Forecasting El Niño-Southern Oscillation (ENSO) with Graph Neural Networks

Ethan Zhang
Dept. of Computer Science
Stanford University
ethanz@stanford.edu

Saahil Sundaresan
Dept. of Computer Science
Stanford University
saahilss@stanford.edu

Yuchen Li
Dept. of Physics
Stanford University
yucli@stanford.edu

1 Introduction

The El Niño-Southern Oscillation (ENSO) is the strongest and arguably most important mode of natural climate variability on Earth. It is primarily associated with the large-scale oscillation of sea surface temperature anomalies in the eastern Tropical Pacific. The warm phase (called El Niño) and cool phase (called La Niña) have been known for centuries to cause shifts in global precipitation and winds, thereby affecting agriculture, fisheries, and water security (McPhaden et al., 2006). Moreover, developing a theory of ENSO has been a central and still ongoing pursuit in the field of climate science (Timmermann et al., 2018). Developing models to forecast ENSO is therefore of both great societal and scientific importance.

Broadly, ENSO forecasting models can be grouped into two categories: dynamical models and statistical models. Dynamical models range in complexity from very simple models that capture the most essential physics to ensembles of global general circulation models (GCMs) that simulate the evolution of the ocean and atmosphere from perturbed initial conditions by integrating fluid flow and thermodynamics equations. Statistical models, which include machine learning approaches, learn spatiotemporal relationships between ocean and atmospheric variables derived either from reanalysis datasets (observationally-constrained state estimates of the Earth) or GCM simulations.

ENSO, like most other atmospheric and oceanic phenomena, has a certain window of predictability before stochastic noise arising from initial condition sensitivity (i.e., chaos) dominates the predictable signal. As such, most models remain skillful (typically defined as $r > 0.5$ where r is the Pearson correlation coefficient between predicted and actual ENSO index timeseries) for lead times of up to a year. Some recent attempts using convolutional neural networks have in fact pushed the maximum skillful lead time up to 22 months (Ham et al., 2019; Wang et al., 2023).

Here, we propose a graph neural network (GNN) approach to forecasting the Oceanic Niño Index (ONI), the primary indicator used to monitor ENSO, from climate data discretized by geolocation and time step. The ONI represents anomalies in sea surface temperatures (SST) from long-term averages in the Niño 3.4 region, which spans (5ºN-5ºS, 120º-170ºW) in the equatorial Pacific Ocean, and is calculated by computing running three-month averages in SST anomalies in the region. An ONI value of greater than 0.5 over five consecutive overlapping three-month periods indicates a warm (El Niño) period, while a value less than -0.5 indicates a cool (La Niña) period; the magnitude of the ONI score is indicative of stronger El Niño/La Niña conditions. The aim of this project is to improve both the accuracy of ONI predictions as well as the skillful lead time measure over existing approaches.

2 Related Works

In recent years, graph neural network-based models have begun achieving considerable success in weather and climate forecasting tasks. One notable example is GraphCast, a medium-range weather prediction model that outperforms traditional dynamical forecasting methods (Lam et al., 2023).

To our knowledge, there only exists one application of GNNs for forecasting ENSO (Cachay et al., 2021), who propose a model to forecast ONI using a grid partition of the world map as its graph nodes and a vector of set climate features across time steps as the node features. The advantage of using a GNN over, say, a CNN (e.g., Ham et al., 2019) to forecast ENSO is that a graph can explicitly model both local and nonlocal dependencies, which are important for climate prediction. Additionally, a graph can directly account for the irregular geometry of the ocean domain (a CNN-based approach would need to mask the land as zeros).

While Cachay et al. (2021) achieve promising results, outperforming Ham et al. (2019) and most dynamical models for sub-6 month lead times, we believe there are improvements that can be made. Critically, their work does not explicitly encode any of the input data's temporal dependency – rather than representing historical data as a set of graphs, all of the training data is instead merged into a single graph representation, with the node features at each time step simply concatenated into a single large vector. Concretely, each node's D-dimensional feature vector $V_i \in \mathbb{R}^D$ is concatenated across all N graph nodes, giving snapshot measurements of the graph node features at time t as $X_t = \{V_1^{(t)}, ..., V_N^{(t)}\} \in \mathbb{R}^{N \times D}$. The input data for predicting the ONI index from w timesteps of historical climate information is simply formatted as a stacked matrix of the X_t's, representation matrix $X \in \mathbb{R}^{N \times wD}$, which clearly lacks any temporal modeling. We detail how we improve upon this weakness, as well as other gaps in the work, in the Results section below.

3 Data

We follow Ham et al. (2019) and Cachay et al. (2021) and use the Global Ocean Data Assimilation System (GODAS) reanalysis (1984–2017) for our model testing. The reanalysis represents the "ground truth" ocean state for the past four decades that is used as input and validation for our model.

Because the observational record of Earth's climate ($\mathcal{O}(100 \text{ years})$) is short compared to the timescale of prediction ($\mathcal{O}(1 \text{ year})$), it is necessary for machine learning approaches to use simulation data derived from climate models to supplant available observational data for training. This can be done either through transfer learning (i.e., pretraining on simulations, then fine-tuning to observations) or simply by naïvely training on simulations.

Our main source for climate model runs is the Coupled Model Intercomparison Project (CMIP; Eyring et al., 2016), a coordinated ensemble of standardized climate model experiments used extensively by the climate science community (for example, it informs the IPCC reports). CMIP features a suite of different dynamical climate models, each of which have ensembles of simulations forced with historical boundary conditions (such as atmospheric CO_2, aerosols, solar radiation, etc.). It is important to note that these simulations, unlike reanalysis, are not explicitly nudged to fit observations, and therefore often suffer from model-specific biases resulting from missing/incorrect physics, discretization, etc. We discuss the implications of this in Section 5.

The final source of data is the Simple Ocean Data Assimilation (SODA) reanalysis dataset, which is an estimation of the true ocean state from 1871 to 1973. This will be used to fine-tune the model after it has been pretrained on CMIP data.

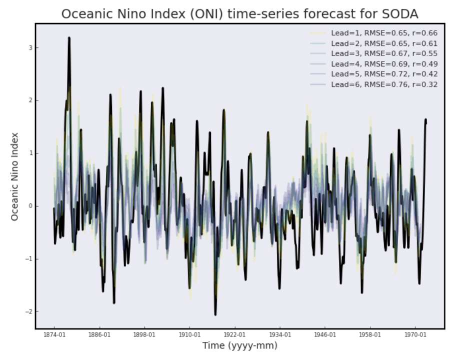
4 Methods

As an initial step, we train a model consisting of a graph convolutional network (GCN) stacked on top of an encoder-decoder recurrent neural network (RNN), hereafter termed GNNRNN. In this preliminary test, we naively train our model entirely on CMIP (simulation) data and test on the SODA dataset.

We aimed to capture the temporal nature of historical climate data and simulation data. As such, we decided to utilize an autoregressive model architecture. GNNRNN's GCN layer embeds $36b$ graphs (where b is the batch size, and 36 months is the input length we chose) into $36b$ vectors $\in \mathbb{R}^g$ (where g is a hyperparameter we set). From there, we feed the batch of 36-month sequences into the encoder portion of the RNN, which further projects the graph embeddings into $\mathbb{R}^e$ (where e is the hidden dimension of the encoder). Lastly, we initialize the decoder with the final hidden state of the encoder ($b \times 2e$ - the factor of 2 stems from the encoder being bidirectional) and autoregressively generate predicted ONI values.

5 Results

We train the above-described GNNRNN for 20 epochs with a learning rate of 0.0005 and batch size b of 32. We chose both g and e (the graph embedding dimension and RNN encoder hidden dimension, respectively) to be 256. We construct our GCN to have seven convolutional layers. A dropout rate of 0.3 was used for the GCN portion. The model was trained using MSE loss.



Our model achieves skillful predictions ($r > 0.5$) for 3 months. Our model's comparatively short skillful prediction window is no doubt in large part due to two main reasons. Firstly, we use of a naive connectivity scheme, where only adjacent nodes possess edges between them (in other words, any given node can have at maximum 4 neighbors - north, south, east, and west). This means that currently, far-away nodes have no capacity to message-pass unless we drastically increase the number of convolutional layers in the GCN module. Secondly, we have yet to attempt to fine-tune our model on a reanalysis dataset, which represents the ground truth far better than the CMIP simulation data we trained on. As such, the current next immediate steps in our work are to:

- Adapt our connectivity scheme to include long-range connections by either integrating domain knowledge of long-range climate teleconnections, or to use meshes of various degrees of resolution, such as is the case in (Lam et al., 2023), in order to facilitate long-range node message passing
- Implement fine-tuning functionality to allow us to align our model to SODA

References

Salva Rühling Cachay, Emma Erickson, Arthur Fender C. Bucker, Ernest Pokropek, Willa Potosnak, Suyash Bire, Salomey Osei, and Björn Lütjens. 2021. The world as a graph: Improving el niño forecasts with graph neural networks.

Veronika Eyring, Sandrine Bony, Gerald A. Meehl, Catherine A. Senior, Bjorn Stevens, Ronald J. Stouffer, and Karl E. Taylor. 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958.

Yoo-Geun Ham, Jeong-Hwan Kim, and Jing-Jia Luo. 2019. Deep learning for multi-year ENSO forecasts. *Nature*, 573(7775):568–572. Publisher: Nature Publishing Group.

Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, Alexander Merose, Stephan Hoyer, George Holland, Oriol Vinyals, Jacklynn Stott, Alexander Pritzel, Shakir Mohamed, and Peter Battaglia. 2023. Graphcast: Learning skillful medium-range global weather forecasting.

Michael J. McPhaden, Stephen E. Zebiak, and Michael H. Glantz. 2006. ENSO as an Integrating Concept in Earth Science. *Science*, 314(5806):1740–1745. Publisher: American Association for the Advancement of Science.

Axel Timmermann, Soon-Il An, Jong-Seong Kug, Fei-Fei Jin, Wenju Cai, Antonietta Capotondi, Kim M. Cobb, Matthieu Lengaigne, Michael J. McPhaden, Malte F. Stuecker, Karl Stein, Andrew T. Wittenberg, Kyung-Sook Yun, Tobias Bayr, Han-Ching Chen, Yoshimitsu Chikamoto, Boris Dewitte, Dietmar Dommenget, Pamela Grothe, Eric Guilyardi, Yoo-Geun Ham, Michiya Hayashi, Sarah Ineson, Daehyun Kang, Sunyong Kim, WonMoo Kim, June-Yi Lee, Tim Li, Jing-Jia Luo, Shayne McGregor, Yann Planton, Scott Power, Harun Rashid, Hong-Li Ren, Agus Santoso, Ken Takahashi, Alexander Todd, Guomin Wang, Guojian Wang, Ruihuang Xie, Woo-Hyun Yang, Sang-Wook Yeh, Jinho Yoon, Elke Zeller, and Xuebin Zhang. 2018. El Niño–Southern Oscillation complexity. *Nature*, 559(7715):535–545. Publisher: Nature Publishing Group.

Haoyu Wang, Shineng Hu, and Xiaofeng Li. 2023. An Interpretable Deep Learning ENSO Forecasting Model. *Ocean-Land-Atmosphere Research*, 2:0012. Publisher: American Association for the Advancement of Science.