Structure of the non writing aspect of the project:

1. Determine which data we want to use (features, spatial extent, temporal extent, augmented data)
2. Perform Proper Orthogonal Decomposition (POD) to reduce the dimensionality of the system. POD looks to convert a high dimensional system into a lower dimensional system by finding a few key “modes” which account for most of the variance in the high dimensional data. A temporally changing vector field f(x,t) can be approximated as f(x,t) ~ a\_1(t)\*phi\_1(x) + a\_2(t)\*phi\_2(x) + … If f(x,t) corresponds to timeseries data of 100 grid cells we may be able to describe most of the change in the system with a\_1(t), and a\_2(t), reducing our 100 variable to just 2. More information on POD at [Proper orthogonal decomposition - Wikipedia](https://en.wikipedia.org/wiki/Proper_orthogonal_decomposition)
3. Use various ML techniques to predict future timeseries data. We will use an RNN, a neural ODE, which acts as an ODE with the derivative of the state vector given by a neural network. We will also use a neural DDE (delay differential equation). The neural DDE can be framed as a neural ODE by changing the state vector of the system to have data from previous time steps. If x\_j is the state vector at iteration j of the neural ODE then the state vector of the neural DDE may be given by [(x\_i)^T, (x\_{i-1})^T, …]

What is ENSO?

ENSO stands for El-Nino Southern Oscillation. It is a nonperiodic oscillatory pattern in sea surface temperature trends in the equatorial Pacific Ocean. It affects the climate of the entire world, generating changes in temperature and precipitation, even in Canda depending on what phase of the ENSO cycle we are in.

What have we done so far for the project?

So far, we have the datasets, we have sea surface temperature, atmospheric temperature, precipitation, and wind speed from 1900 to 2014 from an ensemble of climate reanalysis models.

We have code to crop this data spatially and temporally.

We have code to reduce the dimensionality of the data using POD

We have code to train a neural ODE to predict ENSO trends

We have code to train a neural DDE to predict ENSO trends.

What we need to do

We need feature analysis/augmentation. We need to figure out which features are most important for predicting ENSO. This may include different spatial crops, rolling means, creating smoother data through interpolating between a sparser dataset, different numbers of POD modes, creating new features such as taking the derivatives of a\_i. I assume this will be most of the work we have left on the non writing aspect of the project.

We need to make an RNN

We need to adjust hyperparameters/optimize the neural ODE and neural DDE to get the best performance out of our models