ADVANCING EARTH AND SPACE SCIENCES

Earth's Future

RESEARCH ARTICLE
10.1029/2023EF004188

Projecting Changes in the Drivers of Compound Flooding in Europe Using CMIP6 Models

Tim H. J. Hermans[1], Julius J. M. Busecke[2], Thomas Wahl[3], Víctor Malagón-Santos[4], Michael G. Tadesse[5], Robert A. Jane[3], and Roderik S. W. van de Wal[1,6]

[1]Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, The Netherlands, [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA, [3]Civil, Environmental, and Construction Engineering National Center for Integrated Coastal Research, University of Central Florida, Orlando, FL, USA, [4]Department of Estuarine & Delta Systems, NIOZ Royal Netherlands Institute for Sea Research, Yerseke, The Netherlands, [5]Hazen and Sawyer, Orlando, FL, USA, [6]Department of Physical Geography, Utrecht University, Utrecht, The Netherlands

Key Points:

- We project changes in the joint probability of storm surge and precipitation extremes based on a large ensemble of model simulations from the Coupled Model Intercomparison Project 6
- The joint probability will increase in the northwest and decrease in the southwest of Europe, with an average absolute magnitude of 36%–49%
- Especially under lower emissions, often more than 5 or 6 climate model simulations are needed to draw robust conclusions on these changes

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

T. H. J. Hermans,
t.h.j.hermans@uu.nl

Citation:

Hermans, T. H. J., Busecke, J. J. M., Wahl, T., Malagón-Santos, V., Tadesse, M. G., Jane, R. A., & van de Wal, R. S. W. (2024). Projecting changes in the drivers of compound flooding in Europe using CMIP6 models. *Earth's Future*, *12*, e2023EF004188. https://doi.org/10.1029/2023EF004188

Received 19 OCT 2023
Accepted 10 APR 2024

Author Contributions:

Conceptualization: Tim H. J. Hermans, Thomas Wahl, Roderik S. W. van de Wal
Data curation: Julius J. M. Busecke
Formal analysis: Tim H. J. Hermans, Víctor Malagón-Santos, Robert A. Jane
Investigation: Tim H. J. Hermans, Víctor Malagón-Santos, Robert A. Jane
Methodology: Tim H. J. Hermans, Julius J. M. Busecke, Víctor Malagón-Santos, Robert A. Jane
Resources: Julius J. M. Busecke

Abstract When different flooding drivers co-occur, they can cause compound floods. Despite the potential impact of compound flooding, few studies have projected how the joint probability of flooding drivers may change. Furthermore, existing projections may not be very robust, as they are based on only 5 to 6 climate model simulations. Here, we use a large ensemble of simulations from the Coupled Model Intercomparison Project 6 (CMIP6) to project changes in the joint probability of extreme storm surges and precipitation at European tide gauges under a medium and high emissions scenario, enabled by data-proximate cloud computing and statistical storm surge modeling. We find that the joint probability will increase in the northwest and decrease in most of the southwest of Europe. Averaged over Europe, the absolute magnitude of these changes is 36%–49% by 2080, depending on the scenario. The large-scale changes in the joint probability of extreme storm surges and precipitation are similar to those in the joint probability of extreme wind speeds and precipitation, but locally, differences can exceed the changes themselves. Due to internal climate variability and inter-model differences, projections based on simulations of only 5 to 6 randomly chosen CMIP6 models have a probability of higher than 10% to differ qualitatively from projections based on all CMIP6 simulations in multiple regions, especially under the medium emissions scenario and earlier in the twenty-first century. Therefore, our results provide a more robust and less uncertain representation of changes in the potential for compound flooding in Europe than previous projections.

Plain Language Summary Extreme storm surges, rainfall or river discharge can cause flooding. When these events happen at the same time, even more severe flooding may follow. Climate change could affect the odds that drivers of flooding coincide, potentially leading to larger flood risk. However, few scientists have tried to compute such changes, using only a few different computer models of our climate. Here, we use a much larger set of climate models to compute how the odds that an extreme storm surge coincides with extreme precipitation could change in the future. We find that at the coasts of northwestern Europe, those odds will increase, whereas in southwestern Europe, they will mostly decrease. On average, the changes will be as large as 36%–49% of the current odds, depending on whether the concentration of greenhouse gases in the atmosphere will increase by a medium or a large amount. When we use smaller sets of climate models for our calculations, we get substantially different results in some cases. In conclusion, by using a larger set of climate models than previous studies, we have made more robust computations of how the odds that extreme storm surges and precipitation coincide will change in Europe.

1. Introduction

The co-occurrence or close succession of different flooding drivers like storm surges, rainfall and river discharge has the potential to affect coastal communities more severely than the separate occurrence of these drivers (e.g., Bevacqua et al., 2017; Emanuel, 2017; Kumbier et al., 2018; Paprotny et al., 2018; Ruocco et al., 2011; van den Hurk et al., 2015). For instance, extreme precipitation or river discharge may increase the depth and/or area of flooding due to storm surges and high coastal water levels may hamper storm-water drainage and cause backwater effects. Such combinations of hazard drivers are called compound events (Zscheischler et al., 2018). Since the more traditional univariate analyses that neglect the compounding effects of flooding drivers may underestimate flood risk and the lifetime of adaptation measures to flooding (e.g., Leonard et al., 2014; Moftakhari et al., 2017;

Software: Tim H. J. Hermans, Julius J. M. Busecke, Michael G. Tadesse
Supervision: Thomas Wahl, Roderik S. W. van de Wal
Validation: Tim H. J. Hermans, Víctor Malagón-Santos
Visualization: Tim H. J. Hermans
Writing – original draft: Tim H. J. Hermans, Julius J. M. Busecke, Thomas Wahl, Víctor Malagón-Santos, Robert A. Jane, Roderik S. W. van de Wal
Writing – review & editing: Tim H. J. Hermans, Julius J. M. Busecke, Roderik S. W. van de Wal

Wahl et al., 2015), compound events have received increased attention in the past decade. For instance, the historical dependence between and joint probability of various combinations of flooding drivers has been assessed at local (e.g., Couasnon et al., 2022; Kew et al., 2013; Santos et al., 2021; Zheng et al., 2014), national (e.g., Hendry et al., 2019; W. Wu et al., 2018), continental (e.g., Camus et al., 2021; Ganguli & Merz, 2019; Nasr et al., 2021; Paprotny et al., 2018, 2020; Wahl et al., 2015) and global scales (e.g., Bevacqua, Vousdoukas, Zappa, et al., 2020; Couasnon et al., 2019; Eilander et al., 2020; Lambert et al., 2020; Ridder et al., 2020; Ward et al., 2018), using observations and/or model hindcasts.

In comparison, fewer studies have projected how the potential for compound flooding may change in the future. For instance, a global study projected the joint probability of extreme storm surges and precipitation to decrease in parts of the subtropics and to increase at higher latitudes (Bevacqua, Vousdoukas, Zappa, et al., 2020). For the United States, the joint probabilities of various flooding drivers were projected to increase due to sea-level rise, changes in extreme river discharge and changes in tropical cyclones (Ghanbari et al., 2019; Gori et al., 2022; Moftakhari et al., 2017). For most of Europe, the joint probability of extreme storm surges and precipitation was projected to increase by Bevacqua et al. (2019), predominantly due to the increasing probability of extreme precipitation. However, Ganguli et al. (2020) projected a decrease in the dependence and joint probability of extreme storm surges and river discharge in northwestern Europe. The differences between the projections of these studies are inconsistent with the finding that the joint probability of extreme storm surges and precipitation is generally comparable to that of extreme storm surges and river discharge at small to medium river catchments (Bevacqua, Vousdoukas, Shepherd, & Vrac, 2020).

A common limitation of existing projections of the joint probability of flooding drivers is the small ensembles of global and/or regional climate model simulations on which they are based. For instance, Bevacqua, Vousdoukas, Zappa, et al. (2020) and Ganguli et al. (2020) based their projections on only 5 to 6 models from the Coupled Model Intercomparison Project 5 (CMIP5; Taylor et al., 2012), using only a single, high-emissions scenario simulation per model. Consequently, these projections may be sensitive to the specific models that were used and provide a limited view of the uncertainties related to future emissions, internal climate variability and structural differences between models, especially since the skill of climate models in capturing the atmospheric conditions that may cause compound flooding varies (Ridder et al., 2021; Y. Wu et al., 2021). Some studies used larger multi-model ensembles to project changes in the joint probability of extremes (e.g., Bevacqua et al., 2023; Ridder et al., 2022; Sun et al., 2023), but none included storm surges as a driver.

Furthermore, most projections of the joint probability of extremes in general are based on climate model ensembles that include only one initial-condition simulation per model. However, since co-occurring extremes are rare, estimates of their joint probability are sensitive to internal climate variability when derived from a single simulation, even when using a 50-year period from that simulation (Santos et al., 2021). Hence, as advocated by Bevacqua et al. (2023), projections of the potential for compound extremes would benefit from using single model initial-condition large ensembles (SMILEs). These are ensembles of simulations generated with the same external forcing but initialized at different times, so that internal climate variability has a different phase in each simulation and can be partially averaged out. Consequently, SMILEs can be used to develop more robust projections of the joint probability of extremes (Bevacqua et al., 2023) and to partition the total uncertainty of projections into uncertainties due to emissions scenarios, inter-model differences and internal climate variability (Lehner et al., 2020).

Many global climate models from the current, sixth Coupled Model Intercomparison Project (CMIP6) (Eyring et al., 2016) provide simulations for multiple initial-condition members. Including all these simulations for the analysis of compound flooding is challenging as storm surges and river discharge are not a direct output of global climate models but need to be derived from their simulations offline. This is typically done using computationally demanding hydrodynamic and hydrological models, respectively (e.g., Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020). However, as a computationally more efficient alternative to hydrodynamic modeling, data-driven models have recently been developed to compute storm surges at large spatial scales (Bellinghausen et al., 2023; Bruneau et al., 2020; Tadesse & Wahl, 2021; Tadesse et al., 2020; Tiggeloven et al., 2021). Such statistical models, based on multi-linear regression (MLR) or other machine learning techniques, have been shown to perform similarly to or better than high-resolution hydrodynamic models such as the Global Tide and Surge Model (GTSM) of Muis et al. (2016, 2020, 2023) (Tadesse et al., 2020; Tiggeloven et al., 2021). Therefore, they may also be useful for projecting changes in the joint probability of extreme storm surges and other flooding drivers.

Here, we project changes in the joint probability of extreme storm surges and precipitation and analyze their uncertainty using the simulations of a large ensemble of CMIP6 models, including all initial-condition members available for each model. To derive storm surge information from each simulation, we use the data-driven statistical model of Tadesse et al. (2020), which we will show is well suited for the analysis of the joint probability of storm surge and precipitation extremes. We limit our study to Europe, where data-driven storm surge models generally perform well (Bruneau et al., 2020; Tadesse et al., 2020; Tiggeloven et al., 2021). Storm surges are mainly caused by wind and sea-level pressure. Therefore, the probability of joint extreme wind speed and precipitation events, which can disrupt transport and power systems (e.g., Jaroszweski et al., 2015), is closely related to that of joint storm surge and precipitation extremes and helps to interpret the changes in the latter physically. Therefore, we consider changes in the probability of joint wind speed and precipitation extremes alongside changes in the probability of joint storm surge and precipitation extremes and compare them. Finally, we exploit the large ensemble of CMIP6 simulations to compare the ensemble mean changes to the effect of internal climate variability, partition the uncertainty of our projections and compute the ensemble size required for qualitatively robust projections in different locations.

2. CMIP6 Data and Joint Extremes Analysis

In this section, we explain which CMIP6 simulations we use and how we analyze the changes in the joint probability of extremes in these simulations.

2.1. CMIP6 Data Used

We analyze future changes in the joint probability of extremes for an intermediate and a high emissions scenario (shared socio-economic pathway scenarios SSP2-4.5 & SSP8.5, respectively; Meinshausen et al., 2020). As only few CMIP6 models provide simulations at a sub-daily frequency, we use daily mean CMIP6 simulations. Models are required to provide daily mean sea-level pressure (variable "*psl*"), surface wind speed (variable "*sfcWind*") and precipitation flux (variable "*pr*") output for the historical period (1850–2014) and at least one of the SSP2-4.5 and SSP5-8.5 scenarios (2015–2100). To obtain time series for 1850–2100, each SSP simulation is appended to its corresponding historical simulation. Daily mean wind speed and precipitation flux time series (converted to daily accumulated precipitation) are used to analyze (changes in) the joint probability of wind speed and precipitation extremes (as explained in Sections 2.2 and 2.3), whereas daily mean wind speed and sea-level pressure time series are used as input to the statistical storm surge model (as explained in Section 3). Like Ridder et al. (2022), we use daily mean instead of daily maximum wind speed, as more CMIP6 simulations are available for the former.

For several CMIP6 models, multiple realizations (denoted with "*r*" in the "*ripf*" variant label) are available that have been branched off from their preindustrial control run at different times. Because the phase of internal climate variability differs between these realizations, they can be used to average out part of the changes due to internal climate variability and better isolate the changes due to increasing greenhouse gas concentrations. In contrast to previous projections, we therefore include all available realizations of each CMIP6 model providing the output described above. The resulting data set includes over 20 terabytes of data from 27 different CMIP6 models (see Table 1 for an overview). To process this data efficiently and reproducibly, we use the Analysis-Ready Cloud Optimized CMIP6 data produced by the Pangeo/Earth System Grid Federation (ESGF) Cloud Data Working Group (https://pangeo-data.github.io/pangeo-cmip6-cloud/), held in public Google Cloud Storage. The data sets summarized in Table 1 reflect data sets that were available to download and ingest via the pangeo-forge feedstock (Busecke & Stern, 2023) at the time of writing of this manuscript. The data is analyzed using the code in the CMIP6cex repository (Hermans & Busecke, 2024a), for which the xarray (Hoyer & Hamman, 2017) and xMIP (Busecke et al., 2023) python packages are important building blocks.

Prior to the analysis, we bilinearly interpolated the simulations of each model to a common grid with a $1.5° \times 1.5°$ resolution, using xESFM (Zhuang et al., 2023). A $1.5° \times 1.5°$ grid roughly corresponds with the average resolution of the CMIP6 models (Table 1). The effects of orography and coastlines and mesoscale processes such as fronts and convection may be better resolved by models with a higher resolution, but these typically provide fewer

Table 1
Coupled Model Intercomparison Project 6 Simulations Used

	Model	SSP2-4.5 [#]	SSP5-8.5 [#]	Both [#]	°Lon × °Lat	Reference
1	ACCESS-CM2	5	6	4	1.875 × 1.25	Bi et al. (2020)
2	ACCESS-ESM1-5	38	35	33	1.875 × 1.25	Bi et al. (2020)
3	CanESM5	25	25	25	2.8 × 2.8	Swart et al. (2019)
4	CESM2	2	2	2	1.25 × 0.9	Danabasoglu et al. (2020)
5	CESM2-WACCM	3	3	3	1.25 × 0.9	Danabasoglu et al. (2020)
6	CMCC-ESM2	1	1	1	1.25 × 0.9	Lovato et al. (2022)
7	CMCC-CM2-SR5	1	1	1	1.25 × 0.9	Cherchi et al. (2019)
8	EC-Earth3	59	1	1	0.75 × 0.75	Döscher et al. (2022)
9	EC-Earth3-Veg	1	0	0	0.75 × 0.75	Döscher et al. (2022)
10	FGOALS-g3	1	0	0	2 × 2	L. Li et al. (2020)
11	GFDL-CM4	1	1	1	1 × 1	Held et al. (2019)
12	GFDL-ESM4	1	1	1	1 × 1	Dunne et al. (2020)
13	HadGEM3-GC31-LL	5	4	4	1.875 × 1.25	Andrews et al. (2020)
14	HadGEM3-GC31-MM	0	4	0	0.83 × 0.56	Andrews et al. (2020)
15	INM-CM4-8	1	1	1	2 × 1.5	Volodin and Gritsun (2018)
16	INM-CM5-8	1	1	1	2 × 1.5	Volodin et al. (2017)
17	IPSL-CM6A-LR	11	7	6	2.5 × 1.3	Boucher et al. (2020)
18	KACE-1-0-G	3	3	3	Not reported	Lee et al. (2020)
19	MIROC6	43	50	43	1.4 × 1.4	Tatebe et al. (2019)
20	MIROC6-ES2L	10	1	1	2.8 × 2.8	Hajima et al. (2020)
21	MPI-ESM1-2-LR	24	24	24	1.88 × 1.88	Mauritsen et al. (2019)
22	MPI-ESM1-2-HR	2	2	2	0.93 × 0.93	Mauritsen et al. (2019)
23	MRI-ESM2-0	1	1	1	0.75 × 0.75	Yukimoto et al. (2019)
24	NorESM2-LL	3	1	1	2.5 × 1.88	Seland et al. (2020)
25	NorESM2-MM	2	1	1	1.25 × 0.94	Seland et al. (2020)
26	TaiESM1	1	1	1	1.25 × 0.9	Wang et al. (2021)
27	UKESM1-0-LL	5	5	5	1.875 × 1.25	Sellar et al. (2020)

simulations. Ensemble statistics are computed and displayed on the common 1.5° × 1.5° grid. The regridded simulations are also used as input to the statistical storm surge model (as described in Section 3).

2.2. Definition of Joint Extremes

In this study, we consider two types of compound extremes: (a) the combination of extreme daily mean wind speed and extreme daily accumulated precipitation, and (b) the combination of extreme daily maximum storm surge and extreme daily accumulated precipitation. While compound events can already be impactful if only one of their drivers is extreme (Wahl et al., 2015), we focus on the case in which both drivers are extreme, similar to previous studies (Bevacqua et al., 2019; Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020; Ridder et al., 2022). We define extreme events using a peak-over-threshold (POT) analysis instead of using annual maxima, because this allows us to consider multiple extremes occurring in a single year and avoids including annual maxima that are not extreme.

Previous POT analyses have often used the same threshold percentile or used thresholds resulting in the same number of declustered extremes for each location and variable (e.g., Bevacqua, Vousdoukas, Zappa, et al., 2020; Camus et al., 2021; Ganguli et al., 2020; Hendry et al., 2019; Ridder et al., 2020); a pragmatic approach which we also adopt here. For Europe, Camus et al. (2021) found that using 3 vs 6 declustered extremes per year resulted in similar bivariate dependence patterns for several combinations of compound flooding drivers. Therefore, we use the 98th percentile of daily values as a threshold, which results in a number of extremes slightly higher than 6 per year. Hence, wind speed (w), storm surge (s) and precipitation (p) extremes are defined as $P = p \geq p_{98}$, $W = w \geq w_{98}$ and $S = s \geq s_{98}$, respectively, and joint extreme wind speed and precipitation and joint extreme storm surges and precipitation events as days on which those extremes co-occur ($W \wedge P$ and $S \wedge P$, respectively). As a baseline, we only consider extremes that occur on the same day and do not decluster the extremes prior to the analysis. The sensitivity of our projections to these methods is discussed in Section 5.

2.3. Future Changes in the Joint Probability of Extremes

We analyze the joint probability of extremes empirically by counting the number of joint extremes ($N_{W\wedge P}$ and $N_{S\wedge P}$) and standardizing those numbers by the length of the time period considered, as done by Camus et al. (2021), Couasnon et al. (2019), Hendry et al. (2019), and Ridder et al. (2020, 2022).

2.3.1. Computing Future Changes

To compute the changes in the number of joint extremes that the CMIP6 models simulate ($\Delta N_{W\wedge P}$ and $\Delta N_{S\wedge P}$), we define two 40-year periods centered around 2000 (1981–2020) and 2080 (2061–2100) as the historical and future periods, respectively. We then compute $\Delta N_{W\wedge P}$ (and similarly, $\Delta N_{S\wedge P}$) as the difference in the number of joint extremes between these periods:

$$\Delta N_{W\wedge P} = N_{W\wedge P}^{fut} - N_{W\wedge P}^{hist}, \tag{1}$$

in which the superscripts ^{fut} and ^{hist} mean "evaluated in" the future and historical period, respectively. Importantly, for both the historical and future periods, the number of joint extremes that exceed the historical thresholds w_{98}^{hist} and p_{98}^{hist} are counted. Therefore,

$$N_{W\wedge P}^{fut} = |w^{fut} \geq w_{98}^{hist} \wedge p^{fut} \geq p_{98}^{hist}| \tag{2}$$

and

$$N_{W\wedge P}^{hist} = |w^{hist} \geq w_{98}^{hist} \wedge p^{hist} \geq p_{98}^{hist}|. \tag{3}$$

The same equations are applied to compute $\Delta N_{S\wedge P}$ by replacing wind speed (w) with storm surges (s).

2.3.2. Decomposing Future Changes

Changes in the joint probability of extremes can be decomposed into changes in the marginal distributions of each of the considered variables and changes in the dependence structure between them (see Figure S2 in Supporting Information S1 for a graphical explanation). Using methods similar to those of Bevacqua et al. (2019) and Bevacqua, Vousdoukas, Zappa, et al. (2020), we compute the changes in $N_{W\wedge P}$ (and similarly, $N_{S\wedge P}$) due to changes in the marginal distribution of wind and precipitation (denoted $\Delta N_{W\wedge P}^{w}$ and $\Delta N_{W\wedge P}^{p}$, respectively) as

$$\Delta N_{W\wedge P}^{w} = |w^{hist} \geq w_{U_w}^{hist} \wedge p^{hist} \geq p_{98}^{hist}| - N_{W\wedge P}^{hist} \tag{4}$$

and

$$\Delta N_{W\wedge P}^{p} = |w^{hist} \geq w_{98}^{hist} \wedge p^{hist} \geq p_{U_p}^{hist}| - N_{W\wedge P}^{hist}. \tag{5}$$

Put more simply, we compute how changing the threshold percentile for either wind speed or precipitation extremes affects the number of joint extremes in the historical period. In Equations 4 and 5, the changed threshold percentiles are defined as $U_w = F_w^{fut}\left(w_{98}^{hist}\right)$ and $U_p = F_p^{fut}\left(p_{98}^{hist}\right)$, where F_w^{fut} and F_p^{fut} are the empirical cumulative distribution functions of the wind speed and the precipitation in the future period. Hence, U_w and U_p are the threshold percentiles that the historical threshold values would correspond to in the future. Similarly, we compute the changes in $N_{W\wedge P}$ due to the changes in the two marginal distributions combined, but still with the historical dependence structure, as

$$\Delta N^{w,p}_{W\wedge P} = |w^{hist} \geq w^{hist}_{U_w} \wedge p^{hist} \geq p^{hist}_{U_p}| - N^{hist}_{W\wedge P}. \tag{6}$$

To compute the changes in $N_{W\wedge P}$ due to changes in the dependence between wind speed and precipitation, we simply subtract the changes due to changes in both marginal distributions from the total change:

$$\Delta N^{dependence}_{W\wedge P} = \Delta N_{W\wedge P} - \Delta N^{w,p}_{W\wedge P} \tag{7}$$

Given that U_w and U_p are the threshold percentiles that the historical threshold values correspond to in the future period, the historical threshold values used to compute $\Delta N_{W\wedge P}$ (see Equations 1–3) can be written as $w^{hist}_{98} = w_{U_w}{}^{fut}$ and $p^{hist}_{98} = p_{U_p}{}^{fut}$. Therefore, by using Equation 7 to compute $\Delta N^{dependence}_{W\wedge P}$, we essentially compute the difference in the number of joint extremes between the historical and future periods using the same threshold percentiles to define extremes in both periods, instead of using the same threshold values (as is done to compute $\Delta N_{W\wedge P}$). Since using the same threshold percentiles in both periods means that the number of univariate extremes will not change, the remaining changes in the number of joint extremes must follow from changes in the dependence between the considered variables. Again, $\Delta N_{S\wedge P}$ is decomposed similarly, using s instead of w. Table 2 summarizes the notations defined in this section.

2.4. Computing Ensemble Statistics

To evaluate the ensemble results, we apply model democracy and weight each CMIP6 model equally. This disregards potential differences in model performance, but as we will discuss in Section 6, evaluating model performance in the absence of large initial-condition ensembles for each model is challenging. Indeed, CMIP6 models provide different numbers of initial-condition simulations (see Table 1). To include all initial-condition simulations while weighting models equally, we first average the (changes in the) number of joint extremes over all initial-condition simulations of each model before computing the multi-model ensemble mean and standard deviation. Furthermore, the availability of simulations differs between SSP2-4.5 and SSP5-8.5. To be able to directly compare the projections between the scenarios, we only use the initial-condition simulations that are available for both emissions scenarios in the main text, and provide the projections based on all available simulations per SSP scenario in Supporting Information S1.

In Section 4.4, we compare the magnitude of the ensemble mean changes in the number of joint extremes to the magnitude of the variability in the historical number of joint extremes. As a metric of the effect of internal climate variability on the historical number of joint extremes, we compute the average standard deviation of $N^{hist}_{W\wedge P}$ and $N^{hist}_{S\wedge P}$ between initial-condition simulations using the CMIP6 models that have at least five initial-condition members (see Table 1). To test whether these models are representative for the entire CMIP6 ensemble, more models providing multiple initial-condition simulations are needed. We also use these models to partition the uncertainty in our projections into the uncertainty due to internal variability (I), differences between models (M) and differences between emissions scenarios (S), similar to Lehner et al. (2020). To estimate I, we compute the standard deviation of $\Delta N_{W\wedge P}$ and $\Delta N_{S\wedge P}$ between the initial-condition members of each model, and compute the average of those standard deviations. Next, to estimate M, we compute the standard deviation of $\Delta N_{W\wedge P}$ and $\Delta N_{S\wedge P}$ between the means of the initial-condition members of each model. Then, to estimate S, we compute the standard deviation of the member-mean $\Delta N_{W\wedge P}$ and $\Delta N_{S\wedge P}$ between SSP2-4.5 and SSP5-8.5 for all available CMIP6 models, and compute the average of those standard deviations. The uncertainty due to different emissions scenarios would have likely been larger if we would have also included a lower emissions scenario like SSP1-2.6.

To study how internal climate variability and inter-model differences affect projections of the joint probability of extremes based on small climate model ensembles (e.g., Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020), we compute the probability that the means of such ensembles agree qualitatively with our projections. To this end, we randomly draw up to 5,000 ensembles of size s from all possible combinations of s CMIP6 models (using a single default initial-condition member per model), for $s = 1$ to $s = N_{CMIP6} - 1$. For each s, we then compute the fraction of ensembles for which the ensemble mean $\Delta N_{W\wedge P}$ (or $\Delta N_{S\wedge P}$) has the same sign

Table 2
Notations Defined in Section 2

Notation	Definition
w, s, & p	Wind speed, storm surge and precipitation
W, S, & P	Wind speed, storm surge and precipitation extremes defined as events equal to or higher than the 98th percentile of all wind speed, storm surge and precipitation events ($w \geq w_{98}$, $s \geq s_{98}$ and $p \geq p_{98}$), respectively
$N_{W \wedge P}$	Standardized number of days on which both wind speed and precipitation are extreme
$N^{hist}_{W \wedge P}$	$N_{W \wedge P}$ in the period 1981–2020 as simulated by CMIP6 models
$N^{fut}_{W \wedge P}$	$N_{W \wedge P}$ in the period 2061–2100 as simulated by CMIP6 models
$\Delta N_{W \wedge P}$	The difference between $N^{fut}_{W \wedge P}$ and $N^{hist}_{W \wedge P}$
$\Delta N^{w}_{W \wedge P}$	$\Delta N_{W \wedge P}$ due to changes in the marginal distribution of wind speed
$\Delta N^{p}_{W \wedge P}$	$\Delta N_{W \wedge P}$ due to changes in the marginal distribution of precipitation
$\Delta N^{w,p}_{W \wedge P}$	$\Delta N_{W \wedge P}$ due to changes in the marginal distributions of wind speed and precipitation
$\Delta N^{dependence}_{W \wedge P}$	$\Delta N_{W \wedge P}$ due to changes in the dependence between wind speed and precipitation
$N_{S \wedge P}$	Standardized number of days on which both storm surge and precipitation are extreme
$N^{hist}_{S \wedge P}$	$N_{S \wedge P}$ in the period 1981–2020 as simulated by CMIP6 models
$N^{fut}_{S \wedge P}$	$N_{S \wedge P}$ in the period 2061–2100 as simulated by CMIP6 models
$\Delta N_{S \wedge P}$	The difference between $N^{fut}_{S \wedge P}$ and $N^{hist}_{S \wedge P}$
$\Delta N^{s}_{S \wedge P}$	$\Delta N_{S \wedge P}$ due to changes in the marginal distribution of storm surges
$\Delta N^{p}_{S \wedge P}$	$\Delta N_{S \wedge P}$ due to changes in the marginal distribution of precipitation
$\Delta N^{s,p}_{S \wedge P}$	$\Delta N_{S \wedge P}$ due to changes in the marginal distributions of storm surges and precipitation
$\Delta N^{dependence}_{S \wedge P}$	$\Delta N_{S \wedge P}$ due to changes in the dependence between storm surges and precipitation

as that of the ensemble including all CMIP6 models and initial-condition members. Finally, as an indication for how large ensembles need to be for qualitatively robust projections, we compute the minimum s for which the fraction of ensemble means agreeing in sign is 90% or higher.

3. Modeling Storm Surges

3.1. Training and Application of the Storm Surge Model

To compute storm surges for each CMIP6 simulation in Table 1 we use a MLR model based on the methods of Tadesse and Wahl (2021) and Tadesse et al. (2020), as running a hydrodynamic model for each simulation is computationally infeasible. The MLR model of Tadesse et al. (2020) was trained with daily maximum non-tidal residuals observed at tide gauges (TGs) as predictands and sub-daily surface winds and sea-level pressure from various reanalyzes as predictors. Predictors were used within either 10° by 10° (Tadesse et al., 2020) or 6° by 6° (Tadesse & Wahl, 2021) grids around each TG and lags up to 30 hr between the predictands and predictors were implemented. The daily maximum non-tidal residuals of Tadesse et al. (2020) were obtained by removing the annual mean sea level and predicted astronomical tides from TG records in the GESLA2 data set (Woodworth et al., 2016).

A simpler version of the statistical model was previously applied to compute storm surges for a large ensemble of simulations of the European weather@home atmospheric model (Calafat et al., 2022). This version was trained using daily mean wind speed and sea-level pressure (without lags) from the ERA5 reanalysis (Hersbach et al., 2020) as predictors, in 2° by 2° grids around each TG. Since we will apply the MLR model to daily mean CMIP6 data (see Section 2.1), here, we also use daily mean predictors from ERA5 (1979–2018) to train the MLR model. However, because a larger grid size around each TG leads to a better performance (Tadesse & Wahl, 2021), we use 9° by 9° grids around each TG. To ensure traceability to prior validation and application of the MLR model, we adopted the GESLA2 predictands used by Calafat et al. (2022), Tadesse and Wahl (2021), and Tadesse et al. (2020) for training. The few years of additional data that GESLA3 records (Haigh et al., 2021) would provide are unlikely to substantially affect our results (see Tadesse et al., 2020, and Section 3.2). The resolution of ERA5 (0.25° by 0.25°) is higher than that of the CMIP6 models (see Table 1). Therefore, prior to estimating the regression coefficients with the ERA5 predictors, we coarsen the ERA5 data by bilinearly interpolating it to the same 1.5° by 1.5° grid that the CMIP6 simulations were regridded to (Section 2.1). At this resolution, the 9° by 9° grids around each TG consist of 36 grid cells. Training the storm surge model with coarsened instead of native-resolution ERA5 predictors did not substantially affect its performance.

Other than using different predictor data, we estimate the regression coefficients in the same way as Tadesse and Wahl (2021) and Tadesse et al. (2020). As the flowchart in Figure S1 in Supporting Information S1 shows, the gridded ERA5 data around each TG, including the wind speed squared and cubed, are normalized by removing the time-mean of each variable and scaling them to unit variance. To reduce the dimensionality of the gridded data, the normalized variables are then pooled, after which empirical orthogonal functions (EOFs) are computed. The first EOFs that together explain at least 95% of the variance of the predictor data are then regressed on the daily maximum non-tidal residuals from GESLA2.

To apply the statistical storm surge model to the CMIP6 simulations, the regression coefficients that were estimated with the ERA5 predictor data are multiplied with predictors derived from the regridded CMIP6 simulations (see flowchart in Figure S1 in Supporting Information S1). The CMIP6 predictors are prepared like the ERA5 predictors. For each combined historical and SSP simulation (1850–2100), we take the daily mean wind speed and sea-level pressure gridded around each TG, and also compute the squared and cubed wind speed terms. Subsequently, we normalize these variables, pool them and compute EOFs. As the sign of an EOF is not unique, we flip the sign of an EOF of the CMIP6 predictor data if its spatial pattern better matches that of the corresponding EOF of the ERA5 predictor data when multiplied by -1. For each TG, the first n EOFs of the CMIP6 predictor data around that TG are multiplied with the ERA5-based regression coefficients, where n is the number of EOFs that explained at least 95% of the variance of the ERA5 predictor data around that TG (see Figure S1 in Supporting Information S1). For each simulation, this results in estimates of daily maximum non-tidal residuals at every TG during 1850–2100, which we refer to as storm surges.

By applying a statistical model that is trained on observations to CMIP6 simulations of the future, we implicitly assume that the learned relationships between predictors and predictands will not change. Consequently, while using a statistical model allows us to base our projections on a large ensemble of climate model simulations, we neglect potential changes in how wind and pressure relate to storm surges, for instance due to changes in mean sea level. Future work could test how such effects influence model calibration, for instance by comparing statistical models trained on historical and future hydrodynamic model simulations (e.g., Muis et al., 2020, 2023; Vousdoukas et al., 2018).

3.2. Evaluating the Storm Surge Model

The purpose of using the storm surge model is to analyze the number of joint storm surge and precipitation extremes $N_{S\wedge P}$. Therefore, we evaluate the model by comparing $N_{S\wedge P}$ based on the statistically modeled storm surges ($N_{S_{MLR}\wedge P}$) with that based on the observed daily maximum non-tidal residuals from GESLA2 ($N_{S_{G2}\wedge P}$) (Figure 1). To put this comparison into context, we also evaluate $N_{S\wedge P}$ based on the daily maximum non-tidal residuals from the Coastal Data set for the Evaluation of Climate Impact (CoDEC) ($N_{S_{CoDEC}\wedge P}$). This data set was simulated with the high-resolution GTSM driven by atmospheric forcing from ERA5 (Muis et al., 2020). Furthermore, we also compare $N_{W\wedge P}$ based on daily mean wind speed from ERA5 ($N_{W_{ERA5}\wedge P}$) with $N_{S_{G2}\wedge P}$. In all cases, precipitation comes from ERA5, and we only use the timesteps at which GESLA2 data is available (see Figure S1 in Supporting Information S1 for the temporal coverage).

$N_{S_{G2}\wedge P}$ is relatively large (15–25 joint extremes per decade) at the west and south coasts of Spain, Portugal and France, and at the southwest coast of the UK, while it is relatively small (0–10 joint extremes per decade) at the north coast of Spain and along the North Sea (Figure 1a). This pattern is consistent with the results of previous studies (Bevacqua, Vousdoukas, Zappa, et al., 2020; Couasnon et al., 2019; Hendry et al., 2019; Paprotny et al., 2018). With a correlation of 0.87 and a normalized root mean square error (nRMSE) of 0.36, $N_{S_{MLR}\wedge P}$ agrees relatively well with $N_{S_{G2}\wedge P}$ (compare Figures 1a and 1b), which suggests that the statistical storm surge model predicts the timing of the extremes in the GESLA2 data well. Using 5-fold cross-validation, we verified that the

10.1029/2023EF004188

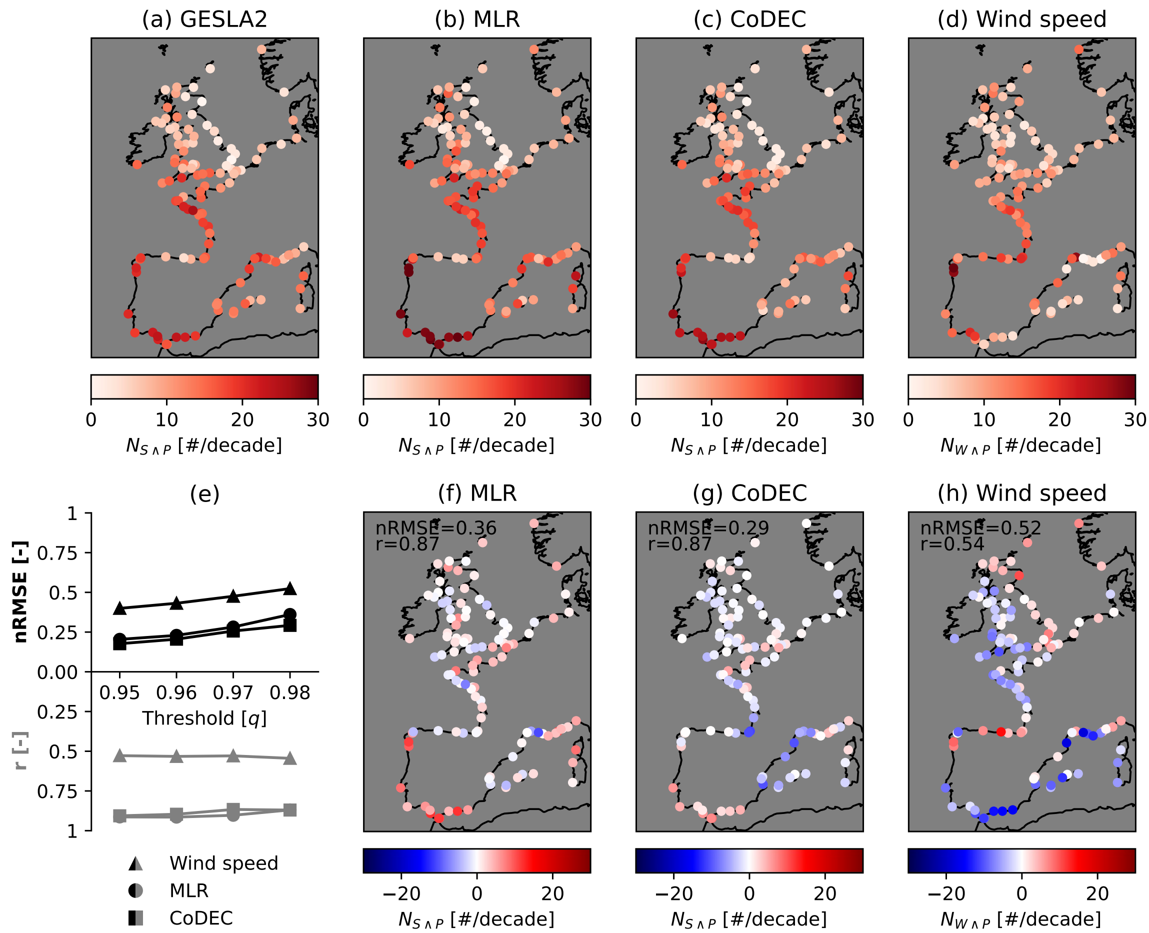

Figure 1. Number of joint extreme (a) storm surges from GESLA2 and precipitation from ERA5 ($N_{S_{G2} \wedge P}$), (b) statistically modeled storm surges and precipitation from ERA5 ($N_{S_{MLR} \wedge P}$), (c) storm surges from Coastal Data set for the Evaluation of Climate Impact (CoDEC) and precipitation from ERA5 ($N_{S_{CoDEC} \wedge P}$) and (d) wind speed and precipitation from ERA5 ($N_{W_{ERA5} \wedge P}$) at GESLA2 tide gauges [#/decade]. (e) The correlation coefficient r and normalized root mean square error (nRMSE) of (b–d) relative to (a) as a function of the extremes threshold percentile. (g–f) $N_{S_{MLR} \wedge P}$, $N_{S_{CoDEC} \wedge P}$ and $N_{W_{ERA5} \wedge P}$ minus $N_{S_{G2} \wedge P}$, respectively. All nRMSEs are normalized by dividing by the mean of $N_{S_{G2} \wedge P}$. All correlation coefficients are statistically significant ($p \ll 0.05$).

agreement between $N_{S_{MLR} \wedge P}$ and $N_{S_{G2} \wedge P}$ does not change much when only considering days that were not used for training. The differences between $N_{S_{MLR} \wedge P}$ and $N_{S_{G2} \wedge P}$ are largest at the south and east coasts of Spain (Figure 1f), where $N_{S_{MLR} \wedge P}$ overestimates $N_{S_{G2} \wedge P}$.

Overall, the number of joint storm surge and precipitation extremes based on the statistically modeled storm surges is very similar to that based on the storm surges simulated with GTSM (compare Figures 1b and 1c). Although the biases of $N_{S_{MLR} \wedge P}$ (nRMSE = 0.36, Figure 1f) are moderately larger than those of $N_{S_{CoDEC} \wedge P}$ (nRMSE = 0.29, Figure 1g), the pattern correlation coefficients are the same ($r = 0.87$). Whereas $N_{S_{MLR} \wedge P}$ overestimates $N_{S_{G2} \wedge P}$ mostly around Spain, $N_{S_{CoDEC} \wedge P}$ underestimates $N_{S_{G2} \wedge P}$ south of France and in the Bay of Biscay by several events per decade. The differences between $N_{S_{G2} \wedge P}$ (Figure 1a) and $N_{W_{ERA5} \wedge P}$ (Figure 1d), which are shown in Figure 1h, are clearly larger. These differences reveal where the joint probability of storm surge and precipitation extremes differs from that of wind speed and precipitation extremes, and therefore where the information on sea-level pressure and the direction of the wind that the statistical storm surge model contains adds value. A comparison between the magnitudes of the biases in Figures 1f and 1h suggests that this is the case for instance at the north coast of Spain, along the Mediterranean Sea, at the west coast of France and around most of the UK. The agreement between $N_{W_{ERA5} \wedge P}$ and $N_{S_{G2} \wedge P}$ improves when we use daily maximum instead of daily mean wind speed, but including daily maximum wind speed as a predictor variable of the MLR model does not lead to a much better performance of $N_{S_{MLR} \wedge P}$.

10.1029/2023EF004188

The temporal coverage of the TG records is limited at several locations (see Figure S1 in Supporting Information S1). Consequently, at some of these locations, the evaluation in Figure 1 is based on only a couple of observed joint storm surge and precipitation extremes. To test whether the results in Figure 1 are robust to using a larger sample size, we repeated the evaluation with lower threshold percentiles (Figure 1e). Other than that the error metrics tend to improve for lower thresholds, which may partially reflect the inclusion of less extreme events, the performance of $N_{S_{MLR} \wedge P}$ relative to that of $N_{S_{CoDEC} \wedge P}$ is not very sensitive to a larger sample size (Figure 1e). Therefore, we conclude that using the statistical storm surge model instead of a hydrodynamic model to analyze $\Delta N_{S \wedge P}$ in CMIP6 simulations is appropriate, especially since a hydrodynamic model would also have to be forced with the relatively low-resolution atmospheric forcing from CMIP6 instead of with the ERA5 forcing used for CoDEC.

As an additional test we evaluated the statistical storm surge model trained with the CoDEC data instead of the daily maxima from GESLA2. We find that with these predictands, the biases of $N_{S_{MLR} \wedge P}$ and $N_{S_{CoDEC} \wedge P}$ are also relatively similar (Figure S3 in Supporting Information S1). Furthermore, applying this version of the model to the CMIP6 simulations did not substantially change the results in Section 4.3. Therefore, future research may use a statistical storm surge model trained with hydrodynamic model simulations to extend our analysis to locations without TGs.

4. Changes in the Number of Joint Extremes

4.1. Wind Speed and Precipitation ($\Delta N_{W \wedge P}$)

Displaying $N_{W_{ERA5} \wedge P}$ for the entire domain, Figure 2a indicates that the observed number of joint wind speed and precipitation extremes is relatively large mainly over west-facing coasts and mountainous regions such as the western Iberian Peninsula, western France, parts of the UK and Norway. In contrast, it is relatively low over Sweden, eastern Spain, southeastern France and the southeastern UK (Figure 2a). To a large degree, $N_{W_{ERA5} \wedge P}$ is consistent with the historical extremal dependency between wind speed and precipitation that has been estimated previously (Martius et al., 2016; Owen et al., 2021). The CMIP6 ensemble mean $N^{hist}_{W \wedge P}$ well approximates this large-scale pattern (Figure 2b), but especially the lower-resolution CMIP6 models do not capture the small-scale imprints of orography and land-sea contrast seen in ERA5 (Figure S4 in Supporting Information S1). Consequently, the ensemble mean $N^{hist}_{W \wedge P}$ is smoother and lower than $N_{W_{ERA5} \wedge P}$ in among others Scotland, Iceland, Norway and Italy.

For both SSPs, the ensemble mean $\Delta N_{W \wedge P}$ shows increases (of up to 4 and 6 per decade under SSP2-4.5 and SSP5-8.5, respectively) in a band extending from the southwest to the northeast, neighbored by decreases (of up to 7 and 11 per decade under SSP2-4.5 and SSP5-8.5, respectively) in the northwest (Iceland) and the south (Bay of Biscay & the Mediterranean Sea) of the domain (Figures 2c and 2d). Averaged over land, the absolute magnitude of the changes is approximately 39% (SSP2-4.5) to 51% (SSP5-8.5) of the historical number of joint extremes. The spatial patterns of $\Delta N_{W \wedge P}$ are similar under SSP2-4.5 and SSP5-8.5, but the magnitude of the changes is larger under SSP5-8.5, reflecting a larger forced response. Correspondingly, the area in which the magnitude of the ensemble mean $\Delta N_{W \wedge P}$ exceeds the standard deviation of the change between the CMIP6 models (shown in Figure S5 in Supporting Information S1) is larger under SSP5-8.5 than under SSP2-4.5 (see stippling in Figures 2c and 2d). If we only include one randomly selected initial-condition member per model, the standard deviation between models increases by approximately 15% (SSP2-4.5) and 11% (SSP5-8.5) on average, reflecting the ensemble uncertainty due to internal variability (see also Section 5). The $\Delta N_{W \wedge P}$ of individual CMIP6 models with only few initial-condition members is clearly more noisy (less spatially coherent) than that of CMIP6 models with more members (Figures S6 and S7 in Supporting Information S1). Compared to the projections of Ridder et al. (2022), our ensemble mean projections seem to indicate larger decreases in southern Spain, larger increases in the east of the UK and smaller increases in the west of the UK. These differences may be related to the different CMIP6 ensembles used, but a more systematic comparison would be needed to confirm this.

As shown in Figure 3, a substantial part of the ensemble mean $\Delta N_{W \wedge P}$ under SSP5-8.5 (and under SSP2-4.5, see Figure S8 in Supporting Information S1) consists of changes in the marginal distribution of precipitation. The ensemble mean $\Delta N^{p}_{W \wedge P}$ (Figure 3a) is positive over most of Europe and negative over the south of the domain; a pattern that is consistent with projections of the magnitude of extreme precipitation (C. Li et al., 2021; Pfahl et al., 2017; Seneviratne, 2021). The increases in extreme precipitation over most of Europe, which would lead to

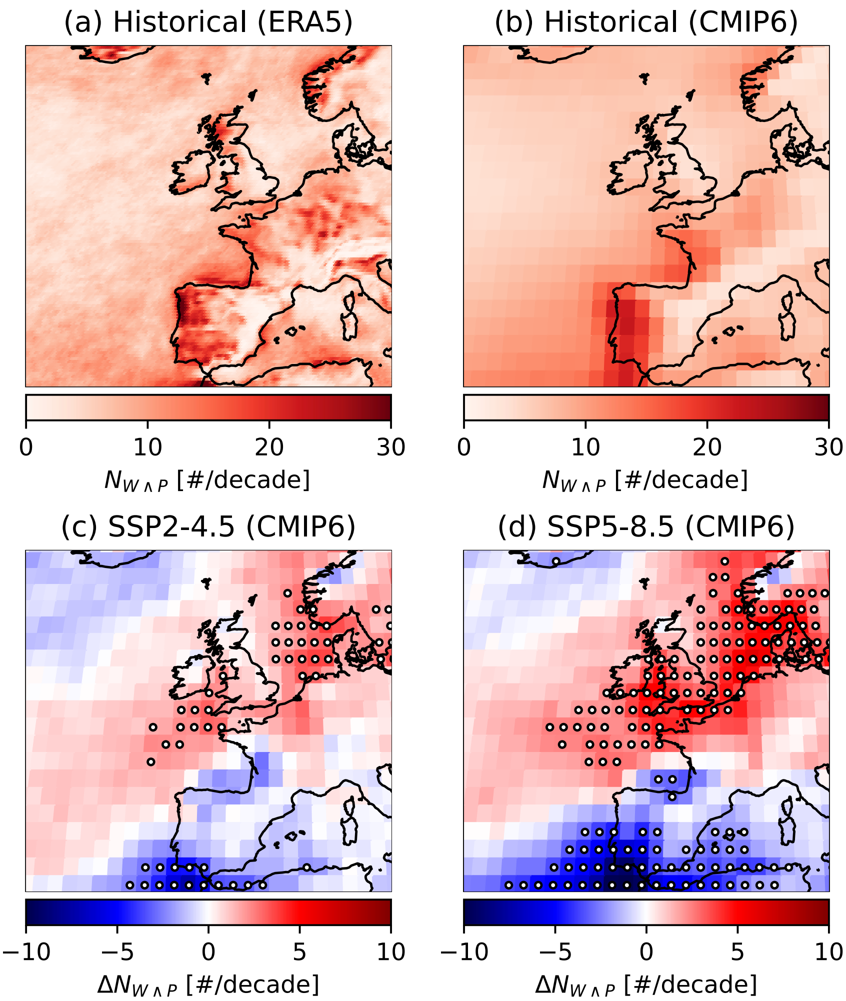

Figure 2. (a) $N_{W \wedge P}$ based on the ERA5 reanalysis (1979–2018), (b) Coupled Model Intercomparison Project 6 (CMIP6) ensemble mean $N^{hist}_{W \wedge P}$ (1981–2020), and (c, d) CMIP6 ensemble mean $\Delta N_{W \wedge P}$ under SSP2-4.5 and SSP5-8.5, respectively (2061–2100 minus 1981–2020). In (c) and (d), the stippling indicates where the absolute ensemble mean change exceeds the standard deviation of the change between models.

a larger number of precipitation events exceeding the historical threshold, are understood to be caused by the increasing moisture carrying capacity of the heating atmosphere. The decreases in the Mediterranean, which would lead to fewer events exceeding the historical threshold in the future, are thought to be caused by dynamic circulation changes such as the projected shift of the North Atlantic storm track (Pfahl et al., 2017). Except in between these regions of increases and decreases, the ensemble mean $\Delta N^{p}_{W \wedge P}$ exceeds the standard deviation between the models.

Changes in the marginal distribution of wind speed, on the other hand, contribute negatively to $\Delta N_{W \wedge P}$ in most of the domain (Figure 3b), except over the North Sea region and Sweden. The pattern of the ensemble mean resembles previously projected changes in storm track density and the wind intensity of extratropical cyclones (Priestley & Catto, 2022; Zappa et al., 2013). In several regions with negative $\Delta N^{w}_{W \wedge P}$, such as in Spain, Portugal, Iceland and Norway, the ensemble mean exceeds the standard deviation across models under SSP5-8.5 (and to a lesser extent under SSP2-4.5, see Figure S8 in Supporting Information S1). Where $\Delta N^{w}_{W \wedge P}$ is positive, this is not the case, similar to what has been found for changes in windstorm damages (Severino et al., 2023).

Together, changes in the marginal distributions of precipitation and wind speed (Figure 3c) explain much but not all of the total change ($\Delta N_{W \wedge P}$, Figure 2d). The remaining differences are caused by changes in the dependence between extreme winds and precipitation ($\Delta N^{dependence}_{W \wedge P}$), which are negative mainly over the northwest of the

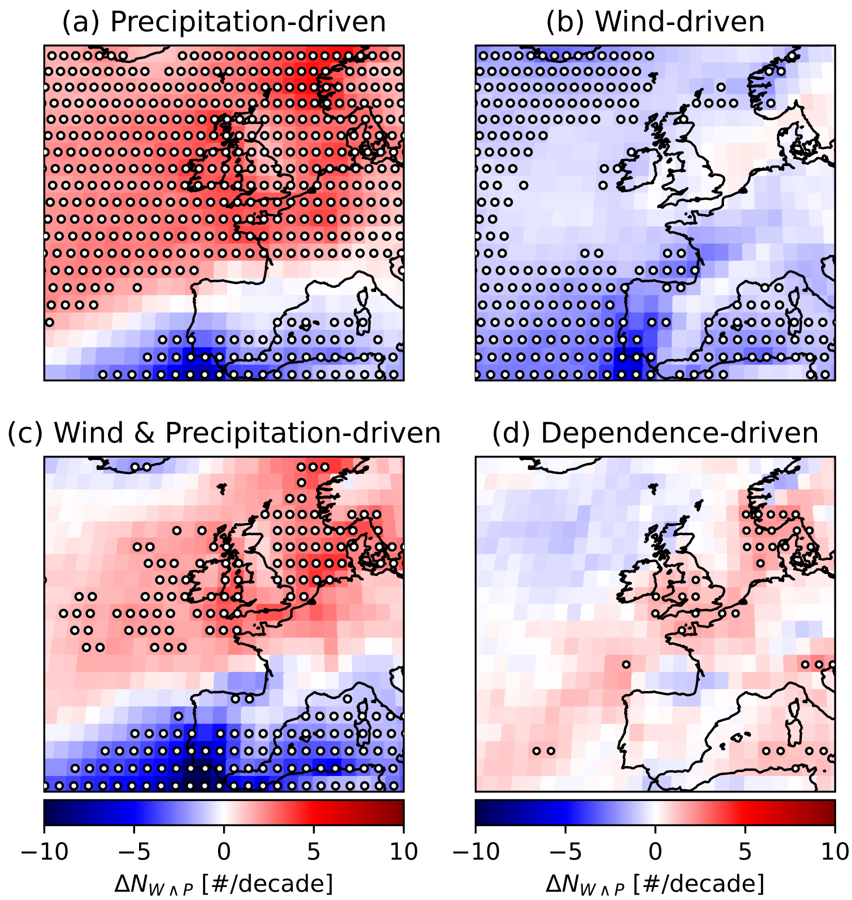

Figure 3. Coupled Model Intercomparison Project 6 ensemble mean $\Delta N_{W\wedge P}$ (under SSP5-8.5, 2061–2100 minus 1981–2020) due to (a) changes in the marginal distribution of precipitation ($\Delta N^{p}_{W\wedge P}$), (b) changes in the marginal distribution of wind speed ($\Delta N^{w}_{W\wedge P}$), (c) changes in the marginal distributions of both precipitation and wind speed ($\Delta N^{w,p}_{W\wedge P}$), and (d) changes in the dependence between precipitation and wind speed $\left(\Delta N^{dependence}_{W\wedge P}\right)$. The stippling indicates where the absolute value of the ensemble mean of each component of $\Delta N_{W\wedge P}$ exceeds the standard deviation of that component between models.

domain, the Bay of Biscay, part of Norway and northern Africa, and positive in most other regions (Figure 3d). While the magnitude of $\Delta N^{dependence}_{W\wedge P}$ is moderate, it is comparable to or higher than that of the other components of $\Delta N_{W\wedge P}$ in several regions. The ensemble mean of $\Delta N^{dependence}_{W\wedge P}$ is lower than its standard deviation between models except south of Norway, in the southern UK and at a few other grid cells (Figure 3d), indicating that the uncertainty in this term due to model differences and internal climate variability is relatively large. In contrast to $\Delta N^{p}_{W\wedge P}$ and $\Delta N^{w}_{W\wedge P}$, the pattern of $\Delta N^{dependence}_{W\wedge P}$ does not clearly relate to the atmospheric changes projected in previous studies. For instance, the dependence-driven increases in $N_{W\wedge P}$ over the Mediterranean appear to be inconsistent with the projected decrease in the frequency of extratropical cyclones over southern Europe (Priestley & Catto, 2022; Zappa et al., 2013), given that in Europe, joint wind speed and precipitation extremes are often associated with extratropical cyclones (Owen et al., 2021). However, as alluded to by Owen et al. (2021), the co-occurrence of wind speed and precipitation extremes also depends on the seasons in which the extremes tend to occur. To better understand the dependence-driven changes, we therefore analyze $\Delta N^{dependence}_{W\wedge P}$ from a seasonal perspective in the next section.

4.2. Seasonal Dependence-Driven Changes

Most of the ensemble mean $\Delta N^{dependence}_{W\wedge P}$ (Figure 3d) consists of changes in autumn (SON) and winter (DJF) (Figures 4a and 4d), which are the seasons in which extratropical cyclones in Europe prevail and extreme wind speed and precipitation are most likely to co-occur (Owen et al., 2021). As explained in Section 2.3.2, to compute

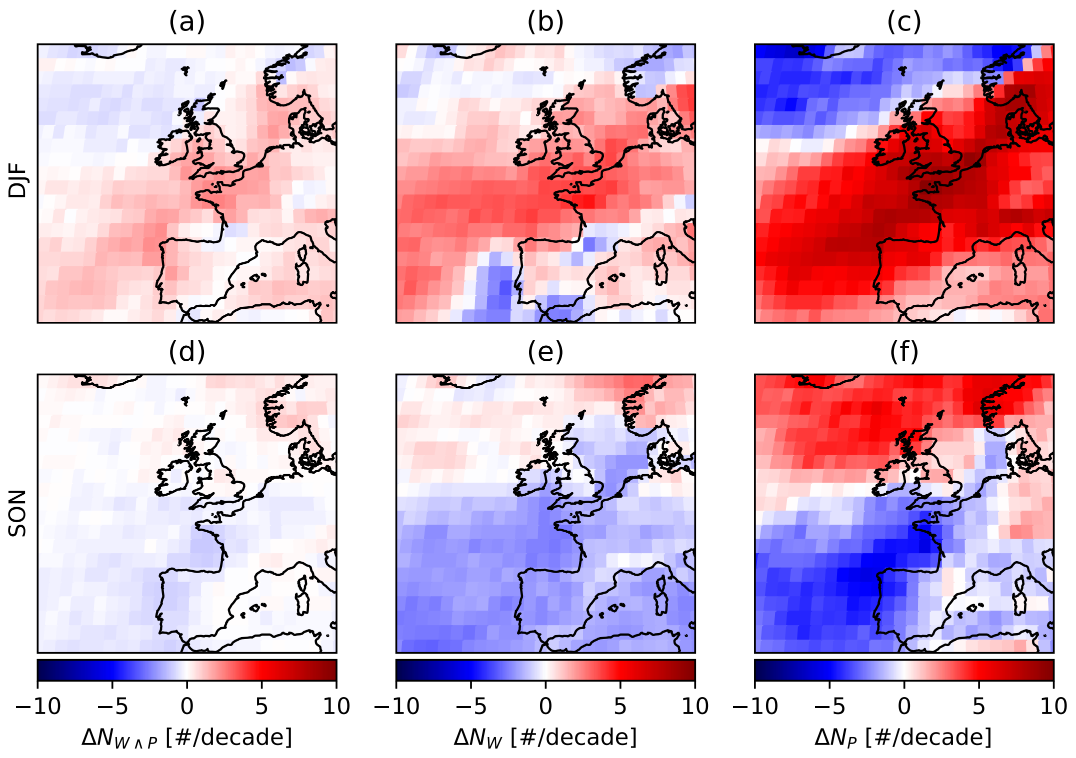

Figure 4. Coupled Model Intercomparison Project 6 ensemble mean changes (under SSP5-8.5, 2061–2100 minus 1981–2020) in the number of joint wind and precipitation extremes ($\Delta N_{W \wedge P}$), univariate wind speed extremes (ΔN_W) and univariate precipitation extremes (ΔN_P), in (a)–(c) winter (DJF) and (d)–(f) autumn (SON), respectively.

$\Delta N^{dependence}_{W \wedge P}$ we use the same threshold percentiles in both the historical and future period, so that the total number of univariate extremes in these periods stays the same. However, the seasons in which the univariate extremes tend to occur can change between these periods (Figures 4b, 4c, 4e, and 4f). The projected shifts in the number of univariate extremes in autumn and winter seem to explain at least part of the ensemble mean $\Delta N^{dependence}_{W \wedge P}$ in these seasons (Figure 4). For example, winter $\Delta N^{dependence}_{W \wedge P}$ is negative in the north and northwest of the domain, where also the numbers of precipitation and wind extremes in winter decrease. Similarly, winter $\Delta N^{dependence}_{W \wedge P}$ is positive over the Mediterranean Sea, where both the numbers of precipitation and wind extremes are simulated to increase in winter (Figures 4a–4c). Furthermore, autumn $\Delta N^{dependence}_{W \wedge P}$ is negative over the Bay of Biscay, consistent with the decreasing numbers of both wind and precipitation extremes in that region (Figures 4d–4f).

The changes in the number of univariate extremes in winter and autumn (Figures 4b, 4c, 4e, and 4f) reflect the changes in the magnitude of extremes in these seasons relative to the other seasons. For instance, if the magnitude of heavy-precipitation events will increase more strongly (or decrease less strongly) in winter than in summer, a larger fraction of the unchanged total number of extreme precipitation events will occur in winter and a smaller fraction in summer. Therefore, even if the frequency of the weather phenomena causing joint extremes, which prevail in autumn and winter, is not projected to increase, a larger fraction of events with strong winds and precipitation in autumn and winter may be classified as compound extreme events if the magnitude of univariate extremes in spring and summer decreases relative to that in autumn and winter. In places where $\Delta N^{dependence}_{W \wedge P}$ is less consistent with the seasonal changes in the number of univariate extremes, such as over the Bay of Biscay in winter (Figures 4a–4c), changes in the frequency of certain weather types may play a larger role.

4.3. Storm Surges and Precipitation ($N_{S \wedge P}$)

Next, we analyze the number of joint storm surge and precipitation extremes, using the storm surges that were statistically derived from the CMIP6 simulations. The CMIP6 ensemble mean $N^{hist}_{S \wedge P}$ (Figure 5a) agrees with $N_{S_{G2} \wedge P}$ (Figure 1a) reasonably well (significant pattern correlation of 0.75 and nRMSE of 0.40). Similarly to

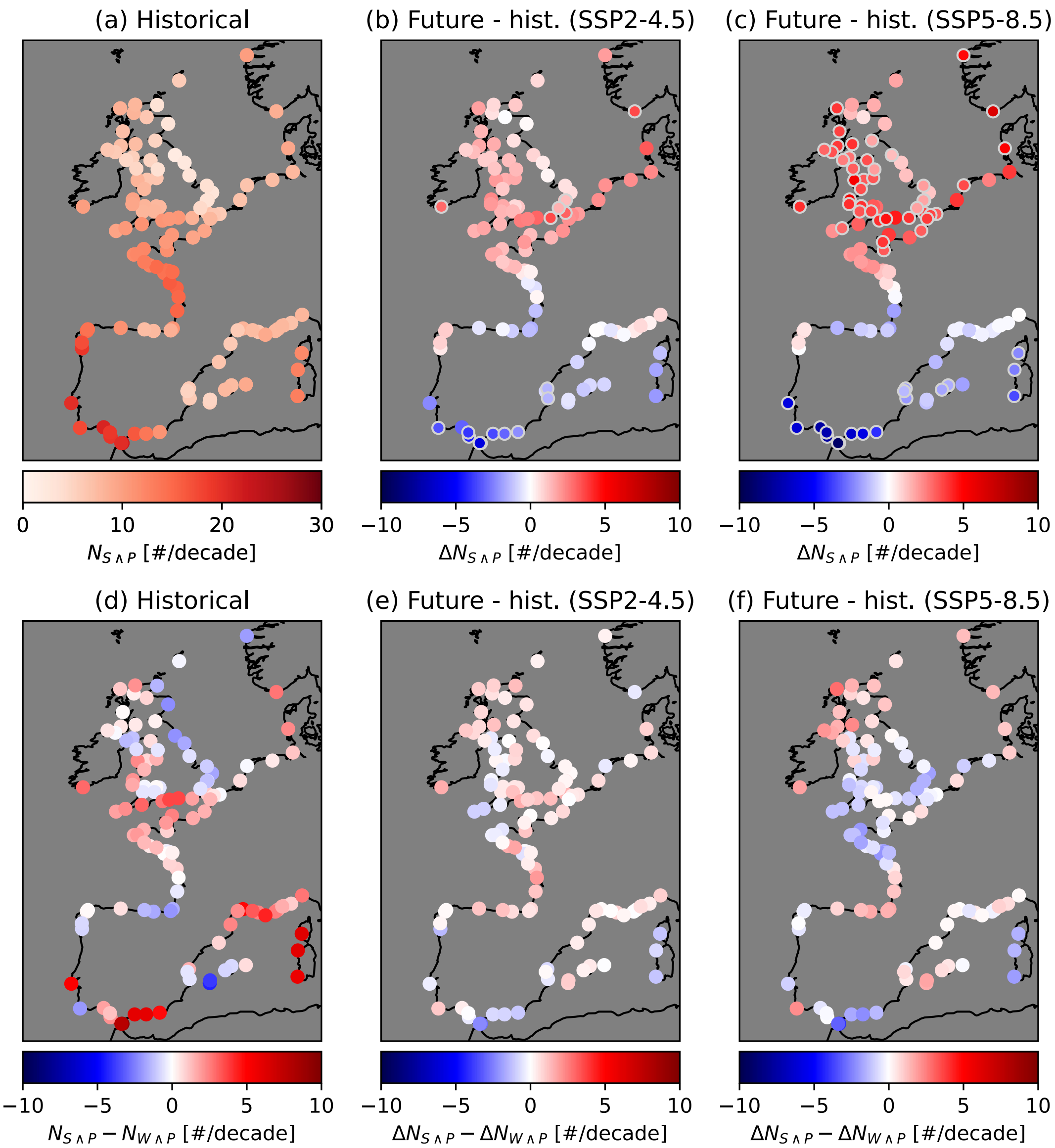

Figure 5. (a) Coupled Model Intercomparison Project 6 (CMIP6) ensemble mean $N^{hist}_{S\wedge P}$ (1981–2020), (b, c) CMIP6 ensemble mean $\Delta N_{S\wedge P}$ under SSP2-4.5 and SSP5-8.5, respectively (2061–2100 minus 1981–2020), (d) the CMIP6 ensemble mean of $N^{hist}_{S\wedge P}$ minus $N^{hist}_{W\wedge P}$, and (e, f) the CMIP6 ensemble mean of $\Delta N_{S\wedge P}$ minus $\Delta N_{W\wedge P}$ under SSP2-4.5 and SSP5-8.5, respectively. In (b) and (c), circles with a gray edge indicate where the absolute ensemble mean change exceeds the standard deviation of the change between models.

$N_{S_{G2}\wedge P}$, the ensemble mean $N^{hist}_{S\wedge P}$ is relatively large at west coasts and relatively small at the east coast of the UK and in northern Spain. However, especially along the northwestern coastline of the Mediterranean Sea and at the east coasts of the UK and France, the ensemble mean $N^{hist}_{S\wedge P}$ tends to underestimate $N_{S_{G2}\wedge P}$ (Figure S9a in Supporting Information S1). While these biases may partially be inherited from the MLR model (see Figure 1f), the ensemble mean $N^{hist}_{S\wedge P}$ also tends to underestimate $N_{S_{MLR}\wedge P}$ in these regions (Figure S9b in Supporting Information S1). Furthermore, a similar underestimation was found for an ensemble of CMIP5 models based on hydrodynamically modeled storm surges (Bevacqua et al., 2019; Bevacqua, Vousdoukas, Zappa, et al., 2020). Hence, part of the biases is related to the differences between the atmospheric forcing of global climate models and ERA5, to which internal variability also contributes. Compared to the ensemble mean $N^{hist}_{W\wedge P}$ (Figure 2b), $N^{hist}_{S\wedge P}$ is larger in the northern Mediterranean, south of Spain and in and around the English Channel, and smaller in the Bay of Biscay and the east of the UK and Spain (Figure 5d). This pattern is very similar to that of the differences

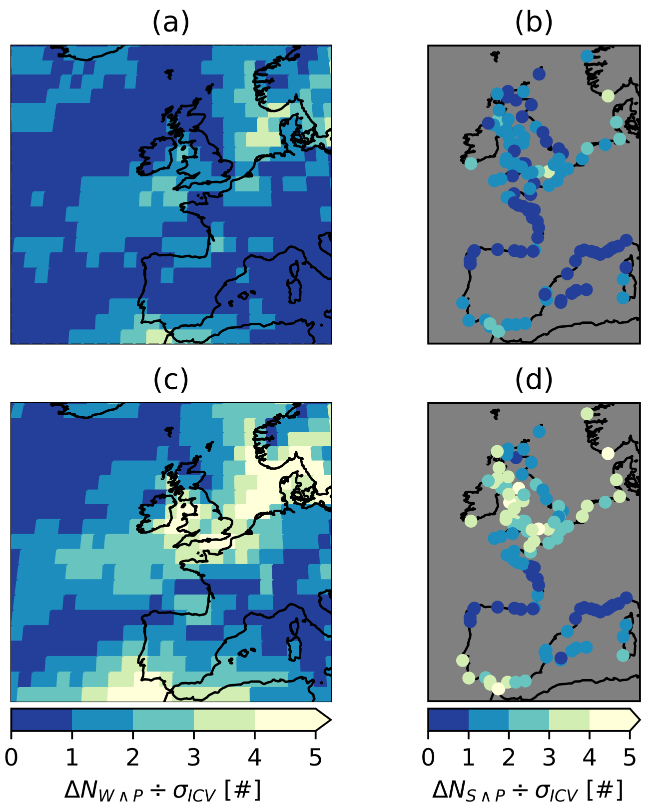

Figure 6. (a) Coupled Model Intercomparison Project 6 (CMIP6) ensemble mean $\Delta N_{W \wedge P}$ under SSP2-4.5 (2061–2100 minus 1981–2020) divided by the average standard deviation of $N^{hist}_{W \wedge P}$ (1981–2020) between initial-condition members of CMIP6 models providing at least five members, (b) CMIP6 ensemble mean $\Delta N_{S \wedge P}$ under SSP2-4.5 (2061–2100 minus 1981–2020) divided by the average standard deviation of $N^{hist}_{S \wedge P}$ (1981–2020) between initial-condition members of CMIP6 models providing at least five members, (c, d) as in (a) and (b), under SSP5-8.5.

between $N_{S_{MLR} \wedge P}$ and $N_{W_{ERA5} \wedge P}$ (Figures 1b and 1d), suggesting that the MLR model indeed translates the atmospheric forcing from ERA5 and CMIP6 models to storm surges similarly.

For both SSPs, we find that the ensemble mean $\Delta N_{S \wedge P}$ (Figures 5b and 5c) is positive at TGs in northwestern Europe and negative at most TGs in southwestern Europe. The largest increases can be seen in the English Channel, at the east coast of the UK and along the southeastern North Sea coast (up to 4 and 6 per decade, under SSP2-4.5 and SSP5-8.5, respectively), and the largest decreases south of Spain (up to 6 and 10 per decade, under SSP2-4.5 and SSP5-8.5, respectively). In these regions, the ensemble mean tends to exceed the standard deviation between models (see Figure S5 in Supporting Information S1), especially under SSP5-8.5 (gray-edged circles in Figures 5b and 5c). If we only include one randomly selected initial-condition member per model, the standard deviation between models increases by approximately 18% (SSP2-4.5) and 15% (SSP5-8.5) on average. West of the UK, north of Spain and at the northern coast of Mediterranean Sea, the ensemble mean $\Delta N_{S \wedge P}$ is relatively small. Averaged over the TGs, the absolute magnitude of the ensemble mean changes is approximately 36% (SSP2-4.5) to 49% (SSP5-8.5) of the historical number of joint extremes.

While their large-scale patterns broadly agree, $\Delta N_{S \wedge P}$ and $\Delta N_{W \wedge P}$ differ by several events per decade in various regions. For instance, in the Bay of Biscay, $\Delta N_{S \wedge P}$ is less negative than $\Delta N_{W \wedge P}$, and at the coast of Scotland, $\Delta N_{S \wedge P}$ is more positive than $\Delta N_{W \wedge P}$ (Figures 5e and 5f). Furthermore, in several locations, the magnitude of the difference between these changes exceeds that of the changes themselves. Although the differences in Figures 5e and 5f inevitably depend on the storm surge model used, their magnitude suggests that changes in the joint probability of wind speed and precipitation extremes are not a great proxy for changes in the joint probability of storm surge and precipitation extremes and that the translation of wind and pressure to storm surges is important to assess the potential for compound flooding. Interestingly, the sign of the differences between $\Delta N_{S \wedge P}$ and $\Delta N_{W \wedge P}$ is not consistent between SSP2-4.5 and SSP5-8.5 everywhere, which could be due to internal variability and/or differences in non-linear changes between the SSPs. Given that both $\Delta N_{S \wedge P}$ and $\Delta N_{W \wedge P}$ are driven by the same large-scale atmospheric circulation changes described in Sections 4.1 and 4.2, their decomposition into changes in marginal distributions and dependence-related changes are also broadly similar (Figure S13 in Supporting Information S1).

4.4. Magnitude Relative to Historical Internal Climate Variability

Figures 6a and 6b show that in most regions, the magnitude of the ensemble mean $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ under SSP2-4.5 is smaller than one or two times the standard deviation of $N^{hist}_{W \wedge P}$ and $N^{hist}_{S \wedge P}$ due to internal climate variability (estimated as explained in Section 2.4). In other words, most of the future changes in the average number of joint extremes projected under SSP2-4.5 are smaller than the temporary deviations from that average that are likely to be seen due to internal climate variability alone. In contrast, over the eastern North Sea and south of Spain, the ensemble mean changes in the number of joint extremes tend to be larger than typical unforced variability (Figures 6a and 6b).

Clearly, the ensemble mean $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ under SSP5-8.5 exceed twice the standard deviation due to internal climate variability in more locations than under SSP2-4.5 (Figures 6c and 6d). For instance, south of Spain, east of the UK and along the southeastern North Sea coastline, the ensemble mean changes are higher than 3 or 4 standard deviations due to internal climate variability. Around the Mediterranean Sea and northeast of the UK, however, the ensemble mean $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ under SSP5-8.5 are still smaller than 2 standard deviations. For $\Delta N_{S \wedge P}$, this is also the case in the Bay of Biscay and west of the English Channel.

5. Uncertainty in the Projections and Sensitivity to Ensemble Size

Variations in the number of joint extremes due to internal climate variability (see Figure 6), as well as inter-model differences and differences between emissions scenarios, introduce uncertainties in the projections of $\Delta N_{W\wedge P}$ and $\Delta N_{S\wedge P}$ (see Figure S14 in Supporting Information S1). Depending on the location, these uncertainties can be large compared to the ensemble mean future changes (compare Figures 2 and 5 with Figure S14 in Supporting Information S1). Therefore, projections of changes in the joint probability of extremes based on only 5 to 6 climate model simulations (e.g., Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020) may change substantially when different models and/or initial-condition members would be used. We investigate this sensitivity by sub-sampling our large CMIP6 ensemble as described in Section 2.4.

We find that at many locations in the northeast of the domain, in the south and west of the UK, in the southeastern North Sea and around the south of Spain, projections based on random subsets of CMIP6 models are more than 90% likely to have the same sign as the projections based on the full CMIP6 ensemble even if the subsets consist of only 5 climate model simulations (Figure 7). This suggests that in these regions, the sign of the projections of previous studies that used 5 to 6 models (e.g., Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020) is relatively insensitive to the specific climate models that were included. In contrast, in the north and east of the UK, over parts of the mainland of Europe, in the Bay of Biscay and along the southern coasts of France and Italy, the probability that projections based on small ensembles differ in sign is often higher than 10% (Figure 7). Therefore, in these locations, it is not unlikely that the projections of previous studies could have had the opposite sign if different climate model simulations would have been used.

Figure 7 also shows that in general, the earlier in the twenty-first century and the lower the emissions scenario, the larger is the ensemble size required for qualitatively robust projections. For instance, for robust projections of $\Delta N_{S\wedge P}$ at the east coast of the UK, more than 10 models are required under SSP2-4.5 (Figure 7d), whereas 5 or

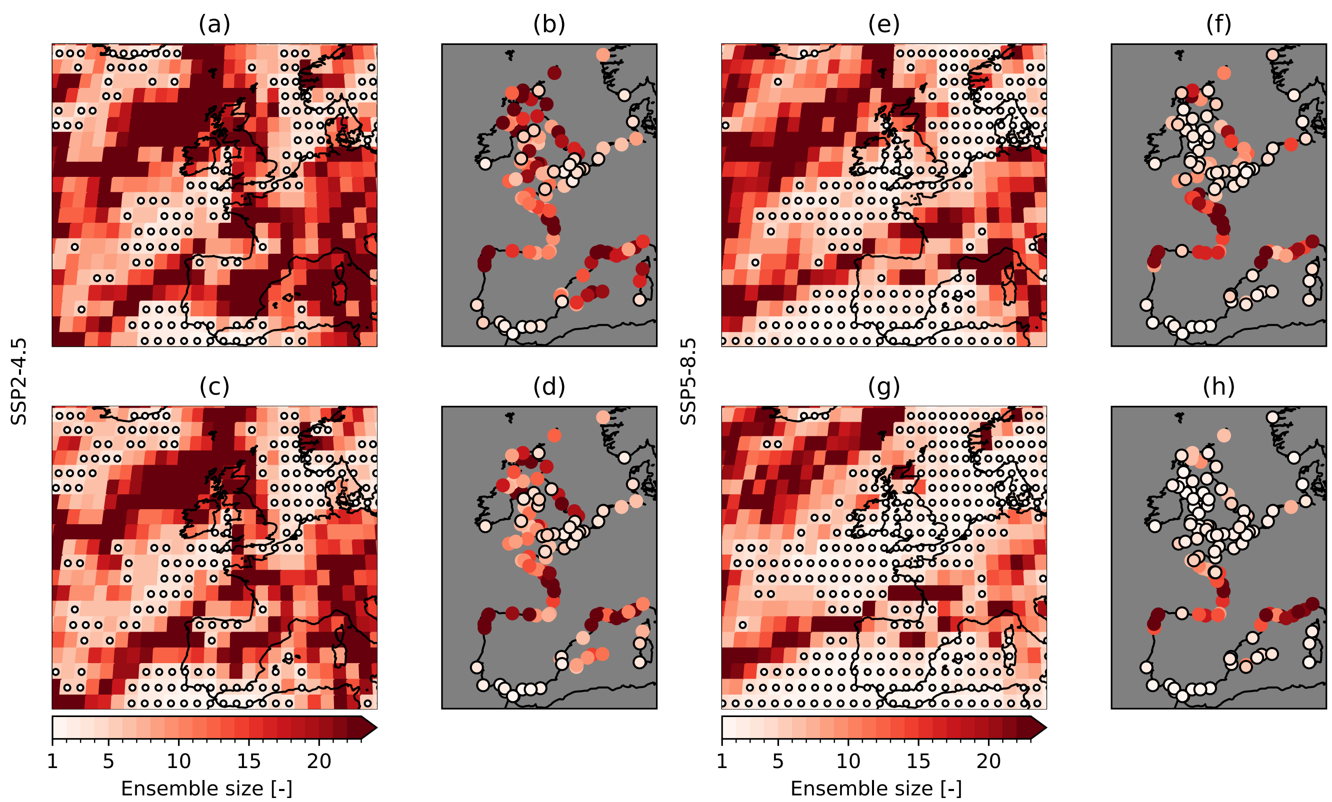

Figure 7. The minimum number of random Coupled Model Intercomparison Project 6 models that an ensemble needs to consist of for its mean (a, b) $\Delta N_{W\wedge P}$ and (c, d) $\Delta N_{S\wedge P}$ under SSP2-4.5 to have a 90% or higher probability of having the same sign as the projections in Section 4, for the periods 2041–2080 and 2061–2100 relative to 1981–2020, respectively, and (e)–(h) as in (a)–(d), under SSP5-8.5. The black-edged circles indicate where the minimum ensemble size is 5 or lower. Only one initial-condition member is used per model.

fewer models suffice under SSP5-8.5 (Figure 7h). This is consistent with the finding that the future changes in $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ are larger relative to the uncertainties due to inter-model differences and internal climate variability under stronger increases in emissions at most locations (as indicated by Figure 6 and the stippling in Figures 2 and 5). However, even under SSP5-8.5, large ensembles are needed in the Bay of Biscay and along the northern coastline of the Mediterranean Sea (Figures 7g and 7h). We note that the required ensemble sizes may change if we would weight models based on their performance and/or interdependence instead of applying model democracy.

While Figure 7 indicates that the climate model simulations used can strongly influence projections of $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ (including their sign), different definitions and analyses of compound extremes can also introduce differences between studies (Camus et al., 2021). When we re-do our projections with the 99th instead of the 98th percentile as the threshold for extremes, include a time-lag of up to 2 days between the extremes, or decluster the extremes using a 3-day window (following Haigh et al., 2016) prior to making the projections, the projections mainly change in magnitude (Figures S15 and S16 in Supporting Information S1). For instance, using a higher threshold percentile or declustering the extremes results in smaller $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$, whereas allowing a time-lag leads to larger $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$. However, the spatial patterns of the changes are not very sensitive to these methods (Figures S15 and S16 in Supporting Information S1), especially when compared to differences in $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ between models and initial-condition members.

6. Discussion and Conclusions

Previous projections of changes in the joint probability of drivers of compound flooding in Europe are based on only 5 to 6 CMIP5 simulations (Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020). In this study, we used a large ensemble of CMIP6 simulations, which we have shown to result in more robust and less uncertain projections. Based on these projections, the joint probability of storm surges and precipitation extremes will increase in the northwest of Europe (e.g., northwest of France, around the North Sea and in the UK), while it will decrease further south (e.g., in most of Spain and around the Mediterranean Sea). The spatial patterns of the ensemble mean change under the two emissions scenarios are similar, but the changes under SSP5-8.5 have a higher absolute magnitude than under SSP2-4.5 (49% vs 36%, on average). Previous studies for Europe only included a high emissions scenario (Bevacqua et al., 2019; Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020).

The changes in the joint probability of storm surge and precipitation extremes have a large-scale pattern similar to the changes in the joint probability of wind speed and precipitation extremes, but locally the differences can be large (e.g., in Scotland and in the Bay of Biscay). Therefore, we conclude that changes in the joint probability of wind speed and precipitation extremes are not always a good indication of changes in the potential for compound flooding. Nevertheless, they help to understand the latter physically. Namely, we find that changes in the marginal distributions of wind speed, storm surges and precipitation strongly resemble previously projected (thermo)dynamic changes of the atmosphere (e.g., Pfahl et al., 2017; Priestley & Catto, 2022; Zappa et al., 2013), as also concluded by Bevacqua, Vousdoukas, Zappa, et al. (2020). Our results additionally reveal that changes in the dependence between the extremes are at least partially related to shifts in the seasons in which the extremes tend to occur (Section 4.2).

Despite several methodological differences, the ensemble mean projections of the joint probability of storm surge and precipitation extremes under SSP5-8.5 seem to agree qualitatively with the projections of Bevacqua, Vousdoukas, Zappa, et al. (2020) in most regions, except southeast of Spain, in the Bay of Biscay and in the north of the Mediterranean Sea. Based on our results in Section 5, it is likely that the differences in these regions are mainly caused by the different and limited number of climate model simulations used by Bevacqua, Vousdoukas, Zappa, et al. (2020). We conclude that especially under not so high emissions scenarios and earlier in the twenty-first century, internal climate variability is large compared to the forced response of the number of joint extremes and relatively large ensembles are needed for qualitatively robust projections. This may partially explain why Ganguli et al. (2020) find decreases in the joint probability of storm surge and river discharge extremes in northwestern Europe for 2055 whereas our results and those of Bevacqua et al. (2019) and Bevacqua, Vousdoukas, Zappa, et al. (2020) indicate increases in the joint probability of storm surge and precipitation extremes in that region. However, this discrepancy could also be related to the fact that Ganguli et al. (2020) used downscaled and bias-corrected instead of raw climate model simulations. We consider it less likely that the differences are caused by their analysis of river discharge instead of precipitation extremes, as both are projected to increase in magnitude over the UK and western Europe (Sante et al., 2021).

As demonstrated in this study, applying large-scale statistical storm surge models such as that of Tadesse et al. (2020) to climate model simulations opens the door to projecting changes in the magnitude and frequency of storm surge extremes based on large ensembles of climate model simulations. This is a promising avenue for future research because projections of extreme storm surges and their co-occurrence with other flooding drivers are sensitive to internal variability and inter-model differences (Bevacqua, Vousdoukas, Zappa, et al., 2020; Calafat et al., 2022; Vousdoukas et al., 2017, 2018, & our results). We conclude that statistically modeled storm surges are appropriate for analyzing the joint probability of storm surge and precipitation extremes, even when using relatively coarse atmospheric forcing as input. Future work could investigate whether using more sophisticated statistical methods and/or weighting the extremes in the training data more strongly (e.g., Bellinghausen et al., 2023) could further improve the results. As the performance of the statistical storm surge model improves when using (time-lagged) sub-daily mean atmospheric forcing (Tadesse et al., 2020), it would also be beneficial if more climate models would provide their output at sub-daily frequencies. Data-proximate cloud computing could help to leverage such large data sets efficiently.

To compute ensemble statistics we weighted each CMIP6 model equally (Section 2.4). However, the performance of the models in simulating numbers of joint extremes varies (Ridder et al., 2021; Y. Wu et al., 2021). Assigning different weights to the CMIP6 models based on their performance may reduce the uncertainty in projections of compound extremes due to inter-model differences (Ridder et al., 2022). As we have shown, though, the historical number of joint extremes on which performance indicators are based is also affected by internal climate variability (Section 4.4). Therefore, skill scores such as those of Ridder et al. (2021) and Y. Wu et al. (2021) would be more reliable if computed using multiple initial-condition simulations per model, which are unfortunately not always available (Table 1).

Like several previous studies (e.g., Bevacqua, Vousdoukas, Zappa, et al., 2020; Ganguli et al., 2020; Gori et al., 2022; Moftakhari et al., 2017), we analyzed changes in the joint probability of two drivers of compound flooding (storm surges and precipitation) but did not consider how these changes in drivers affect the probability and risk of flooding itself. For instance, we excluded astronomical tides. Besides that tides are an additional driver of compound flooding (e.g., Gold et al., 2023; Piecuch et al., 2022), they also modulate the rate of surge-driven flooding and reduce the effective probability of compound surge-pluvial flooding (Bevacqua et al., 2019; Hague et al., 2023). Hence, changes in the joint probability of storm surge and precipitation extremes may have a larger effect on the probability of flooding in the south than in the northwest of Europe, because of the lower tidal ranges (Hague et al., 2023; Merrifield et al., 2013). We also excluded mean sea-level rise, which can increase the frequency of (compound) flooding if coastal flood protection is not adapted accordingly (e.g., Bevacqua et al., 2019; Hermans et al., 2023; Moftakhari et al., 2017).

Furthermore, to project changes in flood risk, additional local information such as on socioeconomic activity, land elevation and protective measures would need to be incorporated. Impact-based thresholds for the absolute magnitude of storm surges, precipitation and their combined effects could help to facilitate this (e.g., Hague et al., 2023; Sweet et al., 2018). To use statistical storm surge models in that context, a more extensive validation in terms of the absolute magnitude of predicted extremes would be warranted. While our projections may not directly reflect changes in compound flood risk, they do show that the potential for compound flooding due to extreme storm surges and precipitation in the northwest of Europe could increase under medium and high emissions scenarios. As we showed, the robustness of such projections hinges on the use of sufficiently large ensembles of climate model simulations.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The code to analyze the CMIP6 simulations on Google Cloud, make the projections and reproduce the figures in this manuscript is available at GitHub (https://github.com/Timh37/CMIP6cex) and published at Zenodo

(Hermans & Busecke, 2024a). The numbers of joint extremes in each CMIP6 simulation that were computed with the code and which form the basis for this study have also been published on Zenodo (Hermans & Busecke, 2024b).

Acknowledgments
T.H.J.H. and V.M.-S. were supported by PROTECT. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 869304, PROTECT contribution number 94. T.H.J.H. also received funding from the NPP programme of NWO. T.W. acknowledges funding from the U.S. National Science Foundation (award numbers 1929382 and 2141461). J.J.M.B acknowledges funding from the Gordon and Betty Moore Foundation (Grant 8434) and the National Science Foundation (Award 2019625). We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the ESGF for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We acknowledge the computing and storage resources provided by the "NSF Science and Technology Center (STC) Learning the Earth with Artificial intelligence and Physics (LEAP)" (Award #2019625). We thank Matthew Priestley, Giuseppe Zappa and Jeremy Röhmer for our useful discussions.

References

Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., et al. (2020). Historical simulations with HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, *12*(6), e2019MS001995. https://doi.org/10.1029/2019MS001995

Bellinghausen, K., Hünicke, B., & Zorita, E. (2023). *Short-term prediction of extreme sea-level at the Baltic Sea coast by Random Forests*. NHESS. https://doi.org/10.5194/nhess-2023-21

Bevacqua, E., Maraun, D., Haff, I. H., Widmann, M., & Vrac, M. (2017). Multivariate statistical modelling of compound events via pair-copula constructions: Analysis of floods in Ravenna (Italy). *Hydrology and Earth System Sciences*, *21*(6), 2701–2723. https://doi.org/10.5194/hess-21-2701-2017

Bevacqua, E., Maraun, D., Vousdoukas, M. I., Voukouvalas, E., Vrac, M., Mentaschi, L., & Widmann, M. (2019). Higher probability of compound flooding from precipitation and storm surge in Europe under anthropogenic climate change. *Science Advances*, *5*(9), eaaw5531. https://doi.org/10.1126/sciadv.aaw5531

Bevacqua, E., Suarez-Gutierrez, L., Jézéquel, A., Lehner, F., Vrac, M., Yiou, P., & Zscheischler, J. (2023). Advancing research on compound weather and climate events via large ensemble model simulations. *Nature Communications*, *14*(1), 2145. https://doi.org/10.1038/s41467-023-37847-5

Bevacqua, E., Vousdoukas, M. I., Shepherd, T. G., & Vrac, M. (2020). Brief communication: The role of using precipitation or river discharge data when assessing global coastal compound flooding. *Natural Hazards and Earth System Sciences*, *20*(6), 1765–1782. https://doi.org/10.5194/nhess-20-1765-2020

Bevacqua, E., Vousdoukas, M. I., Zappa, G., Hodges, K., Shepherd, T. G., Maraun, D., et al. (2020). More meteorological events that drive compound coastal flooding are projected under climate change. *Communications Earth and Environment*, *1*, 47. https://doi.org/10.1038/s43247-020-00044-z

Bi, D., Dix, M., Marsland, S., O'farrell, S., Sullivan, A., Bodman, R., et al. (2020). Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model. *Journal of Southern Hemisphere Earth Systems Science*, *70*(1), 225–251. https://doi.org/10.1071/ES19040

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002010. https://doi.org/10.1029/2019MS002010

Bruneau, N., Polton, J., Williams, J., & Holt, J. (2020). Estimation of global coastal sea level extremes using neural networks. *Environmental Research Letters*, *15*(7), 074030. https://doi.org/10.1088/1748-9326/ab89d6

Busecke, J. J. M., Spring, A., Ritschel, M., Maroon, E., & Nicholas, T. (2023). xMIP (v0.7.1) [Software]. https://doi.org/10.5281/zenodo.7519179

Busecke, J. J. M., & Stern, C. I. (2023). cmip6-leap-feedstock(v0.1.1) [Software]. https://doi.org/10.5281/zenodo.8400921

Calafat, F. M., Wahl, T., Tadesse, M. G., & Sparrow, S. N. (2022). Trends in Europe storm surge extremes match the rate of sea-level rise. *Nature*, *603*(7903), 841–845. https://doi.org/10.1038/s41586-022-04426-5

Camus, P., Haigh, I. D., Nasr, A. A., Wahl, T., Darby, S. E., & Nicholls, R. J. (2021). Regional analysis of multivariate compound coastal flooding potential around Europe and environs: Sensitivity analysis and spatial patterns. *Natural Hazards and Earth System Sciences*, *21*(7), 2021–2040. https://doi.org/10.5194/nhess-21-2021-2021

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., et al. (2019). Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. *Journal of Advances in Modeling Earth Systems*, *11*(1), 185–209. https://doi.org/10.1029/2018MS001369

Couasnon, A., Eilander, D., Muis, S., Veldkamp, T. I. E., Haigh, I. D., Wahl, T., et al. (2019). Measuring compound flood potential from river discharge and storm surge extremes at the global scale and its implications for flood hazard. *Natural Hazards and Earth System Sciences Discussions*. https://doi.org/10.5194/nhess-2019-205

Couasnon, A., Scussolini, P., Tran, T. V., Eilander, D., Muis, S., Wang, H., et al. (2022). A flood risk framework capturing the seasonality of and dependence between rainfall and sea levels—An application to Ho Chi Minh City, Vietnam. *Water Resources Research*, *58*(2), e2021WR030002. https://doi.org/10.1029/2021WR030002

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001916. https://doi.org/10.1029/2019MS001916

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., et al. (2022). The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. *Geoscientific Model Development*, *15*(7), 2973–3020. https://doi.org/10.5194/gmd-15-2973-2022

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., et al. (2020). The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2019MS002015. https://doi.org/10.1029/2019MS002015

Eilander, D., Couasnon, A., Ikeuchi, H., Muis, S., Yamazaki, D., Winsemius, H. C., & Ward, P. J. (2020). The effect of surge on riverine flood hazard and impact in deltas globally. *Environmental Research Letters*, *15*(10), 104007. https://doi.org/10.1088/1748-9326/ab8ca6

Emanuel, K. (2017). Assessing the present and future probability of Hurricane Harvey's rainfall. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(48), 12681–12684. https://doi.org/10.1073/pnas.1716222114

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., et al. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Ganguli, P., & Merz, B. (2019). Extreme coastal water levels exacerbate fluvial flood hazards in Northwestern Europe. *Scientific Reports*, *9*(1), 13165. https://doi.org/10.1038/s41598-019-49822-6

Ganguli, P., Paprotny, D., Hasan, M., Güntner, A., & Merz, B. (2020). Projected changes in compound flood hazard from riverine and coastal floods in Northwestern Europe. *Earth's Future*, *8*(11), e2020EF001752. https://doi.org/10.1029/2020EF001752

Ghanbari, M., Arabi, M., Obeysekera, J., & Sweet, W. (2019). A coherent statistical model for coastal flood frequency analysis under nonstationary sea level conditions. *Earth's Future*, *7*(2), 162–177. https://doi.org/10.1029/2018EF001089

Gold, A., Anarde, K., Grimley, L., Neve, R., Srebnik, E. R., Thelen, T., et al. (2023). Data from the drain: A sensor framework that captures multiple drivers of chronic coastal floods. *Water Resources Research*, *59*(4), e2022WR032392. https://doi.org/10.1029/2022WR032392

Gori, A., Lin, N., Xi, D., & Emanuel, K. (2022). Tropical cyclone climatology change greatly exacerbates US extreme rainfall–surge hazard. *Nature Climate Change*, *12*(2), 171–178. https://doi.org/10.1038/s41558-021-01272-7
Hague, B. S., McGregor, S., Jones, D. A., Reef, R., Jakob, D., & Murphy, B. F. (2023). The global drivers of chronic coastal flood hazards under sea-level rise. *Earth's Future*, *11*(8), e2023EF003784. https://doi.org/10.1029/2023EF003784
Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Hague, B. S., et al. (2021). GESLA Version 3: A major update to the global higher-frequency sea-level dataset. *EarthArXiv Preprint*. https://doi.org/10.31223/X5MP65
Haigh, I. D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., et al. (2016). Spatial and temporal analysis of extreme sea level and storm surge events around the coastline of the UK. https://doi.org/10.1038/sdata.2016.107
Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., et al. (2020). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geoscientific Model Development*, *13*(5), 2197–2244. https://doi.org/10.5194/gmd-13-2197-2020
Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., et al. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, *11*, 3691–3727. https://doi.org/10.1029/2019MS001829
Hendry, A., Haigh, I. D., Nicholls, R. J., Winter, H., Neal, R., Wahl, T., et al. (2019). Assessing the characteristics and drivers of compound flooding events around the UK coast. *Hydrology and Earth System Sciences*, *23*(7), 3117–3139. https://doi.org/10.5194/hess-23-3117-2019
Hermans, T. H., & Busecke, J. J. M. (2024a). CMIP6cex (v1.0.0) [Software]. https://doi.org/10.5281/zenodo.10817904
Hermans, T. H., & Busecke, J. J. M. (2024b). Number of joint precipitation and either wind or storm surge extremes based on CMIP6 simulations [Dataset]. https://doi.org/10.5281/zenodo.10818223
Hermans, T. H., Malagón-Santos, V., Katsman, C. A., Jane, R. A., Rasmussen, D. J., Haasnoot, M., et al. (2023). The timing of decreasing coastal flood protection due to sea-level rise. *Nature Climate Change*, *13*(4), 359–366. https://doi.org/10.1038/s41558-023-01616-5
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803
Hoyer, S., & Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, *5*(1), 10. https://doi.org/10.5334/jors.148
Jaroszweski, D., Hooper, E., Baker, C., Chapman, L., & Quinn, A. (2015). The impacts of the 28 June 2012 storms on UK road and rail transport. *Meteorological Applications*, *22*(3), 470–476. https://doi.org/10.1002/met.1477
Kew, S. F., Selten, F. M., Lenderink, G., & Hazeleger, W. (2013). The simultaneous occurrence of surge and discharge extremes for the Rhine delta. *Natural Hazards and Earth System Sciences*, *13*(8), 2017–2029. https://doi.org/10.5194/nhess-13-2017-2013
Kumbier, K., Carvalho, R. C., Vafeidis, A. T., & Woodroffe, C. D. (2018). Investigating compound flooding in an estuary using hydrodynamic modelling: A case study from the Shoalhaven River, Australia. *Natural Hazards and Earth System Sciences*, *18*(2), 463–477. https://doi.org/10.5194/nhess-18-463-2018
Lambert, E., Rohmer, J., Cozannet, G. L., & van de Wal, R. S. W. (2020). Adaptation time to magnified flood hazards underestimated when derived from tide gauge records. *Environmental Research Letters*, *15*(7), 074015. https://doi.org/10.1088/1748-9326/ab8336
Lee, J., Kim, J., Sun, M. A., Kim, B. H., Moon, H., Sung, H. M., et al. (2020). Evaluation of the Korea Meteorological Administration Advanced Community Earth-System model (K-ACE). *Asia-Pacific Journal of Atmospheric Sciences*, *56*(3), 381–395. https://doi.org/10.1007/s13143-019-00144-7
Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., et al. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth System Dynamics*, *11*(2), 491–508. https://doi.org/10.5194/esd-11-491-2020
Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., Mcinnes, K., et al. (2014). A compound event framework for understanding extreme impacts. https://doi.org/10.1002/wcc.252
Li, C., Zwiers, F., Zhang, X., Li, G., Sun, Y., & Wehner, M. (2021). Changes in annual extremes of daily temperature and precipitation in CMIP6 models. *Journal of Climate*, *34*, 3411–3460. https://doi.org/10.1175/JCLI-D-19
Li, L., Yu, Y., Tang, Y., Lin, P., Xie, J., Song, M., et al. (2020). The Flexible Global Ocean-Atmosphere-Land System Model Grid-Point Version 3 (FGOALS-g3): Description and evaluation. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2019MS002012. https://doi.org/10.1029/2019MS002012
Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., et al. (2022). CMIP6 simulations with the CMCC Earth System Model (CMCC-ESM2). *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002814. https://doi.org/10.1029/2021MS002814
Martius, O., Pfahl, S., & Chevalier, C. (2016). A global quantification of compound precipitation and wind extremes. *Geophysical Research Letters*, *43*(14), 7709–7717. https://doi.org/10.1002/2016GL070017
Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO_2. *Journal of Advances in Modeling Earth Systems*, *11*(4), 998–1038. https://doi.org/10.1029/2018MS001400
Meinshausen, M., Nicholls, Z., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., et al. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, *13*(8), 3571–3605. https://doi.org/10.5194/gmd-13-3571-2020
Merrifield, M. A., Genz, A. S., Kontoes, C. P., & Marra, J. J. (2013). Annual maximum water levels from tide gauges: Contributing factors and geographic patterns. *Journal of Geophysical Research: Oceans*, *118*(5), 2535–25466. https://doi.org/10.1002/jgrc.20173
Moftakhari, H. R., Salvadori, G., AghaKouchak, A., Sanders, B. F., & Matthew, R. A. (2017). Compounding effects of sea level rise and fluvial flooding. *Proceedings of the National Academy of Sciences*, *114*(37), 9785–9790. https://doi.org/10.1073/pnas.1620325114
Muis, S., Aerts, J. C., José, J. A., Dullaart, J. C., Duong, T. M., Erikson, L., 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
Muis, S., Apecechea, M. I., Dullaart, J., de Lima Rego, J., Madsen, K. S., Su, J., et al. (2020). A high-resolution global dataset of extreme sea levels, tides, and storm surges, including future projections. *Frontiers in Marine Science*, *7*, 263. https://doi.org/10.3389/fmars.2020.00263
Muis, S., Verlaan, M., Winsemius, H. C., Aers, J. C. J. H., & Ward, P. J. (2016). A global reanalysis of storm surges and extreme sea levels. *Nature Communications*, *7*(1), 11969. https://doi.org/10.1038/ncomms11969
Nasr, A. A., Wahl, T., Rashid, M. M., Camus, P., & Haigh, I. D. (2021). Assessing the dependence structure between oceanographic, fluvial, and pluvial flooding drivers along the United States coastline. *Hydrology and Earth System Sciences*, *25*(12), 6203–6222. https://doi.org/10.5194/hess-25-6203-2021
Owen, L. E., Catto, J. L., Stephenson, D. B., & Dunstone, N. J. (2021). Compound precipitation and wind extremes over Europe and their relationship to extratropical cyclones. *Weather and Climate Extremes*, *33*, 100342. https://doi.org/10.1016/j.wace.2021.100342
Paprotny, D., Morales-Nápoles, O., & Jonkman, S. N. (2018). HANZE: A pan-European database of exposure to natural hazards and damaging historical floods since 1870. *Earth System Science Data*, *10*(1), 565–581. https://doi.org/10.5194/essd-10-565-2018

Paprotny, D., Vousdoukas, M. I., Morales-Nápoles, O., Jonkman, S. N., & Feyen, L. (2020). Pan-European hydrodynamic models and their ability to identify compound floods. *Natural Hazards*, *101*(3), 933–957. https://doi.org/10.1007/s11069-020-03902-3

Pfahl, S., O'Gorman, P., & Fischer, E. (2017). Understanding the regional pattern of projected future changes in extreme precipitation. *Nature Climate Change*, *7*(6), 423–427. https://doi.org/10.1038/nclimate3287

Piecuch, C. G., Coats, S., Dangendorf, S., Landerer, F. W., Reager, J. T., Thompson, P. R., & Wahl, T. (2022). High-tide floods and storm surges during atmospheric rivers on the US West Coast. *Geophysical Research Letters*, *49*(2), e2021GL096820. https://doi.org/10.1029/2021GL096820

Priestley, M. D. K., & Catto, J. L. (2022). Future changes in the extratropical storm tracks and cyclone intensity, wind speed, and structure. *Weather and Climate Dynamics*, *3*(1), 337–360. https://doi.org/10.5194/wcd-3-337-2022

Ridder, N. N., Pitman, A. J., & Ukkola, A. M. (2021). Do CMIP6 climate models simulate global or regional compound events skillfully? https://doi.org/10.1029/2020GL091152

Ridder, N. N., Pitman, A. J., Westra, S., Ukkola, A., Hong, X. D., Bador, M., et al. (2020). Global hotspots for the occurrence of compound events. *Nature Communications*, *11*(1), 5956. https://doi.org/10.1038/s41467-020-19639-3

Ridder, N. N., Ukkola, A. M., Pitman, A. J., & Perkins-Kirkpatrick, S. E. (2022). Increased occurrence of high impact compound events under climate change. *npj Climate and Atmospheric Science*, *5*(1), 3. https://doi.org/10.1038/s41612-021-00224-4

Ruocco, A. C., Nicholls, R. J., Haigh, I. D., & Wadey, M. P. (2011). Reconstructing coastal flood occurrence combining sea level and media sources: A case study of the Solent, UK since 1935. *Natural Hazards*, *59*(3), 1773–1796. https://doi.org/10.1007/s11069-011-9868-7

Sante, F. D., Coppola, E., & Giorgi, F. (2021). Projections of river floods in Europe using EURO-CORDEX, CMIP5 and CMIP6 simulations. *International Journal of Climatology*, *41*(5), 3203–3221. https://doi.org/10.1002/joc.7014

Santos, V. M., Casas-Prat, M., Poschlod, B., Ragno, E., Hurk, B. V. D., Hao, Z., et al. (2021). Statistical modelling and climate variability of compound surge and precipitation events in a managed water system: A case study in The Netherlands. *Hydrology and Earth System Sciences*, *25*(6), 3595–3615. https://doi.org/10.5194/hess-25-3595-2021

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., et al. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, *13*(12), 6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Sellar, A. A., Walton, J., Jones, C. G., Wood, R., Abraham, N. L., Andrejczuk, M., et al. (2020). Implementation of U.K. Earth System Models for CMIP6. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2019MS001946. https://doi.org/10.1029/2019MS001946

Seneviratne. (2021). *Weather and climate extreme events in a changing climate* (pp. 1513–1766). Cambridge University Press. https://doi.org/10.1017/9781009157896.013

Severino, L. G., Kropf, C. M., Afargan-Gerstman, H., Fairless, C., Vries, A. J. D., Domeisen, D. I. V., & Bresch, D. N. (2023). *Projections and uncertainties of future winter windstorm damage in Europe*. EGUsphere. https://doi.org/10.5194/egusphere-2023-205

Sun, P., Zou, Y., Yao, R., Ma, Z., Bian, Y., Ge, C., & Lv, Y. (2023). Compound and successive events of extreme precipitation and extreme runoff under heatwaves based on CMIP6 models. *Science of the Total Environment*, *878*, 162980. https://doi.org/10.1016/j.scitotenv.2023.162980

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019). The Canadian Earth System Model version 5 (CanESM5.0.3). *Geoscientific Model Development*, *12*(11), 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Sweet, W., Dusek, G., Obeysekera, J., & Marra, J. J. (2018). Patterns and projections of high tide flooding along the U.S. coastline using a common impact threshold.

Tadesse, M., Wahl, T., & Cid, A. (2020). Data-driven modeling of global storm surges. *Frontiers in Marine Science*, *7*, 512653. https://doi.org/10.3389/fmars.2020.00260

Tadesse, M. G., & Wahl, T. (2021). A database of global storm surge reconstructions. *Scientific Data*, *8*(1), 125. https://doi.org/10.1038/s41597-021-00906-x

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., et al. (2019). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, *12*(7), 2727–2765. https://doi.org/10.5194/gmd-12-2727-2019

Taylor, K., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Tiggeloven, T., Couasnon, A., van Straaten, C., Muis, S., & Ward, P. J. (2021). Exploring deep learning capabilities for surge predictions in coastal areas. *Scientific Reports*, *11*(1), 17224. https://doi.org/10.1038/s41598-021-96674-0

van den Hurk, B., van Meijgaard, E., de Valk, P., van Heeringen, K.-J., & Gooijer, J. (2015). Analysis of a compounding surge and precipitation event in The Netherlands. *Environmental Research Letters*, *10*, 35001.

Volodin, E. M., & Gritsun, A. S. (2018). Simulation of possible future climate changes in the 21st century in the INM-CM5 climate model. *Earth System Dynamics*, *9*(3), 1235–1242. https://doi.org/10.1134/S0001433820030123

Volodin, E. M., Mortikov, E. V., Kostrykin, S. V., Galin, V. Y., Lykossov, V. N., Gritsun, A. S., et al. (2017). Simulation of the present-day climate with the climate model INMCM5. *Climate Dynamics*, *49*(11–12), 3715–3734. https://doi.org/10.1007/s00382-017-3539-7

Vousdoukas, M. I., Mentaschi, L., Feyen, L., & Voukouvalas, E. (2017). Earth's future extreme sea levels on the rise along Europe's coasts Earth's Future. *Earth's Future*, *5*, 1–20. https://doi.org/10.1002/eft2.192

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P., & Feyen, L. (2018). Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. *Nature Communications*, *9*(1), 2360. https://doi.org/10.1038/s41467-018-04692-w

Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, *5*(12), 1093–1097. https://doi.org/10.1038/nclimate2736

Wang, Y. C., Hsu, H. H., Chen, C. A., Tseng, W. L., Hsu, P. C., Lin, C. W., et al. (2021). Performance of the Taiwan Earth System Model in simulating climate variability compared with observations and CMIP6 model simulations. *Journal of Advances in Modeling Earth Systems*, *13*(7), e2020MS002353. https://doi.org/10.1029/2020MS002353

Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., et al. (2018). Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. *Environmental Research Letters*, *13*(8), 084012. https://doi.org/10.1088/1748-9326/aad400

Woodworth, P. L., Hunter, J. R., Marcos, M., Caldwell, P., Menéndez, M., & Haigh, I. (2016). Towards a global higher-frequency sea level dataset. *Geoscience Data Journal*, *3*(2), 50–59. https://doi.org/10.1002/gdj3.42

Wu, W., McInnes, K., O'Grady, J., Hoeke, R., Leonard, M., & Westra, S. (2018). Mapping dependence between extreme rainfall and storm surge. *Journal of Geophysical Research: Oceans*, *123*(4), 2461–2474. https://doi.org/10.1002/2017JC013472

Wu, Y., Miao, C., Sun, Y., AghaKouchak, A., Shen, C., & Fan, X. (2021). Global observations and CMIP6 simulations of compound extremes of monthly temperature and precipitation. *GeoHealth*, *5*, e2021GH000390. https://doi.org/10.1029/2021GH000390

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., et al. (2019). The meteorological research institute Earth system model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component. *Journal of the Meteorological Society of Japan*, *97*(5), 931–965. https://doi.org/10.2151/jmsj.2019-051

Zappa, G., Shaffrey, L. C., Hodges, K. I., Sansom, P. G., & Stephenson, D. B. (2013). A multimodel assessment of future projections of North Atlantic and European extratropical cyclones in the CMIP5 climate models*. *Journal of Climate*, *26*(16), 5846–5862. https://doi.org/10.1175/JCLI-D

Zheng, F., Westra, S., Leonard, M., & Sisson, S. A. (2014). Modeling dependence between extreme rainfall and storm surge to estimate coastal flooding risk. *Water Resources Research*, *50*(3), 2050–2071. https://doi.org/10.1002/2013WR014616

Zhuang, J., Dussin, R., Huard, D., Bourgault, P., Banihirwe, A., Raynaud, S., et al. (2023). xESMF: v0.8.2 [Software]. https://doi.org/10.5281/zenodo.8356796

Zscheischler, J., Westra, S., Hurk, B. J. V. D., Seneviratne, S. I., Ward, P. J., Pitman, A., et al. (2018). Future climate risk from compound events. *Nature Climate Change*, *8*(6), 469–477. https://doi.org/10.1038/s41558-018-0156-3