Christian Alameda
Noah Betoshana
Anthony Castillo
Kirk van Gorkom
Lima Zulmai
05/20/24
Neural Networks
Dr. Carter

Predicting Future Aerosol Optical Depth Values with Aeronet Data

Link to Project: https://github.com/CastilloAnthony/Neural_Nets/tree/main

Explanation of primary files in the repository

1. The dataset contains air quality data collected from the publicly available NASA Aerosol Robotic Network, termed Aeronet. It includes multiple atmospheric aerosol properties using a spectral radiometer that measures Sun and sky radiances at a number of fixed wavelengths within the visible and near-infrared spectrum.

2. The ResidualWrapper class is a custom TensorFlow Keras model wrapper that implements a residual connection. Residual connections are often used in deep learning models to help mitigate the vanishing gradient problem by allowing gradients to flow through the network more easily.

3. The DataHandler class is made to perform preprocessing functions on the data, such as data normalization and the integration of a time of day signal. The class also integrates the data into concise time-series windows, and coordinates the flow of data between the windows and the model itself.

4. The WindowGenerator class is made to handle the preparation and processing of time series data for model training. It creates input-output pairs from the raw time series data, which are then used for training, validation, and testing.

5. The Arbiter class is responsible for managing the overall workflow of data preparation, model training, and evaluation. It coordinates the different components such as data loading, window generation, and model training.

6. The Main class is responsible for starting and running the program, it controls how and when the functions of the Arbiter class gets called.
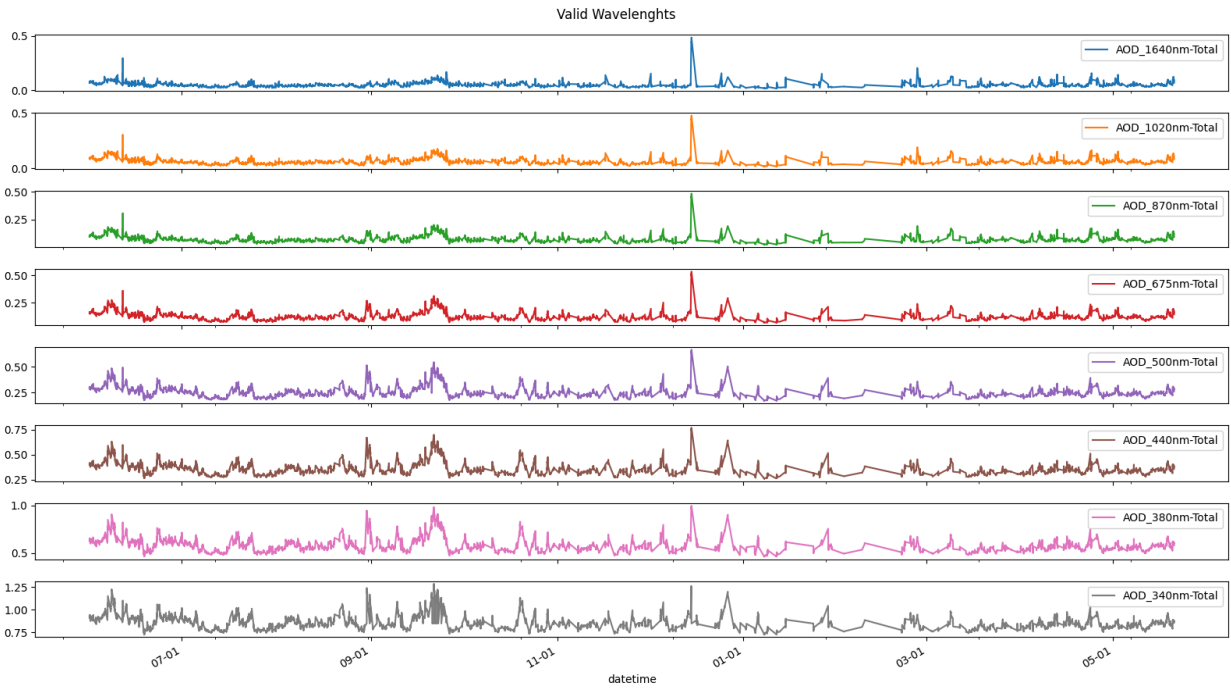
Data Visualizations

Figure 1. Visualization of available valid wavelengths from the dataset.
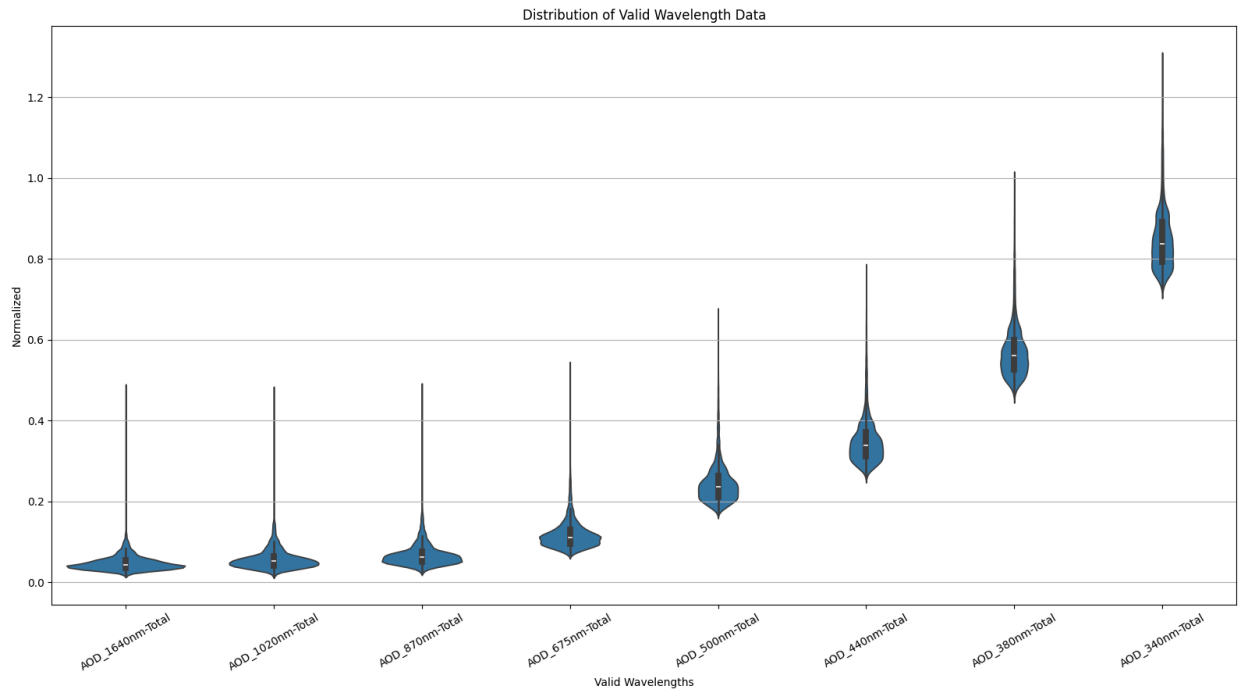

Figure 2. Visual representation of the distribution of data within each valid wavelength.
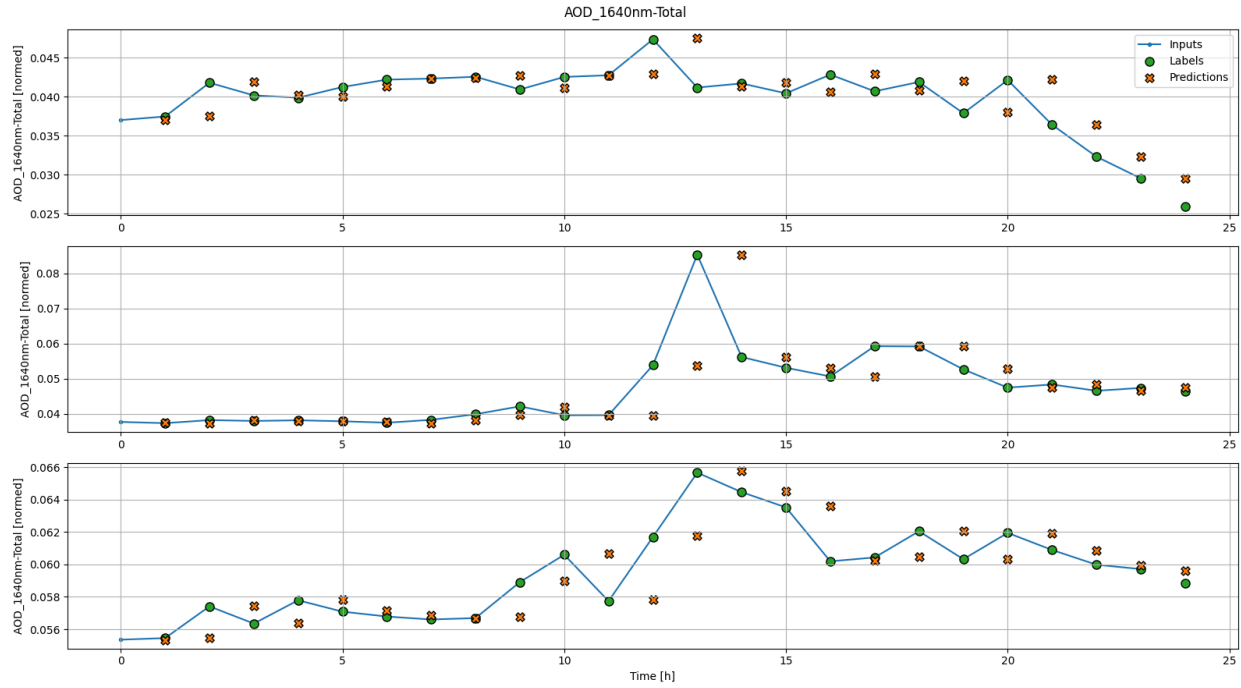
Figure 3. Visualization showing the actual values from within the 1640nm wavelength data versus the predictions made by the model.
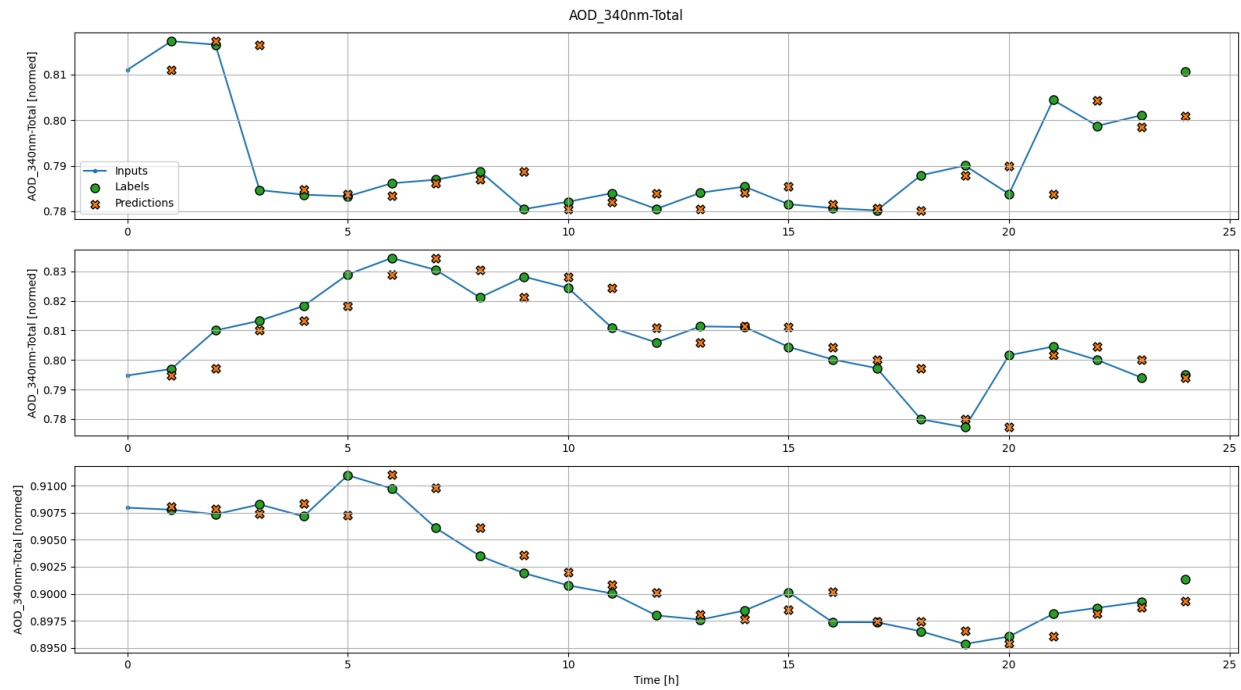

Figure 4. Visualization, similar to above, showing the actual values from within the 340nm wavelength data versus the predictions made by the model.
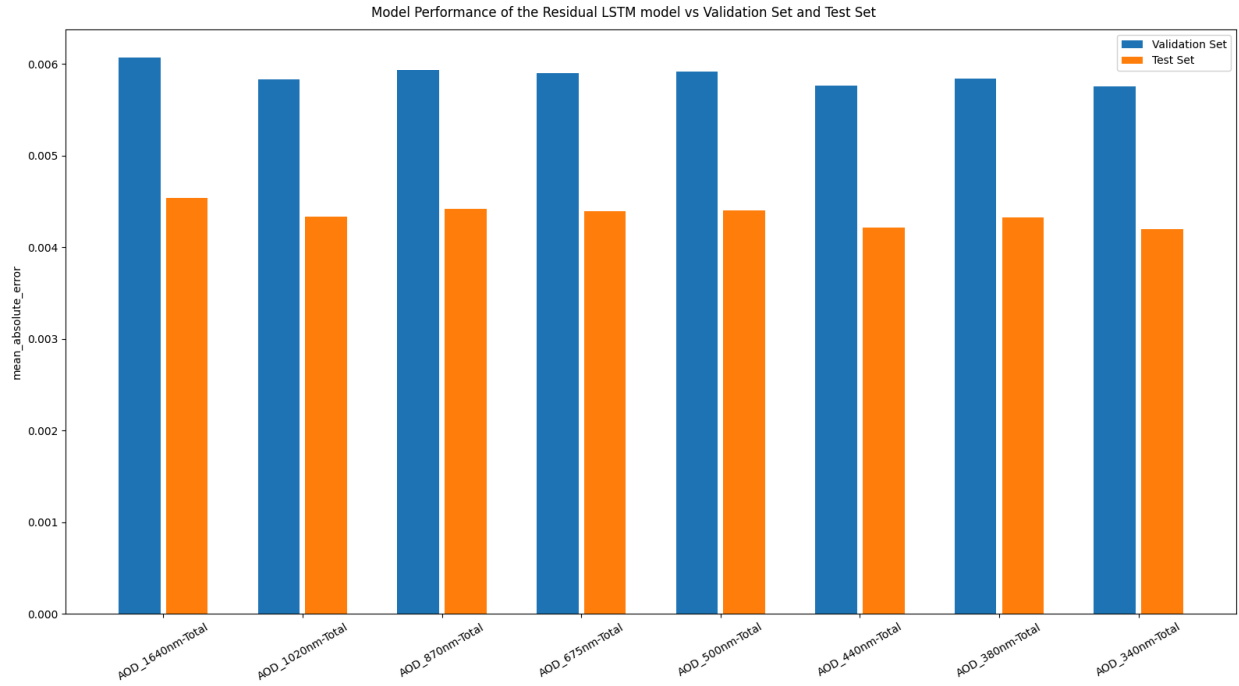
Figure 5. Visualization of the performance of the Residual LSTM model used, tested versus the validation set and the test set (lower mean_absolute_error is better).
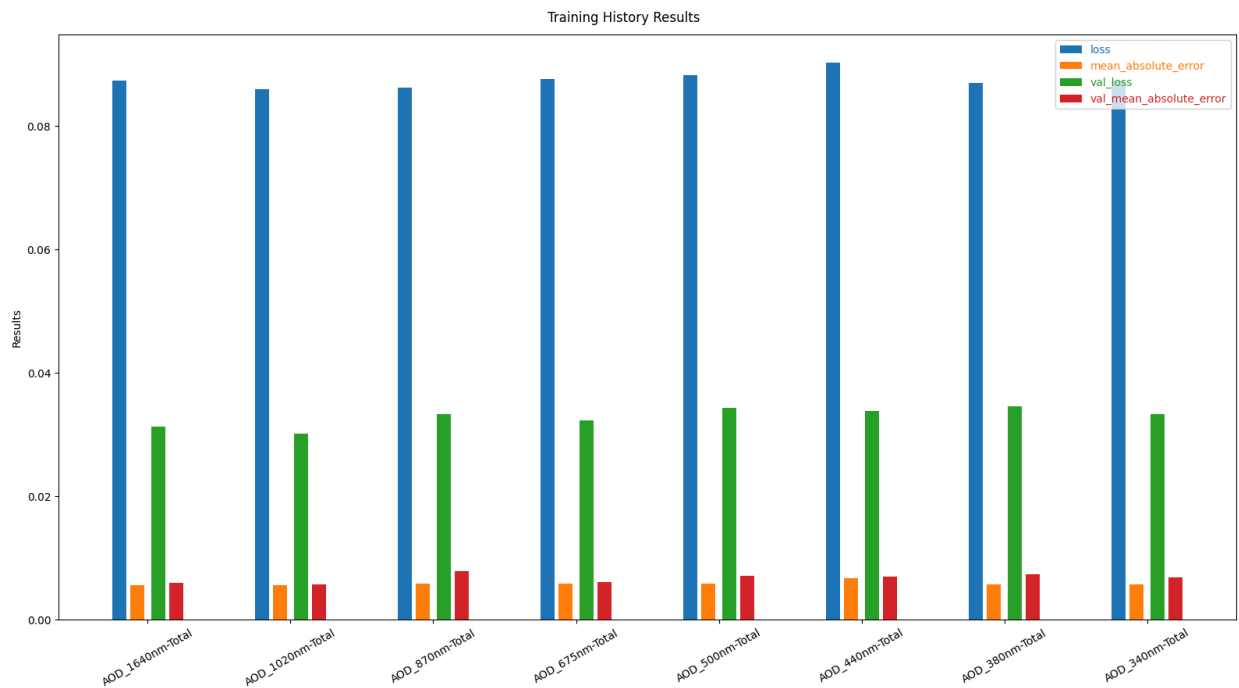

Figure 6. Visual representation of the final results of each model's training history.