Citation (APA)
Vurilj, M. P., Antolínez, J. A. A., & Morales-Nápoles, O. (2026). Spatial Clustering of Sea Level Hydrographs Across the Dutch Coast. In C. Coelho, C. Hallin, F. Sancho, & P. A. Silva (Eds.), *Coastal Dynamics 2025 : Volume 2* (pp. 114-120). (Coastal Research Library; Vol. 42). Springer Nature. https://doi.org/10.1007/978-3-032-15477-4_19

Spatial Clustering of Sea Level Hydrographs Across the Dutch Coast

Mia Pupić Vurilj(✉), José A. A. Antolínez, and Oswaldo Morales-Nápoles

Delft University of Technology, Delft, The Netherlands
`m.pupicvurilj@tudelft.nl`

Abstract. Extreme sea level events pose significant risks to coastal regions, with non-tidal residuals (NTRs) being a primary driver in low-lying areas like the Netherlands, where shallow seas amplify their impact. This study investigates the spatial patterns of NTRs along the Dutch coast using time series clustering on historical NTR hydrographs. The design of hydraulic boundary conditions divides the Netherlands into three coastal regions. To evaluate whether this division sufficiently captures regional variability, three clustering scenarios ($k = 3$, $k = 4$, and $k = 5$) were explored. The analysis identified $k = 5$ as the optimal configuration based on the Davies-Bouldin index. This result emphasized the importance of fine-scale approaches to understanding regional spatial variations in NTR dynamics.

Regional bathymetry and tide-surge interactions were explored as drivers of these spatial patterns. Southern stations near river systems and deeper waters displayed characteristics distinct from northern stations in the Wadden Sea, which are influenced by shallow tidal flats. Analysis of the M2 tidal constituent and the timing of NTR maxima relative to high tides underscored the role of tidal dynamics in shaping spatial clusters.

Future research will focus on integrating spatio-temporal patterns and environmental drivers into clustering methodologies, providing deeper insights for coastal risk management and adaptation strategies.

Keywords: Spatial Clustering · Hydrographs · Storm Surge · Non-tidal Residual

1 Introduction

Extreme sea level events, once occurring once per century, are projected to become annual at over half of all tide gauge locations by the end of the 21st century [1]. This rise poses significant risks to the environment and socioeconomic infrastructure, increasing susceptibility to recurrent coastal floods. To address these challenges, a deeper understanding of extreme sea level events and their dynamics is needed for effective adaptation measures and coastal flood risk management.

With its low-lying coastal areas, the Netherlands is particularly vulnerable to coastal flooding [2]. This underscores the need for detailed analysis of extreme sea level events, including their non-tidal residual (NTR) components. Such events can exhibit non-uniform temporal and spatial patterns, clustering in certain regions and periods, thereby intensifying flood risks. While basin-scale studies of extreme sea levels in the North Sea

C. Coelho et al. (Eds.): CD 2025, COASTALRL 42, pp. 114–120, 2026.
https://doi.org/10.1007/978-3-032-15477-4_19

exist [3, 4], there is a gap in recognizing and analysing the joint spatial and temporal characteristics of NTRs.

The design of hydraulic boundary conditions divides the Netherlands into three coastal regions [5]. This study investigates spatial patterns of measured historical NTR hydrographs to evaluate whether this division sufficiently captures spatial variability. This can advance best practices in local modelling, improve our ability to predict future extreme sea levels, and enhance stakeholders' capacity for risk management and decision-making in coastal areas.

2 Materials and Methods

2.1 Data

Observed sea level data were used for this study. The data were collected from the Rijkswaterstaat, provided within the GESLA-3 (Global Extreme Sea Level Analysis) dataset [6]. Sixteen stations along the Dutch coast were selected, with observational periods ranging from 38 to 68 years and observation frequencies varying over time (Fig. 1). To analyse bathymetry characteristics, we used the dataset measured by the Rijkswaterstaat, referred to as "Vaklodingen" [7].

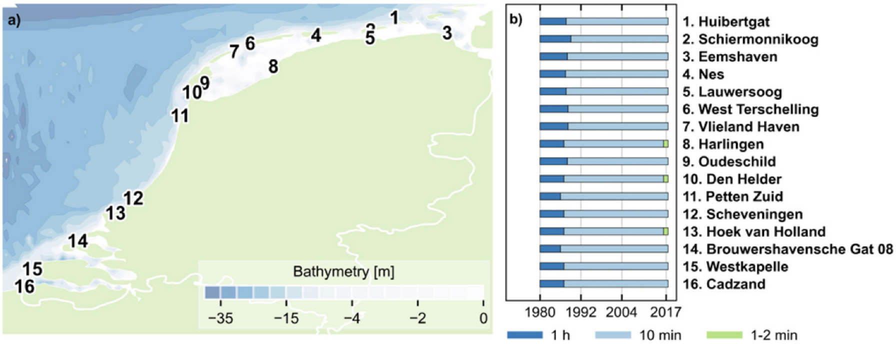

Fig. 1. **a)** Map of the Dutch coast, showing the location of the sixteen GESLA-3 stations used in this study, and **b)** the duration of the sea level records with frequency periods.

2.2 Data Preprocessing and Event Definition

The observed total sea level was searched for outliers, linearly detrended to remove long-term trends caused by sea level rise, and then split into two components: (i) the tidal component, and (ii) the non-tidal residual (NTR). The former was derived using harmonic tidal analysis [8] and the latter was calculated by removing the predicted tidal signal from the observed total sea level. The NTR thus includes water level variations caused by meteorological factors and wave effects, as well as the non-linear interactions between them. Since the tidal phase can be affected by depth changes due to the surge, and vice versa, the NTR also includes the non-linear effect of tide-surge interaction. The data remained in its original frequency, and any gaps were not filled.

Events were identified using the peak over threshold (POT) method on NTR, with a 70th percentile threshold per station. An event was defined from the first threshold-exceeding moment to the last, ensuring a 3-day gap for independence between subsequent events. The relatively low threshold allowed us to include less severe scenarios and enabled us to extract the time series of storm development rather than focusing solely on peak sea level magnitudes. This approach resulted in a total of 45,843 events across stations.

2.3 Spatial Time Series Clustering

To analyse spatial patterns, we defined NTR events across the entire study area by grouping occurrences within a 2-day window. This approach resulted in 3,140 events along the Dutch coast between 1980 and July 2017. However, due to data gaps in the Brouwershavensche Gat 8 tide gauge from Nov 5–17, Nov 26–Dec 8, and Dec 29–31, 2014, several events from that year were excluded. Ultimately, 3,129 NTR events were used for clustering analysis.

The clustering process involved the following steps, adapted from [9]:

1. **Time series Extraction**: For each event, the start time was the earliest occurrence across the coast, and the end time was the latest. NTR time series spanning this interval were extracted for each station.
2. **Normalisation**: Each time series was normalized to a time index ranging from 0 to 1, based on its duration in hours.
3. **Re-gridding**: Time series were re-gridded using linear interpolation to a uniform grid of 536 regular time steps, using the event with the largest time step and duration as a reference.
4. **Standardization**: Time series were standardized to have a mean of 0 and a standard deviation of 1.
5. **Concatenation**: All time series per station were combined into a single dataset with 536 time steps × 3,129 events, resulting in 1,677,144 features.
6. **Dimensionality Reduction**: Applying PCA to the features to reduce data dimensionality, retaining 99% of the variance. This reduced the number of features from 1,677,144 to only 14.
7. **K-medians clustering**: Clustering of the PCA-reduced dataset using three different numbers of clusters, $k = 3$, 4, and 5.

To evaluate different cluster number options, we employed the Davies-Bouldin index. Moreover, we used the Kolmogorov-Smirnov hypothesis test to determine if the empirical distribution functions of the clusters were statistically different.

3 Results and Discussion

Figure 2 presents the spatial clusters of NTR hydrographs. Based on the Davies-Bouldin index, the optimal number of clusters was determined to be $k = 5$. Notably, $k = 3$ was found to be more favourable than $k = 4$. The division into five regions highlights the value of a finer-scale approach for evaluating regional extremes and their underlying drivers. Additionally, the Kolmogorov-Smirnov test confirmed that the empirical distributions of all clusters were statistically distinct across all three clustering scenarios.

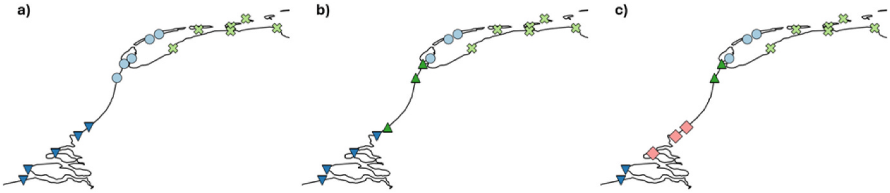

Fig. 2. Spatial clustering results for **a)** k = 3, **b)** k = 4, and **c)** k = 5, with different symbols and colours representing tide gauge stations within each cluster.

3.1 Bathymetry Characteristics and Tide-Surge Interactions

The spatial differences of clusters were analysed with a focus on bathymetry and tide-surge interactions.

Figure 3 illustrates the bathymetry within a 3.33 km (0.03 degrees) radius per station, including mean bathymetry values and a broader regional perspective. While individual stations within clusters do not show many similarities, regional patterns are evident. Southern stations are located near river systems and closer to deep waters, whereas northern stations in the Wadden Sea are situated in shallow waters and tidal flats. Stations such as Petten Zuid and Den Helder show similarities to the nearby western Wadden Sea stations. This showed that regional patterns have a greater influence than localized conditions.

The Dutch coast is primarily influenced by the M2 and S2 tidal constituents. Analysing the M2 phase and amplitude (Fig. 4c and d) revealed that variations aligned closely with the clusters when k = 5. Although the clusters were based on the NTR time series, tidal information was still present, as the tidal phase can be affected by depth changes due to the surge, and vice versa. Tide-surge interactions were evident in the timing of NTR maxima relative to high tides (Fig. 4a). Stations in the northeast Wadden Sea displayed a bimodal distribution of peaks during rising and falling tides, while those in the west Wadden Sea showed a trimodal pattern with peaks during rising, falling, and near high tide. At Petten Zuid, peaks were most frequent during rising tides near high tide, a pattern also seen at Scheveningen and Hoek van Holland, but with more frequent peaks during falling tides. Along the southern Dutch coast, peaks predominantly occurred during rising tides near high tide. Therefore, clusters were likely formed due to the regional bathymetry influence on wave and tidal propagation, affecting NTR.

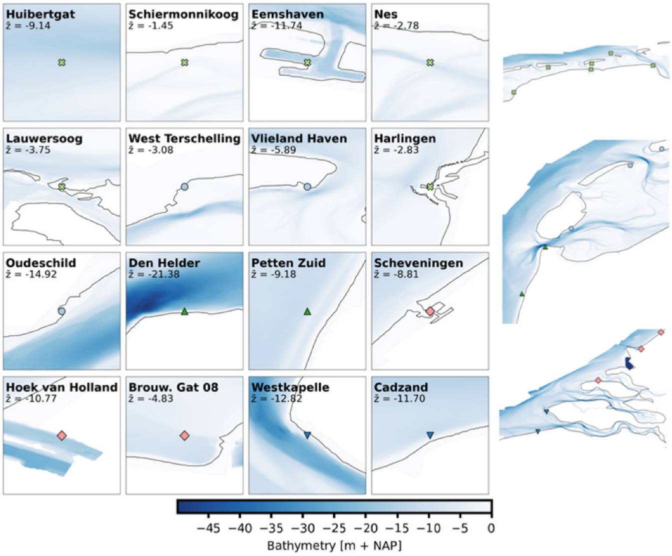

Fig. 3. *Left:* A close-up view showing the location and bathymetry of the sixteen tide gauge stations, with the indicated mean bathymetry ($\overline{z}$) of the area. *Right:* The Dutch coast divided into three regions based on the k = 3 clustering, with tide gauge symbols and colours representing the clusters for k = 5.

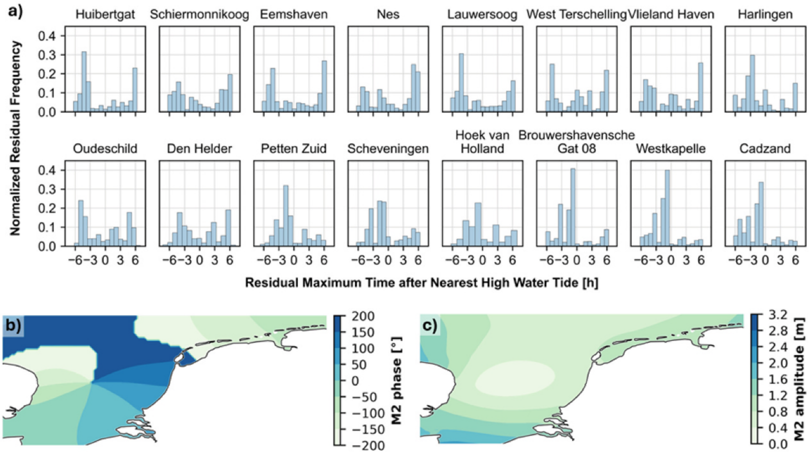

Fig. 4. **a)** Timing of NTR events maximum relative to the nearest high water tide, **b)** M2 phase and **c)** amplitude in the North Sea, as derived from the FES2014 global tide dataset (Lyard et al. 2016).

4 Conclusion

This study analysed the spatial patterns of elevated non-tidal residuals (NTR) along the Dutch coast. A time series clustering technique was employed on NTR time series for three scenarios (k = 3, 4, and 5). The optimal number of clusters was determined to be k = 5, highlighting the importance of finer-scale approaches for understanding regional extremes.

Regional bathymetry and tide-surge interactions emerged as important factors shaping the observed NTR spatial patterns. For example, southern stations near river systems and deeper waters exhibited distinct characteristics compared to northern stations in the Wadden Sea.

While this study focused on spatial clustering, the integration of spatio-temporal patterns is essential for comprehensive coastal management. Thus, future work will explore joint patterns of sea level hydrographs to inform more effective coastal management strategies. Additionally, future research should analyse other environmental drivers as well, such as wind and storm tracks. Drivers could also be incorporated into clustering methodologies to better capture the nuanced spatio-temporal patterns.

Acknowledgements. This research was supported by the Dutch Research Council (NWO) OCENW.M.21.109 'CHANCE: Climate cHange impActs on extreme sea levels iN Coastal watErs' project.

References

1. IPCC (2021) Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press
2. Kron W (2013) Coasts: the high-risk areas of the world. Nat Hazards 66:1363–1382
3. de Valk CF, van den Brink HW (2023) Update van de statistiek van extreme zeewaterstand en wind op basis van meetgegevens en modelsimulaties. KNMI number: TR-406
4. Ganske A, Fery N, Gaslikova L et al (2018) Identification of extreme storm surges with high-impact potential along the German North Sea coastline. Ocean Dyn 68:1371–1382
5. Groeneweg J, Beckers J, Gautier C (2011) A probabilistic model for the determination of hydraulic boundary conditions in a dynamic coastal system. Coastal Eng Proc 1
6. Haigh ID et al (2021) GESLA Version 3: a major update to the global higher-frequency sea-level dataset. Geosci Data J 10(3):293–314
7. Wiegmann E, Perluka R, Oude Elberink S, Vogelzang J (2005) Vaklodingen: de inwintechnieken en hun combinaties, Report AGI-2005-GSMH-012, Rijkswaterstaat, Adviesdienst

8. Bowman W (2014) UTide Python package
9. Pupić Vurilj M, Antolínez JAA, Muis S, Morales Napoles O (2024) Storm surge hydrographs from historical observations of sea level along the Dutch coast in the North Sea. Submitted