Frederi Viens (Statistics & Probability)–Colloquium 2/13/2020 MSU

Title:
Statistics and Bayesian machine learning help reconstruct the Earth's past climate, and make projections to 2100


Abstract:
The framework of Bayesian statistics traces back more than 250 years, but it is only in the most recent decades that computational advances, particularly in hardware, have made it possible to unleash its full power on data analysis. We will allude, through several examples of diverse applications, to how this framework gives rise, in a rather intuitive fashion, to computational supervised machine learning schemes. These allow specific participatory modes of collaboration via prior elicitation, which are particularly useful to situations which are challenging to frequentist frameworks, such as observational studies, environmental space-time data analysis, and studies with multiple competing models. Modulo other expressed interests from interactions with the audience during this presentation, we plan to focus the bulk of our presentation on the use of hierarchical modeling and validation metrics for model selection in paleo-climatology for the late Holocene period 1000-1899, and for how these models can be extended to prediction tools for global temperature estimation to 2100. This work, which is joint with a STT Ph.D. student, two MS students from the Ecole Polytechnique in Palaiseau., France, and one colleague from the University of Costa Rica, concludes that, under realistic scenarios for future greenhouse gas emissions, compared to what the  climate community predicts, global warming will be significantly more severe

under the business-as-usual scenario, and will be less intense after mid century under the most aggressive emissions reductions scenario. Our projections are roughly in agreement with the community's projections under intermediate scenarios. These findings can be explained via our principled quantification of uncertainty (which is build into the Bayesian framework), a methodological aspect which is far from being widely adopted in climate science. Our methodology contains, as a by-product, a clear attribution of factors in what global external forces influence climate change. Time permitting, we will attempt to explain how this research differs from other applications of Bayesian hierarchical modeling, and what commonalities exist, by mentioning work on model averaging for competing models with non-identical application domains in nuclear physics, attribution of factors in hydrology and agro-ecology, predictive analytics for cropping recommendations based on farmer surveys, predictive analytics in the context of workplace satisfaction for nurses' work environment, and hierarchical modeling with time lag distribution in global agricultural total factor productivity. Time probably will not permit discussing computational technicalities behind Bayesian machine learning, including current mathematical questions using analysis on Wiener space which may give rise to highly simplified approximations of these machine-learning algorithms.