SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

Project Title:	Data-driven calibration of stochastic parametrization of IFS using approximate Bayesian computation
Computer Project Account:	spgbdutt
Start Year - End Year :	2020 - 2022
Principal Investigator(s)	Dr. Ritabrata Dutta
Affiliation/Address:	Department of Statistics, Warwick University, UK.
Other Researchers (Name/Affiliation):	Dr. Nils Wedi and Dr. Peter Dueben, ECMWF.

Summary of project objectives

The aim of the project was to develop Bayesian inferential techniques suitable to be applied in the setting of ensemble NWP models. Specifically, this could allow us to perform Bayesian inference of parametrization parameters from observations. Bayesian inference, opposed to frequentist inference, provides a better way to quantify uncertainty starting from previous knowledge. During the project, we have tested different Bayesian inference methodologies for the considered problem and acquired expertise on the meaning of the physical parameters.

Summary of problems encountered

After long investigation, we realised that current standard Bayesian methodologies are unsuitable for tackling tuning of physical parameterization parameters; that is due to the high cost of the NWP models and to their high dimensional parameter and domain space. However, we have leveraged the understanding acquired by studying the considered problem by pivoting our effort to the task of developing data-driven neural-network based probabilistic weather forecasting systems, with a novel training objective based on Scoring Rules, which show superior performance with respect to traditional training methods, lower training cost, and physical motivation. We descrive this in detail in the results section.

Experience with the Special Project framework

Our interaction with the administration regarding access to the supercomputing facility, application procedure or reporting our progress have been very smooth.

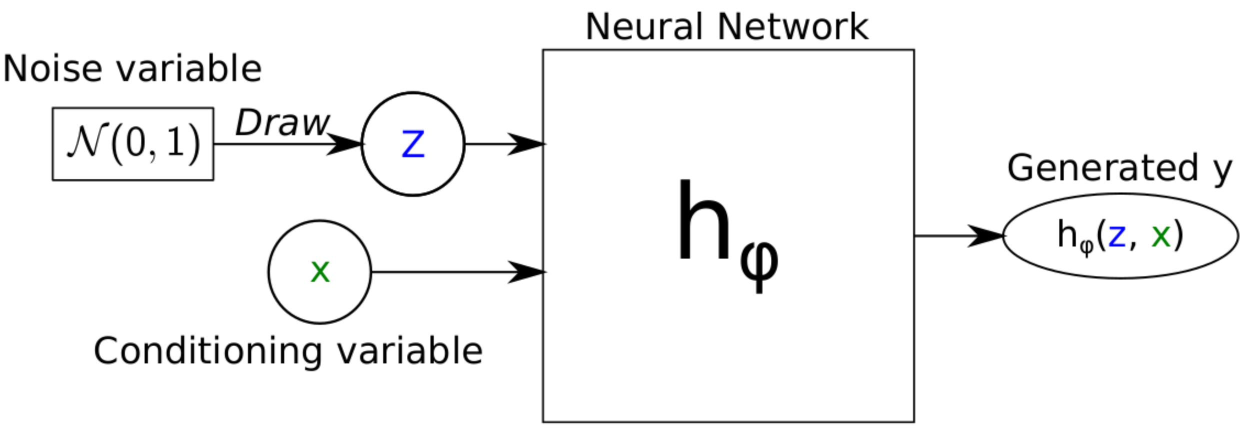
Summary of results

From the last report, we have spent some time thinking about the feasibility of the original project, namely tuning the spread parameters for IFS. We realised that standard Bayesian methodologies are not yet suitable to this task due to the high computational cost and the last parameter and data space. While investigating this problem, however, we became familiar with the concept of Scoring Rules and the theory behind probabilistic forecasting.

Parallely, we have been thinking on how to use neural networks to provide probabilistic forecasts for weather models, using the Weatherbench dataset (https://github.com/pangeo-data/WeatherBench) as a benchmark. We therefore decided to concentrate our efforts on this, being a more feasible task than the originally proposed one and equally impactful. We were therefore able to develop a new training method, based on scoring rules, for generative networks for probabilistic forecasting. In the following, we therefore briefly review our method, which can be found in Pacchiardi et al. (2021, https://arxiv.org/abs/2112.08217).

Generative networks

A generative network parametrizes a probability distribution by transforming a random variable Z with a simple distribution via a learnable transformation. Considering a conditioning variable x, we generate a draw $z \sim Z$ and compute $y = h_\phi(x, z)$, for a learnable transformation h_ϕ. In this way, y is a sample from a random variable $Y|x$ whose distribution is implicitly defined by the transformation and the base distribution for Z. If Z has a distribution which is easy to sample from, it is immediate to obtain a sample for Y. However, the density of Y is in general not available.



Typically, h_ϕ is parametrized by a neural network. An important task is that of fitting the resulting generative distribution to a set of observations. Given that the density of the generative distribution is unavailable, fitting ϕ via Maximum Likelihood is undoable. People usually recur to an *adversarial* formulation, where the generative network is trained in a zero-sum game against a *discriminator* network: the discriminator is tasked with distinguishing between true samples and draws from the generator, while the generative network attempts to fool the discriminator by generating samples which are as close as possible to the true data. At convergence, the generative network generates draws indistinguishable from the truth. However, training the neural networks in this formulation is tricky and requires a lot of hand tuning and computational power. In many cases, additionally, the resulting generative distribution underestimates the uncertainty.

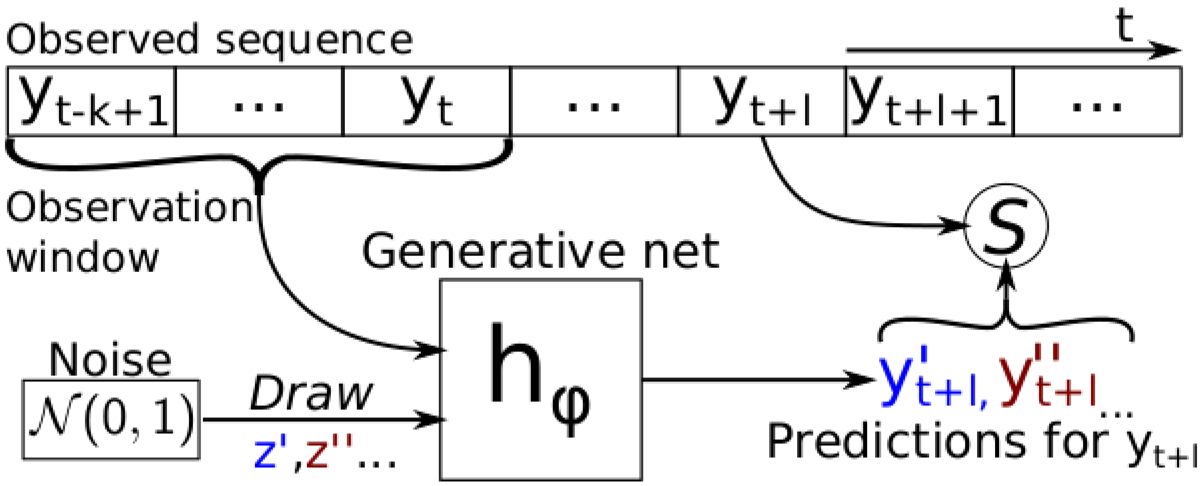
Scoring rules

Scoring rules are widely used by meteorologists to assess the predictive performance of probabilistic weather models; these work by considering the empirical distribution defined by the ensemble and by assessing the precision and dispersion of this distribution with respect to observations. More precisely, for a scoring rule S, $S(P, x)$ represents the penalty given to a probabilistic forecast P for an observation x. If x is drawn from a different distribution Q, the expected scoring rule is $S(P, Q) = E_{X \sim Q} S(P, X)$. A proper scoring rule is such that $S(P, Q)$is minimised in P when $P = Q$. Moreover, S is strictly proper if this is the unique minimum; therefore, strictly proper scoring rules encourage to provide predictions close to the data-generating process.

Scoring rules are commonly used in assessing the performance of probabilistic prediction. In what follows, we describe a training strategy for generative networks for probabilistic forecasting based on scoring rules.

Training generative networks using scoring rules

In Pacchiardi et al. (2021), we considered a conditional generative network tasked with forecasting with lag l given a past observation window of length k. Specifically, the generative network parametrizes therefore a probability distribution $P^{\phi}_{t+l}(\cdot\ |y_{t-k+1:t})$ for the observation y_{t+l}.



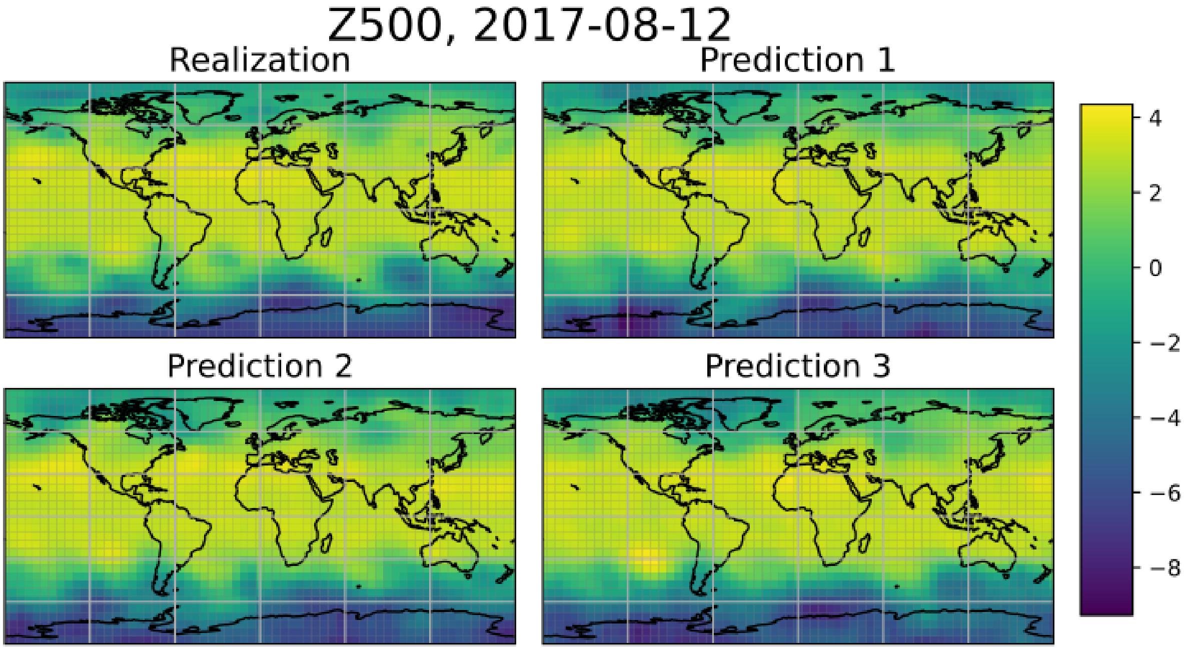
We consider therefore training the conditional neural network by minimising the following quantity:

$$\sum_{t=k}^{T-l} S(P^{\phi}_{t+l}(\cdot\ |y_{t-k+1:t}),\ y_{t+l})$$

where we assume to have an observation of length T. Here, each term of the sum is evaluating the performance of the generative network in forecasting the timestep $t + l$.

With some choices of scoring rules, Pacchiardi et al. 2021 shows how the generative network can be trained by minimising the above quantity with no need of an adversarial network, by only relying on draws from the generative network. We provide theoretical guarantees for the minimizer of the above quantity as the length of the observed time-series goes to infinity, under some stationarity and memory-less assumptions.

In experimental results, we show how the above method performs better than standard adversarial training approaches and is faster and easier to train. The performance is especially improved with regards to uncertainty quantification. As a high-dimensional example, we consider the WeatherBench dataset, for which we develop spatial scoring rules. The following figure shows the actual realisation and three possible forecasts for the WeatherBench dataset, for a specific date in the test set. Notice how the different predictions are similar, but some small scale features are different between the different forecasts.



For more results, please refer to Pacchiardi et al. (2021), which can be found at: https://arxiv.org/abs/2112.08217.

List of publications/reports from the project with complete references

Pacchiardi, L., & Dutta, R. (2021). Generalized Bayesian likelihood-free inference using scoring rules estimators. *arXiv preprint arXiv:2104.03889*.

Pacchiardi, Lorenzo, et al. "Probabilistic Forecasting with Conditional Generative Networks via Scoring Rule Minimization." arXiv preprint arXiv:2112.08217 (2021).

Pacchiardi, L., & Dutta, R. (2022). Likelihood-Free Inference with Generative Neural Networks via Scoring Rule Minimization. arXiv preprint arXiv:2205.15784.

Future plans

We are currently working on improving the Scoring-Rules training methodology, specifically for forecasting spatio-temporal sequences with recurrent neural networks. We envision applying this methodology to post-processing and downscaling tasks as well, in collaboration with domain experts from various meteorological centres.