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, USA
[4]NIOZ Royal Netherlands Institute for Sea Research, Department of Estuarine & Delta Systems, PO Box 140, 4400AC Yerseke, The Netherlands
[5]Hazen and Sawyer, Orlando, FL, USA
[6]Department of Physical Geography, Utrecht University, Utrecht, 3584 CB, The Netherlands

Correspondence: Tim H.J. Hermans (t.h.j.hermans@uu.nl)

Key points

– We project changes in the joint probability of storm surge and precipitation extremes based on a large ensemble of CMIP6 models.

– The probability will increase in the northwest and decrease in the southwest of Europe, with an absolute magnitude of on average 36-49%.

– Because previous projections are based on only few climate model simulations, they can can differ from our projections qualitatively.

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% to 49% by 2080, depending on the scenario. The large-scale changes in the joint probability of extreme wind speed and precipitation are similar, but locally, differences between the changes in the two types of compound extremes can exceed the changes themselves. Due to internal climate variability and inter-model differences, projections based on only 5 to 6 random climate model simulations 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 21st 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% to 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.

Keywords

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., Paprotny et al., 2018; Kumbier et al., 2018; Emanuel, 2017; Bevacqua et al., 2017; 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., Moftakhari et al., 2017; Wahl et al., 2015; Leonard et al., 2014), compound events have received increased attention in the past decade. For instance, the dependence between and joint probability of various combinations of flooding drivers has been assessed at local (e.g., Santos et al., 2021; Kew et al., 2013; Zheng et al., 2014; Couasnon et al., 2022), national (e.g., Hendry et al., 2019; Wu et al., 2018), continental (e.g., Wahl et al., 2015; Paprotny et al., 2018, 2020; Ganguli and Merz, 2019; Camus et al., 2021; Nasr et al., 2021) and global scales (e.g., Ward et al., 2018; Couasnon et al., 2019; Eilander et al., 2020; Bevacqua et al., 2020b; Ridder et al., 2020; Lambert et al., 2020), using observations and/or model hindcasts.

In contrast, few 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 et al., 2020b). 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 (Moftakhari et al., 2017; Ghanbari et al., 2019; Gori et al., 2022). 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 et al., 2020a).

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 et al. (2020b) 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; Wu et al., 2021). Some studies used larger multi-model ensembles to project changes in the joint probability of extremes (e.g., Ridder et al., 2022; Sun et al., 2023; Bevacqua 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 et al., 2020b; 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

(Bruneau et al., 2020; Tadesse et al., 2020; Tadesse and Wahl, 2021; Tiggeloven et al., 2021; Bellinghausen et al., 2023). Such statistical models, for instance based on multi-linear regression 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 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 joint storm surge and precipitation extremes. We limit our study to Europe, where data-driven storm surge models generally perform well (Tadesse et al., 2020; Bruneau 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.

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 (SSP2-4.5) and a high (SSP5-8.5) emissions scenario (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.

115   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 dataset includes over

120   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 / ESGF Cloud Data Working Group (https://pangeo-data.github.io/pangeo-cmip6-cloud/), held in public Google Cloud Storage. The datasets summarized in Table 1 reflect datasets that were available to download and ingest via the pangeo-forge feedstock (Busecke and Stern, 2023) at the time of writing of this manuscript. The data is analyzed using the code in the CMIP6cex repository (Hermans

125   and Busecke, TBD), for which the xarray (Hoyer and 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). Ensemble statistics are computed and displayed on this grid. The regridded simulations are also used as input to the

130   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: (i) the combination of extreme daily mean wind speed and extreme daily accumulated precipitation, and (ii) 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.,

135   2015), we focus on the case in which both drivers are extreme, similar to previous studies (Bevacqua et al., 2019, 2020b; 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

140   declustered extremes for each location and variable (e.g., Ridder et al., 2020; Bevacqua et al., 2020b; Hendry et al., 2019; Camus et al., 2021; Ganguli et al., 2020); a pragmatic approach which we also adopt here. For Europe, Camus et al. (2021) found that using 3 v.s. 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

145    $P = p \geq p_{98}$, $W = w \geq w_{98}$ 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.

Table 1. CMIP6 simulations used.

	Model	SSP2-4.5 [#]	SSP5-8.5 [#]	Both [#]	°Lon x °Lat	Reference
1	ACCESS-CM2	5	6	4	1.875 x 1.25	(Bi et al., 2020)
2	ACCESS-ESM1-5	38	35	33	1.875 x 1.25	(Bi et al., 2020)
3	CanESM5	25	25	25	2.8 x 2.8	(Swart et al., 2019)
4	CESM2	2	2	2	1.25 x 0.9	(Danabasoglu et al., 2020)
5	CESM2-WACCM	3	3	3	1.25 x 0.9	(Danabasoglu et al., 2020)
6	CMCC-ESM2	1	1	1	1.25 x 0.9	(Lovato et al., 2022)
7	CMCC-CM2-SR5	1	1	1	1.25 x 0.9	(Cherchi et al., 2019)
8	EC-Earth3	59	1	1	0.75 x 0.75	(Döscher et al., 2022)
9	EC-Earth3-Veg	1	0	0	0.75 x 0.75	(Döscher et al., 2022)
10	FGOALS-g3	1	0	0	2 x 2	(Li et al., 2020)
11	GFDL-CM4	1	1	1	1 x 1	(Held et al., 2019)
12	GFDL-ESM4	1	1	1	1 x 1	(Dunne et al., 2020)
13	HadGEM3-GC31-LL	5	4	4	1.875 x 1.25	(Andrews et al., 2020)
14	HadGEM3-GC31-MM	0	4	0	0.83 x 0.56	(Andrews et al., 2020)
15	INM-CM4-8	1	1	1	2 x 1.5	(Volodin and Gritsun, 2018)
16	INM-CM5-8	1	1	1	2 x 1.5	(Volodin et al., 2017)
17	IPSL-CM6A-LR	11	7	6	2.5 x 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 x 1.4	(Tatebe et al., 2019)
20	MIROC6-ES2L	10	1	1	2.8 x 2.8	(Hajima et al., 2020)
21	MPI-ESM1-2-LR	24	24	24	1.88 x 1.88	(Mauritsen et al., 2019)
22	MPI-ESM1-2-HR	2	2	2	0.93 x 0.93	(Mauritsen et al., 2019)
23	MRI-ESM2-0	1	1	1	0.75 x 0.75	(Yukimoto et al., 2019)
24	NorESM2-LL	3	1	1	2.5 x 1.88	(Øyvind Seland et al., 2020)
25	NorESM2-MM	2	1	1	1.25 x 0.94	(Øyvind Seland et al., 2020)
26	TaiESM1	1	1	1	1.25 x 0.9	(Wang et al., 2021)
27	UKESM1-0-LL	5	5	5	1.875 x 1.25	(Sellar et al., 2020)

2.3 Future changes in the joint probability of extremes

150 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 Hendry et al. (2019); Camus et al. (2021); Couasnon et al. (2019); Ridder et al. (2020) and Ridder et al. (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

155 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.

160 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).

165 ### 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 Supporting Fig. 2 for a graphical explanation). Using methods similar to those of Bevacqua et al. (2019, 2020b), 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

170 $$\Delta N_{W \wedge P}^{w} = |w^{hist} \geq w_{U_w}^{hist} \wedge p^{hist} \geq p_{98}^{hist}| - N_{W \wedge P}^{hist} \tag{4}$$

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 Eqs. 4 and 5, the changed threshold percentiles are defined as $U_w = $

175 $F_w^{fut}(w_{98}^{hist})$ and $U_p = F_p^{fut}(p_{98}^{hist})$, 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

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

180   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:

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

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 Eqs. 1-3) can be written as $w_{98}^{hist} = w_{U_w}^{fut}$ and $p_{98}^{hist} = p_{U_p}^{fut}$.

185   Therefore, by using Eq. 7 to compute $\Delta N_{W \wedge P}^{dependence}$, 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

190   instead of w. Table 2 summarizes the notations defined in this section.

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.
$N_{W \wedge P}^{hist}$	$N_{W \wedge P}$ in the period 1981-2020 as simulated by CMIP6 models.
$N_{W \wedge P}^{fut}$	$N_{W \wedge P}$ in the period 2061-2100 as simulated by CMIP6 models.
$\Delta N_{W \wedge P}$	The difference between $N_{W \wedge P}^{fut}$ and $N_{W \wedge P}^{hist}$.
$\Delta N_{W \wedge P}^{w}$	$\Delta N_{W \wedge P}$ due to changes in the marginal distribution of wind speed.
$\Delta N_{W \wedge P}^{p}$	$\Delta N_{W \wedge P}$ due to changes in the marginal distribution of precipitation.
$\Delta N_{W \wedge P}^{w,p}$	$\Delta N_{W \wedge P}$ due to changes in the marginal distributions of wind speed and precipitation.
$\Delta N_{W \wedge P}^{dependence}$	$\Delta N_{W \wedge P}$ 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.
$N_{S \wedge P}^{hist}$	$N_{S \wedge P}$ in the period 1981-2020 as simulated by CMIP6 models.
$N_{S \wedge P}^{fut}$	$N_{S \wedge P}$ in the period 2061-2100 as simulated by CMIP6 models.
$\Delta N_{S \wedge P}$	The difference between $N_{S \wedge P}^{fut}$ and $N_{S \wedge P}^{hist}$.
$\Delta N_{S \wedge P}^{s}$	$\Delta N_{S \wedge P}$ due to changes in the marginal distribution of storm surges.
$\Delta N_{S \wedge P}^{p}$	$\Delta N_{S \wedge P}$ due to changes in the marginal distribution of precipitation.
$\Delta N_{S \wedge P}^{s,p}$	$\Delta N_{S \wedge P}$ due to changes in the marginal distributions of storm surges and precipitation.
$\Delta N_{S \wedge P}^{dependence}$	$\Delta N_{S \wedge P}$ due to changes in the dependence between storm surges and precipitation.

2.4 Computing ensemble statistics

Not every CMIP6 model provides the same number of initial-condition simulations (see Table 1). To include all 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 the Supporting Information.

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 $N_{W \wedge P}^{hist}$ and $N_{S \wedge P}^{hist}$ between initial-condition simulations using the CMIP6 models that have at least 5 initial-condition members (see Table 1). 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
205 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.

To study how internal climate variability and inter-model differences affect projections of the joint probability of extremes
210 based on small climate model ensembles (e.g., Ganguli et al., 2020; Bevacqua et al., 2020b), 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 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 as that of the ensemble including all CMIP6 models and initial-condition members. Finally, as an indication for how large
215 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 multi-linear regression (MLR) model based on
220 the methods of Tadesse et al. (2020) and Tadesse and Wahl (2021), 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 reanalyses as predictors. Predictors were used within either $10°$ by $10°$ (Tadesse et al., 2020) or $6°$ by $6°$ (Tadesse and Wahl, 2021) grids around each TG and lags up to 30 hours between the predictands and predictors were implemented. The daily maximum non-
225 tidal residuals of Tadesse et al. (2020), which we also use here, were obtained by removing the annual mean sea level and predicted astronomical tides from TG records in the GESLA2 dataset (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
230 around each TG. Since we use daily mean CMIP6 data, 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 and Wahl, 2021), we use $9°$ by $9°$ grids around each TG. 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

10

235   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 et al. (2020) and Tadesse and Wahl (2021). As the flowchart in Supporting Fig. 1 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.

240   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 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 Supporting

245   Fig. 1). 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

250   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 Supporting Fig. 1). For each simulation, this results in estimates of daily maximum non-tidal residuals at every tide gauge during 1850-2100, which we refer to as storm surges.

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 $N_{S \wedge P}$.

255   Therefore, we evaluate the model by comparing $N_{S \wedge P}$ based on the statistically modelled storm surges ($N_{S_{MLR} \wedge P}$) with that based on the observed daily maximum non-tidal residuals from GESLA2 ($N_{S_{G2} \wedge P}$) (Fig. 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 Dataset for the Evaluation of Climate Impact (CoDEC) ($N_{S_{CoDEC} \wedge P}$). This dataset was simulated with the high-resolution Global Tide and Surge Model (GTSM) driven by atmospheric forcing from ERA5 (Muis et al., 2020). Furthermore, we also compare $N_{W \wedge P}$ based on daily

260   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 Supporting Fig. 1 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 (Fig. 1a). This pattern is consistent with the results of previous studies (Paprotny et al., 2018; Bevacqua

265   et al., 2019; Hendry et al., 2019; Couasnon et al., 2019). 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 Figs. 1a & b), 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 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 (Fig. 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 modelled storm surges is very similar to that based on the storm surges simulated with GTSM (compare Figs. 1b & c). Although the biases of $N_{S_{MLR} \wedge P}$ (nRMSE = 0.36, Fig. 1f) are moderately larger than those of $N_{S_{CoDEC} \wedge P}$ (nRMSE = 0.29, Fig. 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 $N_{S_{G2} \wedge P}$ (Fig. 1a) and $N_{W_{ERA5} \wedge P}$ (Fig. 1d), which are shown in Fig. 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 Figs. 1f & h 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 $N_{S_{MLR} \wedge P}$.

The temporal coverage of the TG records is limited at several locations (see Supporting Fig. 1). Consequently, at some of these locations, the evaluation in Fig. 1 is based on only a couple of observed joint storm surge and precipitation extremes. To test whether the results in Fig. 1 are robust to using a larger sample size, we repeated the evaluation with lower threshold percentiles (Fig. 1e). Other than that the error metrics tend to improve for lower thresholds, which may also partially reflect the inclusion of less extreme events, the performance of $N_{S_{MLR} \wedge P}$ remains similar to that of $N_{S_{CoDEC} \wedge P}$ (Fig. 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 (Supporting Fig. 3). 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 tide gauges.
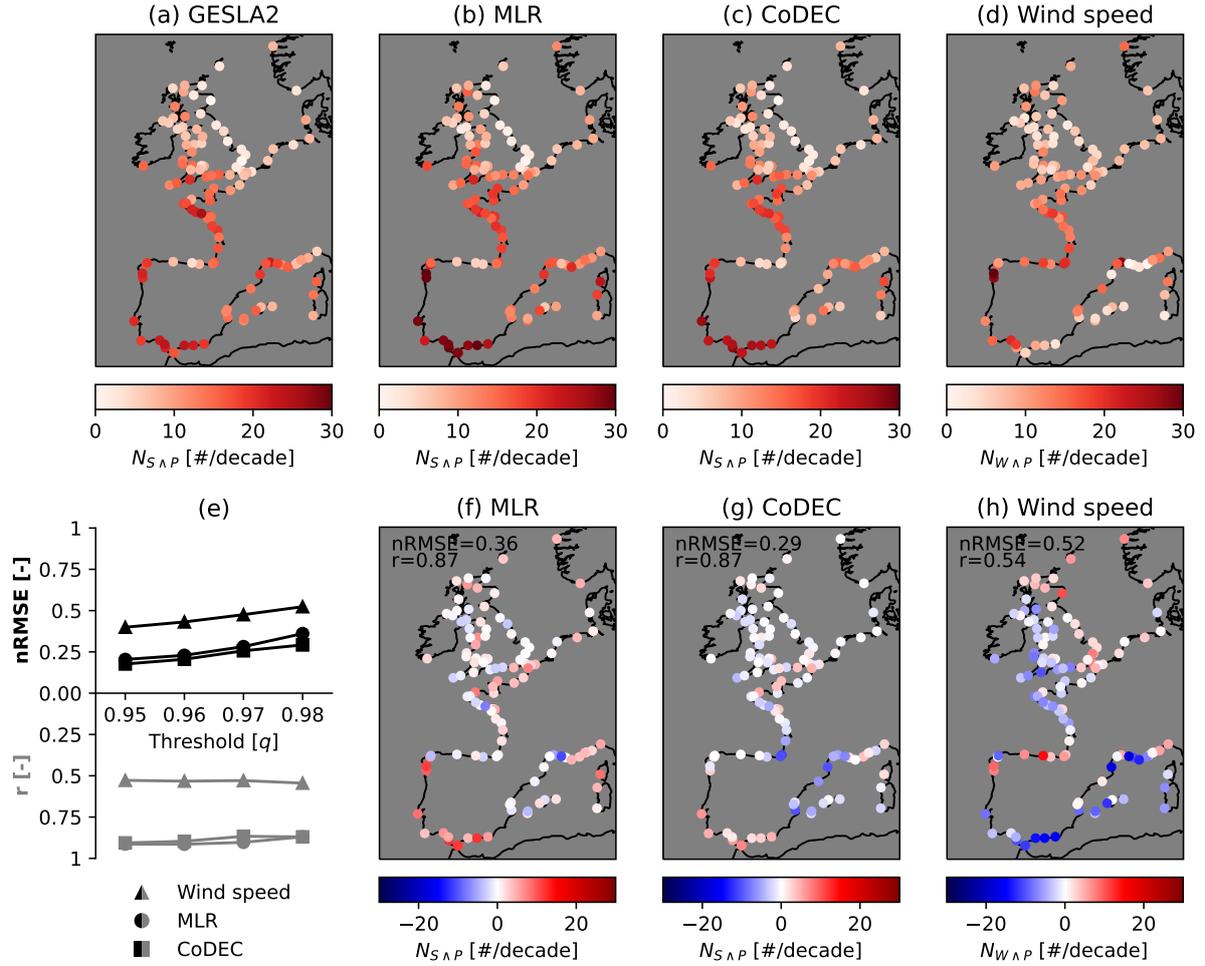
Fig. 1. Number of joint extreme **(a)** storm surges from GESLA2 and precipitation from ERA5 ($N_{S_{G2} \wedge P}$), **(b)** statistically modelled storm surges and precipitation from ERA5 ($N_{S_{MLR} \wedge P}$), **(c)** storm surges from CoDEC and precipitation from ERA5 ($N_{S_{CoDEC} \wedge P}$) and **(d)** wind speed and precipitation from ERA5 ($N_{W_{ERA5} \wedge P}$) at GESLA2 TGs [#/decade]. **(e)** the correlation coefficient r and 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 ($p << 0.05$).

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, Fig. 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

13

and the southeastern UK (Fig. 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 $N_{W \wedge P}^{hist}$ well approximates this large-scale pattern (Fig. 2b), but especially the lower-resolution CMIP6 models do not capture the small-scale imprints of orography and land seen in ERA5 (Supporting Fig. 4). Consequently, the ensemble mean $N_{W \wedge P}^{hist}$ is smoother and lower than $N_{W_{ERA5} \wedge P}$ in among others Scotland, Iceland, Norway and Italy.
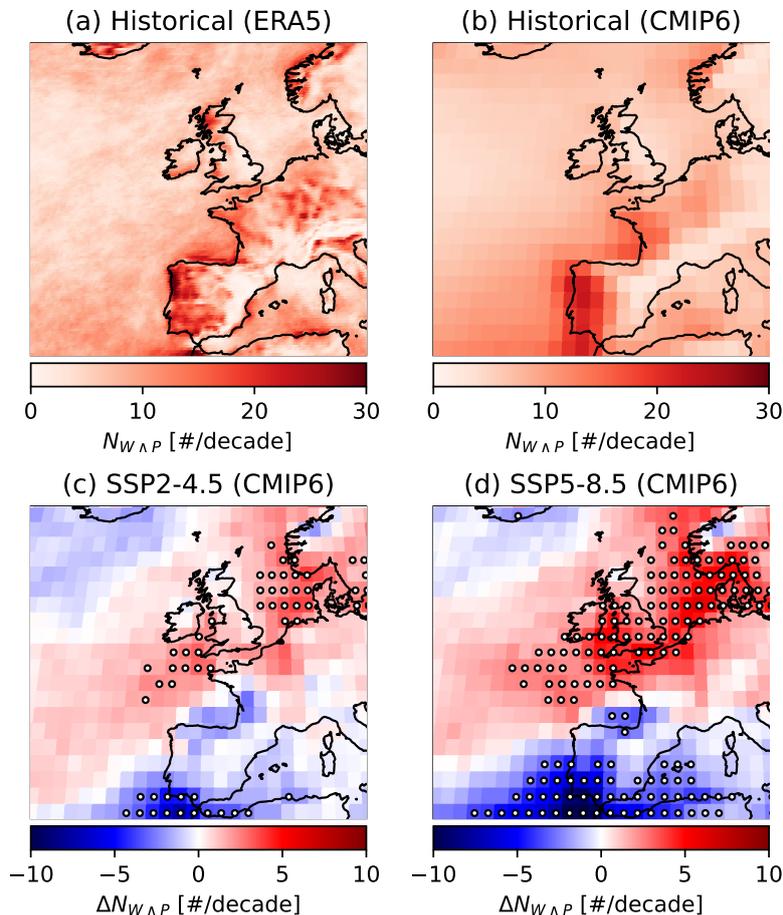
Fig. 2. (a) $N_{W \wedge P}$ based on the ERA5 reanalysis (1979-2018), **(b)** CMIP6 ensemble mean $N_{W \wedge P}^{hist}$, and **(c-d)** CMIP6 ensemble mean $\Delta N_{W \wedge P}$ under SSP2-4.5 and SSP5-8.5, respectively. In (c-d), the stippling indicates where the absolute ensemble mean change exceeds the standard deviation of the change between models.

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, neighboured 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 (Figs. 2c & d). Averaged over land, the absolute magnitude of the changes is approximately 39% (SSP2-4.5) to

310 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 Supporting Fig. 5) is larger under SSP5-8.5 than under SSP2-4.5 (see stippling in Figs. 2c & d). If we only include one randomly selected initial-condition member per model, the standard deviation between models increases by

315 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 (Supporting Figs. 6 & 7). 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

320 ensembles used, but a more systematic comparison would be needed to confirm this.

As shown in Fig. 3, a substantial part of the ensemble mean $\Delta N_{W \wedge P}$ under SSP5-8.5 (and under SSP2-4.5, see Supporting Fig. 8) consists of changes in the marginal distribution of precipitation. The ensemble mean $\Delta N_{W \wedge P}^{p}$ (Fig. 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 (Pfahl et al., 2017; Li et al., 2021; Seneviratne, 2021). The increases in extreme precipitation over most

325 of Europe, which would lead to 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 $\Delta N_{W \wedge P}^{p}$ exceeds the standard deviation between the models.

330 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 (Fig. 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 (Zappa et al., 2013; Priestley and Catto, 2022). In several regions with negative $\Delta N_{W \wedge P}^{w}$, 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 Supporting Fig. 8).

335 Where $\Delta N_{W \wedge P}^{w}$ 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 (Fig. 3c) explain much but not all of the total change ($\Delta N_{W \wedge P}$, Fig. 2d). The remaining differences are caused by $\Delta N_{W \wedge P}^{dependence}$, which is negative mainly over the northwest of the domain, the Bay of Biscay, part of Norway and northern Africa, and positive in most other regions (Fig. 3d).

340 While the magnitude of $\Delta N_{W \wedge P}^{dependence}$ is moderate, it is comparable to or higher than of the other components of $\Delta N_{W \wedge P}$ in several regions. The ensemble mean of $\Delta N_{W \wedge P}^{dependence}$ is lower than its standard deviation between models except south of Norway, in the southern UK and at a few other grid cells (Fig. 3d), indicating that the uncertainty in this term due to model differences and internal climate variability is relatively large. In contrast to $\Delta N_{W \wedge P}^{p}$ and $\Delta N_{W \wedge P}^{w}$, the pattern of $\Delta N_{W \wedge P}^{dependence}$ does not clearly relate to the atmospheric changes projected in previous studies. For instance, the dependence-

345 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 (Zappa et al., 2013; Priestley and Catto, 2022), 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 $\Delta N_{W \wedge P}^{dependence}$ from a

350 seasonal perspective in the next section.
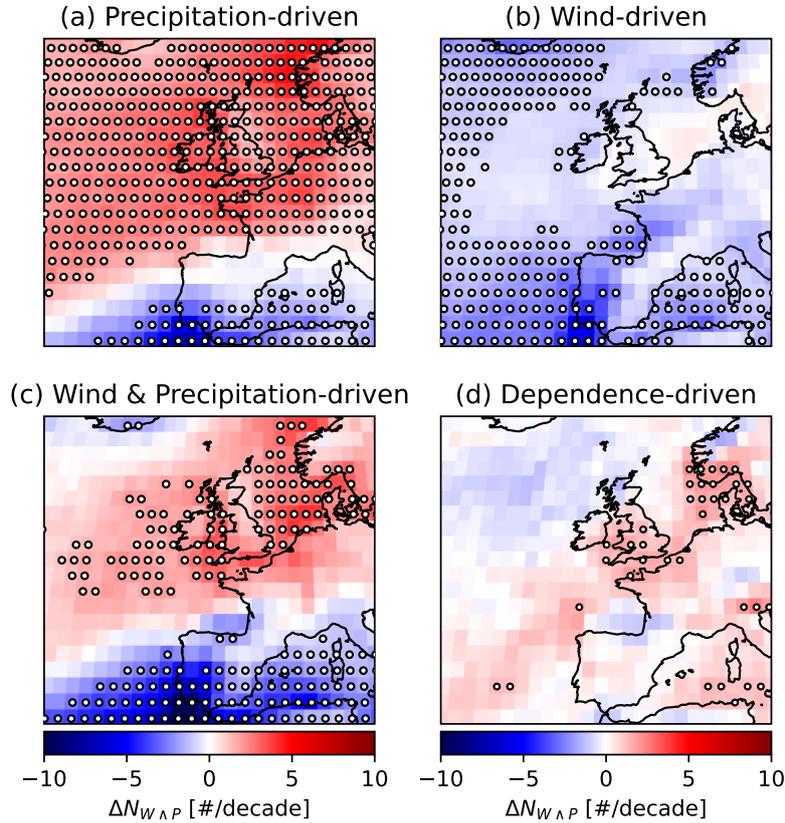
Fig. 3. CMIP6 ensemble mean $\Delta N_{W \wedge P}$ (under SSP5-8.5) due to **(a)** changes in the marginal distribution of precipitation ($\Delta N_{W \wedge P}^{p}$), changes in the marginal distribution of wind speed ($\Delta N_{W \wedge P}^{w}$), changes in the marginal distributions of both precipitation and wind speed ($\Delta N_{W \wedge P}^{w,p}$), and **(d)** changes in the dependence between precipitation and wind speed ($\Delta N_{W \wedge P}^{dependence}$). 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.

4.2 Seasonal dependence-driven changes

Most of the ensemble mean $\Delta N_{W \wedge P}^{dependence}$ (Fig. 3d) consists of changes in autumn (SON) and winter (DJF) (Figs. 4a & d), 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 $\Delta N_{W \wedge P}^{dependence}$ 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 (Figs. 4b, c, e & f). The projected shifts in the number of univariate extremes in autumn and winter seem to explain at least part of the ensemble mean $\Delta N_{W \wedge P}^{dependence}$ in these seasons (Fig. 4). For example, winter $\Delta N_{W \wedge P}^{dependence}$ is negative in the north and northwest of the domain, where also the numbers of precipitation and wind extremes in winter decrease. Similarly, winter $\Delta N_{W \wedge P}^{dependence}$ is positive over the Mediterranean Sea, where both the numbers of precipitation and wind extremes are simulated to increase in winter (Figs. 4a-c). Furthermore, autumn $\Delta N_{W \wedge P}^{dependence}$ is negative over the Bay of Biscay, consistent with the decreasing numbers of both wind and precipitation extremes in that region (Figs. 4d-f).

The changes in the number of univariate extremes in winter and autumn (Fig. 4b, c, e & f) 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 $\Delta N_{W \wedge P}^{dependence}$ is less consistent with the seasonal changes in the number of univariate extremes, such as over the Bay of Biscay in winter (Figs. 4a-c), changes in the frequency of certain weather types may play a larger role.
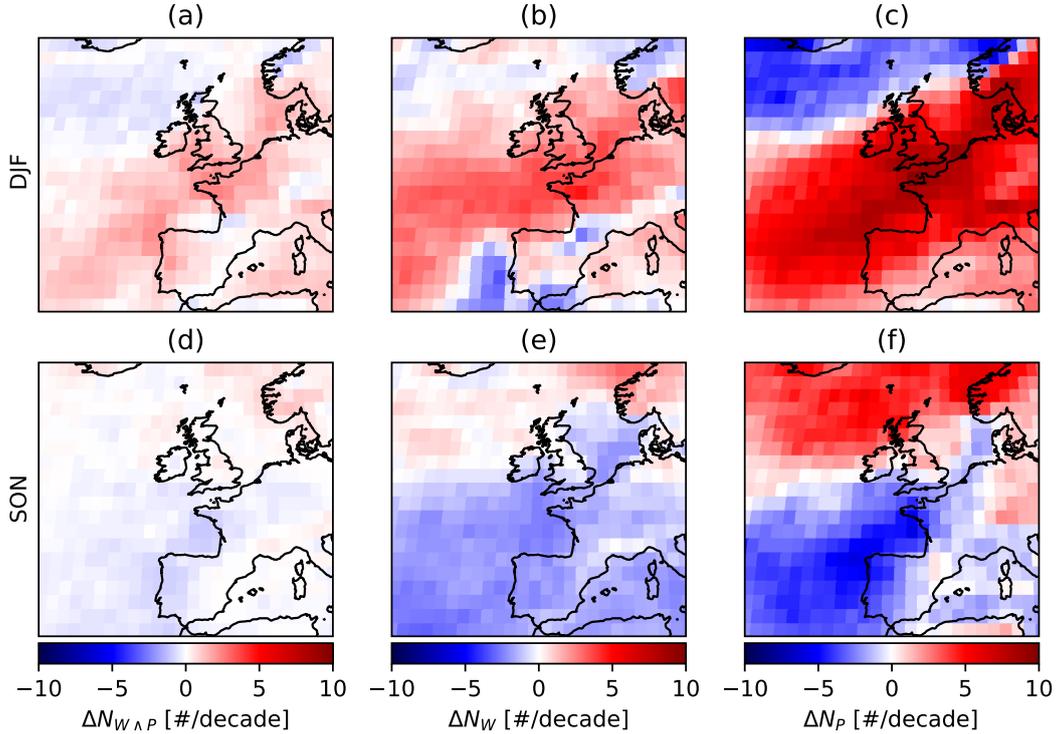
Fig. 4. CMIP6 ensemble mean changes (under SSP5-8.5) in the number of joint wind and precipitation extremes ($\Delta N_{W \wedge P}$), univariate wind speed extremes (ΔN_W) and univariate precipitation extremes (ΔN_P), in **(a-c)** winter (DJF) and **(d-f)** autumn (SON), respectively.

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 $N_{S \wedge P}^{hist}$ (Fig. 5a) agrees with $N_{S_{G2} \wedge P}$ (Fig. 1a) reasonably well (significant pattern correlation of 0.75 and nRMSE of 0.40). Similarly to $N_{S_{G2} \wedge P}$, the ensemble mean $N_{S \wedge P}^{hist}$ 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 $N_{S \wedge P}^{hist}$ tends to underestimate $N_{S_{G2} \wedge P}$ (Supporting Fig. 9a). While these biases may partially be inherited from the MLR model (see Fig. 1f), the ensemble mean $N_{S \wedge P}^{hist}$ also tends to underestimate $N_{S_{MLR} \wedge P}$ in these regions (Supporting Fig. 9b). Furthermore, a similar underestimation was found for an ensemble of CMIP5 models based on hydrodynamically modeled storm surges (Bevacqua et al., 2019, 2020b). 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 $N_{W \wedge P}^{hist}$ (Fig. 2b), $N_{S \wedge P}^{hist}$ 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 (Fig. 5d). This pattern is very similar to that of the differences between $N_{S_{MLR} \wedge P}$ and

18

$N_{W_{ERA5} \wedge P}$ (Figs. 1b & d), 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 $\Delta N_{S \wedge P}$ (Figs. 5b-c) is positive at tide gauges in northwestern Europe and negative at most tide gauges 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 Supporting Figure 5), especially under SSP5-8.5 (grey-edged circles in Figs. 5b-c). 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 tide gauges, 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 $\Delta N_{W \wedge P}$ (Figs. 5e-f). Furthermore, in several locations, the magnitude of the difference between these changes exceeds that of the changes themselves. These differences reflect that winds with extreme speeds and winds that cause extreme storm surges do not necessarily change in the same way, although part of the differences will depend on the storm surge model used. 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 non-linear changes. 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 & 4.2, their decomposition into changes in marginal distributions and dependence-related changes are also broadly similar (Supporting Fig. 13).
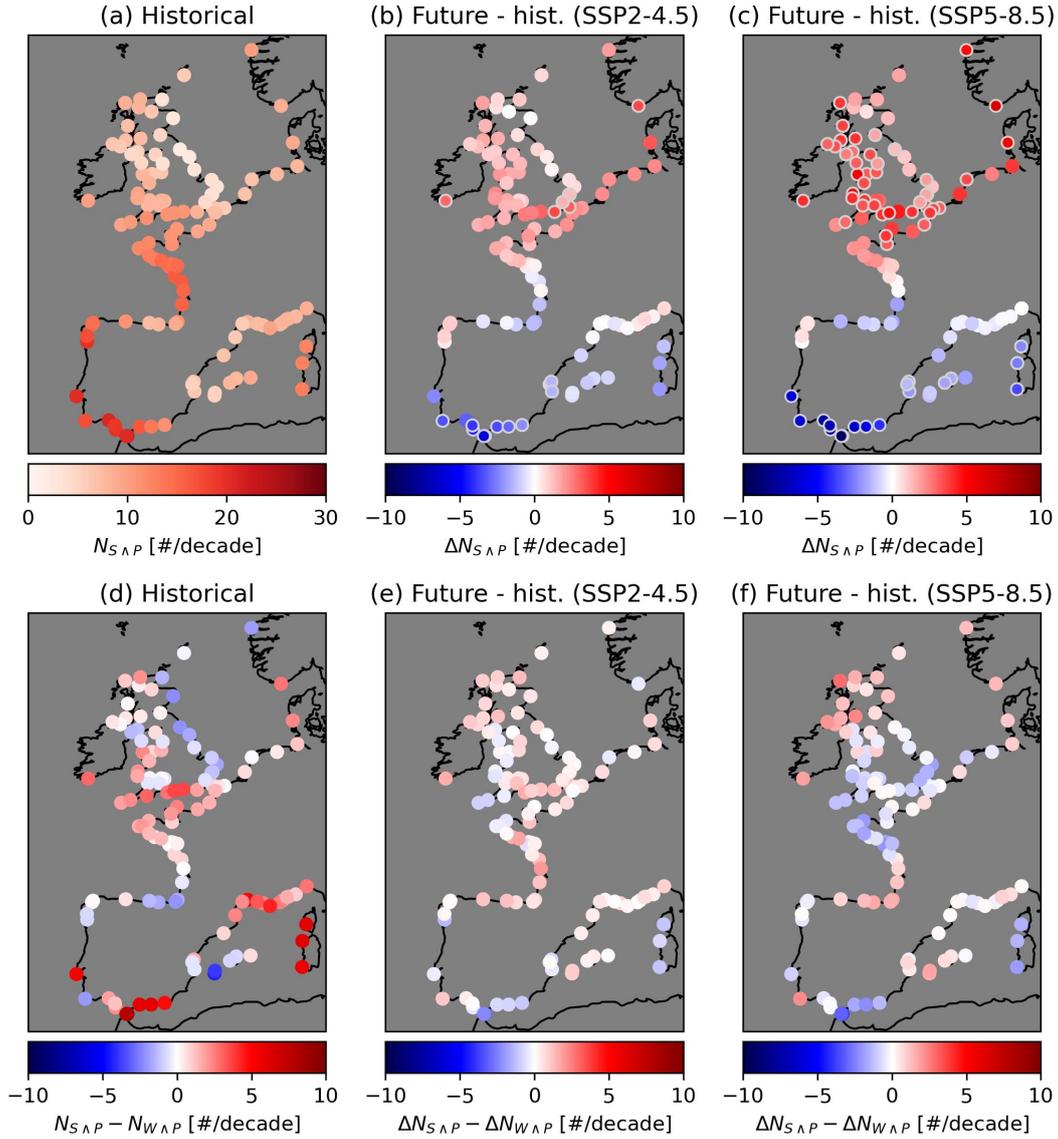
Fig. 5. (a) CMIP6 ensemble mean $N_{S \wedge P}^{hist}$, **(b-c)** CMIP6 ensemble mean $\Delta N_{S \wedge P}$ under SSP2-4.5 and SSP5-8.5, respectively, **(d)** the CMIP6 ensemble mean of $N_{S \wedge P}^{hist}$ minus $N_{W \wedge P}^{hist}$, and **(e-f)** the CMIP6 ensemble mean of $\Delta N_{S \wedge P}$ minus $\Delta N_{W \wedge P}$ under SSP2-4.5 and SSP5-8.5, respectively. In (b-c), circles with a grey edge indicate where the absolute ensemble mean change exceeds the standard deviation of the change between models.

4.4 Magnitude relative to internal climate variability

Figs. 6a & b 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 $N_{W \wedge P}^{hist}$ and $N_{S \wedge P}^{hist}$ due to internal climate variability (estimated as explained in Section 2.4). In other words, most of the future changes in the number of joint extremes projected under SSP2-4.5 are smaller than deviations that are likely to be seen due to internal climate variability alone. This does not necessarily imply that the ensemble mean is caused by internal climate variability, but rather that the average joint probability onto which internal climate variability is superimposed will change. Over the eastern North Sea and south of Spain, unforced deviations in the number of joint extremes are less likely to be larger than the ensemble mean changes (Figs. 6a & b).

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 (Figs. 6c & d). 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.
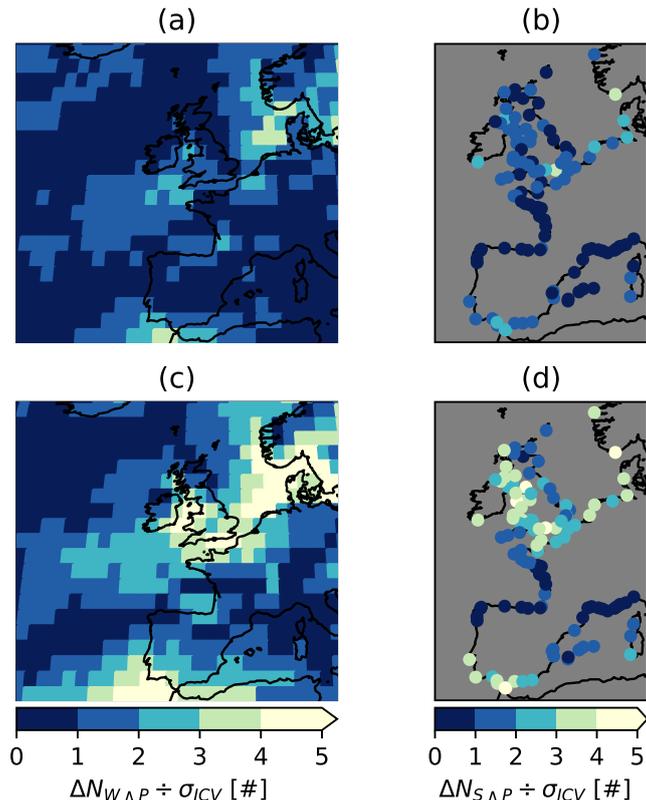
Fig. 6. (a) CMIP6 ensemble mean $\Delta N_{W \wedge P}$ under SSP2-4.5 divided by the average standard deviation of $N_{W \wedge P}^{hist}$ between initial-condition members of CMIP6 models providing at least 5 members, **(b)** CMIP6 ensemble mean $\Delta N_{S \wedge P}$ under SSP2-4.5 divided by the average standard deviation of $N_{S \wedge P}^{hist}$ between initial-condition members of CMIP6 models providing at least 5 members, **(c-d)** as in (a-b), under SSP5-8.5.

5 Uncertainty in the projections and sensitivity to ensemble size

The projections of $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ are affected by uncertainties related to internal climate variability, inter-model differences and differences between emissions scenarios. Based on the CMIP6 models providing at least 5 initial-condition simulations (see Section 2.4), we find that the uncertainties due to inter-model differences and internal climate variability exceed the uncertainty due to differences between SSP2-4.5 and SSP5-8.5 in most of the domain (see Supporting Fig. 14). The emissions scenario uncertainty would have likely been larger if we would have also included SSP1-2.6. Whether the uncertainties due to internal climate variability or inter-model differences are largest varies by location, but the former is largest in more regions under SSP2-4.5 than under SSP5-8.5 (Supporting Fig. 14). To test whether the uncertainty partitioning is representative for the entire CMIP6 ensemble, more models providing multiple initial-condition simulations are needed.

Because the uncertainties due to internal climate variability and inter-model differences are large relative to the ensemble mean changes (compare Figs. 2 & 5 with Supporting Fig. 14), projections of changes in the joint probability of extremes based on only 5 to 6 climate model simulations (e.g., Ganguli et al., 2020; Bevacqua et al., 2020b) may change qualitatively 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. In the northeast of the domain, in the band extending from southwest of the UK to southern Scandinavia, 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 models (Fig. 7). In these regions, the ensemble mean $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ are relatively large and tend to exceed the standard deviation between models (see Figs. 2 & 5).

In contrast, in part of the UK and 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 higher than 10%. In most places, we find that the later in the 21st century and the higher the emissions scenario, the smaller is the ensemble size required for projections of which the sign is insensitive to which CMIP6 models are included (compare the top row with the bottom row and the left two columns with the right two columns of Fig. 7, respectively), as the forced response is larger. For instance, for qualitatively 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 (Fig. 7d), whereas 5 or fewer models suffice under SSP5-8.5 (Fig. 7h). However, even under SSP5-8.5, large ensembles are needed for qualitatively robust projections of $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ in the Bay of Biscay and along the northern coastline of the Mediterranean Sea (Fig. 7h), consistent with the relatively large internal climate variability in these regions (see Fig. 6).

While Fig. 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 two 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 (Supporting Figs. 15 & 16). 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 (Supporting Figs. 15 & 16), especially when compared to differences in $\Delta N_{W \wedge P}$ and $\Delta N_{S \wedge P}$ between models and initial-condition members.
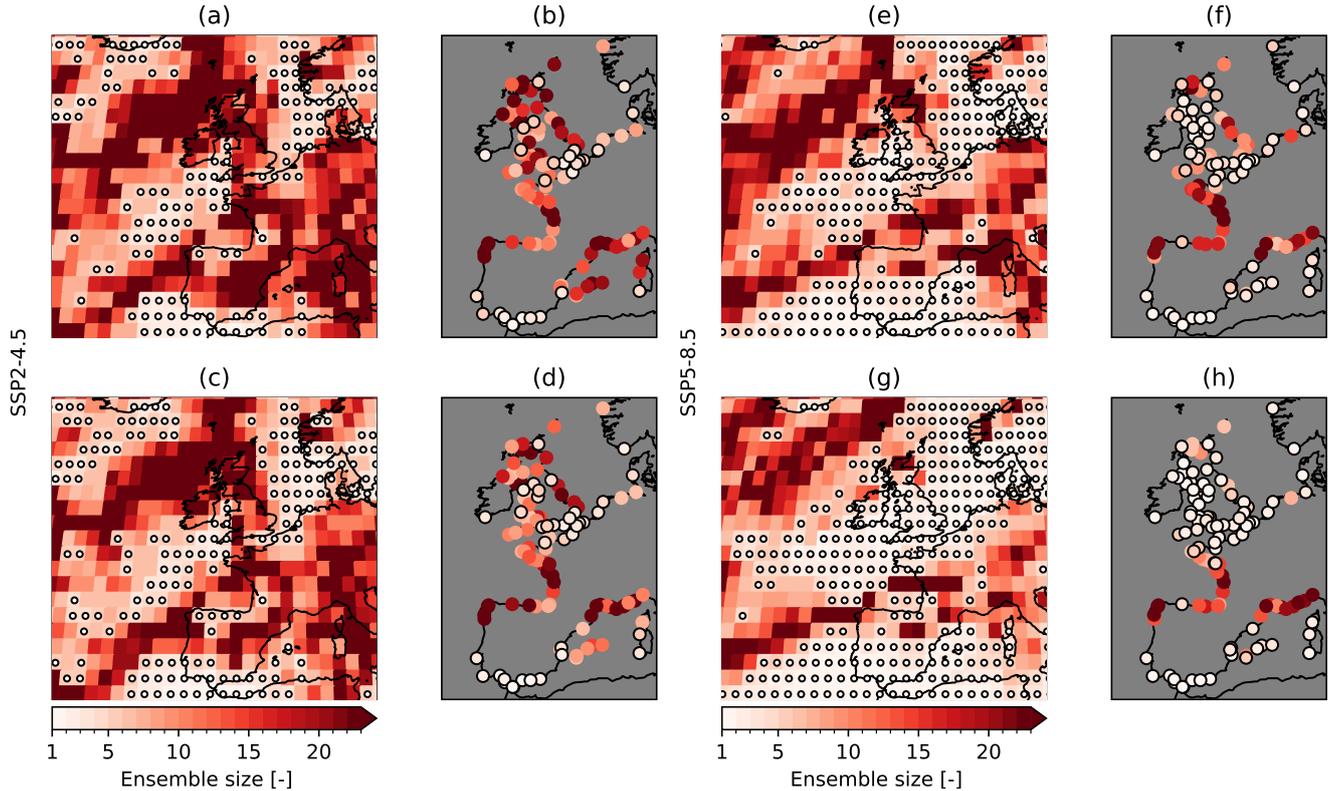
Fig. 7. The minimum number of random CMIP6 models that an ensemble needs to consist of for its mean **(a-b)** $\Delta N_{W \wedge P}$ and **(c-d)** $\Delta N_{S \wedge P}$ under SSP2-4.5 to have a 90% or higher probability of having the same sign as the projections in Section 4, for the periods 2041-2080 and 2061-2100, 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.

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 et al., 2020b; 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 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% v.s. 36%, on average). Previous studies for Europe only included a high emissions scenario (Bevacqua et al., 2020b; 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., Zappa et al., 2013; Priestley and Catto, 2022; Pfahl et al., 2017), as also concluded by Bevacqua et al. (2020b). 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 et al. (2020b) 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 et al. (2020b). We conclude that especially under not so high emissions scenarios and earlier in the 21st 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. (2020a) 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 (Calafat et al., 2022; Vousdoukas et al., 2017, 2018; Bevacqua et al., 2020a, & our results). We conclude that statistically modelled 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 datasets efficiently.

By applying the statistical storm surge model to CMIP6 simulations, we implicitly assume that the observation-based relationships between predictors and predictands will not change, which is not necessarily true. For instance, the effect of sea-level

change on storm surges, which can be simulated with hydrodynamic models (e.g., Muis et al., 2020), cannot easily be included in a statistical model trained with observations. It would be useful to explore if this could be overcome by training statistical models with hydrodynamic model simulations of future changes. Furthermore, to compute ensemble statistics we weighted each CMIP6 model equally (Section 2.4). However, the skill of the models in simulating the number of joint extremes varies (Ridder et al., 2021; Wu et al., 2021). Assigning different weights to the CMIP6 models based on their skill may reduce the uncertainty due to inter-model differences (Ridder et al., 2022). As we have shown, though, the historical number of joint extremes on which such skill scores are based is also affected by internal climate variability (Section 4.4). Therefore, skill scores may need to be computed using multiple initial-condition simulations per model, which has not yet been done.

Like previous studies (e.g., Bevacqua et al., 2020b; Ganguli et al., 2020; Moftakhari et al., 2017; Gori et al., 2022), we have only considered changes in the joint probability of drivers of compound flooding, which do not necessarily imply changes in the risk of compound flooding. To project changes in flood risk, more information such as on local socioecomic activity, land elevation and protective measures would need to be incorporated. Sea-level rise, which we excluded here, can also increase the frequency of (compound) flooding if coastal flood protection is not adapted accordingly (Hermans et al., 2023). While our projections may therefore 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.

References

535   Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., Burke, E., Dittus, A. J., Florek, P., Gray, L. J., Haddad, S., Hardiman, S. C., Hermanson, L., Hodson, D., Hogan, E., Jones, G. S., Knight, J. R., Kuhlbrodt, T., Misios, S., Mizielinski, M. S., Ringer, M. A., Robson, J., and Sutton, R. T.: Historical Simulations With HadGEM3-GC3.1 for CMIP6, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001995, 2020.

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

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

Bevacqua, E., Maraun, D., Vousdoukas, M. I., Voukouvalas, E., Vrac, M., Mentaschi, L., and Widmann, M.: Higher probability of
545   compound flooding from precipitation and storm surge in Europe under anthropogenic climate change, Science Advances, 5, http://advances.sciencemag.org/, 2019.

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

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

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

555
Bi, D., Dix, M., Marsland, S., O'farrell, S., Sullivan, A., Bodman, R., Law, R., Harman, I., Srbinovsky, J., Rashid, H. A., Dobrohotoff, P., Mackallah, C., Yan, H., Hirst, A., Savita, A., Dias, F. B., Woodhouse, M., Fiedler, R., and Heerdegen, A.: 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, 225–251, https://doi.org/10.1071/ES19040, 2020.

560   Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J. L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M. A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J. Y., Guenet, B., Lionel, E. G., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott,
565   F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J. B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002010, 2020.

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

28

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

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

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

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

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A., and Navarra, A.: Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model, Journal of Advances in Modeling Earth 580 Systems, 11, 185–209, https://doi.org/10.1029/2018MS001369, 2019.

Couasnon, A., Eilander, D., Muis, S., Veldkamp, T. I. E., Haigh, I. D., Wahl, T., Winsemius, H., and Ward, P. J.: 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, 2019.

Couasnon, A., Scussolini, P., Tran, T. V., Eilander, D., Muis, S., Wang, H., Keesom, J., Dullaart, J., Xuan, Y., Nguyen, H. Q., Winsemius, 585 H. C., and Ward, P. J.: 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, https://doi.org/10.1029/2021WR030002, 2022.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, 590 M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001916, 2020.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P., 595 Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: 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, https://doi.org/10.1029/2019MS002015, 2020.

600 Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L. P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., Hardenberg, J. V., Hieronymus, J., Karami, M. P., Keskinen, J. P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., Noije, T. V., Nolan, P., O'donnell, D., Ollinaho, P., Oord, G. V. D., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Sager, P. L., Schmith, T., Schrödner, R., Serva, F., 605 Sicardi, V., Madsen, M. S., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geoscientific Model Development, 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

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

610 Emanuel, K.: 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, 12 681–12 684, https://doi.org/10.1073/pnas.1716222114, 2017.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., Taylor, K. E., Dynamique, D. M., Pierre, I., Laplace, S., and Ipsl, L. M. D.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

615 Ganguli, P. and Merz, B.: Extreme Coastal Water Levels Exacerbate Fluvial Flood Hazards in Northwestern Europe, Scientific Reports, 9, https://doi.org/10.1038/s41598-019-49822-6, 2019.

Ganguli, P., Paprotny, D., Hasan, M., Güntner, A., and Merz, B.: Projected Changes in Compound Flood Hazard From Riverine and Coastal Floods in Northwestern Europe, Earth's Future, 8, https://doi.org/10.1029/2020EF001752, 2020.

Ghanbari, M., Arabi, M., Obeysekera, J., and Sweet, W.: A Coherent Statistical Model for Coastal Flood Frequency Analysis Under Nonsta-
620 tionary Sea Level Conditions, Earth's Future, 7, 162–177, https://doi.org/10.1029/2018EF001089, 2019.

Gori, A., Lin, N., Xi, D., and Emanuel, K.: Tropical cyclone climatology change greatly exacerbates US extreme rainfall–surge hazard, Nature Climate Change, 12, 171–178, https://doi.org/10.1038/s41558-021-01272-7, 2022.

Haigh, I. D., Wadey, M. P., Wahl, T., Ozsoy, O., Nicholls, R. J., Brown, J. M., Horsburgh, K., and Gouldby, B.: 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, 2016.

625 Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks, Geoscientific Model Development, 13, 2197–2244, https://doi.org/10.5194/gmd-13-2197-2020, 2020.

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., Shevliakova, E., Winton, M., Zhao, M., Bushuk, M., Wittenberg,
630 A. T., Wyman, B., Xiang, B., Zhang, R., Anderson, W., Balaji, V., Donner, L., Dunne, K., Durachta, J., Gauthier, P. P., Ginoux, P., Golaz, J. C., Griffies, S. M., Hallberg, R., Harris, L., Harrison, M., Hurlin, W., John, J., Lin, P., Lin, S. J., Malyshev, S., Menzel, R., Milly, P. C., Ming, Y., Naik, V., Paynter, D., Paulot, F., Rammaswamy, V., Reichl, B., Robinson, T., Rosati, A., Seman, C., Silvers, L. G., Underwood, S., and Zadeh, N.: 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, 2019.

635 Hendry, A., Haigh, I. D., Nicholls, R. J., Winter, H., Neal, R., Wahl, T., Joly-Lauge, A., and Darby, S. E.: Assessing the charac-
teristics and drivers of compound flooding events around the UK coast, Hydrology and Earth System Sciences, 23, 3117–3139, https://doi.org/10.5194/hess-23-3117-2019, 2019.

Hermans, T., Malagon-Santos, V., Jane, R., Garner, G., Kopp, R., and Slangen, A.: Projections of the Timing of Decreasing Coastal Flood Protection, https://doi.org/10.5281/zenodo.7505441, 2023.

640 Hermans, T. H. and Busecke, J. J. M.: CMIP6cex (v0.1.1), https://doi./TBD, TBD.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Sim-
mons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Vil-

laume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hoyer, S. and Hamman, J.: xarray: N-D labeled arrays and datasets in Python, Journal of Open Research Software, 5, https://doi.org/10.5334/jors.148, 2017.

Jaroszweski, D., Hooper, E., Baker, C., Chapman, L., and Quinn, A.: The impacts of the 28 June 2012 storms on UK road and rail transport, Meteorological Applications, 22, 470–476, https://doi.org/10.1002/met.1477, 2015.

Kew, S. F., Selten, F. M., Lenderink, G., and Hazeleger, W.: The simultaneous occurrence of surge and discharge extremes for the Rhine delta, Natural Hazards and Earth System Sciences, 13, 2017–2029, https://doi.org/10.5194/nhess-13-2017-2013, 2013.

Kumbier, K., Carvalho, R. C., Vafeidis, A. T., and Woodroffe, C. D.: Investigating compound flooding in an estuary using hydrodynamic modelling: A case study from the Shoalhaven River, Australia, Natural Hazards and Earth System Sciences, 18, 463–477, https://doi.org/10.5194/nhess-18-463-2018, 2018.

Lambert, E., Rohmer, J., Cozannet, G. L., and van de Wal, R. S. W.: Adaptation time to magnified flood hazards underestimated when derived from tide gauge records, Environmental Research Letters, 15, 074 015, https://doi.org/10.1088/1748-9326/ab8336, 2020.

Lee, J., Kim, J., Sun, M. A., Kim, B. H., Moon, H., Sung, H. M., Kim, J., and Byun, Y. H.: Evaluation of the Korea Meteorological Administration Advanced Community Earth-System model (K-ACE), Asia-Pacific Journal of Atmospheric Sciences, 56, 381–395, https://doi.org/10.1007/s13143-019-00144-7, 2020.

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., Knutti, R., and Hawkins, E.: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6, Earth System Dynamics, 11, 491–508, https://doi.org/10.5194/esd-11-491-2020, 2020.

Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., Mcinnes, K., Risbey, J., Schuster, S., Jakob, D., and Stafford-Smith, M.: A compound event framework for understanding extreme impacts, https://doi.org/10.1002/wcc.252, 2014.

Li, C., Zwiers, F., Zhang, X., Li, G., Sun, Y., and Wehner, M.: Changes in Annual Extremes of Daily Temperature and Precipitation in CMIP6 Models, Journal of Climate, pp. 3411–3460, https://doi.org/10.1175/JCLI-D-19, 2021.

Li, L., Yu, Y., Tang, Y., Lin, P., Xie, J., Song, M., Dong, L., Zhou, T., Liu, L., Wang, L., Pu, Y., Chen, X., Chen, L., Xie, Z., Liu, H., Zhang, L., Huang, X., Feng, T., Zheng, W., Xia, K., Liu, H., Liu, J., Wang, Y., Wang, L., Jia, B., Xie, F., Wang, B., Zhao, S., Yu, Z., Zhao, B., and Wei, J.: 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, https://doi.org/10.1029/2019MS002012, 2020.

Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., Fogli, P. G., Cherchi, A., Bellucci, A., Gualdi, S., Masina, S., and Navarra, A.: CMIP6 Simulations With the CMCC Earth System Model (CMCC-ESM2), Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2021MS002814, 2022.

Martius, O., Pfahl, S., and Chevalier, C.: A global quantification of compound precipitation and wind extremes, Geophysical Research Letters, 43, 7709–7717, https://doi.org/10.1002/2016GL070017, 2016.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., de-la Cuesta, D. J., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E., Nam, C. C., Notz, D., Nyawira, S. S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J. S., Tian,

F., Voigt, A., Vrese, P., Wieners, K. H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

685

Meinshausen, M., Nicholls, Z., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P., Reimann, S., Smith, S., den Berg, M. V., Velders, G., Vollmer, M., and Wang, R.: The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, GMD, 13, 3571–3605, https://doi.org/10.5194/gmd-13-3571-2020, 2020.

690

Moftakhari, H. R., Salvadori, G., AghaKouchak, A., Sanders, B. F., and Matthew, R. A.: Compounding effects of sea level rise and fluvial flooding, Proceedings of the National Academy of Sciences, 114, 9785–9790, https://doi.org/10.1073/pnas.1620325114, 2017.

Muis, S., Verlaan, M., Winsemius, H. C., Aers, J. C. J. H., and Ward, P. J.: A global reanalysis of storm surges and extreme sea levels, Nature Communications, 7, https://doi.org/10.1038/ncomms11969, 2016.

Muis, S., Apecechea, M. I., Dullaart, J., de Lima Rego, J., Madsen, K. S., Su, J., Yan, K., and Verlaan, M.: A High-Resolution

695

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, 2020.

Muis, S., Aerts, J. C., José, J. A., Dullaart, J. C., Duong, T. M., Erikson, L., Haarsma, R. J., Apecechea, M. I., Mengel, M., Bars, D. L., O'Neill, A., Ranasinghe, R., Roberts, M. J., Verlaan, M., Ward, P. J., and Yan, K.: Global Projections of Storm Surges Using High-Resolution CMIP6 Climate Models, Earth's Future, 11, https://doi.org/10.1029/2023EF003479, 2023.

700

Nasr, A. A., Wahl, T., Rashid, M. M., Camus, P., and Haigh, I. D.: Assessing the dependence structure between oceanographic, fluvial, and pluvial flooding drivers along the United States coastline, Hydrology and Earth System Sciences, 25, 6203–6222, https://doi.org/10.5194/hess-25-6203-2021, 2021.

Owen, L. E., Catto, J. L., Stephenson, D. B., and Dunstone, N. J.: Compound precipitation and wind extremes over Europe and their relationship to extratropical cyclones, Weather and Climate Extremes, 33, https://doi.org/10.1016/j.wace.2021.100342, 2021.

705

Paprotny, D., Morales-Nápoles, O., and Jonkman, S. N.: HANZE: A pan-European database of exposure to natural hazards and damaging historical floods since 1870, Earth System Science Data, 10, 565–581, https://doi.org/10.5194/essd-10-565-2018, 2018.

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

Pfahl, S., O'Gorman, P., and Fischer, E.: Understanding the regional pattern of projected future changes in extreme precipitation, Nature

710

Climate Change, 7, https://doi.org/http://dx.doi.org/10.1038/nclimate3287, 2017.

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

Ridder, N. N., Pitman, A. J., Westra, S., Ukkola, A., Hong, X. D., Bador, M., Hirsch, A. L., Evans, J. P., Luca, A. D., and Zscheischler, J.: Global hotspots for the occurrence of compound events, Nature Communications, 11, https://doi.org/10.1038/s41467-020-19639-3, 2020.

715

Ridder, N. N., Pitman, A. J., and Ukkola, A. M.: Do CMIP6 Climate Models Simulate Global or Regional Compound Events Skillfully?, https://doi.org/10.1029/2020GL091152, 2021.

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

Ruocco, A. C., Nicholls, R. J., Haigh, I. D., and Wadey, M. P.: Reconstructing coastal flood occurrence combining sea level and media

720

sources: A case study of the Solent, UK since 1935, Natural Hazards, 59, 1773–1796, https://doi.org/10.1007/s11069-011-9868-7, 2011.

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

Santos, V. M., Casas-Prat, M., Poschlod, B., Ragno, E., Hurk, B. V. D., Hao, Z., Kalmár, T., Zhu, L., and Najafi, H.: 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, 3595–3615, https://doi.org/10.5194/hess-25-3595-2021, 2021.

Sellar, A. A., Walton, J., Jones, C. G., Wood, R., Abraham, N. L., Andrejczuk, M., Andrews, M. B., Andrews, T., Archibald, A. T., de Mora, L., Dyson, H., Elkington, M., Ellis, R., Florek, P., Good, P., Gohar, L., Haddad, S., Hardiman, S. C., Hogan, E., Iwi, A., Jones, C. D., Johnson, B., Kelley, D. I., Kettleborough, J., Knight, J. R., Köhler, M. O., Kuhlbrodt, T., Liddicoat, S., Linova-Pavlova, I., Mizielinski, M. S., Morgenstern, O., Mulcahy, J., Neininger, E., O'Connor, F. M., Petrie, R., Ridley, J., Rioual, J. C., Roberts, M., Robertson, E., Rumbold, S., Seddon, J., Shepherd, H., Shim, S., Stephens, A., Teixiera, J. C., Tang, Y., Williams, J., Wiltshire, A., and Griffiths, P. T.: Implementation of U.K. Earth System Models for CMIP6, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001946, 2020.

Seneviratne: Weather and Climate Extreme Events in a Changing Climate, pp. 1513–1766, Cambridge University Press, https://doi.org/10.1017/9781009157896.013, 2021.

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

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

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., Salzen, K. V., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geoscientific Model Development, 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Tadesse, M., Wahl, T., and Cid, A.: Data-Driven Modeling of Global Storm Surges, Frontiers in Marine Science, 7, https://doi.org/10.3389/fmars.2020.00260, 2020.

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

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'Ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geoscientific Model Development, 12, 2727–2765, https://doi.org/10.5194/gmd-12-2727-2019, 2019.

Taylor, K., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bull Am Meteorol Soc, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

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

van den Hurk, B., van Meijgaard, E., de Valk, P., van Heeringen, K.-J., and Gooijer, J.: Analysis of a compounding surge and precipitation event in the Netherlands, Environmental Research Letters, 10, 35 001, 2015.

Volodin, E. M. and Gritsun, A. S.: Simulation of Possible Future Climate Changes in the 21st Century in the INM-CM5 Climate Model, Earth System Dynamics, 9, 1235–1242, https://doi.org/10.1134/S0001433820030123, 2018.

Volodin, E. M., Mortikov, E. V., Kostrykin, S. V., Galin, V. Y., Lykossov, V. N., Gritsun, A. S., Diansky, N. A., Gusev, A. V., and Iakovlev, N. G.: Simulation of the present-day climate with the climate model INMCM5, Climate Dynamics, 49, 3715–3734, https://doi.org/10.1007/s00382-017-3539-7, 2017.

Vousdoukas, M. I., Mentaschi, L., Feyen, L., and Voukouvalas, E.: 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, 2017.

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

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

Wang, Y. C., Hsu, H. H., Chen, C. A., Tseng, W. L., Hsu, P. C., Lin, C. W., Chen, Y. L., Jiang, L. C., Lee, Y. C., Liang, H. C., Chang, W. M., Lee, W. L., and Shiu, C. J.: 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, https://doi.org/10.1029/2020MS002353, 2021.

Ward, P. J., Couasnon, A., Eilander, D., Haigh, I. D., Hendry, A., Muis, S., Veldkamp, T. I. E., Winsemius, H. C., and Wahl, T.: Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries, Environmental Research Letters, 13, 084 012, https://doi.org/10.1088/1748-9326/aad400, 2018.

Woodworth, P. L., Hunter, J. R., Marcos, M., Caldwell, P., Menéndez, M., and Haigh, I.: Towards a global higher-frequency sea level dataset, Geoscience Data Journal, 3, 50–59, 2016.

Wu, W., McInnes, K., O'Grady, J., Hoeke, R., Leonard, M., and Westra, S.: Mapping Dependence Between Extreme Rainfall and Storm Surge, Journal of Geophysical Research: Oceans, 123, 2461–2474, https://doi.org/10.1002/2017JC013472, 2018.

Wu, Y., Miao, C., Sun, Y., AghaKouchak, A., Shen, C., and Fan, X.: Global Observations and CMIP6 Simulations of Compound Extremes of Monthly Temperature and Precipitation, GeoHealth, 5, https://doi.org/10.1029/2021GH000390, 2021.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M.: 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, 931–965, https://doi.org/10.2151/jmsj.2019-051, 2019.

Zappa, G., Shaffrey, L. C., Hodges, K. I., Sansom, P. G., and Stephenson, D. B.: A Multimodel Assessment of Future Projections of North Atlantic and European Extratropical Cyclones in the CMIP5 Climate Models*, Journal of Climate, 26, 5846–5862, https://doi.org/10.1175/JCLI-D, 2013.

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

Zhuang, J., Dussin, R., Huard, D., Bourgault, P., Banihirwe, A., Raynaud, S., Malevich, B., Schupfner, M., Filipe, and Levang, S.: xESMF: v0.8.2, https://doi.org/10.5281/zenodo.8356796, 2023.

Zscheischler, J., Westra, S., Hurk, B. J. V. D., Seneviratne, S. I., Ward, P. J., Pitman, A., Aghakouchak, A., Bresch, D. N., Leonard, M., Wahl, T., and Zhang, X.: Future climate risk from compound events, Nature Climate Change, 8, 469–477, https://doi.org/10.1038/s41558-018-0156-3, 2018.

Øyvind Seland, Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y. C., Kirkeväg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geoscientific Model Development, 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

800