) and a *Trend*, taking into account long time variations, irrespective of fouling effects. We use the Seasonal-Trend decomposition using LOESS (STL), known to be robust to outliers, where the seasonal component is allowed to change over time. With this decomposition, we take daily means of a few parameters, namely the mean trend, the mean and standard deviation of the air and water temperatures, the mean solar radiation. The data of 414 days (Dec. 2021 - Apr. 2023) are then used in a random forest regressor, with the air and water temperatures, the solar irradiation and the days passed from the last cleaning as input parameters. Once properly renormalized, 90% of the data are used for train, and 10% for test, chosen randomly. We found out that the fluctuations of water and air temperatures (i.e. the standard deviation), together with the number of days from cleaning are totally irrelevant in the regression, while the solar radiation weights for about 60%, and the mean water and air temperatures for 20% each. Remarkably, the oxygen prediction are accurate within about 4 %. Future investigations will include different methods of decomposition, and analysis of the time signal with recurrent neural networks [O'Donncha(2022)].

References

O'Donncha(2022). O'Donncha F, et al. (2022) A spatio-temporal LSTM model to forecast across multiple temporal and spatial scales. Ecological Informatics 69:101687

Raiteri(2018). Raiteri G, Bordone A, Ciuffardi T, Pennecchi F (2018) Uncertainty evaluation of CTD measurements: a metrological approach to water-column coastal parameters in the gulf of La Spezia area. Measurement 126:156–163

SmartBay. `https://smartbaysteresa.com/en/green-star/`

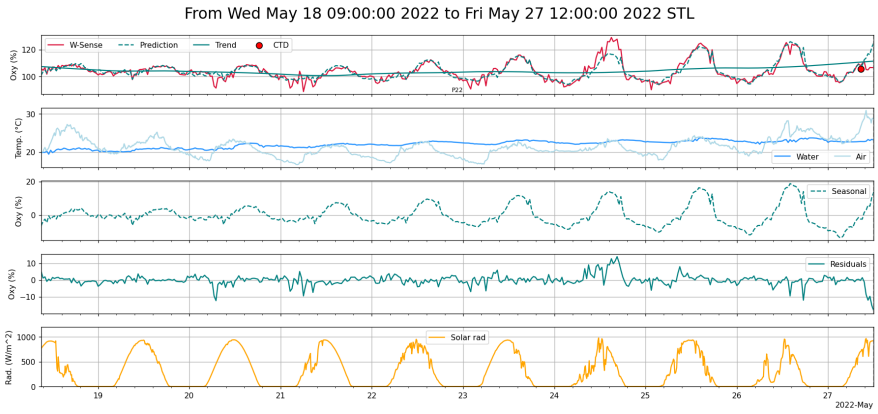


Fig. 1 Decomposition of the Oxygen content in % (red curve) with Trend (solid green line), Seasonal and Residuals. Temperatures of water and air in blue, and solar radiation in gold.

Mathmet Quality Assurance Tools for data, software, and guidelines

Keith Lines, Jean-Laurent Hippolyte, Indhu George and Peter Harris[1], Nicolas Fischer, Diane Gumuchian and Sebastien Marmin[2], Stephen Ellison and Simon Cowen[3], Gertjan Kok[4], Sebastian Heidenreich and Oliver Henze[5], Francesca Pennecchi[6], Alen Bosnjakovic[7], Joao Sousa, Carlos Pires and Olivier Pellegrino[8], and Jörg F. Unger[9]

Key words: Quality Management System, Quality Assurance Tools, trust, data product, software product

1. Introduction

The European Metrology Network for Mathematics and Statistics (Mathmet), provides Quality Assurance Tools (QAT) for download from its website [1]. The aim of these tools is to help ensure fitness for purpose of data sets, software, and guidelines for metrology. This session will present the QAT, discuss its contents and the lessons learned during its development.

A special session of MSMM 2021 [2] provided valuable feedback when the QAT was in development and referred to as a Quality Management System (QMS). The MSMM 2021 session was largely an exercise in presenting the thinking behind

[1] Keith Lines, Jean-Laurent Hippolyte, Indhu George and Peter Harris NPL, Hampton Road, Teddington TW11 0LW, UK, e-mail: keith.lines@npl.co.uk
[2] Nicolas Fischer, Diane Gumuchian and Sebastien Marmin, LNE, 1 rue Gaston Boissier, 75724 Paris Cedex 15, France, e-mail: Nicolas.Fischer@lne.fr
[3] Stephen Ellison and Simon Cowen LGC Limited, Queens Road, Teddington, TW11 0LY, UK, e-mail: S.Ellison@LGCGroup.com
[4] Gertjan Kok, VSL B.V., Thijsseweg 11, 2629 JA, Delft, the Netherlands, e-mail: GKok@vsl.nl
[5] Sebastian Heidenreich and Oliver Henze, PTB, Abbestrasse 2-12, 10587 Berlin, Germany, e-mail: sebastian.heidenreich@ptb.de
[6] Francesca Pennecchi, INRIM, Strada delle Cacce 91, 10135 Turin, Italy, e-mail: f.pennecchi@inrim.it
[7] Alen Bosnjakovic and Merima Čaušević, IMBIH**,** Augusta Brauna 2, 71000 Sarajevo, e-mail: alen.bosnjakovic@met.gov.ba
[8] Joao Sousa, Carlos Pires and Olivier Pellegrino, IPQ, Rua António Gião, 2, 2829-513 Caparica, Portugal, e-mail: jasousa@ipq.pt
[9] Jörg F. Unger, BAM, Unter den Eichen 87, 12205 Berlin, Germany, e-mail: joerg.unger@bam.de

the QAT and gathering ideas. This MSMM 2023 session will provide an update, describing how those ideas were implemented. The overall approach of the QAT, as summarised in [3], will also be presented.

2. QAT for data, software, and guidelines

The QAT consists of separate components for data, software, and guidelines. For data and software, an interactive risk assessment tool linked to supporting documentation guides the user in developing a quality management plan. The activities listed in the plan follow a typical life cycle from requirements capture to design and development, verification, and validation through to release and maintenance. For guidelines, there is a focus on understanding the provenance, bearing in mind that, for example, ISO Standards, OIML Recommendations and BIPM Guides, can be expected to be thoroughly reviewed.

Pragmatism has underpinned the development of the QAT, which sets only high-level requirements on users with the aim to ensure that there are no conflicts with the processes and procedures in-place and adopted by individual members of Mathmet.

Presentations and training material [4], developed to publicise the QAT, will be discussed. Use-cases, updated from MSMM 2021, will illustrate how the QAT can help support the Mathmet Strategic Research Agenda (SRA) [5].

As Mathmet continues and evolves, so should the QAT. Feedback to help improve and encourage adoption of either the QAT or, more importantly, its underlying concepts will be vital.

Acknowledgements

The project 18NET05 MATHMET 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. Mathmet European Metrology Network for Mathematics and Statistics: Quality Assurance Tools. https://www.euramet.org/european-metrology-networks/mathmet/activities/quality-assurance-tools Cited 20 March 2023
2. MSMM 2021 http://www.msmm2021.polito.it/programme Cited 20 March 2023
3. KJ Lines; J-L Hippolyte; I George; PM Harris: A MATHMET Quality Management System for data, software, and guidelines". Acta IMEKO Journal Vol. 11 No. 4. (2022)
4. **TBC.** Link to material generated at Mathmet QAT course held on 22 / 23 March
5. Mathmet Strategic Research Agenda. https://www.euramet.org/european-metrology-networks/mathmet/strategy/strategic-research-agenda Cited 20 March 2023

Myocardial Fibrosis Segmentation from MRF Images

Martin Špendl[1] and Constance G. F. Gatefait[2] and Christoph Kolbitsch[3] and Jure Žabkar[4]

Key words: Magnetic Resonance Fingerprinting, Myocardial Fibrosis, Deep Learning

1. Introduction

Magnetic Resonance Fingerprinting (MRF) is a recent multiparametric quantitative MRI method, which can provide numerous parameters such as the relaxation times T1 and T2 [Ma(2013)]. In contrast to a classic MR protocol, MRF acquisition parameters are not fixed but vary throughout the scanning process. Thus, unique temporal signal curves are obtained for each combination of T1 and T2 values at the end of the scan. Those curves, named "fingerprints", are then voxel-wise matched to a pre-calculated dictionary, simulating the signal evolution for diverse tissue types (i.e. different combinations of T1 and T2). T1 and T2 quantitative maps are then obtained [Ma(2013), Bipin Mehta(2019)]. MRF can be used for cardiac imaging applications to help detect myocardium diseases, including fibrosis [Liu(2018), Hamilton(2020)].

Myocardial fibrosis is characterized by increased collagen production in tissues, leading to a stiffening of the cardiac muscle. Those scarred tissues also lead to elevated T1 and T2 values. Quantitative MRI is a non-invasive way to quantify it in order to treat it and prevent related diseases [Bing(2019), Liu(2017)].

The task we address in this experiment resembles an anomaly detection task, where fibrosis represents the anomaly of a normal heart MRF scan. In a classic anomaly detection problem, the task is to detect an anomalous sample among mostly anomaly-free samples. Detection algorithms thus focus on detecting rare patterns and variations in shapes. In medicine, the problem is reversed: healthy subjects are not routinely scanned with expensive medical imaging techniques such as MRF and only the patients suspected of having a disease are scanned with MRF. Our data set resembles such discrepancy by having 197 cases with an anomaly and only 3 without it. Due to a small size and disproportional amount of anomaly in the data set, state-of-the-art anomaly detection approaches are not directly applicable to our use case.

2. Experiments

Our in-situ data set consists of 200 simulations of MRF acquisition and reconstruction of artificial hearts based on the XCAT phantom [Segars(2010)]. Each simulation sample is represented by two 120x120 grey-scale images of T1 and T2 reconstruction signals. The XCAT phantom assumes a patient with focal fibrosis in the myocardium in 197 out of 200 cases; 3 cases represent a healthy myocardium without fibrosis. The fibrosis was simulated with increased T1 and T2 values compared to healthy myocardium.

---

[1] Martin Špendl
Faculty of Computer and Information Science, University of Ljubljana, Slovenia, ms4681@student.uni-lj.si
[2] Constance G. F. Gatefait
Physikalisch-Technische Bundesanstalt (PTB), Berlin, Germany, constance.gatefait@ptb.de
[3] Christoph Kolbitsch
Physikalisch-Technische Bundesanstalt (PTB), Berlin, Germany, christoph.kolbitsch@ptb.de
[4] Jure Žabkar
Faculty of Computer and Information Science, University of Ljubljana, Slovenia, jure.zabkar@fri.uni-lj.si

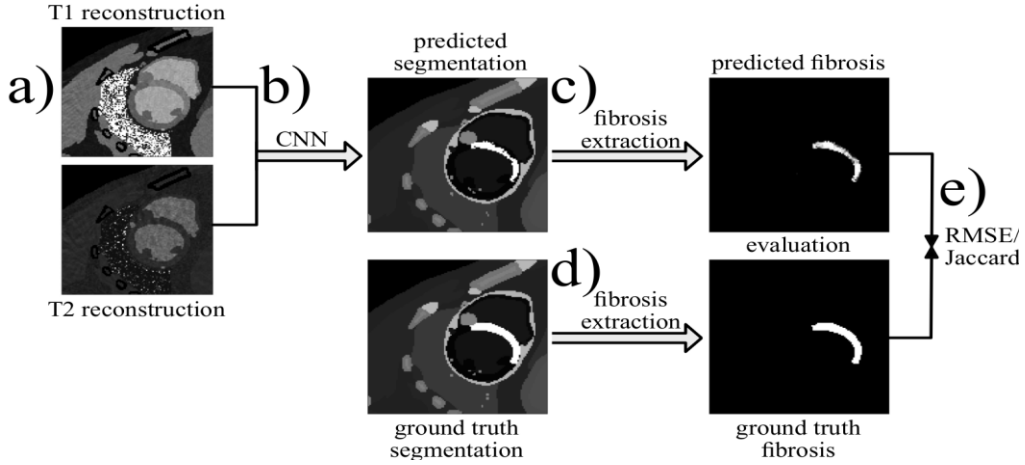


Figure 1: Schema of a workflow for model evaluation. a) Lung denoising of test images, b) model prediction, c) extraction of fibrosis pixels in the prediction, d) extraction of fibrosis pixels in the ground truth segmentation, e) evaluation of fibrosis prediction using Jaccard similarity index or RMSE

Our segmentation model was based on the U-Net encoder-decoder architecture with slight modifications from the model in the RIAD method [Zavrtanik(2021)]. Inputs to the model are T1 and T2 images, and the target variable is the per-pixel segmentation matrix. We evaluate models in a 2-fold cross-validation. Using the ground truth segmentation annotations for test samples, we construct a binary matrix depicting pixels with fibrosis labels. For model evaluation, we use the Jaccard similarity index and RMSE between ground truth fibrosis segmentation and predicted segmentation. For the Jaccard similarity index, the threshold of 0.5 is applied to the predicted fibrosis segmentation to get a set of predicted pixels. Intersection and union are calculated with a set of ground truth fibrosis pixels. We calculate RMSE between ground truth segmentation of fibrosis and predicted values. Calculated error is a single value estimate of model performance on fibrosis segmentation.

3. Conclusion

Our approach to assigning uncertainty to fibrosis segmentation is an important step toward clinical usage. Our model achieves reliable performance even in samples that are challenging to predict by a human observer. We show that using both T1 and T2 images can improve the detection capabilities of the model. We believe the approach has a great potential in future transfer learning tasks for in-vivo samples.

References

1. Ma, D., Gulani, V., Seiberlich, N., Liu, K., Sunshine, J.L., Duerk, J.L., Griswold, M.A.: Magnetic resonance fingerprinting. Nature 495(7440), 187–192 (2013)
2. Bipin Mehta, B., Coppo, S., Frances McGivney, D., Ian Hamilton, J., Chen, Y., Jiang, Y., Ma, D., Seiberlich, N., Gulani, V., Alan Griswold, M.: Magnetic resonance fingerprinting: a technical review. Magnetic resonance in medicine 81(1), 25–46 (2019)
3. Liu, Y., Hamilton, J., Rajagopalan, S., Seiberlich, N.: Cardiac magnetic resonance fingerprinting: technical overview and initial results. JACC: Cardiovascular Imaging 11(12), 1837–1853 (2018)
4. Hamilton, J.I., Jiang, Y., Eck, B., Griswold, M., Seiberlich, N.: Cardiac cine magnetic resonance fingerprinting for combined ejection fraction, t1 and t2 quantification. NMR in Biomedicine 33(8), e4323 (2020)
5. Bing, R., Dweck, M.R.: Myocardial fibrosis: why image, how to image and clinical implications. Heart 105(23), 1832–1840 (2019)
6. Liu, T., Song, D., Dong, J., Zhu, P., Liu, J., Liu, W., Ma, X., Zhao, L., Ling, S.: Current understanding of the pathophysiology of myocardial fibrosis and its quantitative assessment in heart failure. Frontiers in physiology 8, 238 (2017)
7. Segars, W.P., Sturgeon, G., Mendonca, S., Grimes, J., Tsui, B.M.: 4d xcat phantom for multimodality imaging research. Medical physics 37(9), 4902–4915 (2010)
8. Zavrtanik, V., Kristan, M., Skočaj, D.: Reconstruction by inpainting for visual anomaly detection. Pattern Recognition 112, 107706 (2021)

Towards Reference Point Cloud Generation for Data Fusion in Dimensional Metrology

Louis-Ferdinand Lafon[1,2], Nabil Anwer[2], Charyar Mehdi-Souzani[2], Alain Vissiere[1] and Hichem Nouira[1]

Key words: Dimensional metrology, data fusion, registration, reference data

1. Introduction

Spatial data fusion is widely applied in dimensional metrology (quality control) and computer vision applications. Data fusion is defined as the combination of data from several sensors into a common representational format to obtain more complete, detailed and accurate measurements [Weckenmann(2009)]. The resulting data is used for comparison with CAD models, to verify the conformance to the design tolerance specifications of the manufactured parts, or for 3D reconstruction.

One of the processes in data fusion is point cloud registration. It aims to align measurements, initially in the sensors coordinate systems, then into the coordinate system of the measured part [Zhu(2019)]. Depending on whether two or more point clouds are aligned, the algorithm is classified as pairwise or groupwise, and the estimated transformation can be classified as either rigid or non-rigid. The most commonly used pairwise algorithm is the Iterative Closest Point (ICP) and it is subject to many variants to improve its robustness and computational time [Jiang(2020)].

2. Literature on the evaluation of point cloud registration

The robustness, speed and accuracy of registration algorithms are impacted by large relative transformations, measurement noise, partial overlaps, density variations and symmetries in the data. In the literature, the accuracy of these algorithms is evaluated with synthetic point clouds based on pre-generated data sets, usually from a sampling of CAD models. Data augmentation is performed on these point clouds to study the impact of different transformations, noise, overlaps, etc., on the registration accuracy [Zhu(2019)].

Point clouds have a predefined relative transformation, and the accuracy of the registration is often quantified by the difference between the predefined and estimated relative transformation, or the difference in point positions. This metric, as well as the cost functions to minimize in ICP variants, tend to approximate the orthogonal deviation of the points from the underlying surface. However, no constraints have been applied to the point cloud generation. Constraints could ensure that the minimum deviations are achieved at the predefined relative transformation to qualify the generated point clouds as reference data.

[1] Laboratoire Commun de Métrologie (LNE-CNAM), 1 Rue Gaston Boissier, 75015, Paris, France, e-mail: louis-ferdinand.lafon@lne.fr

[2] Université Paris-Saclay, ENS Paris-Saclay, université Sorbonne Paris Nord, LURPA, 4 avenue des Sciences, 91190, Gif-sur-Yvette, France

Similar constraints have been used in point cloud generation to evaluate the fitting of form by association [Forbes(2011)]. For these evaluations, reference data are inputted to the software under test and the returned value is compared to the reference value with known uncertainty. Afterwards, the software under test is validated based on a performance measure and degree of difficulty.

3. Proposed reference data generation

In this study, a metric and a reference data generation method are proposed for the point cloud registration process. The metric is based on quadratic orthogonal deviations $\boldsymbol{d}$ between estimated point positions in one point cloud and the underlying surface in the other point cloud. Therefore, the registration process is formulated as a Least Squares Orthogonal Distance Regression (LSODR) problem. Then, the reference data generation is based on a nullspace approach to obtain a misaligned point cloud pair with known predefined rigid transformation $\boldsymbol{T}$ (Fig. 1).

This method is applied on ICP variants to evaluate their accuracy, and the uncertainty on the estimated transformation is evaluated by Monte Carlo simulations. The generalization of this data generation method is also discussed for non-rigid and groupwise registrations with the aim to evaluate additional data fusion processes in future work.

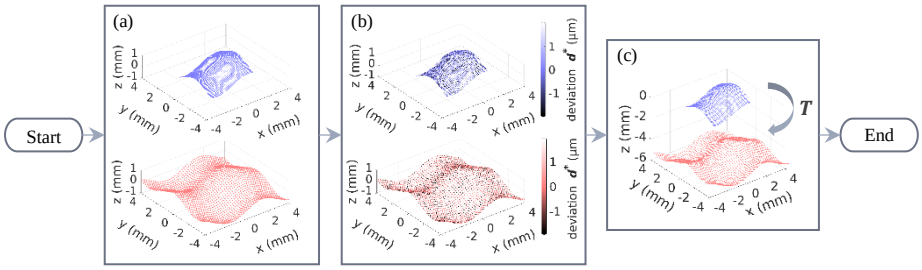


Figure 1: Reference point cloud pair generation on a sinusoidal surface: (a) points sampling; (b) addition of constrained orthogonal deviations $\boldsymbol{d}^*$; (c) application of a predefined relative transformation $\boldsymbol{T}$.

Acknowledgements This project (20IND07 TracOptic) 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. A. B. Forbes and H. D. Minh. 'Form Assessment in Coordinate Metrology'. In: Approximation Algorithms for Complex Systems. Springer Proceedings in Mathematics. Springer, Berlin, Heidelberg, 2011, pp. 69-90. doi:10.1007/978-3-642-16876-5_4
2. X. Jiang, M. Liu, Y. Huang, and R. Luo, 'Review on improved algorithms based on ICP algorithm', ICCEIC 2020, Nov. 2020, pp. 185–189. doi:10.1109/ICCEIC51584.2020.00043.
3. A. Weckenmann et al., 'Multisensor data fusion in dimensional metrology', CIRP Annals, vol. 58, no. 2, pp. 701– 721, Jan. 2009, doi:10.1016/j.cirp.2009.09.008.
4. H. Zhu et al., 'A Review of Point Set Registration: From Pairwise Registration to Groupwise Registration', Sensors, vol. 19, no. 5, Art. no. 5, Jan. 2019, doi:10.3390/s19051191.

Simulation of acquisition process in Magnetic Resonance Imaging to support standardization

Riccardo Ferrero[1], Marta Vicentini[1], Alessandra Manzin[1] and the rest of the iMet-MRI Consortium[2]

Key words: Quantitative Magnetic Resonance Imaging, Bloch Equations, Synthetic Medical Imaging, Numerical Models

1. Introduction

Magnetic Resonance Imaging (MRI) is a widely diffused, non-invasive diagnostic imaging technique used in clinics to acquire detailed information about anatomy and functions of different organs, in both health and disease conditions [Keenan (2018)]. MRI provides excellent anatomical details due to its high soft-tissue contrast and the possibility to differentiate between several types of tissues by means of different acquisition protocols. However, MRI measurements are strongly susceptible to differences in the image acquisition parameters and used scanners, making complicated the comparison of data obtained in large-scale trials and multi-site studies.

To validate the techniques adopted for the analysis of the acquired images, reference data (ground truth) are required. These can be provided by computer-generated virtual models (digital phantoms) with known anatomy and physiological functions, to be used as a gold standard for evaluating and improving MRI devices, image processing and reconstruction techniques.

In the framework of the European Metrological Project 20NRM05 iMET-MRI "Improved Metrology for Quantitative MRI" [Hall (2022)], we have developed an MRI simulator based on the Bloch equations to produce realistic images, first, as a ground truth for signal generation and image production process and, second, as a validation tool for quantitative analysis. The solver is parallelized and developed to run on a CUDA compatible GPU-accelerated environment.

2. Methodology and results

The MRI solver we have implemented enables us to compute, in a voxel-based domain representing the digital phantom, the Bloch equations:

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \gamma \mathbf{M} \times \mathbf{B} + \frac{M_0 - M_z}{T_1}\mathbf{e}_z - \frac{M_x}{T_2}\mathbf{e}_x - \frac{M_y}{T_2}\mathbf{e}_y \quad (1)$$

$$\mathbf{B}(\mathbf{r},t) = B_0\mathbf{e}_z + \left[\mathbf{G}(t)\cdot\mathbf{r}\right]\mathbf{e}_z + \mathbf{B}_1(t) \quad (2)$$

[1] Riccardo Ferrero e-mail: r.ferrero@inrim.it · Marta Vicentini e-mail: m.vicentini@inrim.it ·Alessandra Manzin, e-mail: a.manzin@inrim.it
INRiM, Strada delle Cacce 61, Torino, Italy

[2] Improved Metrology for quantitative MRI, http://empir.npl.co.uk/imet-mri/

where $\mathbf{M} = (M_x, M_y, M_z)^T$ is the macroscopic nuclear magnetization vector within a voxel, M_0 is the amplitude of the magnetization at equilibrium and γ is the gyromagnetic ratio. These equations describe the precessional motion of **M**, under the effect of a static magnetic field $\mathbf{B}_0$, a radio-frequency (RF) contrast-generating magnetic field pulse sequence $\mathbf{B}_1$ and a set of imaging/encoding time-dependent magnetic field gradients **G** [Webb (2016)]. The solver allows the simulation of the entire MRI signal generation process, calculating the k-space data that can then be converted into an image via inverse Fourier transform.

To each voxel of the digital phantom we assign, as material properties, the two relaxation times T_1 and T_2, and the proton density. The exciting and encoding fields are inputted to the solver with a Python script based on the PyPulseq library [Ravi (2019)]. The magnetization components are acquired at precise time intervals producing the MRI signal, which permits to generate the synthetic image.

As an example, we reproduce the behaviour of a reference phantom with known material properties under the effect of a simple MR sequences, with the aim of recovering the T_2 material parameter. To this aim, we calculate a series of spin-echo images, acquired at different echo time values TE (5 ms, 25 ms, 50 ms and 100 ms). Each image has a resolution of 256 x 256 pixels. Using the grey scale levels, a parameter fitting analysis has been performed to reconstruct the value of T_2 of each tissue-like material. As shown in Fig. 1, the true T_2 values are accurately reconstructed ,also considering the pixel-pixel variance.

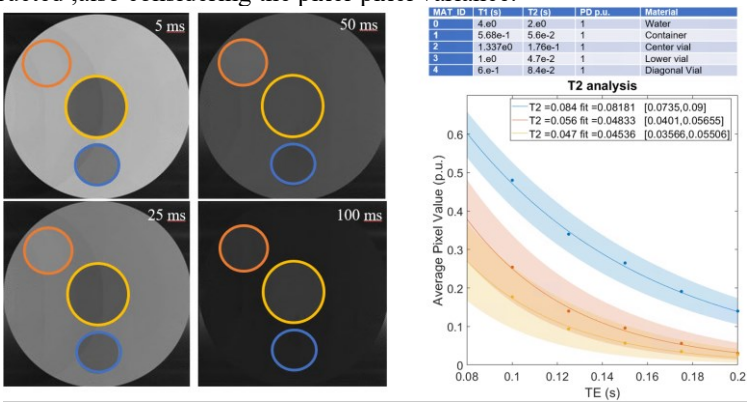


Figure 1: Images from a spin-echo sequence applied on a digital phantom containing three objects with different T_2 values, calculated at echo time TE of 5 ms, 25 ms, 50 ms and 100 ms. The T_2 values are reconstructed with the estimated error depending on the variation inside the material circled region.

References

[Hall (2022)] Hall M, et al (2022) 283 - IMET-MRI – Update on European project aiming at improving metrology in quantitative MRI. Physica Medica 104:S40–S41.

[Keenan (2018)] Keenan KE, et al (2018) Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom. Magn. Reson. Med. 79(1):48–61,

[Ravi (2019)] Ravi KS, et al (2019) Pypulseq: A python package for mri pulse sequence design. J. Open Source Softw. 4(42):1725, DOI 10.21105/joss.01725

[Webb (2016)] Webb AG (ed) (2016) Magnetic Resonance Technology. New Developments in NMR, The Royal Society of Chemistry

Monte Carlo simulations for uncertainty estimation of error separation techniques

Saint-Clair T. Toguem[1] and Charyar Mehdi-Souzani[2] and Nabil Anwer[2] and Hichem Nouira[1]

Key words: Monte Carlo simulations, Uncertainty, Error separation technique, r Roundness.

1. Introduction

The circularity is an essential geometrical characteristic for assessing the roundness of critical mechanical components such as ball bearings rolling elements and piston-cylinder. However, various sources of error limit its assessment. Precision design principles [Leach2018] applied in the development of accurate circularity measuring devices are often completed by error separation techniques (ESTs). Although ESTs leads to enhanced measurement uncertainty, they still have residual errors due to positioning errors, environmental disturbances and probe non-linearity.

Traditional methods for quantifying these errors effects provide limited information and rely on assumptions that can bias uncertainty estimates. Monte Carlo simulations are a better tool for accurately estimating measurement uncertainty and identifying individual contributions. This study applies Monte Carlo simulations to the multi-step EST for roundness measurements. The aim is to identify the most important contributors to the uncertainty and to improve the accuracy.

2. Monte Carlo simulations, discussions and conclusion

The measurement datasets (measured part form p, spindle synchronous/asynchronous error motion B) are generated by simulation with harmonics of 1µm amplitude and frequencies in the range of [2, 15] upr. The main error sources and corresponding distributions for the multi-step EST, including the measured part positioning errors, probe nonlinearity and thermal drift [Cappa2014], are selected in the range of 0.02°, ±5 nm and ±0.5 °C respectively, to match the actual experimental observations. The calculations are performed over 200 000 iterations (Figure 1 (a)-(c)).

[1]Laboratoire Commun de Métrologie, Laboratoire National de Métrologie et d'Essais (LNE-CNAM), 1 Rue Gaston Boissier, 75015 Paris, France, e-mail: saint-clair.toguemtagne@lne.fr
[2]Université Paris-Saclay, Université Sorbonne Paris Nord, ENS Paris-Saclay, LURPA, 91190, Gif-sur-Yvette, France

As expected [Cappa2014], the results given in Figure 1 (d)-(i), show that the measurement uncertainty is inversely proportional to the number of measurements N. The uncertainty associated with multi-step can be reduced by a factor of ~1.5, by doubling the number of measurements under similar environmental conditions. It can be noted that the probe nonlinearity stands as the major contributor to the uncertainty in the given measurement conditions. Further analysis includes the examination of different experimental conditions and ESTs such as reversal and multiprobe.

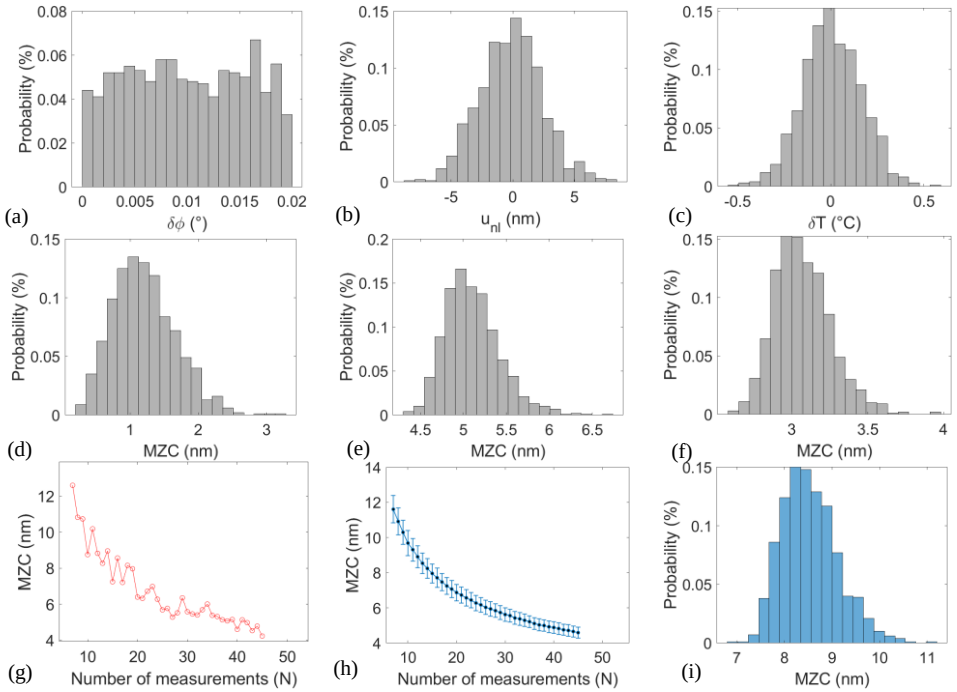


Figure 1: MC simulation results; distributions of error sources and their effects for *N=12*: (a) positioning errors, (b) nonlinearity, (c) thermal gradient, (d) positioning error effect, (e) nonlinearity effect, (f) thermal gradient effect; (g) resulting measurement error, (h) Std of resulting measurement error and (i) distribution of resulting measurement error

References

1. Leach, R., Smith, S. T.: Basics of Precision Engineering : CRC Press, 2018. — Google-Books-ID: 3GdQDwAAQBAJ — ISBN 978-0-429-88744-4
2. Cappa, S., reynaerts, D., AL-Bender, F.: A sub-nanometre spindle error motion separation technique. In: Precision Engineering Bd. 38 (2014), Nr. 3, S. 458–471
3. Neugebauer, M.: Uncertainty analysis for roundness measurements by the example of measurements on a glass hemisphere. In: Measurement Science and Technology Bd. 12, IOP Publishing (2000), Nr. 1, S. 68–76

Reference data for Electrical Resistance Tomography

Alessandro Cultrera, Gabriele Germito, Luca Callegaro

Key words: Electrical Resistance Tomography, Inverse Problems, Reference Data, Open Data.

Abstract

The assessment of the electrical properties of thin-film and nano-structured materials over large area (e.g. like graphene and topological insulators, nano-wire networks, conductive polymers) is of primary importance for the up-scaling of laboratory advances in these fields [1, 2, 3]. Electrical Resistance Tomography (ERT) is a technique that allows for the mapping of the electrical conductivity of thin film materials by performing only boundary electrical measurements [4]. In particular, the ERT boundary measurements consist in a series of four-terminal resistance measurements performed over a number of contacts, following a suitable measurement sequence. The 1D measurements vector, is then used as the input quantity for a suitable solver to retrieve a 2D conductivity distribution array; this calculation represents an ill-posed inverse problem [5].

Many implementations of inverse solvers, some of which of particular interest for ERT in material science, have been proposed [6]. Since ERT represents both a challenge in terms of measurement science and of algorithms development, the availability of open and interchangeable reference datasets would be of great interest for a fruitful exchange among different scientific communities. The last draft

Alessandro Cultrera
INRIM, strada delle cacce, 91 10135, Torino, Italy, e-mail: a.cultrera@inrim.it

Gabriele Germito
INRIM, strada delle cacce, 91 10135, Torino, Italy, e-mail: g.germito@inrim.it

Luca Callegaro
INRIM, strada delle cacce, 91 10135, Torino, Italy, e-mail: l.callegaro@inrim.it

of the MATHMET's Strategic Research Agenda[1], sec. 4.2.1, cites ERT as one of the techniques that would take advantage form the availability of such a reference dataset[2]. At the conference we will present a first draft of a machine-readable reference dataset structure for ERT measurements on thin film materials, based on `XLM` language. Reference data will include ERT measurements, along with a set of metadata, such as information about the samples (geometry, materials), the measurement protocol used to perform the ERT multiterminal measurements, the measurements uncertainty, instrumentation specifications.

Acknowledgements The work has been realized within the Joint Resarch Projects 20FUN03 COMET and 20FUN06 MEMQUD. These projects received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme. This work is also funded by the EU under the HORIZON EUROPE Project 101091997 — EMPHASIS.

References

1. P. Bøggild, "Research on scalable graphene faces a reproducibility gap," *nature communications*, vol. 14, no. 1, p. 1126, 2023.
2. T. Sannicolo, M. Lagrange, A. Cabos, C. Celle, J.-P. Simonato, and D. Bellet, "Metallic nanowire-based transparent electrodes for next generation flexible devices: a review," *Small*, vol. 12, no. 44, pp. 6052–6075, 2016.
3. G. Tarabella, D. Vurro, S. Lai, P. D'Angelo, L. Ascari, and S. Iannotta, "Aerosol jet printing of pedot: Pss for large area flexible electronics," *Flexible and Printed Electronics*, vol. 5, no. 1, p. 014005, 2020.
4. A. Cultrera, D. Serazio, A. Zurutuza, A. Centeno, O. Txoperena, D. Etayo, A. Cordon, A. Redo-Sanchez, I. Arnedo, M. Ortolano, *et al.*, "Mapping the conductivity of graphene with electrical resistance tomography," *Scientific reports*, vol. 9, no. 1, p. 10655, 2019.
5. M. Vauhkonen, D. Vadász, P. A. Karjalainen, E. Somersalo, and J. P. Kaipio, "Tikhonov regularization and prior information in electrical impedance tomography," *IEEE transactions on medical imaging*, vol. 17, no. 2, pp. 285–293, 1998.
6. Z. Cui, Q. Wang, Q. Xue, W. Fan, L. Zhang, Z. Cao, B. Sun, H. Wang, and W. Yang, "A review on image reconstruction algorithms for electrical capacitance/resistance tomography," *Sensor Review*, vol. 36, no. 4, pp. 429–445, 2016.

[1] "MATHMET Strategic Research Agenda: Supporting the EMN Mathmet by addressing EURAMET's key priorities and challenges in mathematics and statistics in metrology"

[2] Deliverable D3 on prioritisation task of the EMPIR project 18NET05 MATHMET.

The thermal dynamics of a brake pad, and the estimation of its thermal parameters

Francesca Collini, Stefano Serra, Pietro Macchi, Enrico Bibbona

Key words: Brake Pads, mathematical model, braking pad thermal dynamics, parameter estimation, statistical inference ...

It is well known that repeated or sustained application of the brakes can heat up the pads and the rotor so much that the braking performances get dramatically reduced. This phenomenon, called fading, could be easily prevented if the rotor temperature was known while driving. Unfortunately, such a temperature can be measured only using experimental setups not feasible for on board applications. Despite the main physical laws governing heat transfer are well known, a reliable mathematical model to predict the temperature of interest from the dynamics of the car is not currently available. The main difficulties are both the intrinsical aleatority of breaking events and the lack of any information about the temperatures inside the braking system in working conditions (high temperatures, high pressure, wear, dust...) that makes it impossible to tune the model so that it does not drift away from real data.

A recent breakthrough in this field is the invention of GALT. Smart Pad by ITT that makes possible to measure the temperature at the backplate of the Pad. Despite it is not exactly the temperature of interest, we aim at using it to estimate the rotor temperature. The heat equation, indeed, can be solved to derive the entire temperature

Francesca Collini
DISMA, Politecnico di Torino, e-mail: francesca.collini@polito.it

Stefano Serra
ITT Friction Technologies, e-mail: stefano.serra@itt.com

Pietro Macchi
ITT Friction Technologies, e-mail: pietro.macchi@itt.com

Enrico Bibbona
DISMA, Politecnico di Torino, e-mail: francesca.collini@polito.it

profile of the pad, given as side conditions both the measured backplate temperature and the estimated energy instantaneously dissipated in the braking events. In view of an application of such model to predict the temperature of interest, we propose a method to solve the inverse problem, i.e. to infer the (some) thermal parameters of the friction material, knowing the temperature of some points inside the friction material itself (including backplate and pad/rotor interface temperatures). This talk will describe such first step towards our goal, that also provides us with a validation of the model.

Employing machine learning models to enhance the prediction of cocrystals formation

Rebecca Birolo[1] and Alessandro Cossard[1] and Lorenzo Castellino[1] and Eugenio Alladio[1] and Michele Remo Chierotti[1]

Key words: Cocrystals, Machine Learning, Prediction, Pharmaceutics

1. Abstract

Due to their promising performance for data-driven prediction, data-driven machine learning (ML) methods have become increasingly popular in the chemical and pharmaceutical fields. Among the several application areas, algorithms such as Partial Least Squares – Discriminant Analysis (PLS-DA), Random Forest (RF), Neural Networks (NN) and Support Vector Machines (SVM) exhibit promising results in the context of the prediction of cocrystal formation (Mswahili, 2021).

Pharmaceutical cocrystals are crystalline materials composed of at least two molecules, i.e., an active pharmaceutical ingredient (API) and a coformer, assembled by non-covalent forces. Cocrystallization is successfully applied to improve the physicochemical properties of APIs, such as solubility, dissolution profile, pharmacokinetics, and stability. However, choosing the ideal coformers is a challenging task in terms of time, effort and laboratory resources (Cerreia Vioglio, 2017) (Dai, 2018).

Several computational tools and ML models were proposed to mitigate the problem. However, the challenge of achieving a robust and generalizable predictive method is still open (Devogelaer, 2020) (Hao, 2022).

In this study, we propose a new approach to quickly predict the formation of cocrystals, employing the PLS-DA, RF, SVM and NN algorithms. The models were based on a several training sets of cocrystallization containing both positive and negative experimental outcomes. At the same time, the features were specially selected from a variety of molecular descriptors to explain the phenomenon of

[1]Rebecca Birolo
University of Turin, Via P. Giuria 7, 10125, Torino, Italy rebecca.birolo@unito.it
[1]Alessandro Cossard
University of Turin, Via P. Giuria 7, 10125, Torino, Italy alessandro.cossard@unito.it
[1]Lorenzo Castellino
University of Turin, Via P. Giuria 7, 10125, Torino, Italy lorenzo.castellino@unito.it
[1]Eugenio Alladio
University of Turin, Via P. Giuria 7, 10125, Torino, Italy eugenio.alladio@unito.it
[1]Michele Remo Chierotti
University of Turin, Via P. Giuria 7, 10125, Torino, Italy michele.chierotti@unito.it

cocrystallization and, in parallel, molecular descriptors are being tested as predictors.

The proposed ML models showed, on average, cross-validation accuracy above 70%. Furthermore, this approach was successfully applied to drive the cocrystallization experimental tests of 2-phenylpropionic acid, showcasing the high potential of the ML models in practice.

References

1. Cerreia Vioglio, P., Chierotti, M.R., Gobetto, R. Pharmaceutical Aspects of Salt and Cocrystal Forms of APIs and Characterization Challenges. Adv. Drug Deliv. Rev. (2017) doi: 10.1016/j.addr.2017.07.001.
2. Dai, X.L., Chen, J.M., Lu, T.B. Pharmaceutical Cocrystallization: An Effective Approach to Modulate the Physicochemical Properties of Solid-State Drugs. Cryst. Eng. Comm. (2018) doi: 10.1039/C8CE00707A.
3. Devogelaer, J., Meekes, H., Tinnemans, P., Vlieg, E., Gelder, R. Co-crystal Prediction by Artificial Neural Networks. Chem. Int. Ed. (2020) doi: 10.1002/anie.202009467.
4. Hao, Y., Hung, Y. C., Shimoyama, Y. Investigating Spatial Charge Descriptors for Prediction of Cocrystal Formation Using Machine Learning Algorithms. Cryst. Growth. Des. (2022) doi: 10.1021/acs.cgd.2c00812.
5. Mswahili, M.E., Lee, M.J.; Martin, G.L., Kim, J.; Kim, P., Choi, G.J., Jeong, Y.S.: Cocrystal Prediction Using Machine Learning Models and Descriptors. Appl. Sci. (2021) doi: 10.3390/app11031323.