

NEURAL PROCESSES FOR CLIMATE EMULATION

Jack Kai Lim *

Hacıoğlu Data Science Institute
University of California, San Diego
jklim@ucsd.edu

ABSTRACT

Climate modeling has long been limited by the amount of resources needed to use legacy built Earth system models. As such, exploration into the many different and possible emission pathways have been relegated to one-dimensional impulse response or simple pattern scaling models, neither of which are capable of accounting for finer details when emulating the Shared Socioeconomic Pathways. This paper introduces a new benchmark using a (think of a word) model called a Neural Process in the hopes to improve upon the benchmarks that have been established in Watson-Parris et al. (2022). In the hopes that these Deep Learning models are capable of emulating the response of the full complexity of the Earth System Models to forcings using a fraction of the time and resources while also being more complex than simple pattern scaling models. These models would predict annual global mean global distributions of temperature, diurnal temperature range and precipitation given the emissions and concentrations of Carbon Dioxide (CO_2), Methane (CH_4), Sulfur Dioxide (SO_2) and Black Carbon (BC).

1 INTRODUCTION

Climate Scientists have already generated different emission pathways that are used to predict the future climate. These pathways are called the Shared Socioeconomic Pathways (SSPs). However, these models are computationally expensive and require a lot of resources to run, which limits the amount of exploration that can be done. Which is specifically important for policymakers as they need to assess different social and economic impacts under different emission scenarios to reduce the effect of climate change and achieve the goal. Watson-Parris et al. (2022) introduced a benchmark using different machine learning models that emulate different emissions scenarios are able to generate predictions of global temperature, diurnal temperature, and precipitation (including extreme precipitation). This paper looks to build upon it by introducing a new benchmark using a Neural Process model.

For the task at hand, we looked into a few implementations. A base Neural Process (NP) model Gordon et al. (2020) which takes a sparse representation of the data and learns the distribution of the data. A Convolutional Neural Process (CNP) which introduces the use of Convolutional layer in the Neural Process, allowing the model to learn the spatial dependencies of the data and also a SNP which introduces the use of a Spherical Convolutional layer in the Neural Process, allowing the model to learn the spatial dependencies of the data in a spherical manner which fits the idea of the Earth being a sphere.

2 DATA

Following the work of Watson-Parris et al. (2022), this paper will use the same training data from the Norwegian Earth System Model (NorESM2) which is generated data from simulations performed by the NorESM2 model. This was done as part of the sixth coupled model intercomparison project (CMIP6 ; Eyring et al., 2016). The data is used by the policymakers when deciding climate policies so that the emissions data that come from the NorESM2 and CMIP allow the improved

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies. Funding acknowledgements go at the end of the paper.

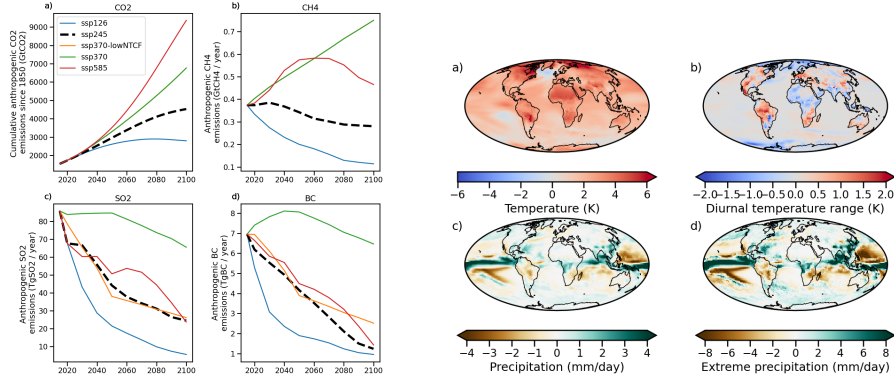


Figure 1: (a) Line Graphs of Input Variables from Multiple SSP (b) Map of Output Variables from SSP245

ClimateBench Plus models to predict outcomes of different possible scenarios that align with the policymakers’ want to reduce climate change. The data that is extracted from NorESM2 and CMIP are netcdfs which are multi-dimensional containing data on every latitude and longitude for the emissions and aerosols that we are looking at i.e Carbon Dioxide, Methane, Sulfur Dioxide, and Black Carbon and span from the 1850s to the end of 2100.

The input variables include emission data of **CO₂**, **SO₂**, **CH₄**, and **Black Carbon** by different emission scenarios data and historical data. We took the global average of input variables CO₂ and CH₄ by year. We convert the units of SO₂ and Black Carbon into Tera gram and take the global sum. Through this process, we acquire the patterns of four different input variables from 2020 to 2100. The output variables are temperature (TAS), Precipitation (PR), 90th percentile precipitation (PR90), and daily diurnal temperature range (DTR).

DATA PREPROCESSING

For the base Neural Process we used a Sparse Representation of the data in order to reduce the data into 1 dimension. Where the context points are a mask of 10% to 30% of the training data from the geographical grid cells. The target points are the rest of the data. The data is then normalized to have a mean of 0 and a standard deviation of 1.

$$x \rightarrow (\text{lat, lon, time, aerosols input}) \quad y \rightarrow (\text{temperature, precipitation, DTR, PR90})$$

For the CNP models we use a multidimensional representation similar to the one in the Covolutional Neural Network baseline model in Watson-Parris et al. (2022). Where the dimensions are:

$$x \rightarrow (\text{time, aerosols input, lat, lon}) \quad y \rightarrow (\text{temperature, precipitation, DTR, PR90})$$

NEURAL PROCESSES

A Neural Process is a class of neural latent variable models which combines the best of both a Gaussian Process (GP) Rasmussen, Williams, et al. (2006) and a Neural Network. The Neural Process (NP) Garnelo et al. (2018) define a distribution over functions like GPs and are capable of rapid adaptation to new observations, and can also estimate the uncertainties. While at the same time are computationally efficient like Neural Networks and are able to adapt their priors to the data, while as the same time also having the flexibility to model complex functions.

A Neural Process is made up of two components, the encoder and the decoder. The encoder takes in the context points and encodes them into a latent representation. The decoder then takes the latent representation and the target points and decodes them into a distribution over the target points.

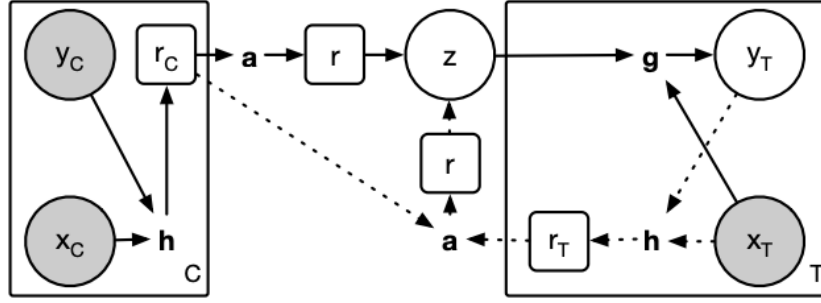


Figure 2: Neural Process Model taken from Garnelo et al. (2018)

With the flexibility of the Neural Process this allowed the construction of another variant of the Neural Process called the Convolutional Neural Process (CNP) Gordon et al. (2020). The CNP is a variant of the NP that uses a Convolutional Neural Network (CNN) to encode the context points. This allows the model to learn the spatial dependencies of the data. The CNP is able to learn the spatial dependencies of the data and is able to predict the target points based on the context points.

3 METHODS

In this paper we used three different variations of the Neural Process Model to predict the global mean temperature, diurnal temperature range, and precipitation. The models are as follows:

BASE NEURAL PROCESS

For a baseline model, we used a Neural Process model which takes in a sparse representation of the data, which reduces the higher dimensional spatial temporal data into a 1 dimensional representation. This simplifies the task removing the spatial dependencies and temporal dependencies of the data and making them independent of each other. This model is used to see how well the model can predict the data without the spatial and temporal dependencies.

The model takes in a sparse representation of the data 2 is passed through a linear Neural Network to encode the data into a latent representation with the context points. The latent representation is then passed through another linear Neural Network to decode the data into a distribution over the target points.

The architecture of the model is as follows:

CONVOLUTIONAL NEURAL PROCESS

The Convolutional Neural Process is a variant of the Neural Process that uses a Convolutional Neural Network to encode the context points. This allows the model to learn the spatial dependencies of the data. The model is able to learn the spatial dependencies of the data and is able to predict the target points based on the context points.

The model takes in a multidimensional representation of the data 2 is passed through a Convolutional Neural Network to encode the data into a latent representation with the context points. The latent representation is then passed through another Convolutional Neural Network to decode the data into a distribution over the target points.

The architecture of the model is as follows:

Details on the Neural Network layers can be found here

SPHERICAL CONVOLUTIONAL NEURAL PROCESS

Finally, we introduce the Spherical Convolutional Neural Process which is a variant of the Convolutional Neural Process that uses a Spherical Convolutional Neural Network to encode the context

Model	Temperature	Precipitation	DTR	PR90
Base NP	x			
CNP	x			
Spherical CNP	x			
GP	x			
CNN	x			

Table 1: Comparison of the Neural Process Models against the Benchmark Models

points. With the hopes to improve upon the Convolutional Neural Process by taking into account the spherical nature of the Earth when learning the spatial dependencies.

The model takes the same data as the Convolutional Neural Process and the architecture of the model is as follows:

4 EVALUATION AGAINST BENCHMARK MODELS

Following Watson-Parris et al. (2022) we will also be using the $NRMSE_t$ from the paper which is defined as follows,

$$NRMSE_s = \sqrt{\langle (|x_{i,j,t}|_t - |y_{i,j,t,n}|_{n,t})^2 \rangle / \langle |y_{i,j}|_{t,n} \rangle} \quad (1)$$

$$NRMSE_g = \sqrt{\langle (|x_{i,j,t}|_t - |y_{i,j,t,n}|_{n,t})^2 \rangle / \langle |y_{i,j}|_{t,n} \rangle} \quad (2)$$

$$NRMSE_t = NRMSE_s + \alpha \times NRMSE_g \quad (3)$$

Where $NRMSE_s$ is the global mean root-mean squared error, and $NRMSE_g$ is $NRMSE$ in the global mean. The equation also includes a weighing function to take the decreasing grid-cell area towards the north and south poles which is defined as follows,

$$\langle x_{i,j} \rangle = \frac{1}{N_{lat}N_{lon}} \sum_i^{N_{lat}} \sum_j^{N_{lon}} \cos(lat(i))x_{i,j} \quad (4)$$

and the co-efficient α is chosen to be 5 from the ClimateBench paper, in order to provide equal weightage between the measures.

The reason, we are using this evaluation metric is to be able to compare directly the performances of our models against the models from the paper. And using the same evaluation metric will give us the most 1-1 comparison.

5 RESULTS

Using the evaluation metric from 4 we are able to compare the performance of the models against the benchmark models from the ClimateBench paper. The results are as follows:

6 DISCUSSION

neural processes benefit are does not require a lot of data, can be used for small datasets, can be used for high dimensional data, can be used for sequential data. Also good for few shot learning, works due to the context points.

7 CONCLUSION

REFERENCES

Marta Garnelo, Jonathan Schwarz, Dan Rosenbaum, Fabio Viola, Danilo J. Rezende, S. M. Ali Eslami, and Yee Whye Teh. Neural processes, 2018.

Table 2: Sample table title

PART	DESCRIPTION
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

Jonathan Gordon, Wessel P. Bruinsma, Andrew Y. K. Foong, James Requeima, Yann Dubois, and Richard E. Turner. Convolutional conditional neural processes, 2020.

Carl Edward Rasmussen, Christopher KI Williams, et al. *Gaussian processes for machine learning*, volume 1. Springer, 2006.

Duncan Watson-Parris, Yuhao Rao, Dirk Oliv  ,   yvind Seland, Peer Nowack, Gustau Camps-Valls, Philip Stier, Shahine Bouabid, Maura Dewey, Emilie Fons, et al. Climatebench v1. 0: A benchmark for data-driven climate projections. *Journal of Advances in Modeling Earth Systems*, 14(10):e2021MS002954, 2022.

A MODEL ARCHITECTURES

A.1 CONVOLUTIONAL NEURAL PROCESS

A.2 SPHERICAL CONVOLUTIONAL NEURAL PROCESS