The objective of this project was to predict future values of El Niño characteristics using recursive equation models, data trees, and an evolutionary algorithm. In addition, the project attempted to find how number of generations and size of the tree influenced accuracy of the predictions.

In order to generate a final model and test its accuracy, a Python program was written and run. First, the program generated a number of models using a data tree. The program then tested all the potential models against training data. Poor models were removed, new models were generated as replacements, and random changes were carried out. This process of testing and changing the set of models, called a generation, was repeated various times. Finally, the best model throughout the entire process was outputted, and its accuracy was measured by comparing its predictions to testing data.

3 different characteristic values of El Niño were tested. These values measured air pressure, sea surface temperature, and precipitation levels each in different areas. For each value, the program was used to generate and test models after 10, 50, and 100 generations. Models created with different tree sizes were also generated.

Ultimately, this program had success in predicting future values of 2 El Niño characteristics, but also did not succeed at predicting the temperature characteristic. In addition, number of generations had little effect on accuracy, while increased tree size led to increased accuracy.

The successes of the model did satisfy the objective and also indicate potential for improvements in future testing. Furthermore, the program’s generality indicates possible success both in predicting other El Niño characteristics and in predicting any other time series as well.

The objective of this project was to predict future values of El Niño characteristics using recursive equation models, data trees, and an evolutionary algorithm. In addition, the project attempted to find how number of generations and size of the tree influenced accuracy of the predictions.

In order to generate a final model and test its accuracy, a Python program was written and run. First, the program generated a number of models using a data tree. The program then tested all the potential models against training data. Poor models were removed, new models were generated as replacements, and random changes were carried out. This process of testing and changing the set of models, called a generation, was repeated various times. Finally, the best model throughout the entire process was outputted, and its accuracy was measured by comparing its predictions to testing data.

3 different characteristic values of El Niño were tested. These values measured air pressure, sea surface temperature, and precipitation levels each in different areas. For each value, the program was used to generate and test models after 10, 50, and 100 generations. Models created with different tree sizes were also generated.

Ultimately, this program had success in predicting future values of 2 El Niño characteristics, but also did not succeed at predicting the temperature characteristic. In addition, number of generations had little effect on accuracy, while increased tree size led to increased accuracy. The successes of the model did satisfy the objective and also indicate potential for improvements in future testing. Furthermore, the program’s generality indicates possible success both in predicting other El Niño characteristics and in predicting any other time series as well.