Weekly Climate Indices: Generation and Prediction

Dimensionality Reduction of Surface Temperature & Prediction of North Atlantic Oscillation

JK Hunt, James Schull, Javier Echevarría Cuesta | {jkhunt, jschull, javierec}@stanford.edu

Problem Outline

Climates and local climatic phenomena are enormously complex and intricate systems, often approximated by indices. We investigated methods for the generation and prediction of *weekly* climate indices, using 70 years worth of weekly-averaged global surface air temperature and pressure data to train and evaluate our models.

Our project was comprised of two parts: firstly, the generation of weekly surface skin temperature (SKT) indices, and secondly, the prediction of weekly North Atlantic Oscillation indices.

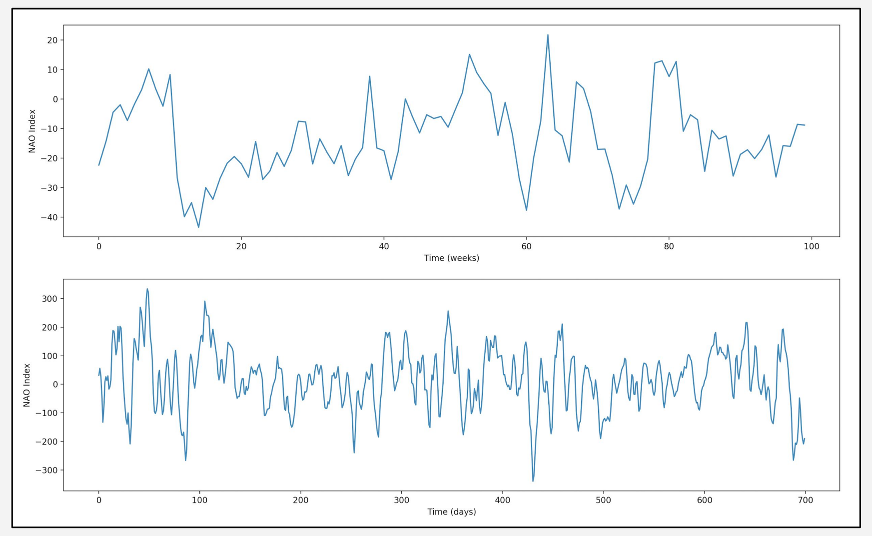
The generation of SKT indices allowed us to identify and visualize regions causally related to near-future temperatures in Peru and Northern California. Modelling with an LSTM and ARIMA model allowed us to reasonably predict weekly NAO, despite properties of weekly-averaged NAO indices that make prediction unusually difficult.

Data

	Part 1: Index Generation (SKT)	Part 2: Index Prediction (NAO)
Dataset	NCEP/NCAR Reanalysis Surface Skin Temperature 1948-2009 Weekly Averaged	NCEP/NCAR Reanalysis 250 hPA 1948-2009 Weekly Averaged
Features	(3731 x 192 x 96) time series matrix representing grid points across Earth of surface temperature	(3731 x 192 x 96) time series matrix representing grid points across Earth of atmospheric pressure
Labels Dataset	Weekly averaged surface temperature for Peru and California	Weekly averaged NAO index
Labels Shape	(3731 x 1) time series vector	(3730 x 1) time series vector

Part 2: Predicting NAO Indices

In the second part of our project, we aimed to predict the NAO on a weekly averaged timescale (shown to the right), and in particular wanted to incorporate time series data in order to identify patterns in oscillation. However, it is apparent that the weekly averaged index does not exhibit the same clear oscillating structure that we can see in daily index, which influenced our decision to use global time series pressure data as an additional feature.



NAO index plot against time (top = weeks, bottom = days)

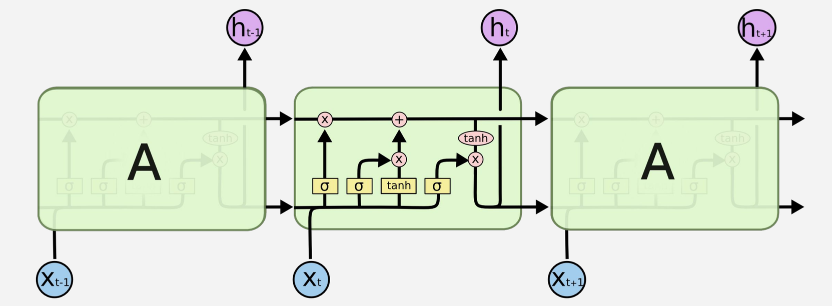
We trained and evaluated two models in particular: ARIMA (AutoRegressive Integrated Moving Average), and an LSTM recurrent neural network.

ARIMA

$$y_t = k + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \epsilon_t + m_1\epsilon_{t-1} + m_2\epsilon_{t-2} + \ldots + m_q\epsilon_{t-q}$$

p = 4 and *q = 7* are hyperparameters that represent autoregressive and moving average order respectively

LSTM



Hyperparameters:
- 50 unit hidden layer
- tanh activation
- 30 epochs
- 10 weeks of lag
- PCA reduction of pressure data to 10 components

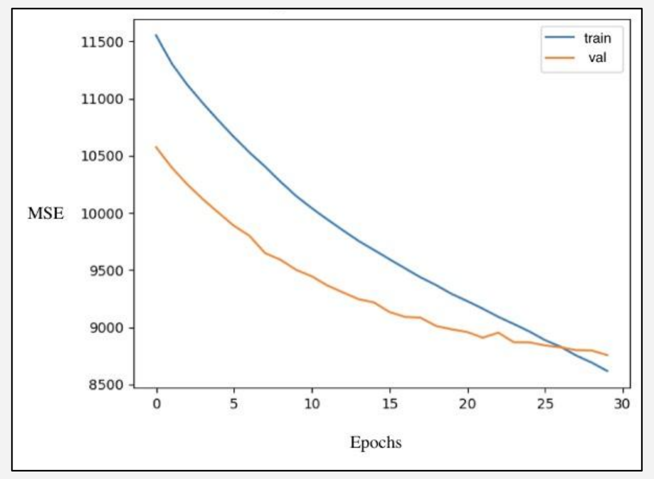
When training, we experimented with different methods of incorporating global pressure data as a feature in our models. This included using: the full non pre-processed grid, the top *k* components from PCA reduction, and climate indices as generated by the technique described in part 1 (using pressure data instead of temperature data).

Results & Discussion

Model	RMSE
Baseline (persistence)	106.6
LSTM	93.57
ARIMA	90.12

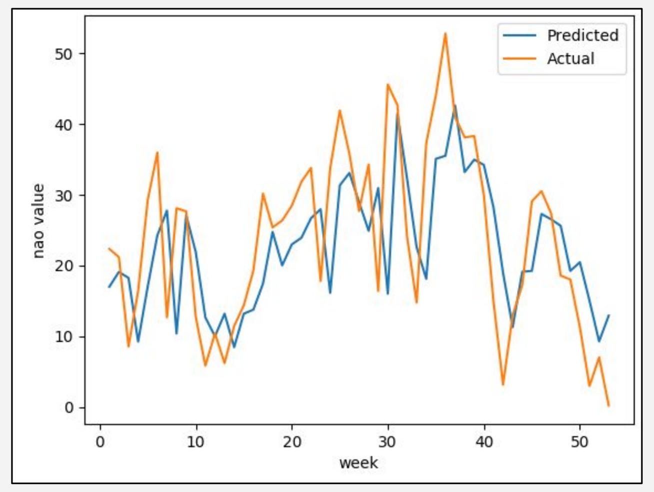
We performed grid searches over both the ARIMA and LSTM models in order to optimize hyperparameters.

For the LSTM, we found that 50 hidden units and a tanh activation function, trained for 30 epochs with 10 weeks of lag and a pressure dataset reduced to 10 dimensions via PCA gave optimal performance. For ARIMA model, an AR order of 4 weeks and an MA order of 7 weeks were optimal.



LSTM Train/Validation Loss per Epoch

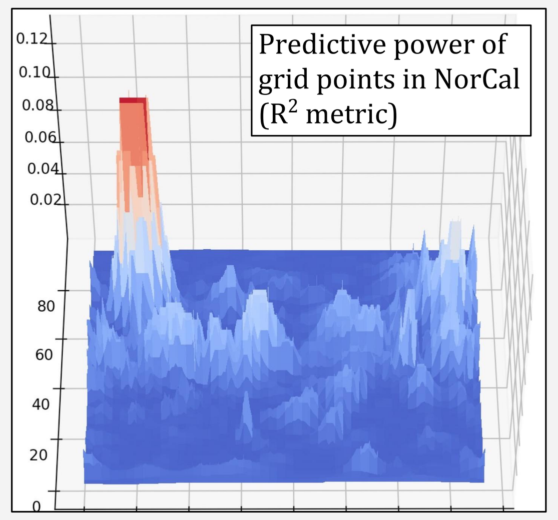
While our models outperformed the baseline by a reasonable amount, it was difficult to improve performance much further. We believe that this is a strong reflection of the difficulty of predicting the weekly average of a climate phenomenon that oscillates on a non-weekly basis. The weekly-averaged NAO time series appeared far noisier than the daily series, and consequently, train performance may well not correlate strongly with test set performance. While weekly NAO indices can be modelled, they are difficult to model well, and this is a property that we would expect to extend to any weekly-averaged index of a climatic oscillation with a non-weekly period.



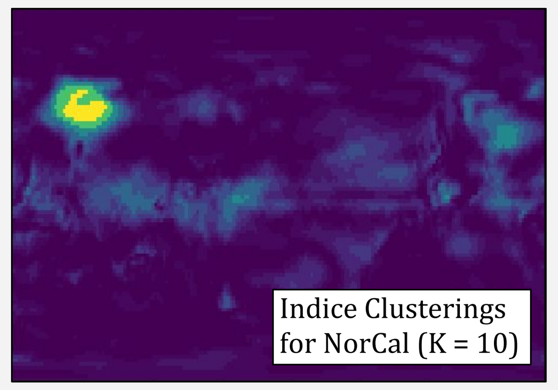
1 Week Lead Forecasting using ARIMA

Part 1: Generation of Surface Skin Temperature Indices

In order to generate the climate indices used to predict surface air temperature at a given location, the first step was to identify the predictive power of every grid point with respect to the location. To do this, we performed linear regression with each individual grid point across the entire time series to predict the surface air temperature in Northern California, and set test R^2 as the determinant of predictive power. In the second step, we ran K-means to cluster the grid points by a custom distance metric that incorporated geographical distance and R2 score, and found that this served as an effective dimensionality reduction technique for temperature prediction, identifying causal regions.



Predictive power of grid points in NorCal (R^2 metric)

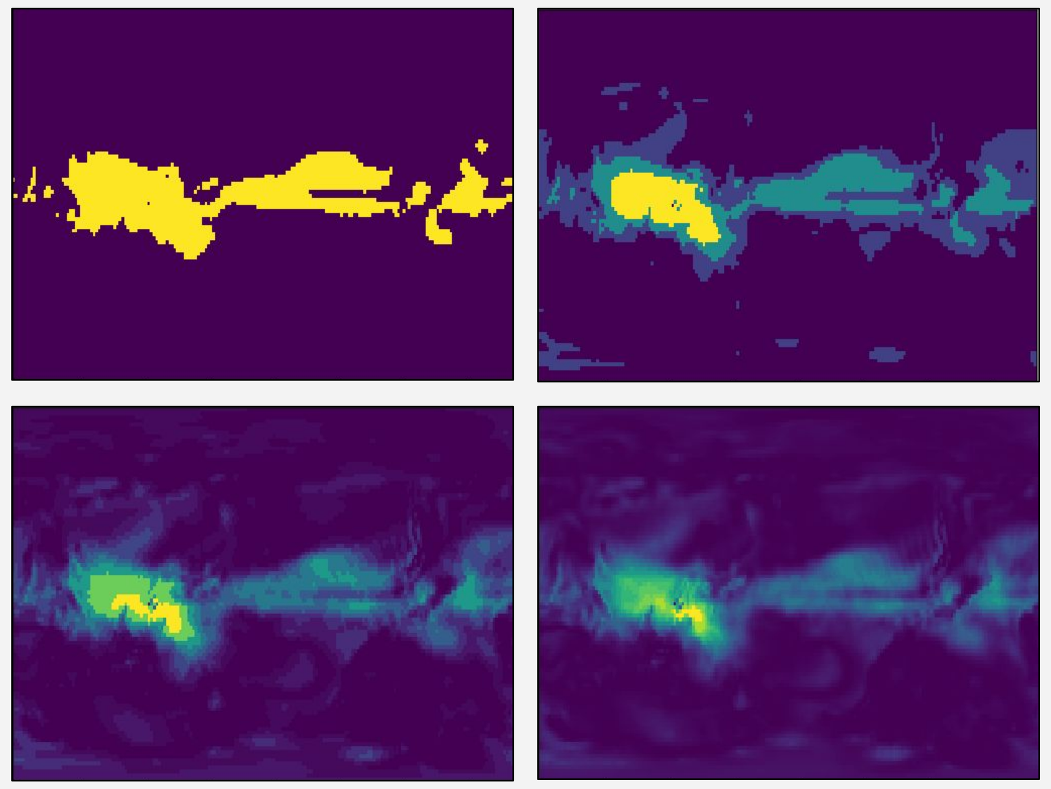


Indice Clusterings for NorCal (K = 10)

We repeated our indices generation for Peru, and the maps on the right display our indice results for different centroid cluster numbers (K = 2, 4, 10, 100).

We see that a small number of clusters well represent the predictability of the grid, and do not lose much information.

Ultimately, to evaluate the performance of these climate indices, we ran linear regression using each the average of all grid point temperatures per centroid as features (these constitute our indices), and output a temperature prediction in NorCal/Peru (respective R^2 of 0.25/0.51).



Indice Clusterings for Peru (K = 2, 4, 10, 100) (left to right, top to bottom)

Future Work

Given more time, we would like to experiment with different approaches to week-forward predictions, using daily as well as weekly input data, and integrating other climate features alongside pressure. We would also like to further explore the influence of different dimensionality reduction methods on our models' forecasting - given that PCA reduction and climate indices both improved performance, we believe there are further refinements that could be made here.

On a broader scale, we believe that accurate weekly climate forecasts would be very useful for further work in many different supervised learning contexts, and we would also like to explore that area.

References

[1] Maximilian Evans, Jasdeep Singh (Dec. 2017), *Unsupervised Machine Learning for Long Range Climate Prediction*, Stanford University.
[2] Michael S Steinbach et al. (Dec. 2003), *Discovery of climate indices using clustering*, University of Minnesota.
[3] Shijin Yuan et al. (May 2019), *Prediction of North Atlantic Oscillation Index with Convolutional LSTM Based on Ensemble Empirical Mode Decomposition*, School of Software Engineering, Tongji University, Shanghai.
[4] A.A. Scaife et al. (2014), *Skilful Long Range Prediction of European and North American Winters*, Geophys. Res. Lett., 41, 2514–2519