Spectral wave model of the Southern North Sea

This README provides practical and background information on the dataset. The dataset can be cited as:

> DHI. (2025). Spectral wave model of the Southern North Sea: MIKE 21 SW model setup, outputs and observation data (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.17099653

See the license for details on data usage.

> **Important Disclaimer**
> This model setup is a **highly simplified version** of a hindcast model developed by DHI. The mesh has been reduced from over **80,000** to approximately **8,000 elements**, and input data is coarser in both time and space. As a result, model accuracy is lower — e.g., typical RMSE for significant wave heights increases from ~20 cm to **30 cm**.
>
> This model is **not calibrated** for operational use. It is intended for **educational and research purposes only**, and the results **must not** be used for decision-making. The performance shown here does **not** reflect that of DHI's high-resolution models.

Intended use

This dataset is designed to support educational, research, and exploratory activities, including:

- Experimenting with coastal and ocean modeling using MIKE 21 SW.
- Validating spectral wave model outputs with observational data.
- Developing methods for automatic model parameter calibration.
- Building data-driven models, including machine learning surrogates of MIKE simulation results.

Folder structure

The repository is organized in the following way

- README.md
- license
- model
  - MIKE SW model setup file (.sw)
  - log file (.log)

- input
  - meteorological forcing (wind)
  - boundary conditions (spectral wave energy density)
  - mesh file (.mesh)
  - area file with water levels and currents (.dfsu)
- observations
  - csv files containing observational data for model validation
- code
  - Jupyter notebooks exploring the dataset
  - Jupyter notebooks for model validation using observations
  - a requirements file (.txt)
- figures
  - figures generated by notebooks
- output_sample
  - a reduced MIKE model result data file (.dfsu and .dfs0)

Separately from the current repository, you can find the output zip file (in Zenodo):

- output
  - MIKE model result data (.dfsu and .dfs0)

Introduction

This folder contains data for a MIKE 21 Spectral Wave Model flexible mesh simulation in the southern part of the North Sea. The model domain spans from the English Channel along the southern and eastern English coast, part of the northern French coast, and the coasts of Belgium, the Netherlands, northwest Germany, and southwest Denmark.

This model setup and dataset represent a simplified version of the work performed by DHI for the Rijksdienst voor Ondernemend Nederland (RVO). In that work, DHI provides metocean condition predictions for the IJmuiden Ver (IJB) Offshore Wind Farm Zone. More details are available in the blogpost here and the scientific report.

This repository includes

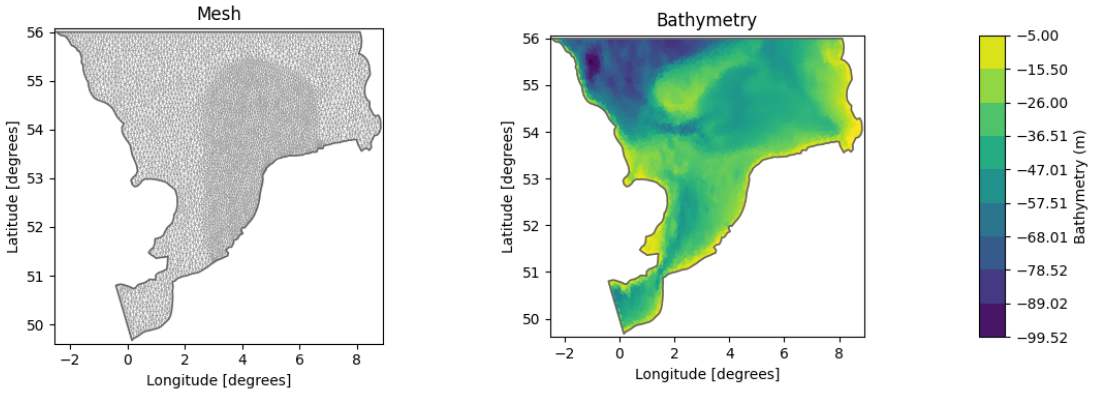
- Model setup files, input data, and observation data.
- Simulation result, wind input data, and the hydrodynamic area input file are available on Zenodo https://doi.org/10.5281/zenodo.17099653.

The dataset covers the period 2022-01-01 to 2023-12-31. Since the simulation includes a spin-up period, it is recommended to discard the first two days of result data for analysis.

The MIKE 21 Spectral Wave Model

MIKE 21 SW is a modelling system based on a flexible mesh approach. The modelling system has been developed for applications within oceanographic, coastal and estuarine environments. The governing equation is the wave action balance equation formulated in either Cartesian or spherical coordinates. More information can be found in the documentation.

The mesh consists of triangular elements, with higher resolution in areas of interest. For each mesh node, a depth value is assigned by interpolating bathymetric data. The mesh is shown in the figure below and consists of approximately 8000 elements covering an area of about 24.000 km^2.



A simulation can be performed with the MIKE modelling software, while supplying the following data for driving the model:

- the spectral wave energy density on the North boundary
- the spectral wave energy density on the South boundary
- the water level and currents in the domain
- wind velocities in the domain.

This input data can be found in the /input folder. Model parameters, e.g. bed resistance constant, can be calibrated by comparing the model simulations to observational data.

The model outputs are provided both as point timeseries (dfs0) and as time-varying full-domain results (dfsu) including the variables in the table below.

Variable	Unit	Acronym
Significant wave height	meter	Hm0

Variable	Unit	Acronym
Peak wave period	second	Tp
Zero crossing period	second	T02
Mean wave direction	degree	Mdir

These results are provided as a separate file on Zenodo. After running a simulation, a log file is generated. For the benchmark simulation, this log file is available in the /model folder and contains details on parameter settings and runtime.

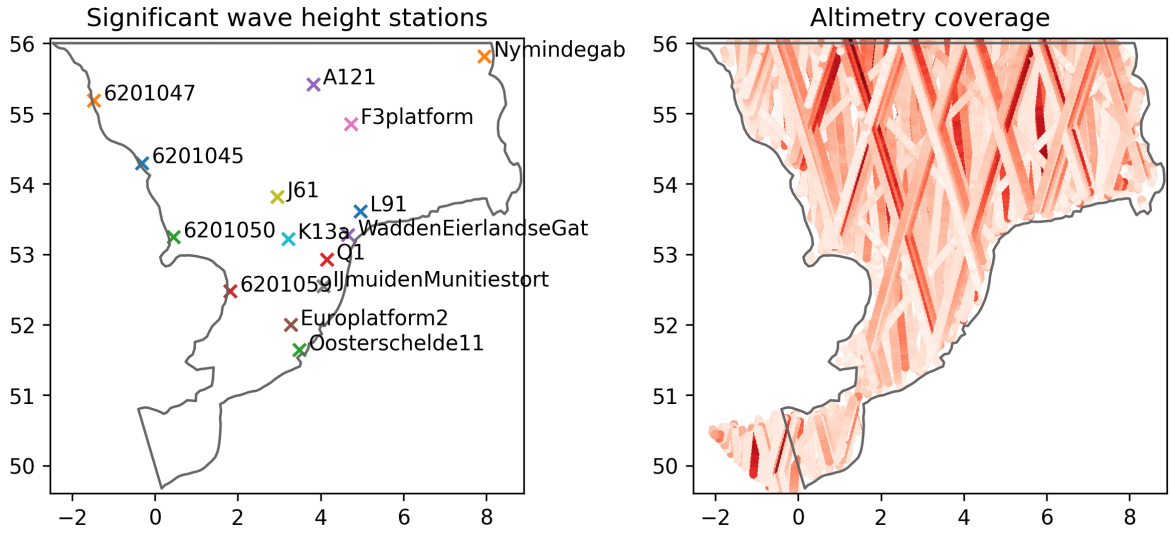
Running the MIKE software requires a license. Read more about obtaining a license and about academic licensing. If you already have a license, read the step-by-step guide to learn how to run simulations.

The input and output data are given as dfs files, which can be processed using the MIKE IO Python package developed by DHI.

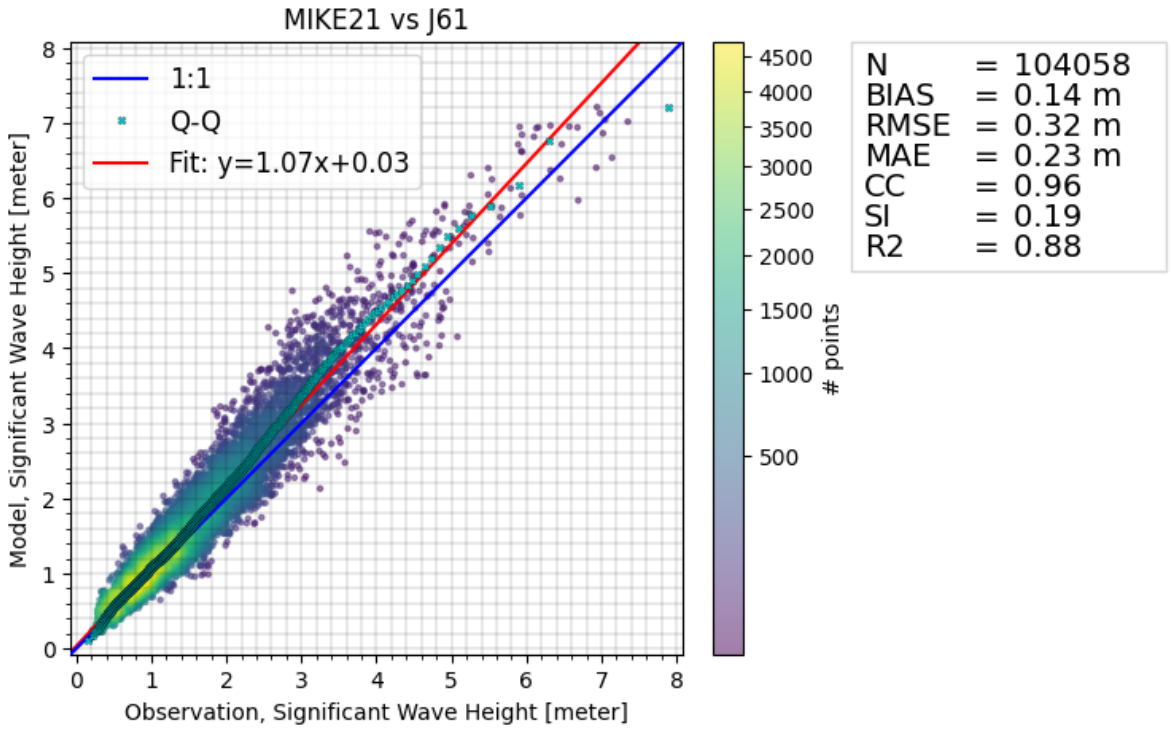
Model validation

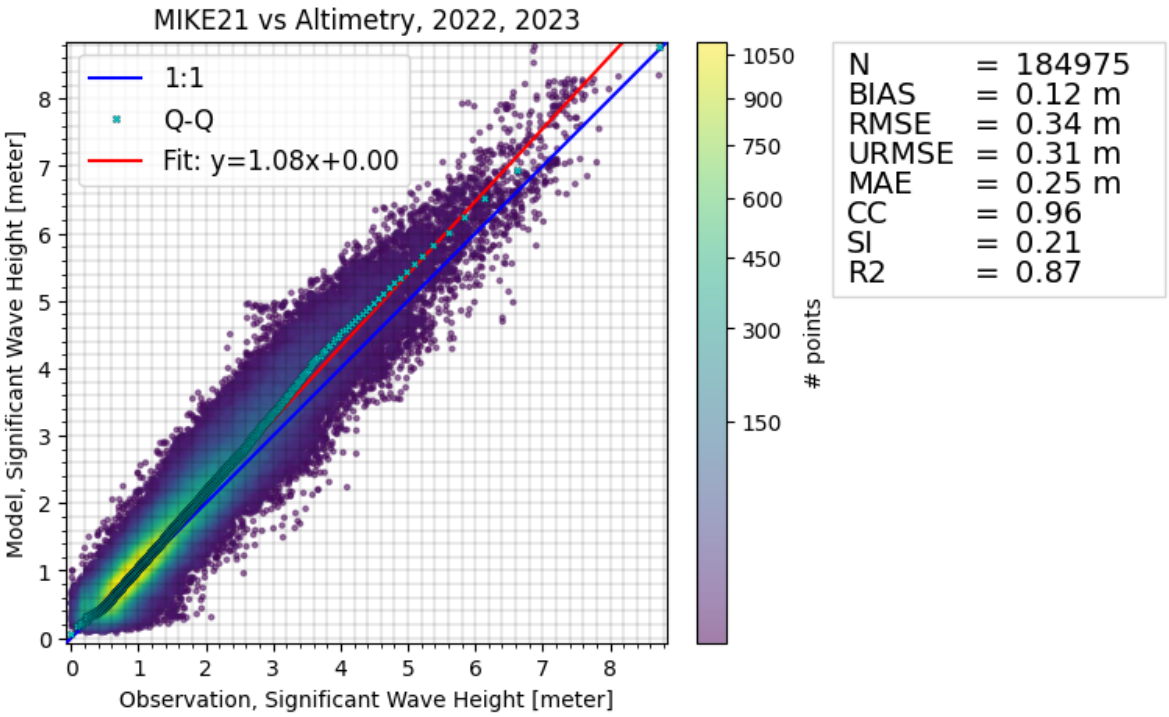
The ModelSkill Python package developed at DHI can be used to validate model outputs by comparing them to observational data.

In this case, the model is validated using both in-situ measurements from stations in the Southern North Sea and satellite altimetry data. This validation is performed in the Jupyter notebooks located at `code/model_validation_*.ipynb`. The leftmost figure below shows the location of the measurement stations, while the rightmost figure shows the spatial coverage of the altimetry data.



Below is an example comparison between observational data from a measurement station and the MIKE 21 simulation, followed by a comparison with satellite altimetry data. Many more similar plots can be found in the validation notebooks `code/model_validation_*.ipynb`.





Data processing

Only active stations and buoys were included in the analysis, and all timestamps were converted to UTC. The following processing steps were applied to specific observation types. - **Wave height observations**: A few outliers and zero observations (caused by a station malfunction) were removed. - **Peak wave period observations**: Values were only consideret when the significant wave height is larger than 1.0 m, as lower wave heights may result in inaccurate measurements of the peak wave period due to the discrete nature of the parameter. - **Mean wave direction observations**: Treated as directional data, with circular statistics applied when calculating error metrics. - **Altimetry data**: Only data points marked as "good" by quality flags are included.

Data sources

Several data sources contribute to this case, as outlined below. The data sources are categorized into three sections: data necessary to run the model, observation data for model validation, and altimetry data from satellite observations.

Data sources necessary to run the model

Data	Source	Citation	License
Mesh	The mesh is developed by DHI.	Repository citation, see the top	CC BY-NC 4.0
Bathymetry	The bathymetry is based on: 1: the EMODnet Bathymetry Digital Terrain Model (DTM) 2020. 2: Rijkswaterstaat's bathymetric measurements along the Dutch coasts 3: FUGRO's bathymetric measurements at the HKZ, HKN and HKW wind farm zones 4: MMT SWEDEN AB's bathymetric survey for the TNW wind farm zone 5: GEOxyz's bathymetric survey of IJmuiden Ver wind farm zone (alpha and beta) 6: FUGRO's bathymetric survey of IJmuiden Ver Wind Farm Zone (gamma). The bathymetry data is interpolated onto the mesh.	1: EMODnet Bathymetry Consortium (2020). EMODnet Digital Bathymetry (DTM 2020).EMODnet Bathymetry Consortium https://doi.org/10.12770/bb6a87dd-e579-4036-abe1-e649cea9881a For the rest, more information can be found in: DHI, "Metocean Assessment Modelling Report, IJmuiden Ver Wind Farm Zone", Dec 2023, URL: https://offshorewind.rvo.nl/file/download/bfa49f34-f894-4562-882f-eb1a8b7497e9/ijv_20231222-dhi-metocean-modelling.pdf (accessed: 2024-11-13)	The data has been interpolated to the mesh resolution such that the original resolution is no longer represented nor reproduceable from the attached mesh. Hence, the mesh with the depth estimates is shared under the same licence as the rest of this repository, namely CC BY-NC 4.0.

Data	Source	Citation	License
Wind and Air pressure	A subsampled version of the high-resolution ERA5+Weather Research & Forecasting model (WRF) data used for the RVO project. The data is subsampled in space to 16km and temporally to hourly data. Air pressure is provided, but not required for the SW model. It is included because it is necessary for the HD model, which generates AreaHD, that is an input to the SW model.	Repository citation, see the top.	Only non-commercial use: CC BY-NC 4.0.

Data	Source	Citation	License
Boundary condition	The boundary conditions are represented by a 36x36 matrix describing the spectral energy across both direction and magnitude. The data is downsampled to include spectra in three points in the North and one spectrum in the South (covering the English Channel). The data is generated using the DHI global wave model dataset (GWM). Read more in the Dataset Details.	Repository citation, see the top.	Only non-commercial use: CC BY-NC 4.0.
Water level and current velocities	The water level and current velocities in each element on the grid is generated by a MIKE 21 HD flexible mesh hydrodynamic model simulation of the same area. The data is collected in the AreaHD file.	Repository citation, see the top.	Only non-commercial use: CC BY-NC 4.0.

Observation data

The observational data come from in-situ measurement stations provided by the Copernicus Marine Service (CMEMS). These near real-time ocean observations, including parameters such as significant wave height, peak wave period and mean wave direction, were accessed through the CMEMS open download service and cover the period 2022–2023.

Data	Source	Citation	License
Measurement stations	Data from the stations is retrieved through the Copernicus Marine Service (CMEMS, or Copernicus Marine Environmental Monitoring Service). More specifically, the open download service dashboard, which can be found here. The physical parameter list can be found here.	Copernicus Marine Service, "Global Ocean In-Situ Near Real-Time Observations", https://doi.org/10.48670/moi-00036 (accessed: 2025-09-16).	See full license here. Excerpt: "The Licensee is hereby granted a worldwide, non exclusive, royalty free, perpetual licence, (subject to the terms and conditions of this agreement) to: (a) make and use such reasonable copies of Copernicus Marine Service Products for internal use and back up purposes; (b) modify, adapt, develop, create and distribute Value Added Products or Derivative Work from Copernicus Marine Service Products for any purpose; (c) redistribute, disseminate any Copernicus Marine Service Product in their original form via any media."

Altimetry data

The altimetry data is based on satellite observations and was downloaded via DHI's services for marine observation data. The satellites and data sources are listed below:

Satellite name	Owner	Link
Sentinel-3A (3a)	European Space Agency (ESA)	https://sentiwiki.copernicus.eu/web/s3-altimetry-instruments
Sentinel-3B (3b)	European Space Agency (ESA)	https://sentiwiki.copernicus.eu/web/s3-altimetry-instruments
Sentinel-6A (6a)	European commission, ESA, EUMETSAT, NASA and NOAA	https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-6/data-products
CryoSat-2 (c2)	European Space Agency (ESA)	https://earth.esa.int/eogateway/catalog/cryosat-products
SARAL (sa)	French National Space Agency (CNES) and Indian Space Research Organisation (ISRO)	https://podaac.jpl.nasa.gov/dataset/ALTIKA_SARAL_L2_OST_XOGDR
Jason-3 (j3)	NASA, CNES, EUMETSAT, NOAA	https://sealevel.jpl.nasa.gov/missions/jason-3/summary/