**Title**

A machine learning model for the estimation of hourly non-tidal water levels in the Dutch coastal zone: Based on satellite altimetry observations and pressure and wind fields from ERA5

**Author**

[Schijvenaars, Sofie](https://repository.tudelft.nl/islandora/search/author%3A%22Schijvenaars%2C%20Sofie%22?collection=education)

**Date**

2024-06-20

**Abstract**

Globally, coastal communities face increasing risks from climate-related hazards such as flooding, shoreline erosion, and salt intrusion. These hazards pose threats to both people and their environment, with extreme sea level events increasing these risks. Satellite altimetry allows for global observation of the sea level, reaching remote regions that are not covered by unevenly distributed tide gauges, as these are concentrated in densely populated regions of Western cultures. However, their 10- to 35-day repeat cycles complicate the capture of extreme sea level events. Machine learning offers a promising approach to combine direct satellite altimetry observations with ERA5 pressure and wind speed fields into a data-driven model. As opposed to global and regional numerical models, which require substantial time and expertise to develop, machine learning models are time efficient and require relatively low effort to develop and expand.

This study presents a shallow neural network that effectively estimates hourly non-tidal water levels in the Dutch coastal zone, using X-TRACK retracked and reprocessed satellite altimetry observations and ERA5 hourly pressure and wind speed fields. Reprocessed satellite altimetry observations from 11 missions are used to provide more accurate coastal observations. Tide gauge records are used as ground truth. Both tide gauge and satellite altimetry data are corrected for harmonic tidal signals before training. A 48-hour time window is applied, using all data from 48 hours to 1 hour prior to the estimates as input into the network. The area of interest covers most of the North Sea, from the Strait of Dover to the northern North Sea, excluding the Danish and Norwegian coasts. The neural network is trained and tested at three locations: Scheveningen, Vlissingen, and the Europlatform.

Results show that the neural network can estimate hourly non-tidal water levels with mean squared errors ranging from 0.011 to 0.018 m, mean absolute errors from 0.078 to 0.101 m and standard errors from 0.100 to 0.134 m. K-fold cross-validation with K=4 indicates high robustness, with mean squared errors varying by 0.004 m, mean absolute errors by 0.012 m and standard errors by 0.017 m. The model performs best for hourly and high water levels at the Europlatform and worst for high water levels at Scheveningen. This is partly due to the location of the Scheveningen tide gauge in a harbour with more localised disruptions of the water level compared to the tide gauge at the Europlatform. The ERA5 longitudinal wind speed component contributes most to the estimation of non-tidal water levels, accounting for $\pm$18\% of all weights corresponding to the input variables. Key regions for the estimation of non-tidal water levels include the Dutch coast and the northern North Sea.

When compared to a local numerical model, the developed neural network does not perform with the same accuracy. However, several upsides of the model are identified, such as high computational efficiency for single locations and easy implementation options for refinement of the model. Recommendations for future research focus mostly on improving the model's performance on high water levels and applicability to different regions.