Post-doctoral offer: Objectivation in Bayesian modeling for interpretable decision support

July 2024

1 Context

The Sustainable Energies Chair of the *Ecole Polytechnique* is offering a 1-year post-doc position, renewable once and starting before December 2024, on a subject related to improving the objectivity of Bayesian modeling rules in view of supporting interpretable decision. The situations concerned with this quantification problem of uncertain knowledge are numerous, and are particularly motivated within the framework of this work by applications interesting the field of decarbonated energies.

The technical motivation for this work stems from the research program proposed in [4], that extends the proposals made in [3]. Bayesian modeling choices play an important role in the context of forecasting and decision support through statistical learning applied to costly or rare data. These are characteristic of the risk situations affecting the industrial world. Their promise is to be able to model the uncertain knowledge required to inform models in addition to available data, or even in their absence in the event of a regime change. Major applications related to decarbonated energies are the following (among others):

- rapid location of nuclear material in waste packages, in order to apply focused spectrometry techniques to identify fission products;
- the reliability of important or even critical industrial components, such as steam generators, batteries, valves, etc. ;
- quantification of the intensity of extreme natural phenomena (torrential rain, marine submersion, floods, etc.);
- calibration of technico-economic models used to optimize the design and operation of energy parks, particularly when the depth of history is shallow (e.g. offshore park deployments).

Objectivation aims to respond to the obstacles that are generally blamed on the corpus of existing Bayesian methods, despite their huge number [7], and which still limit their use in proposal submitted to safety and control authorities in the energy sector: the low repeatability of methodologies, the lack of control over subjective elements in modeling choices, and the lack of interpretability of models. See [12] for more details. An additional difficulty is to propose calibration rules based on the repeatability of experiments.

2 Work program

The work will aim to bring together a set of known "methodological constraints", still separated, into a single approach that will extend methodologies already established for sub-families of models (in particular exponential families and conjugate models). Such constraints are, for instance, related to prior-data conflict [2, 10] or q-vague convergence [1].

This approach will thus focus on defining a "modeling continuum" limited by objective reference priors [8, 13] and so-called "Posterior Priors" models resulting from the application of Bayes' theorem [11, 6, 5]. Such approaches are interpretable because they consider that the available information can be assimilated to that provided by data not directly known, but that it can be manipulated by explicit approximation techniques outside conjugate families [3]. Such approaches participate to provide rules for clarifiying the meaning of Prior Effective Sample Sizes (PESS), the definition of which is still the subject of debate in the community (e.g., [9]).

The resulting methodology will be submitted to leading scientific journals (e.g. *Bayesian Analysis*) and re-used in many sectors beyond the energy industry.

3 Supervision

The work will be supervised by Professor Josselin Garnier (École Polytechnique / CMAP-Centre de Mathématiques Appliquées), in collaboration with Dr. Nicolas Bousquet, senior researcher at EDF R&D. CMAP and EDF have been working together for a very long time. CEA DES, a frequent partner of EDF and CMAP, is also keen to participate in discussions during the post-doctoral work.

Contacts

`josselin.garnier@polytechnique.edu`

`nicolas.bousquet@edf.fr`

4 Application

This o
ffer is available until early 2025.

The candidate should have a PhD thesis in statistics or applied mathematics, with a good knowledge of Bayesian statistics. A good knowledge of mathematical tools related to the approximation of probability distributions and non-convex optimization would be a plus.

The candidate will join CMAP's SIMPAS (*Statistique Apprentissage Simulation Image*) team at École Polytechnique, located in Palaiseau, France. École Polytechnique specializes in science and engineering. CMAP conducts theoretical and numerical research on mathematics in interaction with other sciences (biology, economics, computer science, mechanics, physics, etc) or in connection with industrial or societal applications. Its specialties are numerical analysis, scientific computing, control, artificial intelligence, modeling, optimization, probability, signals, statistics, etc.

The candidate will become involved in the *Uncertainty Quantification* thematic network (formerly known as GDR MASCOT-NUM). In addition, the candidate will benefit from an industrial environment strongly interested in this work, and from a French and international network in Bayesian modeling (especially, the Bayesian Group at SFdS, the APPLIBUGS Group, the ISBA Community).

5 Remuneration

The successful candidate will receive an annual salary around 55,000 euros (depending on experience) and travel assistance in the Paris region. He/she will be given the means to take part in national and international conferences.

References

[1] C. Bioche and P. Druilhet. Approximation of improper priors. *Bernoulli*, 22:1709–1728, 2016.

[2] N. Bousquet. Diagnostics of prior-data agreement in applied Bayesian analysis. *Journal of Applied Statistics*, 35:1011–1029, 2008.

[3] N. Bousquet. Towards new formal rules for informative prior elicitation ? a discussion on "specifying prior distributions in reliability applications". *Applied Stochastic Models in Business and Industry*, pages 1–11, 2023.

[4] N. Bousquet. *Contributions to the statistical quantification of uncertainties affecting the use of numerical models.* HDR, Sorbonne Université', 2024.

[5] N. Bousquet, M. Fouladirad, A. Grall, and C. Paroissin. Bayesian Gamma processes for optimizing condition-based maintenance under uncertainty. *Applied Stochastic Models in Business and Industry*, 31:360–379, 2015.

[6] D. Fouskakis, I. Ntzoufras, and K. Perrakis. Power-expected-posterior priors for generalized linear models. *Bayesian Analysis*, 13:721–748, 2016.

[7] A. Gelman and Y. Yao. Holes in bayesian statistics. *J. Physics G*, 48(1):014002, dec 2020.

[8] R. Kass and L. Wasserman. The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91:1343–1370, 1996.

[9] B. Neuenschwander, S. Weber, H. Schmidli, and A. O'Hagan. Predictively Consistent Prior Effective Sample Sizes. *Biometrics*, 76(2):578–587, 2020.

[10] D. Nott, M. Seah, L. Al-Labadi, M. Evans, H. Khoon Ng, and B.-G. Englert. Using Prior Expansions for Prior-Data Conflict Checking. *Bayesian Analysis*, 16(1):203 – 231, 2021.

[11] J. Pérez and J. Berger. Expected posterior prior distributions for model selection. *Biometrika*, 89:491–511, 2002.

[12] J. Sprenger. Bayesianism vs. Frequentism in statistical inference. In *Handbook of the Philosophy of Probability*, pages 382–405. Alan Hajek and Chris Hitchcock (eds.), Oxford University Press, 2016.

[13] A. Van Biesbroeck, C. Gauchy, C. Feau, and J. Garnier. Reference prior for bayesian estimation of seismic fragility curves. *Probabilistic Engineering Mechanics (in press)*, 2024.