Predicting the Precipitation Error in the Southwestern U.S. Between the Green's Function and a Coupled Climate Model

Marc J. Alessi

Homework 1, Due: Friday, November 10

Collaborators: Nicole, Elisa, Yiyu(, Charlie (huge thank you to Charlie))

1. Motivation and Problem Statement

Atmosphere-ocean general circulation models (AOGCMs) are complex models that predict the Earth's climate given a set of governing equations and conservation laws of physics. The earliest versions of these models, with only the atmosphere, were first developed in the 1960's to study how the climate would respond to an applied forcing (e.g., anthropogenic greenhouse gas emissions). For these simple models, it was easy to attribute a change in the climate (e.g., precipitation) to a change in forcing (e.g., prescribed sea surface temperature). Today, with the addition of ocean models, land models, ice models, carbon models, and many complex parameterizations, attribution in AOGCMs has become nearly impossible. For example, a change in precipitation in the Southwestern United States (SWUS) could be due to a change in local soil moisture, a change in local evapotranspiration, a change in the sea surface temperature (SST) off the California coast, or even a change in SST in the West Pacific warm pool. We use a Green's function (GF) approach to attribute changes in SWUS precipitation to changes in local or remote SST.

A GF linearizes the climate system between a forcing and a climate response. In our case, the forcing is a change in SST, and the climate response is precipitation. The GF not only demonstrates what regions of SST change SWUS precipitation is most sensitive to (Figure 1a), but it can also calculate the SWUS precipitation change given an SST pattern (Figure 1b). In this way, the GF is an interpretable simple climate model, since we can both calculate precipitation change and attribute that change to specific SST regions. However, the GF is still a linearization of the climate system, so it misses important nonlinearities. For example, an increase in SST in the West Pacific can lead to greater absolute changes in the midlatitude Rossby wave train than a decrease in SST in the West Pacific (this is due to the saturation of moist static energy in the upper Tropical troposphere, but I will not discuss this here). In other words, changes in SWUS precipitation do not necessarily scale with changes in SST. Because of this, we attempt here to develop a convolutional neural network (CNN) that predicts the precipitation error between the GF prediction and the AOGCM prediction given an SST pattern. The error between the GF and AOGCM prediction could explain important nonlinearities in the climate system. We could then use explainable AI (in the future) to identify what regions of SST are most helpful in predicting this error, i.e., identifying which SST regions are most important for nonlinearities in the climate system.

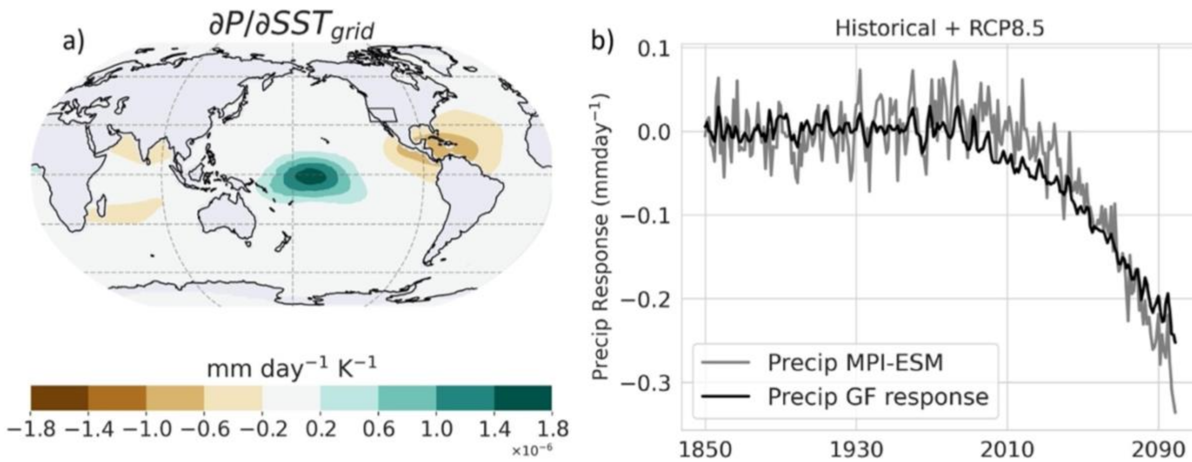

Figure 1. a) The annual-mean SWUS precipitation response per unit SST warming in each grid box. Warming in green (brown) areas leads to a wetting (drying) of the SWUS (region outlined by box). b) The SWUS precipitation response of the Green's function convolved with the MPI-ESM Grand Ensemble SST pattern (black). The SWUS precipitation output from the MPI-ESM Grand Ensemble is also shown (gray). The annual-mean SWUS v-wind response at 200mb (left) and the annual-mean u-wind response at 250mb (right) per unit SST warming in each grid box.

2. Method

To train, validate, and test our CNN, we use data from the Max Planck Institute Earth System Model Grand Ensemble (MPI-GE). The MPI-GE has 100 ensemble members, each running from 1850-2100 under the RCP8.5 (high emissions) scenario. Here, we only use 25 of the ensemble members (to save time in training the model).

2.1. Data description

The SST of each year and ensemble member is first convolved with the SWUS precipitation GF (Figure 1a) to get the annual-mean SWUS precipitation (in units of mm/day). The GF is created through perturbing an atmospheric GCM with many anomalous SST patches. The GF is then created by calculating all nonlocal and local responses to all nonlocal and local SST changes. We then subtract this SWUS precipitation value from what the AOGCM predicts for the same year and ensemble member, giving us our predictand (difference between AOGCM and GF prediction). The predictor is the SST pattern for each year and ensemble member. This is fed into the CNN to predict the error (predictand).

2.2. Pre-processing and data preparation

We first detrended the MPI-GE SST and precipitation data and replaced all nans (land) with zeros. We also normalized the predictor and predictand data before training the CNN. We split the data by randomly selecting 30 ensemble members for training (15 ensemble members * 250 years = 7,500 samples), 10 ensemble members for validation, and 10 ensemble members for testing.

2.3. Model setup

We use a convolutional neural network consisting of a convolutional layer followed by a max pooling layer. This is then repeated. We then flatten the data and use two dense layers. The output layer is a single value (regression task) that predicts the error between the GF and AOGCM. This process is shown in Figure 2. Originally, we had an extra convolutional layer and repeated the convolutional/max pooling layers three times instead of two, but this unfortunately slowed the training of the model significantly.

The CNN has a kernel size of 5 and padding is set to ‘same,’ meaning that there are zeros added to the edges of the SST maps in order to not shorten the map dimensions. Our dense layers are made up of 16 and 8 nodes, and the activation function for all layers was set to ‘relu.’ Originally, I received a list of settings from Libby which included a learning rate of 0.000005, but we found this to be too small of a value. A value of 0.001 trained the model much faster and did not seem to lead to overfitting (in our original tests). We set the number of epochs to 30 through trial and error (20 was too small, but 40 was too much). Our loss function is mean square error, and the batch size is 32.

Note that we solved an error in the code a few hours before the project was due, so we did not have enough time to tune significantly. Using only 15 ensemble members of data requires 1.5 hours of training.

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_1 (InputLayer)        [(None, 96, 192, 1)]      0

 layer (Layer)               (None, 96, 192, 1)        0

 conv2d (Conv2D)             (None, 48, 96, 1)         1025

 max_pooling2d (MaxPooling2D  (None, 24, 48, 1)        0
 )

 conv2d_1 (Conv2D)           (None, 12, 24, 1)         1025

 max_pooling2d_1 (MaxPooling  (None, 6, 12, 1)         0
 2D)

 flatten (Flatten)           (None, 72)                0

 dense (Dense)               (None, 16)                1168

 dense_1 (Dense)             (None, 8)                 136

 dense_2 (Dense)             (None, 1)                 9

=================================================================
Total params: 3,363
Trainable params: 3,363
Non-trainable params: 0
```

Figure 2. Convolutional neural network setup.

3. Results

The loss for both training and validation is plotted in Figure 3. The CNN appears to be overfitting the training data quite significantly. This is shown by the training loss decreasing more than the validation loss, while also being significantly less than the validation loss. Furthermore, it appears 30 epochs were not enough to train this model, given that the training loss line is not stabilizing around a loss value. We did experiment with multiple epochs, pooling sizes, number of convolutional layers, and different learning rates, but more testing of hyperparameters should be carried out.

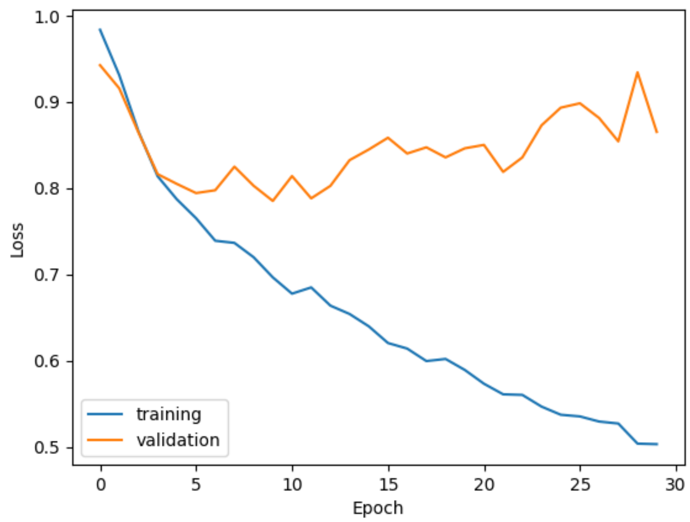

Figure 3. Training and validation loss (mean square error) for each epoch.

We also plot the truth (AOGCM – GF) vs. what the CNN model predicts for the testing data (Figure 4). There is some correlation (value not calculated) between what the CNN predicts for precipitation error and what the truth is, demonstrated by values generally following a 1:1 line.

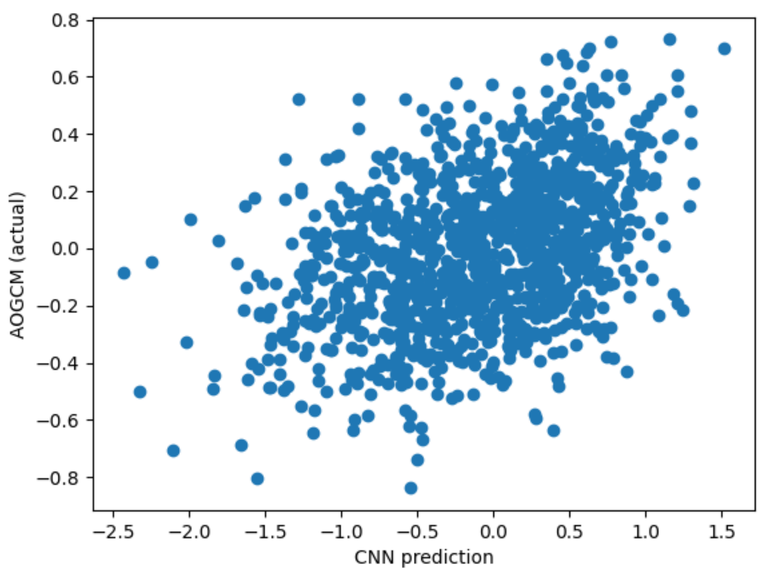

Figure 4. AOGCM – GF SWUS precipitation response (truth) compared to what the CNN predicts (model). Some correlation exists, though the model can be improved significantly.

4. Discussion

Unfortunately, despite much testing, our model continued to overfit (Figure 3), meaning that there is not much skill in the CNN predicting the error between the GF and AOGCM SWUS precipitation. Over the next few weeks, we will continue to work on this issue by adding more data (using all the ensemble members), developing a less complex model (removing one convolutional layer), and/or implementing a regularization technique.

We think the model performs better than a climatological baseline in error, i.e., the difference between what the AOGCM predicts and precipitation climatology. This is because we see variation in SWUS precipitation from the CNN, and often this can predict what the actual error is between the AOGCM and GF (Figure 4).

It is possible that we will not be able to better predict the error between AOGCM SWUS precipitation and GF SWUS precipitation. This is because the error between the two represents the atmospheric internal variability, which we are unable to predict from the SST pattern. Put another way, the GF predicts a precipitation response in the SWUS given an SST pattern. However, AOGCMs predict a precipitation response to the SST pattern *and* internal atmospheric variability, and this atmospheric variability is unrelated to the pattern of SST. This is the main reason why a difference exists between GF and AOGCM SWUS precipitation predictions. On the other hand, this error could also provide clues to the percent of precipitation that affects the SUWS that is influenced directly by the pattern of SST (what the GF predicts) and influenced by the atmosphere (the difference between the GF and the AOGCM). If the CNN model perfectly predicted the error in SUWS precipitation given the SST pattern, then this would indicate that atmospheric variability has very little influence over the SWUS.

5. Conclusion

We find that a convolutional neural network can predict some of the error between SWUS precipitation response according to the GF and according to the AOGCM (Figure 4). The lack of skill is most likely due to the existence of internal atmospheric variability; it is unlikely the CNN predicts the precipitation response of the AOGCM given only the SST pattern, since atmospheric variability exists in addition to oceanic variability (i.e., not all atmospheric variability is driven by SST). However, the fact that the CNN can predict some of the error indicates that the model has learned some nonlinearities of the climate system. This could prove helpful using explainable AI in attributing nonlinear precipitation responses to SST changes.

Github repository: https://github.com/mja244/ats780/tree/main