Calibrating stochastic models for understanding solar activity

Research Group Biomedical Simulation

Dr. Simone Ulzega
Senior Research Scientist,
ulzg@zhaw.ch

Research project
Bayesian Inference with Stochastic Models (BISTOM)

Lead:
Dr. Carlo Albert (EAWAG, ETHZ)

Role ZHAW:
Development of inference algorithms, parallelization for high-performance computing

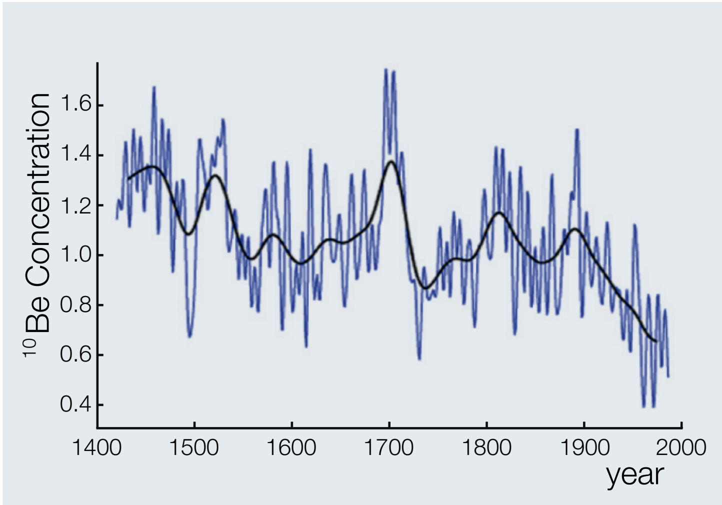
Duration:
2 years
(01.04.2018–31.03.2020)

Partners:
SDSC, Eawag, ZHAW, Università della Svizzera Italiana (USI), Paul Scherrer Institute (PSI)

Funding:
Swiss Data Science Center (SDSC, ETHZ)

One of the most fundamental questions in essentially all applied sciences is how to predict the dynamics of complex systems. Complex systems are everywhere, in chemistry, biology, physics, engineering, economy, environmental, life and social sciences.

To tackle complexity, scientists often have to resort to simplified conceptual models that incorporate only a small selection of system variables and parameters. In such framework, only a few dominant dynamic processes occurring on our observation scale are described by deterministic differential equations, while all other processes are included in the model as *noise.* This naturally leads to *stochastic differential equation models.* Stochastic models take uncertainties intrinsic to dynamic processes into account thus providing more realistic descriptions of real systems. However, for making reliable probabilistic predictions, model parameters need to be soundly calibrated to measured data and their uncertainty needs to be quantified. Parameter inference, as this process is called, is a fundamental problem in data-driven modeling. Bayesian statistics is a consistent framework for parameter inference where knowledge about model parameters is conveniently expressed through probability distributions and updated using measured data. However, Bayesian inference with non-trivial stochastic models can become mathematically and computationally extremely challenging, and it is therefore hardly ever applied. In recent years, sophisticated and scalable algorithms have emerged, which have the potential of making Bayesian inference for complex stochastic models feasible, even for very large data sets. In the framework of a 2-year project funded by the Swiss Data Science Center (SDSC), in collaboration with Eawag, USI and PSI, we will explore the power and versatility of two classes of Bayesian inference algorithms, that is, Approximate Bayesian Computation (ABC) and Hamiltonian Monte Carlo (HMC) methods. The former is well-known and technically easy to apply but yields only approximate results, while the latter requires much more tailoring to a particular problem but has the potential of yielding exact results. The HMC algorithm, as recently proposed by Carlo Albert, Simone Ulzega and Ruedi Stoop (Albert et al. Phys. Rev. E 93, 2016), is raising great attention in various scientific communities due to its exceptional efficiency and high parallelizability. An efficient parallelization of the HMC algorithm, in particular, will be of paramount importance for making Bayesian inference amenable to a «Big Data» context.



Berillium-10 concentration (10^4 atoms/gram of ice) in polar ice cores, Greenland. Solar activity varies inversely with the concentration of the radioisotopes.

We will focus on a real case study in solar physics. Time-series of cosmogenic radionuclides, that is, radioactive Carbon-14 and Berillium-10 nuclei produced in the Earth's atmosphere by galactic cosmic rays and stored in wood and polar ice cores, are an exceptional proxy for solar activity on multi-millennial time-scales. Cosmic rays are in fact modulated by solar magnetic fields and the production rates of these isotopes is thus modulated by the solar magnetic activity. These time-series exhibit a number of interesting and mostly not-yet-understood features such as stable cycles and intermittency.
Solar physicists have put a lot of effort into the development of stochastic solar dynamo models, which need to be calibrated to the observations. Parameter inference for stochastic dynamo models on long time-series of radionuclides is currently an open and highly topical question in solar physics. Achieving more reliable predictions of the solar activity may have important implications also for our understanding of the Earth's climate. ■