# Final_Project Report

Ananya, Rosie, Sameer

2025-04-04

## Introduction

### Motivation and Background

Our group, composed of students from the Coastal Marine Systems and Energy concentrations within the MEM program, chose to explore tidal power due to its strong relevance to our academic and professional interests. Tidal energy presents a compelling case as a clean and sustainable energy source, producing zero greenhouse gas emissions while avoiding the environmental drawbacks of traditional hydro-power, such as dams and large reservoirs. Beyond its ecological benefits, tidal power boasts a high energy density and efficiency—unlike intermittent sources like wind and solar, ocean currents remain relatively constant, ensuring continuous power generation without heavy reliance on costly storage systems. This reliability makes tidal energy a valuable asset in balancing energy grids and complementing other renewables. Additionally, the predictability of tidal currents enhances its appeal as a stable energy solution. While the U.S. currently lacks commercially operational tidal power sites, growing interest and investment, including pilot projects in Maine and New York, signal its potential as a key player in the renewable energy landscape. (Sources of references: altenergymag.com)

### Project Objectives and Relevance

In this project, we aim to evaluate and compare tidal power potential across three U.S. sites at varying latitudes and coastal conditions. Our assessment focuses on Wave Energy Flux (measured in watts per square meter), which is calculated by Wave Power Density multiplying by Wave Energy Period. We will analyze seasonal and long-term trends in these time-series datasets, apply predictive modeling using training and testing datasets, and compare forecasting results to determine which site offers the most stable and highest Wave Energy Flux. By doing so, we hope to contribute meaningful insights into the feasibility and optimization of tidal energy as a sustainable power source. (Sources of data: NREL Marine Energy)

## Dataset information

### Data selection and extraction

We first picked 3 sites/ locations/ areas tend to be some of the locations referred to for its tidal power potential in the general Internet searches. These areas are distinct by latitudes and longitudes, as well as different coasts and climatic environments. These three sites are:
- Gulf of Maine (43.68° N, 69.77 ° W)
- ~ 3 km distance into ocean near Miami, Florida (25.74 ° N, 80.11 ° W)
- Cook Inlet, Alaska (59.37° N, 152.64° W)

With the lat-long of these sites, we extracted raw data of 1979-2010 Wave Energy Density and Energy period data in NREL marine energy atlas: http://nrel.gov/marine-energy-atlas/data-viewer

### Data Wrangling & General Steps of Analysis

1. Calculate the Wave Energy Flux data based on the raw data we extracted.

2. Wave Energy Flux datasets for all 3 sites' are aggregated to daily and monthly view (original series).

3. Split the current full dataset using 80/20 rule into training and testing datasets.

4. Conduct the original time series analysis on all of the sites.

5. Run various models to compare predicting/ forecasting performance on the Maine dataset.

6. Get the top 3 performing models based on the Maine Wave Energy Flux forecasting
7. Use the top 3 performance models (Arima + Fourier k-4, TBATs. and STL+ETS) for the other 2 sites Wave Energy Flux.

8. Compare Wave Energy Flux Results of the 3 sites and determine which site has more potential.

### Initial Summary Statistics/ Visualizations of the Datasets

We do a preliminary analysis of the three sites based on summary statistics for monthly mean wave power. We see drastic contracts among them, with the Gulf of Maine delivering the most energetic and variable resource, averaging 8,443.4 W/m. On the other hand, Cook Inlet's sheltered basin gives us only ~16.2 W/m of monthly wave power on average. Thus, based on these observations, we first test out forecasting models using the Gulf of Maine site, and then choose the top three performing models on the other two sites for a final comparison of tidal power potential.

## Initial Summary Statistics

Table 1: Summary statistics of monthly mean wave power by site

| Site | Mean (W/m) | Median (W/m) | SD (W/m) | Min (W/m) | Max (W/m) |
|---|---|---|---|---|---|
| Alaska | 16.2 | 11.5 | 15.3 | 0.0 | 105.1 |
| Florida | 3422.6 | 3167.6 | 2271.9 | 268.7 | 12334.1 |
| Maine | 8443.4 | 7346.8 | 4897.7 | 1528.5 | 28885.9 |

Here are the first ten rows of our Maine dataset

## Analysis

### Maine

As concluded from the initial summary statistics, we start off with Maine as a test site. The location we chose for Maine is 3 kilometers offshore with reliable wave and wind currents and thus, a good starting point to test our ability to forecast tidal power.

We proceeded with looking at the time series plots for both daily and monthly mean wave power. We also looked at the ACF and PACF plots of both series, so that we can make an informed decision on which time series (daily or monthly) to pursue.

From Figure 1, we see that while the daily series provides a more granular view of the fluctuations in mean wave power, it is subject to extreme outliers that can affect our forecasting ability. Additionally, when compared to the monthly series, there is a lot more noise in the graph, making it harder to discern trends.

The monthly series on the other hand shows a clearer seasonal pattern. Further, by averaging out to a monthly series, the impact of outliers is diminished.

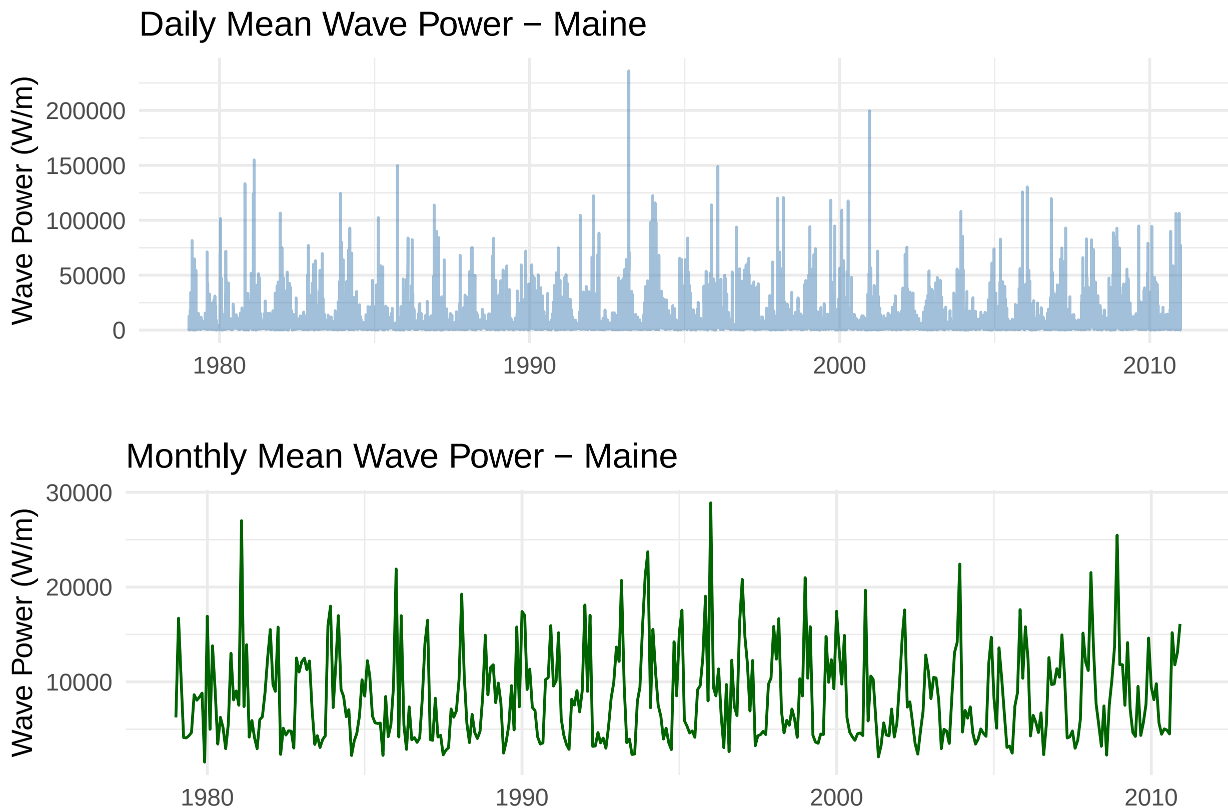
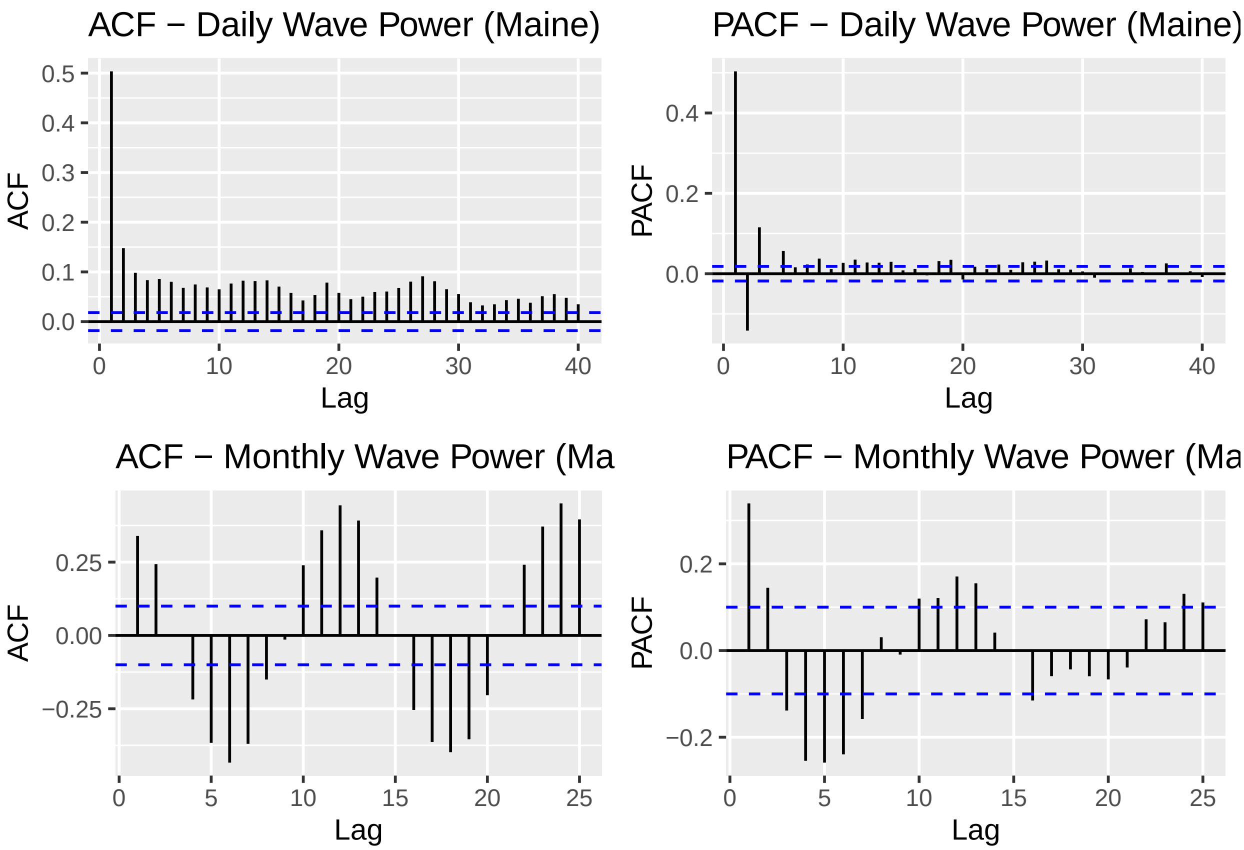



Figure 1: Figure 1. Plots for mean wave power in Maine (Daily and Monthly)

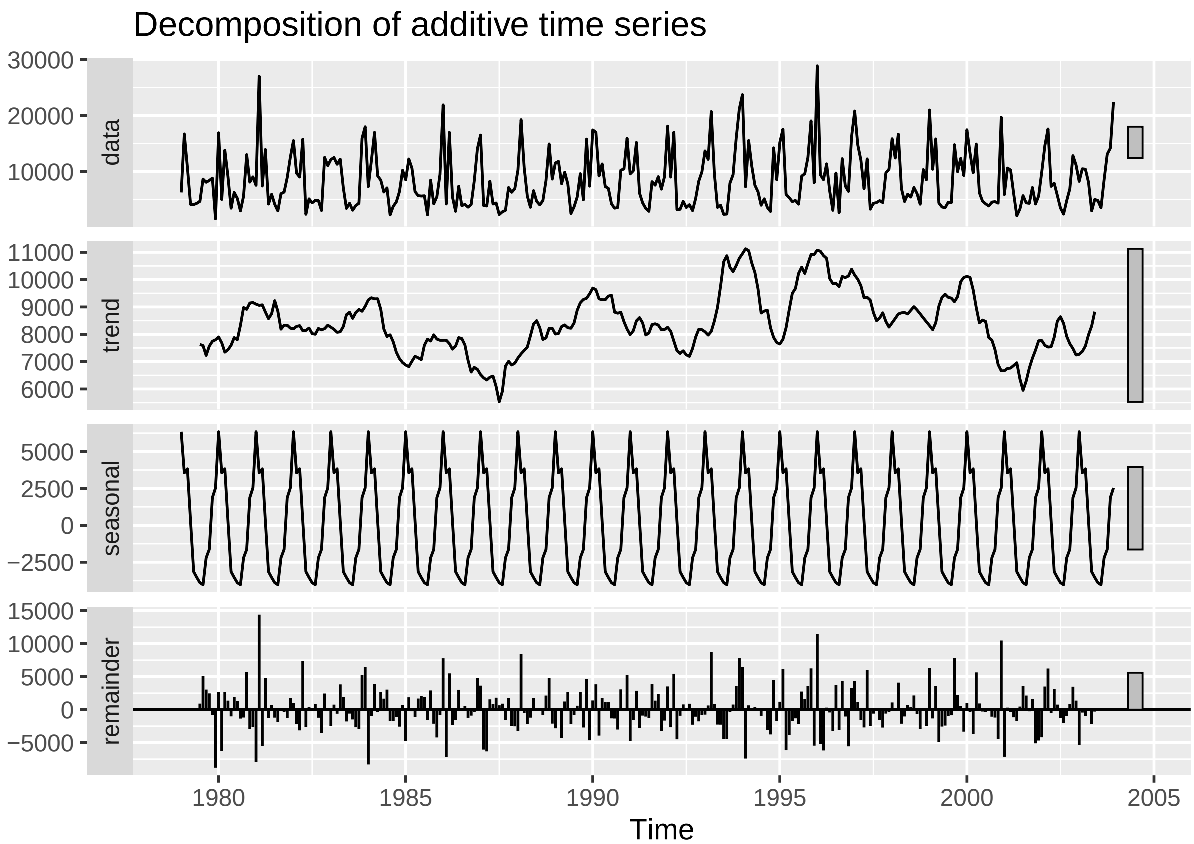
From Figure 2, we see that both ACF and PACF plots for the daily series have a sharp drop-off after lag 1, suggesting that past daily mean wave power values do not carry much signal for future values. For the monthly mean wave power, the ACF and PACF plots shows strong seasonality with the wave pattern. The autocorrelation persists over time, which will be ideal for ARIMA/SARIMA and other seasonal models.




**Preparing the data**

Thus, we will proceed with forecasting for the monthly series. We start by splitting our monthly mean wave power datasets using the 80-20 rule: 80% for training, 20% for testing. Our training dataset for Maine contains mean wave power from January 1979 to December 2003 and our testing dataset for Maine contains mean wave power from January 2004 to December 2010.

We used the decompose() function on our training dataset and performed an additive deomposition. From Figure 3, we can confirm the strong seasonality seen in mean wave power in Maine. We also see that mean wave power peaked from 1995 to around 2000 in Maine. This could be due to climate and oceanographic factors (wind speed, underwater currents) that warrant a further look. Since tidal energy is subject to a whole host of climate and oceanographic factors, our remainder series shows some significant spikes at irregular intervals which could be caused by specific weather events. Further research into Maine's short-term and long-term weather patterns is recommended for future forecasting projects.




**Start of forecasting models**

We proceed by setting the forecasting length to be from January 2004 to December 2010 (testing dataset). We will evaluate our forecasts against values found in our testing datasets.

For forecasting, we trained and tested the following models: - Seasonal Naïve: Used as a preliminary baseline due to strong seasonal patterns in monthly wave power

- ARIMA + Fourier: Applied using auto.arima() function with a fourier term
- ETS + STL: Implements seasonal trend decomposition (Figure 3) and performs exponential smoothing using stlf() function
- SARIMA: Applied a seasonal ARIMA model (due to strong seasonality present) using the auto.arima() function
- ES: Exponential smoothing to use recent, past values of monthly wave power to forecast future ones
- StructTS: Uses a state-space framework to model unobserved components of monthly tidal series. Applied using a Basic Structural Model (BSM).
- TBATS: Used to model complex seasonal patterns present in tidal power
- Neural Network + XREG: Single, hidden-layer neural network (with our training dataset as an external regressor).

Table 1 displays our accuracy metrics for our models.

Judging from the RMSE and MAPE, our 3 best models are STL + ETS, TBATS and ARIMA + Fourier. This is in line with our hypothesis that models that capture the complex seasonality of tidal power would perform best.

```
## The best model by RMSE is: ARIMA+Fourier
```

Table 2: Forecast Accuracy for Monthly Wave Power - Maine

| | ME | RMSE | MAE | MPE | MAPE | ACF1 | Theil's U |
|---|---|---|---|---|---|---|---|
| SNAIVE | -822.52074 | 4290.692 | 3202.526 | -19.62825 | 42.60721 | 0.06549 | 0.89498 |
| SARIMA | -217.59468 | 4026.117 | 3221.739 | -29.70686 | 50.01054 | 0.32049 | 0.83462 |
| STL+ETS | -192.97893 | 3730.735 | 2709.232 | -17.21843 | 36.39588 | 0.18831 | 0.70042 |
| ARIMA+Fourier | 736.38317 | 3690.025 | 2601.569 | -5.30294 | 31.99013 | 0.20083 | 0.72561 |
| ES | 4686.01400 | 5864.526 | 4762.718 | 78.82107 | 80.54519 | -0.00122 | 1.84991 |
| StructTS | 4686.01400 | 5864.526 | 4762.718 | 78.82107 | 80.54519 | -0.00122 | 1.84991 |
| TBAT | 569.08601 | 3728.780 | 2690.142 | -8.37019 | 33.99924 | 0.17281 | 0.70592 |
| NN | -61.90436 | 4012.568 | 2846.284 | -13.62389 | 36.60158 | 0.15156 | 0.73000 |

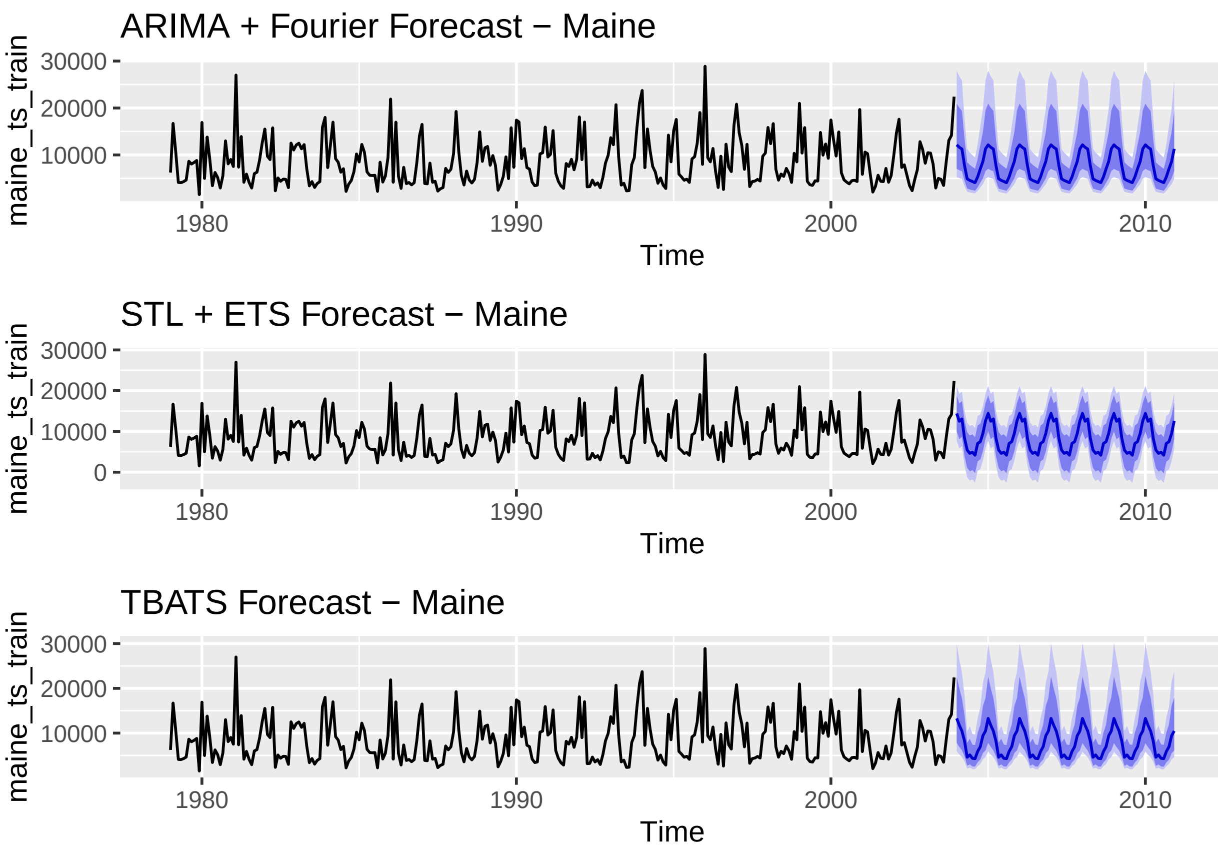
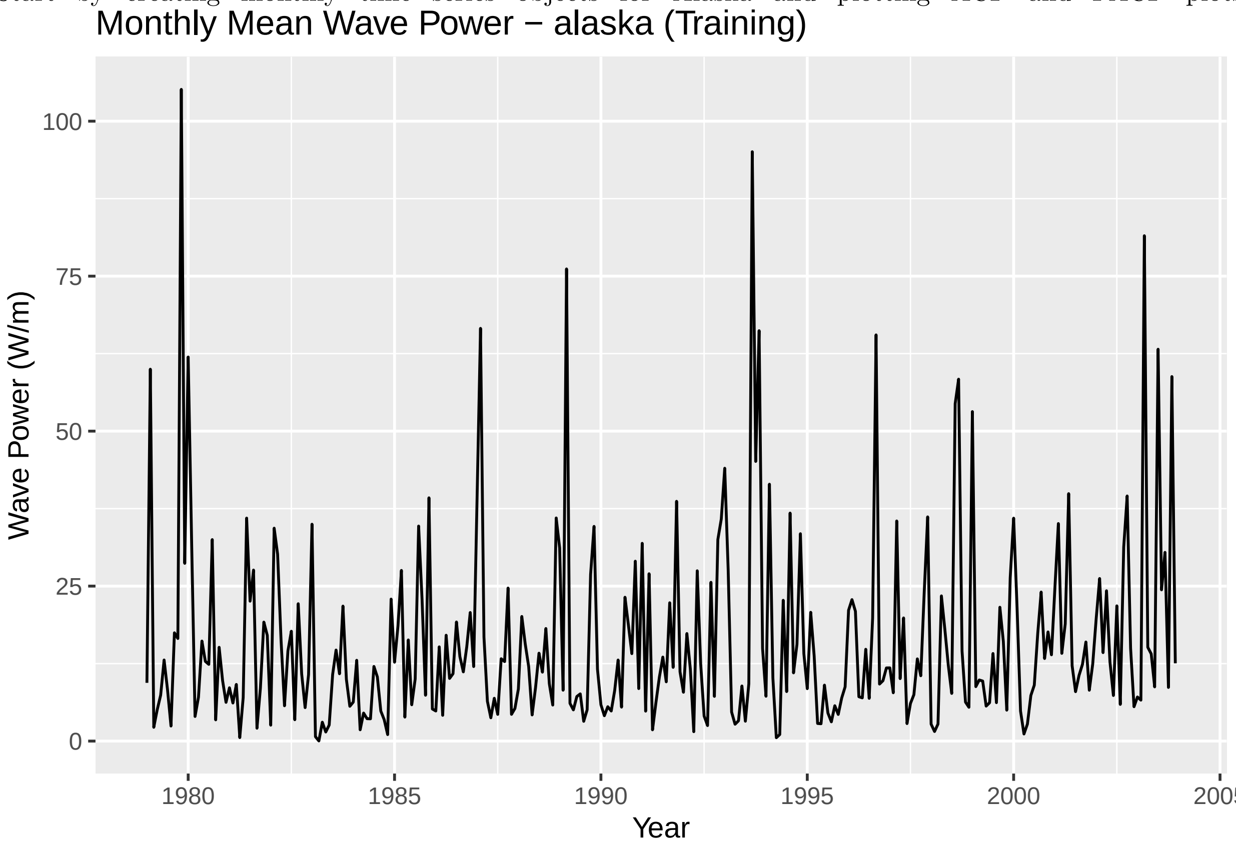
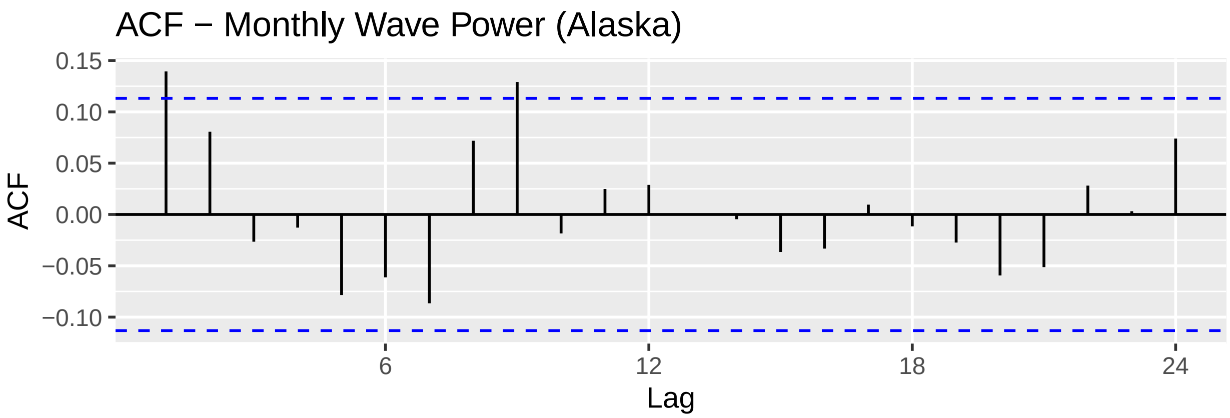


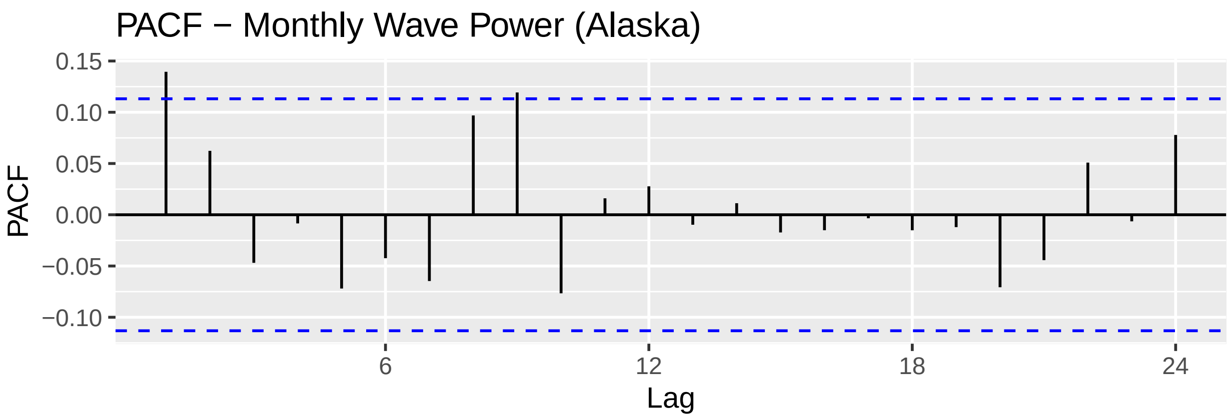

Figure 2: Figure 4. Forecast comparisons of our top 3 models

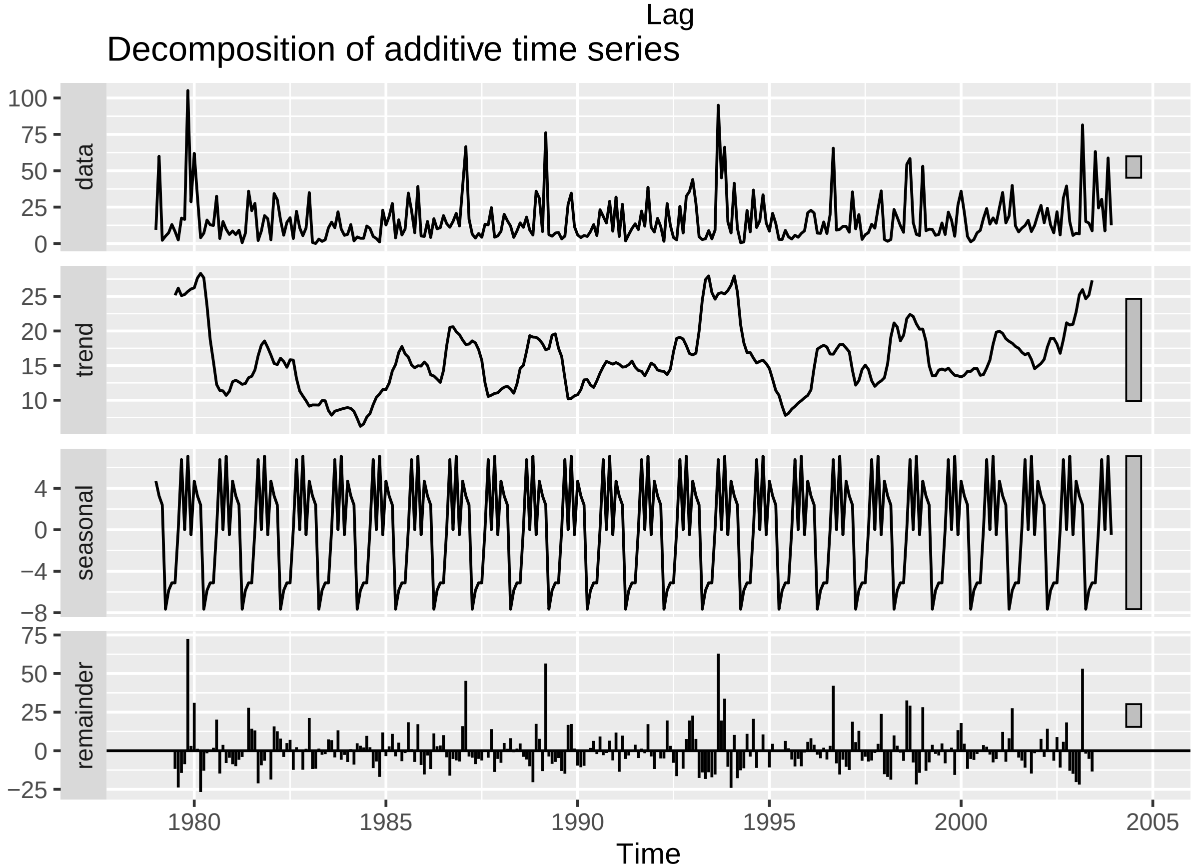
### Alaska

Start by creating monthly time series objects for Alaska and plotting ACF and PACF plots






Observations: The significant spike at lag 1 in both ACF and PACF strongly suggests an AR(1) component. Also, the significant spikes at lag 12 in both ACF and PACF indicate a strong seasonal autoregressive component with a period of 12 months (SAR(1) with a seasonal lag of 12).

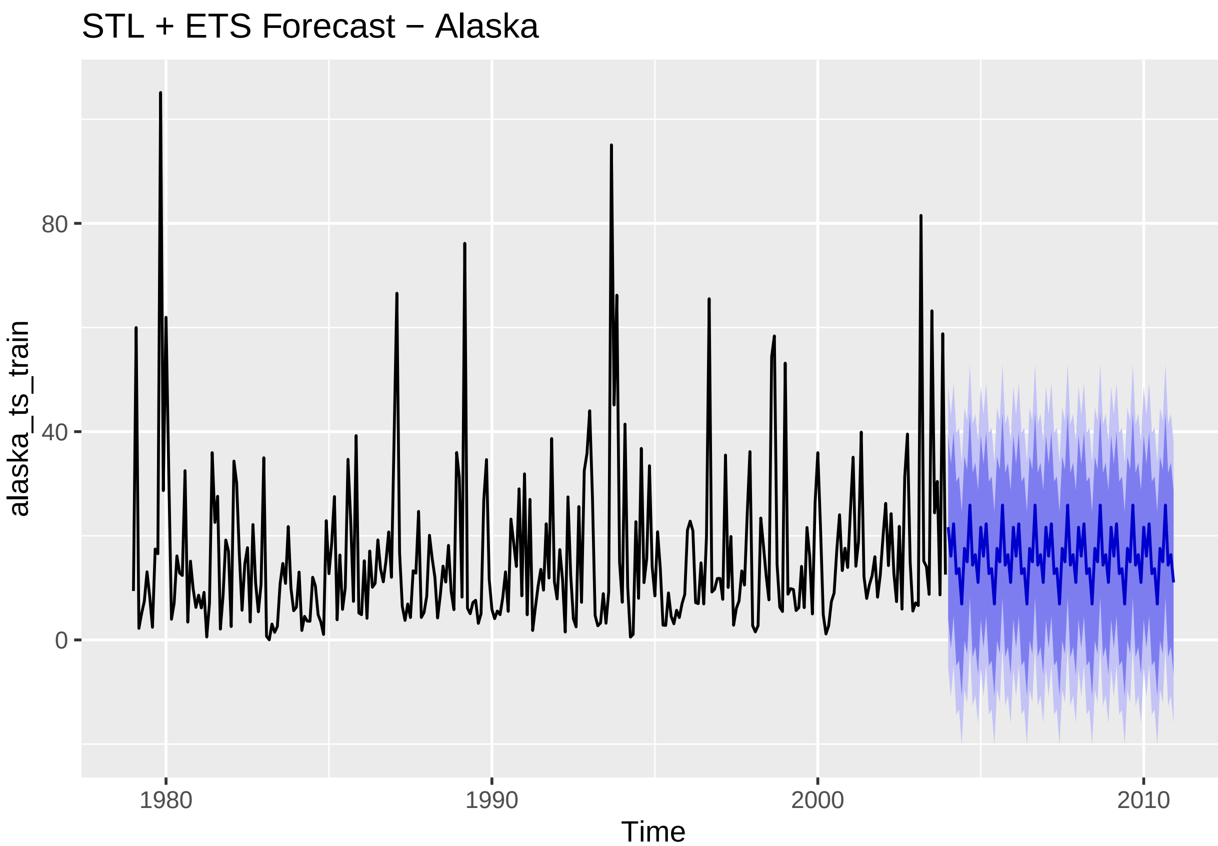
```
## Warning in adf.test(alaska_ts_train): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  alaska_ts_train
## Dickey-Fuller = -7.3596, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary
```

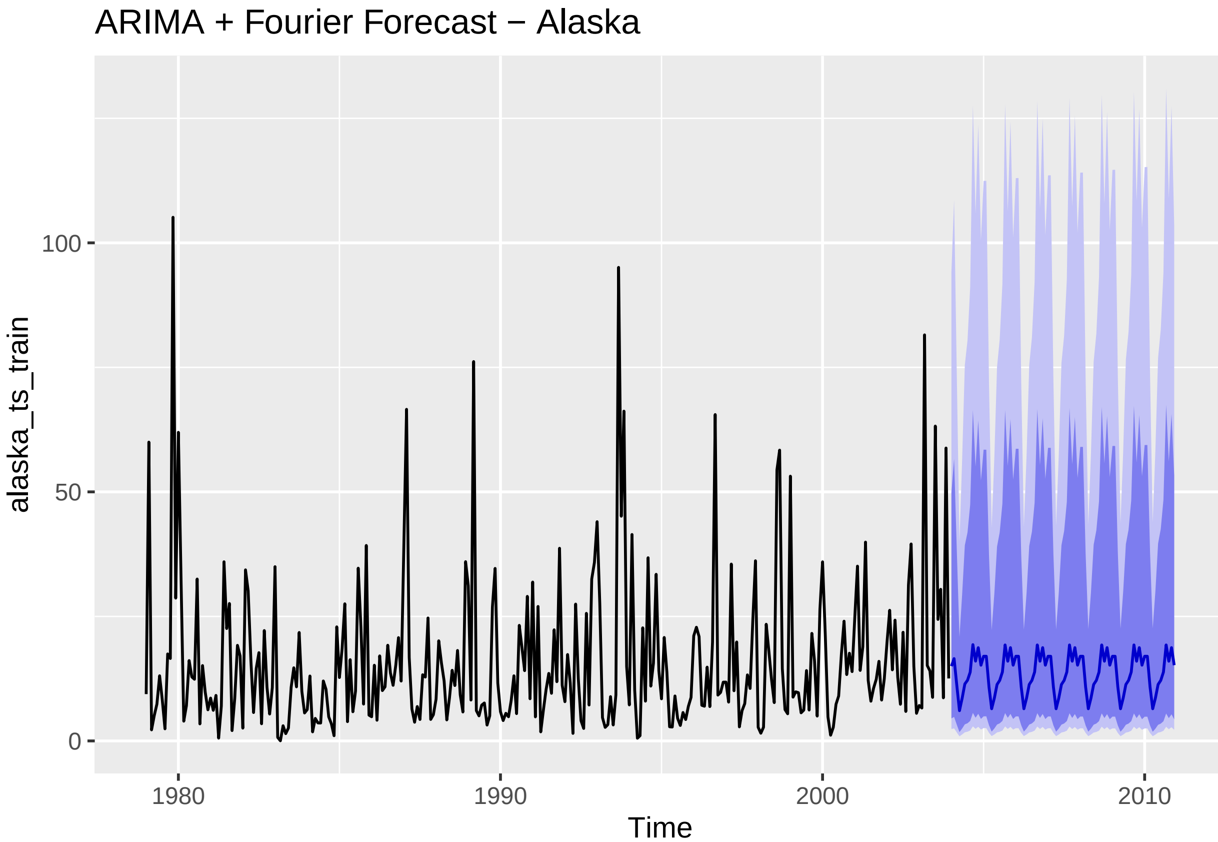
Observations: ADF test returns a p-value of 0.01, which is smaller than our chosen significance level of 0.05. We reject the null hypothesis that the Alaska mean monthly wave power series has a unit root and thus, the series is likely stationary and does not need differencing.

Proceed with using our 3 chosen models on Alaska

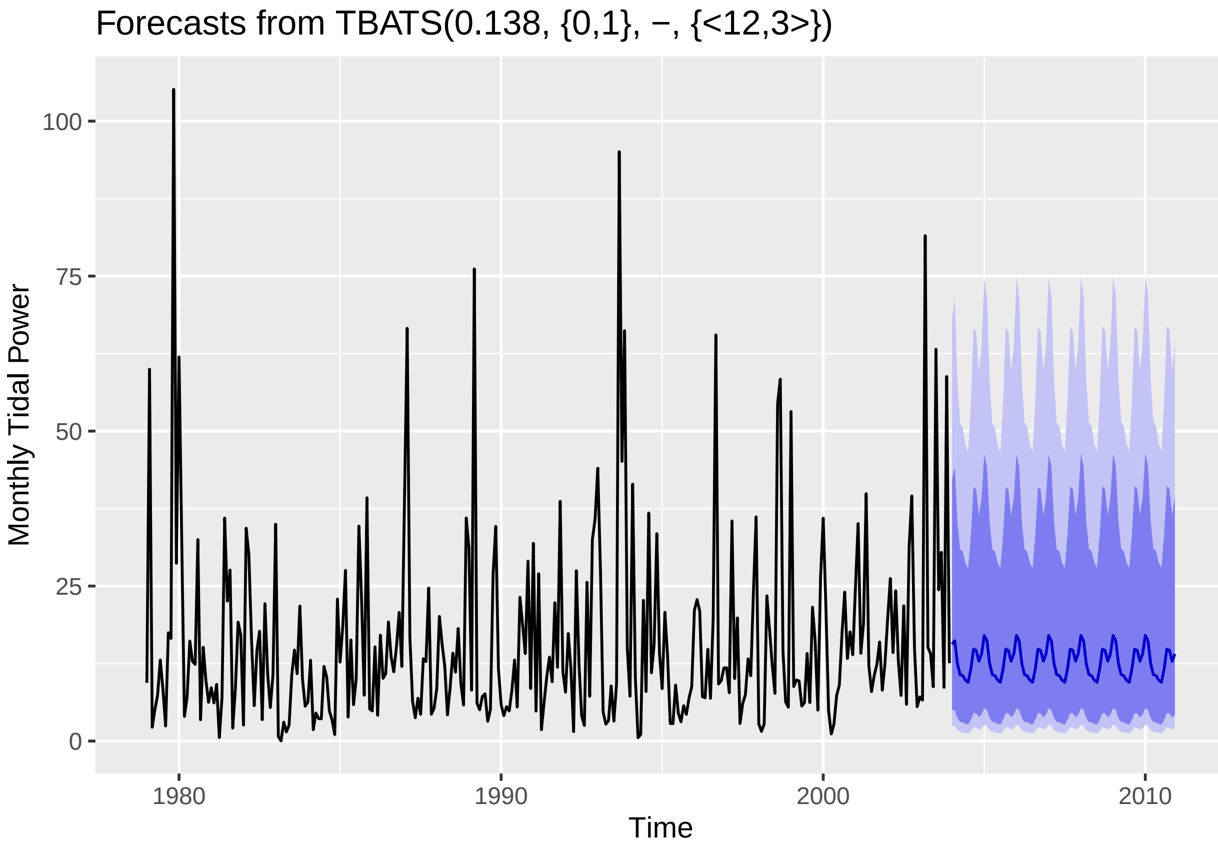
**Model 1: STL decomposition + ETS**

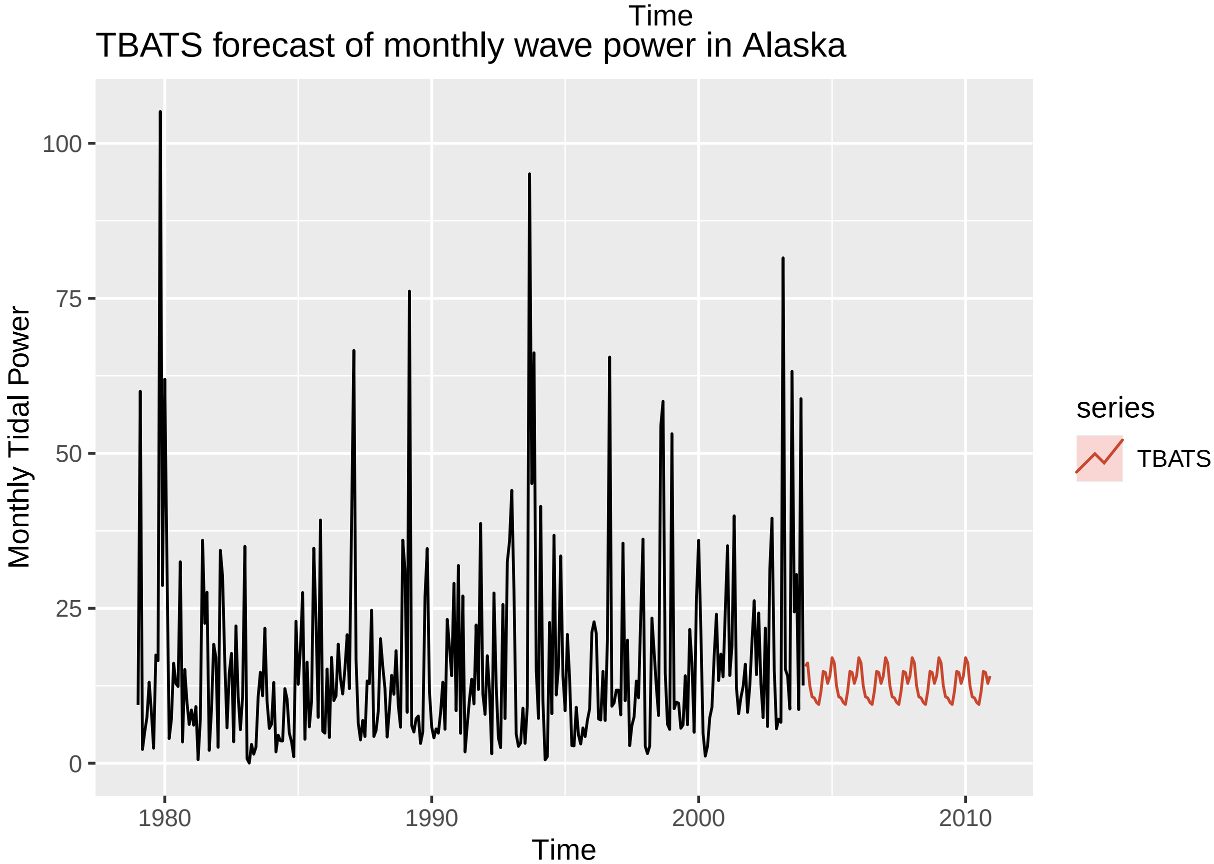


**Model 2: ARIMA + Fourier terms**



**Model 3: TBATS**



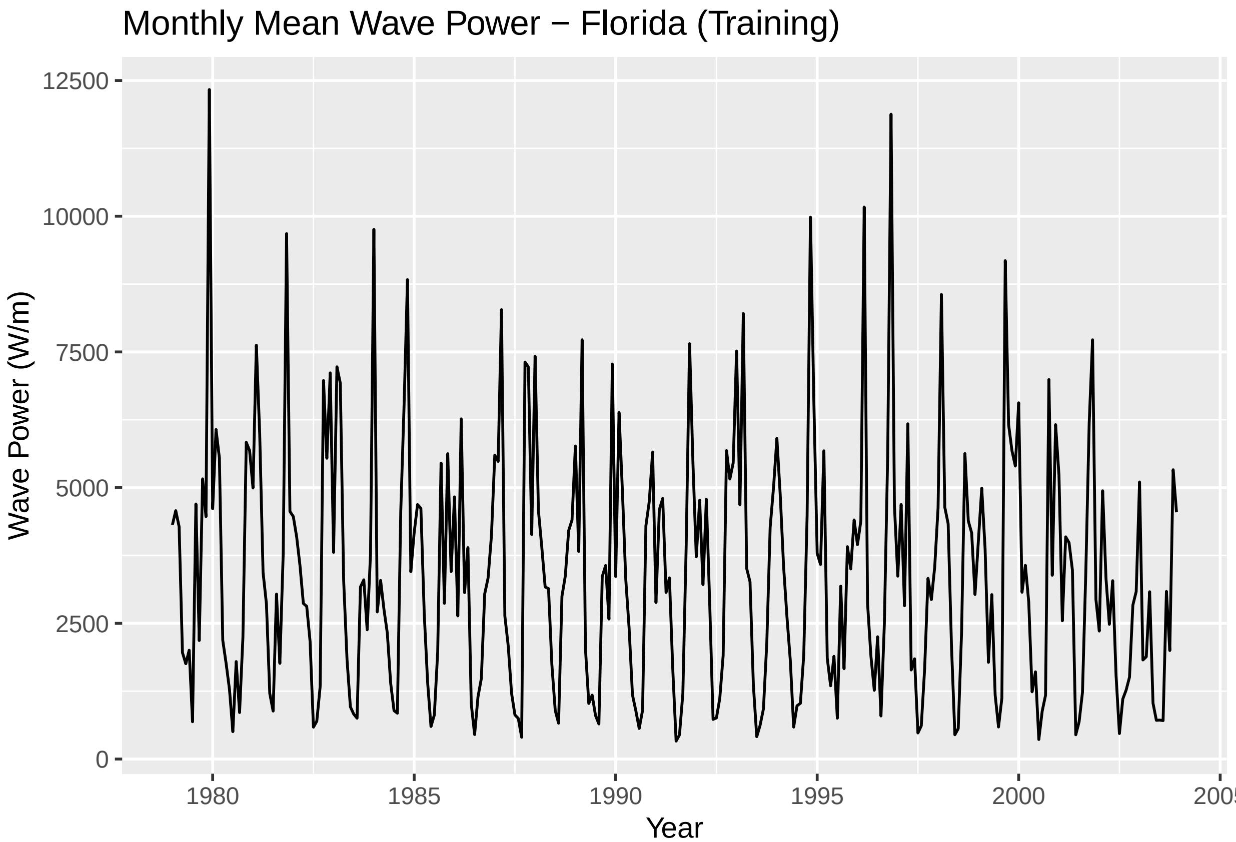
```
## The best model by RMSE is: ARIMA+Fourier
```

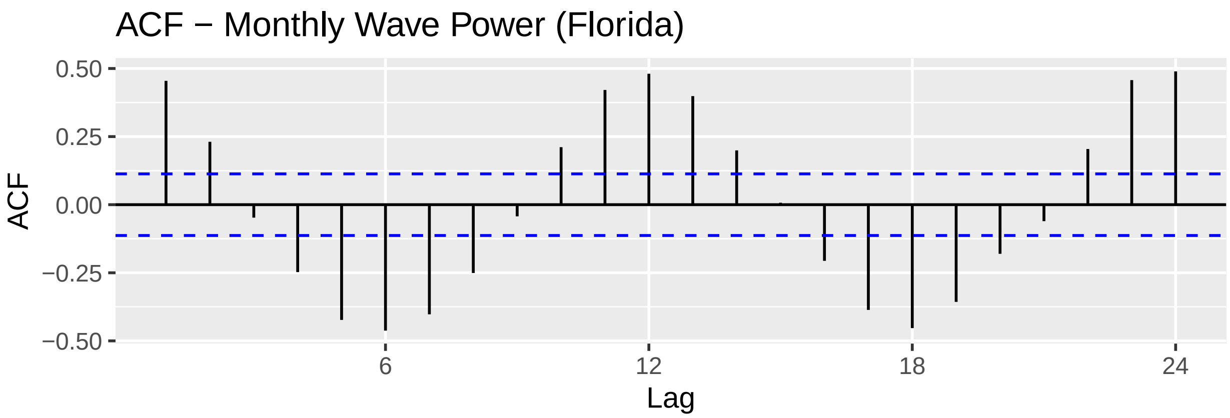
Table 3: Forecast Accuracy for Monthly Wave Power - Alaska

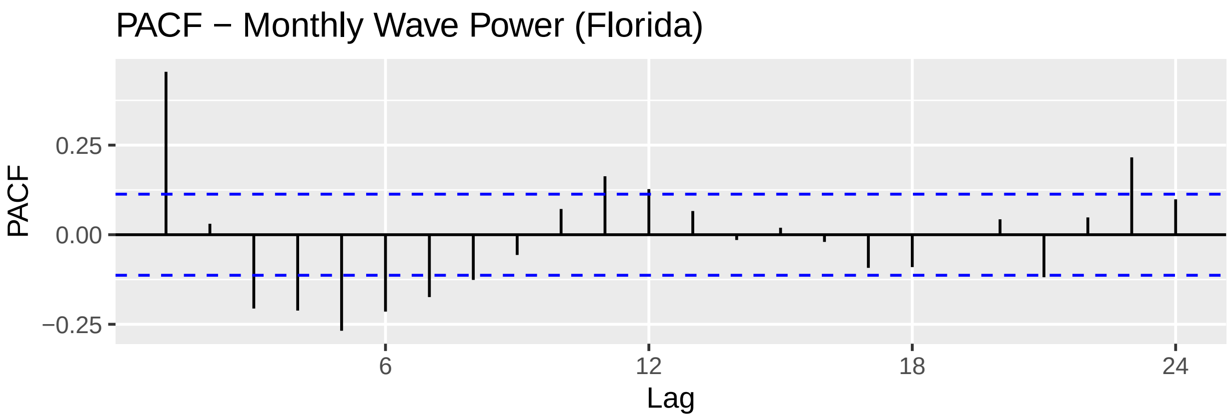
| | ME | RMSE | MAE | MPE | MAPE | ACF1 | Theil's U |
|---|---|---|---|---|---|---|---|
| STL+ETS | 0.41065 | 14.12462 | 10.06050 | -118.37052 | 148.4886 | 0.24409 | 0.44005 |
| ARIMA+Fourier | 2.70368 | 13.68095 | 9.00760 | -79.91860 | 115.1316 | 0.19990 | 0.39252 |
| TBAT | 3.71047 | 14.20461 | 8.81815 | -71.65192 | 107.4045 | 0.18163 | 0.41341 |

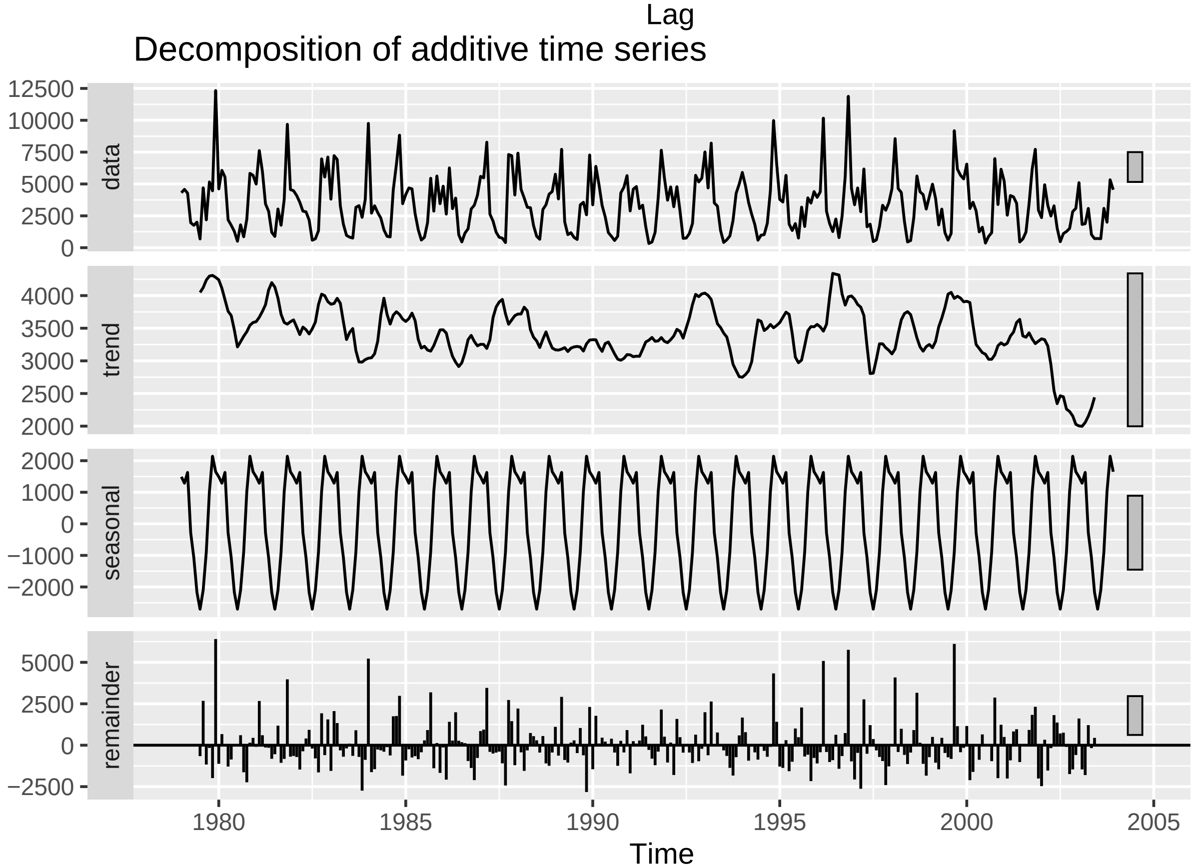
### Florida

Start by creating monthly time series objects for Florida and plotting ACF and PACF plots






Observations: The significant spike at lag 1 in both ACF and PACF strongly suggests an AR(1) component. Also, the repeating seasonal patterns at lag 12 for the ACF suggest strong yearly seasonality.

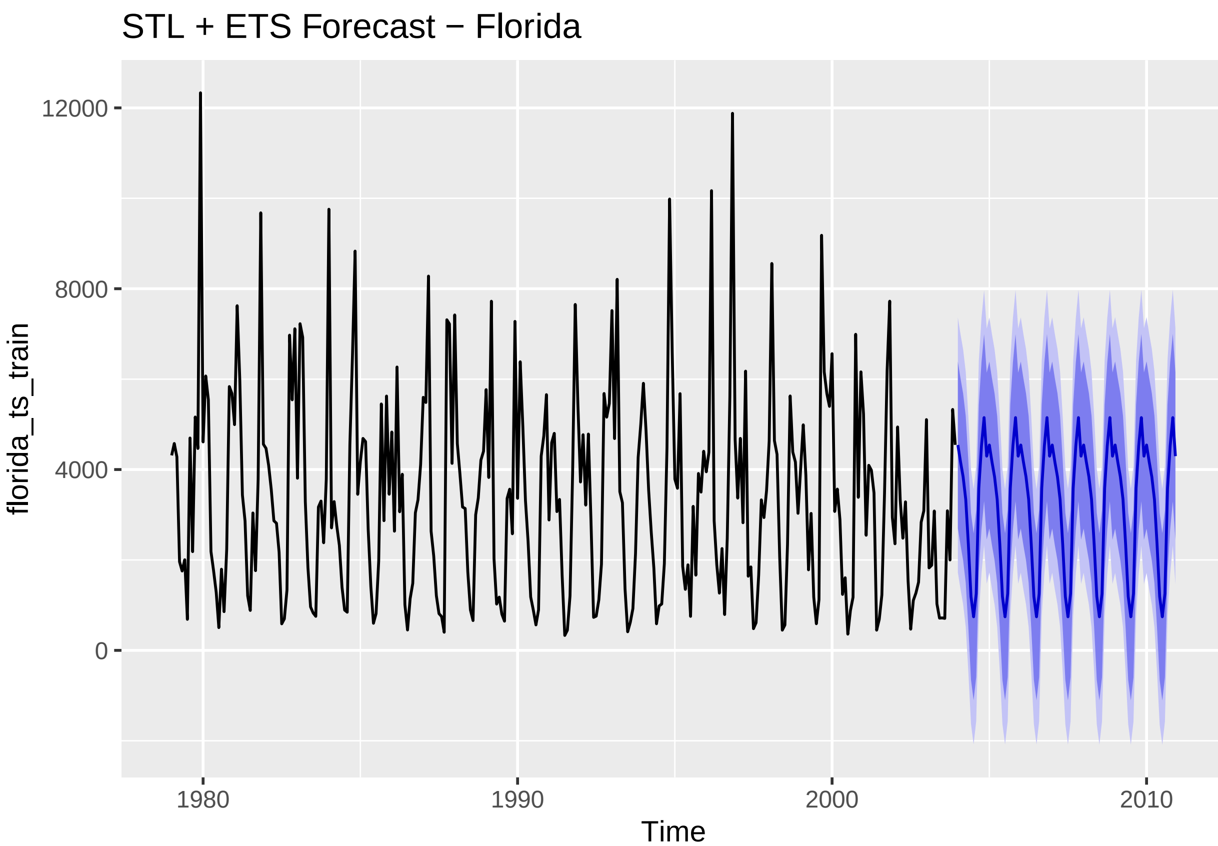
```
## Warning in adf.test(florida_ts_train): p-value smaller than printed p-value

##
##  Augmented Dickey-Fuller Test
##
## data:  florida_ts_train
## Dickey-Fuller = -12.062, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary
```

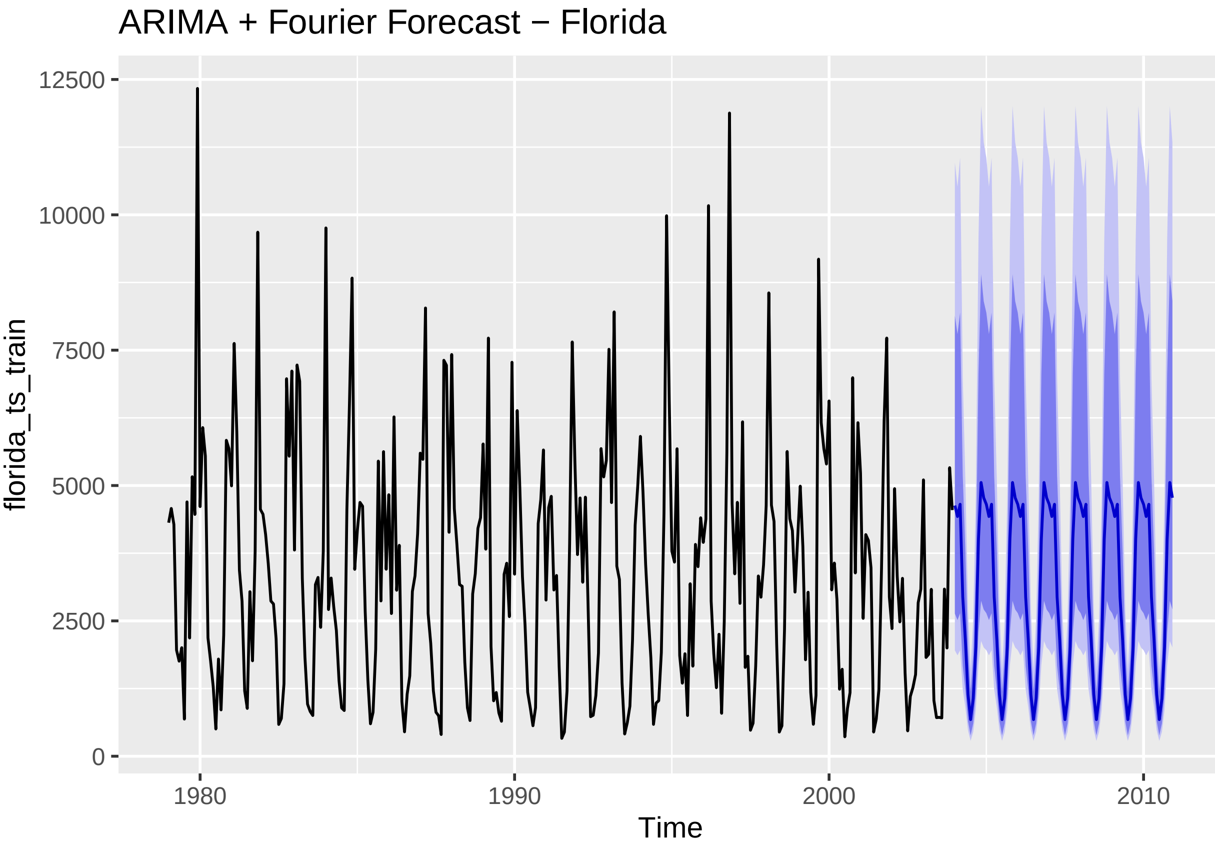
Observations: ADF test returns a p-value of 0.01, which is smaller than our chosen significance level of 0.05. We reject the null hypothesis that the Alaska mean monthly wave power series has a unit root and thus, the series is likely stationary and does not need differencing.

Proceed with using our 3 chosen models on Florida

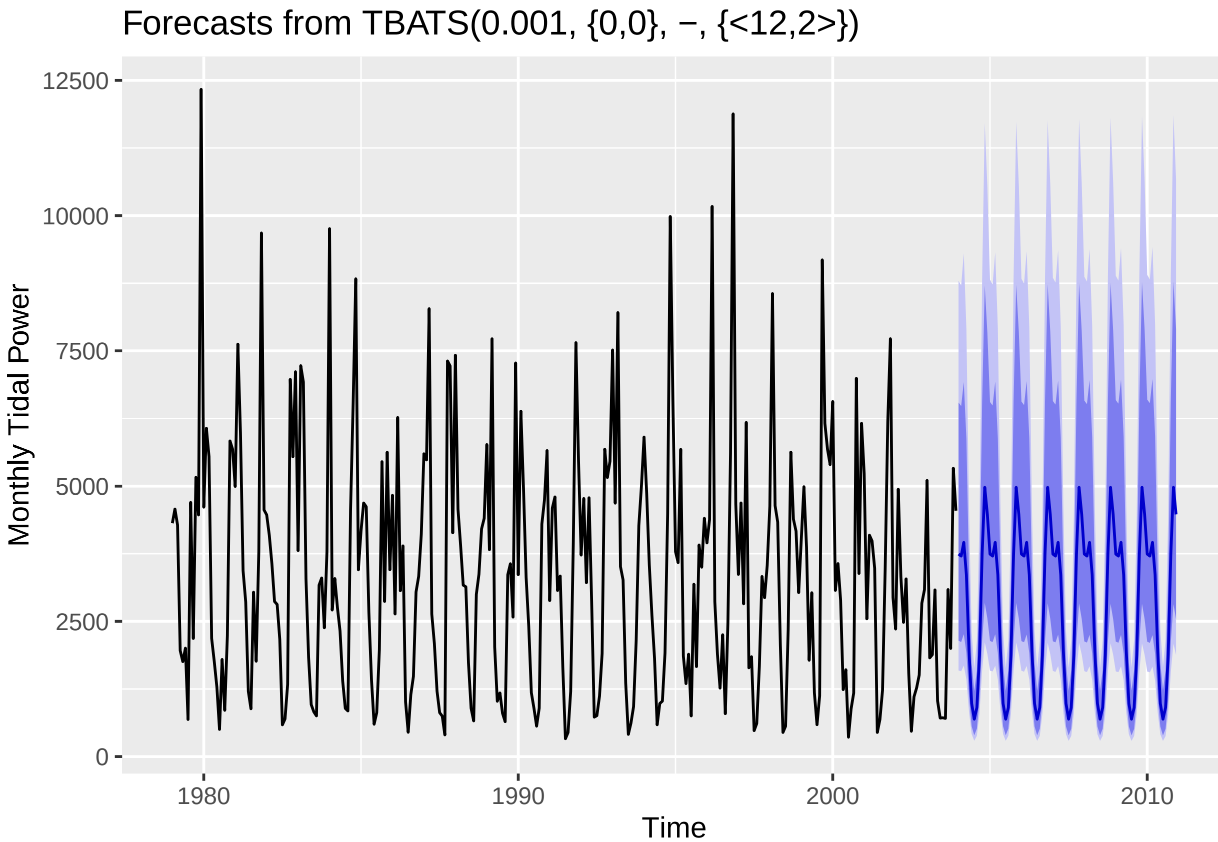
**Model 1: STL + ETS**

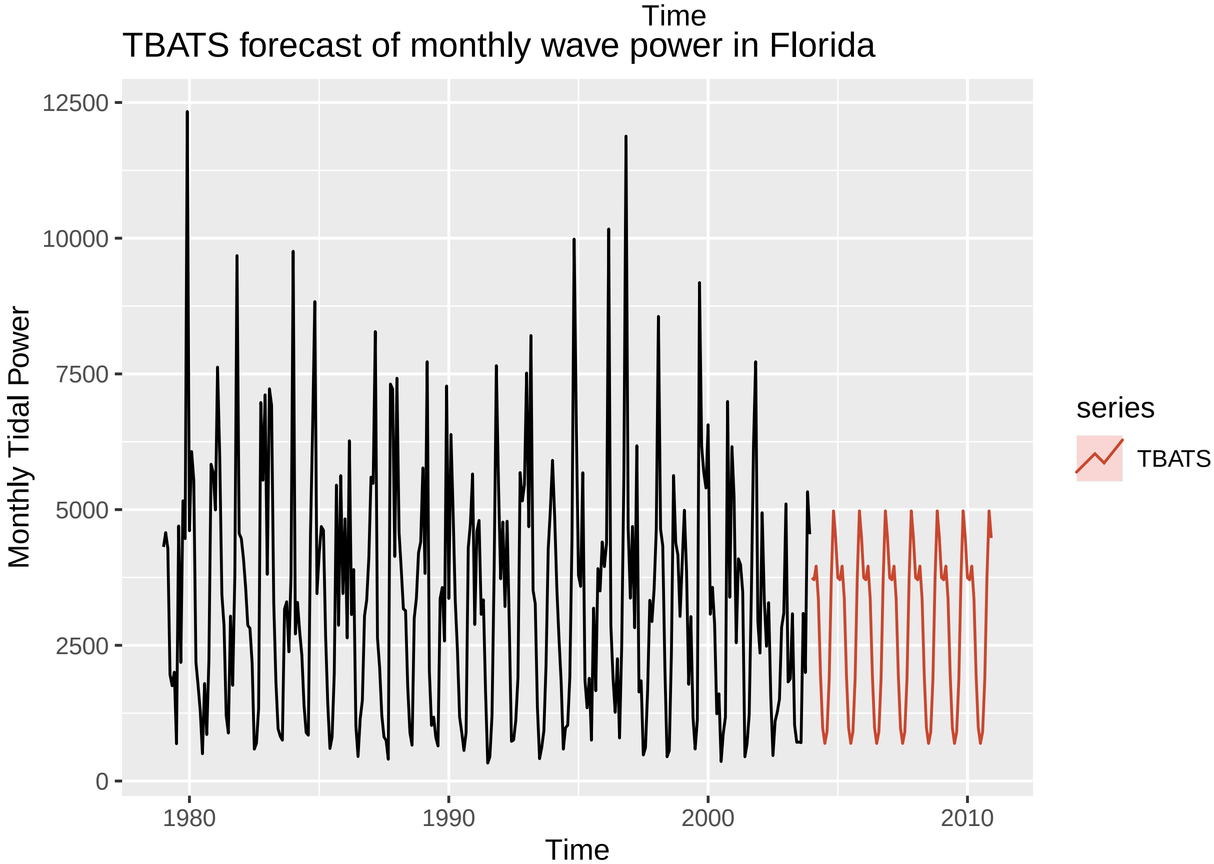


**Model 2: ARIMA + Fourier terms**



**Model 3 TBATs**



```
## The best model by RMSE is: STL+ETS
```

Table 4: Forecast Accuracy for Monthly Wave Power - Florida

| | ME | RMSE | MAE | MPE | MAPE | ACF1 | Theil's U |
|---|---|---|---|---|---|---|---|
| STL+ETS | 260.0615 | 1884.797 | 1242.601 | -22.94636 | 48.47699 | 0.04601 | 0.78809 |
| ARIMA+Fourier | 377.5698 | 2019.667 | 1254.629 | -13.78776 | 42.55154 | 0.03404 | 0.85821 |
| TBAT | 644.5316 | 2081.479 | 1271.549 | -4.49300 | 39.17815 | 0.02116 | 0.86973 |

# Comparing the three sites

Observation: We see that Maine not only delivers the highest monthly wave power, but also results in the most reliable forecasts as it has the lowest relative error (MAPE%). Alaska's Cook Inlet site is highly unpredictable, with a MAPE of 115%, essentially meaning that our one-year ahead forecast could be no better than random guesses!

Limitations and Future Analyses: - Currently our models are purely univariate as we use the monthly mean wave power as a proxy for tidal power potential. However, there are a lot of other factors that affect the wave patterns, such as wind speed, sea surface temperature, weather events etc. These could have significant impacts on the forecasts, especially at highly variable sites. Also, our current dataset spans the time period from 1979-2010, which does not capture the recent climate change impacts. So, future analyses would involve digging up more recent data and redoing our analysis based on that to see how the conditions have changed. Apart from extending our models temporally, we could also incorporate a wider spatial component, wherein we take the wave potential from a defined area rather than point estimates.

# List References and Datasets:


- Evaluating the present and future potential of tidal power in the U.S. AltEnergyMag. (n.d.). https://www.altenergymag.com/article/2024/10/evaluating-the-present-and-future-potential-of-tidal-power-in-the-us/43464

- Lavaa, A. (2023, April 24). Tidal Energy Diagram: The story of tidal power based on diagrams - industrial manufacturing blog. linquip. https://www.linquip.com/blog/tidal-energy-diagram-story-of-tidal-power/

- Marine Energy Atlas. Innovative Data Energy Applications. (n.d.). https://maps.nrel.gov/marine-energy-atlas/data-viewer/download?vL=WavePowerMerged&b=%5B%5B-150.717111%2C43.537926%5D%2C%5B-114.418117%2C51.672586%5D%5D


Table 5: Forecast performance by site

| Site | Best Model | Mean Power (W/m) | RMSE | MAPE (%) |
|---|---|---|---|---|
| Maine | ARIMA+Fourier | 8443.415 | 3690.025 | 31.990 |
| Alaska | ARIMA+Fourier | 16.232 | 13.681 | 115.132 |
| Florida | STL+ETS | 3422.569 | 1884.797 | 48.477 |