Natural Hazards (2026) 122:196
https://doi.org/10.1007/s11069-025-07906-9

ORIGINAL PAPER

Tail dependence of surge height and wind speed along the Dutch coast for storm clusters from large simulated datasets

Paulina E. Kindermann[1,2] · **José A. Á. Antolínez**[1] · **Oswaldo Morales-Nápoles**[1]

Received: 16 May 2025 / Accepted: 19 December 2025

Abstract
This study explores the statistical dependence between wind speed and surge height along the Dutch coast using a large synthetic dataset. Storms were clustered based on wind direction, tidal offset, wind rotation, tidal peak, surge and wind exceedance duration, resulting in 16 clusters per wind direction and per location. Apart from wind direction, comparing clusters revealed a limited impact of clustering based on these storm characteristics on the choice of the best-fitting copula model, suggesting sub-clustering may not be necessary for accurately representing the statistical dependence between extreme wind speeds and surge heights. The BB8 copula generally provided the best fit to the data. However, the observed upper tail dependence did not decrease to zero, particularly for western to northern wind directions, indicating non-negligible dependence in joint extremes of wind speed and surge height. Therefore, applying the BB8 copula (or any other copula model without upper tail dependence) may lead to underestimation of the flood risk, when applied in probabilistic analyses. The findings from this study provide valuable insights for refining hydraulic load models for reliability assessments and design of flood defenses.

Keywords Upper tail dependence · Storm surges · Copulas · Flood safety

✉ Paulina E. Kindermann
p.e.kindermann@tudelft.nl

José A. Á. Antolínez
J.A.A.Antolinez@tudelft.nl

Oswaldo Morales-Nápoles
O.MoralesNapoles@tudelft.nl

[1] Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands

[2] HKV Lijn in Water, Lelystad, The Netherlands

Published online: 21 February 2026

1 Introduction

The Netherlands consists of approximately 800 km of coastline, characterized by ca. 300 km of dunes, 460 km of dikes and six coastal barriers, to protect the low-lying, densely populated hinterland, as illustrated in Fig. 1 (HWBP 2024). The safety and reliability of these flood defenses is crucial, since the Dutch coast is affected by extratropical cyclones crossing over the North Sea, that, on average every year, cause extreme waves and storm surges over 1.5 m (Rijkswaterstaat 2024). Since 2017, the safety standards for flood defenses are defined in terms of required annual probabilities of flooding. These safety standards are based on a risk approach (probability of flooding times consequences), which translates to expected consequences. Since consequences are not equal everywhere, due to e.g. land elevation and accessibility for evacuation, the required annual probability of flooding varies spatially (Kok et al. 2016). For the Dutch coast, it means that flood defenses should withstand sea levels and wave heights with a return period varying between 300 to 1,000,000 years (Jonkman and Schweckendiek 2015), as indicated in Fig. 1.

According to current, Dutch guidelines for design and reliability assessments of flood defenses, these design conditions with large return periods were derived from statistical extrapolation of observations (Chbab and de Waal 2017). Wind speeds and sea levels were translated to wave conditions at the toe of the dike using SWAN model simulations, cali-

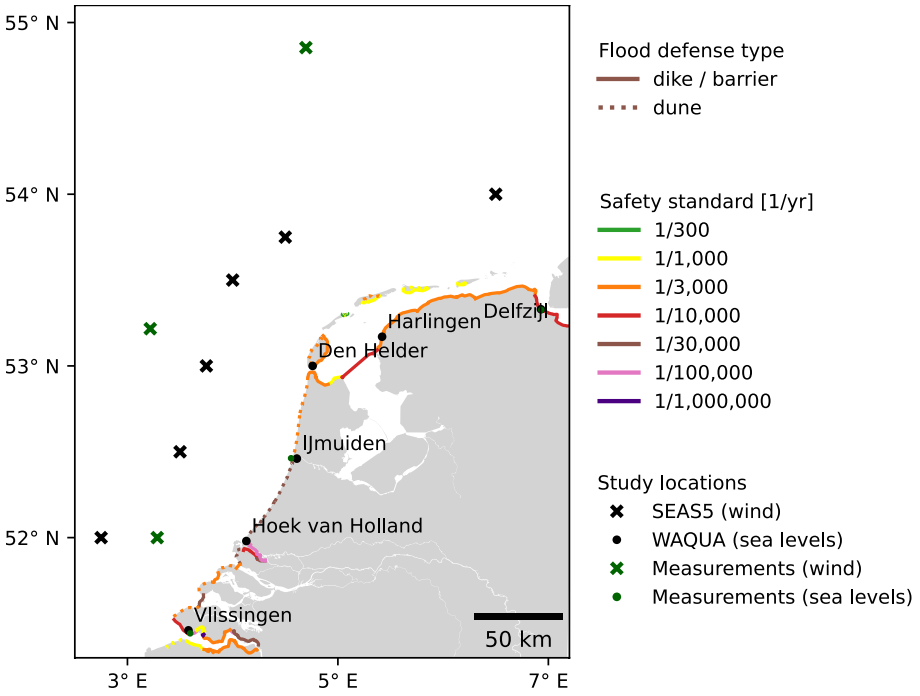


Fig. 1 Map of coastal flood defenses along the Dutch shore, where the line style indicates the type of flood defense and the color indicates the safety standard, expressed in a required maximum annual probability of flooding [1/yr] (HWBP 2024). The selected study locations are indicated by black and green crosses (SEAS5 and measured wind data respectively) and black and green dots (WAQUA-DCSM5 and measured sea level data respectively)

brated with observations (Gautier and Groeneweg 2012). Sea levels are a superposition of astronomical tide and storm surge. Since wind speeds and sea levels are physically linked, the current design guidelines consist of a correlation model for these two variables, that was again fitted to observations. Besides, assessment procedures require information on durations of hydraulic loads as input, because failure of coastal flood defenses is dominated by time-dependent erosion processes. Trapezium-shaped storm surge hydrographs were derived based on observations, with fixed base and top duration and fixed timing with respect to the tidal signal, depending on the location (Chbab 2015).

Despite the availability of about 100 to 150 years of observations in the Netherlands (Rijkswaterstaat 2024), the derivation of these design conditions—with return periods of more than 1000 years—introduces uncertainties due to limited data length relative to target return periods, data inhomogeneities, and missing storm observations (Haigh et al. 2023). Besides, the current guidelines rely on assumptions, such as presuming that a 1000-year storm event typically has the same storm duration as more frequent events, which cannot be verified with the limited data length. Observations of extreme events are scarce by definition, while natural variability of storms in the North Sea is large (Pupić Vurilj et al. 2025). To illustrate, the devastating storm surge in 1953 lasted for about 50 h, while storm Pia, that initiated automated closure of the Maeslant barrier in December 2024 for the first time in history, only lasted for about 20 h (Deltacommissie 1961; Rijkswaterstaat 2024). Due to the scarcity of data for such extreme events, relatively large uncertainties are incorporated into the return values of wind speeds and sea levels, leading to conservative design conditions. A representative correlation model for wind speeds and sea levels is difficult to determine for the joint upper tail, and representative values for storm characteristics—such as storm durations, and the timing relative to tides—are in practice often treated as deterministic values.

Recent advances in numerical weather prediction enable reanalysis and hindcast datasets to emerge as a standard alternative to observational data (Bauer et al. 2015). These datasets provide continuous atmospheric and oceanographic information, spanning more than 60 years and providing detailed information of storm characteristics. Atmospheric reanalysis is being used to analyze storm activity (e.g., Krieger et al. 2024; Karwat et al. 2022; Meyer and Gaslikova 2024) and to generate wave and water level reanalysis (e.g., Muis et al. 2023; Dullaart et al. 2023; Camus et al. 2024). These hydrodynamic reanalysis datasets have great potential to contribute to the understanding of extreme storms, and their natural variability in both space and time, particularly in regions where observational data are not available. This understanding is crucial for deriving realistic boundary conditions for design of flood defenses and other coastal infrastructure. For example, Dullaart et al. (2023) have developed a method to generate a global dataset of storm tide hydrographs, based on time series of storm surges and tides derived from the Global Tide and Surge Model (GTSM) forced with the ERA5 reanalysis wind and pressure fields, for the period from 1980 to 2017. Similarly, Camus et al. (2011) utilized wave reanalysis data from the SIMAR-44 database (1958–2001), to statistically describe the multivariate wave climate in Galicia, Spain. However, the utility of reanalysis datasets is limited by the quality and time span of atmospheric reanalysis (Bauer et al. 2015). Therefore, they do not fully address the challenge, as they still rely on approximately 60 years of data to extrapolate beyond 1000-year return periods, which introduces uncertainty.

Together with improved atmospheric reanalysis (National Center for Atmospheric Research Staff 2022; Hersbach et al. 2017), short, mid, and recently long range forecast-

ing have been made available, the latter providing operational long-range seasonal forecast data, such as the seasonal forecast system (SEAS5) by the European Centre for Medium-range Weather Forecasts (ECMWF) (ECMWF 2021). These long-range forecasts offer an opportunity to overcome the challenge of limited datasets by supplying thousands of climate realizations, which can help reduce uncertainty in the estimation of extreme events with return periods exceeding 1000 years. van den Brink (2020) has recently used wind and pressure fields from SEAS5 to produce a large dataset of simulated sea levels for the North Sea, using the Dutch Continental Shelf Model (DCSM) model. The potential of using this extensive dataset to derive higher-precision estimates for the return levels of extreme wind speeds and sea levels, is currently being explored by the KNMI (the Royal Netherlands Meteorological Institute) and Rijkswaterstaat, which is the executive organization of the Ministry of Infrastructure and Water Management (De Valk and Van den Brink 2023). In De Valk and Van den Brink (2024), it is argued that even if the wind stress from present-day weather prediction models slightly deviates from observations, these data are suitable for estimating the shape of the upper tail of the distribution function of wind stress, and that this extends to simulated data of water level along the Dutch coast. The study proposes an update of return levels for wind speed and sea level for several coastal stations, conditional on wind direction and omnidirectional, with reduced standard deviations. Additionally, a Gaussian copula is proposed to model their mutual correlation, with the parameter fitted per wind direction (De Valk and Van den Brink 2023). As mentioned before, this correlation is needed to derive the joint distribution of wind speed and sea level, which is then used to model the wave statistics at the toe of the dike.

From this extensive dataset by KNMI, not only the one-dimensional marginal statistics of wind speed and sea level can be updated. It also enables us to study other storm characteristics in more detail such as the duration and timing of the storm surge with respect to tide (Caspers et al. 2025). The study by De Valk and Van den Brink (2023) focuses on the characterization of one-dimensional margins and correlations conditional on wind direction, the effect of these other storm characteristics have not been studied. Using the same dataset by KNMI, Geerse et al. (2022) proposed a new, stochastic hydraulic load model for the reliability assessments of Dutch coastal flood defenses. This model includes a stochastic description of storm durations and the timing with respect to tide, and their effect on the marginal distributions of wind speed and *surge height* (instead of sea level). The correlation between wind speed and surge heights was analyzed and the study proposed a Gumbel copula, with parameters fitted per wind direction, in contrast to the Gaussian copula as proposed by De Valk and Van den Brink (2023). However, both studies do not evaluate any other copula models except for the Gaussian or Gumbel copula and also both studies only conditionalize on wind direction before fitting copula models to the storm data. Diakomopoulos et al. (2024) have studied other copula models, such as the BB1 copula, for modeling the correlation between extreme surges and peaks in wind, but the effect of other storm characteristics on the copula model that best fits the data has not been studied. In Geerse and Diermanse (2006), however, other copula models (Gaussian, Gumbel, Frank and Clayton) were evaluated, for observational data of wind speed and sea level at Hoek van Holland for the period of 1979 to 2002 and wind direction 330°N. The Gumbel copula was found to be the best-fitting model to describe the correlation, which was the reason to apply the Gumbel copula in Geerse et al. (2022) for the simulated data. However, only one wind direction was

studied and the limited amount of observational data points did not contain any events in the failure domain of coastal flood defenses.

Summarizing, the effects of other storm characteristics (such as storm duration and the timing between high tide and storm peak) on the correlation structure of wind speed and sea level have not been studied, due to the limited amount of observational data. Besides, only two copula models (Gaussian and Gumbel) were evaluated for modeling the correlation between wind speeds and sea levels or surge heights using the extensive simulation dataset by KNMI. With these long time series available, it is now possible to cluster storm events based on multiple storm characteristics and to evaluate and compare joint distributions for each cluster. Therefore, the current study follows up on the two aforementioned studies by De Valk and Van den Brink (2023) and Geerse et al. (2022). The objective is to investigate the statistical dependence between wind speed and surge height for different storm types (based on characteristics), that are relevant for extreme sea levels along the Dutch coast. The large dataset by KNMI of simulated sea levels and wind speed is used to select extreme storm events for the analysis. Storm types are clustered and for each cluster the statistical dependence between wind speed and surge height is evaluated using multiple copula models. The study reveals whether conditionalizing on other storm characteristics than wind direction, affects the correlation structure between wind speed and surge height. These insights can potentially contribute to future improvements of the stochastic description of hydraulic loads for reliability assessments of coastal flood defenses along the North Sea shore (Bakker et al. 2025).

2 Data and methods

Long-range seasonal forecast data of wind and sea levels were used (Sect. 2.1), from which relevant storm events were selected, using peaks-over-threshold. Each selected storm event consists of the temporal evolution of wind speed, wind direction, astronomical tide, sea level and residual surge. Each storm event was parametrized using several storm characteristics, like the maximum wind speed during the storm and the exceedance duration of the surge height above a relative level (Sect. 2.2). Then, storm events were clustered, using some of the storm characteristics (Sect. 2.3). For each storm cluster, the statistical dependence between maximum wind speed and maximum surge height during the storm event was analyzed, with a focus on investigating upper tail dependence (Sect. 2.4.1). Several copula models were fitted to the data corresponding to each cluster to determine the best fit (Sect. 2.4.2). Besides, the marginal distributions of wind speed and surge height are analyzed for each cluster. An overview of the method is given in Fig. 2.

2.1 SEAS5 and WAQUA-DCSMv5 simulation data

For the wind, the SEAS5 dataset by ECMWF is used (ECMWF 2021). The seasonal forecasts are initiated on the first of every month since November, 2017, and are run for the upcoming seven months. Each ensemble consists of 51 members. SEAS5 is a global model with a spatial resolution of 35 km and a temporal resolution of 6 h. For quantifying the bias and assessing the skill of the seasonal forecasting system, re-forecasts (also known as hindcasts) were produced for the period from 1981 until 2016. The hindcasts consist

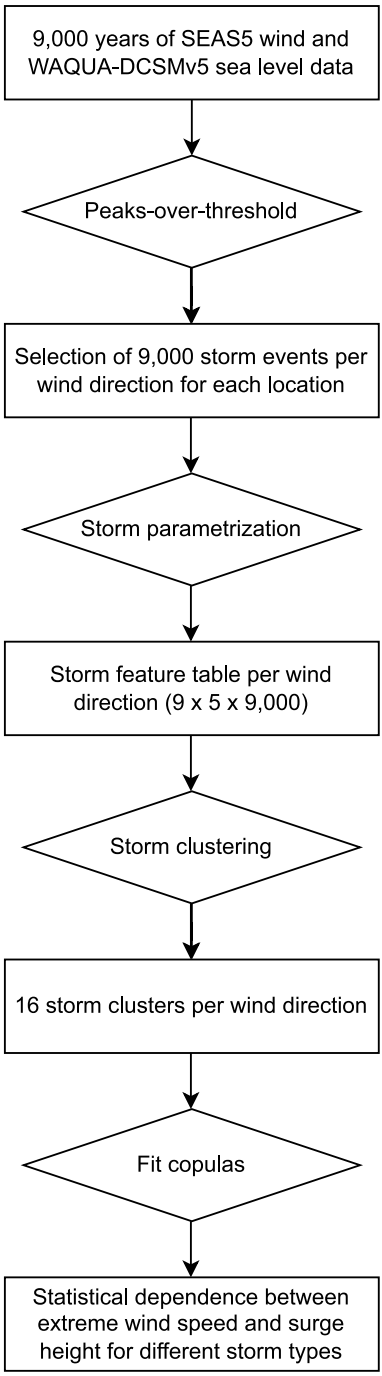


Fig. 2 Overview of the methodology, in which rectangular boxes illustrate data and diamond-shaped boxes illustrate methods

of 51 ensemble members for the months February, May, August and November, and 25 members for the other months. Biases were removed by considering only model anomalies with respect to the model mean state. Specifically, the values of the forecast ensemble are compared to the values of a climate reference ensemble, and the differences between model forecast and model climate were assessed and plotted (ECMWF 2021).

From a statistical perspective, each ensemble member represents a possible climate scenario, enabling their combination into a single, extensive dataset. With the years 1981 to 2017 consisting of 404 ensemble members each ($4 \times 51 + 8 \times 25$) and the years 2018 to 2023 consisting of 612 ensemble members each, it adds up to 18,620 ensemble members, with a length of seven months each. To prevent mutual dependence among ensemble members, the first month of each ensemble is excluded, which results in a dataset with a length equivalent to 9310 years (De Valk and Van den Brink 2023, 2024; Benito et al. 2025).

Subsequently, KNMI has run the WAQUA-DCSMv5 model, with the 6-hourly mean sea level pressure and 10-m wind stress from each SEAS5 ensemble member as input, and the corresponding sea levels in the North Sea as output. The astronomical tide at the open boundaries of the model are derived by harmonic expansion using ten harmonic constituents. The so-called Dutch Continental Shelf Model (DCSM) is a numerical model that solves the two-dimensional shallow-water equations on a $\frac{1}{8}^\circ \times \frac{1}{12}^\circ$ spatial grid (Gerritsen et al. 1995). DCSM, initially developed in the 1980s (Verlaan et al. 2005), is now in its sixth generation being used for operational forecasting along the Dutch coast. Due to the increased computational demands of newer versions, De Valk and Van den Brink (2023) have used the fourth generation of the model (WAQUA-DCSMv5) to generate the sea level dataset corresponding to the SEAS5 ensemble members.

2.1.1 Data comparison

In recent years, KNMI has validated the dataset, using DCSMv6 (with a five times higher resolution than DCSMv5), a different climate model (HARMONIE, with a higher spatial and temporal resolution than SEAS5), and observational data. As a result, a bias correction of 10% increase was applied for the SEAS5-wind stress input for WAQUA-DCSMv5. After bias correction, KNMI demonstrated that, combined with observations, the SEAS5-WAQUA-DCSMv5 dataset enables robust and reliable estimation of extreme return values of wind speed and sea level, with reduced statistical uncertainty thanks to the large amount of data (van den Brink 2020; De Valk and Van den Brink 2023, 2024). Similar conclusions were reported by Benito et al. (2025). A detailed validation of extreme quantiles was already performed by De Valk and Van den Brink (2024) and is therefore not repeated here.

However, the upper tail dependence between wind speed and surge height has not yet been explicitly compared to observational data. To address this, we compare empirical tail dependence coefficients (see Sect. 2.4 for the definition of upper tail dependence) from pairs of extreme sea level and wind speed from observations at three coastal stations with those from the SEAS5-WAQUA-DCSMv5 dataset. Sea level measurements are taken from tide gauges at Vlissingen, IJmuiden and Delfzijl (indicated by green dots in Fig. 1), and wind measurements are taken from nearby off-shore platforms that have the longest measured record: Europlatform, K13, F3-FB-1, respectively (indicated by green crosses). The overlapping observation period spans 1996–2023 for Vlissingen and IJmuiden and 1994–2023 for Delfzijl. For each location, we selected approximately 50 storm events with the most

extreme sea levels, together with the concomitant wind peaks, providing a balance between retaining a sufficient sample size and focusing on the most extreme events relevant for tail dependence analysis. For this, sea level thresholds of NAP+3.1 m, NAP+2.05 m and NAP+2.9 m were applied for Vlissingen, IJmuiden and Delfzijl, respectively. We selected wind speed and sea level pairs from the SEAS5-WAQUA-DCSMv5 dataset using the same thresholds. Fig. 3 presents the empirical tail dependence coefficients for wind speed and sea level pairs from: (i) the observations (1996–2023, in blue), (ii) SEAS5 data (1981–2023, in orange), and (iii) SEAS5 data for the overlapping period (1994–2023 or 1996–2023, in green). The shaded area indicates the 95%-Wald asymptotic confidence interval (as explained in Appendix C.2.1).

All datasets show non-zero tail dependence coefficients for quantile levels $t \approx 0.75{-}0.95$ ($t - 1 \approx 0.25{-}0.05$), indicating a clear degree of upper tail dependence across all datasets. For most thresholds, the tail dependence coefficients derived from SEAS5 lie within the 95%-confidence interval of those estimated from observations, suggesting that the SEAS5-WAQUA-DCSMv5 dataset adequately captures the dependence between wind speed and sea level. Furthermore, these results emphasize the need for a long synthetic dataset, such as SEAS5-WAQUA-DCSMv5, for a robust estimation of tail dependence between extreme wind speed and surge height: estimates based on observations are highly unstable and uncertain, and tend to approach zero beyond $t \approx 0.95$, due to the limited number of extreme events in the observational record.

In addition, we compared histograms of three key storm characteristics between observations and the synthetic dataset, as shown in Appendix A. The wind and surge duration histograms are generally comparable, whereas the tidal offset histograms differ more noticeably. In both the observations and the SEAS5-WAQUA-DCSMv5 dataset, the tidal offset histograms show a non-uniform distribution with peaks at specific tidal offset values. However, the values of these peaks differ between the two datasets, which may be attributed to uncertainties in the tidal signal extraction and tide-surge interactions.

2.2 Storm selection and parametrization

For the current study, we selected six locations along the Dutch coast for the analyses: Vlissingen, Hoek van Holland, IJmuiden, Den Helder, Harlingen and Delfzijl. Sea level time series were extracted from the closest WAQUA-DCSMv5 output grid cells, and the

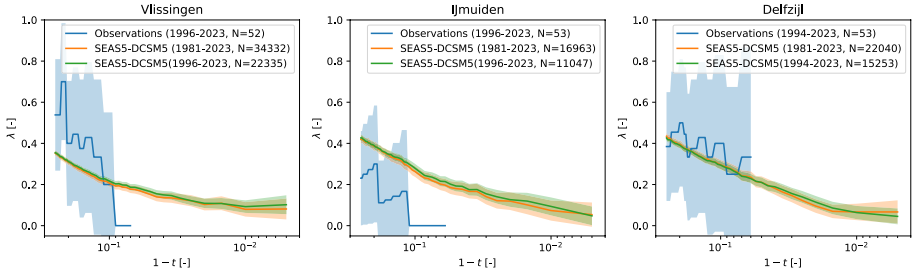


Fig. 3 Comparison of empirical upper tail dependence coefficient λ for different sample fractions t, for pairs of extreme sea level and wind speed, including a 95%-confidence interval. Different colors indicate the different datasets. Each panel represents one location. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than three

wind time series were selected from nearby SEAS5 grid cells. The SEAS5 grid cells were selected about 50 km from the coastline, to ensure that the surroundings are not affected by land surface. The selected SEAS5 locations are indicated as black crosses and the WAQUA-DCSMv5 locations as dots in Fig. 1. The selected combinations of wind and sea level data locations are consistent with locations that were selected by De Valk and Van den Brink (2023) and Caspers et al. (2025).

Subsequently, we translated the wind stress to wind speed, using a Charnock constant of 0.025. This value was chosen based on comparisons with observations in literature (Fig. A.2 in Geerse et al. (2022) and Fig. 6 in De Valk and Van den Brink (2023)). The 6-hourly wind speed time series were interpolated to a 10-min time step using Spine interpolation. The wind direction was calculated from the wind speed vectors and we derived the residual surge time series, being the difference between the sea level and the astronomical tide.

Since the wind speed is one of the main drivers of extreme sea levels, we selected storms from the SEAS5-DCSMv5 time series using a peaks-over-threshold (POT) based on wind speed. A window of 1.5 days before and after each selected wind speed peak was applied to ensure independence of the selected storm events. This choice follows from the current Dutch design guidelines, which prescribe storm durations between 44 and 54 h (i.e., 22 to 27 h before and after the peak), depending on the location (Chbab 2015). These durations are based on analyses of historical storm events. A comparison between observed data and SEAS5 simulations indicated slightly longer storm durations in SEAS5 (see Appendix A); therefore, the window was extended to ensure the independence of events. The threshold was determined iteratively, conditional on the wind direction during the peak wind speed, such that 9000 storms were selected for each wind direction and each location, consistent with the approach in Caspers and Kindermann (2023) and Geerse et al. (2022). As explained in Sect. 2.1, the length of the dataset is approximately 9000 years, and therefore the selection roughly corresponds to storm events with a return period of one year, conditional on the wind direction. The impact of the chosen threshold was evaluated by comparing scatter plots of wind speed-surge height pairs obtained using an alternative threshold (see Appendix B). The comparison indicates that the upper-tail correlations are very similar for both thresholds, suggesting that the results are robust to the specific threshold choice.

The wind directions are defined as 16 directional sectors (*N, NNE, NE,... , NNW*) with a width of $22.5°$, for example *W* corresponds to $270 \pm 11.25°$ with respect to North. The wind speed thresholds that resulted in the selection of 9000 storm events per wind direction are indicated in Fig. 4, by the length of the radial bars. Each radial bar indicates a wind direction sector and the color of the bar indicates the mean of the maximum (skew) surge height during the storm events corresponding to the directional sector. The figure reveals that wind speeds are generally higher for southern to western wind directions, while western to northern wind directions drive the higher storm surges.

For each selected storm event, the time series of wind speed, wind direction, sea level, residual surge and astronomical tide were extracted, for a time period of three days around the selected peak. Since the wind speed is the selection variable in the POT, the time index is defined such that the selected wind speed peak occurs at time step zero. The time series of each storm event were parametrized using the following storm characteristics: water level w, residual surge height s, skew surge height k, tidal peak a, tidal offset φ, surge duration d_s, wind speed u, wind direction r, wind duration d_u and wind rotation ψ. An example of one selected storm event is presented in Fig. 5, including these storm characteristics. The time

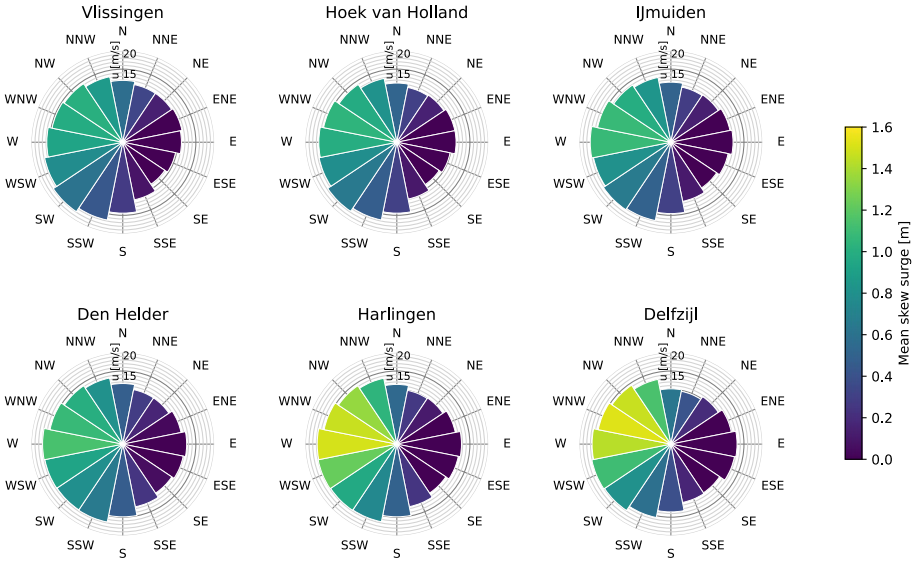


Fig. 4 Wind speed thresholds and mean surge heights per wind direction sector and location. The length of the bins illustrates the POT-threshold of the wind speed per wind direction and the color illustrates the average of the skew surge during the storm events, corresponding to the direction sector

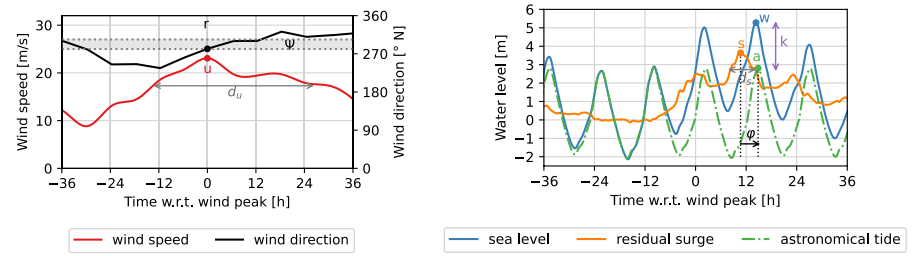


Fig. 5 Example of a selected storm event, including definitions of storm characteristics, such as the tidal offset φ and the wind duration d_u. Modified from Caspers et al. (2025). **a** Temporal evolutions of wind speed (red line) and wind direction (black line). **b** Temporal evolutions of wind speed (red line) and wind direction (black line)

series of wind speed are parametrized by the maximum wind speed u during the storm event (i.e. the selected peak in the POT), in meters per second, and by the duration d_u that the wind speed exceeds the 75%-level of u, in hours. The wind direction time series is parametrized by the wind direction r at time step zero in degrees with respect to North and by the wind rotation ψ. ψ is defined as the duration in hours that the wind direction resides within the corresponding wind direction sector during the storm event. The maximum sea level w is expressed in meters with respect to the local Dutch reference level (Normaal Amsterdams Peil (NAP)). The maximum residual surge height during the storm event is s in meters, the corresponding skew surge height k in meters and the closest astronomical high tide is defined as a in meters with respect to NAP. The exceedance duration d_s is the duration in

hours that the residual surge exceeds the 75%-level of s. The tidal offset φ is the time difference in hours between the maximum surge height s and the closest astronomical high tide t, and is defined positive if the surge peak precedes the tidal peak (φ is positive in the example of Fig. 5). Correlations and interactions exist between many of these storm characteristics, which play an important role in driving extreme sea levels and in the resulting shapes of the storm hydrographs (Caspers et al. 2025).

Not all storm characteristics are included as features in the clustering procedure. Since the objective is to investigate how different storm characteristics influence the dependence between extreme wind speed and surge height, these two variables are excluded from the clustering itself. Furthermore, w, s and k are highly correlated and therefore omitted to avoid redundancy. Since the wind direction is known to be a dominant factor determining the dependence between wind speed and surge heights and their marginal distributions, storms are pre-clustered by wind direction r, resulting in 9000 storms per wind direction, as mentioned earlier. Easterly wind directions (NNE,..., SSE) are excluded, as they do not generate storm surges in the North Sea basin. As a result, five features are used for clustering: φ, d_s, a, d_u and ψ, resulting in nine (one per wind direction) $(N \times n)$-dimensional dataset, with $N = 9000$ the number of selected storm events per wind direction, and $n = 5$ the number of features.

2.3 Storm clustering

Storms were clustered using unsupervised learning (K-Means (Jin and Han 2010), and the maximum dissimilarity algorithm (MDA)) on the storm characteristics. Other clustering techniques were explored as well ('manual' clustering, similar to the work in Caspers et al. (2025), Self-Organizing Maps and the Maximum Dissimilarity algorithm), based on storm characteristics and on the full time evolutions, similar to Pupić Vurilj et al. (2025), and resulted in similar conclusions. However, this paper focuses on the results using K-Means.

The K-Means algorithm divides the $(N \times n)$-dimensional storm data into K clusters. Each cluster j is defined by centroid $\vec{c_j}$ and contains the data points $\vec{x}$ that are closest to it, and $\vec{c}$ and $\vec{x}$ have the same dimensions. The clustering procedure starts with initialization of the centroids $\{\vec{c_1^0}, \ldots, \vec{c_M^0}\}$. Various methods are possible for initialization. In this study, we apply the maximum dissimilarity algorithm (MDA) to define the initialization vector (Kennard 1968). MDA is a method to select a representative subset of size K from the total dataset of N storms. The subset is initialized by selecting one data point from the total dataset. The rest of the $K - 1$ elements of the subset are selected iteratively, by calculating the dissimilarity between each remaining data point in the total dataset and the elements of the subset. The most dissimilar data point is then transferred to the subset. There are multiple ways to define the initial data point in the subset and the definition of 'most dissimilar'. In this study, we have used the Python-implementation of MDA by Camus et al. (2011).

Before clustering, the input data—consisting of the nine sets (one for each wind direction) of 9000 data points (storm events) characterized by five storm features—are normalized. Normalizing is required, because if features have different scales, features with larger magnitudes will dominate the clustering process.

In each iteration z, the K-Means algorithm calculates the Euclidean distance between each data point $\vec{x_i}$ and each centroid $\vec{c_j^z}$, and assigns every data point to the cluster j whose corresponding centroid is closest to it (in other words, $j = \min \|x_i - c_j^z\|, j = 1, \ldots, K$).

Subsequently, the centroid is updated to be the mean of subset C_j of the m_j assigned data points of the corresponding cluster j. Thus, in each iteration z, the centroids are moved such that the overall intra-cluster distance eventually converges and the subsets of data points belonging to each cluster are stable. An elaborate description about KMA can be found in Jin and Han (2010).

2.3.1 Cluster evaluation

We determined the optimal amount of clusters by evaluating the Within-Cluster Sum of Squares (WCSS) and the Between-Cluster Sum of Squares (BCSS). WCSS measures the variance within each cluster and therefore quantifies how close the data points are to their cluster centroid. BCSS measures the variance between cluster centroids and the overall mean of the dataset and therefore quantifies how well-separated the clusters are.

Figure 6 shows the WCSS and BCSS scores for a range of cluster sizes K in dashed and solid lines respectively, for each wind direction and all six coastal stations. Based on these graphs, we decided that 16 is an adequate amount of clusters, for which the within-cluster variability is sufficiently low, while the between cluster variability is large enough. This value is indicated by the red dotted line.

The resulting clusters that follow from the K-Means algorithm are analyzed in terms of the statistical dependence between maximum wind speed u and maximum (residual) surge height s of the storm events belonging to the cluster.

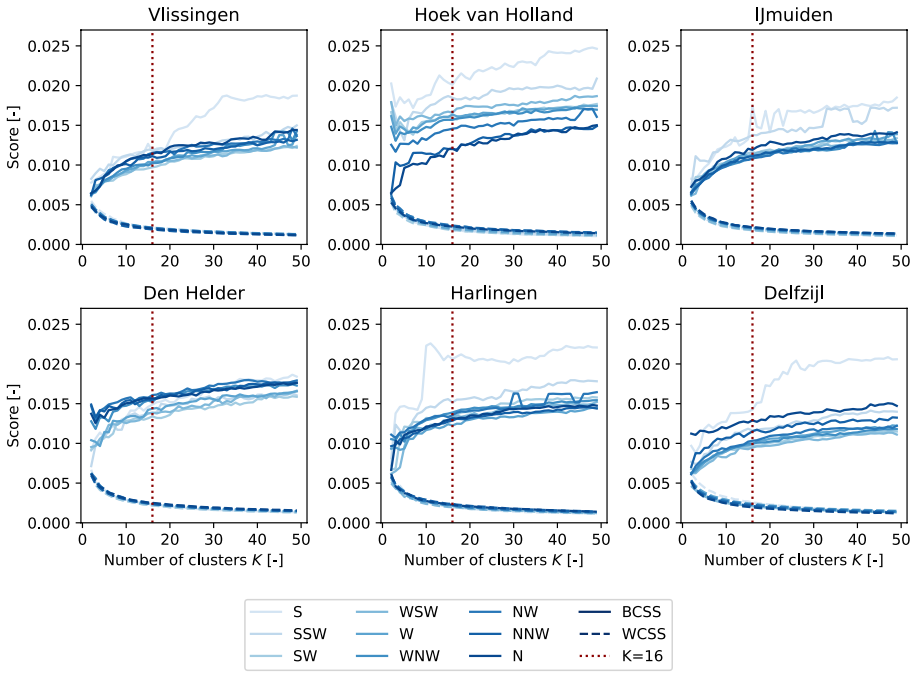


Fig. 6 Scores for WCSS (dashed lines) and BCSS (solid lines), for each wind direction, depicted in shades of blue. Each panel represents one of the six locations. The dotted red line indicates the chosen number of clusters, $k = 16$

2.4 Statistical dependence

For each of the 9x16 clusters, the upper tail dependence between maximum wind speed u and surge height s was analyzed (see Sect. 2.4.1). This analysis of upper tail dependence for the different clusters was used to gain a first impression of the relation between extreme wind speed and extreme surge height for different storm types (e.g. short versus long durations). Then, several copula models were fitted to the storm clusters, and their performance was evaluated using two goodness of fit tests, as elaborated in Sect. 2.4.2.

2.4.1 Upper tail dependence

The analysis is restricted to the bi-variate case. Loosely, tail dependence measures the strength of dependence between large values of two random variables. Specifically, it quantifies how likely it is for one variable to realize a large percentile given that another variable realizes that large percentile. As the objective of this study is related to flood risk assessment for the Dutch coast, we focus on upper tail dependence, which expresses the probability associated to the joint occurrence of extreme events. In particular, we focus on extremely high wind speed u and high percentiles of the concomitant peak surge height s. In this study, we evaluate two measures for quantifying the upper tail dependence.

Firstly, upper tail dependence can be quantified by computing semi-correlations, which are the correlations in the four quadrants of the standard normal transform of variable pairs (Z_1, Z_2). For upper tail dependence, ρ_U, the correlation in the upper right quadrant, is considered, which is defined as Pearson's correlation coefficient for (z_1, z_2)-pairs with $z_1 > 0$ and $z_2 > 0$.

Secondly, we compute the upper tail dependence coefficient, or extremal dependence coefficient λ_U. For two random variables (X_1, X_2), λ_U is defined as (Coles et al. 1999):

$$\lambda_U = \lim_{t \to 1} Pr\left(X_2 > F_2^{-1}(t) | X_1 > F_1^{-1}(t)\right) \tag{1}$$

where F_1 and F_2 are the cumulative distribution functions (CDF) of X_1 and X_2, F_1^{-1} and F_2^{-1} the respective inverse CDF and t is a threshold representing the selected sample fraction. Here, the non-identically distributed pairs (X_1, X_2) were transformed to uniform margins (U_1, U_2), using the empirical CDF $\hat{F}_1$ and $\hat{F}_2$ and t is taken arbitrarily close to 1. λ_U is hence computed empirically. $\lambda_U = 0$ suggests that large percentiles of the random variables (when $1 - t \to 0$) are unlikely to occur at the same time. This is often referred to as asymptotic independence, while $0 < \lambda_U < 1$ implies that there is some positive probability that large values of the two random variables under investigation can occur simultaneously. If variables are asymptotically dependent, incorrectly assuming an asymptotically independent model can lead to underestimation of the joint risk of extreme events (Coles et al. 1999), and vice versa. Upper tail dependence coefficients presented in this study are accompanied by Wald confidence intervals, to take into account the uncertainty due to limited sample size, as described in Appendix C.2.1.

2.4.2 Copula modeling

Copula functions are a tool to separate the specification of marginal distributions and the dependence structure. The copula approach for dependence modeling is based on Sklar's theorem, which, in the bi-variate case, states that if H is any joint distribution function of continuous variables X_1 and X_2, with uni-variate marginal distributions F_1 and F_2, then a so-called Copula exists, such that (Nelsen 2006):

$$H(x_1, x_2) = C\left(F_1(x_1), F_2(x_2)\right). \tag{2}$$

C is the copula to model the dependence between X_1 and X_2, which is defined in the unit square $I^2 = [0, 1]^2$. There are several types of copula families known in literature. Within this study, the following parametric copulas are considered: the Gaussian copula, Clayton (1978), Gumbel (1960), Frank (1979), Joe, BB1, BB6, BB7, and BB8 (Joe 1997), and the Student's t copula (e.g., Genest et al. 2007). Among these copulas, the Gumbel and Joe are most suitable to model upper tail dependence, while the Clayton copula allows for lower tail dependence. The Gaussian and Frank copulas have no tail dependence. The other copula models are more flexible and allow for both upper and lower tail dependence. A detailed overview of the mathematical properties of these copulas can be found in Nelsen (e.g., 2006); Salvadori et al. (e.g., 2007); Joe (e.g., 1997).

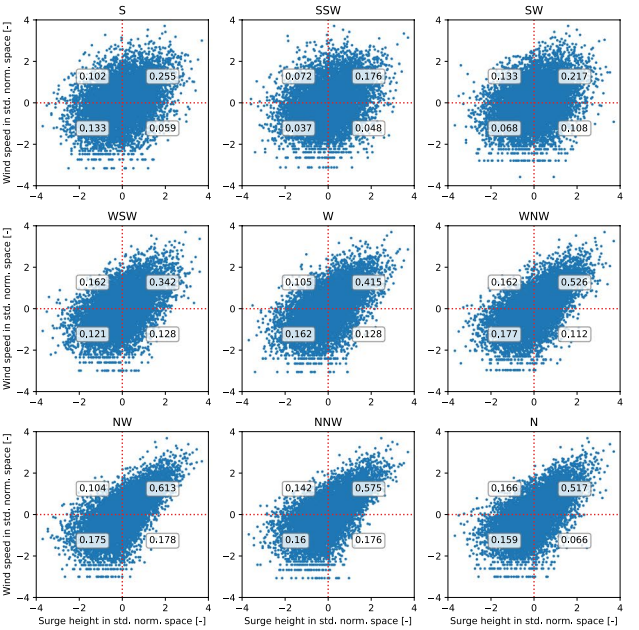
For each cluster, the copula model parameters are fitted to the corresponding pairs of u and s using the pyvinecopulib package in Python. Then, goodness of fit tests are applied to the candidate copulas, to verify their suitability in modeling the dependence structure of the wind speed and surge height pairs. Two methods are applied to identify the best copula fit for the data: the corrected Akaike Information Criterion (AIC) (Akaike 1998) and the Cramér-Von Mises statistics (Genest et al. 2009). Since the cluster size varies, the corrected AIC is applied, which corrects for sample size. These two metrics are used complementary to each other, since the Cramér-Von Mises evaluates global deviations between the empirical and fitted copula (i.e. *how well* a copulate fits the data), while the AIC compares the copula models with each other (i.e. *which* copula model fits best). Both metrics are briefly introduced in Appendix C.

Results are compared to the analysis of upper tail dependence in terms of the semi-correlations ρ_{NE} and the tail dependence coefficient λ. By evaluating different clusters and copulas, it becomes clear whether subdividing storms has any effect on the dependence between wind speed u and surge height s and thus on the choice of correlation model that is most suitable to model their dependence.

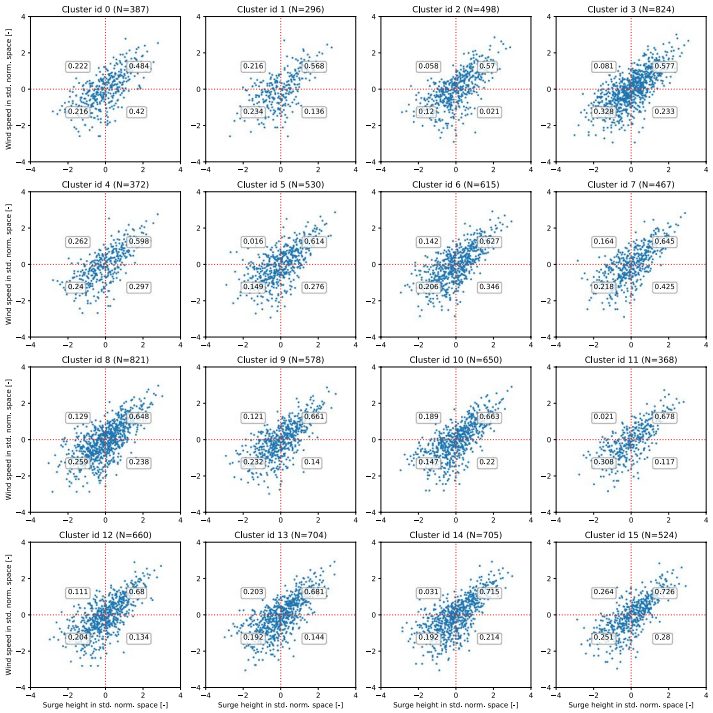
3 Results

3.1 Evaluation of statistical dependence

After clustering the storm events, the initial step involved analyzing the semi-correlations between the pairs of wind speed u and surge height s belonging to each cluster, with a focus on the upper tail dependence. First, Pearson's correlation coefficient was computed in each quadrant within the standard normal space. Figure 7a illustrates pair plots of u and s in



(a)



(b)

Fig. 7 Pair plots for surge height and wind speed in the standard normal space at location IJmuiden, including semi-correlations based on Pearson's correlation coefficient. **a** Per relevant wind direction. **b** Clusters within wind direction NW, sorted by Pearson's correlation coefficient in the upper tail, ρ_{NW}

the standard normal space for the nine relevant wind directions at location IJmuiden. The Pearson's correlation coefficient is shown for each quadrant. It is evident from the figure that the upper tail dependence increases from southern to northern wind directions, with the highest upper tail dependence for wind direction *NW* ($\rho_{NW} = 0.613$). These observations are consistent with physical expectations based on the North Sea's geography: when winds blow from the northwest, they drive water into the funnel-shaped basin of the North Sea, causing elevated surges along the shore.

For each wind direction, storms were clustered using K-Means, based on five storm features: the tidal offset φ, surge exceedance duration d_s, wind exceedance duration d_u, tidal peak a and wind rotation ψ. As mentioned, this resulted in 16 clusters per wind direction per location. The semi-correlations were computed for the pairs of u and s corresponding to each cluster. Figure 7b provides the pair plots for the 16 clusters corresponding to wind direction *NW* at IJmuiden, as an example. The upper tail dependence varies between 0.484 and 0.726 across clusters within this wind direction, however it appears more or less consistent. In order to see if we can identify some patterns between these correlations and the storm characteristics corresponding to a cluster, we look at Fig. 8. The first five panels correspond to the five clustering features (a, φ, d_u, ψ, d_s) with colors representing the centroid value of the respective feature. Then, we have added two panels for maximum wind speed u and surge height s, which were not used as clustering features. For these two features, colors

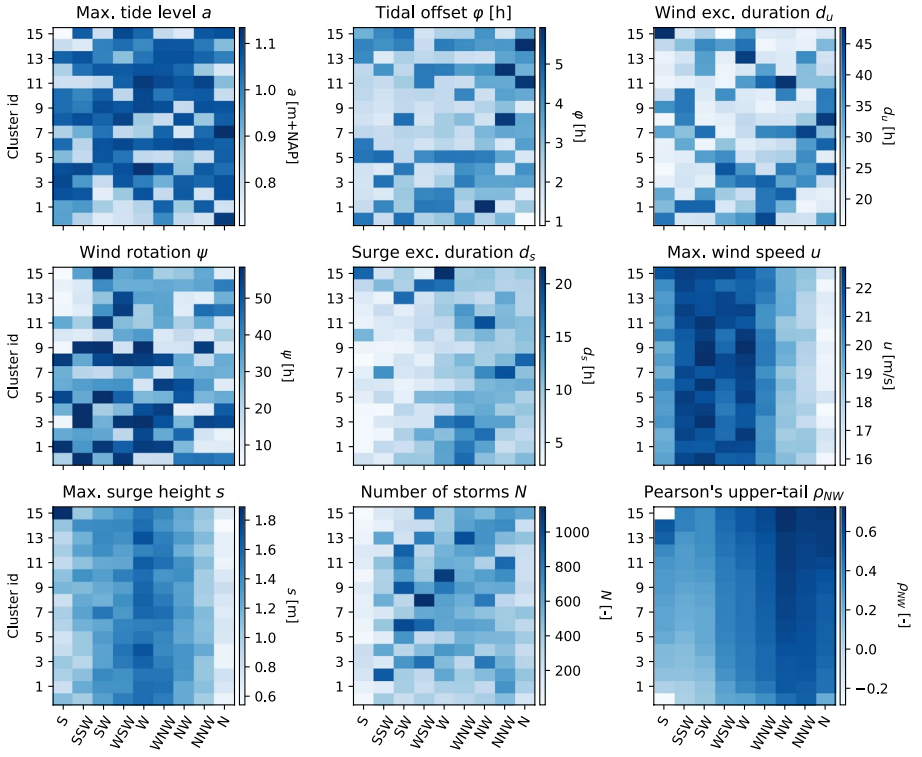


Fig. 8 Overview of the clusters for location IJmuiden. Each panel refers to one storm feature, the cluster size and Pearson's correlation coefficient. Each column in a panel shows the 16 clusters belonging to a specific wind direction. The color indicates the cluster's centroid value of the corresponding feature

indicate the mean values over the storm events belonging to each cluster. In addition, the number of storms N in each cluster is shown in the lower center panel. The lower right panel presents the upper tail Pearson's correlation coefficient ρ_{NW} for the pairs of u and s of the respective cluster. Within each panel, columns represent the nine relevant wind directions and rows are the clusters within the wind direction. Clusters are ordered by ρ_{NW} within the wind direction. So, for wind direction *NW* the values in the lower right panel correspond to the ρ_{NW}-values from Fig. 7b. The upper tail dependence appears more or less consistent across clusters within a wind direction, with a slight tendency towards stronger correlations for shorter wind durations d_u (upper right corner of the upper right panel). Also higher *absolute* tidal offsets φ (i.e. occurrence of the maximum surge height close to low tide, middle row of Fig. 7b) reveal slightly higher correlation (upper right corner of the upper center panel). The same figures are presented for the other locations in Appendix C.1. All locations show the strongest upper tail dependence for wind direction north-west, with Pearson's correlation coefficient $\rho_{NW} \in (0.53, 0.62)$, and consistency across clusters with slightly higher correlation coefficients for sub-clusters with short duration d_u and large absolute tidal offset φ, with Pearson's correlation coefficient up to 0.76.

The smaller number of storms in clusters associated with wind direction *S* can be attributed to the removal of events containing NaN-values for one or more storm features, prior to clustering. Southern storms often generate limited storm surge responses at the Dutch coast, making it impossible to define the surge duration d_s, which are thus set to NaN. This variation in cluster size is considered when assessing uncertainty in the subsequent tail dependence coefficients and copula model fitting, by presenting confidence intervals.

Second, the empirical upper tail dependence coefficient λ, as defined in Eq. (1), was calculated for pairs of u and s corresponding to each cluster, for various sample fraction thresholds t close to the upper limit. The results are shown in Fig. 9 for location IJmuiden. Each plot contains the results for the 16 clusters of the corresponding wind direction (depicted in shades of blue), plus results for all storms belonging to that wind direction (in black). λ is only displayed when there are at least five data points above the percentile threshold, which explains why some lines terminate earlier than others (due to variations in cluster size). The shaded areas indicate the 95%-Wald asymptotic confidence interval. Since the sample size decreases for increasing threshold t, the uncertainty of the estimated λ becomes larger/the confidence intervals become wider from left to right.

Consistent with the analysis of Pearson's correlation coefficient in the upper tail (ρ_{NW}), the upper tail dependence appears to be stronger for northern wind directions. In general, the λ-values of the clusters (blue lines) follow the same trend and are more or less in the same order of magnitude as for the total wind direction cluster (black lines). For all clusters, the upper tail dependence λ decreases for decreasing sample fraction (i.e. increasing values of u and s). Still, the value of λ is considerably higher than zero for the total western to northern wind directions, including confidence intervals, implying asymptotic dependence between wind speed and surge height. The same can be observed for other locations, as presented in Appendix C.

3.2 Comparison of copula models

Next, copula models were fitted to pairs of wind speed u and surge height s in the uniform space, for storm subsets corresponding to each cluster. Based on the findings from the pre-

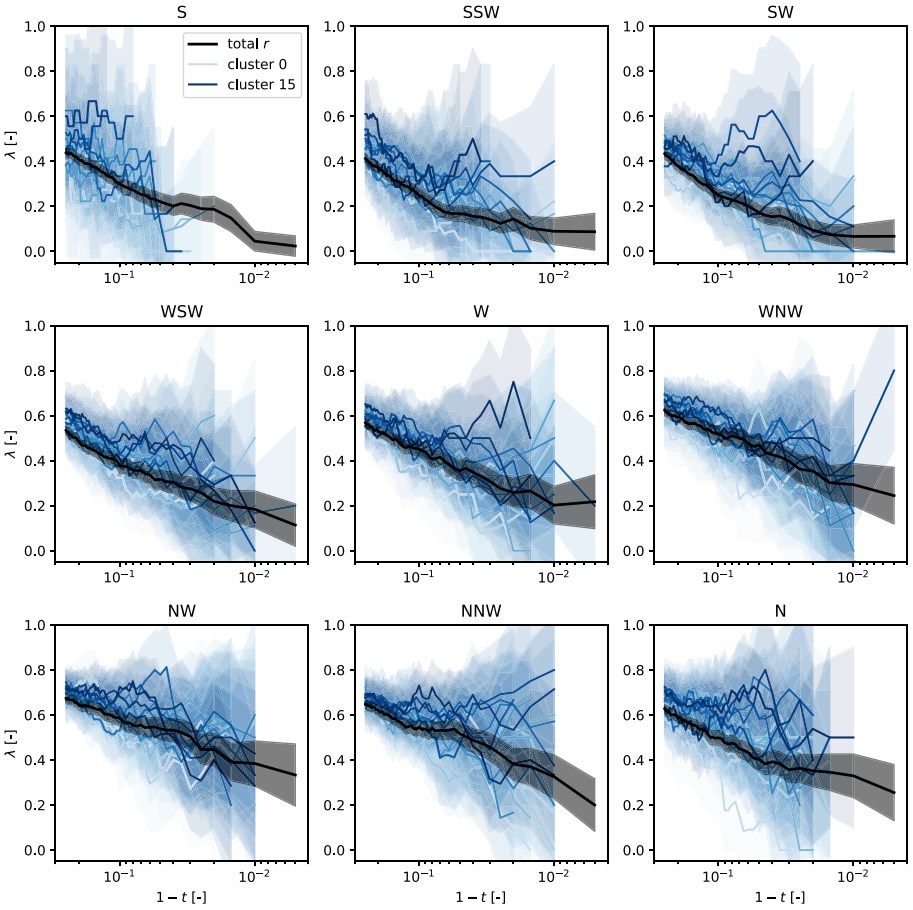


Fig. 9 Upper tail dependence λ for pairs of wind speed u and surge height s, for different sample fraction thresholds t, at location IJmuiden, with the x-axis being $1 - t$ on a log-scale. Each panel represents one of the relevant wind directions. In blue, λ is presented per cluster and in black for all storm events within the corresponding wind direction. The shaded bands represent confidence intervals based on Wald statistics. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

vious section, we expect that clusters associated with western to northern wind directions are best modeled by copulas exhibiting some degree of upper tail dependence, such as the Gumbel copula. Figure 10 presents the CvM scores for various copula models, for each cluster, organized by wind direction r. Again, the clusters are represented by shades of blue and the black line indicates the scores for the total storm set corresponding to a wind direction r. The lower the S_{CvM} the better. Note that the S_{CvM} is sensitive to sample size, which explains why the scores can deviate between clusters. However, since we want to compare various copula model fits for one cluster, this is not so relevant. The best-fit copula (i.e. lowest S_{CvM}) for each cluster is marked by a red (or black) square.

Overall, it is observed that, according to the Cramér–von Mises criterion, the BB8 copula is the best fit for all total wind direction storm groups (depicted in black in Fig. 10) and also for most clusters within each wind direction, at IJmuiden. Second-best scores the Gumbel

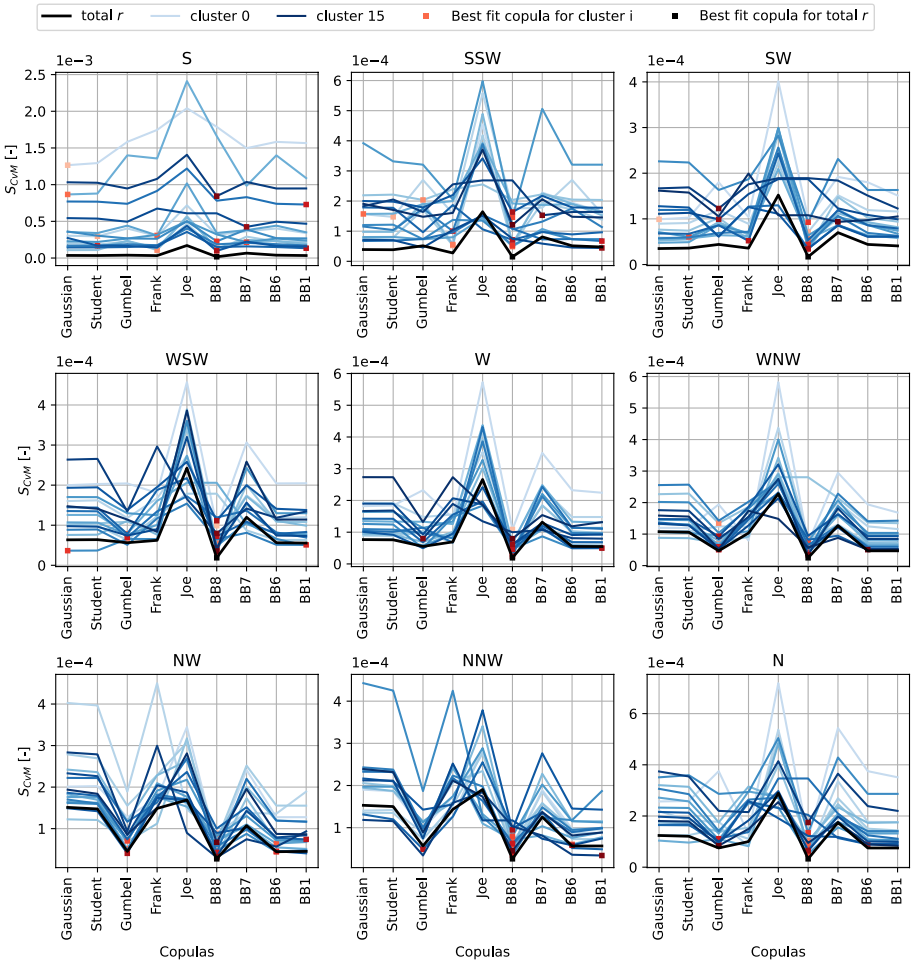


Fig. 10 Cramér–von Mises criterion for various copula models, fit to pairs of wind speed u and surge height s corresponding to the 16 storm clusters at location IJmuiden. Each plot shows the results for clusters corresponding to that wind direction. The blues indicate the 16 clusters and the black all storm events within the corresponding the wind direction. The squares indicate the copulas that best fit the corresponding data

copula for wind directions *WNW* to *N*, which is in line with the expectations based on the analysis of upper tail dependence. For wind directions *S* to *W*, results are more scattered. The results show that the copula that best fits all storms within a wind direction (depicted in black) is generally also a suitable model for most of its clusters. In Fig. 31 of the appendix, the AICc scores are presented in the same way for location IJmuiden and Appendix D contains results for other locations, in terms of CvM and AICc scores, and tables with parameters for the copulas that best fit the data corresponding to each wind direction. Results appear to be consistent for all locations and both metrics. In contrast, Geerse et al. (2022) proposed a Gumbel copula to model the dependence between wind speed and surge height, while De Valk and Van den Brink (2023) suggested a Gaussian copula to model the depen-

dence between wind speed and *sea level*. However, neither study evaluated various copula models in detail and data was only clustered based on wind direction.

For flood defense design, we are mainly aiming for a good fit in the upper most tail. Therefore, we compared the upper tail dependence, expressed in λ, for three of the fitted copula models: the BB8, Gumbel and Gaussian copula. Since we concluded previously that sub-clusters show more or less the same behavior as the total wind direction groups, the comparison of tail dependence is only made for the total wind direction. Results are presented in Fig. 11 for location IJmuiden. From literature it is known that the Gaussian and BB8 copula have no upper tail dependence when $t \to 1$ (except for $\delta = 1$ for the BB8 copula parameter). Each panel shows the results for one wind direction, with in blue empirical λ-values for the data, in black for the fitted BB8, in orange for the fitted Gaussian, in gray for the fitted Gumbel copula and in red for a fitted BB8-copula with δ forced to 1. Note that the blue lines are the same as the black lines in Fig. 9, however the thresholds t now range towards even smaller sample fractions, in order to illustrate how the behavior of different copula models evolves towards the upper tail limit. Again, 95%-Wald asymptotic confidence intervals are included. Note that for λ's estimated from the fitted copulas, the uncertainty is not increasing since these estimates are parametric.

The figure shows that the Gumbel copula and BB8 copula with $\delta = 1$ generally overestimate the upper tail dependence of the data, while the fitted Gaussian and BB8 copulas (slightly) underestimate the upper tail dependence. As was previously concluded from the goodness-of-fit test, the BB8 copula resembles the upper tail dependence of the data the best within the data domain. However, when extrapolating to more extreme events for western to northern wind directions ($t \to 1$), the fitted BB8 copula seems to underestimate the dependence between wind speed and surge height. The same holds for the other locations, as presented in Appendix D.

4 Discussion

This study follows a systematic approach as depicted in Fig. 2, to model the dependence between extreme wind speeds and surge heights for various storm clusters. Several methodological choices were made throughout the process, each of which may influence the results. The first step was selecting storm events from the SEAS5-WAQUA dataset using Peaks-Over-Threshold using a fixed time window of three days and a wind direction-dependent threshold. This ensures the selection of storms with a return period of at least one year, conditional on wind direction (i.e. 9000 storms per wind direction). Since higher wind speeds are more common from southern to western directions, this approach results in the selection of less extreme storms (in absolute terms) for northern wind directions. However, from a probabilistic perspective, the storm selection is consistent over wind directions. Also, selecting 9000 storms per wind direction allows for further sub-clustering based on other storm features while maintaining a sufficient number of data points. Besides, a sensitivity analysis of the time window and thresholds in the POT approach indicated that these choices are robust.

Each storm was then parametrized to predefined storm features and subsequently clustered using five key features (after pre-clustering based on wind direction r): tidal offset (φ), tidal peak (a), surge exceedance duration (d_s), wind exceedance duration (d_u) and

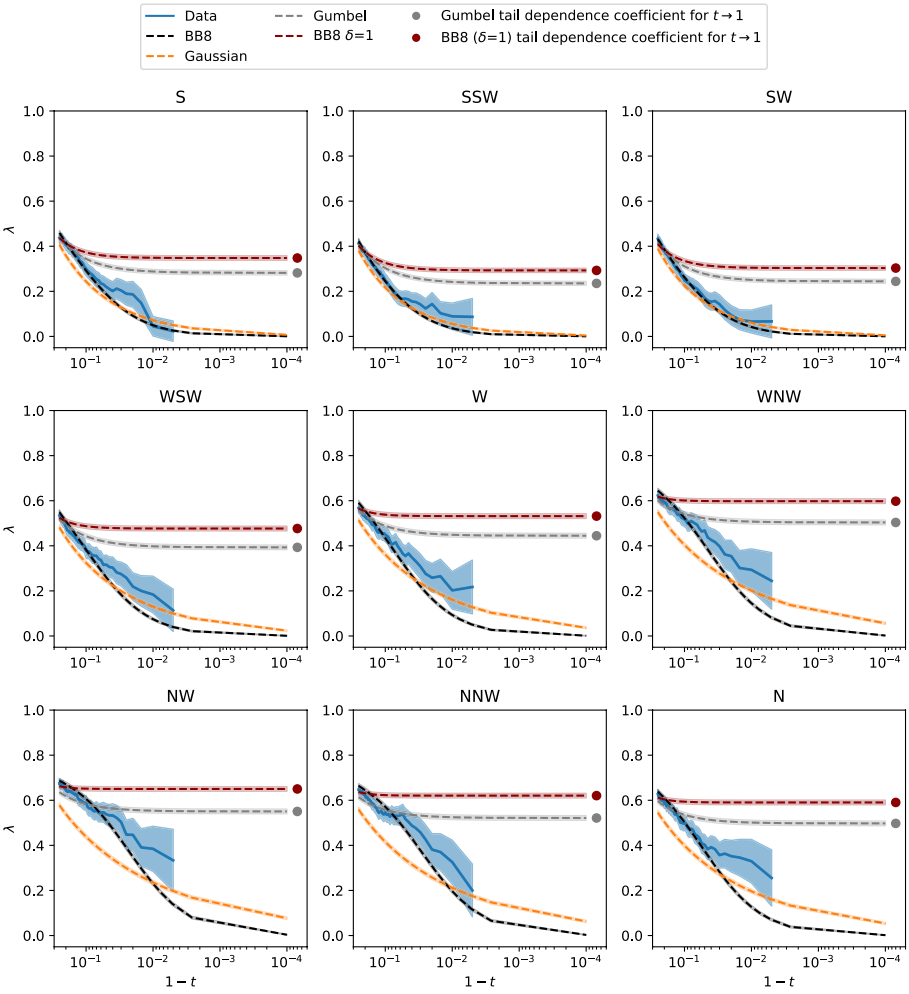


Fig. 11 Tail dependence for wind speed and surge height for different sample fractions from the storm data (in blue), compared to a fitted BB8 copula (dashed black), a Gaussian copula (dashed orange), a Gumbel copula (dashed gray) and a BB8 copula with δ forced to 1, per wind direction at location IJmuiden. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

wind rotation (ψ). These features were chosen based on expert knowledge and insights from Caspers et al. (2025) . It is important to note that the wind and surge exceedance durations were defined on a *relative* level, which inherently introduces a negative correlation between wind speed magnitude and duration—i.e., higher wind speed peaks lead to a high threshold, which is generally exceeded for a shorter time, and vice versa. Alternatively, the clustering could have been performed directly on the storm hydrographs, without prior parameterization, similar to Pupić Vurilj et al. (2025). However, given the aim of aligning the insights from this study with the Dutch framework for flood defense assessments, the current feature-based approach was considered more suitable. In contrast to Caspers et al. (2025), we applied K-means clustering instead of manual clustering, to allow more flex-

ibility in the clustering process and to evaluate whether other features than tidal offset and surge exceedance duration may play an important role.

4.1 Reliability of the results

A key distinction from current design practice is the use of a synthetic dataset. The synthetic dataset provides far more information on extreme storm events than observational data, that are traditionally used to derive design hydraulic loads. However, it is important to acknowledge that simulated data contain biases and model uncertainties. De Valk and Van den Brink (2024) demonstrated that for return values of wind speed, the SEAS5 data is sufficiently accurate compared to observations and other models, making it suitable for deriving design values, when accounting for the model uncertainty. For sea levels, a bias was identified in the DCSMv5 data, and therefore De Valk and Van den Brink (2024) proposes that return values should be estimated using a combination of observational and synthetic data. With respect to the dependence between wind speed and sea level, our results show that the upper tail dependence estimates derived from the SEAS5/DCSMv5 dataset closely match those obtained from observations, supporting the reliability of the synthetic data for the purposes of this study. In fact, this comparison emphasizes the importance of long synthetic datasets for obtaining robust estimates of dependence in the extremes, as reflected by large confidence intervals and unstable estimates based on observations (Fig. 3).

For this study, time series of wind and sea levels were extracted for point locations from SEAS5 wind fields and DCSMv5 sea levels. A limitation of using point data is that spatial effects (e.g. large-scale depressions) cannot be identified. Another important assumption in this study is the selection of wind locations northwest of the sea level point locations, which is based on the well-established relevance of this direction in Dutch flood risk assessments. While this choice aligns with the dominant storm direction, it may have influenced the resulting dependence structures between wind and surge. To assess the robustness of the results, a sensitivity analysis considering alternative locations is recommended for future work.

It should be noted that the uncertainty in the estimation of tail dependence coefficients and copula models is strongly influenced by the cluster size. Smaller sample sizes lead to larger variability and less reliable estimates, particularly in the upper tail where data are sparse. In this study, the sample size varies across clusters, which directly affects the robustness of the estimated dependence measures. To make this explicit, confidence intervals have been included where possible to indicate the associated uncertainty in the estimates. For this, Wald statistics have been applied. Bootstrapping has been tested as well and gives similar results (compare Fig. 3), however with much larger computation times.

4.2 Skew surge versus residual surge versus sea level

The dependence between extreme wind speeds and surge heights was analyzed separately for each storm cluster. A key distinction from similar studies for the Dutch coast (e.g. De Valk and Van den Brink 2024) and current Dutch design practice (e.g. Vrouwenvelder et al. 2003; Geerse 2002) is the use of storm surge rather than total sea level. Current design practice typically relies on total sea level, since it is directly observable, whereas estimating the storm surge requires an additional step to account for the (timing of) astronomical

tide. Since storm surge directly results from wind forcing, analyzing the statistical dependence between wind speed and surge height provides a more natural and direct relationship. Moreover, using total sea level introduces additional variability due to the timing relative to the astronomical tide. To illustrate this, Fig. 12 presents three storm hydrographs at Hoek van Holland with varying tidal offset. Despite having the same maximum surge height, the resulting maximum total sea levels differ substantially, due to the timing with respect to high tide. In current design practice, such scenarios would be assigned different probabilities of occurrence, even though they are caused by the same (residual) surge height s (with different probabilities for the occurrence of the tidal offset value).

In this study, we mainly considered the residual surge height s, defined as the difference between the total sea level and the astronomical tide. Alternatively, we could have studied the dependence between wind speed and *skew* surge k, defined as the difference between the highest observed sea level and the highest predicted tide within a tidal cycle. Both skew surge and residual surge are common in storm surge studies, each with distinct advantages. The skew surge is often used to avoid tide-surge interactions (Batstone et al. 2013, e.g.). However, skew surge compresses the surge into a single value per tidal cycle, omitting information on timing and duration. Residual surge, in contrast, provides a full time series of the non-tidal signal, allowing for detailed analysis of both tidal offset and surge duration. Given our focus on how storm characteristics influence the tail dependence between wind speed and surge height, the residual surge was the more appropriate choice. To test the sensitivity of this choice, we also examined pair plots between wind speed and skew surge at Vlissingen (Fig. 20, Appendix C.1). The results show similar upper-tail correlation coefficients, suggesting that our findings would remain largely unchanged if skew surge had been used instead.

4.3 Practical implications

In the current hydraulic load models used for reliability assessments of Dutch coastal flood defenses, the statistical dependence between wind speed u and sea level w is required for deriving local wave conditions associated with combinations of these two variables. To assess the impact of different copula models on design hydraulic loads, we investigated the effect on the significant wave height H_s at the location Hoek van Holland for wind direction *NW*. We applied the cumulative distribution functions (CDFs) of wind speed u and sea level w that are applied in the official Dutch flood defense assessment and design

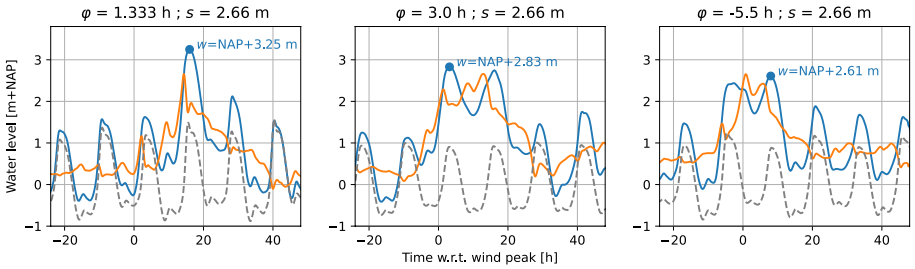


Fig. 12 Illustration of the effect of the tidal phase φ on the resulting maximum sea level w for the same maximum surge height s. The subplots show the hydrographs for a storm event corresponding to $\varphi = 1.333$, 3.0 and $-$ 5.5 h, which are common values for location Hoek van Holland

software Hydra-NL (Duits and Kuijper 2019), ensuring consistency with current practice. Note that this is a pragmatic choice, since these CDFs followed from observations and are thus not consistent with the fitted copulas from the SEAS5-WAQUA dataset used in this study. Also, for this analysis we fitted copula models to pairs of u and w (instead of s). For wind direction *NW*, we sampled paired values of u and w from fitted BB8, Gaussian and Gumbel copulas. For each pair, the corresponding significant wave height H_s was derived from the wave interpolation grid used in Hydra-NL for this specific location and direction. The CDFs of u and w, together with the wave interpolation grid, are provided in Fig. 40 in Appendix E.

Figure 13 compares the resulting exceedance frequency plots of significant wave height for the different copula models. The results indicate that the choice of dependence model has a notable effect: for a return period of 10^4 years, the difference between the Gaussian and Gumbel copulas amounts to approximately 40 cm in significant wave height. Compared to the empirical exceedance frequency curve (following from the empirical copula of SEAS5-WAQUA (u,w)-pairs), applying a Gaussian copula would result in an underestimation of the resulting wave height, while a Gumbel copula may be a slight overestimation.

The impact of choosing the right copula model is non-negligible in the context of flood defense design, highlighting the importance of accurately characterizing the dependence structure between wind speed and sea level. The long synthetic dataset used in this study enables to model this dependence with greater confidence than was previously possible based on observations.

It is important to note that this analysis is indicative rather than definitive. As mentioned, the wave interpolation grid and CDFs used here were not derived from the same SEAS5-DCSM5 dataset that underlies the fitted copula models. Moreover, the SWAN computations on which the interpolation grid is based do not account for variations in storm characteristics such as wind duration or tidal offset. A more consistent assessment would therefore require new wave simulations that explicitly incorporate these storm features by using the large synthetic SEAS5-WAQUA dataset. While this lies beyond the scope of the present study, we plan to address this in future work.

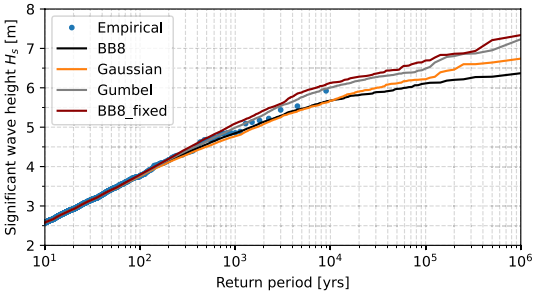


Fig. 13 Comparison of resulting exceedance frequency plots for significant wave height at location Hoek van Holland for wind direction *NW* between various copula models for the dependence between extreme wind speed and surge height

5 Conclusions

The clustering procedure yielded valuable insights into storm behavior and the influence of various characteristics, such as wind direction, tidal offset, and wind exceedance duration, on the statistical description of extreme wind speed and surge height. The clustering approach based on K-Means produced 16 clusters per wind direction and location.

Analysis of the upper tail dependence between wind speed and surge height, using Pearson's semi-correlations and the tail dependence coefficient λ, revealed the strongest upper tail dependence for north-western wind directions, with Pearson's correlation coefficient, σ_{NW}, up to 0.63 across locations. While λ decreased towards the upper tail limit, it did not reach zero for western to northern wind directions, indicating asymptotic dependence between wind speed and surge height. Sub-clusters within a wind direction generally followed similar dependence patterns as the total wind direction cluster, but in some cases with a slightly stronger upper tail dependence for shorter wind exceedance durations and larger (absolute) tidal offsets, with σ_{NW} up to 0.80. Overall, the limited variation in dependence structure across sub-clusters suggests that subdivision based on other storm characteristics than wind direction may not be necessary for representative dependence modeling.

These small variations among sub-clusters with respect to dependence was also reflected in the performance of different copula models: the copula that best fits all storms within a wind direction is generally also a good fit for most of its sub-clusters. The BB8 copula emerged as the best-fit model for most clusters and sub-clusters based on the CvM and AIC criteria. However, the BB8 copula does not model upper tail dependence, unless $\delta = 1$. Comparisons showed that both the fitted BB8 and Gaussian copula tend to underestimate the observed tail dependence toward the upper limit, whereas Gumbel and BB8 with $\delta = 1$ tended to overestimate it. Given the need for safe estimates in flood defense design, overestimation is often preferable over underestimation of upper tail dependence. To improve the dependence modelling in current design practice, the local wave conditions consistent with the SEAS5-WAQUA dataset are required, which we intend to model in the future.

A key strength of this work lies in the use of a large synthetic dataset, which consists of detailed information about extreme storm behaviors. This dataset allowed for a more comprehensive analysis of extreme events, than from the limited amount of observational data. The synthetic data enabled robust clustering and copula model fitting, ultimately enhancing our ability to investigate and interpret the dependencies among variables during extreme conditions.

Appendix A: Data comparison

See Figs. 14, 15, 16, and 17.

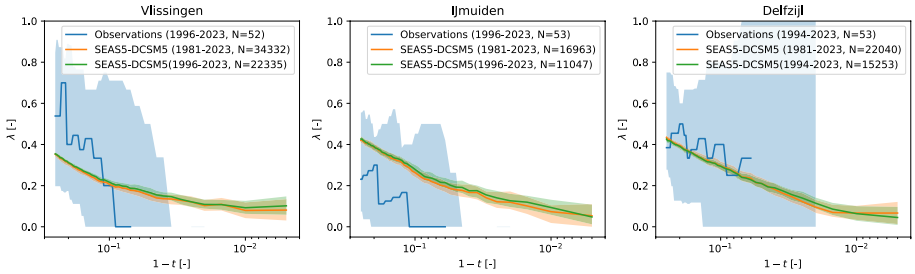


Fig. 14 Comparison of empirical upper tail dependence coefficient λ for different sample fractions t, for pairs of extreme sea level and wind speed, including a 95%-confidence interval based on bootstrapping. Different colors indicate the different datasets. Each panel represents one location. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than three

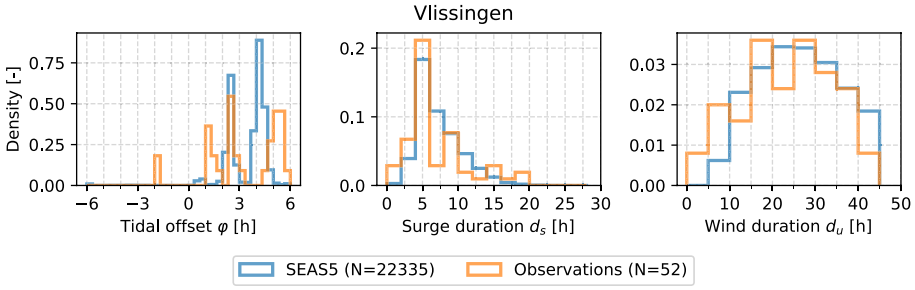


Fig. 15 Comparison of histograms for tidal offset φ, surge duration d_s and wind duration d_u between observational data (orange) and the SEAS5-WAQUA-DCSMv5 data (blue) for location Vlissingen

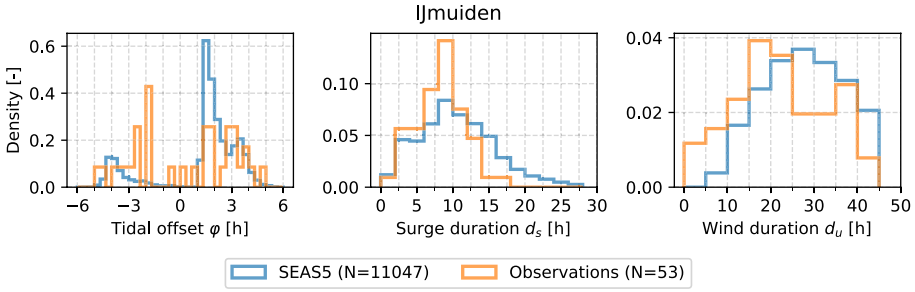


Fig. 16 Comparison of histograms for tidal offset φ, surge duration d_s and wind duration d_u between observational data (orange) and the SEAS5-WAQUA-DCSMv5 data (blue) for location IJmuiden

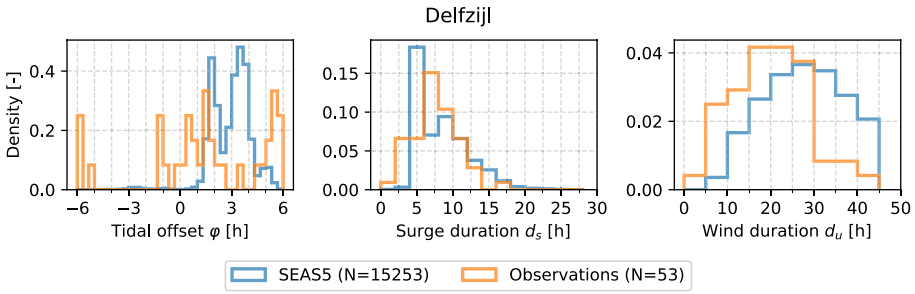


Fig. 17 Comparison of histograms for tidal offset φ, surge duration d_s and wind duration d_u between observational data (orange) and the SEAS5-WAQUA-DCSMv5 data (blue) for location Delfzijl

Appendix B: Sensitivity analysis for the POT-threshold

In this study, a wind speed threshold varying by wind direction was applied, resulting in a selection of 9000 storm events per direction. To assess the sensitivity of this choice, we compared pair plots of maximum wind speed (u) and maximum surge height (s) obtained using an alternative threshold definition. In the alternative approach, 9000 storm events in total were selected based on a sea level threshold. Fig. 18 illustrates this comparison for Den Helder. Pairs selected using the sea level threshold are shown in red, and those based on the wind speed thresholds in blue. The Pearson correlation coefficients in the upper tail are found to be similar across wind directions. However, the sample size varies considerably between directions when using the sea level threshold, which is undesirable for the purpose of comparing upper-tail dependence across wind directions. Therefore, the use of varying wind speed thresholds per direction is considered a suitable choice.

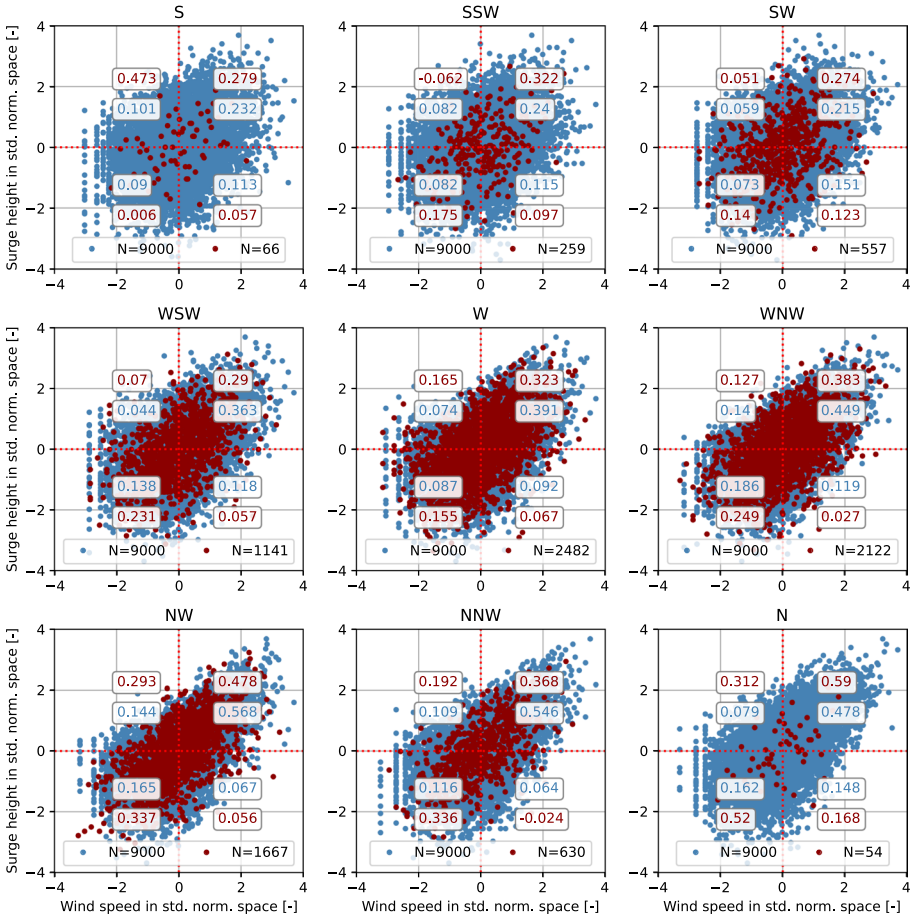


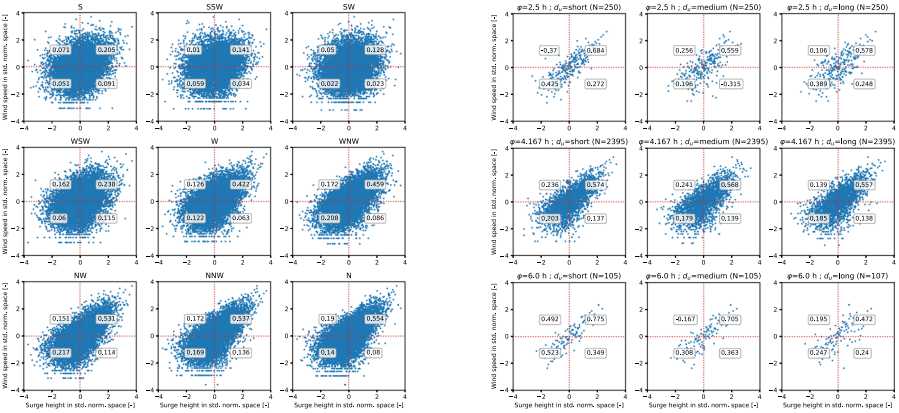
Fig. 18 Comparison of pair plots in the standard normal space

Appendix C: Upper tail dependence

C.1: Pearson's correlation coefficients

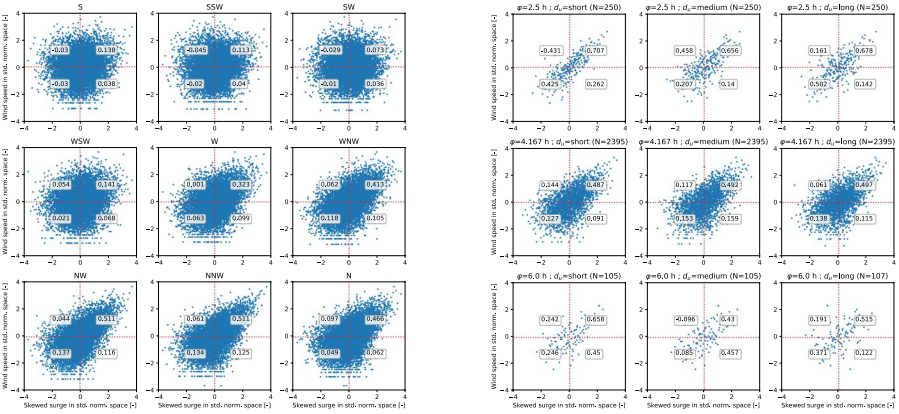
C.1.1: Pair plots

See Figs. 19, 20, 21, 22, 23, and 24.



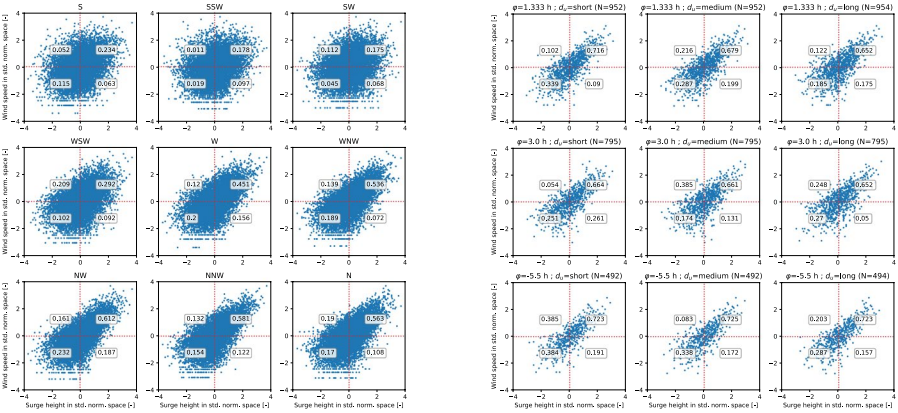
(a) (b)

Fig. 19 Pair plots for surge height and wind speed in the standard normal space at location Vlissingen. **a** Per western wind direction. **b** K-Means sub-clusters within wind direction NNW



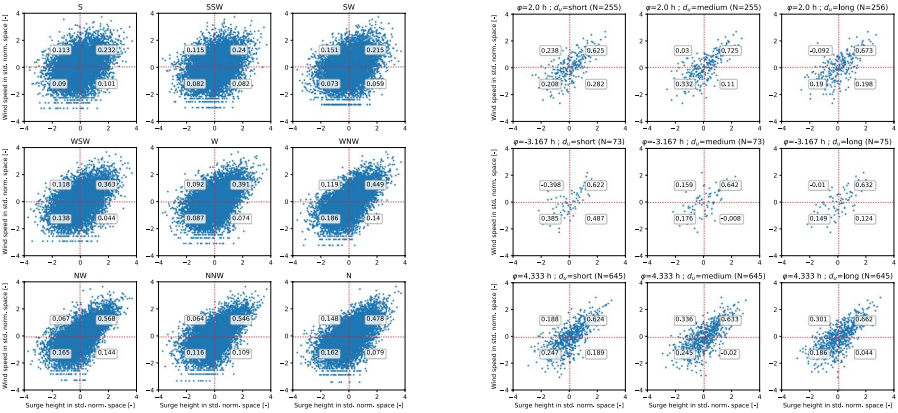
(a) (b)

Fig. 20 Pair plots for skew surge k and wind speed u in the standard normal space at location Vlissingen. **a** Per western wind direction. **b** K-Means sub-clusters within wind direction NNW



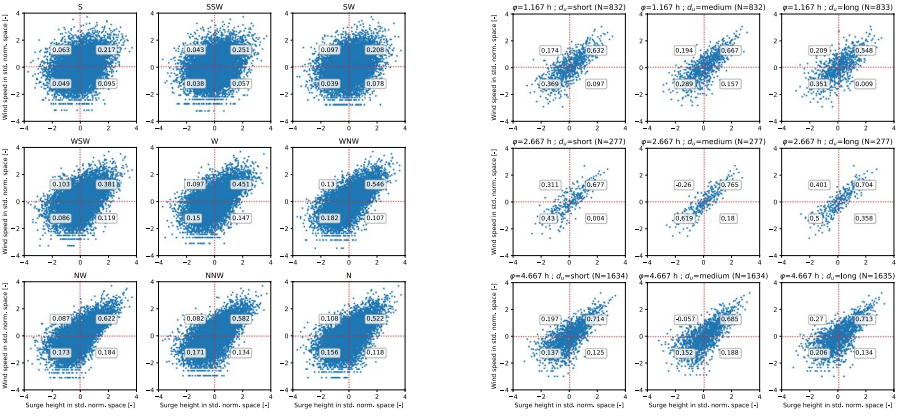
(a) (b)

Fig. 21 Pair plots for surge height and wind speed in the standard normal space at location Hoek van Holland. **a** Per western wind direction. **b** K-Means sub-clusters within wind direction NW



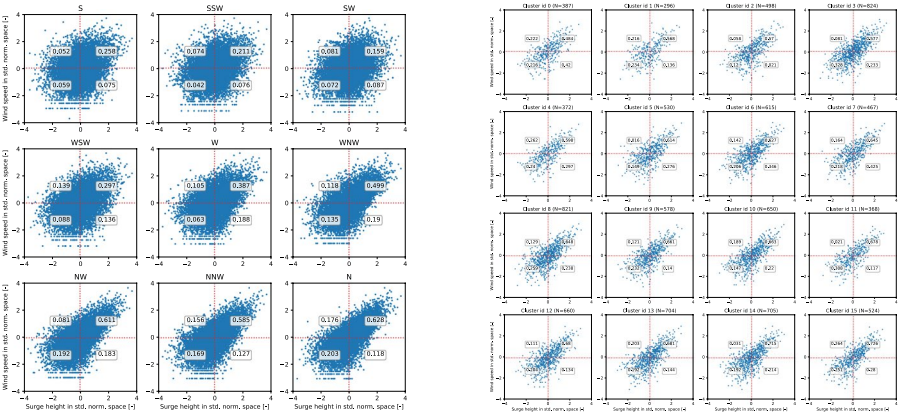
(a) (b)

Fig. 22 Pair plots for surge height and wind speed in the standard normal space at location Den Helder. **a** Per western wind direction. **b** K-Means sub-clusters within wind direction NW



(a) (b)

Fig. 23 Pair plots for surge height and wind speed in the standard normal space at location Harlingen. **a** Per western wind direction. **b** K-Means sub-clusters within wind direction NW



(a) (b)

Fig. 24 Pair plots for surge height and wind speed in the standard normal space at location Delfzijl. **a** Per western wind direction. **b** K-Means sub-clusters within wind direction N

C.1.2: Storm characteristics

See Figs. 25, 26, and 27.

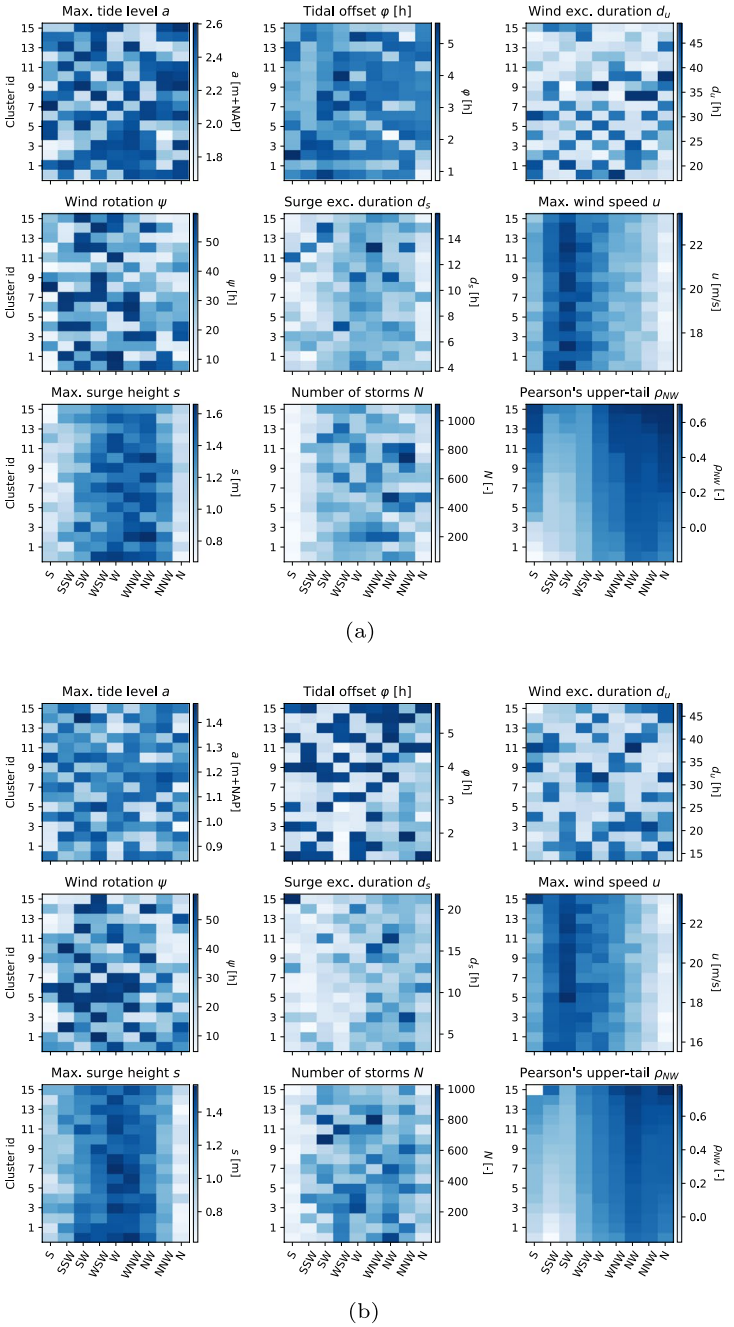


Fig. 25 Overview of the clusters for location **a** Vlissingen and **b** Hoek van Holland. Each panel refers to one storm feature, the cluster size and Pearson's correlation coefficient. Each column in a panel shows the 16 clusters belonging to a specific wind direction. The color indicates the cluster's centroid value of the corresponding feature

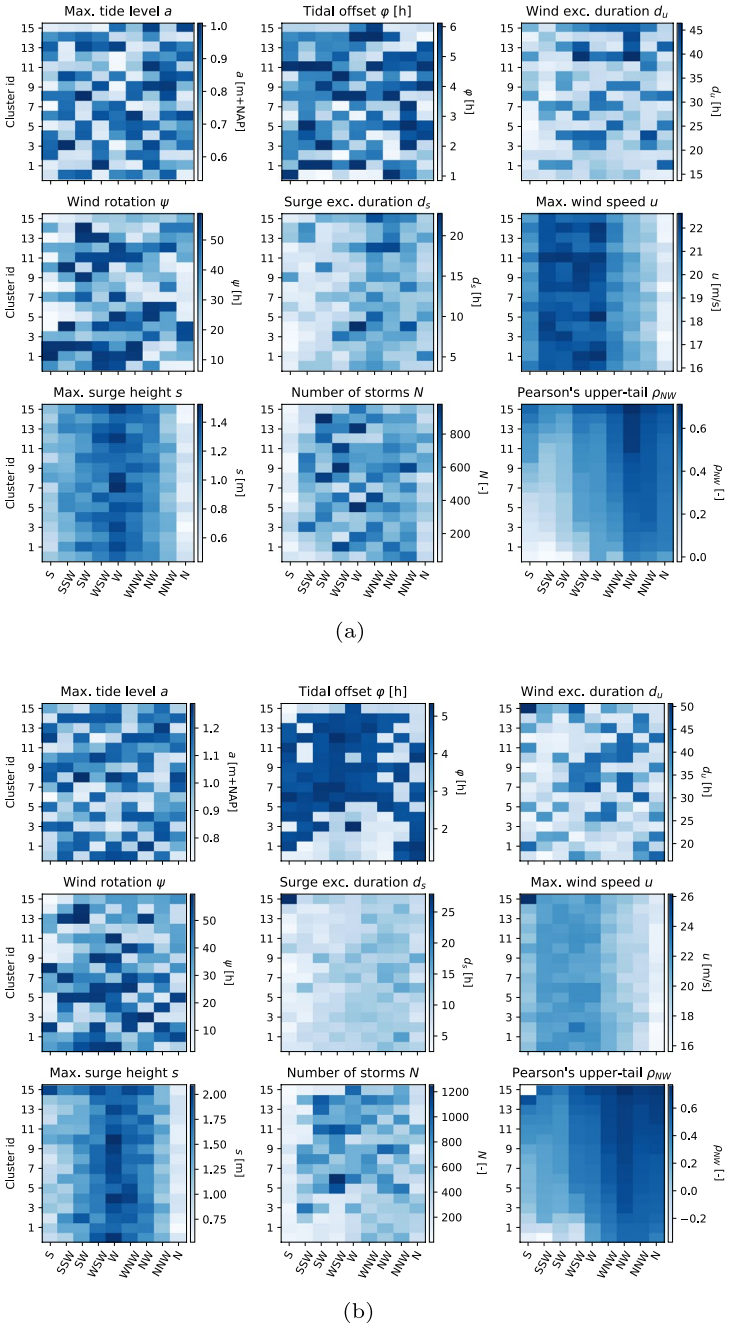


Fig. 26 Overview of the clusters for location **a** Den Helder and **b** Harlingen. Each panel refers to one storm feature, the cluster size and Pearson's correlation coefficient. Each column in a panel shows the 16 clusters belonging to a specific wind direction. The color indicates the cluster's centroid value of the corresponding feature

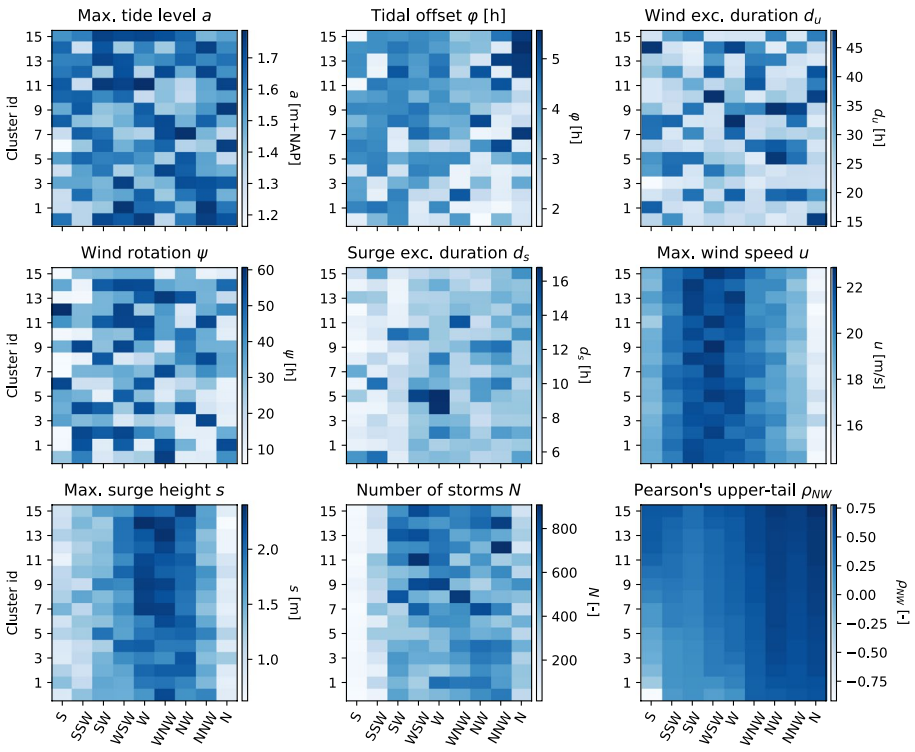


Fig. 27 Overview of the clusters for location Delfzijl. Each panel refers to one storm feature, the cluster size and Pearson's correlation coefficient. Each column in a panel shows the 16 clusters belonging to a specific wind direction. The color indicates the cluster's centroid value of the corresponding feature

C.2: Upper tail dependence coefficient

C.2.1: Confidence intervals

Wald asymptotic confidence interval is applied to estimate the two-sided confidence interval (Agresti and Coull 1998):

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

where $\hat{p}$ denotes the estimated probability (in this case the upper tail dependence coefficient λ), $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution and n is the sample size used to estimate $\hat{p}$.

C.2.2: Results

See Figs. 28, 29, and 30.

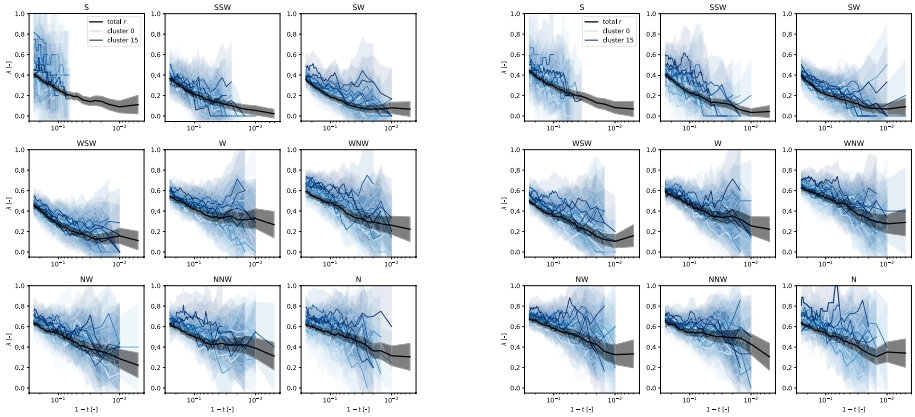


Fig. 28 Upper tail dependence λ for pairs of wind speed u and surge height s, for different sample fraction thresholds t, at location **a** Vlissingen and **b** Hoek van Holland, with the x-axis being $1 - t$ on a log-scale. Each panel represents one of the relevant wind directions. In blue, λ is presented per cluster and in black for all storm events within the corresponding wind direction. The shaded bands represent confidence intervals based on Wald statistics. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

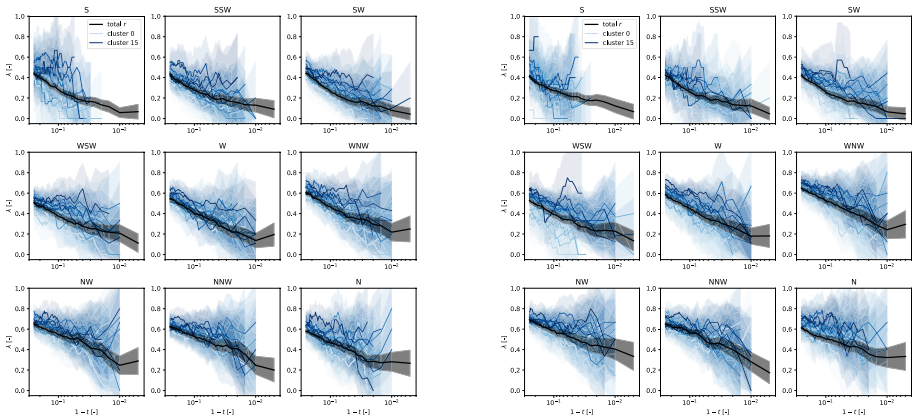


Fig. 29 Upper tail dependence λ for pairs of wind speed u and surge height s, for different sample fraction thresholds t, at location **a** Den Helder and **b** Harlingen, with the x-axis being $1 - t$ on a log-scale. Each panel represents one of the relevant wind directions. In blue, λ is presented per cluster and in black for all storm events within the corresponding wind direction. The shaded bands represent confidence intervals based on Wald statistics. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

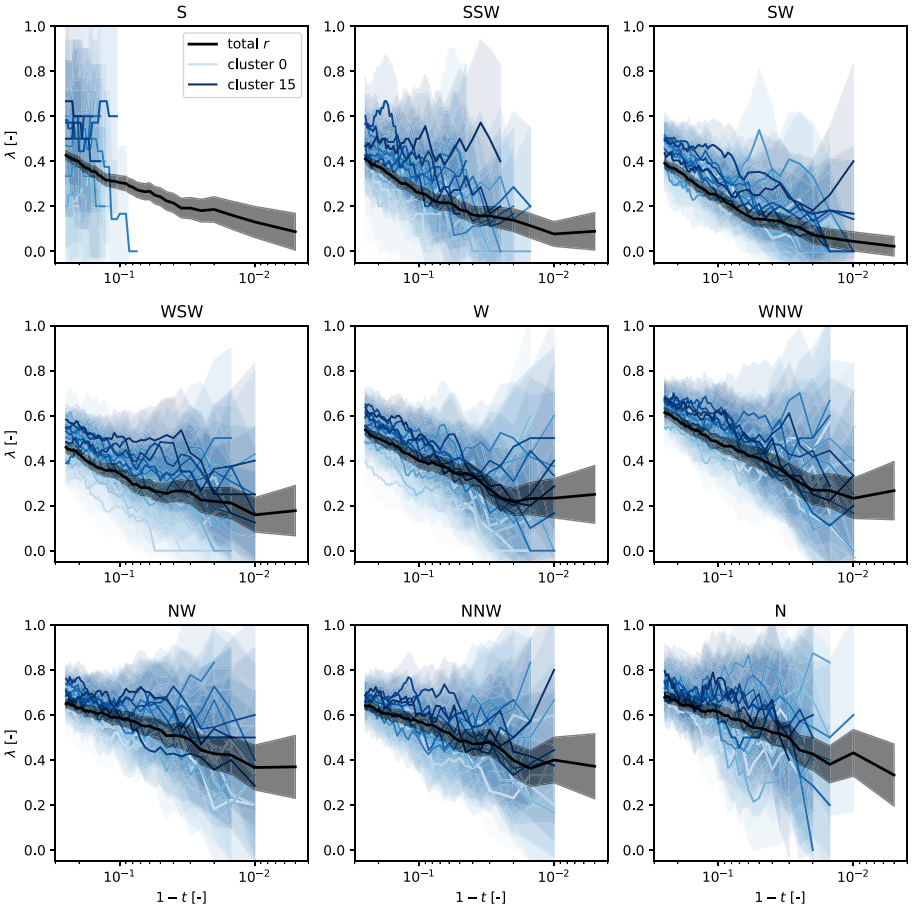


Fig. 30 Upper tail dependence λ for pairs of wind speed u and surge height s, for different sample fraction thresholds t, at location Delfzijl, with the x-axis being $1 - t$ on a log-scale. Each panel represents one of the relevant wind directions. In blue, λ is presented per cluster and in black for all storm events within the corresponding wind direction. The shaded bands represent confidence intervals based on Wald statistics. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

Appendix D: Copula modeling

D.1: Copula definitions

The BB8 copula is defined as

$$C^{\mathrm{BB8}}(u,v;\theta,\delta) = \delta^{-1}\left[1-\left\{1-\left[1-(1-\delta)^{\theta}\right]^{-1}\left[1-(1-\delta u)^{\theta}\right]\left[1-(1-\delta v)^{\theta}\right]\right\}^{(1/\theta)}\right], \quad \text{(D1)}$$

where $\theta \geq 1$ and $0 \leq \delta \leq 1$. The independent copula is obtained as $\theta \to 1$ or $\delta \to 0$. BB8 copula decays to the Joe copula when δ is equal to 1, while the Frank copula is obtained as $\theta \to \infty$, and the single parameter can be calculated by using the formula $1 - (1 - \delta)\theta$. BB8 copula does not exhibit tail dependence except when δ is equal to 1.

D.2: Goodness-of-fit tests

D.2.1: Cramér–Von Mises

The Cramér–von Mises (CvM) criterion is a statistical measure used to assess the goodness of fit of a model. In the context of copulas, it evaluates how well the copula's estimated dependence structure matches the empirical dependence structure observed in the data.

$$S_{CvM} = n \int_{[0,1]^2} \left(C_n(u_1, u_2) - C_\theta(u_1, u_2)\right)^2 \, \mathrm{d}C_\theta(u_1, u_2) \tag{D2}$$

where C_θ is the parametric copula and C_n is the empirical copula.

D.2.2: AIC

The Akaike Information Criterion (AIC) is defined as

$$\mathrm{AIC} = -2 \cdot \log L + 2k \tag{D3}$$

where $\log L$ is the log-likelihood of the copula model and k is the number of estimated parameters in the model.The log-likelihood of the copula model is defined as

$$\log L(\theta) = \sum_{i=1}^{n} \log c_\theta(U_{i1}, U_{i2}) \tag{D4}$$

D.3: Copula evaluation

See Figs. 31, 32, 33, 34, 35, 36 and Tables 1, 2, 3, 4, 5, 6.

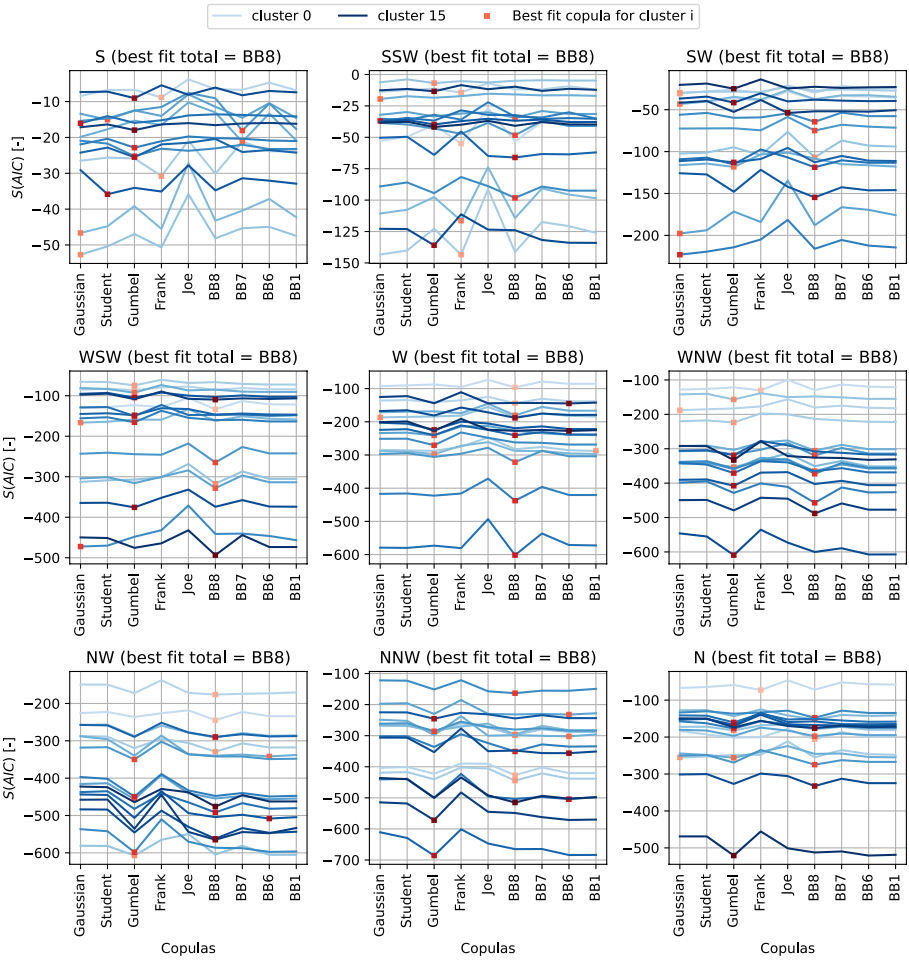


Fig. 31 AICc scores for various copula models, fit to pairs of wind speed u and surge height s corresponding to the 16 storm clusters at location IJmuiden. Each plot shows the results for clusters corresponding to that wind direction. The blues indicate the 16 clusters and the black all storm events within the corresponding the wind direction. The squares indicate the copulas that best fit the corresponding data

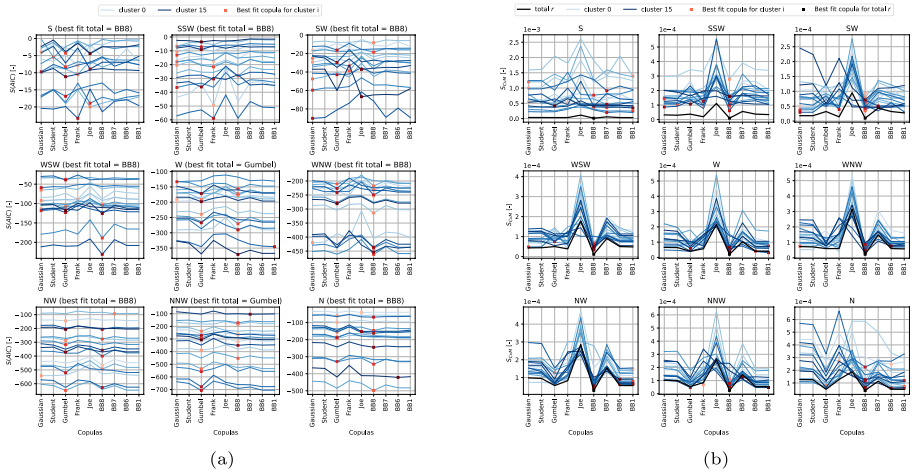


Fig. 32 Best-fit copula models according to **a** the AICc and **b** the Cramér–von Mises criterion for all sub-clusters at location Vlissingen

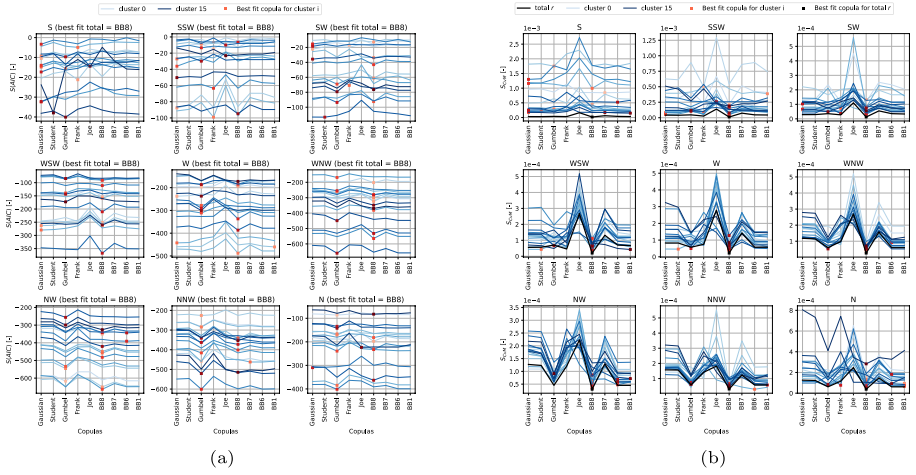


Fig. 33 Best-fit copula models according to **a** the AICc and **b** the Cramér–von Mises criterion for all sub-clusters at location Hoek van Holland

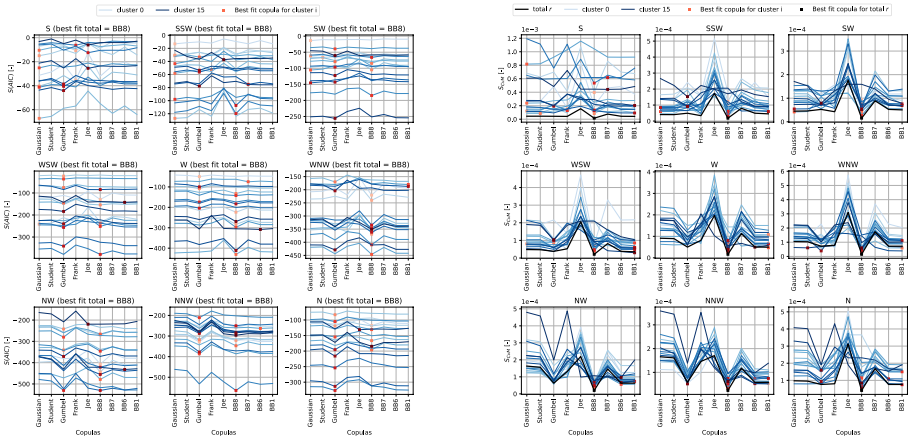


Fig. 34 Best-fit copula models according to **a** the AICc and **b** the Cramér–von Mises criterion for all sub-clusters at location Den Helder

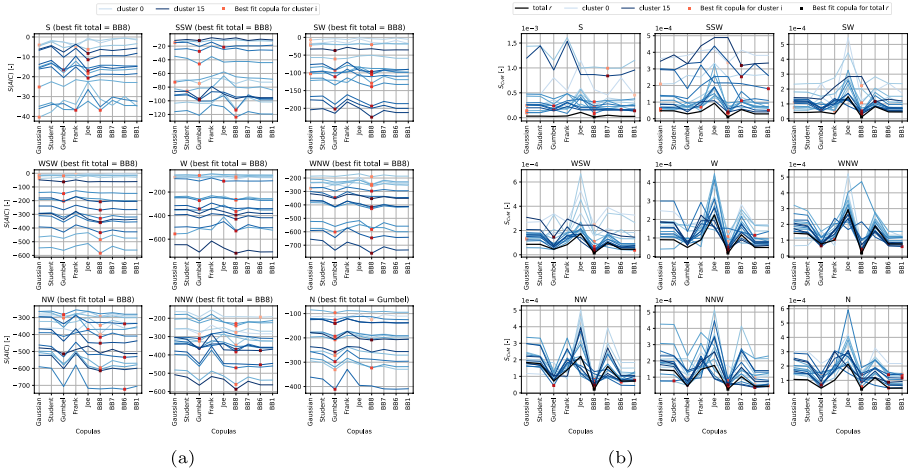


Fig. 35 Best-fit copula models according to **a** the AICc and **b** the Cramér–von Mises criterion for all sub-clusters at location Harlingen

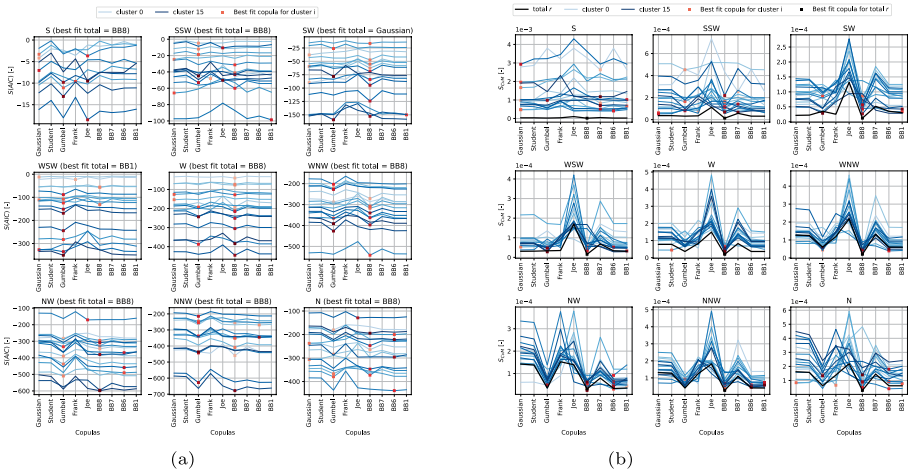


Fig. 36 Best-fit copula models according to **a** the AICc and **b** the Cramér–von Mises criterion for all sub-clusters at location Delfzijl

Table 1 Parameters of the copula models that best fit the data corresponding to each relevant wind direction at location IJmuiden

Index	CvM	AIC
180	BB8, (2.14 ; 0.82)	BB8, (2.14 ; 0.82)
202.5	BB8, (2.19 ; 0.75)	BB8, (2.19 ; 0.75)
225	BB8, (2.02 ; 0.80)	BB8, (2.02 ; 0.80)
247.5	BB8, (2.77 ; 0.83)	BB8, (2.77 ; 0.83)
270	BB8, (2.99 ; 0.84)	BB8, (2.99 ; 0.84)
292.5	BB8, (3.00 ; 0.90)	BB8, (3.00 ; 0.90)
315	BB8, (3.00 ; 0.94)	BB8, (3.00 ; 0.94)
337.5	BB8, (2.86 ; 0.94)	BB8, (2.86 ; 0.94)
0	BB8, (3.15 ; 0.88)	BB8, (3.15 ; 0.88)

Table 2 Parameters of the copula models that best fit the data corresponding to each relevant wind direction at location Vlissingen

Wind direction r	CvM	AIC
180	BB8, (1.76 ; 0.85)	BB8, (0.85)
202.5	BB8, (1.78; 0.77)	BB8, (1.78 ; 0.77)
225	BB8, (1.75 ; 0.73)	BB8, (1.75 ; 0.73)
247.5	BB8, (2.35 ; 0.77)	BB8, (2.35 ; 0.77)
270	BB8, (2.63 ; 0.87)	Gumbel, (1.51)
292.5	BB8, (3.49 ; 0.81)	BB8, (3.49 ; 0.81)
315	BB8, (3.43 ; 0.86)	BB8, (3.43 ; 0.86)
337.5	BB8, (3.22 ; 0.88)	Gumbel, (1.74)
0	BB8, (2.59 ; 0.95)	BB8, (2.59 ; 0.95)

Table 3 Parameters of the copula models that best fit the data corresponding to each relevant wind direction at location Hoek van Holland

Index	CvM	AIC
180	BB8, (2.10 ; 0.81)	BB8, (2.10 ; 0.81)
202.5	BB8, (1.94 ; 0.78)	BB8, (1.94 ; 0.78)
225	BB8, (1.77 ; 0.80)	BB8, (1.77 ; 0.80)
247.5	BB8, (2.94 ; 0.76)	BB8, (2.94 ; 0.76)
270	BB8, (3.14 ; 0.85)	BB8, (3.14 ; 0.85)
292.5	BB8, (3.28 ; 0.88)	BB8, (3.28 ; 0.88)
315	BB8, (3.30 ; 0.92)	BB8, (3.30 ; 0.92)
337.5	BB8, (2.86 ; 0.94)	BB8, (2.86 ; 0.94)
0	BB8, (2.96 ; 0.92)	BB8, (2.96 ; 0.92)

Table 4 Parameters of the copula models that best fit the data corresponding to each relevant wind direction at location Den Helder

index	CvM	AIC
180	BB8, (2.01 ; 8.33e−01)	BB8, (2.01 ; 8.33e−01)
202.5	BB8, (2.13 ; 7.96e−01)	BB8, (2.13 ; 7.96e−01)
225	BB8, (2.18 ; 7.94e−01)	BB8, (2.18 ; 7.94e−01)
247.5	BB8, (2.42 ; 8.56e−01)	BB8, (2.42 ; 8.56e−01)
270	BB8, (2.54 ; 8.69e−01)	BB8, (2.54 ; 8.69e−01)
292.5	BB8, (3.45 ; 8.27e−01)	BB8, (3.45 ; 8.27e−01)
315	BB8, (3.07 ; 9.16e−01)	BB8, (3.07 ; 9.16e−01)
337.5	BB8, (2.59 ; 9.45e−01)	BB8, (2.59 ; 9.45e−01)
0	BB8, (3.06 ; 8.58e−01)	BB8, (3.06 ; 8.58e−01)

Table 5 Parameters of the copula models that best fit the data corresponding to each relevant wind direction at location Harlingen

Index	CvM	AIC
180	BB8, (1.71 ; 8.85e−01)	BB8, (1.71 ; 8.85e−01)
202.5	BB8, (1.65 ; 9.01e−01)	BB8, (1.65 ; 9.01e−01)
225	BB8, (2.01 ; 8.13e−01)	BB8, (2.01 ; 8.13e−01)
247.5	BB8, (2.34 ; 8.92e−01)	BB8, (2.34 ; 8.92e−01)
270	BB8, (2.75 ; 8.78e−01)	BB8, (2.75 ; 8.78e−01)
292.5	BB8, (3.55 ; 8.67e−01)	BB8, (3.55e+00 ; 8.67e−01)
315	BB8, (3.41 ; 9.18e−01)	BB8, (3.41e+00 ; 9.18e−01)
337.5	BB8, (2.83 ; 9.43e−01)	BB8, (2.83e+00 ; 9.43e−01)
0	BB8, (2.67 ; 9.25e−01)	Gumbel, (1.66e+00)

Table 6 Parameters of the copula models that best fit the data corresponding to each relevant wind direction at location Delfzijl

Index	CvM	AIC
180	BB8, (1.68e+00 ; 9.20e−01)	BB8, (1.68e+00 ; 9.20e−01)
202.5	BB8, (1.79e+00 ; 8.59e−01)	BB8, (1.79e+00 ; 8.59e−01)
225	BB8, (2.05e+00 ; 7.57e−01)	Gaussian, (2.91e−01)
247.5	BB8, (2.25e+00 ; 8.37e−01)	BB1, (3.30e−02 ; 1.32e+00)
270	BB8, (2.20e+00 ; 9.13e−01)	BB8, (2.20e+00 ; 9.13e−01)
292.5	BB8, (2.87e+00 ; 8.94e−01)	BB8, (2.87e+00 ; 8.94e−01)
315	BB8, (2.73e+00 ; 9.58e−01)	BB8, (2.73e+00 ; 9.58e−01)
337.5	BB8, (2.94e+00 ; 9.34e−01)	BB8, (2.94e+00 ; 9.34e−01)
0	BB8, (3.38e+00 ; 9.27e−01)	BB8, (3.38e+00 ; 9.27e−01)

D.4: Copula comparison of tail dependence

See Figs. 37, 38, and 39.

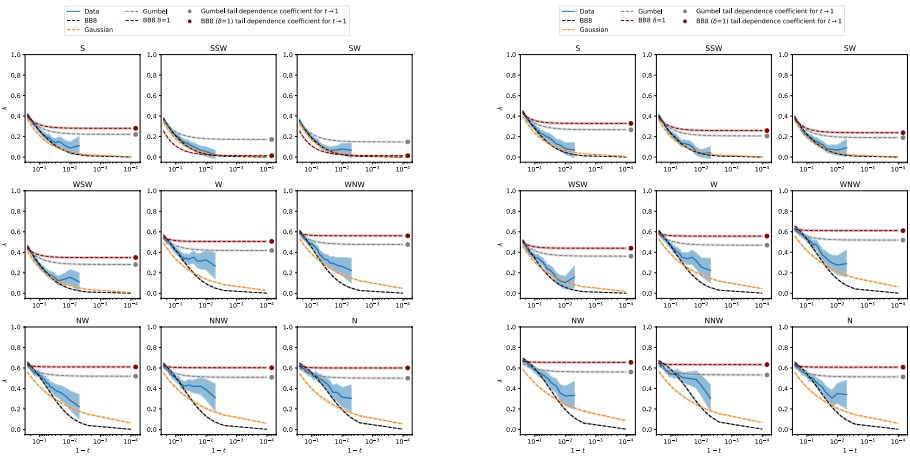


Fig. 37 Tail dependence for wind speed and surge height for different percentile thresholds from the storm data (in blue), compared to a fitted BB8 copula (dashed black), a Gaussian copula (dashed orange) and a Gumbel copula (dashed gray) per wind direction at location Vlissingen (**a**) and Hoek van Holland (**b**). Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

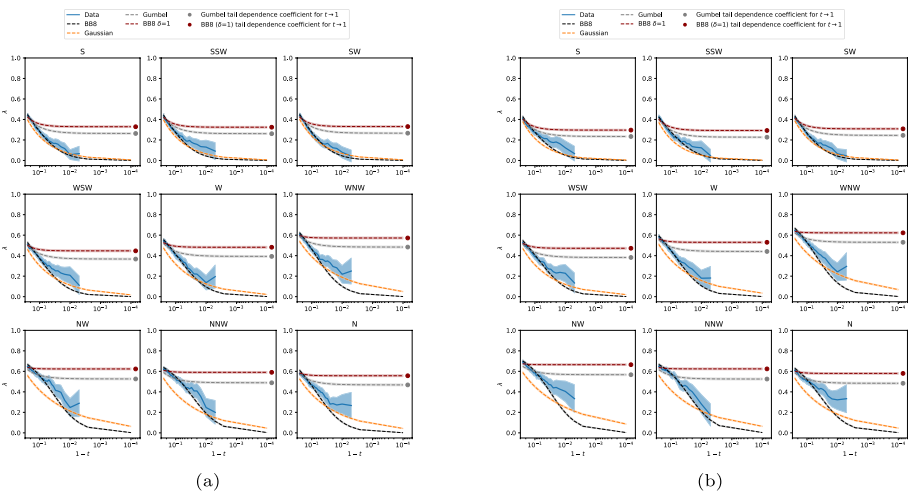


Fig. 38 Tail dependence for wind speed and surge height for different percentile thresholds from the storm data (in blue), compared to a fitted BB8 copula (dashed black), a Gaussian copula (dashed orange) and a Gumbel copula (dashed gray) per wind direction at location Den Helder (**a**) and Harlingen (**b**). Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

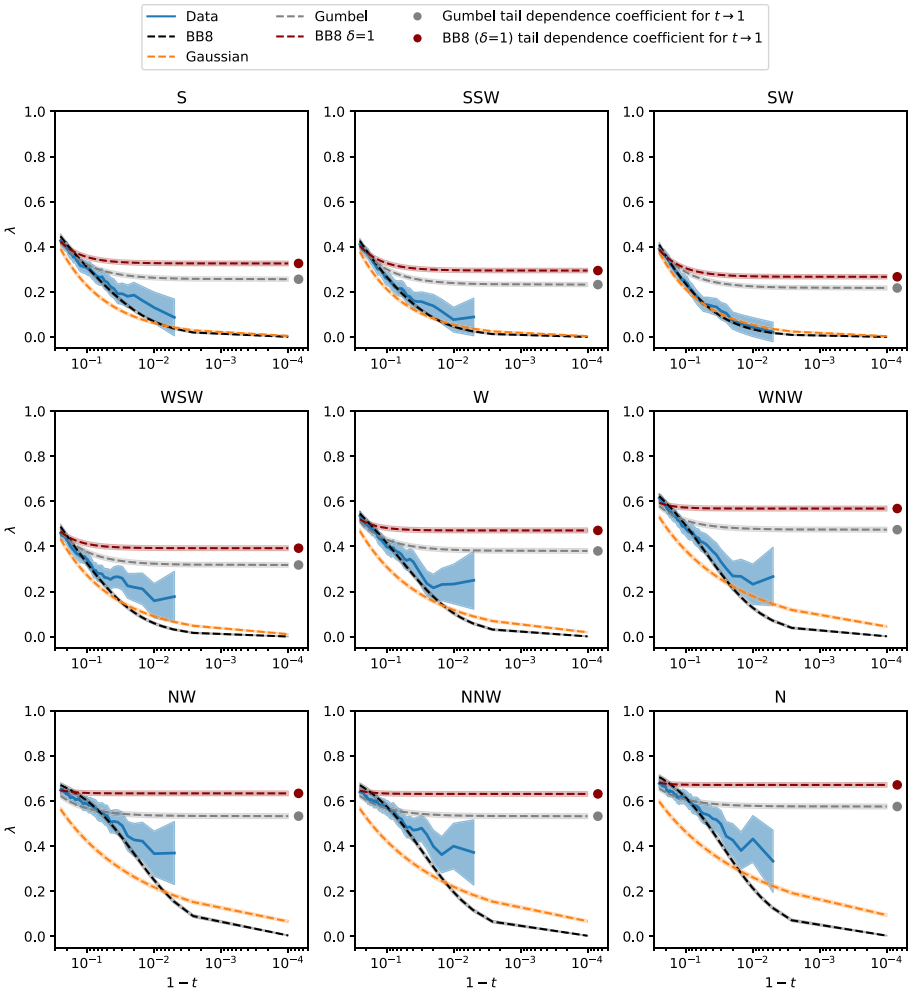


Fig. 39 Tail dependence for wind speed and surge height for different percentile thresholds from the storm data (in blue), compared to a fitted BB8 copula (dashed black), a Gaussian copula (dashed orange) and a Gumbel copula (dashed gray) per wind direction at location Delfzijl. Tail dependence coefficient curves are truncated when the number of data pairs becomes smaller than five

Appendix E: Hydra-NL data

See Fig. 40.

Acknowledgements The authors would like to thank Henk van den Brink and Cees de Valk from KNMI for providing the SEAS5-WAQUA dataset and for reviewing this manuscript. Also, the authors would like to thank the editors and the anonymous reviewers for their valuable insights during the review process. This research was conducted within the ODDS TKI-project, and supported by Rijkswaterstaat, TU Delft, TNO, HKV and Deltares.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Paulina Kindermann. The first draft of the manuscript was written

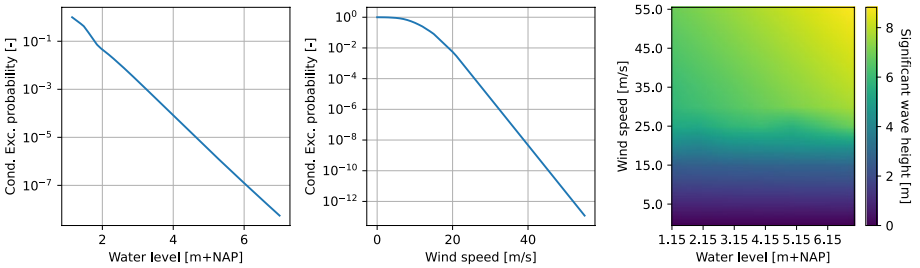


Fig. 40 Hydra-NL CDFs for wind speed and sea level for wind direction *NW* at Hoek van Holland (left and center panels) and the corresponding interpolation grid for deriving the significant wave height H_s (right panel)

by Paulina Kindermann and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research was lead and supported by the Extreme Sea Levels Impact on Flood Defenses in Coastal Areas (ODDS) consortium (TU Delft, Rijkswaterstaat, TNO, HKV and Deltares) nr: TDEL/2023/070-TU10, of The Delta Technology Knowledge and Innovation Top Consortium program (TKI).

Data availability Are available upon request.

Declarations

Conflict of interest Paulina Kindermann is employed as a consultant at HKV for one day a week. José A. A. Antolinéz and Oswaldo Morales-Nápoles declare they have no relevant financial interests to disclose.

References

Agresti A, Coull BA (1998) Approximate is better than exact for interval estimation of binomial proportions. Am Stat 52(2):119–126

Akaike H (1998) Information theory and an extension of the maximum likelihood principle. In: Selected papers of Hirotugu Akaike. Springer, pp 199–213

Bakker AM, Rovers DL, Mooyaart LF (2025) Storm surge clusters, multi-peak storms and their effect on the performance of the Maeslant storm surge barrier (The Netherlands). J Mar Sci Eng 13(2):298

Batstone C, Lawless M, Tawn J et al (2013) A UK best-practice approach for extreme sea-level analysis along complex topographic coastlines. Ocean Eng 71:28–39. https://doi.org/10.1016/j.oceaneng.2013.02.003.

Bauer P, Thorpe A, Brunet G (2015) The quiet revolution of numerical weather prediction. Nature 525(7567):47–55. https://doi.org/10.1038/nature14956

Benito I, Eilander D, Kelder T et al (2025) Pooling seasonal forecast ensembles to estimate storm tide return periods in extra-tropical regions. J Geophys Res: Oceans 130(8):e2025JC022614

Camus P, Mendez FJ, Medina R et al (2011) Analysis of clustering and selection algorithms for the study of multivariate wave climate. Coast Eng 58(6):453–462. https://doi.org/10.1016/j.coastaleng.2011.02.003

Camus P, Haigh ID, Quinn N et al (2024) Tracking the spatial footprints of extreme storm surges around the coastline of the UK and Ireland. Weather Clim Extremes 44:100662. https://doi.org/10.1016/j.wace.2024.100662
Caspers J, Kindermann P (2023) Hydraulic load model for the Dutch shore-analysis about the relevance of time aspects. Technical Report PR4529.20, HKV Lijn in Water, Lelystad
Caspers J, Kindermann P, Rongen G, et al (2025) Storm surge hydrographs and characteristics along the Dutch coast. An analysis from simulation data (manuscript under review)
Chbab H (2015) Waterstandsverlopen kust. wettelijk toetsinstrumentarium wti-2017. Technical report, Deltares
Chbab H, de Waal H (2017) Achtergrondrapportage hydraulische belastingen. wettelijk beoordelingsinstrumentarium 2017. Technical report, Deltares
Clayton DG (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65(1):141–151
Coles S, Heffernan J, Tawn J (1999) Dependence measures for extreme value analyses. Extremes 2:4 (365):339-365. https://doi.org/10.1023/A:1009963131610
De Valk C, Van den Brink H (2023) Update van de statistiek van extreme zeewaterstand en wind op basis van meetgegevens en modelsimulaties. Technical Report TR-406, KNMI. https://www.knmi.nl/kennis-en-datacentrum/publicatie/update-van-de-statistiek-van-extreme-zeewaterstand-en-wind-op-basis-van-meetgegevens-en-modelsimulaties (in Dutch)
De Valk C, Van den Brink H (2024) An appraisal of the value of simulated weather data for quantifying coastal flood hazard in The Netherlands. Nat Hazards Earth Syst Sci. https://doi.org/10.5194/egusphere-2024-912
Deltacommissie (1961) Onderzoekingen betreffende de opzet van het deltaplan en de gevolgen van de werken, deel 5. Technical report, Deltacommissie
Diakomopoulos F, Antonini A, Bakker AMR et al (2024) Probabilistic characterizations of flood hazards in deltas: application to Hoek van Holland (Netherlands). Coast Eng 194:104603
Duits M, Kuijper B (2019) Hydra-nl systeemdocumentatie. Technical Report PR4022.10, HKV Lijn in water
Dullaart JC, Muis S, De Moel H et al (2023) Enabling dynamic modelling of coastal flooding by defining storm tide hydrographs. Nat Hazard 23(5):1847–1862. https://doi.org/10.5194/nhess-23-1847-2023
ECMWF (2021) SEAS5 user guide. Technical report, ECMWF
Frank MJ (1979) On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. Aequ Math 19:194–226
Gautier C, Groeneweg J (2012) Achtergrondrapportage hydraulische belasting voor zee en estuaria. Technical report, Deltares
Geerse C (2002) Bivariate correlatiemodellen met exponenti le en asymptotisch exponentiële marginale verdelingen. Technical report, RIZA
Geerse C, Diermanse F (2006) Correlaties en meerdimensionale statistiek. Technical report, HKV Lijn in water and WL—Delft Hydraulics
Geerse C, Caspers J, Kindermann P (2022) Hydraulic load model for the Dutch shore. Technical Report PR4529.20, HKV Lijn in water
Genest C, Favre AC, Béliveau J et al (2007) Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. Water Resources Res 43(9):W09401
Genest C, Rémillard B, Beaudoin D (2009) Goodness-of-fit tests for copulas: a review and a power study. Insurance Math Econ 44(2):199–213. https://doi.org/10.1016/j.insmatheco.2007.10.005
Gerritsen H, De Vries H, Philippart M (1995) The Dutch continental shelf model. Coast Estuar Stud 47:425–467. https://doi.org/10.1029/CE047P0425
Gumbel EJ (1960) Bivariate exponential distributions. J Am Stat Assoc 55(292):698–707
Haigh ID, Marcos M, Talke SA et al (2023) GESLA version 3: a major update to the global higher-frequency sea-level dataset. Geosci Data J 10(3):293–314
Hersbach H, Bell B, Berrisford P et al (2017) Complete ERA5 from 1940: fifth generation of ECMWF atmospheric reanalyses of the global climate. https://doi.org/10.24381/cds.143582cf. https://cds.climate.copernicus.eu/datasets/reanalysis-era5-complete?tab=overview
HWBP (2024) Nationaal Basisbestand Primaire Waterkeringen. https://waterveiligheidsportaal.nl/
Jin X, Han J (2010) K-means clustering. In: Sammut C, Webb GI (eds) Encyclopedia of machine learning. Springer, Boston, MA, pp 563–564. https://doi.org/10.1007/978-0-387-30164-8_425
Joe H (1997) Multivariate models and multivariate dependence concepts. CRC Press, Boca Raton
Jonkman B, Schweckendiek T (2015) Developments in levee reliability and flood risk analysis in the Netherlands. In: Geotechnical safety and risk V. https://doi.org/10.3233/978-1-61499-580-7-50
Karwat A, Franzke C, Blender R (2022) Long-term trends of northern hemispheric winter cyclones in the extended ERA5 reanalysis. J Geophys Res: Atmos 127(22):e2022JD036952

Kennard R (1968) Computer aided design of experiments. In: Technometrics. American Statistical Association, Alexandria, VA, p 423

Kok M, Jongejan R, Nieuwjaar M et al (2016) Fundamentals of flood protection. Ministry of Infrastructure and the Environment and the Expertise Network for Flood Protection

Krieger D, Brune S, Baehr J et al (2024) Improving seasonal predictions of German bight storm activity. Nat Hazard 24(4):1539–1554. https://doi.org/10.5194/nhess-24-1539-2024

Meyer EMI, Gaslikova L (2024) Investigation of historical severe storms and storm tides in the German bight with century reanalysis data. Nat Hazard 24(2):481–499

Muis S, Aerts JC, Antolínez JA et al (2023) Global projections of storm surges using high-resolution CMIP6 climate models. Earth's Future 11(9):e2023EF003479. https://doi.org/10.1029/2023EF003479

National Center for Atmospheric Research Staff (eds) (2022) The climate data guide: climate forecast system reanalysis (CFSR). https://climatedataguide.ucar.edu/climate-data/climate-forecast-system-reanalysis-cfsr

Nelsen RB (2006) An introduction to copulas, 2nd edn. https://doi.org/10.1515/demo-2017-0006

Pupić Vurilj M, Antolínez JA, Muis S et al (2025) Storm surge hydrographs from historical observations of sea level along the Dutch North Sea coast. Nat Hazards 121:14147

Rijkswaterstaat (2024) Mean water level data (with respect to normaal amsterdams peil in m). https://waterinfo.rws.nl/

Salvadori G, Michele CD, Kottegoda NT et al (2007) Extremes in nature: an approach using copulas. Springer, Berlin

van den Brink H (2020) Het gebruik van de ecmwf seizoens-verwachtingen voor het berekenen van de klimatologie van extreme waterstanden langs de nederlandse kust. Technical Repot TR-385, KNMI. https://www.knmi.nl/kennis-en-datacentrum/publicatie/het-gebruik-van-de-ecmwf-seizoensverwachtingen-voor-het-berekenen-van-de-klimatologie-van-extreme-waterstanden-langs-de-nederlandse-kust (in Dutch)

Verlaan M, Zijderveld A, De Vries H et al (2005) Operational storm surge forecasting in the Netherlands: developments in the last decade. Philos Trans R Soc A Math Phys Eng Sci 363(1831):1441–1453. https://doi.org/10.1098/RSTA.2005.1578

Vrouwenvelder A, Steenbergen H, Diermanse F (2003) Belastingmodellen westerschelde/waddenzee vergelijking hydra-k en pc-ring (inclusief resultaten aanvullende studie). Technical report, TNO