EMPIR Call 2017 – Industry, Fundamental, Normative and Research Potential

Selected Research Topic number: **SRT-f01**
Version: 1.0

Title: Large-scale statistical data analysis and computation for metrology

Abstract
New measurement modalities, such as quantitative imaging in medical physics, inverse problems in nanometrology, optical form metrology, macroscopic flow measurements, state estimation in the electrical power grid, earth observation and sensor networks for environmental and climate monitoring, produce large data sets and involve large numbers of parameters. Such systems require large-scale methods for uncertainty quantification (enabling them to be traceable to SI) and large-scale data analysis, such as machine learning technologies or Bayesian approaches, in order to extract new knowledge about the underlying problems.

Keywords
Statistical data analysis, Bayesian methods, uncertainty quantification, Monte Carlo methods, simulation, inverse problems, machine learning, Big Data

Background to the Metrological Challenges
In recent years, metrology has expanded to new fields to address societal challenges relating to energy and sustainability, climate and environmental monitoring, life sciences and health, using new measurement modalities such as imaging, spectroscopy, earth observation and sensor networks. In these applications, the "measurand" consists of a large, and sometimes huge, number of unknown quantities to be inferred from a large amount of measurements. In addition, the high-dimensional raw data need to be pre-processed, cleaned and classified before analysis methods can be used to extract the underlying measurement information.

A central feature of metrology is that all measurement results must be accompanied by an uncertainty statement. The concept of traceability depends fundamentally on uncertainty quantification. The treatment of measurement models and the evaluation of associated measurement uncertainties are summarised in the GUM. However, the GUM approach essentially involves uncertainty propagation through a known, forward model. Most measurement problems are in fact inverse problems, for which Bayesian inference methods are more appropriate. These and other current statistical methods are well suited to deal with small to modest scale applications. However, they are not directly applicable for the emerging large-scale tasks which require a substantial extension of approaches for data analysis.

Objectives
Proposers should address the objectives stated below, which are based on the PRT submissions. Proposers may identify amendments to the objectives or choose to address a subset of them in order to maximise the overall impact, or address budgetary or scientific / technical constraints, but the reasons for this should be clearly stated in the protocol.

The JRP shall focus on the development of new, large-scale methods for uncertainty quantification to enable such systems deliver measurement results traceable to the SI.

The specific objectives are

1. To identify and prioritise metrological applications that can benefit from the proposed large-scale data analysis and computation methodologies.
2. To develop large-statistical data analysis methods (e.g. estimation, Bayesian inference, machine learning) and computational tools (e.g. optimisation, regression, Monte Carlo sampling, Approximate Bayesian Computation) that have the potential to be relevant for metrology.

3. To develop and implement methods for comprehensive uncertainty quantification in large-scale systems and networks and provide associated guidelines, software and case studies.
4. To build a network of excellence of researchers from academia, industry and metrology organisations to benefit from and steer the project – if suitable under the umbrella of the European Centre for Mathematics and Statistics in Metrology (MATHMET).

Proposers shall give priority to work that aims at excellent science exploring new techniques or methods for metrology and novel primary measurement standards, and brings together the best scientists in Europe and beyond, whilst exploiting the unique capabilities of the National Metrology Institutes and Designated Institutes.

Proposers should establish the current state of the art, and explain how their proposed project goes beyond this. In particular, proposers should outline the achievements of the EMRP project NEW04 "Novel mathematical and statistical approaches to uncertainty evaluation" and how their proposal will build on those.

EURAMET expects the average EU Contribution for the selected JRPs in this TP to be 1.5 M€, and has defined an upper limit of 1.8 M€ for this project.

EURAMET also expects the EU Contribution to the external funded partners to not exceed 40 % of the total EU Contribution to the project.

Potential Impact

Proposals must demonstrate adequate and appropriate participation/links to the "end user" community, describing how the project partners will engage with relevant communities during the project to facilitate knowledge transfer and accelerate the uptake of project outputs. Evidence of support from the "end user" community (e.g. letters of support) is also encouraged.

You should detail how your JRP results are going to:

- Address the SRT objectives and deliver solutions to the documented needs,

You should detail other impacts of your proposed JRP as specified in the document "Guide 4: Writing Joint Research Projects (JRPs)"

You should also detail how your approach to realising the objectives will further the aim of EMPIR to develop a coherent approach at the European level in the field of metrology and include the best available contributions from across the metrology community. Specifically the opportunities for:

- improvement of the efficiency of use of available resources to better meet metrological needs and to assure the traceability of national standards
- the metrology capacity of EURAMET Member States whose metrology programmes are at an early stage of development to be increased
- organisations other than NMIs and DIs to be involved in the work

Time-scale

The project should be of up to 3 years duration.