### **Time series Forecasting Tidal Power Generation: A comparative study with Machine Learning and Quantile Regression**

### **Project Overview**

### **Background**

Tidal and wave energy are promising sources of renewable energy, offering a continuous and predictable power supply compared to other intermittent sources like solar and wind. However, the efficiency of wave energy systems depends heavily on accurately predicting ocean wave characteristics such as Significant Wave Height (SWH) and Mean Wave Period (MWP). These parameters influence the power output and operational planning of tidal power systems.

The variability in wave conditions presents a challenge in maintaining stable power generation, making accurate forecasts essential for efficient integration into power grids. Traditional forecasting models, while effective to some extent, often struggle with capturing the complex temporal dependencies and uncertainties in oceanographic data. This project aims to address these challenges by developing and comparing time series forecasting models using machine learning techniques and quantile regression. The integration of these methods is intended to not only improve predictive accuracy but also provide insights into the range and uncertainty of forecasts, thus supporting better decision-making in wave power generation.

**1. Aim**

This research aims to develop and compare models for forecasting SWH and MWP using historical oceanographic data. A combination of machine learning methods and quantile regression will be applied to provide point and interval estimates of wave conditions. The goal is to assess the effectiveness of each approach in forecasting tidal power generation, contributing to improved planning and optimization of wave energy systems.

**2. Research Questions**

The project seeks to address the following research questions:

1. How accurately can we predict SWH and MWP using time series data and machine learning models?
2. How does quantile regression enhance the understanding of uncertainty in SWH and MWP forecasts?
3. What oceanographic variables significantly impact day-ahead forecasts of SWH and MWP?
4. Can the combined approach improve the practical application of wave forecasts in tidal power generation?

**3. Objectives**

1. Develop and compare predictive models for SWH and MWP using time series data:
   1. Utilize variables like wave velocity, SST, wind speed, and tidal height.
   2. Compare the performance of traditional time series models with machine learning methods.
2. Implement quantile regression to provide prediction intervals for SWH and MWP:
   1. Focus on key quantiles (e.g., 10th, 50th, 90th percentiles).
3. Analyze model performance using metrics like RMSE, MAE, and interval coverage:
   1. Assess the accuracy and reliability of forecasts for practical energy planning.
4. Provide insights into optimizing wave energy generation based on predictive models.

### **Project Plan**

**1. Project Timeline**

| **Week** | **Task Description** | **Specific Tasks** |
| --- | --- | --- |
| **0** | **Literature Review** | Conduct background studies and refer similar literature |
| **1** | **Data Collection** | Gather historical oceanographic data for SWH and MWP, including relevant features; ensure data is cleaned and properly formatted for analysis. |
| **2** | **Data Preprocessing** | Handle missing values and outliers in the dataset; normalize or standardize the feature variables as needed. |
| **3** | **Exploratory Data Analysis (EDA)** | Conduct descriptive statistics to summarize key variables; visualize data distributions, trends, and correlations using plots. |
| **4** | **Feature Engineering for SWH** | Create lag features and rolling statistics for SWH; introduce interaction features and cyclic transformations for time-related variables. |
| **5** | **Model Selection for SWH** | Implement classical time series models (ARIMAX, SARIMAX) for baseline comparisons; begin implementing machine learning models (Random Forest, XGBoost and Quantile regression) for SWH forecasting. |
| **6** | **Model Development for SWH** | Train and evaluate selected machine learning models; perform hyperparameter tuning to optimize model performance. |
| **7** | **Performance Evaluation for SWH** | Evaluate SWH models using RMSE, MAE, and correlation metrics; conduct time series cross-validation to assess model generalization. |
| **8** | **Feature Engineering for MWP** | Repeat feature engineering steps for MWP similar to SWH; ensure to create lag and interaction features relevant to MWP. |
| **9** | **Model Selection for MWP** | Implement baseline models (ARIMA, SARIMA) for MWP; begin developing machine learning and deep learning models (LSTM, GRU) for MWP forecasting. |
| **10** | **Model Development for MWP** | Train selected models for MWP and perform hyperparameter tuning; validate models using appropriate metrics for MWP. |
| **11** | **Comparative Analysis of SWH and MWP Models** | Compare the performance of SWH and MWP models using visualizations and summary statistics; prepare a report on findings and insights from both forecasting models. |
| **12** | **Final Review and Documentation** | Compile and finalize project documentation, including methodology, results, and conclusions; prepare presentation slides summarizing key findings for stakeholder review. |

### **Data Management Plan**

**1. Data Collection and Storage**

**Data Sources:** The dataset consists of two years of hourly data, including SWH, MWP, and other oceanographic variables.

Website Link: https://cds.climate.copernicus.eu

**Data Storage:** All data will be stored securely on a cloud platform (e.g.,Google Cloud) with restricted access to ensure data integrity and confidentiality.

**Data Format:** Data will be stored in CSV and Parquet formats for easy processing and analysis.

**Version Control:** Version control for code and model iterations using Git ensures traceability and reproducibility of results.

2**. Ethical Considerations**

**Transparency:** Clearly communicate the limitations and uncertainties associated with model predictions.

**Environmental Impact:** Consider how forecasting can improve the environmental sustainability of wave energy systems.

**Fairness:** Address potential biases in model predictions to ensure accuracy across different time periods and regions.

**3. Resources Needed**

* **Computing Resources:** High-performance computing for training deep learning models.
* **Software:** Python with libraries like TensorFlow, Keras, scikit-learn, and statsmodels.

### **Conclusion**

This PDM plan outlines a comprehensive approach to forecasting Significant Wave Height (SWH) and Mean Wave Period (MWP) using multivariate time series analysis. By developing robust predictive models, this project aims to significantly improve the management of wave energy systems and enhance the integration of wave power into the energy grid. The successful implementation of this project has the potential to contribute to sustainable energy solutions and advance the field of ocean engineering.

### **References**

1. **Holthuijsen, L. H. (2007). *Waves in Oceanic and Coastal Waters*.** Cambridge University Press.
2. **Meinshausen, N. (2006). *Quantile regression forests*.** *Journal of Machine Learning Research*, 7, 983-999.