

Modeling and Simulation

**Student Id:** 21110214.

**Submitted to:** Dr Murad Yaghi.

**Submitted by:** Saif Kamal Salim Haddad.

**Submitted on:** 29-1-2025.

Table of Contents

[Task 1: 3](#_Toc187437172)

[**Q1: Introduction -1:** 3](#_Toc187437173)

[**Q2: Introduction -2:** 4](#_Toc187437174)

[**Q3: Work done to solve ocean wave height:** 5](#_Toc187437175)

[**Q4: Workflow for Solving the Ocean Wave Height Prediction Problem Using Modeling and Simulation:** 8](#_Toc187437176)

[Task 2: 10](#_Toc187437177)

[**Q1:** **Describing and Analyzing the Data** 10](#_Toc187437178)

[**Q3: evaluate the accuracy of the standard model:** 14](#_Toc187437179)

[**Q4: Set of Mathematical Equations Describing the Ocean Wave Model:** 15](#_Toc187437180)

[**Q5: Theoretical Principles of Each Parameter in the Wave Height Model:** 19](#_Toc187437181)

[Task 3: 26](#_Toc187437182)

[**Q1: Different optimization techniques to tune the simulation parameter:** 26](#_Toc187437183)

[**Q3: Comparison and Analysis of Standard Modeling Techniques Using Optimization Methods:** 26](#_Toc187437184)

[Task 4: 28](#_Toc187437185)

[**Q1: Workflow of different optimization techniques:** 28](#_Toc187437186)

[References: 34](#_Toc187437187)

[**Task 1:** 34](#_Toc187437188)

# **Task 1:**

## **Q1: Introduction -1:**

**Problem:**

Ocean waves are basically known as disturbances in the ocean that affect maritime navigation adversely because of the endangering of safety and efficiency. Because of many factors such as wind speed, air temperature, and sea surface temperature, the wave height may change suddenly and thus threaten the stability of ships, cause delays in operations, and increase fuel consumption. Wave height estimation is hence important to mitigate these threats.

**The Need for Modeling and Simulation:**

Modeling and simulation techniques help in the study and forecasting of ocean waves. The generation and transportation of waves in different conditions can be visualized on computers. The computers use such models for generating a very important prediction of parameters related to average wave height, maximum wave height, and significant wave height. Knowledge gathered through these models is of utmost help in:

Identification of and avoidance of locations which, under wave conditions, are hazardous: in other words, optimization of routing. Routing is basically the route which, while being plotted out for safety, also has considerations for efficiency related to fuel consumption. The real-time decision actually updates the forecast continuously with respect to any changeable condition of the ocean.

**Understanding the Dynamics of Waves:**

Wind speed, air and sea surface temperatures, and the average wave period are some of the variables that affect ocean waves. They can show extremely erratic patterns or periodic waves. For example :

Air temperature affects the production of wind and, therefore, atmospheric pressure. The sea surface temperature determines wave energy and the strength of the ocean currents. Wave height and wave formation are directly related to the wind speed.

**Modeling and Dynamics Applications**

The modeling and simulation efforts may focus on the following using the dataset that has been provided:

Creating models to forecast wave metrics including average, maximum, and important wave heights is known as wave height prediction.

Impact analysis is the process of evaluating how environmental factors, such as temperature and wind speed, affect wave dynamics.

Applications in Real Time: Putting dynamic models into practice for wave forecasting and decision-making in real time.

## **Q2: Introduction -2:**

**Importance and Benefits:**

**Wave Forecasting:**

Using the modeling part gives us the ability to have accurate predictions of some of the main wave parameters like wave height, the period of the wave, and the direction of the wave, so when we have an accurate prediction of these parameters, we can have a safe navigation and route optimization which will minimize the risk. (incois.gov, 2023)

**Improve the Safety and Minimize the Risk:**

This means providing high safety for the ships and also for the people that are on the ship; this can be achieved using modeling and simulation models, which will guarantee to minimize the risk of accidents and also enable the ship captain to be ready for negative advance conditions and situations all this just by training the simulation on real scenarios and then simulate any hard or rare conditions that might happen in the oceans which will maximize the safety of the ship and the people on the ship by avoiding hazardous areas. (Pinto et al., 2023) (incois.gov, 2023)

**Minimize the Cost:**

Modeling and Simulating the Ocean Wave Height is very helpful in minimizing the cost since the logic says that the ship needs to consume more fuel when the waves are high, so in case the model gives results that the waves will be high, the captain will optimize the route to another way that the waves are low to minimize the fuel consumption, which this will lead to cost saving and also will protect the environment because we will have lower emissions. (incois.gov, 2023)

## **Q3: Work done to solve ocean wave height:**

**Paper One:**

"Prediction of Significant Wave Height Using Machine Learning and Its Applications" (Afzal et al., 2023)

**Overview:** The study in this paper was made to talk about the challenge of predicting significant wave height (SWH) with also their return periods, this has been done using machine learning techniques while using a generalized extreme value (GEV) theory.

**Techniques and Algorithms:** The researchers that have worked on this paper have used algorithms like linear regression (LR), artificial neural networks (ANN), and support vector machines (SVM), the performance measure that they used was R2

**Results:** Based on the comparison between the three models’ performance, the results were that the SVM is the best among all the other models, the SVM model has achieved an accuracy of 99.80% in predicting significant wave height.

**Paper Two:**

"Using Machine Learning Techniques to Predict Significant Wave Height Compared with Parametric Methods"(Salah & Elbessa, 2024)

**Overview:** Since the SVM model is one of the best models that suit this case and is also easy to use with low computational power, the study in this paper was made to see and evaluate what the best kernel function of SVM model in predicting significant wave height and also and compares their performance with commonly used parametric models.

**Techniques and Algorithms:** The researchers have used the Support Vector Machine (SVM) Model using different types of kernel functions, which were Linear Kernel, Sigmoid Kernel, Radial Basis Function (RBF) Kernel, and Polynomial Kernel, on the other hand Parametric Models which were Pierson Moskowitz (P-M) Model, Shore Protection Manual (SPM) Model and Coastal Engineering Manual (CEM) Model

**Results:** The SPM and CEM parametric models were more accurate than the SVM models with Linear and Sigmoid kernels, but overall performance, the SVM models with RBF and Polynomial kernels were the most accurate overall, since this model had the best generalization errors between all models evaluated.

**Paper Three:**

"Significant Wave Height Prediction Based on Wavelet Graph Neural Network"

(Delong Chen, 2021)

**Overview:** This study suggests using a Wavelet Graph Neural Network (WGNN) to handle both short-term and long-term relationships efficiently since they want to see if the WGNN will improve predictive performance and will be more accurate.

**Techniques and Algorithms:** The researchers have used multiple things, the first thing was Wavelet Transform, which was used to break down the original SWH time series data into components at multiple frequency levels, The second thing was Graph Neural Network (GNN), which was used because of the complex spatial-temporal dynamics found in ocean wave data can be captured by GNNs because of their ability to represent complicated linkages and dependencies.

**Results:** The proposed method, WGNN, has been compared against a range of deep learning models, traditional numerical models, and some general machine learning techniques. Experimental results demonstrate that WGNN is far better than those models with regards to yielding higher accuracy for SWH predictions. Improved performance underlines the level at which graph neural networks and wavelet decomposition collaborate to capture the complex nature of ocean waves.

**Paper Four:**

"Significant Wave Height Prediction Based on VMD-SA-MLP-BP"

(Yining Wu, 2023)

**Overview:**. Using the traditional models could be so hard in this situation since these models are non-stationary and complex, but this paper will combine VMD, SA, and MLP-BP into an innovative approach to improve accuracy.

**Techniques and Algorithms:** V**ariational Mode Decomposition** VMD is employed to decompose the SWH time series data into multiple intrinsic mode functions (IMFs), also Simulated Annealing SA was used because it is a probabilistic optimization algorithm that has been used to tune the weights and biases of the MLP-BP neural network in an effort to find the best parameter configuration for better performance and generalization of the model. Also, The MLP-BP neural network serves as the predictive model, processing the decomposed sub-modes (IMFs) to forecast SWH. The backpropagation algorithm adjusts the network's weights to minimize prediction errors.

**Results:** The proposed VMD-SA-MLP-BP model was compared with other models, including SSA-MLP-BP, PSO-MLP-BP, and SA-MLP-BP.

These results indicate that the VMD-SA-MLP-BP model outperformed the other models, demonstrating superior accuracy in predicting significant wave heights.

**Paper Five:**

"A Significant Wave Height Data-Driven Modeling for Digital Twins of Marine Environment" (Arman Neyestani, 2024)

**Overview:** The authors in this paper try to provide an overview of the development and implementation of a Predictive Modeling Framework for the forecasting of Significant Wave Height (SWH) (VHMO) concerning marine environments, taking into consideration all the benefits of using Digital Twins (BTs). This model gives the average height of one-third of the highest waves recorded and is considered a key indicator for maritime operations and safety.

**Techniques and Algorithms:** In the framework of the research, the authors applied the neural network GRU to model and predict (SWH) (VHMO). The model was trained and tested with three different wave heights from in situ sensors from Tarragona, Barcelona, and the observatory EMSO-OBSEA in the western Mediterranean. The methodology involved a heavy pre-processing of data by normalizing and creating sequences, later training and testing a model capable of making effective predictions of significant wave heights. Results from the residual analysis show that there is a good amount of outliers behind the threshold criteria established from the residuals' statistical distribution.

**Results:** The Pearson's correlation coefficients are 0.93, 0.95, and 0.88 for the buoys Tarragona, Barcelona, and OBSEA, respectively, indicating a very good accuracy and robustness with respect to data gaps. These results prove that the model is powerful enough concerning predicted accuracy to monitor and perform better decision-making in maritime operations.

**Paper Six:**

"A Simulation Study of Significant Wave Height Retrieval from Bistatic Scattering of Signals of Opportunity" (Paul A. Hwang, 2022)

**Overview:** In this paper, approaches are developed based on simulations to determine Significant Wave Height from satellite signal microwave scattering. It has been demonstrated that the width of the Doppler frequency spectrum can serve as an effective metric in estimating SWH.

**Techniques and Algorithms:** Simulation-based analysis of Doppler frequency spectrum width to correlate with SWH.

The investigation is about the relationship between SWH and Doppler frequency spectrum width. To facilitate the evaluation of the change in the spectrum width about the variation of the wave height, the study of the Doppler spectrum generated by microwave signals scattered off the ocean surface was carried out. The presented approach is tested for a wide range of wind speeds, 5-21 m/s; inverse wave ages, 0.8-2.0; and ocean depths, from 3 m to ∞.

**Results:** Simulations show that, for all considered water depths, wave ages, and wind speeds, the width of the Doppler frequency spectrum is related to the SWH in a rather robust way. This result opens perspectives for remote sensing applications in oceanography since it implies that the width of the Doppler spectrum can serve as a reliable indicator for the retrieval of the SWH under various oceanic conditions.

## **Q4: Workflow for Solving the Ocean Wave Height Prediction Problem Using Modeling and Simulation:**

**Step 1: Define the Problem:**

Predict and simulate ocean wave heights based on wind speed using a discretized wave height model. The goal is to optimize model parameters (𝛼,β) to achieve the best predictive accuracy.

**Step 2: Literature Review and Related Work:**

We reviewed past work on wave height prediction to understand existing approaches and identify methods that could enhance our model.

**Step 3: Model Equation (Differential Equations)**

**-Wave Growth Equation:**

A black text with a white background

Description automatically generated

This equation describes how the wave height (H) changes continuously over time (t).

-𝑈(𝑡) is the wind speed or wind energy applied to the wave system at time 𝑡.

-𝐻(𝑡) is the wave height at time 𝑡.

-𝛼 (Growth Coefficient) represents the rate at which wind energy contributes to increasing wave height.

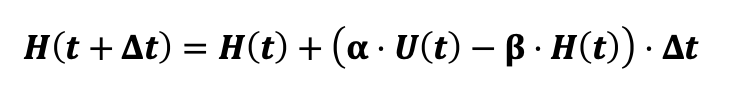
-𝛽 (Dissipation Coefficient) represents the rate at which wave energy is lost.

**-Growth Term:**

(𝛼 ⋅ 𝑈 𝑡 ⋅ 𝛥𝑡): Represents the increase in wave height due to wind energy input.

**-Dissipation Term:**

(𝛽 ⋅ 𝐻 𝑡 ⋅ 𝛥𝑡): Represents the decrease in wave height due to energy losses.

**Discretizing the Equation**: Using Euler's method:

**Step 4: Parameter Initialization:**

**Parameters:**

Growth Coefficient (α) = 0.5 from my mind.

Dissipation Coefficient (β) = 0.5 from my mind.

**Initial Conditions:**

Initial wave height (𝐻0) = 0.75 which is the first observed wave height in the dataset.

**Simulation Settings:**

Time step (Δt): 1 hour.

Total simulation time: 11271.0 hours, 11271.0/24=469.625 days (In a clear way, it’s = 1 year and 4 months).

**Step 5: Model Simulation:**

I will simulate wave heights using the discretized equation, and then I will validate the predictions by comparing them with the actual wave heights using Mean Squared Error (MSE), Mean Absolute Error (MAE), and also by calculating the Residuals, and then I will Visualization them for a clear Comparison.

**Step 6: Optimization of Parameters:**

I will optimize the parameters, which are Alfa and Beta, using optimization methods such as the Simulated Annealing, which will calculate the MSE for each iteration to tell if it gets the best parameter values also I will use the Genetic Algorithmoptimizer  
  
**Step 7: Validation and Visualization:**

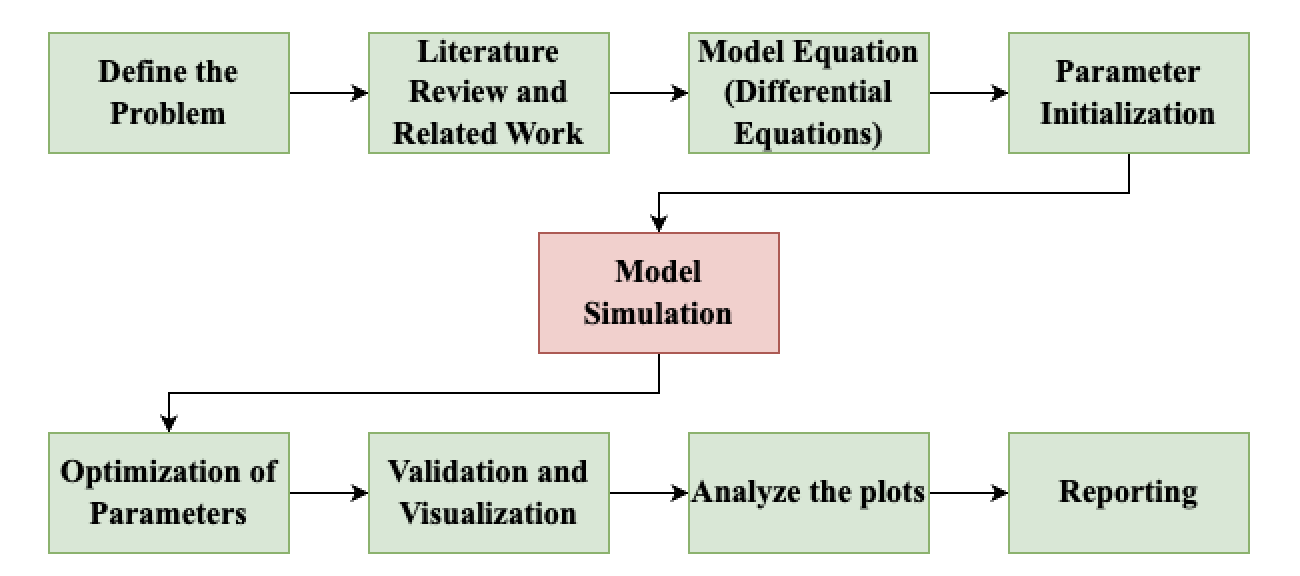
I will use the results of the best parameters to run the model and then do some visualizations such as plot observed vs. predicted wave heights and display residual distributions and cost function trends.

**Step 8: Analyze the plots:**

Get insights from the plots and the visualization that I have done

**Step 9: Reporting:**

Write about the insights and discuss the role of wind speed and its affection on wave dynamics, and discuss the importance of parameter optimization for accuracy improvement.



# **Task 2:**

## **Q1:** **Describing and Analyzing the Data**

**Columns description:**

**Time**: Timestamp indicating the date and time of the measurement.

**Air Temperature (deg\_C)**: Air temperature measured in degrees Celsius at the time of recording.

**Sea Surface Temperature (deg\_C)**: Temperature of the sea surface in degrees Celsius at the time of recording.

**Relative Wind Speed (m s-1)**: Wind speed relative to a reference point, measured in meters per second.

**year\_month**: Year and month of the data.

**Wave Height Average (m)**: Average wave height during the specified time interval, measured in meters.

**Wave Height Maximum (m)**: Maximum wave height recorded during the specified time interval, measured in meters.

**Wave Height Significant (m)**: Significant wave height, a statistical measure typically representing the mean of the highest one-third of wave heights, measured in meters.

**Wave Period Average (sec)**: Average wave period, indicating the average time interval between consecutive wave crests, measured in seconds.

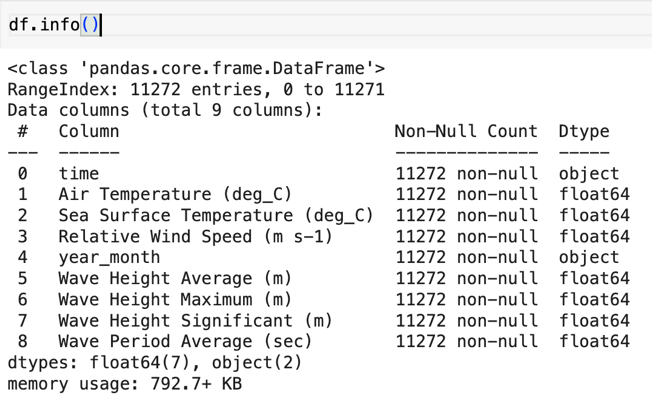
**EDA:**

**-**The shape of the CSV file is nine columns with 11272 rows.

A close-up of a number

Description automatically generated

-Also, we can see that I have checked the columns data type by using (info()), and I found that all the columns are (float) except the time and the year\_month columns are objects because they are saved in the column as a date format.

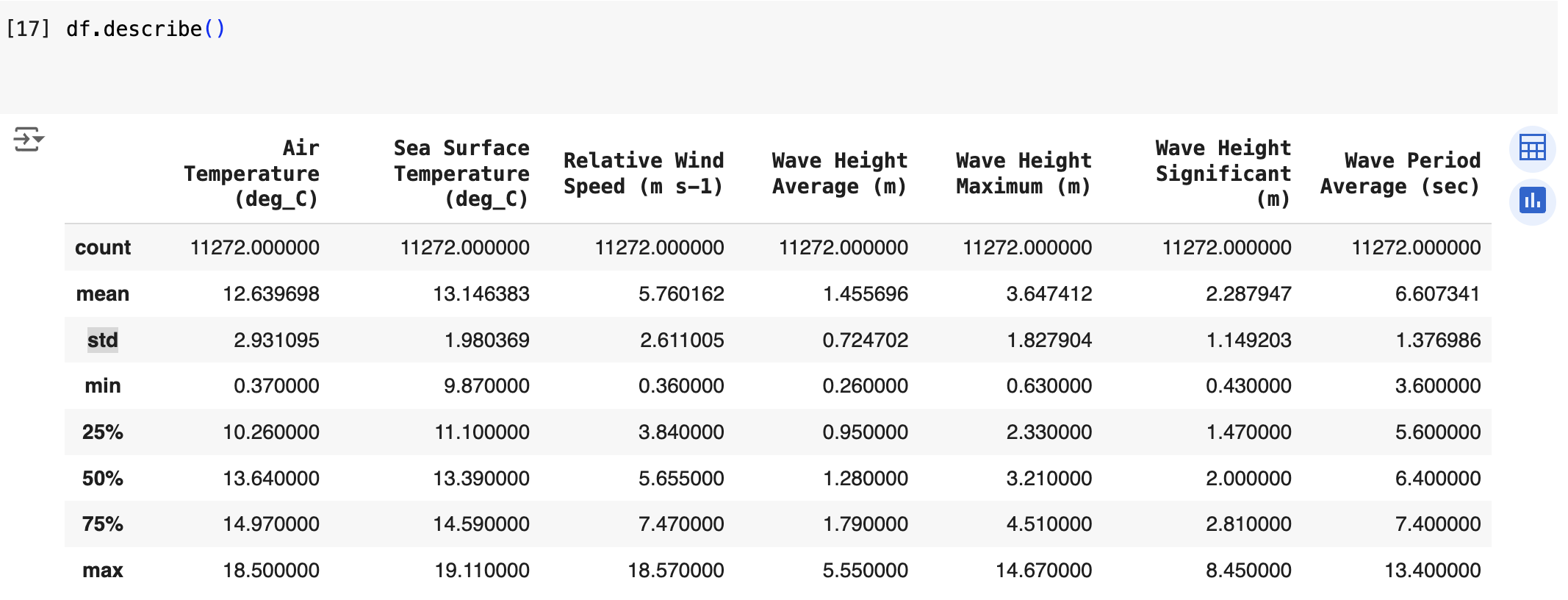


-I have checked if there are null values in the data set but the results show that there are no null values in the data set

A screenshot of a weather forecast

Description automatically generated

-I have used (.describe) to see each column’s specifications.



Based on the count all columns have 11,272 values, meaning no missing data).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Air Temperature (deg\_C)** | **Sea Surface Temperature (deg\_C)** | **Relative Wind Speed (m s-1)** | **Wave Height Average (m)** | **Wave Height Maximum (m)** | **Wave Height Significant (m)** | **Wave Period Average (sec)** |
| **Mean** | 12.64°C | 13.15°C | 5.76 m/s | 1.46 m | 3.65 m | 2.29 m | 6.61 |
| **Range** | From 0.37°C to 18.5°C | From 9.87°C to 19.11°C | From(0.36 m/s) to (18.57 m/s). | From 0.26 m to 5.55 m | From 0.63 m to 14.67 m | From 0.43 m to 8.45 m. | From 3.6 to13.4 seconds. |
| **Quartiles**  (Half the data is between) | 10.26°C and 14.97°C | 11.1°C and 14.59°C | 3.84 m/s and 7.47 m/s | 0.95 m and 1.79 m | 2.33 m and 4.51 m | 1.47 m and 2.81 m. | 5.6 seconds and 7.4 seconds. |

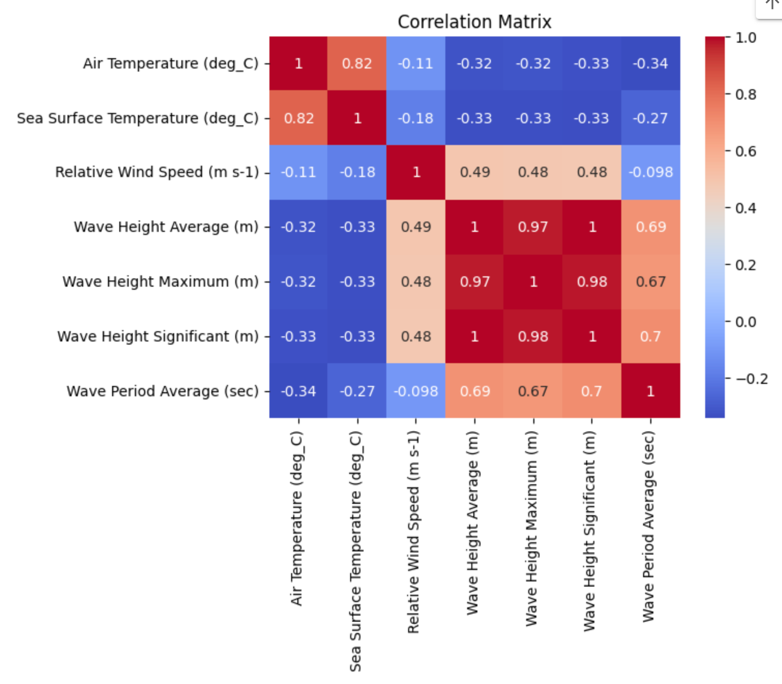
**Main Finding:**

The air temperature and sea surface temperature are close; this shows a relationship between them.

Average wind speeds are moderate (mean: 5.76 m/s), but extreme events (up to 18.57 m/s) occur occasionally.

Average waves are around 1.46 m, but significant wave heights can reach up to 8.45 m, with maximums as high as 14.67 m, which is a big problem that affects safety.

-I have checked the correlation between all the numeric columns.



**Main Findings:**

Strong interdependence among wave height metrics and their relationship with wave periods

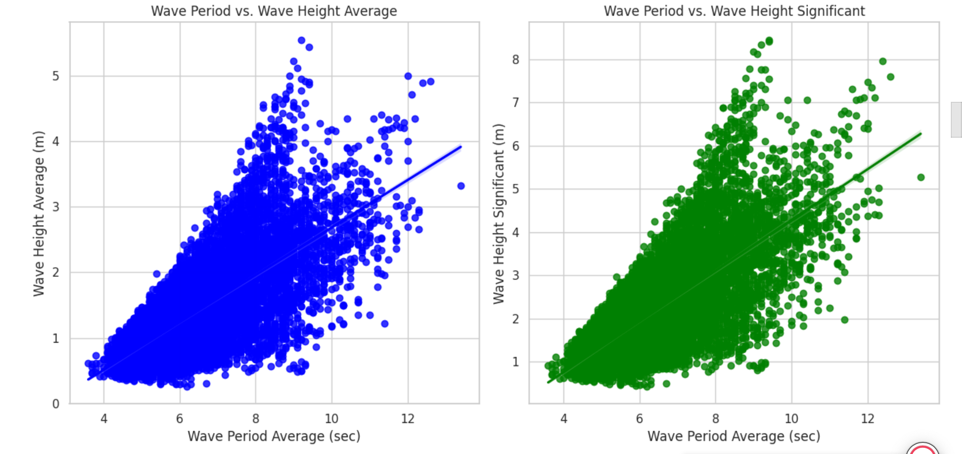
Wind speed influences wave heights moderately but has minimal impact on wave periods.

Wave heights are strongly related to all three columns (average, maximum, and significant), this shows that they track each other.

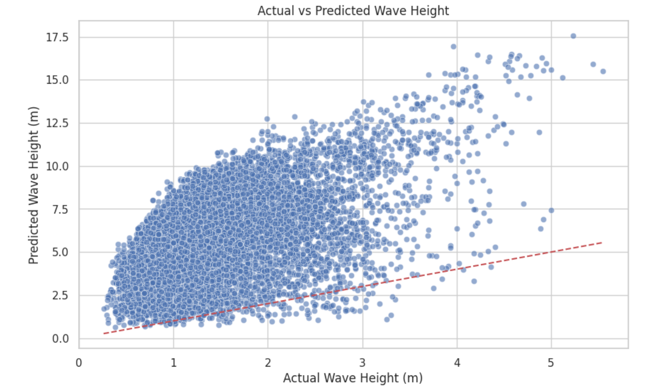
The air temp have a strong relationship with the sea surface temp.

Wave period has a moderate relationship with wave heights; this shows larger waves tend to have longer periods.

We can prove the last finding with this plot.



## **Q3: evaluate the accuracy of the standard model:**



**Actual vs. Predicted Wave Height (Scatter Plot):**

In this plot, we can see a big difference between the

models predictions and the actual data, we can see

this is because the points are too far from the red

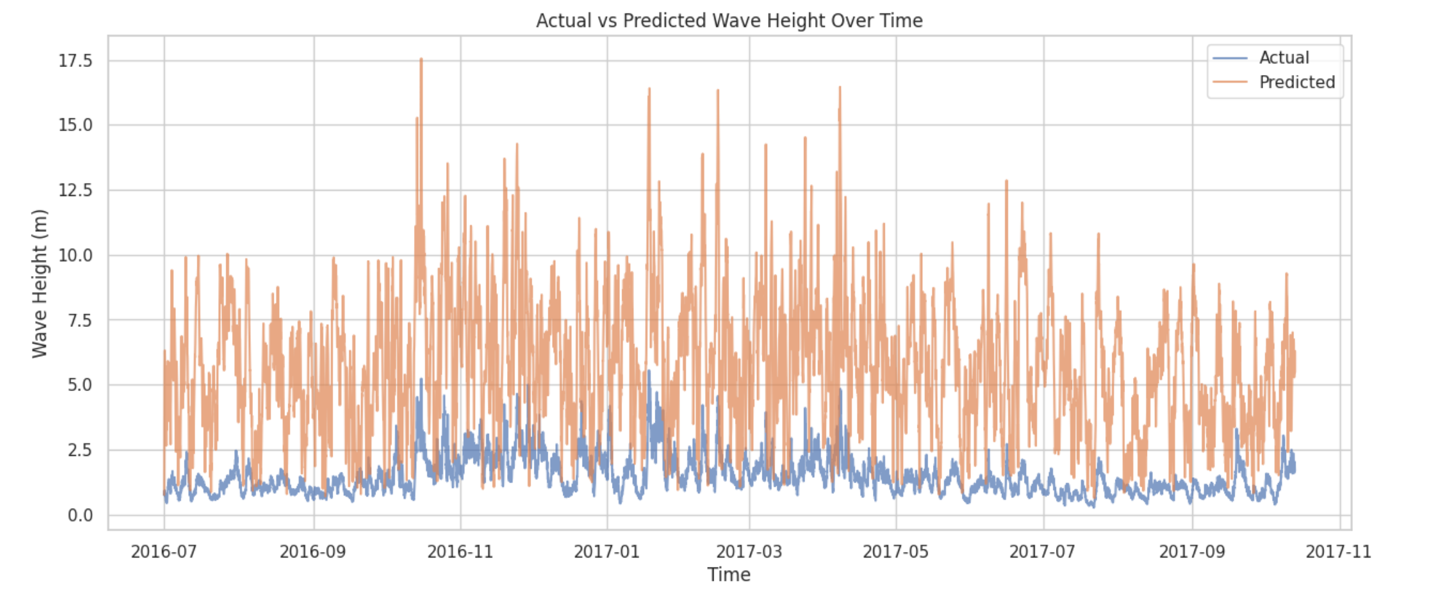
line, especially for higher values, which shows that

the model has overestimated the prediction.

**Actual vs. Predicted Wave Height Over Time (Time Series Plot):**

In this plot, we can see a big difference between the

models predictions (Orange) and the actual data (Blue), which shows an overestimation of the actual wave heights, which we can see from the big difference between the two lines (Orange and Blue).



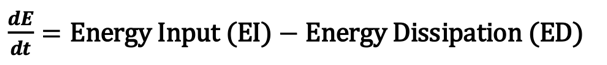
**Overall Evaluation:**

The standard model has performed very weakly, and the accuracy was very limited in predicting the wave heights. Also, all the plots show that the model has overestimated the values, especially the big values. This shows that the standard model requires a lot of improvements to the model and the model parameters

## **Q4: Set of Mathematical Equations Describing the Ocean Wave Model:**

The equations of the ocean wave height model are derived using the energy balance principle. These equations describe the dynamics of the increase of wave height in time and, in fact, explain how wind energy interacts with ocean waves. Now we discuss the equation in detail in parts:

**1-The Energy Balance Equation:**



In this equation, E = total energy of the wave system, and (dt/dE) defines the rate of change of energy with time. Energy input EI gives the wave energy to expand and is determined mainly by the wind speed, on the other hand, energy dissipation ED is the loss of energy that causes wave breaking, friction, and other dissipative processes.

**2. Simplified Wave Height Equation:**

In this model, it gives a simple and clear equation of the energy balance equation to describe the wave height (H) as a function of time:

A black text on a white background

Description automatically generated

**This equation has two parts:** The two parts together give the rate of change of wave height (dH/dt).

**Wave Growth:**

This represents the contribution of wind energy to wave height growth.

**α:** Growth coefficient, indicating how effectively wind energy is transferred to the waves.

**U (t):** Wind speed at time 𝑡. Faster winds cause more energy, which leads to higher wave growth rates.

**Wave Dissipation:**

A black letter h on a white background

Description automatically generated

This represents energy loss from the wave system.

**β:** Dissipation coefficient, indicating how quickly energy is lost due to factors like wave breaking or friction, which will represent the decrease in wave height

**H (t):** Wave height at time t. Larger waves lose energy more rapidly.

**3. Discretized Form:**



**The main variables in this equation:**

**H (t):** The wave height at the current time step.

**𝐻𝑡+Δ𝑡:** The wave height at the next time step.

**Δt:** The time step size

In this equation, The growth term and dissipation term are added to the current wave height (Ht), so this equation allows for simulating wave dynamics over time.

The equation explains how wave height is decreased or increased by dissipative processes and wave growth by wind energy. The wave system's behavior is affected by the balance between these forces:

-Wave height rises as growth takes over.



>

-Wave height stabilizes as the system reaches balance.



 =

-Wave height decreases as Dissipation takes over.



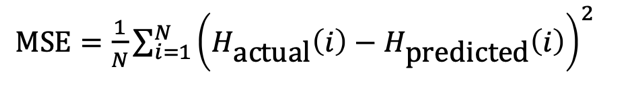
<

**The coefficients Alfa and Beta:**

-α determines how efficiently wind speed contributes to wave growth. It depends on any environmental things that might increase the growth.

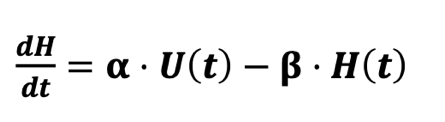
-β controls the rate of energy dissipation. It depends on any environmental things that might decrease the growth.

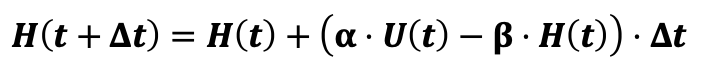
**4. Mean Squared Error (MSE) equation:**



Using the MSE, we can do coefficient calibration, the smaller result of MSE means the better coefficient values

**More details about the equation:**

 The equation was converted from continuous form to discretized form.



**Why Not Use the Continuous Equation Directly in Simulations?**

The equation is an ordinary differential equation (ODE).

Solving ODEs analytically involves integrating the equation, which is only possible for simple cases.

real-world data, like Ut (time-varying wind speed), makes it impossible to find a closed-form solution.

Ut (wind speed) changes over time, and it’s provided as discrete data points.

Continuous equations cannot directly handle such time-dependent, while the discretized form accommodates these changes easily.

Computers operate in discrete time steps. They cannot simulate continuous changes, so the equation must be transformed into a form that works iteratively (step-by-step).

The discretized form calculates H iteratively at each time step, this matches the time intervals of your dataset.

wind speed (Ut) might be recorded hourly, so we only have data for specific time points.

Discretized equations are computationally simple and efficient for long simulations with thousands of time steps.

**What is an Ordinary Differential Equation?**

A function (H(t)), which describes wave height as it changes over time (t).

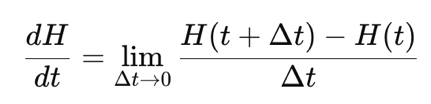
A derivative (dH/dt), which represents the rate at which the wave height is changing at any given time.

**How did I solve the Ordinary Differential Equation problem?**

I used Euler’s Formula, which is is a numerical technique used to solve ordinary differential equations (ODEs) step by step. It offers an iterative way to simulate a system's behavior, such as the wave height equation in your situation.

This formula solved the problem we have, which is the wind speed data recorded at discrete time intervals (hourly), not continuously.

**How did I apply the Euler’s formula?**



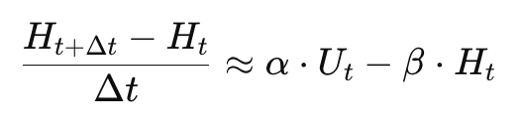
-Comes from the definition of the derivative in calculus.

-This is the foundational concept for describing how a quantity (in this case, wave height H) changes over time (t).

A black text on a white background

Description automatically generated

I put the dH/dt in the first equation, tell it gets in this form.



The last form I get into:

A black text with a plus and u

Description automatically generated

**These steps allowed me to:**

-Simulate the wave height iteratively for each time step (Δt).

-Incorporate real-world wind speed data (Ut).

-Predict wave height (Ht) over the entire duration of my dataset.

## **Q5: Theoretical Principles of Each Parameter in the Wave Height Model:**

A black text with a plus and u

Description automatically generated

This equation is a discretized form of the wave height model, which was derived using Euler’s formula

**Parameter by Parameter:**

**1-Current Wave Height (H(t)):**

Represents the wave height at the current time step t, this parameter is considered as the first main parameter in this equation next step (t + Δt)

**Reminding note:**

**Β:** is the dissipation coefficient, which controls how quickly wave energy is lost, It acts to reduce the wave height.

**𝐻(𝑡):** Is the wave height at the current moment.

**How does H(t) Affect Dissipation?**

If H(t) is large (the wave is tall), the dissipation term (−β⋅H(t)) becomes more negative.

A larger dissipation term (−β⋅H(t)) minimizes more from the wave height growth, reducing the ability of the wave height to increase further.

**Main insights:**

Tall waves (H(t) is large) experience stronger dissipation, which

slows their growth in the next time step (H(t+Δt).

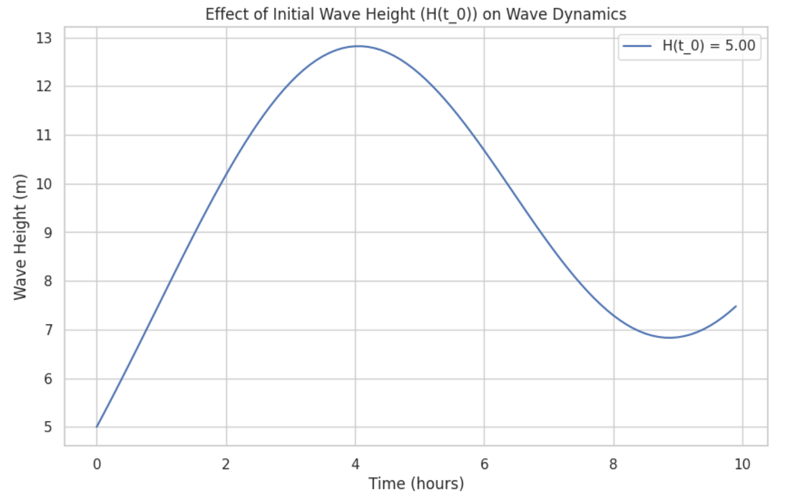
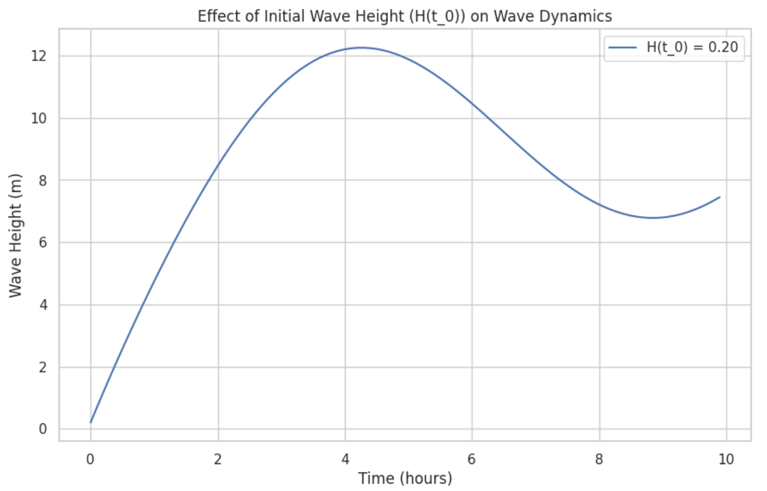
This is why the system naturally balances itself: as waves grow taller, dissipation increases, preventing runaway growth.

When H(t) is small, dissipation is weak, so waves grow faster

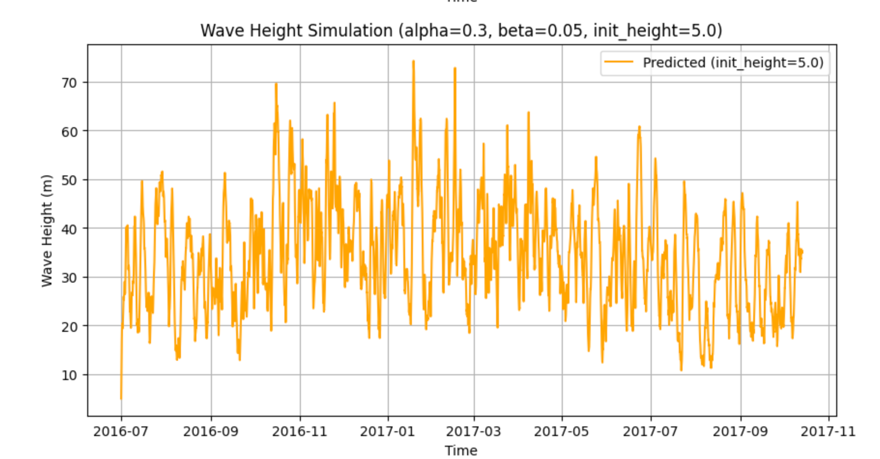
A higher H(t) increases the dissipation term (−β⋅H(t)), which acts to counterbalance the growth term (α⋅U(t)). This mechanism prevents wave height from growing uncontrollably and ensures a natural equilibrium in the system.

If waves grow too tall, dissipation increases and slows their growth.

If waves are small, dissipation is weak, allowing them to grow more easily.

****





**H(t0)=0.20:**

-The wave starts from a small height of 0.2 meters.

-Since the initial wave height (H(t0)) is small, dissipation (−β⋅H(t)) is

weak at the beginning.

-This allows the growth term (α⋅U(t)) to dominate, resulting in faster wave growth.

**H(t0)=5.00:**

-The wave starts from a significant height of 5 meters.

-It grows more slowly compared to the first case and reaches a peak of around 13 meters before dissipation reduces the height.

**2. Growth Coefficient (α):**

This parameter represents the efficiency of energy transfer from the wind (U(t)) to the wave height (H).



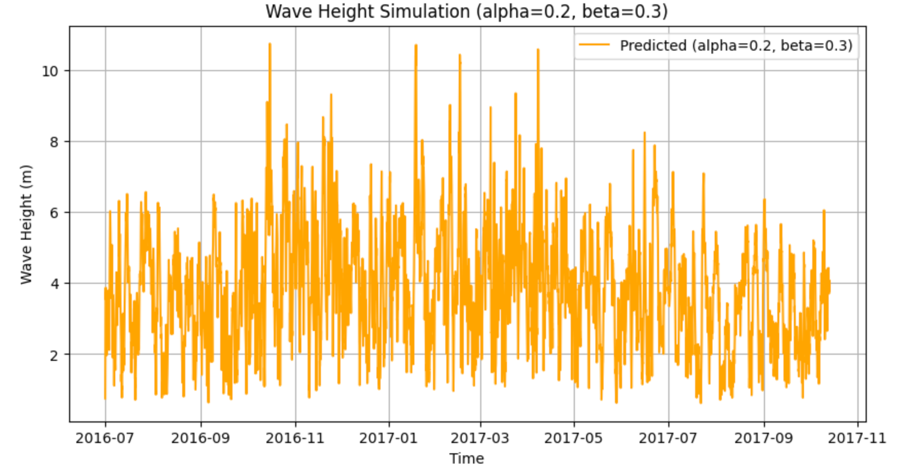
is the growth term, driving wave height increase.

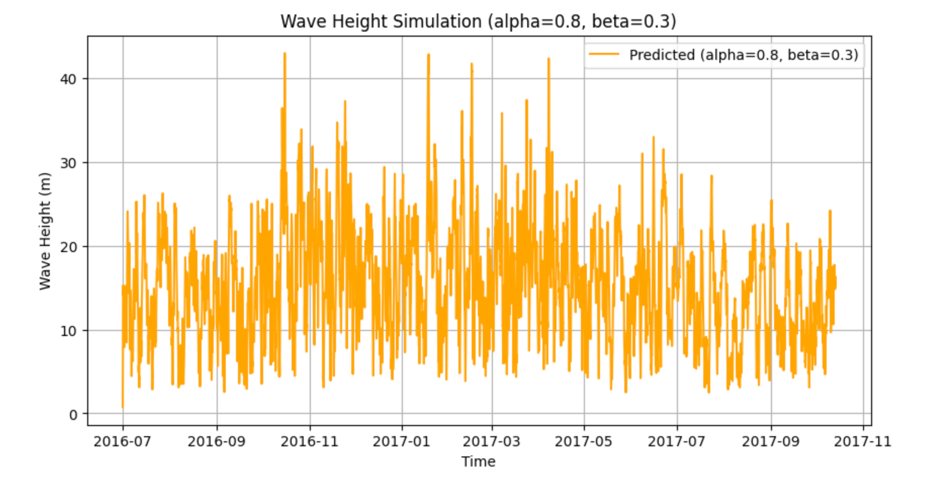
When this coefficient is large, it will increase the energy transfer from wind to waves, resulting in faster wave growth.

A larger growth coefficient (α) allows more energy transfer from the wind, resulting in taller waves for the same wind speed.

When α increases, the system reacts more strongly to changes in wind speed (U t), causing waves to grow or shrink faster.

A low coefficient value slows the rate of wave growth.





Alfa= 0.20: The wave height grows moderately and reaches a peak of around 7 meters.

Alfa= 0.80: The wave height grows much faster, reaching a higher peak of around 20 meters.

The difference in wave height peaks between the two plots directly reflects the change in α.

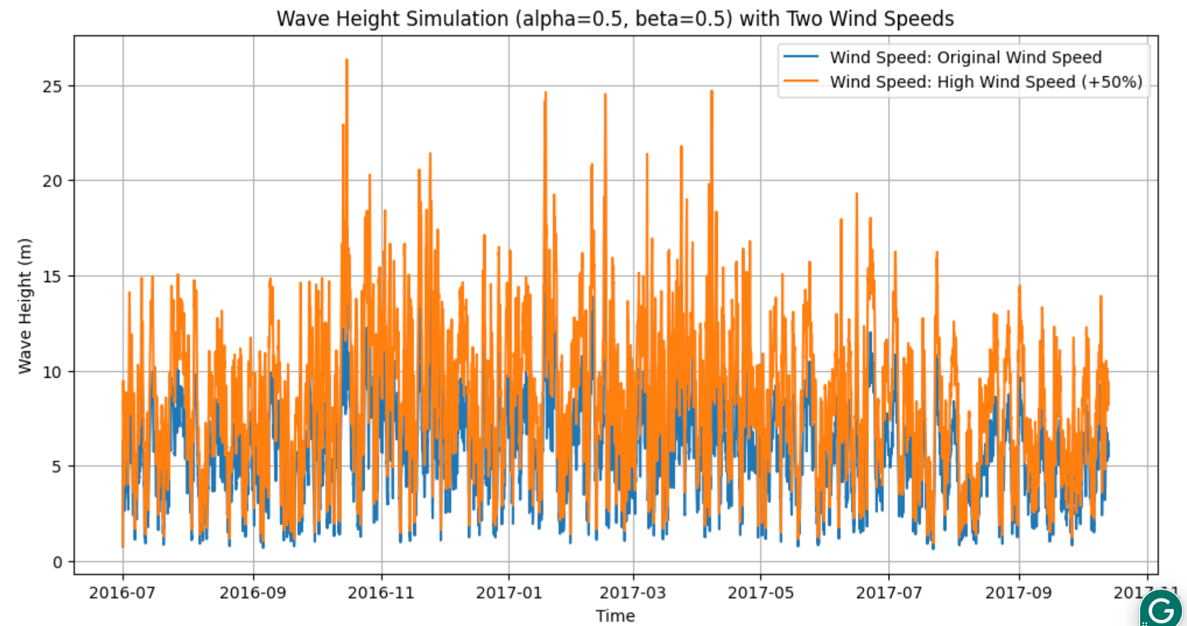
When α increases, the system reacts more strongly to wind energy, resulting in taller waves.

**3. Wind Speed (U(t)):**

Wind speed (Ut) is one of the main parameters that driving force that supplies energy to the system, causing the wave height to grow.

High wind speed The growth term (α⋅Ut) becomes larger, resulting in faster wave growth.

Low wind speed makes the growth term weakens, and dissipation is in charge, causing wave height to decrease over time.



**U(t)=(U\_t):**

-The growth term is relatively small (α⋅small U\_t).

-Dissipation (−β⋅H(t)) will be in charge earlier, limiting wave height growth.

**U(t)=(U\_t \*1.5):**

-The growth term is much larger (α⋅Big U\_t).

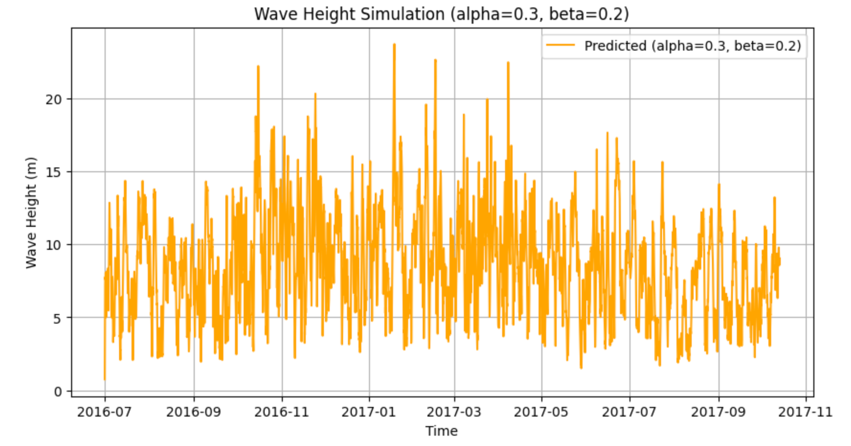
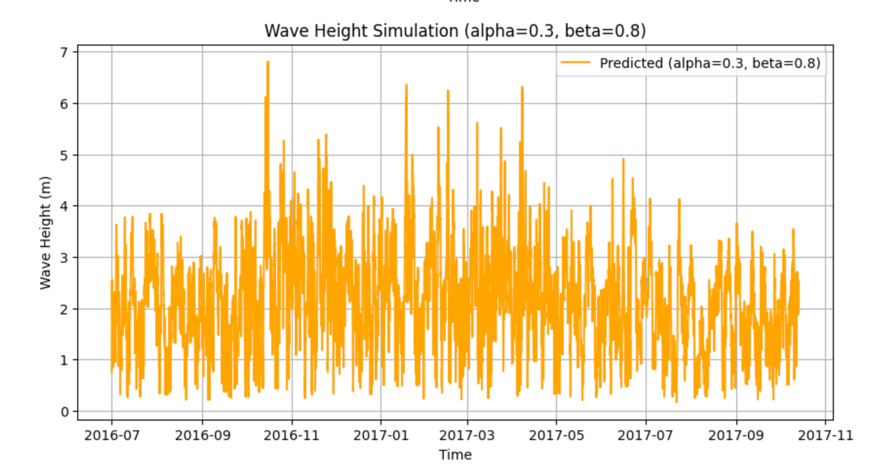
-Wave height increases quickly

**4. Dissipation Coefficient (β):**

Wave energy loss caused by the dissipation mechanism is represented by the parameter β, Higher β values indicate faster dissipation.

If this coefficient is high, Increases the dissipation term (−β⋅Ht), leading to slower wave growth and faster wave damage, the system becomes less sensitive to wind speed changes.

If this coefficient is low, Slows the rate of energy loss, allowing waves to grow taller and persist longer.



**Beta = 0.20:**

Dissipation is weak (−0.20⋅H(t)).

Growth is stronger, resulting in taller waves and slower wave damage.

**Beta = 0.80:**

Dissipation is strong (−0.80⋅H(t)).

Smaller peak and shorter-lived waves.

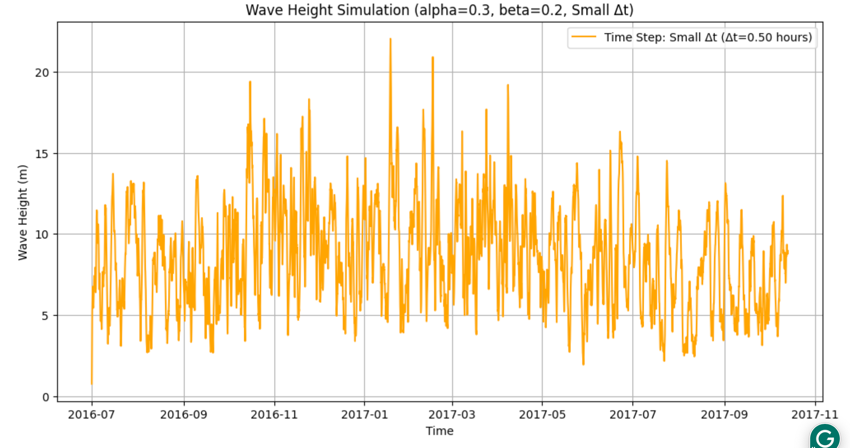
Wave growth is slower, and wave damage after the peak is faster.

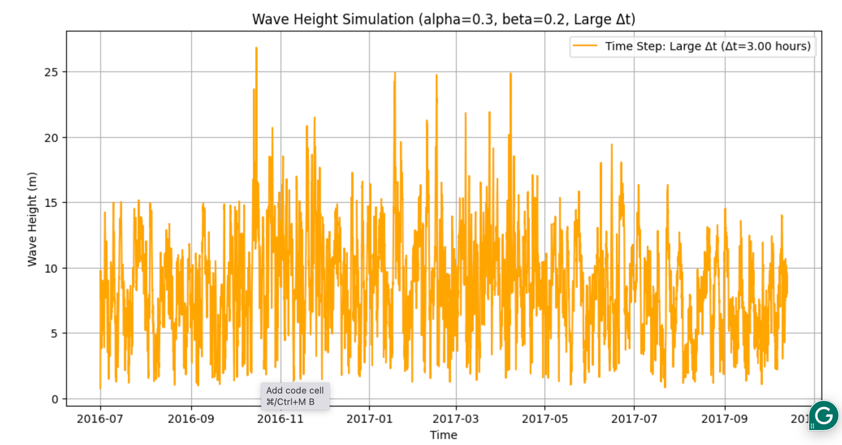
**5. Time Step (Δt):**

Represents the size of the time interval over which changes in wave height are calculated.

A small value of time step improves the accuracy of the simulation by better capturing the dynamic behavior of the system.

A large value of time step, Oversimplifies the dynamics, potentially leading to inaccurate results.





Δt=0.50:

The small step size ensures that changes in H(t) are calculated precisely.

Results in a smooth and accurate depiction of wave height dynamics.

Δt=3.00:

The large step size averages out the intermediate changes in H(t).

This causes the wave height to converge more quickly but at the cost of losing alot of details.

|  |  |  |
| --- | --- | --- |
|  | **Description** | **What happened in case of increasing** |
| **H(t)** | Current wave height. | Higher values increase dissipation, slowing growth. |
| **U(t)** | Wind speed driving wave growth. | High wind speed The growth term (α⋅Ut) becomes larger, resulting in faster wave growth. |
| **α** | Growth coefficient controlling wind energy transfer. | Faster wave growth, more sensitivity to wind. |
| **β** | Dissipation coefficient controlling energy loss. | Slower growth, faster damaging waves. |
| **Δt** | Time step size. | Smaller steps improve accuracy |

# **Task 3:**

## **Q1: Different optimization techniques to tune the simulation parameter:**

I have used Simulated Annealing and Sine Cosine Optimization, which are applied in the code file.

## **Q3: Comparison and Analysis of Standard Modeling Techniques Using Optimization Methods:**

**Model 1: Simulated Annealing Optimization:**

Best Alfa (α)=0.0177, Beta (β)=0.0686

MSE: 0.3053

MAE: 0.4249

Best Cost: 3441.4522

**Model 2: Sine Cosine Optimization (SCA):**

Best α=0.0127, β=0.0500

MSE: 0.3006

MAE: 0.4144

Best Cost: 3388.0331

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Best α** | **Best β** | **Best Cost** | **MSE** | **MAE** |
| **Simulated Annealing Optimization** | 0.0177 | 0.0686 | 3441.4522 | 0.3053 | 0.4249 |
| **Sine Cosine Optimization (SCA)** | 0.0127 | 0.0500 | 3388.0331 | 0.3006 | 0.4144 |

**Main insights:**

The SCA algorithm achieves better parameter optimization (Alfa and Beta) as it results in lower MSE, MAE, and Best Cost compared to Simulated Annealing.

This indicates that SCA finds a more optimal balance between growth (Alfa) and dissipation (Beta).

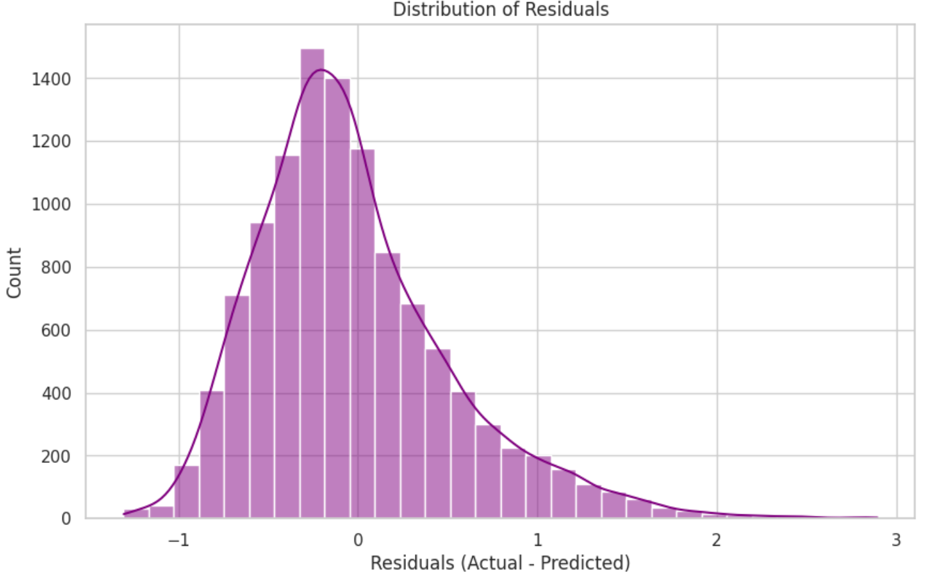
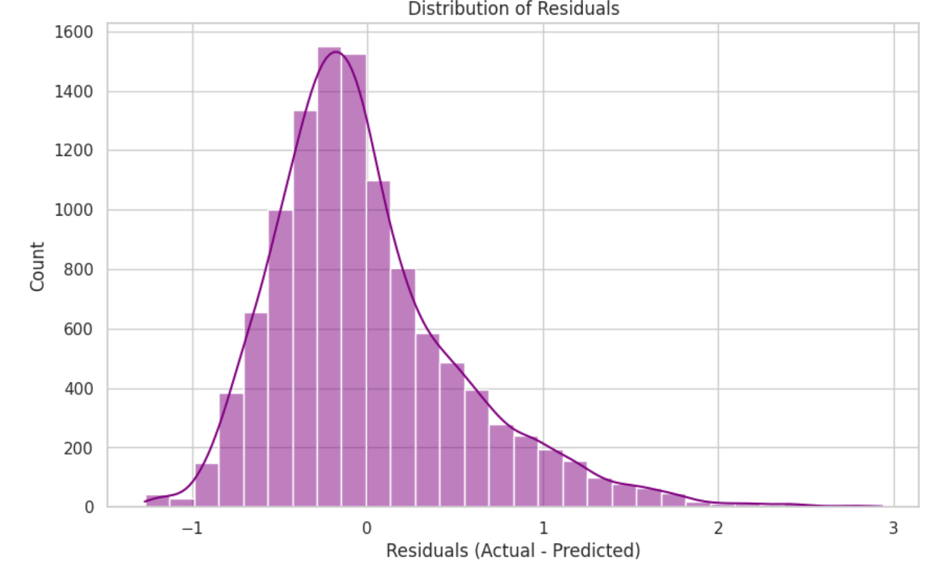
SCA has a little better MSE and MAE compared to Simulated Annealing.

The tighter clustering in the Actual vs Predicted scatter plot and compact residual distribution further reinforce SCA’s improved accuracy.

**Conclusion:**

As a result, we can see that performance by comparing the Simulated Annealing and the

Sine Cosine optimization techniques, we can see the SCA was more effective for my data set, which is the Ocean Waves. The SCA has achieved better α=0.0127 and β=0.0500, I decided that it is better because this algorithm has lower Mean Squared Error (MSE = 0.3006), lower Mean Absolute Error (MAE = 0.4144), and Best Cost (3388.0331). on the other hand the SA with α=0.0177 and β=0.0686 has achieved higher error metrics (MSE = 0.3053, higher MAE = 0.4249, and Best Cost = 3441.4522)



# **Task 4:**

## **Q1: Workflow of different optimization techniques:**

**1-Simulated Annealing:**

It is a probabilistic optimization technique that takes inspiration from the physical annealing process in metallurgy, where a material is subjected to heating to a high temperature followed by a gradual cooling process to get a minimum internal energy. In optimization, SA is employed to determine the global minimum of a cost function by studying the parameter space and sequentially improving the answer.

**1-Initialization parameters:**

The optimization begins by initializing the parameters alpha and beta, which are

(initial\_alpha and initial\_beta), Then the initial temperature (temp) is set to a high value (t\_0) which allows the algorithm to explore the parameter space broadly at the start, then the minimum temperature (temp\_min) is defined to terminate the optimization once the search has sufficiently narrowed, and a list cost\_history is created to track the progress of optimization over iterations



**2- Start Optimization in the while loop:**

The algorithm runs a while loop as long as the temperature is greater than temp\_min.

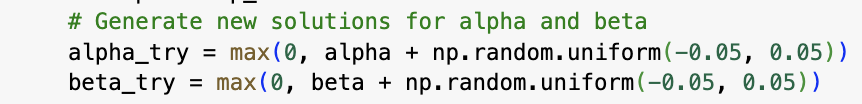
This ensures sufficient exploration of the parameter space at higher temperatures and progressively focuses on fine-tuning as the temperature decreases.

A black text with a white background

Description automatically generated

**3- Generating new Solutions for Alpha and Beta:**

At each iteration, the algorithm generates new trial values for α and β by applying small random adjustments. These new values introduce variability to explore new parts of the parameter space and prevent the algorithm from getting stuck in local minima.



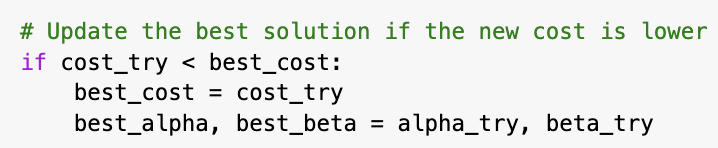
**4- Evaluating Costs:**

The cost\_function calculates the error for the current solution (cost\_current) and the new solution (cost\_try). These costs measure how well the parameters perform in predicting the observed wave heights.

A close up of words

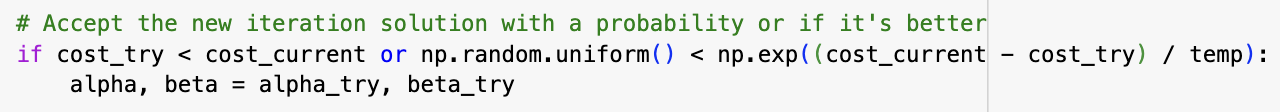
Description automatically generated

**5- Updating the Best Solution:**

****If the trial solution improves the cost (cost\_try < best\_cost), the algorithm updates best\_cost, best\_alpha, and best\_beta to save the new best solution found during the process.

**6- Accepting new iteration Solutions:**

The algorithm always accepts trial solutions with lower costs (cost\_try < cost\_current). However, if this iteration cost is worse, it is accepted with a probability based on the temperature. This mechanism allows the algorithm to escape local minima by occasionally accepting worse solutions.



**7- Tracking the cost:**

At each iteration, the best cost is appended to the cost\_history list. This provides a detailed record of the optimization progress and can be used for analysis and visualization.

**8-Cooling the Temperature:**

A close-up of a word

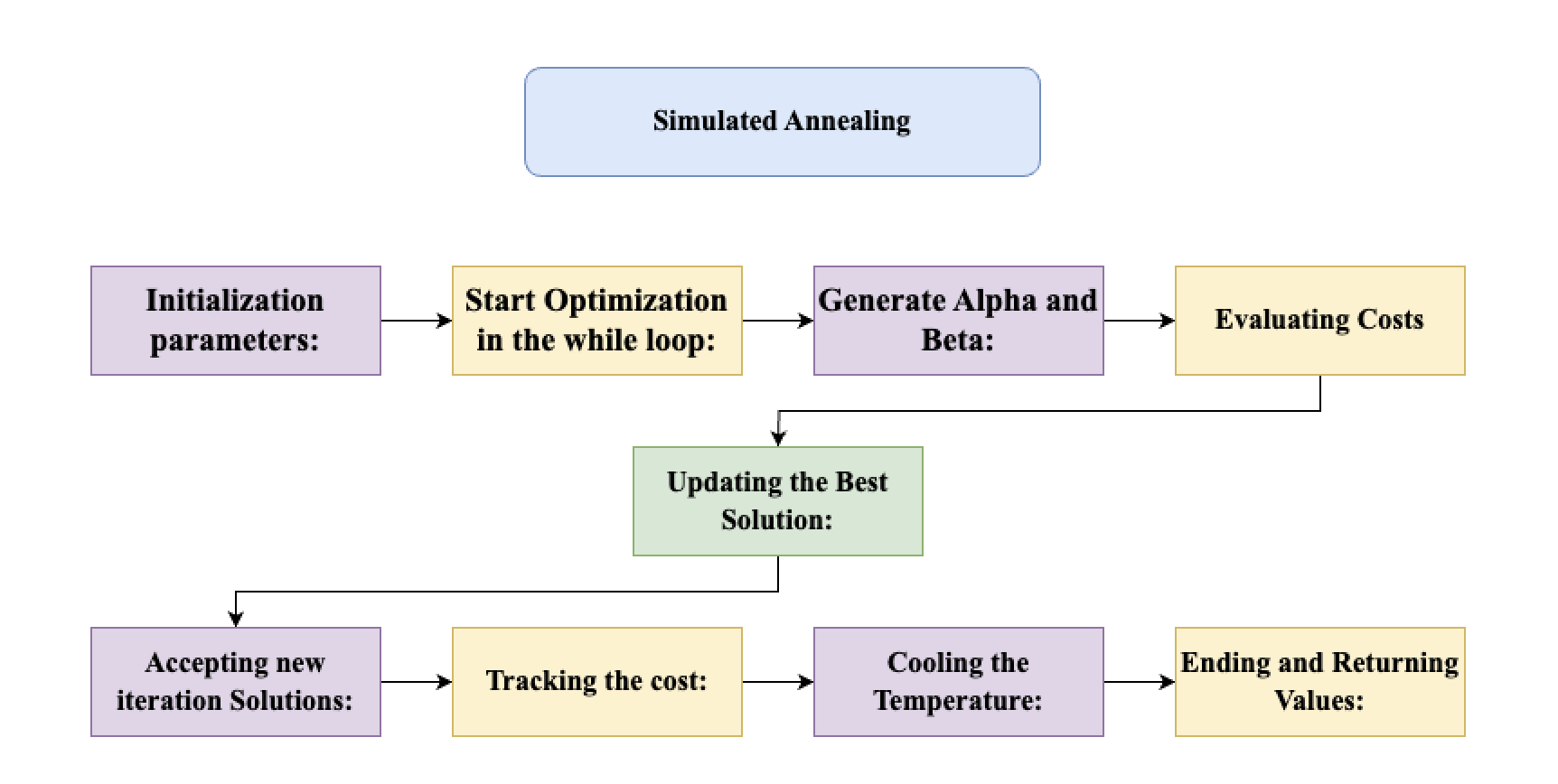
Description automatically generatedThe temperature is sequentially reduced by multiplying it with a cooling rate (alpha\_cooling).

**9- Ending and Returning Values:**

The loop ends when the temperature drops below temp\_min; then, the algorithm returns the optimized parameters (best\_alpha and best\_beta), the best cost achieved, and the cost\_history, summarizing the optimization process.

A close up of a text

Description automatically generated



**2-Sine Cosine Optimization:**

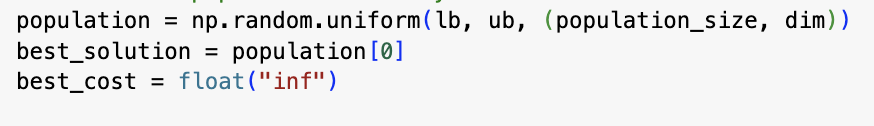
The sine cosine algorithm is a population-based optimization algorithm inspired by the

mathematical features of sine and cosine functions. It iteratively converges a population of solutions towards the update of the global minimum of a cost function.

**1-Initialization parameters:**

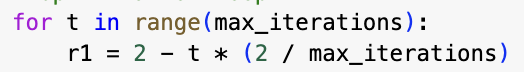
First, this algorithm sets the lower and upper bounds for the parameters Alpha and

Beta. In other words, these define the range of possible values that those parameters will have. To have some variance in the search space, it initializes a population of population\_size solutions randomly within those bounds. It initializes best\_cost as infinity (Infinity) and the first solution in the population as the initial best\_solution. This ensures that the starting values are updated by any improvements in the optimization process.



**2-Optimization Loop (for t in range(max\_iterations)):**

The main loop runs for a fixed number of iterations (max\_iterations). The algorithm makes use of a variable, r1, that ranges from 2 to 0 as iterations go on. However, lower values in later iterations promote mining by being focused on fine-tuning close to the ideal solution, whereas higher values of r1 in earlier iterations encourage discovery with more substantial changes possible.



**Substep 2.1: Update Each Solution (for i in range(population\_size)):**

Within each iteration, the algorithm loops over each solution in the

population. For each solution, it updates the parameter values (α and β) using sine and cosine functions.

R1 is a random value between 2 and 0, at the beginning of the simulation it allows searching in a variety of values, at the end of the simulation it will focus around the best values



**Substep 2.2: Update Each Parameter (for j in range(dim):)**

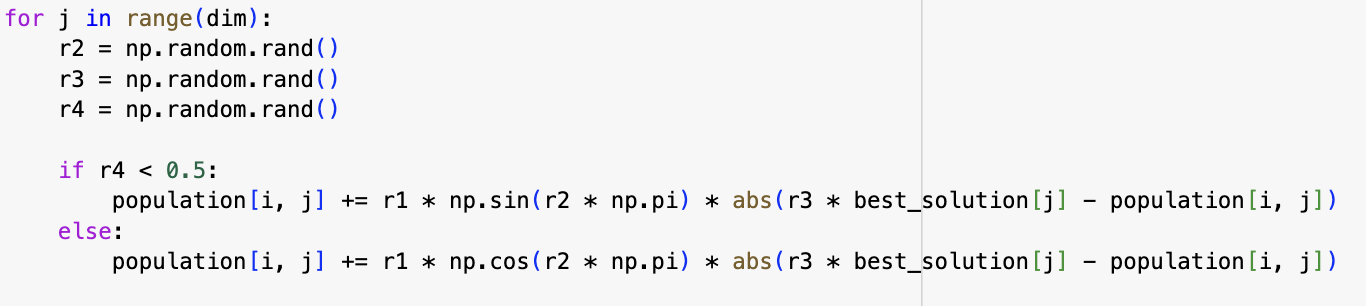
The parameter is updated using the sine function if r4<0.5 and 𝑟4<0.5:

The r4 has a probability of 0.5, which means 50% will use cosine and 50% will use sine, so the r4 is the parameter that decides to use the sine or the cosine.

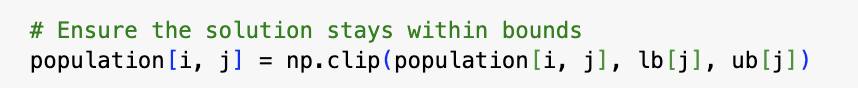
If r4<0.5, the sine function is used to update the parameter:

If r4≥0.5, the cosine function is used:

The updated parameter value is clipped to remain within the specified bounds (lb and ub).In order to explore the parameter space and direct the search for the optimal solution, this procedure adds randomization.



**Substep 2.3: Make sure the solution is between boundaries:**



**3-Evaluate the Solution:**

After updating the parameters, the algorithm calculates the cost of the new solution using the cost\_function. This evaluates how well the updated parameters fit the observed data.



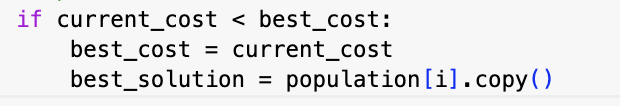
**4-Update the Best Solution:**

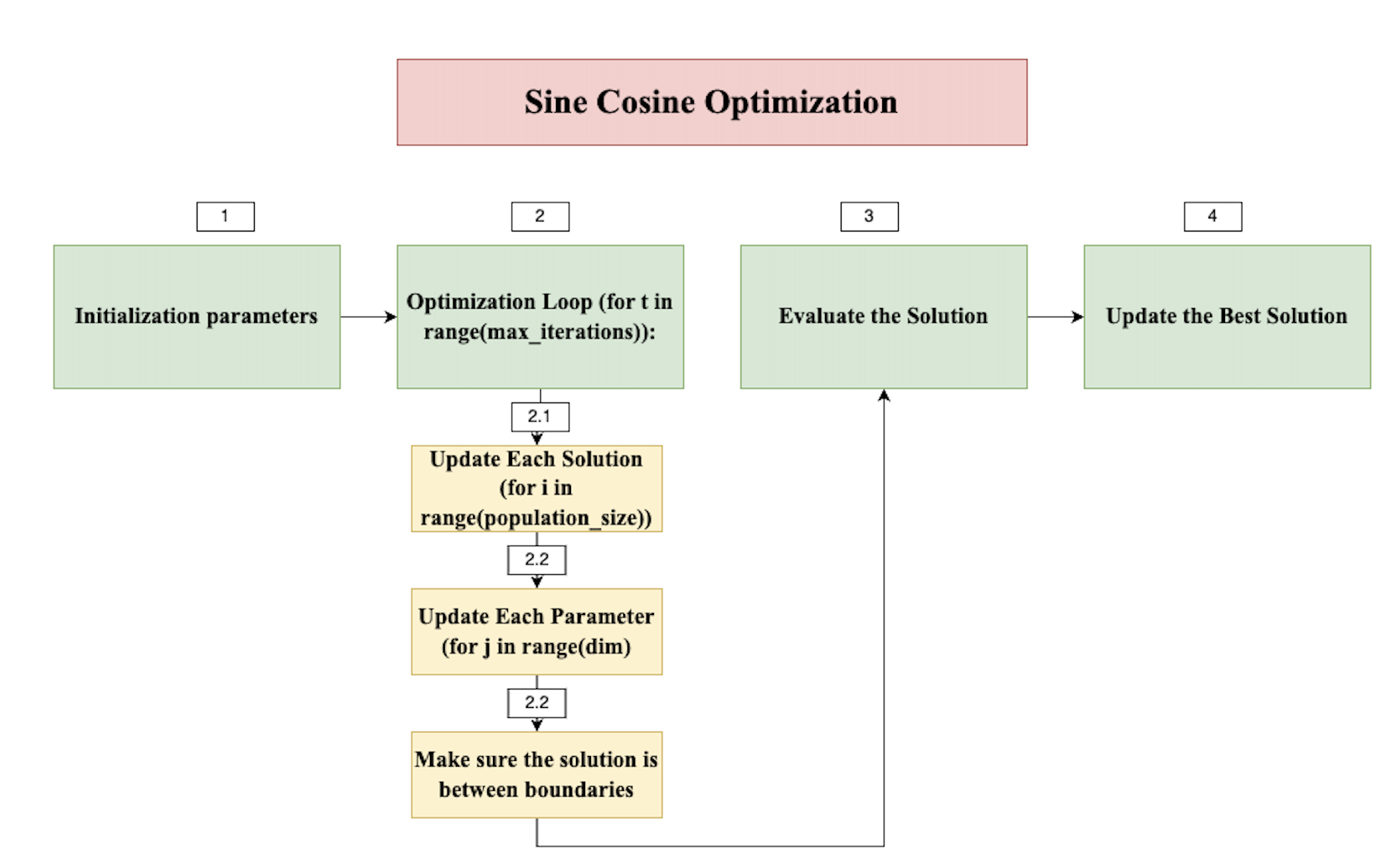
The algorithm will be modified and updated if the cost of the new solution is better than the best\_cost:

best\_cost = the new cost.

best\_solution = updated solution.

By doing this, the algorithm will definitely constantly track the parameters that perform the best throughout the population and iterations.





## Q2: analyze and compare the performance of the produced models

|  |  |  |
| --- | --- | --- |
| **Model** | **MSE** | **MAE** |
| **SINDy** | 0.286047 | 0.363496 |
| **Sine Cosine Optimization** | 0.3006 | 0.4133 |
| **Simulated Annealing Optimisation** | 0.3053 | 0.4249 |

**SINDy Model:**

**MSE**: 0.286047

**MAE**: 0.363496

**Performance**:

The lowest MSE and MAE values among all three models indicate that the SINDy model provides the best overall performance.

**Sine Cosine Optimization (SCA):**

**MSE**: 0.3006

**MAE**: 0.4133

**Performance**:

Slightly higher MSE and MAE compared to SINDy, indicating marginally less accurate predictions.

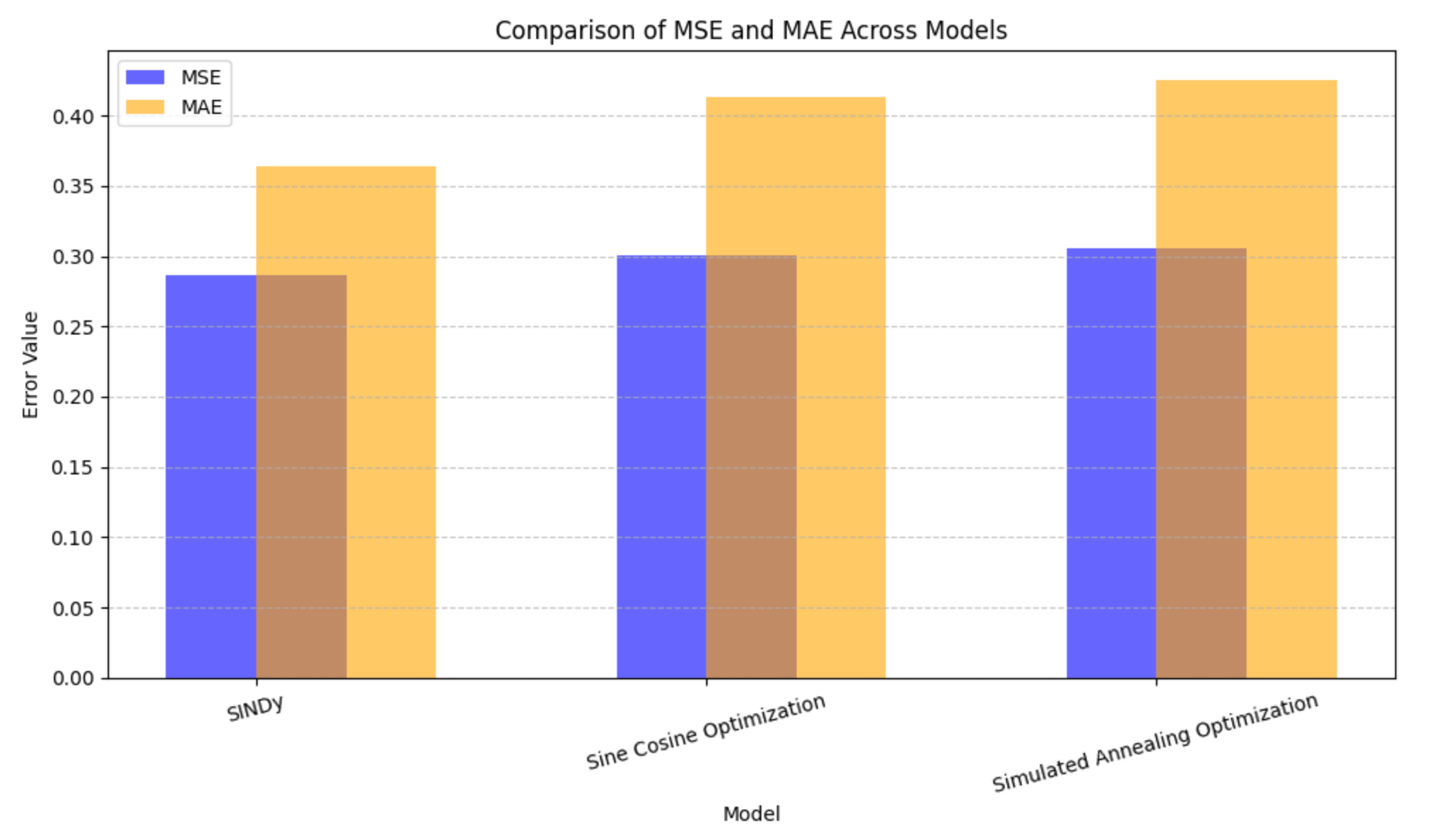
**Simulated Annealing Optimization (SAO):**

**MSE**: 0.3053

**MAE**: 0.4249

**Performance**:

The highest MSE and MAE among the models suggest that SAO is the least accurate.



# **References:**

## **Task 1:**

Q2:

Pinto, P.M.G.M. et al. (2023) Predicting significant wave height with artificial neural networks in the South Atlantic Ocean: A hybrid approach - ocean dynamics, SpringerLink. Available at: https://link.springer.com/article/10.1007/s10236-023-01546-y?utm\_source=chatgpt.com (Accessed: 30 December 2024).

incois.gov (2023) Ocean Wave Modeling and Forecasting System. Available at: https://incois.gov.in/documents/Wave\_forecasting\_system\_Sandhya.pdf (Accessed: 30 December 2024).

Q3:

Afzal, M.S. et al. (2023) Prediction of significant wave height using machine learning and its application to extreme wave analysis - Journal of Earth System Science, SpringerLink. Available at: https://link.springer.com/article/10.1007/s12040-023-02058-5 (Accessed: 02 January 2025).

Salah, H. and Elbessa, M. (2024) Using machine learning techniques to predict significant wave height compared with parametric methods, Science Publishing Group. Available at: https://sciencepublishinggroup.com/article/10.11648/j.eas.20240905.12?utm\_source=chatgpt.com (Accessed: 02 January 2025).

D. Chen, F. Liu, Z. Zhang, X. Lu and Z. Li, "Significant Wave Height Prediction based on Wavelet Graph Neural Network," 2021 IEEE 4th International Conference on Big Data and Artificial Intelligence (BDAI), Qingdao, China, 2021, pp. 80-85, doi: 10.1109/BDAI52447.2021.9515293.

Y. Wu, R. Zhang, F. Liu, Z. Liu, J. Wu and H. Dong, "Significant Wave Height Prediction Based on VMD-SA-MLP-BP," 2023 China Automation Congress (CAC), Chongqing, China, 2023, pp. 7262-7267, doi: 10.1109/CAC59555.2023.10451885.

A. Neyestani, D. M. Toma, A. Falahzadeh, P. Daponte, J. Del Rio Fernandez and L. De Vito, "A Significant Wave Height Data-Driven Modeling for Digital Twins of Marine Environment," 2024 IEEE International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters (MetroSea), Portorose, Slovenia, 2024, pp. 495-500, doi: 10.1109/MetroSea62823.2024.10765714.

P. A. Hwang, J. D. Ouellette, J. V. Toporkov and J. T. Johnson, "A Simulation Study of Significant Wave Height Retrieval From Bistatic Scattering of Signals of Opportunity," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 8017205, doi: 10.1109/LGRS.2021.3096871.