ORIGINAL PAPER

Empirical evaluation of windstorm losses and meteorological variables over the Netherlands

M. d. S. Fonseca Cerda[1] · H. de Moel[1] · D. van Ederen[1,2] · J. C. J. H. Aerts[1,3] · W. J. W. Botzen[1] · T. Haer[1]

Abstract

This study investigates windstorm impacts by combining high-resolution wind hazard data with a unique asset-level insurance loss dataset, specifically focusing on the Netherlands. We conduct statistical analyses to associate wind hazard characteristics with spatial data on windstorm losses at various spatial aggregation levels (four-digit to nationwide postal codes). Different wind hazard intensities (e.g. maximum wind gust, maximum hourly wind speed) are derived using meteorological data from 2017 to 2021 (the same period as the loss data). This data is based on station and downscaled ERA5 reanalysis data. Results show that the recorded gust has a good correlation with damage components (r = 0.41–0.61). The downscaled reanalysis data on gust and daily maximum (hourly mean) wind speed also have a good correlation (r = 0.38–0.59), albeit a bit smaller than the observed gust. When comparing different levels of aggregated data (PC4—four-digit postal code, PC2—two-digit postal code, and NL—national level), the correlation between claim and loss ratios becomes more pronounced as the level of aggregation increases. In addition, at the aggregated data level of two-digit postal codes, we see a wind speed threshold (around the 98th percentile of the records, ~ 22 m/s), where both losses and reported claims begin to rise as wind speed increases. Nevertheless, with lower wind speeds, damages and reported claims become meaningful using more aggregated data (NL). Our findings highlight the complex link between hazard and damage variables for windstorm losses, offering valuable insights for insurance portfolios, risk assessment, and management.

1 Introduction

As one of the most expensive natural disasters in Europe, extreme windstorms have substantial societal impacts (Koks and Haer 2020; Severino et al. 2024). In 2007–2022, 10 out of the 20 most expensive natural hazard events in the Netherlands were related

---

✉ M. d. S. Fonseca Cerda
  m.d.s.fonseca.cerda@vu.nl

1   Institute for Environmental Studies (IVM), VU Amsterdam, Amsterdam, The Netherlands

2   Achmea, Zeist, The Netherlands

3   Deltares, Delft, The Netherlands

to windstorms (Verbond van Verzekeraars 2017, 2023). Despite being among the most damaging and frequent weather extreme events, quantifying windstorm impacts is challenging due to limitations in the available data. For instance, the length of hazard observations is usually short, and damage data is usually postdisaster insurance data, which is not publicly available (Koks and Haer 2020; Kunreuther and Michel-Kerjan 2007). Additionally, even though there is low confidence in the observed trends in extreme storms associated with extreme hail and wind, according to an IPCC report (2021) and Little et al. (2023), the current expectation is to have a slightly increased frequency and amplitude for both hazards in northern, central, and western Europe.

Regional- and large-scale studies have explored windstorm risks in Europe (Bonazzi et al. 2012; Della-Marta et al. 2009, 2010, 2012; Donat et al. 2010a, b, 2011a, b; Koks et al. 2019; Koks and Haer 2020; Pinto et al. 2007; Schwierz et al. 2010). These studies estimate potential losses based on the statistical storm loss model developed by Klawa and Ulbrich (2003). The model analyses the hazard–loss relationship as a function of wind gust and annual insurance loss. Studies usually use either daily maximum wind speed or wind gust speed when available, and both variables provide acceptable results (Donat et al. 2010a).

Both wind speed variables, daily maxima and gust, are available at the point locations of weather stations. Alternatively, reanalysis data can be used, which provides a larger coverage of average wind speeds. However, in reanalysis data, gust speed estimations are based on parametrisations, which are calibrated differently for each region based on their specific point observations (Sheridan 2011). These gust estimates are difficult to compare on a larger scale. Thus, to overcome spatial and model heterogeneity, Klawa and Ulbrich (2003) have proposed normalising wind gust speeds to local extreme wind gust speeds.

Moreover, in addition to wind speed variables, the existing literature has not evaluated the suitability of different high-resolution meteorological variables associated with windstorms and the appropriate intensity threshold to estimate potential losses. Such an evaluation could give insights into potential application of proxies to further evaluate windstorm losses, for example, when evaluating future scenarios due to climate change. In addition to providing a better understanding of the wind-hazard variables, high-resolution damage data can improve the evaluation of the hazard–loss relationship. The hazard–loss relations are usually derived using aggregated annual losses or aggregated event losses and not high-resolution (e.g. object-level) damage data, mainly because detailed damage loss data is insurance companies' proprietary information. However, studies have demonstrated the benefit of applying high-resolution damage data to estimate windstorm damage (Donat et al. 2011a).

In this study, we have access to a unique high-resolution dataset on windstorm losses for the Netherlands from Achmea over 2017–2022. Combining this loss data with windstorm hazard data from the Royal Netherlands Meteorological Institute (KNMI) provides a unique opportunity to closely analyse the spatiotemporal hazard–loss patterns of windstorm damage. The primary goals of this research are to define the best windstorm hazard variable for assessing windstorm damages and to outline the parameters that should be used, such as the threshold at which losses start to occur. For these purposes, we use high-resolution windstorm loss data from Achmea and statistically assess which hazard variable fits best with the loss data while evaluating the spatial resolution of aggregated losses.

This study is organized as follows: Sect. 2 describes the data and methods used. Section 3 presents the main results, and Sect. 4 discusses these findings and concludes the study.
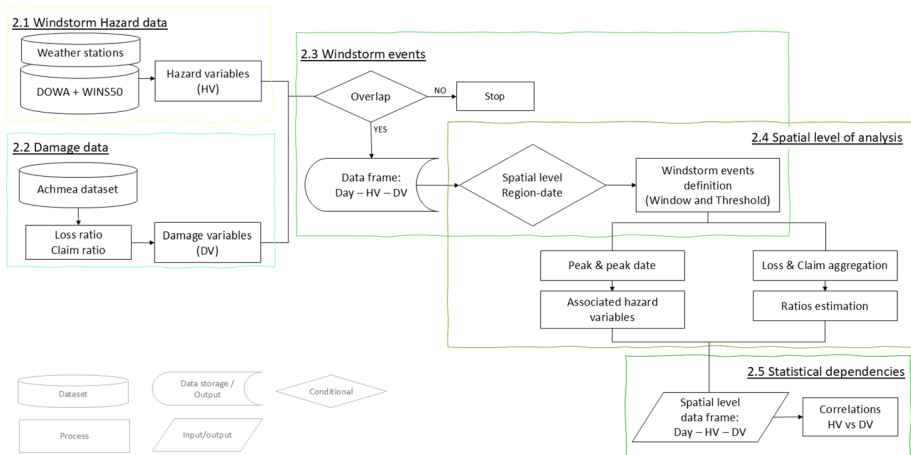
Fig. 1 Methodology. Steps 2.1 and 2.2 of the general workflow involve preprocessing windstorm hazard and damage data, followed by steps 2.3 and 2.4, which match the hazard and damage data for each of the spatial levels of aggregation and define windstorm events based on a wind gust threshold and a window. Finally, step 2.5 estimates the statistical dependencies between hazard and damage data

2 Data and methods

Figure 1 shows the overall methodology to assess windstorm losses in the Netherlands. First, we analyse windstorm data and extract different hazard variables (Sect. 2.1). Furthermore, historic windstorm loss data for private households (buildings) in the Netherlands is available from Achmea's non-life insurance division for 2017–2022 (Sect. 2.2). By defining windstorm events (Sect. 2.3), we limit our dataset to days exceeding a certain wind gust speed threshold. Additionally, we aggregate loss data at different spatial levels (Sect. 2.4). Next, we test the dependencies between different hazard variables and the event loss data using statistical methods such as Spearman correlations and linear regressions (Sect. 2.5).

2.1 Windstorm hazard data

Wind speed is the most commonly used variable to assess windstorm risk and studies have proved that it is a suitable variable (Bonazzi et al. 2012; Della-Marta et al. 2009, 2010, 2011b, 2012; Donat et al. 2010a, b, 2011a; Klawa and Ulbrich 2003; Koks et al. 2019; Koks and Haer 2020; Pinto et al. 2007; Schwierz et al. 2010). However, we also evaluate other meteorological variables such as air pressure, temperature, and relative humidity in this paper. We include such variables to assess their potential use as complementary variables and/or proxies (e.g. in climate change and future scenario evaluation). We use three datasets on wind hazards; Table 1 summarises them.
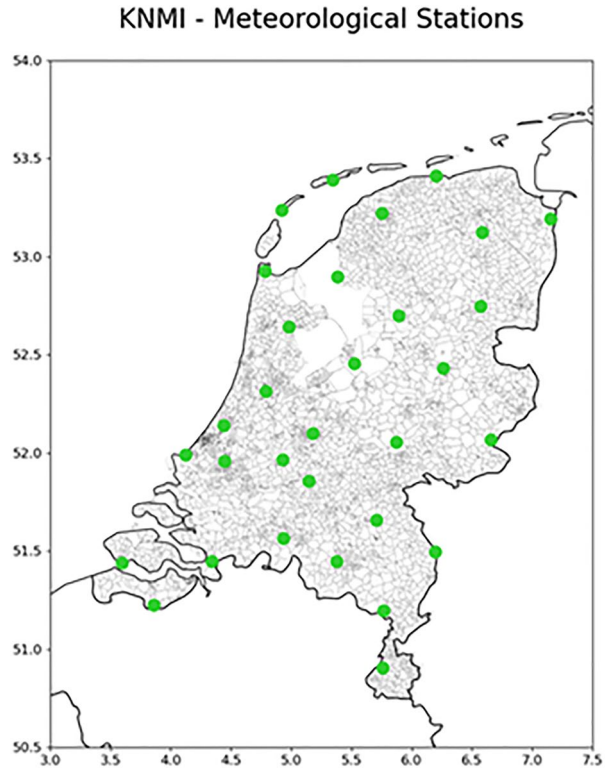
2.1.1 Daily weather station data

We use daily weather data from the Royal Netherlands Meteorological Institute (KNMI) measuring stations. The daily weather data provides weather observations such as temperature, sunshine duration, cloud cover, visibility, air pressure, precipitation, and wind. While

Table 1 KNMI windstorm datasets

Dataset	DOWA	WINS50	Weather stations
Name	Dutch Offshore Wind Atlas	Winds of the North Sea	Daily observations
Period	2008–2019	2019–2021	2017–2021
Spatial Resolution	2.5 km grid	2.5 km grid	31 stations
Covered area	$\sim 1970 \times \sim 1970$ km^2 (Lon: 1°–10°, Lat: 50°–55°)	$\sim 1970 \times \sim 1970$ km^2 (Lon: 1°–10°, Lat: 50°–55°)	~ 1350 km^2 mean area per station
Vertical profile	10, 20, 40, 60, 80, 100, 120, 140, 150, 160, 180, 200, 220, 250, 300, 500, 600 m	10, 20, 40, 60, 80, 100, 120, 140, 150, 160, 180, 200, 220, 250, 300, 500, 600 m	10 m
Temporal resolution	Hourly	Hourly	Daily
Models and data	ERA5, downscaled with HARMONIE-AROME	ERA5, downscaled with HARMONIE-AROME	Average of the 24-h observations Maximum and minimum measured values
Variables:			
Degrees	Wind direction (D)	Wind direction (D)	
m/s	Wind speed (F)	Wind speed (F)	Wind speed (S–F)
m/s	Parameterised wind gust (G)	Parameterised wind gust (G)	Wind gust speed (S–G)
Pa	Air pressure (P)	Air pressure (P)	
%	Relative humidity (Q)	Relative humidity (Q)	
K	Temperature (T)	Temperature (T)	Temperature (S–T)

Fig. 2 KNMI Meteorological Stations in the Netherlands

approximately 50 stations on both land and sea are available covering different periods (KNMI, Daggegevens Van Het Weer in Nederland, n.d.), we only use the 31 inland stations (Fig. 2) that have records over the period of interest (2017–2021) for which we have loss data (see Sect. 2.2). From the station data, we use temperature (S–T), daily mean, maximum and minimum windspeed (S–F), and maximum wind gust (S–G), all measured at an elevation of 10 m (KNMI, Daggegevens Van Het Weer in Nederland n.d.). We observe that the maximum wind gust provided by KNMI is rounded to the nearest integer.

2.1.2 Dutch offshore wind atlas (DOWA) and winds of the North Sea in 2050 (WINS50)

Two different downscaled reanalysis datasets on wind hazards are processed by KNMI: (1) the Dutch Offshore Wind Atlas (DOWA; Wijnant et al. 2019) and (2) the Winds of the North Sea in 2050 (WINS50; Baas 2024). The DOWA data provides hourly average values of wind climatology for the Netherlands at a high resolution (2.5×2.5 km^2). The data is based on 11 years (2008–2019) of ERA5 reanalysis data, which is downscaled using the weather model HARMONIE. The DOWA wind data is extensively validated using observations from satellite data, data from offshore stations, lidar measurements, and data from the Cabauw tower measurement station in the Netherlands (Wijnant et al. 2019). WINS50 is an extension of DOWA by three more years (2019–2021) and is made in the same way as DOWA. However, for WINS50, an updated version of the HARMONIE model is used to downscale the global ERA5 reanalysis to hourly information at a higher resolution (van Stratum et al. 2022). Combining these hazard datasets allows covering a period of 14 years

(2008–2021), which includes the period of interest (2017–2021) for which we have loss data. Consequently, the DOWA dataset is used for the years 2017–2018 and the WINS50 dataset for 2019–2021.

From the downscaled reanalysis data, we use wind direction, wind speed, air pressure, relative humidity, and temperature. Even though all variables are available at eight elevations (10, 20, 40, 60, 80, 100, 150, and 200 m), we use the variables at a 10 m elevation, aligning with standard practices, and since we did not find significant differences for elevations ranging from 10 to 200 m (not shown). Finally, we also compute the daily mean, maximum, and minimum of each hourly variable to have the same temporal resolution as the weather station data.

2.1.3 Gust parametrisation

Wind gust speed is unavailable in the downscaled reanalysis datasets. Therefore, the wind gust time series are estimated using the daily maxima wind speed and following the approach by van den Brink (2020). This method uses the shear between the wind at two heights (z_1 and z_2, with $z_1 < z_2$):

$$G(z) = U(z) + \ln(\alpha)\frac{U(z_2) - U(z_1)}{\ln(z_2) - ln(z_1)} \tag{1}$$

Here $\alpha = e^{2.4*2.65*0.41}$, G is the gust, and U is the daily maxima wind speed, both at a certain height above the surface (z). The heights z_1 and z_2 that are used are 10 m and 200 m, respectively.

After preprocessing the downscaled reanalysis datasets, we derive the following daily hazard variables: daily maximum wind gust (G); daily mean, maximum, and minimum of the hourly mean wind speed (F); daily mean of the hourly wind direction (D); daily mean, maximum, and minimum of the hourly mean air pressure (P); daily mean, maximum, and minimum of the hourly mean relative humidity (Q); and daily mean, maximum, and minimum of the hourly mean temperature (T). These variables, in combination with the weather station variables–temperature (S–T), daily mean, maximum and minimum wind-speed (S–F), and maximum wind gust (S–G)—are the final hazard variables at a daily temporal resolution that are used in further analysis.

2.2 Damage data

To represent wind damage, we select four variables, namely absolute losses, loss ratio, number of claims, and claim ratio.

2.2.1 Absolute losses

Postdisaster loss data from the Dutch insurance company Achmea is used to represent the historical losses for the residential sector (buildings). Even though Achema does not insure every building in the Netherlands (see Appendix A), this database represents a major subset of the larger insurance market, since Achmea approximately insures one-fourth of the non-life market in the Netherlands according to Fitch Ratings (2023). The full dataset consists of 297,352 entries and covers 6 years, from 1 January 2017 to 31

Table 2 Storm claims dataset

Dataset	Private buildings	
	Full	Processed
Period	01/01/2017	01/01/2017
	31/12/2022	31/12/2021
Claims	221,693	101,057
Damage (€ million)	325.70	115.59
Total policy holders and total insured value per year		
	Policy holders	Insured value (€ million)
2017	28,282	6993.64
2018	70,894	19,301.99
2019	36,154	9924.79
2020	44,877	12,467.60
2021	41,486	6993.64

December 2022, and it contains all claims and policy contract data (e.g. location and building characteristics) for different insurance coverages.

We preprocess the datasets by only considering 221,693 entries over 5 years, from 1 January 2017 to 31 December 2021, due to the hazard data availability (see Sect. 2.1) and by filtering the claims that are solely related and reported as 'storms', considering 111,423 entries where a claim has been made. This 'storm' label includes both winter and summer events that can be associated either with extra tropical cyclones or convective storms. We only use the following five variables from the Achmea dataset: postal code, claim date, damage burden, deductible, and insured value.

Some remarks about the insurance variables are as follows: (1) the claim date can differ from the actual date of damage since the policy holder is the one to report this; (2) if the claim has been settled, the damage burden is equal to the final amount paid to the policy holder; (3) if the claim has not been settled yet, the damage burden will reflect the best estimate; and (4) the dataset is truncated at the deductible, that is, if damage is equal to 0 (zero) in the dataset, it can indicate errors, can mean the claim has been made but the total damage is not higher than the deductible and the policy holder has covered it, or can also include other non-damage-related adjustments that were below the deductible. Therefore, we do not know how many observations are below the deductible. Given these factors, we assume that the total damage is the reported damage plus the deductible, as the policy holder is obliged to pay a (non-constant) deductible every time a storm causes damage. Furthermore, when preprocessing the damage data, we observe negative values for the total damage (reported damage plus deductible). Therefore, we apply an additional filter to avoid having negative values, leading to a final dataset with 101,057 claims (see Table 2).

All entries contain a six-digit postal code; however, to limit the computational time, we use the aggregated losses. We apply three different spatial aggregation levels (see Sect. 2.4). This usage allows for an evaluation of the relation between the hazard variables and losses at different spatial resolutions.

2.2.2 Loss ratio, number of claims, and claim ratio

It is common practice to use loss ratios to depict losses (Endendijk et al. 2023; Held et al. 2013; Khanduri and Morrow 2003). We calculate loss ratios as the ratio between claims and the total reconstruction value of the building, which equals its insured value. Using loss ratios controls for inflation rates, socioeconomic factors such as increases in wealth or changes in the number of insurance policies, and building characteristics such as building surface (Botzen et al. 2010; Donat et al. 2010a, 2011a, 2012; Pinto et al. 2007). As we are using aggregated loss data at three different postal code levels, we calculate the loss ratio using the accumulated insured claims and the accumulated total insured value for each area level. The total insured values also include the insured value of the policy holders without a claim. By including not only the total insured value of reported claims but also the insured value of policy holders without a claim, we account for the differences in the distribution of the insurance portfolio per area.

In addition to the reported losses, we study the total number of claims per postal code area per storm event. Based on the number of claims, calculate the ratio between the number of claims and the number of total policy holders, also known as the claims ratio (Prahl et al. 2015). This ratio provides insights into the risk exposure of a portfolio (Prahl et al. 2015; Welker et al. 2016). After preprocessing the damage data, we derive datasets that correspond to postal code areas, which contain aggregated damage, loss ratio, number of reported claims, and claim ratio over 5 years (2017–2021).

2.3 Windstorm events

The Achmea loss dataset includes a large number of reports associated with low damage over the years. Records of low losses could be caused by low wind speeds but also by errors in cause attribution within the claim entry, inaccurate event date recording, or delayed claim reporting. To remove noise from the hazard–loss dataset, we only consider events that exceed a predetermined threshold. In existing studies, this threshold is usually considered to be the 98th percentile of the daily maximum wind gust speed, assuming that the local conditions (buildings) are adapted to the wind intensities and only 2% of the windstorms cause damage (Bonazzi et al. 2012; Della-Marta et al. 2010; Donat et al. 2010a, 2011a, 2012; Klawa and Ulbrich 2003; Pinto et al. 2007; Schwierz et al. 2010). While Schwierz et al. (2010) have used 15 m/s as a condition for the minimum peak wind gust speed, other studies have used 9 m/s or population-weighted averaged wind speeds within a certain area (e.g. municipality level) to avoid low local 98th percentiles of wind speeds (Jaison et al. 2024; Karremann et al. 2014; Little et al. 2023). According to our dataset, the 98th percentile is almost equivalent to 22 m/s wind gust speed, and the mean 98th percentile is 21.48 m/s with a standard deviation of 1.85 m/s across the 31 meteorological stations. However, we observe damage incidents and claim reports at lower wind speeds. Considering these findings, we select a lower threshold of 15 m/s based on the wind gust speed recorded in the 31 meteorological stations (see Appendix B). The 15 m/s wind gust speed corresponds to a mean 84th percentile with a standard deviation of 7th percentile across the stations.

We set a 5-day window that begins 1 day before and ends 3 days after the occurrence of a 'peak' event as detected in the hazard data. The wind gust peak is the maximum recorded wind gust speed within a 5-day window. In cases where there are consecutive peaks over

multiple days, we retain only the largest peak to define the 5-day period. With the peak date serving as a reference for the daily hazard variables (minimum, mean, and maximum) associated with that speed. Damage variables (loss and claims) are summed over the 5 days, providing a comprehensive view of damage accumulation during the event. Since the claim date can differ from the actual date of damage (see Sect. 2.2), this window allows processing claims and captures the immediate effects prior to the damage report. It also accounts for scenarios in which the storm happens during the night. The 5-day period reduces temporal uncertainty by balancing data sufficiency and statistical robustness in the evaluation of hazard–damage relationships.

2.4 Spatial level of analysis

All damage data entries include a six-digit postal code, as explained in Sect. 2.2. To reduce computational time and evaluate different spatial levels, we aggregate damage and hazard variables into three different spatial levels (Fig. 3), using the standard postal code (PC) areas as a guide: four-digit postal codes (PC4; neighbourhood level), two-digit postal codes (PC2), and NL (national level).

Several processing processes are carried out to evaluate the hazard–damage relationship per PC area.

First, we obtain the centroid coordinates for each PC4 area from the Statistics Netherlands (CBS) dataset. We use a kd-tree algorithm to match each PC4 area to two hazard datasets, downscaled reanalysis and meteorological stations. This allows for efficient nearest neighbour searches, which are required to match PC4 areas with information from hazard datasets. As a result, this algorithm matches each PC4 area, based on its centroid, to the closest cell of downscaled reanalysis data as well as the nearest weather station. The distances between the PC4 area's centroid and the downscaled reanalysis cells are small (maximum ~ 3 km), while the distances to meteorological stations range from 0.52 to 81.61 km.

Second, losses and claims for a specific day that are reported on PC6 levels are aggregated to PC4. This aggregation results in a new damage and claims dataframe organised by PC4 areas and days.

Third, for each PC4 date, daily hazard variables (minimum, mean, and maximum; see Sect. 2.1) from the closest downscaled reanalysis cells and meteorological stations (see
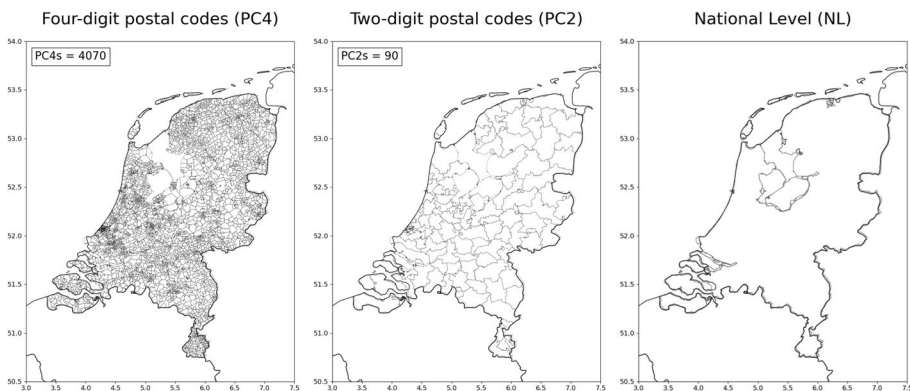

Fig. 3 Spatial levels of analysis, from left to right: 4,070 four-digit postal codes (PC4), 90 two-digit postal codes (PC2), and national level (NL)

first step) are added to the dataframe. Losses and claims for the PC2 and NL spatial levels are obtained by aggregating results from the PC4 area within each respective level. For the hazard variables, we calculate the mean, maximum, and minimum across the PC2 and NL levels.

Fourth, we apply the threshold and 5-day window, as discussed in Sect. 2.3, for each spatial level of analysis (PC4, PC2, and NL).

Fifth, the loss and claim ratios are calculated using the total insured value and policy holders in the corresponding spatial level of analysis (PC4, PC2, and NL, respectively).

After completing these steps, the datasets are reduced to 27,635 PC4 5-day-event; 5,338 PC2 5-day-event; and 212 NL 5-day-event combinations.

2.5 Statistical dependencies

We estimate the statistical dependencies between the different windstorm variables (daily maximum wind gust; daily mean wind direction; daily mean, maximum, and minimum wind speed; air pressure; relative humidity; and temperature; Sect. 2.1) and the damage variables, including damage, number of claims, loss, and claim ratios (Sect. 2.2). These statistical dependencies show which hazard variable best explains the losses. Using the Spearman rank correlation coefficients r_s in Eq. 2, ranging from -1 to 1, we measure the degree of association between two variables and the direction of the relationship (Schober et al. 2018). The absolute values indicate the strength of the relationship between the two variables x and y. In the case of a positive correlation, the two variables tend to increase or decrease simultaneously, resulting in a positive value. Conversely, for a negative correlation, one variable tends to increase when the other decreases (Chok 2010). If the correlation coefficient value is close to 0, the relationship between the variables is weak (Chen and Popovich 2002). This correlation does not carry any assumptions about the distribution of the data (Chen and Popovich 2002; Moran 2022), but the variables should have a monotonic relationship, meaning that as the value of one variable increases or decreases, the other variable consistently increases or decreases as well. The following equation is used to calculate the coefficient:

$$r_s = \frac{\sum_{i=1}^{n}((rank(x_i) - \overline{rank(x)})(rank(y_i) - \overline{rank(y)}))}{\sqrt{\sum_{i=1}^{n}(rank(x_i) - \overline{rank(x)})^2 \sum_{i=1}^{n}(rank(y_i) - \overline{rank(y)})^2}} \quad (2)$$

Additionally, a bivariate linear regression is used to evaluate the differences between the spatial levels. When there is a nonlinear relationship between the independent (hazard) and dependent (damage) variables, one popular method for handling these circumstances in regression models is to logarithmically modify the variables. The effective connection becomes nonlinear despite maintaining the linear model when one or more variable logarithms are used (Benoit 2011). Therefore, the following log-linear regression is used:

$$\log \widehat{Y}_i = \alpha + \beta X_i \quad (3)$$

In the above equation, $\log Y$ is the natural logarithm of the damage variable (damage, loss ratio, claims, or claim ratio) and X the hazard variable (see list of variables in Sect. 2.1). Moreover, α is the intercept, the expected value of $\log Y$ for $X = 0$. The average change in $\log Y$ for each unit increase in X is known as the slope (β). The slope estimate not

only indicates the direction and intensity of the linear relationship between X and $\log Y$ but also provides an interpretation of how $\log Y$ varies as X grows (Montgomery et al. 2021).

Although values of $\log Y$ may be predicted using this equation given a value of X, the use of linear regression in this study is not to predict but to measure the strength of the hazard–loss relationship. In order to gain sufficient prediction power, a multivariate regression should be used.

3 Results

3.1 Hazard–damage relationships

Figure 4 shows Spearman correlations between hazard (columns) and damage (rows) variables. We show the results for the PC4 aggregation level. Because the pattern is consistent across all spatial levels, we show the results for the three spatial levels in Appendix C.

As expected, we find that variables related to temperature (T), relative humidity (Q), and pressure (P) do not correlate well with the damage variables, although minimum pressure (Pmin) and maximum relative humidity (Qmax) have a reasonable correlation with the number of claims and claim ratio (up to − 0.35 and 0.35, respectively). This correlation makes sense because, for instance, low-pressure systems are associated with storm systems in the Netherlands. On the other hand, relative humidity increases whereas the temperature decreases when there is a windstorm. Various wind variables show a relatively good correlation with the damage variables (0.37–0.61). Mean wind speed has a weak correlation (<0.20) and should, thus, not be used in wind risk assessments. Specifically, daily maximum wind speed (Fmax) and wind gust speeds (G) show a strong correlation with the damage variables. The difference in correlations for maximum wind speed and gusts is very small. This finding indicates that, while wind gusts theoretically cause the most damage to houses, maximum wind speed is a useful proxy.

When comparing the station data (labelled 'S' in Fig. 4) with the downscaled reanalysis data, we find that wind gusts perform slightly better when taken from the nearest station as compared to the parameterised gust from the downscaled reanalysis data. The wind gust recorded at stations is also likely the closest to reality as it is directly measured, whereas the gust from the downscaled reanalysis data is a parametrisation and, therefore, involves uncertainties (van den Brink 2019; see Appendix D). Nevertheless, it is interesting to see that station data still performs best overall given its lower spatial coverage (31 stations for the whole of the Netherlands) and using the nearest station for any PC4. An explanation for the better performance of the wind gust recorded at stations, could be the flat topography of the Netherlands, which leads to more homogeneous wind speeds across the country.

	Dmean	Gust	Fmax	Fmean	Fmin	S-Gust	S-Fmax	S-Fmean	S-Fmin	Pmax	Pmean	Pmin	Qmax	Qmean	Qmin	Tmax	Tmean	Tmin	S-Tmax	S-Tmean	S-Tmin
Damage	0.07	0.38	0.37	0.07	-0.12	0.41	0.30	0.05	-0.10	-0.08	-0.15	-0.23	0.20	0.14	0.08	-0.08	-0.16	-0.19	-0.05	-0.15	-0.17
Loss_ratio	0.11	0.42	0.41	0.10	-0.12	0.44	0.37	0.11	-0.08	-0.07	-0.16	-0.25	0.26	0.19	0.12	-0.12	-0.19	-0.22	-0.10	-0.18	-0.18
Claims	0.11	0.59	0.56	0.20	-0.11	0.61	0.47	0.17	-0.07	-0.11	-0.24	-0.35	0.31	0.23	0.17	-0.11	-0.21	-0.25	-0.08	-0.19	-0.22
Claim_ratio	0.14	0.54	0.54	0.17	-0.12	0.56	0.49	0.18	-0.06	-0.10	-0.22	-0.34	0.35	0.28	0.18	-0.17	-0.24	-0.28	-0.15	-0.22	-0.23

Fig. 4 Spearman correlation between four damage variables (rows) and 21 hazard variables for the PC4 level. The hazard variables are coded as follows: Wind direction (D), wind speed (F), wind gust speed (G), air pressure (P), relative humidity (Q), and air temperature (T). S refers to station data, while all other hazard data are derived from the downscaled reanalysis datasets (Sect. 2.1)

(Agüera-Pérez 2012; Daly 2006; Gastón et al. 2008; Shen et al. 2018). For maximum wind speed, we see that the downscaled reanalysis dataset is better to use compared to the station data. This outcome is probably related to the better spatial coverage of the downscaled reanalysis dataset versus the stations, 2.5 km grid versus 31 stations in the whole country. The difference in the performance between maximum wind speed and wind gust speed from the downscaled reanalysis dataset might be because the downscaled reanalysis wind speed has been validated using observations. On the other hand, the downscaled reanalysis wind gust has been computed using a parametrisation and, therefore, more uncertainties might be involved in this computation.

Analysing the damage variables and their correlation with maximum wind speed and wind gusts, we observe that the loss ratio has a higher correlation (0.41–0.42) compared to the absolute damage (0.38–0.37). On the one hand, this finding is expected as loss ratios perform better when they are standardised with respect to the absolute damage (Endendijk et al. 2023; Botzen et al. 2010). On the other hand, the difference is not so significant. Notably, the damage variables related to the amount of claims, both the absolute amount of claims and the claim ratio, yield considerably better correlation coefficients (a few tenths of points). Specifically, the total number of claims gives the best correlation (0.61 with wind gusts). This finding implies that with increasing wind speeds, the number of buildings that are damaged increases.

3.2 Spatial levels evaluation

Similarly to the previous section, we also determine the correlation coefficients at the other spatial levels (see Appendix C). The results are generally in line with the PC4 results presented in Sect. 3.1, Fig. 4. However, we find that for the downscaled reanalysis data, the parameterised gust performs best for PC4, but when going to coarser resolutions (PC2 and NL), Fmax performs better. Furthermore, for the station data, the recorded gust performs noticeably better than the maximum wind speed. This finding could be due to the uncertainties related to the gust parametrisation: whereas for lower wind gust speeds (5–15 m/s) the parametrisation is mostly overestimated, for high wind gust speeds (15–40 m/s), it is underestimated (see Appendix D).

Figure 5 shows the scatter plots of loss (upper) and claim (bottom) ratios on the y-axis versus wind gust speed (blue: gust from stations, green: parameterised wind gust) on the x-axis at various spatial levels (from left to right: PC4 areas, PC2 areas, and NL). To improve clarity and facilitate a comparison between spatial levels, the ratios and the linear fit are plotted on a logarithmic scale. The 15 m/s threshold is applied to the wind gust speed measured at the stations (blue); therefore, lower wind speeds are observed for the parameterised wind gust (green). See Appendix D for more information about the difference in wind gust speed between station and downscaled reanalysis data.

The scatter plots show a similar trend when we study the differences by aggregation level. As we proceed from PC4 to NL (left to right), the correlation between the variables and the slope of the log-linear regression increase and the cloud of observations becomes narrower, leading to a clearer trend. Furthermore, the slopes of the linear fits go from 0.14–0.12 to 0.31–0.30. Although the relationship between the ratios and wind speed is less clear at the more localised PC4 level, particularly at lower wind speeds, a direct and stronger association is evident at NL, revealing that higher wind speeds increase both the loss and claim ratios. One possible explanation is that aggregating the data balances out differences in the portfolio, such as coverage, building, and environmental characteristics. For PC4, the differences in
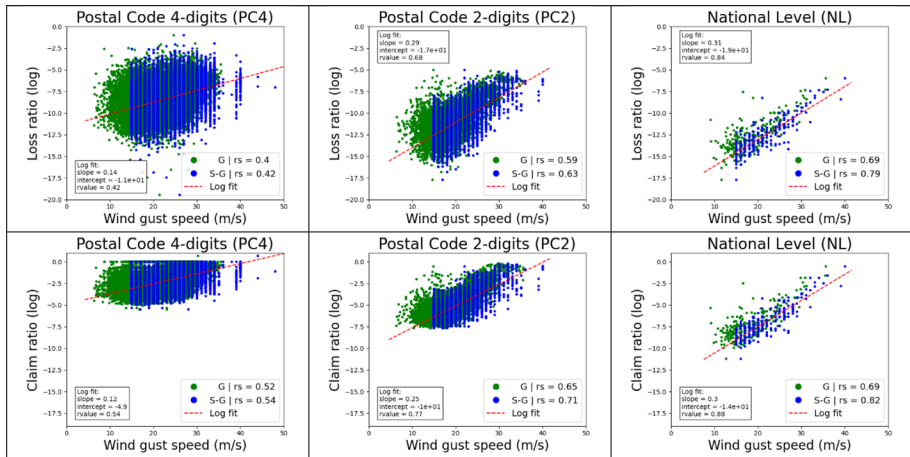
Fig. 5 Scatter plots of the 5-day windstorm events at different spatial levels (right: 27,635 PC4 5-day event, middle: 5338 PC2 5-day event, left: 212 NL 5-day event). Y-axis, plotted in log base 10, loss ratio (upper) and claim ratio (bottom), X-axis wind gust speed from stations (blue), and downscaled reanalysis (green) data. The log-linear fits, using the ratios and wind gust speed from stations combination (blue), are shown (red)

the building characteristics, environmental conditions, and number of policy holders between areas could be significant. On the other hand, at NL, we analyse the entire portfolio as one, considering the total number of policy holders and its total insured value.

When studying the PC4 level, claim ratios are closer to 0, or 1 in natural values as the graph is plotted on a log base 10 scale, than the loss ratios. This finding means that, while damage per PC4 may not be equal to the total insured value, especially at low wind speeds, the number of recorded claims may be close to the total number of policy holders at high wind speeds, consequently indicating damage to several buildings. This observation may be an indication that, while stronger winds may cause more damage to several buildings, the severity of damage to each structure may not change considerably.

Further analysis of the scatter plots reveals a wind speed threshold, particularly at the PC2 aggregation level, where both loss and claim ratios rise substantially as wind gust speed reaches approximately 22 m/s. This finding aligns with recognised criteria like the 98th percentile, which are frequently employed in comparable research (Bonazzi et al. 2012; Della-Marta et al. 2010; Donat et al. 2010a, 2011a; Klawa and Ulbrich 2003; Leckebusch et al. 2007; Pinto et al. 2007, 2012; Priestley et al. 2018; Severino et al. 2023; Schwierz et al. 2010). There is no significant increase in damage nor in the number of reported claims with the increase in wind speed per PC4 or PC2 below this threshold. However, at higher aggregated levels such as NL, the contribution of lower wind speeds to overall losses becomes obvious, emphasising the need to consider a wider range of wind speeds when analysing national losses or insurance portfolios.

4 Discussion and conclusions

Our research investigates windstorm impacts and the relationship between various wind and weather variables and damage characteristics in a given insurance portfolio. We employ two hazard datasets, namely recordings from 31 meteorological stations and downscaled reanalysis. Despite the low spatial resolution of the meteorological stations, the recorded gust has a good correlation with damage components (r = 0.41–0.61). Furthermore, the parameterised gust and daily maximum (hourly mean) wind speed from the reanalysis data also have a good correlation (r = 0.38–0.59), albeit significantly smaller than the reported gust. This finding is especially crucial when it comes to data availability. Reanalysis data is becoming increasingly popular and is generally available on a large scale in Europe (Balsamo et al. 2015; Bollmeyer et al. 2015; Gonzalez-Aparicio et al. 2017), with the potential for regional applications when downscaled and made available at a higher resolution. Additionally, the density of meteorological stations varies per country, which leads to differences in the spatial data quality of meteorological variables, and the topography of the research area is also an essential consideration (Daly 2006; Jaison et al. 2024). The high correlation with meteorological stations in this study could be attributed to the flat topography of the Netherlands (Agüera-Pérez 2012; Gastón et al. 2008; Shen et al. 2018). However, it may be difficult to attribute this correlation to topography in more complex orographic areas, and reanalysis data mixed with local measurements may provide better estimates (Carvalho et al. 2013).

The majority of the existing research has used either the daily maximum or wind gust speed where available, yielding reasonable results in loss models (Bonazzi et al. 2012; Della-Marta et al. 2010; Donat et al. 2010a, 2011a, 2012; Jaison et al. 2024; Karremann et al. 2014; Klawa and Ulbrich 2003; Little et al. 2023; Pinto et al. 2007; Schwierz et al. 2010). While these studies do not analyse the performance of the variables, they assess the correctness of the outcomes when modelling the losses. Some studies (Donat et al. 2010a) examine the accuracy of findings obtained with either the daily maximum wind speed or the wind gust, demonstrating that both variables perform adequately. This finding is consistent with our results, which indicate that both variables can be used. Although we find no significant variations between the wind speeds (gust or maximum), our study underlines the importance of consistency in hazard variable selection. The wind gust speed is higher than the daily maxima wind speed; therefore, variations in variable usage for calibrating loss–hazard functions and estimating losses might result in damages being overestimated or underestimated.

We investigate the relationship between the loss and claim ratios and the corresponding wind gust speed. Overall, there is a stronger link with the claim ratio than with the loss ratio, particularly for less aggregated data (PC4). This finding could be due to a greater quantity of damage across several buildings in addition to the exacerbated damage to particular structures. However, a more granular analysis, one including building characteristics, is needed to evaluate the damage changes and vulnerabilities of particular structures (Donat et al. 2011a).

To analyse windstorm risk, it is necessary to evaluate both losses and the number of claims, as the latter has a larger correlation with wind speed. When comparing different levels of aggregated data (PC4, PC2, and NL), the correlation between claim and loss ratios becomes more pronounced as the aggregation level increases. The high correlation between wind speed and total losses and claims at NL suggests that when specific damage (or hazard) information is unavailable, an analysis at a coarse resolution can provide useful

insights into the overall expected damages, especially one that has similar spatial coverage and building characteristics within the portfolio. The finding that aggregating data has explanatory power could be useful for insurance (and reinsurance) businesses that could use their entire portfolio to determine insurance premiums and/or their own risk.
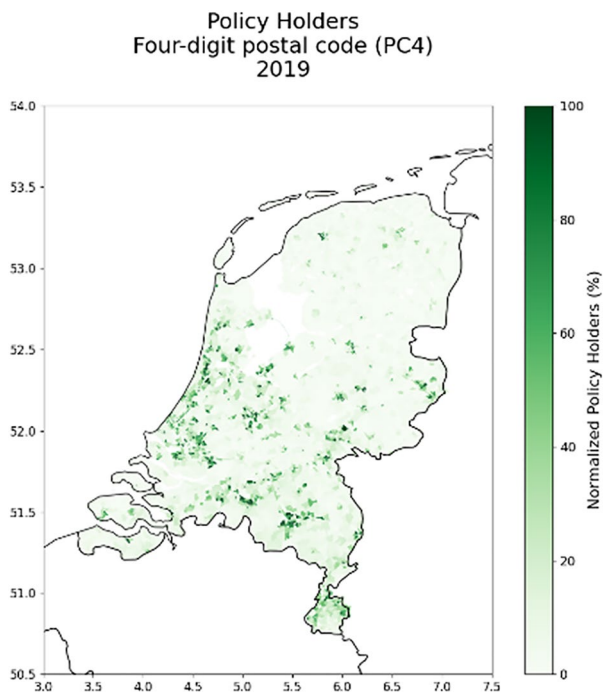
Additionally, at the aggregated data level of PC2, we see a wind speed threshold around the 98th percentile of the records, ~22 m/s, where both losses and reported claims begin to rise as wind speed increases. This criterion is consistent with previous research (Bonazzi et al. 2012; Della-Marta et al. 2010; Donat et al. 2010a, 2011a; Klawa and Ulbrich 2003; Leckebusch et al. 2007; Pinto et al. 2007, 2012; Priestley et al. 2018; Severino et al. 2023; Schwierz et al. 2010). Nevertheless, when evaluating more aggregated data (NL), we see that damages and reported claims become meaningful at lower wind speeds. While the damage per building may not be considerable at lower wind speeds, noticeably more claims are reported for these lower wind speeds, resulting in larger aggregated damage. Consequently, the 98th percentile-based models have limited applicability in the lower loss range due to their incapacity to capture losses associated with lower wind speeds (Jaison et al. 2024; Prahl et al. 2015).

Our findings highlight the complex link between hazard and damage variables in an insurance portfolio, offering valuable insights for risk assessment and management. Future studies should address limitations associated with the inclusion of building characteristics as explanatory variables, with a focus on increasing damage estimation at a finer asset level by accounting for additional building information. After using high-resolution data to evaluate building attributes and vulnerabilities for establishing a hazard–loss model, the model can easily be applied at a relatively coarse aggregation level, including the claim ratio. This finding is especially important given the limited (or nonexistent) availability of high-resolution postdisaster insurance data. As a result, the significant association between total loss and wind speed in a specific event at NL demonstrates that valid results may be obtained despite coarse resolution. Furthermore, future research should evaluate the co-occurrence of meteorological drivers such as wind and precipitation or wind and hail. Other studies have focused on the statistical relationship between damage data and meteorological parameters such as temperature (Botzen and Bouwer 2016). Therefore, correlations with other meteorological parameters such as air pressure, temperature, and relative humidity could provide significant information concerning prospective applications for assessing future climatic scenarios.

Appendix A

Figure 6 shows the distribution of policy holders across four-digit postal code (PC4) areas for 2019. The number of policy holders has been normalised by dividing the number of policy holders in a specific PC4 region by the area and multiplying by 100.

Fig. 6 Distribution of policy holders across four-digit postal code (PC4) areas for 2019

Appendix B

Figure 7 shows the different peaks associated with using the 98th percentile of the recorded wind gust speed of Cabauw station (red) or a 15 m/s wind gust speed (blue) as a threshold. Additionally, Figs. 8 and 9 show the timeseries of 2 years (2017 and 2019, respectively) of the maximum wind gust speed and the associated damage variables at the national level. Finally, Fig. 10 shows a threshold evaluation using the percentage of damage, claims, and windstorm days (blue, green, and red respectively; Y-axis) excluded in the analysis for different wind gust thresholds (X-axis). By using a wind gust threshold of 15 m/s instead of 22 m/s, approximately 8% of the damage and claims are included in the analysis. More importantly, this threshold also captures nearly 35% of low-intensity, high-frequency events (i.e. windstorm days), which might be particularly relevant for certain stakeholders, such as insurance companies.
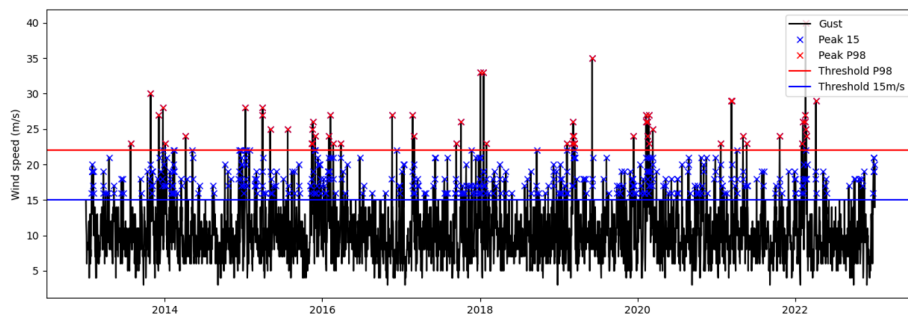

Fig. 7 Wind gust speed time series at Cabauw meteorological station. Threshold and peaks are represented for the 98th percentile (red) and 15 m/s (blue)
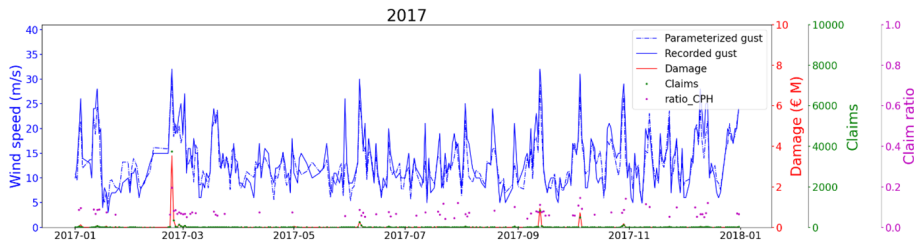
Fig. 8 Maximum wind gust speed time series at the national level based on downscaled reanalysis data (dotted blue line) and meteorological stations (continuous blue line), as well as the associated national damage (€ M; red), number of claims (green), and claim ratio (magenta), 2017
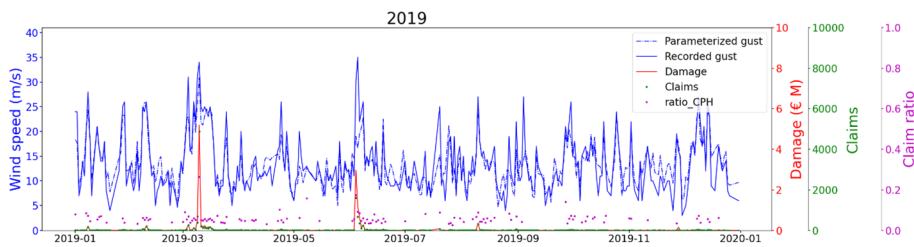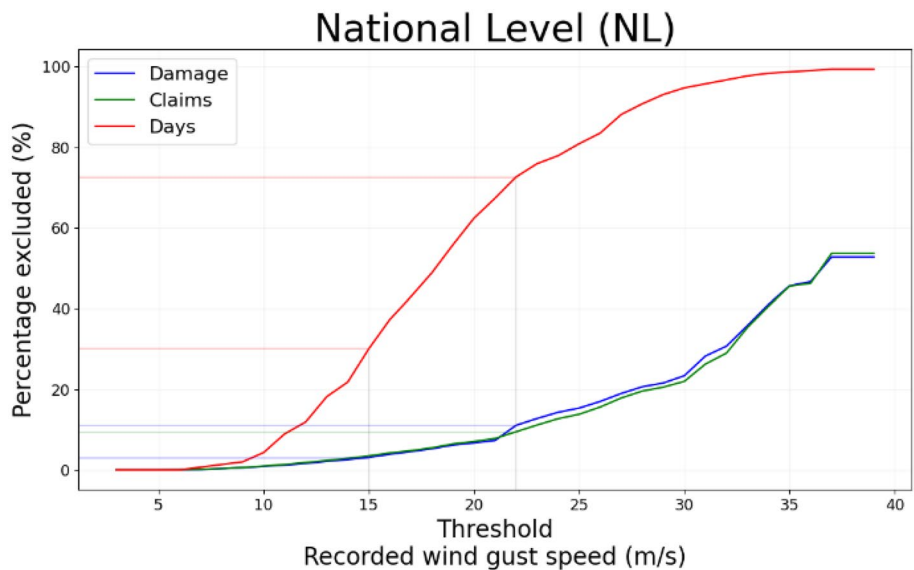

Fig. 9 Maximum wind gust speed time series at the national level based on downscaled reanalysis data (dotted blue line) and meteorological stations (continuous blue line), as well as the associated national damage (€ M; red), number of claims (green), and claim ratio (magenta), 2019

Fig. 10 Threshold evaluation: percentage of damage, claims, and windstorm days (blue, green, and red respectively; T-axis) excluded in the analysis for different wind gust thresholds (X-axis)

Appendix C

Figure 11 shows correlations between hazard (columns) and damage (rows) variables. We show the results for PC4 (upper), PC2 (middle), and NL (bottom) spatial levels. The pattern is consistent across all spatial aggregation levels (see Sect. 3.1).

		Dmean	Gust	Fmax	Fmean	Fmin	S-Gust	S-Fmax	S-Fmean	S-Fmin	Pmax	Pmean	Pmin	Qmax	Qmean	Qmin	Tmax	Tmean	Tmin	S-Tmax	S-Tmean	S-Tmin
PC4	Damage	0.07	0.38	0.37	0.07	-0.12	0.41	0.30	0.05	-0.10	-0.08	-0.15	-0.23	0.20	0.14	0.08	-0.08	-0.16	-0.19	-0.05	-0.15	-0.17
	Loss_ratio	0.11	0.42	0.41	0.10	-0.12	0.44	0.37	0.11	-0.08	-0.07	-0.16	-0.25	0.26	0.19	0.12	-0.12	-0.19	-0.22	-0.10	-0.18	-0.18
	Claims	0.11	0.59	0.56	0.20	-0.11	0.61	0.47	0.17	-0.07	-0.11	-0.24	-0.35	0.31	0.23	0.17	-0.11	-0.21	-0.25	-0.08	-0.19	-0.22
	Claim_ratio	0.14	0.54	0.54	0.17	-0.12	0.56	0.49	0.18	-0.06	-0.10	-0.22	-0.34	0.35	0.28	0.18	-0.17	-0.24	-0.28	-0.15	-0.22	-0.23
		Dmean	Gust	Fmax	Fmean	Fmin	S-Gust	S-Fmax	S-Fmean	S-Fmin	Pmax	Pmean	Pmin	Qmax	Qmean	Qmin	Tmax	Tmean	Tmin	S-Tmax	S-Tmean	S-Tmin
PC2	Damage	0.08	0.55	0.56	0.40	0.06	0.60	0.45	0.32	0.06	-0.17	-0.29	-0.36	0.16	0.08	0.04	-0.05	-0.09	-0.11	-0.03	-0.08	-0.09
	Loss_ratio	0.13	0.59	0.59	0.44	0.09	0.63	0.51	0.37	0.10	-0.18	-0.29	-0.37	0.19	0.11	0.08	-0.08	-0.11	-0.12	-0.07	-0.10	-0.10
	Claims	0.08	0.64	0.65	0.48	0.07	0.71	0.55	0.40	0.09	-0.19	-0.31	-0.40	0.16	0.06	0.03	-0.06	-0.10	-0.13	-0.05	-0.09	-0.12
	Claim_ratio	0.15	0.65	0.66	0.50	0.10	0.71	0.58	0.43	0.13	-0.20	-0.32	-0.41	0.20	0.11	0.09	-0.10	-0.13	-0.14	-0.09	-0.12	-0.13
		Dmean	Gust	Fmax	Fmean	Fmin	S-Gust	S-Fmax	S-Fmean	S-Fmin	Pmax	Pmean	Pmin	Qmax	Qmean	Qmin	Tmax	Tmean	Tmin	S-Tmax	S-Tmean	S-Tmin
NL	Damage	0.05	0.70	0.71	0.54	-0.02	0.80	0.63	0.50	0.09	-0.30	-0.43	-0.59	0.21	0.01	-0.12	0.01	-0.00	-0.00	0.02	0.00	0.04
	Loss_ratio	0.07	0.69	0.70	0.53	-0.02	0.79	0.63	0.49	0.08	-0.29	-0.43	-0.59	0.21	0.01	-0.10	0.02	0.01	0.01	0.03	0.01	0.05
	Claims	0.07	0.70	0.71	0.54	-0.05	0.83	0.66	0.50	0.08	-0.29	-0.42	-0.59	0.21	-0.01	-0.15	0.02	-0.00	-0.02	0.03	0.00	0.02
	Claim_ratio	0.10	0.69	0.71	0.53	-0.05	0.82	0.66	0.49	0.05	-0.29	-0.43	-0.58	0.22	0.00	-0.14	0.04	0.02	0.01	0.05	0.02	0.04

Fig. 11 Spearman correlation between four damage variables (rows) and 21 hazard variables (columns) for four-digit postal codes (PC4; upper), two-digit postal codes (PC2; middle), and national level (NL; bottom). The hazard variables are coded as follows: Wind direction (D), wind speed (F), wind gust speed (G), air pressure (P), relative humidity (Q), and air temperature (T). S means station data, while all other hazard data are derived from the downscaled reanalysis datasets (Sect. 2.1)

Appendix D

The differences between the parametrisation and the recorded wind gust speed are shown in Fig. 12 for the damage–hazard dataset match at a four-digit postal code level. Whereas the recorded wind gust speed is based on the 31 meteorological stations, the parametrised wind gust speed is estimated using the daily maximum wind speed at heights 10 m and 200 m from the downscaled reanalysis data and following the approach by van den Brink (2020); see Sect. 2.1 (Fig. 13).
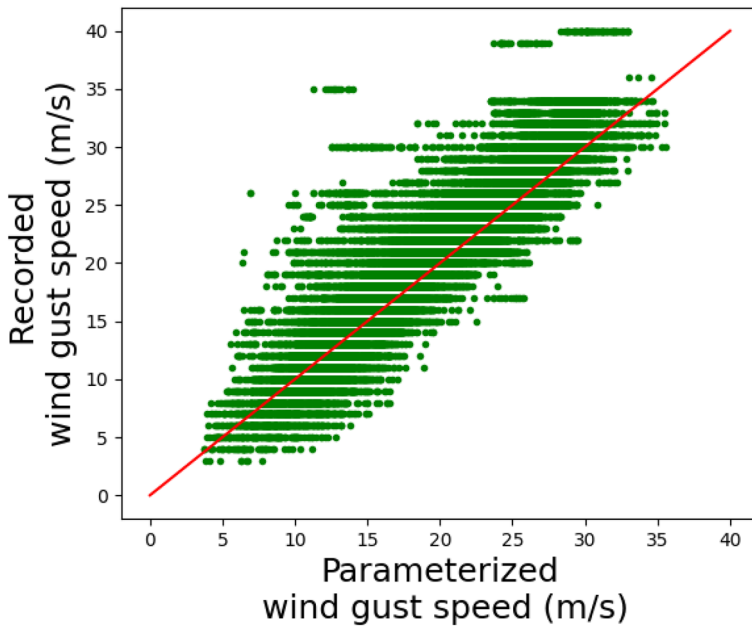
Fig. 12 Scatter plot of the parameterized wind gust (X-axis) and the recorded gust at meteorological stations (Y-axis); the 1:1 line is shown (red)
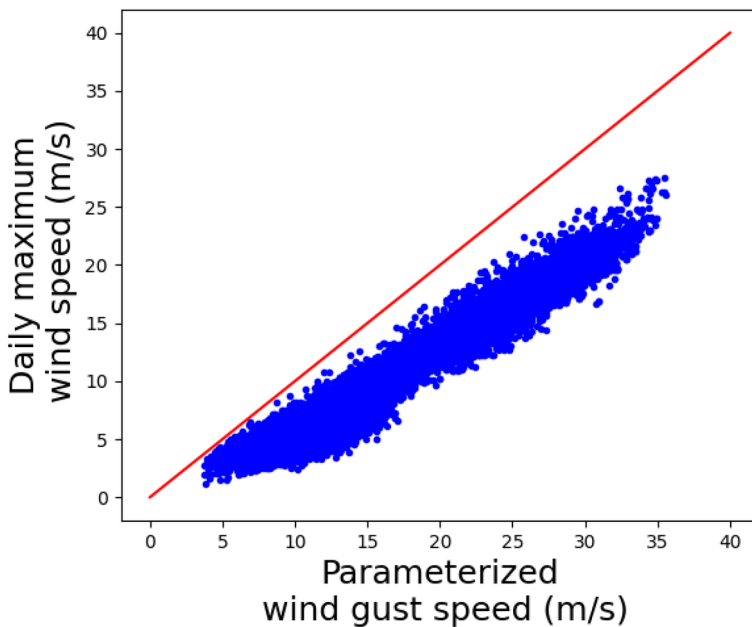

Fig. 13 Scatter plot of the parameterized wind gust (X-axis) and the daily maximum wind speed at 10 m (Y-axis); the 1:1 line is shown (red)

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval The study followed the accepted principles of ethical and professional conduct throughout the research work. No animals or human participants were involved in this research.

References

Agüera-Pérez A, Palomares-Salas JC, de la Rosa JJG, Ramiro-Leo JG, Moreno-Muñoz A (2012) Basic meteorological stations as wind data source: a mesoscalar test. J Wind Eng Ind Aerodyn 107:48–56

Baas, P. (2024). Winds of the North Sea in 2050—Public Final Report. In: Winds of the North Sea in 2050—Public Final Report. Whiffle. https://www.wins50.nl/downloads/wins50_endreport_final_20240403.pdf

Balsamo G, Albergel C, Beljaars A, Boussetta S, Brun E, Cloke H, Dee D, Dutra E, Muñoz-Sabater J, Pappenberger F, Vitart F (2015) ERA-Interim/Land: a global land surface reanalysis data set. Hydrol Earth Syst Sci 19(1):389–407

Benoit K (2011) Linear regression models with logarithmic transformations. Lond Sch Econ 22(1):23–36

Bollmeyer C, Keller JD, Ohlwein C, Wahl S, Crewell S, Friederichs P, Keune J, Kneifel S, Pscheidt I, Steinke S (2015) Towards a high-resolution regional reanalysis for the European CORDEX domain. Q J R Meteorol Soc 141(686):1–15

Bonazzi A, Cusack S, Mitas C, Jewson S (2012) The spatial structure of European wind storms as characterized by bivariate extreme-value Copulas. Nat Hazard Earth Syst Sci 12(5):1769–1782

Botzen WJW, Bouwer LM, Van den Bergh JCJM (2010) Climate change and hailstorm damage: empirical evidence and implications for agriculture and insurance. Resour Energy Econ 32(3):341–362

Botzen WJW, Bouwer LM (2016) Weather indicators for insured hailstorm damage to motor vehicles and potential climate change impacts. Geneva Pap Risk Insur Issues Pract 41:512–527

Carvalho D, Rocha A, Santos CS, Pereira R (2013) Wind resource modelling in complex terrain using different mesoscale–microscale coupling techniques. Appl Energy 108:493–504

Chen PY, Popovich PM (2002) Correlation: parametric and nonparametric measures, vol 139. Sage, London

Chok NS (2010) Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data. Doctoral dissertation, University of Pittsburgh

Daly C (2006) Guidelines for assessing the suitability of spatial climate data sets. Int J Climatol J R Meteorol Soc 26(6):707–721

Della-Marta PM, Mathis H, Frei C, Liniger MA, Kleinn J, Appenzeller C (2009) The return period of wind storms over Europe. Int J Climatol J R Meteorol Soc 29(3):437–459

Della-Marta PM, Liniger MA, Appenzeller C, Bresch DN, Köllner-Heck P, Muccione V (2010) Improved estimates of the European winter windstorm climate and the risk of reinsurance loss using climate model data. J Appl Meteorol Climatol 49(10):2092–2120

Donat MG, Leckebusch GC, Wild S, Ulbrich U (2010a) Benefits and limitations of regional multi-model ensembles for storm loss estimations. Clim Res 44(2–3):211–225

Donat MG, Leckebusch GC, Pinto JG, Ulbrich U (2010b) Examination of wind storms over Central Europe with respect to circulation weather types and NAO phases. Int J Climatol 30(9):1289–1300

Donat MG, Pardowitz T, Leckebusch GC, Ulbrich U, Burghoff O (2011a) High-resolution refinement of a storm loss model and estimation of return periods of loss-intensive storms over Germany. Nat Hazard Earth Syst Sci 11(10):2821–2833

Donat MG, Leckebusch GC, Wild S, Ulbrich U (2011b) Future changes in European winter storm losses and extreme wind speeds inferred from GCM and RCM multi-model simulations. Nat Hazard Earth Syst Sci 11(5):1351–1370

Endendijk T, Botzen WW, de Moel H, Aerts JC, Slager K, Kok M (2023) Flood vulnerability models and household flood damage mitigation measures: An econometric analysis of survey data. Water Resour Res 59(8):e2022WR034192

Fitch Ratings (2023) Rating report. In: fitchratings.com, pp 1–3 [Report]. https://www.achmea.nl/-/media/achmea/documenten/investors/ratings-fitch/19-07-2023-achmea-rating-report.pdf

Gastón M, Pascal E, Frias L, Marti I, Irigoyen U, Cantero E, Lozano S, Loureiro Y (2008) Wind resources map of Spain at mesoscale. Methodology and validation. In: Proceedings of European wind energy conference

Gonzalez-Aparicio I, Monforti F, Volker P, Zucker A, Careri F, Huld T, Badger J (2017) Simulating European wind power generation applying statistical downscaling to reanalysis data. Appl Energy 199:155–168

Held H, Gerstengarbe FW, Pardowitz T, Pinto JG, Ulbrich U, Born K, Donat MG, Karremann MK, Leckebusch GC, Ludwig P, Nissen KM, Österle H, Prahl BF, Werner PC, Befort DJ, Burghoff O (2013) Projections of global warming-induced impacts on winter storm losses in the German private household sector. Clim Change 121:195–207

IPCC (2021) Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, Zhou B (eds)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391, pp https://doi.org/10.1017/9781009157896

Jaison A, Sorteberg A, Michel C, Breivik Ø (2024) Assessment of wind–damage relations for Norway using 36 years of daily insurance data. Nat Hazard Earth Syst Sci 24(4):1341–1355

Karremann MK, Pinto JG, Reyers M, Klawa M (2014) Return periods of losses associated with European windstorm series in a changing climate. Environ Res Lett 9(12):124016

Khanduri AC, Morrow GC (2003) Vulnerability of buildings to windstorms and insurance loss estimation. J Wind Eng Ind Aerodyn 91(4):455–467

Klawa M, Ulbrich U (2003) A model for the estimation of storm losses and the identification of severe winter storms in Germany. Nat Hazard Earth Syst Sci 3(6):725–732

KNMI - Daggegevens van het weer in Nederland. (n.d.-b). https://www.knmi.nl/nederland-nu/klimatologie/daggegevens

Koks EE, Haer T (2020) A high-resolution wind damage model for Europe. Sci Rep 10(1):6866

Koks EE, Rozenberg J, Zorn C, Tariverdi M, Vousdoukas M, Fraser SA, Hall JW, Hallegatte S (2019) A global multi-hazard risk analysis of road and railway infrastructure assets. Nat Commun 10(1):2677

Kunreuther H, Michel-Kerjan E (2007) Climate change, insurability of large-scale disasters and the emerging liability challenge. NBER Working Paper No. 12821, National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w12821

Leckebusch GC, Ulbrich U, Fröhlich L, Pinto JG (2007) Property loss potentials for European mid-latitude storms in a changing climate. Geophys Res Lett 34:L05703. https://doi.org/10.1029/2006GL027663

Little AS, Priestley MD, Catto JL (2023) Future increased risk from extratropical windstorms in northern Europe. Nat Commun 14(1):4434

Montgomery DC, Peck EA, Vining GG (2021) Introduction to linear regression analysis. Wiley, Hoboken

Moran M (2022) Correlation (Pearson, Kendall, Spearman). Statistics Solutions. Retrieved August 05, 2022, from https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/correlation-pearson-kendall-spearman/

Pinto JG, Fröhlich EL, Leckebusch GC, Ulbrich U (2007) Changing European storm loss potentials under modified climate conditions according to ensemble simulations of the ECHAM5/MPI-OM1 GCM. Nat Hazard Earth Syst Sci 7(1):165–175

Pinto JG, Karremann MK, Born K, Della-Marta PM, Klawa M (2012) Loss potentials associated with European windstorms under future climate conditions. Clim Res 54(1):1–20

Prahl BF, Rybski D, Burghoff O, Kropp JP (2015) Comparison of storm damage functions and their performance. Nat Hazard Earth Syst Sci 15(4):769–788

Priestley MD, Dacre HF, Shaffrey LC, Hodges KI, Pinto JG (2018) The role of serial European windstorm clustering for extreme seasonal losses as determined from multi-centennial simulations of high-resolution global climate model data. Nat Hazard Earth Syst Sci 18(11):2991–3006

Schober P, Boer C, Schwarte LA (2018) Correlation coefficients: appropriate use and interpretation. Anesth Analg 126(5):1763–1768

Schwierz C, Köllner-Heck P, Zenklusen Mutter E, Bresch DN, Vidale PL, Wild M, Schär C (2010) Modelling European winter wind storm losses in current and future climate. Clim Change 101:485–514

Severino LG, Kropf CM, Afargan-Gerstman H, Fairless C, De Vries AJ, Domeisen DI, Bresch DN (2024) Projections and uncertainties of winter windstorm damage in Europe in a changing climate. Nat Hazard Earth Syst Sci 24(5):1555–1578

Shen X, Zhou C, Fu X (2018) Study of time and meteorological characteristics of wind speed correlation in flat terrains based on operation data. Energies 11(1):219

Sheridan P (2011) Review of techniques and research for gust forecasting and parameterisation. Met Office, Exeter

van den Brink HW (2019) An effective parametrization of gust profiles during severe wind conditions. Environ Res Commun 2(1):011001

van Stratum B, Theeuwes N, Barkmeijer J, van Ulft B, Wijnant I (2022) A one-year-long evaluation of a wind-farm parameterization in HARMONIE-AROME. J Adv Model Earth Syst 14(7):e2021MS002947

Verbond van Verzekeraars (2017) Hoofd boven Water-Verzekeren van schade in een veranderend klimaat. Retrieved from https://www.verzekeraars.nl/media/3643/hoofd-boven-water-verzekeren-van-schade-in-een-veranderend-klimaat.pdf

Verbond van Verzekeraars (VvV) (2023) Klimaatschademonitor versie 2 2-5-2023 https://bipublic.verzekeraars.nl/Home/ShowReport/5beb3391-1660-4b40-a9d5-6982d8cae31f

Welker C, Martius O, Stucki P, Bresch D, Dierer S, Brönnimann S (2016) Modelling economic losses of historic and present-day high-impact winter windstorms in Switzerland. Tellus A Dyn Meteorol Oceanogr 68(1):29546

Wijnant I, van Ulft B, van Stratum B, Barkmeijer J, Onvlee J, de Valk C, Knoop S, Kok S, Marseille G, Baltink HK, Stepek A (2019) The Dutch Offshore Wind Atlas (DOWA): description of the dataset. Royal Netherlands Meteorological Institute, Ministry of Infrastructure and Water Management, De Bilt