Earth's Future

RESEARCH ARTICLE
10.1029/2020EF001752

Projected Changes in Compound Flood Hazard From Riverine and Coastal Floods in Northwestern Europe

Poulomi Ganguli[1,2], Dominik Paprotny[2], Mehedi Hasan[2], Andreas Güntner[2,3], and Bruno Merz[2,3]

[1]Agricultural and Food Engineering Department, Indian Institute of Technology Kharagpur, Kharagpur, India, [2]German Research Centre for Geosciences (GFZ), Potsdam, Germany, [3]Institute of Environmental Sciences and Geography, University of Potsdam, Potsdam, Germany

Key Points:
- We combine high-resolution projected storm surges with localized sea level rise and projected river floods to assess compound flood hazard in the RCP8.5 scenario
- We find decreasing compound flood hazard for 66% of the locations across northwestern Europe
- Models reproduce upper tail dependence between storm surge and river floods, yet discrepancies exist in capturing dependence strength

Supporting Information:
- Supporting Information S1

Correspondence to:
P. Ganguli,
pganguli@agfe.iitkgp.ac.in

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

Received 13 AUG 2020
Accepted 28 OCT 2020
Accepted article online 5 NOV 2020

Author Contributions:
Conceptualization: Bruno Merz
Data curation: Dominik Paprotny, Mehedi Hasan, Andreas Güntner
Methodology: Dominik Paprotny, Mehedi Hasan, Andreas Güntner, Bruno Merz
Project administration: Bruno Merz
Resources: Dominik Paprotny, Mehedi Hasan, Andreas Güntner, Bruno Merz
Software: Dominik Paprotny, Mehedi Hasan, Andreas Güntner
Supervision: Bruno Merz
(continued)

Abstract Compound flooding in coastal regions, that is, the simultaneous or successive occurrence of high sea levels and high river flows, is expected to increase in a warmer world. To date, however, there is no robust evidence on projected changes in compound flooding for northwestern Europe. We combine projected storm surges and river floods with probabilistic, localized relative sea-level rise (SLR) scenarios to assess the future compound flood hazard over northwestern coastal Europe in the high (RCP8.5) emission scenario. We use high-resolution, dynamically downscaled regional climate models (RCM) to drive a storm surge model and a hydrological model, and analyze the joint occurrence of high coastal water levels and associated river peaks in a multivariate copula-based approach. The RCM-forced multimodel mean reasonably represents the observed spatial pattern of the dependence strength between annual maxima surge and peak river discharge, although substantial discrepancies exist between observed and simulated dependence strength. All models overestimate the dependence strength, possibly due to limitations in model parameterizations. This bias affects compound flood hazard estimates and requires further investigation. While our results suggest decreasing compound flood hazard over the majority of sites by 2050s (2040–2069) compared to the reference period (1985–2005), an increase in projected compound flood hazard is limited to around 34% of the sites. Further, we show the substantial role of SLR, a driver of compound floods, which has frequently been neglected. Our findings highlight the need to be aware of the limitations of the current generation of Earth system models in simulating coastal compound floods.

1. Introduction

Compound flooding in coastal regions, that is, the simultaneous or successive occurrence of high sea levels and high river flows, is a significant hazard. By 2100, globally over 600 million people residing on coasts could be exposed to compound floods under RCP8.5 (Kulp & Strauss, 2019). Coastal compound flooding may be a consequence of two mechanisms: First, extreme coastal water levels may affect river flows by backwater effects or by reversing the seaward flow of rivers (Zhang et al., 2019). Rivers located in low-lying (below 10 m above mean sea level) regions of northwestern Europe (Figure 1 in Hoitink & Jay, 2016) are likely to be impacted by this mechanism. Sea-level rise (SLR) increases the water levels in the mouth of rivers (Nicholls et al., 2011), thus increasing the probability of extreme coastal water levels. Second, severe storm episodes may lead to extreme storm surges, and at the same time, to heavy precipitation and high river runoff (Kew et al., 2013). Hence, the main drivers of coastal compound flooding are storm surges, SLR and river floods, which in turn, are driven by a range of processes in the atmosphere and oceans, and on the land surface. Quantifying compound flood hazard under climate change poses a particular challenge at the intersection of climate dynamics and statistical analysis, as the estimation of event probabilities is not only affected by climate projection uncertainties in relation to the different drivers but also in relation to the interdependencies between them (Zscheischler & Seneviratne, 2017).

A number of studies, at local to continental scales across Europe (Bevacqua et al., 2019; Ganguli & Merz, 2019a, 2019b; Kew et al., 2013; Klerk et al., 2015; Petroliagkis, 2018; Svensson & Jones, 2002), the United States (Moftakhari et al., 2017, 2019), Australia (Zheng et al., 2014), China (Lian et al., 2013; Tu et al., 2018), India (Mohanty et al., 2020), Bangladesh (Ikeuchi et al., 2017), and the entire globe (Couasnon et al., 2020; Eilander et al., 2020; Ward et al., 2018) have analyzed compound effects of high coastal water levels (or storm surges and/or tides) and river discharges (or extreme precipitation as proxy

Visualization: Dominik Paprotny, Andreas Güntner, Bruno Merz
Writing – review & editing: Dominik Paprotny, Mehedi Hasan, Andreas Güntner, Bruno Merz

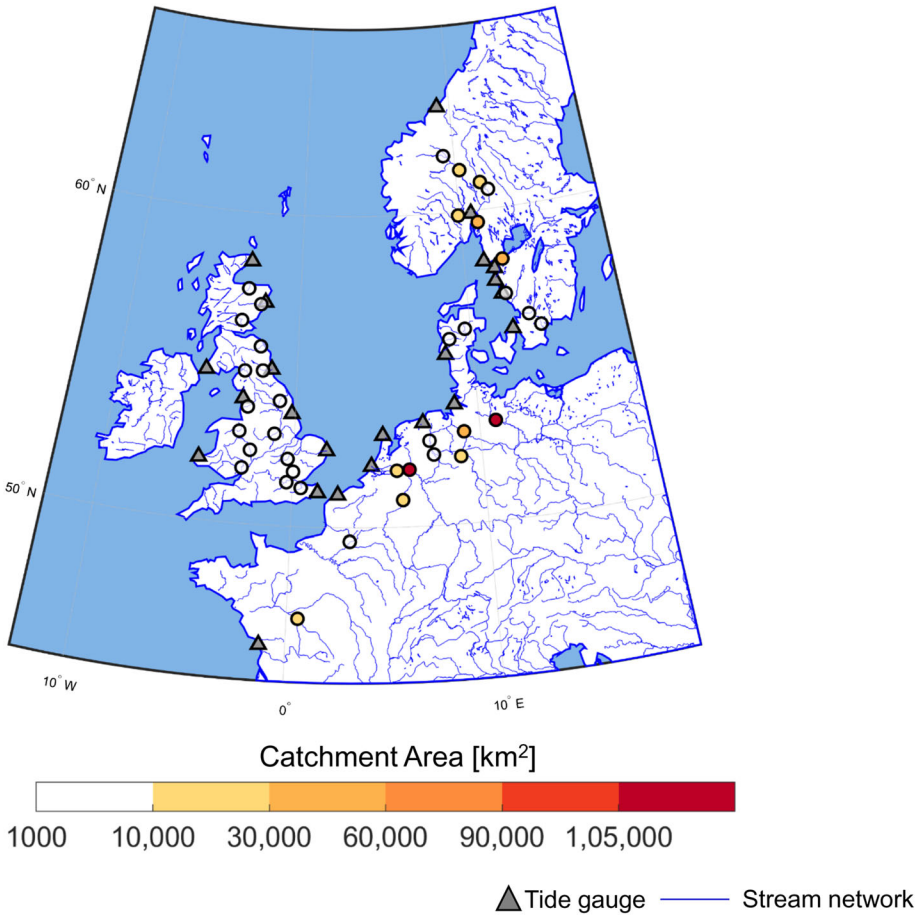

Figure 1. Map of tide (gray triangles) and stream gauge (colored circles) locations. The shades of the circles are proportional to catchment area (in km^2) with darker (lighter) hues indicating larger (smaller) catchments.

for river discharge). Challenges to compound flood hazard assessment at a large scale under projected climate involve precisely estimating the interaction between the various drivers given the uncertainty of Earth system models (Deser et al., 2020). As an initial proof of concept, only four published studies have assessed future changes. These are concentrated around Europe (Bevacqua et al., 2019; Kew et al., 2013; Klerk et al., 2015) and the United States (Moftakhari et al., 2017).

Although a limited number of studies have projected changes in compound floods, these studies (Bevacqua et al., 2019; Kew et al., 2013; Klerk et al., 2015; Moftakhari et al., 2017) have carefully attempted to evaluate the uncertainty. However, earlier studies (Bevacqua et al., 2019; Kew et al., 2013; Klerk et al., 2015) have used different spatial scales, different time periods, variations in the underlying drivers, and different flood sampling schemes (annual maxima or peak over threshold), making it difficult to draw conclusions about future compound flooding for Europe. At the large river basin scale, a few European studies (Kew et al., 2013; Klerk et al., 2015) have analyzed future changes in compound floods from North Sea storm surges and river floods from the Rhine basin considering multiple time lags. While Kew et al. (2013) has used north-northwesterly winds over the North Sea as a proxy for storm surges and multiple-day precipitation as a proxy for peak discharge, Klerk et al. (2015) has implemented the Delft Continental Shelf Model to simulate high storm surges at the Hoek van Holland tide gauge and a lumped conceptual hydrological model to simulate daily runoff. Moftakhari et al. (2017) has assessed compound floods resulting from future SLR and river discharge scenarios at eight coastal-estuarine systems over the Pacific and Atlantic coasts of the United States. Despite their focus on uncertainty quantification, projections of changes in storm surges, one of the drivers, remain elusive. On a continental scale, to our knowledge, only a single study (Bevacqua et al., 2019) covering a large portion of European coasts (including the Mediterranean and Baltic sea regions) has

evaluated future trends in compound floods. However, this study investigated the concurrence of high sea levels and precipitation-induced pluvial flooding.

Second, owing to computational complexity, studies of future compound flood hazards include projection scenarios based on coarse spatial resolutions and simplified assumptions to model individual drivers of compound floods. For instance, in Kew et al. (2013), the forcing variables for simulating extreme North Sea surges were derived from the ESSENCE data set, generated from ECHAM5/MPI-OM coupled general circulation models (GCM) with a horizontal resolution of 180 km (T63 grids). Likewise, in Bevacqua et al. (2019) the meteorological forcing was directly taken from the output of six GCMs available at the CMIP5 archive, for which neither any downscaling nor bias correction was implemented. However, modeling storm surges in shallow coastal areas requires high-resolution meshes and meteorological forcing to capture the effects of complex coastal bathymetry and topography (Marsooli et al., 2019; Muis et al., 2016). The same holds for extreme precipitation causing river floods as another driver of coastal compound floods (Prudhomme et al., 2002). The coarse resolution of GCMs possibly introduces large uncertainty in the representation of the drivers of compound floods. In addition, assessing compound flooding based on coarse spatial resolutions of the hydrometeorological forcing may underestimate extremes (Marsooli et al., 2019; Muis et al., 2016). Further, despite strong confidence in SLR (Kopp et al., 2014), only a small number of studies have assessed the contribution of SLR to compound floods over the United States (Moftakhari et al., 2017), Europe (Bevacqua et al., 2019; Klerk et al., 2015), and Bangladesh (Ikeuchi et al., 2015) coasts. A few regional-scale assessments (Ikeuchi et al., 2015; Klerk et al., 2015) have assumed a spatially homogeneous SLR (i.e., the same increase in sea levels among different locations). However, SLR may stem from various factors, such as nonuniform changes in ocean dynamics, heat and salinity content, glacial isostatic adjustment, and vertical land motion at individual locations (Kopp et al., 2014). Hence, localized SLR projections are required to represent the spatial variability of SLR. Further, Moftakhari et al. (2017) has considered a stationary hydrologic regime (i.e., floods do not change between past and projected periods), which is not a realistic assumption in a changing climate (Stocker et al., 2013).

Third, unlike univariate extremes, the joint occurrence of multiple drivers can be measured by upper tail dependence, which takes into account the dependence at the tail of the probability distributions. While the state-of-the-art literature on projected changes in compound floods has mainly assessed the central dependence among the underlying drivers, this measure only considers the probability that one variable exceeds its median given that the other variable also exceeds its median. Ignoring tail dependence among compound flood drivers may underestimate hazards associated with low-probability but high-impact events.

Understanding the mechanisms underpinning compound floods, under consideration of the interactions among the drivers, especially at the tail of the distribution, in a high-resolution climate change projection, is of great importance to manage and minimize risks associated with climate change. To summarize, there is no study available at the continental scale in general and northwestern Europe in particular that estimates possible future changes in compound floods, driven by storm surges, SLR and river peak discharge, in an ensemble of high-resolution regional climate models (RCMs). Further, the heterogeneity of the few available studies make it impossible to derive solid statements for the European coasts. To fill these gaps, here we quantify the future (2040–2069) changes in compound flood hazard resulting from high coastal water level, SLR, and peak fluvial discharge over northwestern Europe. These projections include an uncertainty assessment and are compared to a reference period (1981–2005). Northwestern Europe comprises one of the densely populated major economic hubs (McMaster, 2016) (Figure 1 and Tables S1 and S2 in the supporting information) and is extremely vulnerable to severe storm-induced compound flooding (Kew et al., 2013; Ward et al., 2018). We apply a hydrodynamic approach to simulate annual maxima surge and WaterGAP Global Hydrology Model (WGHM) to compute peak river runoff (or streamflow) from individual river basins and analyze the joint frequency of the compound floods (section 2.3). Using high-resolution (0.11° or 12.5 km), dynamically downscaled meteorological forcing from the World Climate Research Program's CORDEX (COordinated Regional Climate Downscaling EXperiment) framework (Dosio, 2016) for Europe, we simulate projected surge. An analysis of Europe-wide precipitation climatology (Prein et al., 2016) has shown a clear improvement in 0.11° CORDEX-RCM simulations in reproducing mean and extreme climatology for all regions and seasons due to improved representation of the orography and surface characteristics, and better resolved physical processes, such as convection. We evaluate the performance of RCM-forced compound flood climatology against observations; such evaluations were largely ignored in

earlier assessments (Bevacqua et al., 2019; Kew et al., 2013). Further, we integrate our results with the physically based, probabilistic SLR projection, labeled here as K14 (Kopp et al., 2014) for the RCP8.5 high emission scenario that represents a possible increase in radiative forcing of 8.5 W m^{-2} by 2100 (section 2.3). Our study assesses, for the first time, how projected climate change will affect the joint frequency of compound floods over the coastal regions of northwestern Europe involving all three drivers (i.e., localized SLR, storm surges, and river floods) based on high-resolution, dynamically downscaled climate realizations.

2. Data and Methods

2.1. Hydrometeorological Data

Hourly sea level observations (in meter) for the reference period, 1981–2005 for northwestern Europe (approximately 46–66°N and −12.5°W to 19°E) were obtained from 23 tide gauges (Table S1) archived at Global Extreme Sea Level Analysis version II database (Woodworth et al., 2016). Each tide gauge contains high-quality sea-level records with less than 6 months of missing data during the reference period. Relative sea-level data at Hoek van Holland tide gauge was obtained from Directorate for Public Works and Water Management, Rijkswaterstaat, the Netherlands. Most of the tide gauges contain records of hourly resolution. A few tide gauges have higher sampling frequencies, such as tide gauges from the UK and Norway with sampling frequencies of 15 and 10 min from 1993 and 2001 onward, respectively. For consistency, observations of higher sampling frequencies were transformed to hourly resolution by calculating the median of the subhourly values (Wahl et al., 2017). We obtain daily continuous river discharge data from 39 medium to large-sized river basins with catchment area between 1,000 and 160,800 km^2 (70% of stations with catchment area <10,000 km^2; Table S2) from the Global Runoff Data Centre (GRDC; Grabs, 1997). Each of these stream gauges are located within 200 km radius from the tide gauges (Ganguli & Merz, 2019a, 2019b). The use of gauge-based observations facilitates model validations in the historical (1981–2005) period.

To simulate surge and streamflow in the reference (1981–2005) and projected (2040–2069) periods, we use high-resolution (0.11° in a rotated-pole grid, that is, about 12 km) time series from the Euro CORDEX data set archived at https://esg-dn1.nsc.liu.se/search/esgf-liu/. Given the available data, we obtain 3-hourly sea level pressure ("psl") and 6-hourly eastward ("uas") and northward ("vas") near-surface wind, bias-corrected daily precipitation ("prAdjust") and daily temperature ("tasAdjust"; Dosio, 2016) from a single RCM: SHMI's Rossby Centre regional atmospheric model (RCA4; Strandberg et al., 2015) forced by five host GCMs participating in CMIP5, namely, CNRM-CERFACS-CNRM-CM5, ICHEC-EC-EARTH, IPSL-IPSL-CM5A-MR, MOHC-HadGEM2-ES, and MPI-M-MPI-ESM-LR. For each host GCM, the first initial condition ensemble (*r1i1p1*) has been used except for the ICHEC-EC-EARTH model. Here, the *r12i1p1* initial condition ensemble has been used, as this is the only realization, which provides bias-corrected precipitation and temperature. All simulations have the same physical parameterization (*p1*) and initialization method (*i1*) but differ in initial states (i.e., *r1* and *r12*), indicating internal variability. The pressure and wind data come in different resolutions, but as the model needs a consistent time step for the inputs, we force the hydrodynamic model with 6-hourly pressure and wind data.

Although temperature and wind speed are typically well simulated by regional climate models, considerable biases exist for precipitation (Schoetter et al., 2012). Further, bias-corrected meteorological forcing substantially improves hydrologic predictions (Roy et al., 2017; Teutschbein & Seibert, 2013). Hence, to simulate daily river discharge, we use bias-corrected precipitation and temperature time series (Dosio, 2016), respectively. As downwelling radiation, which is required for the streamflow simulation, is currently not available from the RCM runs, we obtain downwelling longwave and shortwave radiation data from reanalysis, WFDEI-GPCC (Weedon et al., 2014), available at 0.5°. To force the hydrodynamic model Delft3D, we use the climate model output available at 0.11° resolution. However, since the grid resolution of the hydrological model WGHM is 0.5° (Müller Schmied et al., 2014), the meteorological forcing data are regridded to 0.5° to match the resolution of the hydrologic model.

2.2. Sea Level Projections

Following earlier studies (Kopp et al., 2014; Kulp & Strauss, 2019; Lin et al., 2016; Lin & Shullman, 2017; Lowe et al., 2001; Marsooli et al., 2019; Sterl et al., 2009), we assume a negligible influence of mean SLR on surge heights (SLR = 0) during the reference (1981–2005) period. For the future, since SLR projections

are not available from CORDEX simulation, we use a downscaled localized SLR scenario, K14 (Kopp et al., 2014) under the RCP8.5 scenario. K14 offers physical model-based projections based on IPCC AR5 scenarios. Apart from ice sheet components for Greenland, West Antarctic, and the East Antarctic, K14 incorporates all spatially explicit submodels for climatic components of SLR (such as glacier and ice cap surface mass balance, oceanographic processes, land water storage) as well as nonclimatic components (such as global isostatic adjustment, sediment compaction, and tectonics). Leveraging tide gauge observations from the Permanent Service for Mean Sea Level (PMSL), K14 estimates historical SLR rates and generates projections on a 2° global grid that intersects with the world coastline. Further, it offers projections for each RCP, based on 10,000 Monte Carlo samples of SLR at each PMSL tide gauge location.

The SLR projections are based on single realizations of 29 GCMs and represent a 19-year running average value available for each decade since the year 2010 (i.e., average over 2001–2019) until 2300. These SLR projections assume the year 2000 as the baseline year (i.e., SLR = 0), which we treat as a part of the reference period with respect to sea level for hazard assessment. The coastal flood height should be estimated by combining distributions of projected surge height and SLR, projected by the same climate model since a change in storm climatology and SLR both are affected by the large-scale environment (Lin & Shullman, 2017). However, neither probabilistic SLR projections based on individual RCM nor GCM realizations are currently available in the literature. Hence, following the earlier studies (Lin et al., 2016, 2019; Lin & Shullman, 2017; Marsooli et al., 2019), we combine projected surge height with composite SLR distributions for individual GCM-RCM realizations and their multimodel ensemble (discussed later). For the future period, we compute median values of decadal localized SLR projections between 2040 and 2069 (i.e., $2040 \leq s < 2069$, where s indicates a random variable, SLR) at individual tide gauge locations before adding it to the projected surge.

2.3. Methods

2.3.1. Statistical Postprocessing of Compound Flood Drivers

While we force dynamically downscaled (as a preprocessor) climate model simulations to reduce meteorological input uncertainty and to minimize a possible error propagation through hydrodynamic and hydrologic models, we incorporate a statistical postprocessor, which effectively reduces the uncertainty in ensemble climate projections (Kang et al., 2010). Following earlier studies (Lin et al., 2016, 2019; Roy et al., 2017), we bias correct the simulated surge and discharge climatology during the reference period using a simple quantile mapping (or CDF matching) approach, in which a quantile of the present-day simulated distribution is replaced by the same quantile of the present-day observed distribution preserving its statistical moments (Maraun, 2016). Given a historical surge or discharge time series x, the method is formulated as

$$\widetilde{x}_{m-c} = F^{-1}_{o-c}[F_{m-c}(x_{m-c})] \tag{1}$$

where $\widetilde{x}$ is the bias-corrected surge or discharge time series, $F(\cdot)$ is the Cumulative Distribution Function (CDF) and $F^{-1}(\cdot)$ is the inverse CDF of either observation ("o") or model ("m") for the historical period or current climate condition ("c"). We use the Gaussian kernel density function with the bandwidth according to Silverman's rule of thumb to estimate the Probability Density Function (PDF) and its inverse PDF (MATLAB "ksdensity" function). To prevent negative values for discharge, we restrict the kernel density estimator to positive values by setting the ksdensity function parameter to "positive."

Equation 1 assumes only changes in the mean but stationarity in the variance and skew of the distribution. This, however, may not hold since climate model projection may lead to changes in higher order moments as well. In order to incorporate information from the distribution of the model projection, the difference between CDFs of the future and the reference period is taken into account. To bias correct the hourly surge simulated by the Delft3D hydrodynamic model, in the projected ("p") period, we apply an equidistant quantile mapping by adding an additive factor (Yang et al., 2018) as below:

$$\widetilde{x}_{m-p} = x_{m-p} + F^{-1}_{o-c}\left[F_{m-p}\left(x_{m-p}\right)\right] - F^{-1}_{m-c}\left[F_{m-p}\left(x_{m-p}\right)\right] \tag{2}$$

To avoid the problem of negative values in the projected discharge time series simulated by WGHM, we use an equiratio quantile mapping by applying a multiplicative factor (Yang et al., 2018) as below:

$$\widetilde{x}_{m-p} = x_{m-p} \times \frac{F^{-1}_{o-c}\left[F_{m-p}\left(x_{m-p}\right)\right]}{F^{-1}_{m-c}\left[F_{m-p}\left(x_{m-p}\right)\right]} \quad (3)$$

Equation 2 is analogous to the bias correction of temperature, where an additive adjustment is usually applied to correct the mean biases of the model, whereas Equation 3, that is, scaling adjustment by a multiplicative factor, is often utilized for precipitation to avoid negative values. Figures S1 and S2 compare the performance of RCM-forced annual maxima surge and peak discharge time series relative to observations, with and without bias correction for selected tide and stream gauges. The bias correction clearly improves the reproduction of the variability of the compound flood drivers. Figure S3 compares the variability of reference (1981–2005) versus projected (2040–2069) annual maxima surge and corresponding peak fluvial discharge after bias correction across all GCM-RCM realizations for selected tide and stream gauges. There is no clear change between the reference and future periods for the annual maxima surge heights (Figure S3a). In contrast, future peak discharge (Figure S3b) tends to be higher than the observed streamflow, which agrees with earlier assessments over northwestern Europe (Bosshard et al., 2014; Lobanova et al., 2018).

2.3.2. Modeling High Coastal Water Level

To account for the nonstationarity resulting from SLR and interannual variability in hourly sea level observations, we subtract the annual average sea level on a year-to-year basis (Muis et al., 2016; Wahl et al., 2017). We use the MATLAB UTide package (Codiga, 2011) for a year-to-year harmonic tidal analysis with all 67 tidal constituents (Wahl et al., 2015, 2017). The advantage of this tool is its ability to extract tidal constituents from sea level records with uneven temporal sampling and data gaps. Then we compute hourly skew surge as the difference between the maximum observed high water (i.e., detrended time series) and the maximum predicted tidal water level within a tidal cycle for each tide gauge record (Wahl et al., 2017).

To simulate surge with meteorological forcing from GCM-RCM runs, we run the hydrodynamic model Delft3D-FLOW (Delft3D-FLOW, 2014), which uses the depth-averaged shallow water equations. Since we are mostly interested in the meteorological phenomena that drive compound floods, we focus on modeling surges and do not simulate tides. Further, the simulated storm surge does not include the surge component due to wave setup. The review of the literature (Lowe et al., 2001; Sterl et al., 2009) shows a negligible influence of mean sea level rise on surge height; hence, following these studies, the mean sea level rise was not included during the calibration of the hydrodynamic model. The hydrodynamic model was previously calibrated (Paprotny et al., 2019, 2020) with a computational grid of 0.11° with the same spatial extent as the EURO-CORDEX domain using observed skew surges. The model was validated using reanalysis data ERA-Interim for the period 1979–2014 and showed a satisfactory model performance (Paprotny et al., 2019, 2020; Figure 2, right panel of Paprotny et al., 2019 shows the calibrated hydrodynamic model performance, which should have been placed in Figure 3, left panel of the paper. Further, in Paprotny et al., 2019, the relative error of surge output was found to vary between −25% to 3% for different tide gauges, which was obtained without applying any statistical postprocessor. We argue that we have bias corrected the surge to ensure better model performance). Details of the model parameters (such as wind drag coefficients and channel roughness) and boundary conditions are discussed in Paprotny et al. (2019, 2020). The simulation is driven by 6-hourly sea level pressure and winds. To accurately simulate annual maxima storm surges, the time step of the calculation is kept as 30-min. The modeled surge output is outputted at every 6-hr interval.

Assuming a negligible nonlinear interaction between surge, S and SLR, R for the RCP8.5 scenario, following the earlier literature (Lin et al., 2016, 2019; Marsooli et al., 2019), we estimate the future coastal flood elevation (H_f) as the sum of the future surge elevation and the SLR scenario. As both variables are given as PDFs, we employ a convolution integral operation to obtain the PDF associated with surge and SLR. Assuming surge and SLR be two independent continuous random variables with PDFs $f_S(s)$ and $f_R(r)$, and $H_f = S + R$, yields the PDF of H_f:

$$f_{H_f}\left(h_f\right) = f_S(s) * f_R(r) = \int_{-\infty}^{\infty} f_S\left(h_f - r\right) f_R(r)\mathrm{d}r \quad (4)$$

Our choice of neglecting nonlinear interactions between water level components is justified as follows: (1) Besides a few regional exceptions (Devlin et al., 2017; Zijl et al., 2019), the assumption of nonlinear

interactions between the water level components are relaxed in continental-scale assessments (Losada et al., 2013; Paprotny et al., 2020) and climate change projections (Lowe et al., 2001; Sterl et al., 2009; Vousdoukas et al., 2018), since the resulting error is negligible as compared to the other sources of uncertainty, such as the cascading uncertainty in numerical modeling chains and climate model experiments (Vousdoukas et al., 2018). (2) Simulating water level resulting from nonlinear interaction of streamflow and coastal water level integrating tide, surge, and wave setup would require a fully coupled modeling approach (Santiago-Collazo et al., 2019), which goes beyond the computational capabilities for a continental-scale assessment.

2.3.3. Modeling Streamflow

We use the global hydrological model, WGHM version 2.2d (Döll et al., 2003; Müller Schmied et al., 2020) to simulate current and future streamflow at daily time steps. WGHM simulates storage and flux components of the continental water cycle with a spatial resolution of 0.5° (~55 km). WGHM is run with daily reanalysis-based WFDEI-GPCC (Watch Forcing Data based on ERA-Interim) (Weedon et al., 2014) meteorological forcing. The model is calibrated (Müller Schmied et al., 2014) based on observed river discharge at stations around the world by tuning the runoff coefficient parameter. Figure S4 compares the performance of WFDEI-forced WGHM runs with and without bias correction using an overall score, *OS*, which is the median of runoff signatures and temporal variability (Text S1.1). $0.8 < OS < 1$ indicates an excellent performance; $0.6 < OS < 0.8$ represents good performance, and $0.4 < OS < 0.6$ shows moderate performance. Without bias correction, we note (from Figure S4a) a moderate to good performance especially for stream gauges located in southern UK. Bias correction clearly improves the overall performance for the gauges across the southern and western coasts of the UK, Germany, Denmark, and the Swedish coast. Figure S5 shows the heat map of the performance of bias-corrected daily discharge obtained from RCM-driven WGHM runs for each stream gauge location during the reference period, indicating that most of the sites show a good to excellent model performance with *OS* larger than 0.6.

2.3.4. Identification of Compound Event

Following the earlier literature (Ganguli & Merz, 2019a, 2019b), we identify compound floods when hourly annual maxima surge events coincide with d day (where $d = 1,.., n$, which is proportional to the catchment area) lagged peak river discharge within ±7 days of the surge event. To account for the delay between storm surges and river peak flows, which are caused by the same meteorological event, we shift the discharge time series by a lag time based on catchment-specific watershed response time. Following the earlier literature (Berghuijs et al., 2019; Merz et al., 2018), we consider two events, that is, a coastal surge and a river peak flow, as a compound event when they occur within a window of ±7 days. This window is motivated by the fact that two events do not need to occur at exactly the same day to enhance their impacts. For instance, the first event could weaken flood defense systems or stress disaster management capacities, posing a problem when the second event occurs with a few days delay.

While other statistical methods, such as *r*-largest methods or peak-over-threshold approach (Wahl et al., 2017), could be applied to sample surge by choosing more than one event per year, we prefer annual maxima since this typically ensures the *iid* (independent, identically distributed) assumption of extreme value statistics. The annual maxima method is also easier to apply without the need of choosing an appropriate threshold from surge time series over a large number of tide gauges and across a suite of climate model run (Muis et al., 2016).

2.3.5. Calculation of Basin Lag Time (Watershed Response Time)

The inclusion of basin lag time is motivated by the delay between storm surges and river floods when both are caused by the same large-scale meteorological event. A lag time of a few days has been suggested for establishing a moderate to a strong correlation between surge and river discharge (Petroliagkis, 2018). We adopt the watershed response time (d), representing the time needed for rainfall to propagate through the catchment to the downstream streamflow gauge (Ganguli & Merz, 2019a, 2019b):

$$d = 2.51 A_d^{0.4}[\text{hrs}] = 0.11 A_d^{0.4}[\text{days}] \quad (5)$$

Where A_d denotes the catchment area (in km^2). For each tide gauge-stream gauge combination, we search for the highest peak discharge of the d-shifted streamflow time series within an interval of ±7 days of the occurrence of the annual maxima surge event.

2.3.6. Upper-Tail-Dependence Coefficient

The upper-tail-dependence coefficient (UTDC) between surge and peak discharge is determined using a nonparametric tail-dependence metrics, namely, Capéraá-Fougéres-Genest (CFG) estimator, λ_{CFG}. We estimate $\widehat{\lambda}_{CFG}$ as follows (Frahm et al., 2005)

$$\widehat{\lambda}_{CFG} = 2 - 2\exp\left\{\frac{1}{n}\sum_{i=1}^{n}\log\left[\sqrt{\log\left(\frac{1}{u_i}\right)\log\left(\frac{1}{v_i}\right)}\Big/\log\left(\frac{1}{\max(u_i, v_i)^2}\right)\right]\right\} \quad (6)$$

Where,u_i and v_i denote the empirical cumulative distributions of the compound flood drivers and n indicates the number of samples. We determine the empirical UTDC by calculating the empirical cumulative distribution of each of the variables and then substituting the corresponding values in Equation 6. We determine statistical significance (p value < 0.10) of UTDC estimates by drawing 10,000 random bootstrap samples and then calculating the p value of the test using a standard percentile-based approach.

2.3.7. Marginal Distribution Fit of Correlated Compound Flood Drivers

For the marginal distributions of annual maxima surge and peak discharge the Gamma, Log-normal, Log-logistic, and Generalized Extreme Value (GEV) are used. While the parameters of the first three distributions are determined using the method of maximum likelihood, the GEV parameters are estimated via Bayesian inference combined with Differential Evaluation Markov Chain (DE-MC) Monte Carlo (MC) simulation (Ganguli & Merz, 2019b) owing to the fact that the application of maximum likelihood estimator in small samples ($n \leq 50$) often leads to an absurd value of the GEV shape parameter. The use of a Bayesian prior distribution restricts the shape parameter to a physically consistent range (Martins & Stedinger, 2000). The DE-MC simulations are run for 3,000 iterations of multiple ($n = 5$) parallel chains and the convergence of MC simulation is checked by a scale reduction factor $\left(\widehat{R}\right)$ that suggests the value of $\widehat{R}$ should remain below the threshold value of 1.1. The parameters of the GEV distribution are derived by computing the 50th percentile of *postburn in* (2001st–3000th) random draw.

We select the marginal distribution fit based on the corrected Akaike Information Criterion (AIC_c) (Lee & Ghosh, 2009). The validity of the best fitted distribution is evaluated using Kolmogorov-Smirnov and Cramer-von-Mises goodness-of-fit statistics at a 10% significance level. Figures S6 and S7 show the probability-probability plots of empirical versus simulated compound flood drivers for the historical and projected periods, which indicates good performance of the modeled distribution. Since SLR variates are available in a discrete form as an output from a Monte Carlo simulations, following the previous studies (Lin et al., 2016, 2019; Lin & Shullman, 2017; Marsooli et al., 2019), the marginal distribution of SLR is estimated using nonparametric kernel density estimators. We use the Gaussian kernel density function to fit PDF of SLR whereas the bandwidth is obtained using Silverman's rule of thumb.

2.3.8. Copula-Based Dependence Modeling

To model the joint distribution, we select the tide gauge-stream gauge pairs for which empirical upper tail dependence coefficient, UTDC between surge and peak discharge, $\lambda_{emp} > 0$ and absolute signal (simulated change) to noise (simulated variability) ratio, $|SNR| > 1$ for both historical and projected periods. We relax the statistical significance level to include all sites that show positive upper tail dependency, even when the total correlation value is weakly negative or insignificant. This also allows us to include gauges that show a large climate change signal over its variability (or noise).

Based on analytical tractability and ability to model upper tail dependence, we use three copula families: Gumbel-Hougaard, flipped Clayton and Student's t. We estimate the copula parameters using the maximum pseudo-likelihood method. We evaluate the adequacy of the copula model in capturing the upper tail dependence using Mean Error to Standard Error (MESE) statistics (Frahm et al., 2005) between empirical UTDC estimate, $\widehat{\lambda}_{emp}$ and simulated UTDC from the parametric family of the copula, $\widehat{\lambda}_{C_\theta}$ (see Equation 7). For comparison, synthetic bivariate samples are simulated from the copula family, C_θ with the same length as the observation, $\mathbf{X}_{i,d} = \{X_{i,1}, X_{i,2}\}$. The computation of $\widehat{\lambda}_{C_\theta}$ is repeated for $N = 500$ random runs. MESE statistics is then computed as the absolute deviations between $\widehat{\lambda}_{emp}$ and the mean of $\widehat{\lambda}_{C_\theta}$, $\boldsymbol{\mu}\left(\widehat{\lambda}_{C_\theta}\right)$ for all 500 runs normalized by its standard deviation, $\boldsymbol{\sigma}\left(\widehat{\lambda}_{C_\theta}\right)$

$$MESE = \frac{\boldsymbol{abs}\left|\widehat{\lambda}_{emp} - \boldsymbol{\mu}\left(\widehat{\lambda}_{C_\theta}\right)\right|}{\boldsymbol{\sigma}\left(\widehat{\lambda}_{C_\theta}\right)} \tag{7}$$

Out of these three copula families, we select the one that results in the lowest value of MESE statistics for a given location. Further, we assess the goodness-of-fit of the selected copula model using parametric bootstrap-based Cramér-von Mises test statistics at a 10% significance level (Genest et al., 2009). Figure S8 shows the performance of empirical versus the best selected copula models in terms of tail dependency for historical and projected periods. The scatter close to the 1:1 line indicates a good performance of the selected copula.

2.3.9. Determination of Joint Return Period

Considering the dependency between annual maxima surge and associated peak river discharge, we select the 95th percentile value of each of the compound flood drivers to compute the joint return period. For a given compound flood event pair, the bivariate joint return periods (JRP) of the cooccurrence or successive occurrence (within ±7 days' time window) of annual maxima surge and peak discharge (i.e., both) exceeding 95th percentile threshold using "AND" operator can be expressed as

$$\begin{aligned} T(s,\,q)_{AND} &= \frac{1}{P(S > s \,\text{and}\, Q > q)} = \frac{1}{1 - [P(S \le s) + P(Q \le q) - P(S \le s,\, Q \le q)]} \\ &= \frac{1}{1 - F_S(s) - F_Q(q) + C_{SQ}(s,\,q)} \end{aligned} \tag{8}$$

Where $F_S\,(s)$ and $F_Q\,(q)$ are the marginal distributions of the surge and discharge, respectively, and $C_{SQ}\,(s,q)$ is the copula-based joint distribution.

Accounting for the 50th percentile (median) SLR condition, the joint return period of coastal flood elevation and river discharge for the projected period using "AND" operator

$$T'\left(h_f,',q\right)_{AND} = \frac{1}{P\left(H_f' > h_f' \,\text{and}\, Q > q\right)} = \frac{1}{1 - F_{H_f'}\left(h_f'\right) - F_Q(q) + C_{H_f'Q}\left(h_f', q\right)} \tag{9}$$

Where $h_f' = s_{95} + r_{50}$; s_{95} and r_{50} correspond to 95th percentile projected surge and 50th percentile SLR values for a tide gauge. $F_{H_f}{}'\left(h_f'\right)$ is the CDF value corresponding to h_f' and obtained from interpolation of coastal flood elevation with its CDF, $F_{H_f}\left(h_f\right)$. $F_{H_f}\left(h_f\right)$ is obtained using numerical integration of $f_{H_f}\left(h_f\right)$ as obtained from Equation 4.

3. Results

3.1. Simulation and Projection of Compound Flood Drivers

A comparison of the univariate return level plots between observed and modeled surge heights during the reference period, 1981–2005 (when SLR = 0; Figures 2 and S9), shows a good agreement for most of the selected tide gauges, except in Dover, where return levels of surge heights are overestimated by four out of five models. This is due to the complex topography of the Dover Straits that limits the ability of hydrodynamic model to adequately simulate the flow of water through such narrow straits. Except for a few models, the return level plots show a marginal increase in surge heights for the projected time period. Considering the contribution of SLR (Figures 2 and S9) along with surge, the return levels of projected coastal flood elevation are significantly higher across all gauges. This suggests that SLR plays an important role in future compound floods in northwestern Europe and should not be neglected.

The univariate return level plots of river floods (Figure 3) demonstrate a good agreement between the median of the multimodel ensemble and the observed peak discharge for most of the selected stream gauges. The observed peak discharge is underestimated at the gauge Gothenburg (River Lagan, Figure 3e) and overestimated at the gauge Dover (River Thames, Figure 3f). GHM often tends to underestimate the peak discharge of river basins at high latitudes (Gudmundsson et al., 2012). The underestimation at Gothenburg and other

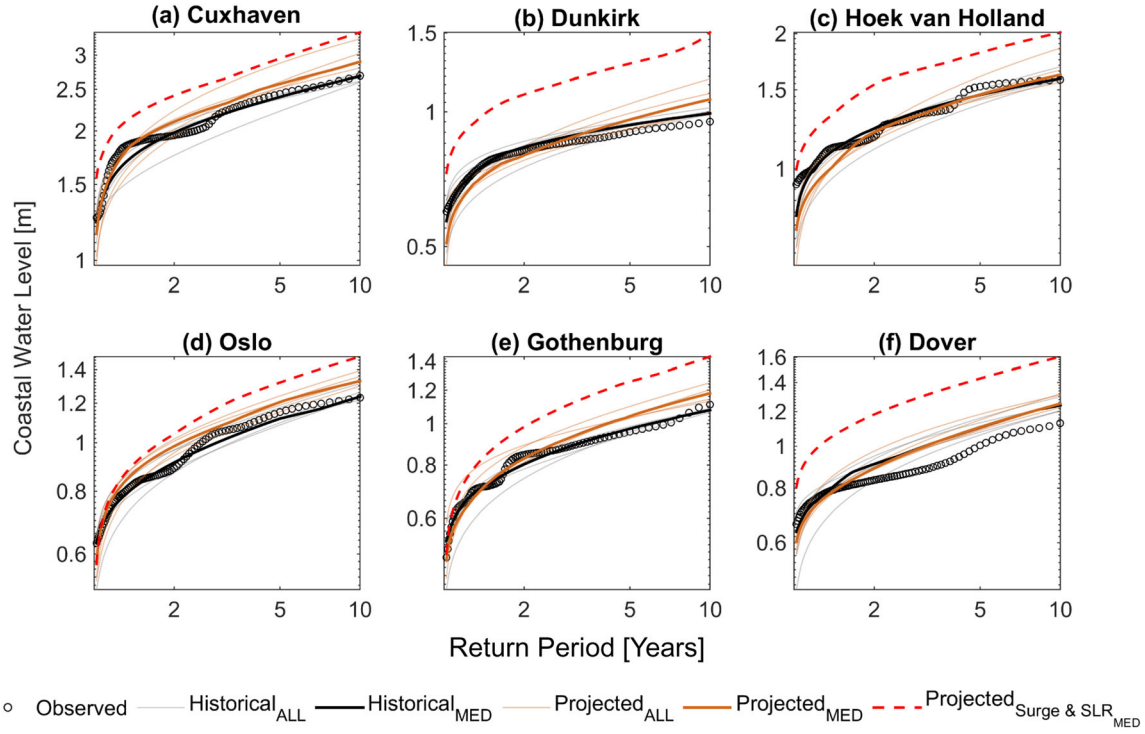

Figure 2. Comparison of historical (1981–2005) versus projected (2040–2069) extreme coastal water levels with and without sea level rise at selected tide gauges in northwestern Europe. (a) Cuxhaven, (b) Dunkirk, (c) Hoek van Holland, (d) Oslo, (e) Gothenburg, (f) Dover. Observations of annual maxima are shown as circles, simulations from individual climate model simulations (ALL) are shown as thin solid lines (historical in gray; projected in light brown). The median (MED) of the multimodel RCM ensemble is shown as thick solid lines (historical in black; projected in dark brown). It is the median of the single GCM-RCM realizations for each return period. The return levels of projected surge from the multimodel RCM ensemble considering the effect of SLR is shown using thick dashed red lines (surge and SLR_{MED}).

high-altitude gauges is likely a consequence of biases in the forcing data, for instance, the local orographic effects on precipitation for a region with complex topography cannot be resolved well with the large grid cells of RCM simulations (Gudmundsson et al., 2012). For the future, a marked increase in peak discharge is seen at most stream gauges, especially for the larger rivers Rhine (gauge Hoek van Holland) and Elbe (gauge Cuxhaven). Overall, we find that future peak discharge (Figure 3) tends to be higher than the observed streamflow, which agrees well with earlier assessments over northwestern Europe (Lobanova et al., 2018; Schneider et al., 2013).

3.2. Spatial Variability in Interdependence Between Compound Flood Drivers

Next, we analyze the dependence between compound flood drivers. Since central dependence measures (such as rank correlation coefficients) are more suitable to reflect dependence at the centre of the distribution (Ganguli & Merz, 2019a), we use an upper tail dependence coefficient (UTDC; section 2.3) to assess the dependence between annual maxima surge and peak river discharge (Figure 4). In the observations (Figure 4a), we find stronger positive UTDC values along the north shore of Scotland; the dependence pattern weakens gradually toward the south. The strong dependence along the north shore of Scotland has been attributed to orographically enhanced precipitation owing to the presence of hills on the northern side and cyclone traveling from northeastward to the north (Svensson & Jones, 2004). Statistically significant, rather strong positive UTDC values (i.e., $\lambda \geq 0.4$) are also observed for the German tide gauge Cuxhaven, Danish tide gauge Esbjerg, Dutch tide gauge Delfzijl, and Swedish tide gauge Stenungsund. Weakly negative to insignificant positive dependence values are mostly concentrated in southeast England, which could be due to recurrent dry winter and persistent groundwater droughts leading to multiyear hydrological droughts (Kendon et al., 2013). The values of observed UTDC vary between −0.12 and 0.61 (Figure 4a), whereas the values of RCM-modeled UTDC during the reference period (1981–2005) vary between 0.15 and 0.70

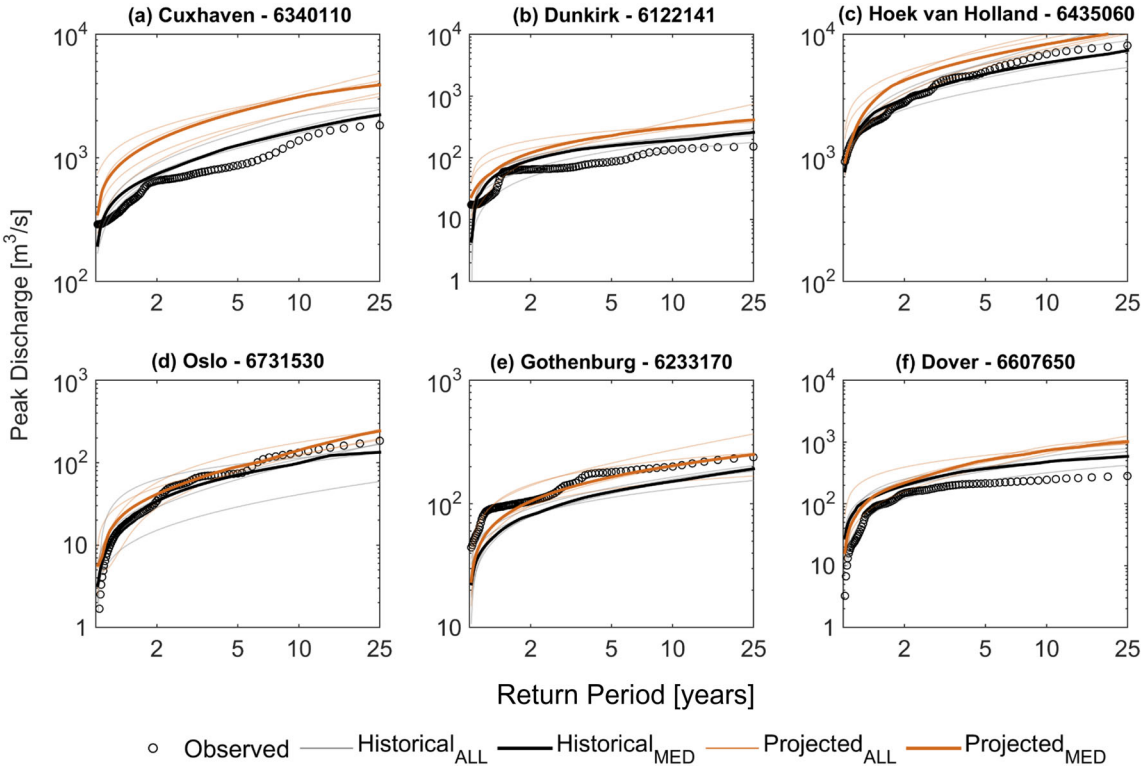

Figure 3. Comparison of historical (1981–2005) versus projected (2040–2069) peak river discharge at selected streamflow gauges in northwestern Europe. (a) Cuxhaven, (b) Dunkirk, (c) Hoek van Holland, (d) Oslo, (e) Gothenburg, (f) Dover. The observed return levels are shown using circles (in black), simulated return levels forced with individual RCMs are shown using thin solid lines (historical in gray; projected in light brown). The return levels of the multimodel ensemble is shown using thick solid lines (historical in black; projected in dark brown). The multimodel return level curve is obtained as the median of each return level curve of the different GCM-RCM realizations.

(Figures 4b and 4d). Overall, the strength of dependence weakens in the projected period (2040–2069) among all models (Figure 4c) and ranges between −0.17 and 0.51. The spatial pattern of the mean UTDC values, that is, averaged across climate realizations, for the reference period (Figures 4a and 4b) shows mostly an overestimation relative to the observations, along the western coast of the United Kingdom (UK), France, Germany, and the Netherlands. However, there are also a few locations where the dependence is underestimated, in particular along the north shore of Scotland and a few gauges in Scandinavian coasts (Figures 4a and 4b). These regions are characterized by hills and high plateau in the north of Scotland and by deep valleys and fjords around the Scandinavian coasts. The underestimation in dependence patterns is likely caused by this pronounced orography, which is not well simulated by many climate models (Gudmundsson et al., 2012).

The spatial pattern of the mean UTDC for the projected period indicates a weakening of the dependence strength with a decrease in UTDC for 93% of the stream gauges (Figures 4c and 4d). The decrease in mean UTDC ranges between −3.5% and −84.6%. A few gauges in the UK and Denmark show an increase in the mean UTDC in the range from 0.3% to 34%. Among the proposed mechanisms for a weakening of dependence strength is a decrease in wind speed in large parts of northwestern Europe (the North Atlantic region) associated with a poleward shift of the North Atlantic jet (Gonzalez et al., 2019; Kjellström et al., 2018). Pooling RCM realizations, 92% (66 out of 72) and 78% (56 out of 72) of the tide gauge–stream gauge pairs show an absolute signal (simulated change) to noise (simulated variability) ratio (SNR) >1 during the historical and projected periods, respectively, indicating that the simulated change is larger than the simulated variability among models (Figures 4a and 4c). This implies a higher model agreement in the reference period compared to the projected period. This decrease in SNR (Figure S10) from reference to projected period is seen over large portions of northwestern Europe, except for gauges across the southeast UK and Scotland (Figure 1), where SNR value tends to be greater than 3.

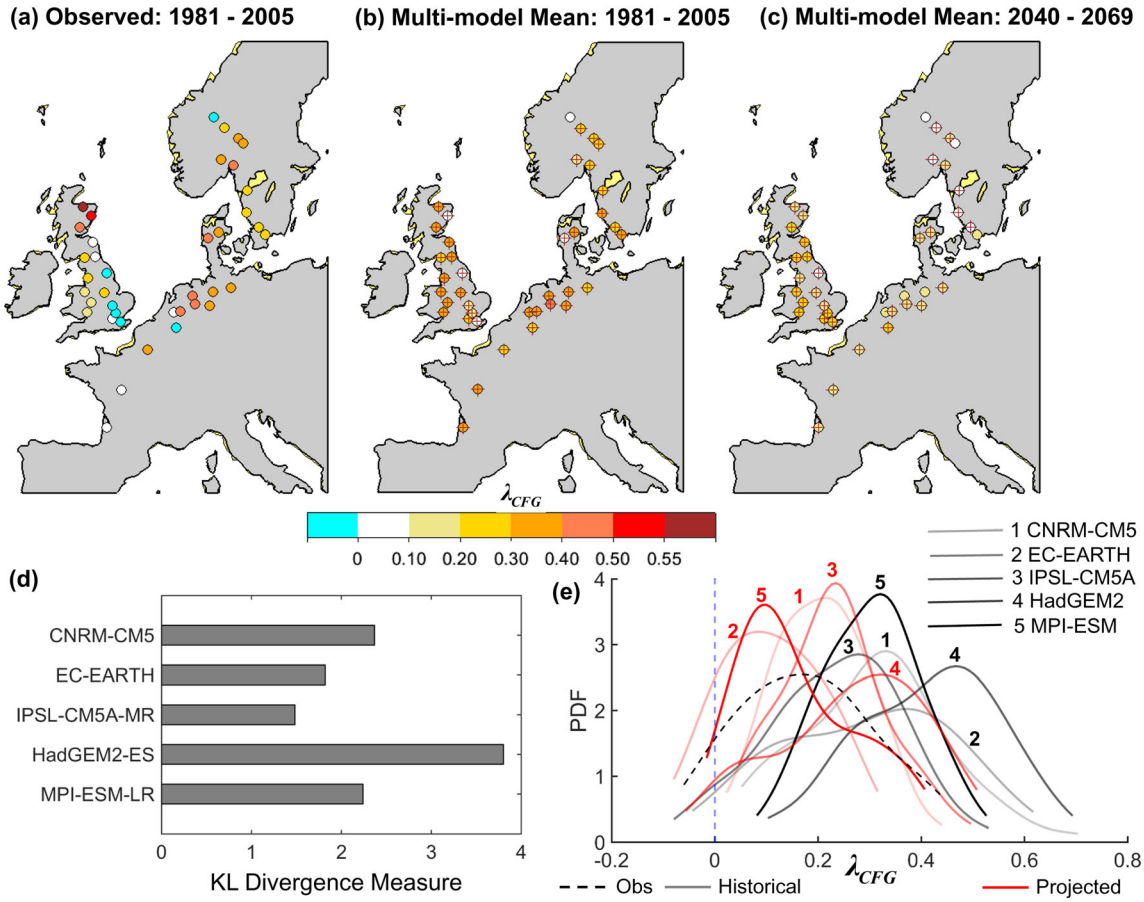

Figure 4. Dependence between annual maxima surge and peak river discharge for observation and GCM-RCM ensembles. Spatial distribution of upper tail dependence coefficient, λ_{CFG} is shown for (a) observation, (b) simulation for the reference period 1981–2005, and (c) the projected period 2040–2069. Circles show stream gauge locations. Shaded circles, ranging from yellow to red, indicate the fraction of models that show significant correlation at 0.10 level if this fraction is 60% or larger (i.e., ≥3/5 models agree in dependence pattern). The circles in white indicate insignificant positive correlation. The circles in cyan show a weak negative dependence, indicating an inverse relation between them that does not lead to a compound event. Locations with |SNR| > 1 (signal-to-noise ratio) are marked with a "plus" sign. (d) Kullback-Leibler (KL) divergence metric between observed and historical upper tail dependence for sites with SNR > 1; the lower the KL divergence measure, the better the match to the empirical distribution. (e) probability density functions (PDF) of upper tail dependence coefficients λ_{CFG} for observations (dashed line), individual GCM-RCM runs during the reference (shaded solid lines in gray) and projected (shaded solid lines in red) periods for sites where SNR > 1.

Next, we compare the UTDC values between observations and simulations for the historical (Figure 4d) and projected period (Figure 4e), considering only those stream gauges for which SNR > 1. In Figure 4d, we show the performance of individual RCM realizations in simulating the observed dependence using the Kullback–Leibler (KL) divergence metric (Thorarinsdottir et al., 2013), which quantifies the divergence between the probability distributions of observation and simulation. While the PDFs of the models tend to encompass the observed dependence (Figure 4e), they differ in their shape and width relative to observations, which may result from model structural differences and relative roles of internal variability (Deser et al., 2020). All models overestimate the dependence strength. RCA4 forced with HadGEM2-ES shows the largest discrepancy between observed and simulated dependence with the largest divergence value as well as the strongest overestimation, whereas IPSL-CM5A performs the best. The changes in simulated standard deviations of UTDC, forced with the RCM ensemble, are up to ±23% of the observed. The biases of the different GCM-RCM realizations show that the driving GCMs have some impact on the dependence pattern. The model structural differences emerge because of limitations in the process representation, such as the parametrization of convection (Prein et al., 2015; Púčik et al., 2017; Rajczak & Schär, 2017). Finally, except HadGEM2-ES and MPI-ESM-LR, three out of five models show statistically significant (p value < 0.05 in two-sample Kolmogorov-Smirnov test) leftward shifts in projected PDFs relative to their historical

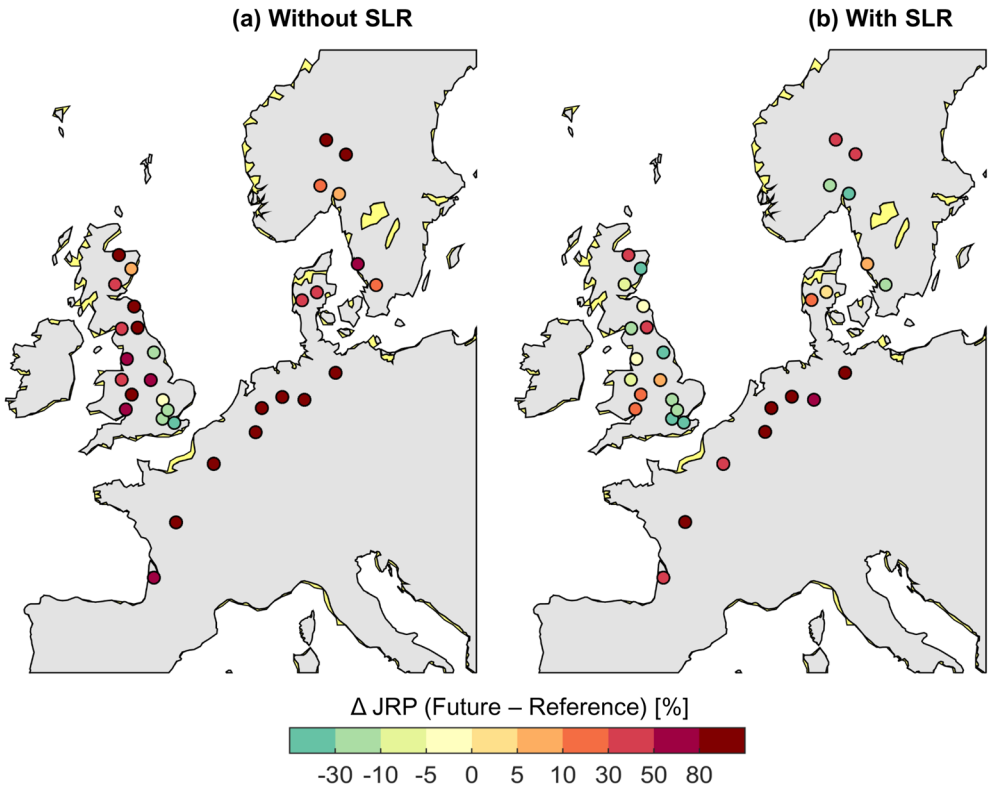

Figure 5. Projected changes in joint return periods. Changes (2040–2069 minus 1981–2005) in joint return periods (JRPs) for the 95th percentile of storm surge and peak river flow without (a) and with (b) the effect of probabilistic sea level rise (SLR) under K14 scenario. The changes in joint return periods are mapped for gauges with |SNR| > 1 owing to the highest agreement among climate models at these locations. Values show the changes in joint return periods (in %) derived from multimodel median climate realization with positive (negative) values showing a decrease (increase) in hazard, relative to the reference period (1981–2005).

simulation, indicating a substantial decrease in the strength of dependence in the projected period (Figure 4e).

An overestimation in the dependence strength by the models highlights the need to be careful when deriving the interdependence between compound flood drivers from climate model realizations. This cautionary note seems particularly valuable, as most of the earlier assessments have not validated their simulated dependence strength (Bevacqua et al., 2019; Kew et al., 2013), which is a robust statistic affecting the occurrence frequency of compound floods.

3.3. Joint Frequency Analysis and Effect of SLR on Compound Flooding

We model the dependence between extreme surge and peak river discharge using a copula-based multivariate distribution (section 2.3) for gauges that show positive UTDC and SNR > 1. Next, we focus on the 95th percentile (occurring on average once in every 20 years) of the compound flood drivers, storm surge, and river peak. The joint return periods of such events, that is, storm surge and river peak flow are both higher than the 95th percentile, for most gauges between 20 and 50 years in the reference period (Figure S11a). This implies a considerable dependence between surge and peak discharge.

Projected changes in compound flood hazard, that is, changes in the mean of the multimodel ensemble relative to the reference period, are shown in Figure 5 for the different streamflow locations and in Figure 6b (indicated by the density plot in black) as PDF across all locations. These changes vary between −41% and 291% across the locations. For instance, selecting those locations with a Joint Return Period (JRP) of 60 years for the 95th percentile of the compound flood drivers in the reference period, this event would occur on average every 35 and 76 years, respectively, in the future period. Neglecting SLR (Figures 5a and 6b),

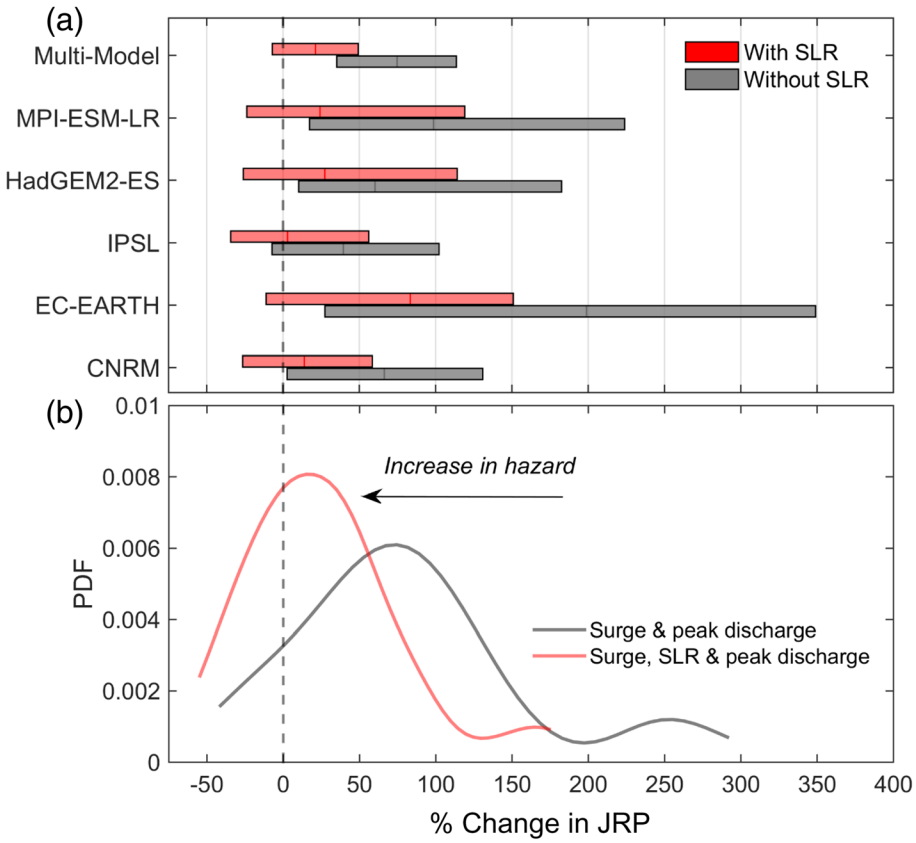

Figure 6. Interquartile range and density functions for the projected changes in joint return periods. (a) Interquartile range of changes in joint return periods (in %) with (red) and without (gray) SLR effect for individual climate models and the multimodel ensemble. The uncertainty, expressed as interquartile range, of changes in joint return periods is larger for individual GCM-RCM realization than for the multimodel ensemble. The median changes are shown using the vertical line within the horizontal boxplot. (b) Kernel density functions of the median of the multimodel RCM ensemble. Incorporating localized SLR, shifts the PDF substantially to the left, indicating the important role of SLR in future compound flood hazard estimation.

anthropogenic climate forcing decreases the flood hazard for the large majority of locations. 87% of the gauges show positive changes in JRPs, out of which 66% show a more than 50% increase in JRP. This increase in JRPs is mainly a consequence of the weakening in dependence strength between storm surges and river peaks (Zscheischler & Seneviratne, 2017) by the 2050s (Figure 4c). On the other hand, only 13% of the gauges show more severe compound flood hazard. Their projected decreases in JRPs range from −0.5% to −41%, with a comparatively more robust decrease for gauges in southeast England. These sites show |SNR| > 1 in RCM simulations (Figures 4b and 4c), indicating a strong agreement among the climate models.

When the effect of SLR is incorporated (Figures 5b and 6b), we find again that anthropogenic climate change tends to decrease compound flood hazard over northwestern Europe. The majority of locations (66%) shows positive changes in JRPs, that is, the given combination of storm surge and river peak flow occurs less often. However, the PDF of changes in JRPs is clearly shifted toward a less favorable situation compared to the situation without SLR (Figure 6b). Of the locations, 34% show negative changes in JRPs, indicating more severe hazard, compared to only 13% when we do not consider the effect of SLR. Hence, SLR is an important driver of compound flood hazard in northwestern Europe and should not be neglected. Stream gauges with more severe hazard by 2050s are mostly concentrated along the coasts of Great Britain (including southeast England, Wales, and Scotland), and a few of them are located along the Scandinavian coast. We find a similar trend in JRPs, when forced with individual RCM realizations (Figures S11b and S11c). When the RCP8.5 RCM ensembles are pooled together (Figure 6a), we observe a high intermodel spread of projected changes

in JRPs without considering SLR. This spread tends to decrease when considering the SLR. Further, the uncertainty of the multimodel ensemble is smaller as compared to the single RCM realizations for both with and without the SLR. The smaller spread of multimodel ensembles results from error cancelation (i.e., the inclusion of an apparent poor model can still improve the skill of the multimodel projection), and leads to greater consistency and increased reliability as shown in earlier studies (Deser et al., 2020; Hagedorn et al., 2005; Tebaldi & Knutti, 2007).

With SLR, the future change in JRPs lies between −7% and 48% (interquartile spread) with a median increase of 22% (Figure 6a) for locations with $|SNR| > 1$. Only considering the compounding effects of storm surge and peak discharge, the median increase in JRP is 75% (with 36% to 112% interquartile spread).

The PDFs of the projected changes in JRPs of the median of the multimodel ensemble show an elongated right tail (Figure 6b). This distortion arises from the large differences in JRP values between the future and reference periods in the ICHEC-EC-EARTH model. Considering SLR, the PDF shows a significant leftward shift (two-sample Kolmogorov-Smirnov test comparing the two PDFs shows p value ≤ 0.05) with an extended right tail and a little change on the lower tail. This shift in PDF is an indication of "changed symmetry" in combination with the "shifted mean/and variance" (changes in median and variance between the two distributions are statistically significant at 5% level using the Wilcoxon rank-sum test and the F test for the homogeneity of variance), as reported in the IPCC AR5 report (page 134, Figure 1.8 in the IPCC Fifth Assessment Report: physical science basis, Stocker et al., 2013), suggesting considering a median SLR projection (section 2.3) favors more frequent occurrence of compound floods than neglecting the SLR.

4. Discussion and Conclusions

We assess future compound floods due to high sea levels, resulting from storm surge together with probabilistic SLR projection, and peak fluvial discharge under the highly energy-intensive RCP8.5 scenario and identify potentially vulnerable sites across northwestern Europe. While past studies relied on GCMs (Bevacqua et al., 2019; Bloemendaal et al., 2019; Lin et al., 2019; Marsooli et al., 2019) and were, thus, limited due to the very coarse spatial model resolution, this study for the first time uses a suite of dynamically downscaled regional climate model simulations archived at EURO-CORDEX domain to analyze future compound flood hazards and associated uncertainty. While most of the earlier assessments on future trends in compound floods were limited to only one (Kew et al., 2013; Klerk et al., 2015) or a few gauges (Moftakhari et al., 2017), we analyze 72 tide gauge-stream gauge pairs representing the entire northwestern Europe. Further, unlike previous assessments that considered either a spatially homogeneous SLR (Ikeuchi et al., 2015; Klerk et al., 2015) or stationary fluvial flood regime (Moftakhari et al., 2017), we consider physical-model-based, probabilistic and localized SLR projections (Kopp et al., 2014) and simulate fluvial peak discharge by forcing a global hydrological model with downscaled climate model simulations (Dosio, 2016).

Our results contrast with a previous study (Bevacqua et al., 2019) that suggests an increased probability of compound floods, in particular for northern Europe, due to higher sea levels and heavier precipitation in a warming climate. As they included precipitation as one of their compound flood drivers, it is not clear whether their result has implications for river flooding or is limited to pluvial floods. Further, they have used meteorological forcing directly from GCMs to simulate coastal water levels and did not bias correct the forcing variables. While the limited ability of GCMs to resolve fine-scale processes may lead to underestimation of extremes, such as storm surge heights (Bloemendaal et al., 2019; Muis et al., 2016) and fluvial peak discharges (Prudhomme et al., 2002), RCMs are expected to provide an added value in the frequency distribution of local weather anomalies (Laprise, 2008). Unlike earlier assessments (Bevacqua et al., 2019; Kew et al., 2013), in which no validation was performed, our validation, comparing dynamically downscaled regional climate model simulations with observations, reveals an overestimation in the strength of dependence between the compound flood drivers. This could be a consequence of the models' structural uncertainties and limitations in their parameterizations (Fowler et al., 2005). This overestimation should motivate careful evaluations of future assessments on how well model simulations agree with observations.

A few caveats should be considered so that our results are not overinterpreted. Although a recent study (Zeppetello et al., 2019) has shown that the downwelling longwave radiation response to greenhouse

forcing is controlled primarily by an increase in surface temperature, due to a lack of available RCM projections, we assume that the future downwelling longwave and shortwave radiation is the same as in the historical period. Second, due to the limited number of available ensemble members, we limit our analysis to a 30-year window in the future using a single RCM, RCA4 driven by a limited number of host-GCMs barring us to assess the full spectrum of multimodel response variability and the internal climate variability (Deser et al., 2020), which contributes significantly to projection uncertainty, especially at regional scales. Finally, the spatial resolution of WGHM is limited to 0.5°; however, this has been widely used to explore hydrologic response to climate change at a continental scale (Gosling et al., 2017; Lobanova et al., 2018; Prudhomme et al., 2014).

Global warming is expected to affect compound flood hazard in different ways: through rising sea levels, changes in precipitation and river discharge, and changes in storm activity, affecting the magnitude and timing of these flood drivers and the interactions between them. Our climate projection analysis demonstrates the intricate nature of changes in compound floods: Changes in surge elevations are expected to be small with no clear trend in most of the locations, whereas trends in river peaks tend to increase for northwestern Europe. The dependence strength between surges and river peaks shows a clear downward change, decreasing the compound flood hazard, whereas SLR is a strong driver of increasing compound floods. When these drivers and their interactions are acting together, we find mostly a decrease in coastal compound flood hazard in northwestern Europe. The substantial role of SLR in future compound flood hazard is in agreement with an earlier study for New York City (Garner et al., 2017), where the projected SLR dominated the overall increase in coastal flood levels, hence increasing coastal flood hazards, whereas the storm intensity and shifts in storm tracks showed minimal changes. However, this study was solely focused on coastal flood hazards and did not investigate trends in compound floods and the role of SLR in modulating compound flood hazard. In future research, the proposed modeling framework could be combined with hydrodynamic inundation modeling to assess the inundation extent and depth in the low-lying delta regions. Another potential research avenue is to integrate projections of compound flood hazards with expected damages and possible adaptation costs to inform risk reduction decisions (Kreibich et al., 2017; Rosner et al., 2014). Although the present work is applied to the locations of tide and stream gauges in northwestern Europe, the methodology can be extended to any geographical region across the globe and over finer-resolution data grids. The findings are based on state-of-the-art regional climate models and available greenhouse gas emission scenarios but may be subjected to change with new scenarios and socioeconomic pathways in the future. Given the cascade of uncertainty across model chains, the results presented herein should be interpreted with caution. Nevertheless, the obtained insights will assist modelers to identify model deficiencies to reduce uncertainty associated with projected changes in coastal compound floods. A close coordination between regional climate modelers and hydrologists is required to further improve model parameterizations, so that the coupled modeling framework is capable of capturing hydrologically relevant metrics for end-user applications.

Data Availability Statement

Observed daily streamflow and hourly sea level data are obtained from GRDC data center, (https://www.bafg.de/GRDC/EN/01_GRDC/13_dtbse/database_node.html) and GESLA (http://gesla.org/) website. The high-resolution EURO-CORDEX data set used in this study was downloaded from the Earth System Grid Federation (ESGF) website: https://esg-dn1.nsc.liu.se/search/esgf-liu/ (as accessed in June 2019). The bias-adjusted RCM data for precipitation and temperature were downloaded from the European Commission Joint Research Centre website (https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/LISCOAST/10011/LATEST/EURO-CORDEX/EUR-11/SMHI-RCA4/) (as accessed in June 2019). The Delft3D code is freely available at this site (https://oss.deltares.nl/web/delft3d). The GEV analysis with parameter estimation in a Bayesian framework was performed using the MATLAB-based NEVA toolbox, available at the University of California, Irvine website (http://amir.eng.uci.edu/neva.php). The code for generating localized sea-level rise projections is available in the LocalizeSL (https://github.com/bobkopp/LocalizeSL). The time series of the hourly surge and the daily river discharge developed during this study are archived at the GFZ data services website with this identifier (http://doi.org/10.5880/GFZ.4.4.2019.003). Codes used for this work are available from the authors upon request.

Acknowledgments

Dr. P. Ganguli, is a recipient of the *Humboldt Research Fellowship* for early career researchers from Alexander von Humboldt Foundation, Germany. Support from the foundation and the host institute—GFZ German Research Centre for Geosciences, Potsdam, Germany is deeply acknowledged. The first author of the manuscript would like to thank Dr. Thomas Wahl at the University of Central Florida and Dr. Ivan D Haigh at the University of Southampton for their fruitful suggestions on the computation of skew surge from tide gauge records. The first author would like to thank Dr. Alessandro Dosio at the European Commission Joint Research Centre for his suggestions on implementation of dynamically downscaled bias-adjusted precipitation and temperature data at a WGHM scale as well as Dr. Robert Kopp of Rutgers University and Dr. Ning Lin of Princeton University for helpful correspondence on implementation of localized SLR. Open access funding enabled and organized by Projekt DEAL.

References

Berghuijs, W. R., Allen, S. T., Harrigan, S., & Kirchner, J. W. (2019). Growing spatial scales of synchronous river flooding in Europe. *Geophysical Research Letters*, *46*, 1423–1428. https://doi.org/10.1029/2018GL081883

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

Bloemendaal, N., Muis, S., Haarsma, R. J., Verlaan, M., Irazoqui Apecechea, M., de Moel, H., et al. (2019). Global modeling of tropical cyclone storm surges using high-resolution forecasts. *Climate Dynamics*, *52*(7–8), 5031–5044. https://doi.org/10.1007/s00382-018-4430-x

Bosshard, T., Kotlarski, S., Zappa, M., & Schär, C. (2014). Hydrological climate-impact projections for the Rhine River: GCM–RCM uncertainty and separate temperature and precipitation effects. *Journal of Hydrometeorology*, *15*(2), 697–713.

Codiga, D. L. (2011). Unified tidal analysis and prediction using the UTide Matlab functions. Graduate School of Oceanography, University of Rhode Island Narragansett, RI.

Couasnon, A., Eilander, D., Muis, S., Veldkamp, T. I. E., Haigh, I. D., Wahl, T., et al. (2020). Measuring compound flood potential from river discharge and storm surge extremes at the global scale. *Natural Hazards and Earth System Sciences*, *20*(2), 489–504. https://doi.org/10.5194/nhess-20-489-2020

Delft3D-FLOW, D (2014). *Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments* (p. 690). The Netherlands: User Manual Delft3D-FLOW.

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., et al. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286. https://doi.org/10.1038/s41558-020-0731-2

Devlin, A. T., Jay, D. A., Talke, S. A., Zaron, E. D., Pan, J., & Lin, H. (2017). Coupling of sea level and tidal range changes, with implications for future water levels. *Scientific Reports*, *7*(1), 17021.

Döll, P., Kaspar, F., & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: Model tuning and validation. *Journal of Hydrology*, *270*(1–2), 105–134.

Dosio, A. (2016). Projections of climate change indices of temperature and precipitation from an ensemble of bias-adjusted high-resolution EURO-CORDEX regional climate models. *Journal of Geophysical Research: Atmospheres*, *121*, 5488–5511. https://doi.org/10.1002/2015JD024411

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

Fowler, H. J., Ekström, M., Kilsby, C. G., & Jones, P. D. (2005). New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. *Journal of Hydrology*, *300*(1–4), 212–233. https://doi.org/10.1016/j.jhydrol.2004.06.017

Frahm, G., Junker, M., & Schmidt, R. (2005). Estimating the tail-dependence coefficient: Properties and pitfalls. *Insurance: Mathematics & Economics*, *37*(1), 80–100.

Ganguli, P., & Merz, B. (2019a). Extreme coastal water levels exacerbate fluvial flood hazards in northwestern Europe. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-49822-6

Ganguli, P., & Merz, B. (2019b). Trends in compound flooding in northwestern Europe during 1901–2014. *Geophysical Research Letters*, *46*, 10,810–10,820. https://doi.org/10.1029/2019GL084220

Garner, A. J., Mann, M. E., Emanuel, K. A., Kopp, R. E., Lin, N., Alley, R. B., et al. (2017). Impact of climate change on New York City's coastal flood hazard: Increasing flood heights from the preindustrial to 2300 CE. *Proceedings of the National Academy of Sciences*, *114*(45), 11,861–11,866. https://doi.org/10.1073/pnas.1703568114

Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics & Economics*, *44*(2), 199–213.

Gonzalez, P. L. M., Brayshaw, D. J., & Zappa, G. (2019). The contribution of North Atlantic atmospheric circulation shifts to future wind speed projections for wind power over Europe. *Climate Dynamics*, *53*(7–8), 4095–4113. https://doi.org/10.1007/s00382-019-04776-3

Gosling, S. N., Zaherpour, J., Mount, N. J., Hattermann, F. F., Dankers, R., Arheimer, B., et al. (2017). A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1 °C, 2 °C and 3 °C. *Climatic Change*, *141*(3), 577–595. https://doi.org/10.1007/s10584-016-1773-3

Grabs, W. (1997). Report on the third meeting of the GRDC steering committee: Koblenz. Germany.

Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., et al. (2012). Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. *Journal of Hydrometeorology*, *13*(2), 604–620. https://doi.org/10.1175/JHM-D-11-083.1

Hagedorn, R., Doblas-Reyes, F. J., & Palmer, T. N. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus A: Dynamic Meteorology and Oceanography*, *57*(3), 219–233.

Hoitink, A. J. F., & Jay, D. A. (2016). Tidal river dynamics: Implications for deltas. *Reviews of Geophysics*, *54*, 240–272. https://doi.org/10.1002/2015RG000507

Ikeuchi, H., Hirabayashi, Y., Yamazaki, D., Kiguchi, M., Koirala, S., Nagano, T., et al. (2015). Modeling complex flow dynamics of fluvial floods exacerbated by sea level rise in the Ganges–Brahmaputra–Meghna Delta. *Environmental Research Letters*, *10*(12), 124,011. https://doi.org/10.1088/1748-9326/10/12/124011

Ikeuchi, H., Hirabayashi, Y., Yamazaki, D., Muis, S., Ward, P. J., Winsemius, H. C., et al. (2017). Compound simulation of fluvial floods and storm surges in a global coupled river-coast flood model: Model development and its application to 2007 cyclone Sidr in Bangladesh. *Journal of Advances in Modeling Earth Systems*, *9*, 1847–1862. https://doi.org/10.1002/2017MS000943

Kang, T.-H., Kim, Y.-O., & Hong, I.-P. (2010). Comparison of pre- and post-processors for ensemble streamflow prediction. *Atmospheric Science Letters*, *11*(2), 153–159. https://doi.org/10.1002/asl.276

Kendon, M., Marsh, T., & Parry, S. (2013). The 2010–2012 drought in England and Wales. *Weather*, *68*(4), 88–95.

Kew, S. F., Selten, F. M., Lenderink, G., & Hazeleger, W. (2013). The simultaneous occurrence of surge and discharge extremes for the Rhine delta. *Natural Hazards and Earth System Sciences*, *13*(8), 2017–2029.

Kjellström, E., Nikulin, G., Strandberg, G., Christensen, O. B., Jacob, D., Keuler, K., et al. (2018). European climate change at global mean temperature increases of 1.5 and 2°C above pre-industrial conditions as simulated by the EURO-CORDEX regional climate models. *Earth System Dynamics*, *9*(2), 459–478. https://doi.org/10.5194/esd-9-459-2018

Klerk, W. J., Winsemius, H. C., Verseveld, W. J., van Bakker, A. M. R., & Diermanse, F. L. M. (2015). The co-incidence of storm surges and extreme discharges within the Rhine–Meuse Delta. *Environmental Research Letters*, *10*(3), 035005. https://doi.org/10.1088/1748-9326/10/3/035005

Kopp, R. E., Horton, R. M., Little, C. M., Mitrovica, J. X., Oppenheimer, M., Rasmussen, D. J., et al. (2014). Probabilistic 21st and 22nd century sea-level projections at a global network of tide-gauge sites. *Earth's Future*, *2*, 383–406. https://doi.org/10.1002/2014EF000239

Kreibich, H., Botto, A., Merz, B., & Schröter, K. (2017). Probabilistic, multivariable flood loss modeling on the mesoscale with BT-FLEMO. *Risk Analysis*, *37*(4), 774–787. https://doi.org/10.1111/risa.12650

Kulp, S. A., & Strauss, B. H. (2019). New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. *Nature Communications*, *10*(1), 1–12.

Laprise, R. (2008). Regional climate modelling. *Journal of Computational Physics*, *227*(7), 3641–3666.

Lee, H., & Ghosh, S. K. (2009). Performance of information criteria for spatial models. *Journal of Statistical Computation and Simulation*, *79*(1), 93–106. https://doi.org/10.1080/00949650701611143

Lian, J., Xu, K., & Ma, C. (2013). Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network: A case study of Fuzhou City, China. *Hydrology and Earth System Sciences*, *17*(2), 679.

Lin, N., Kopp, R. E., Horton, B. P., & Donnelly, J. P. (2016). Hurricane Sandy's flood frequency increasing from year 1800 to 2100. *Proceedings of the National Academy of Sciences*, *113*(43), 12,071–12,075.

Lin, N., Marsooli, R., & Colle, B. A. (2019). Storm surge return levels induced by mid-to-late-twenty-first-century extratropical cyclones in the northeastern United States. *Climatic Change*, *154*(1–2), 143–158.

Lin, N., & Shullman, E. (2017). Dealing with hurricane surge flooding in a changing environment: Part I. Risk assessment considering storm climatology change, sea level rise, and coastal development. *Stochastic Environmental Research and Risk Assessment*, *31*(9), 2379–2400.

Lobanova, A., Liersch, S., Nunes, J. P., Didovets, I., Stagl, J., Huang, S., et al. (2018). Hydrological impacts of moderate and high-end climate change across European river basins. *Journal of Hydrology: Regional Studies*, *18*, 15–30. https://doi.org/10.1016/j.ejrh.2018.05.003

Losada, I. J., Reguero, B. G., Méndez, F. J., Castanedo, S., Abascal, A. J., & Mínguez, R. (2013). Long-term changes in sea-level components in Latin America and the Caribbean. *Global and Planetary Change*, *104*, 34–50. https://doi.org/10.1016/j.gloplacha.2013.02.006

Lowe, J. A., Gregory, J. M., & Flather, R. A. (2001). Changes in the occurrence of storm surges around the United Kingdom under a future climate scenario using a dynamic storm surge model driven by the Hadley Centre climate models. *Climate Dynamics*, *18*(3–4), 179–188.

Maraun, D. (2016). Bias correcting climate change simulations-a critical review. *Current Climate Change Reports*, *2*(4), 211–220.

Marsooli, R., Lin, N., Emanuel, K., & Feng, K. (2019). Climate change exacerbates hurricane flood hazards along US Atlantic and gulf coasts in spatially varying patterns. *Nature Communications*, *10*(1), 1–9.

Martins, E. S., & Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, *36*(3), 737–744.

McMaster, I. (2016). North Sea in Numbers: North Sea Region 2020 (pp. 1–69). University of Strathclyde, Glasgow, UK: European Policies Research Centre. Retrieved from https://cpmr-northsea.org/download/north-sea-in-numbers-north-sea-region-2020/?wpdmdl=1957&ind=1537795964963

Merz, B., Dung, N. V., Apel, H., Gerlitz, L., Schröter, K., Steirou, E., & Vorogushyn, S. (2018). Spatial coherence of flood-rich and flood-poor periods across Germany. *Journal of Hydrology*, *559*, 813–826.

Moftakhari, H., Schubert, J. E., AghaKouchak, A., Matthew, R. A., & Sanders, B. F. (2019). Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries. *Advances in Water Resources*, *128*, 28–38. https://doi.org/10.1016/j.advwatres.2019.04.009

Moftakhari, H. R., Salvadori, G., AghaKouchak, A., Sanders, B. F., & Matthew, R. A. (2017). Compounding effects of sea level rise and fluvial flooding. *Proceedings of the National Academy of Sciences*, *114*(37), 9785–9790.

Mohanty, M. P., Sherly, M. A., Ghosh, S., & Karmakar, S. (2020). Tide-rainfall flood quotient: An incisive measure of comprehending a region's response to storm-tide and pluvial flooding. *Environmental Research Letters*, *15*(6), 064029. https://doi.org/10.1088/1748-9326/ab8092

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., et al. (2020). The global water resources and use model WaterGAP v2.2d: Model description and evaluation. *Geoscientific Model Development Discussions*, 1–69. https://doi.org/10.5194/gmd-2020-225

Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C., & Ward, P. J. (2016). A global reanalysis of storm surges and extreme sea levels. *Nature Communications*, *7*, 11969.

Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Flörke, M., & Döll, P. (2014). Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration. *Hydrology and Earth System Sciences*, *18*(9), 3511–3538.

Nicholls, R. J., Marinova, N., Lowe, J. A., Brown, S., Vellinga, P., de Gusmão, D., et al. (2011). Sea-level rise and its possible impacts given a 'beyond 4°C world' in the twenty-first century. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1934), 161–181. https://doi.org/10.1098/rsta.2010.0291

Paprotny, D., Morales-Nápoles, O., Vousdoukas, M. I., Jonkman, S. N., & Nikulin, G. (2019). Accuracy of pan-European coastal flood mapping. *Journal of Flood Risk Management*, *12*(2), e12459. https://doi.org/10.1111/jfr3.12459

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

Petroliagkis, T. I. (2018). Estimations of statistical dependence as joint return period modulator of compound events – Part 1: Storm surge and wave height. *Natural Hazards and Earth System Sciences*, *18*(7), 1937–1955. https://doi.org/10.5194/nhess-18-1937-2018

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., et al. (2016). Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: High resolution, high benefits? *Climate Dynamics*, *46*(1–2), 383–412. https://doi.org/10.1007/s00382-015-2589-y

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., et al. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, *53*, 323–361. https://doi.org/10.1002/2014RG000475

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., et al. (2014). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences*, *111*(9), 3262–3267. https://doi.org/10.1073/pnas.1222473110

Prudhomme, C., Reynard, N., & Crooks, S. (2002). Downscaling of global climate models for flood frequency analysis: Where are we now? *Hydrological Processes*, *16*(6), 1137–1150. https://doi.org/10.1002/hyp.1054

Púčik, T., Groenemeijer, P., Rädler, A. T., Tijssen, L., Nikulin, G., Prein, A. F., et al. (2017). Future changes in European severe convection environments in a regional climate model ensemble. *Journal of Climate*, *30*(17), 6771–6794. https://doi.org/10.1175/JCLI-D-16-0777.1

Rajczak, J., & Schär, C. (2017). Projections of future precipitation extremes over Europe: A multimodel assessment of climate simulations. *Journal of Geophysical Research: Atmospheres*, *122*, 10,773–10,800. https://doi.org/10.1002/2017JD027176

Rosner, A., Vogel, R. M., & Kirshen, P. H. (2014). A risk-based approach to flood management decisions in a nonstationary world. *Water Resources Research*, *50*, 1928–1942. https://doi.org/10.1002/2013WR014561

Roy, T., Serrat-Capdevila, A., Gupta, H., & Valdes, J. (2017). A platform for probabilistic multimodel and multiproduct Streamflow forecasting. *Water Resources Research*, *53*, 376–399. https://doi.org/10.1002/2016WR019752

Santiago-Collazo, F. L., Bilskie, M. V., & Hagen, S. C. (2019). A comprehensive review of compound inundation models in low-gradient coastal watersheds. *Environmental Modelling & Software*, *119*, 166–181. https://doi.org/10.1016/j.envsoft.2019.06.002

Schneider, C., Laizé, C. L. R., Acreman, M. C., & Flörke, M. (2013). How will climate change modify river flow regimes in Europe? *Hydrology and Earth System Sciences*, *17*(1), 325–339. https://doi.org/10.5194/hess-17-325-2013

Schoetter, R., Hoffmann, P., Rechid, D., & Schlünzen, K. H. (2012). Evaluation and bias correction of regional climate model results using model evaluation measures. *Journal of Applied Meteorology and Climatology*, *51*(9), 1670–1684.

Sterl, A., Brink, H., van den Vries, H., de Haarsma, R., & van Meijgaard, E. (2009). An ensemble study of extreme storm surge related water levels in the North Sea in a changing climate. *Ocean Science*, *5*(3), 369–378.

Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., et al. (2013). *Climate change 2013: The physical science basis. Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)*. New York: Cambridge University Press.

Strandberg, G., Bärring, L., Hansson, U., Jansson, C., Jones, C., Kjellström, E., et al. (2015). CORDEX scenarios for Europe from the Rossby Centre regional climate model RCA4. SMHI.

Svensson, C., & Jones, D. A. (2002). Dependence between extreme sea surge, river flow and precipitation in eastern Britain. *International Journal of Climatology*, *22*(10), 1149–1168.

Svensson, C., & Jones, D. A. (2004). Dependence between sea surge, river flow and precipitation in South and West Britain. *Hydrology and Earth System Sciences*, *8*(5), 973–992.

Tebaldi, C., & Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 2053–2075.

Teutschbein, C., & Seibert, J. (2013). Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrology and Earth System Sciences*, *17*(12), 5061–5077.

Thorarinsdottir, T. L., Gneiting, T., & Gissibl, N. (2013). Using proper divergence functions to evaluate climate models. *SIAM/ASA Journal on Uncertainty Quantification*, *1*(1), 522–534.

Tu, X., Du, Y., Singh, V. P., & Chen, X. (2018). Joint distribution of design precipitation and tide and impact of sampling in a coastal area. *International Journal of Climatology*, *38*, e290–e302.

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

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., & Slangen, A. B. (2017). Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nature Communications*, *8*, 16075.

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

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

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-interim reanalysis data. *Water Resources Research*, *50*, 7505–7514. https://doi.org/10.1002/2014WR015638

Woodworth, P. L., Hunter, J. R., Marcos, M., Caldwell, P. C., Menendez, M., & Haigh, I. D. (2016). GESLA (global extreme sea level analysis) high frequency sea level dataset - version 2 [data set]. Retrieved from https://www.bodc.ac.uk/resources/inventories/edmed/report/6562/

Yang, X., Wood, E. F., Sheffield, J., Ren, L., Zhang, M., & Wang, Y. (2018). Bias correction of historical and future simulations of precipitation and temperature for China from CMIP5 models. *Journal of Hydrometeorology*, *19*(3), 609–623.

Zeppetello, L. R. V., Donohoe, A., & Battisti, D. S. (2019). Does surface temperature respond to or determine downwelling longwave radiation? *Geophysical Research Letters*, *46*, 2781–2789. https://doi.org/10.1029/2019GL082220

Zhang, F., Lin, B., & Sun, J. (2019). Current reversals in a large tidal river. *Estuarine, Coastal and Shelf Science*, *223*, 74–84. https://doi.org/10.1016/j.ecss.2019.04.017

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

Zijl, F., Muis, S., Verlaan, M., & Herman, P. (2019). Rapid decline in tidal amplitude along the Dutch coast. In Geophysical Research Abstracts (Vol. 21).

Zscheischler, J., & Seneviratne, S. I. (2017). Dependence of drivers affects risks associated with compound events. *Science Advances*, *3*(6), e1700263. https://doi.org/10.1126/sciadv.1700263