Supplement S1. Validation of annual maxima daily discharge and storm surge

For this study we select the simulated runoff from the JULES model, routed with the CaMa-Flood model based on the performance tests presented in Beck et al. (2017) and Schellekens et al. (2017). Here, we complement these tests by looking at the rank correlation coefficient and the absolute average lag in the timing of the annual maxima in observations with a record length of at least 20 years. We also apply similar performance tests for the storm surge variable.

We use the Spearman's rank correlation, a nonparametric measure for monotonic relationships between two variables. The rank correlation coefficient is equivalent to the Pearson's product moment correlation, ρ, applied to the ranks of the annual maxima both observed (X_o) and simulated (X_s), such that:

$$r_{X_o,X_s} = \rho(X_o, X_s)$$

We also calculate a simple metric often used in flood forecasting studies, the Hit Rate, H, but applied to the date of the annual maximum. This corresponds to the probability of detection of the date of the annual maxima. We assume that the simulated date of the annual maximum D_S^i is correctly represented if it is within ± 3 days of the observed annual maximum D_O^i in the i-th year considered:

$$H = \frac{\sum_{i=1}^{N}(D_S^i \cap D_O^i)}{N}$$

S1.1 Annual maxima of daily discharge

We compare the performance of the modelled annual maxima of daily discharge with discharge observations from the Global Runoff Data Base (GRDB) from the Global Runoff Data Centre[1] . We follow a similar procedure as described in Zhao et al. (2017) to select stations in near-natural areas, and therefore minimise anthropogenic influence on the measured discharge. A catchment is selected if less than 2% of its upstream area is subject to irrigation, if the total reservoir capacity in the catchment is less than 10% of its long-term mean annual discharge, if its catchment area is at least 1000 km^2 or higher, and if the record length is at least 20 years with a minimum completeness of 75% per year within the period 1980-2014. This leads to the selection of 1116 stations, shown in Figure S1 and S2. The timing of the simulated discharge annual maxima compared with observations varies greatly globally. We find a median hit rate of 0.21 (min:0, max:0.79, s.d.:0.18) and a median rank correlation coefficient is 0.57 (min: -0.35, max: 0.96, s.d.: 0.22).

[1] The Global Runoff Data Centre, 56068 Koblenz, Germany www.bafg.de/GRDC/EN/Home/homepage_node.html

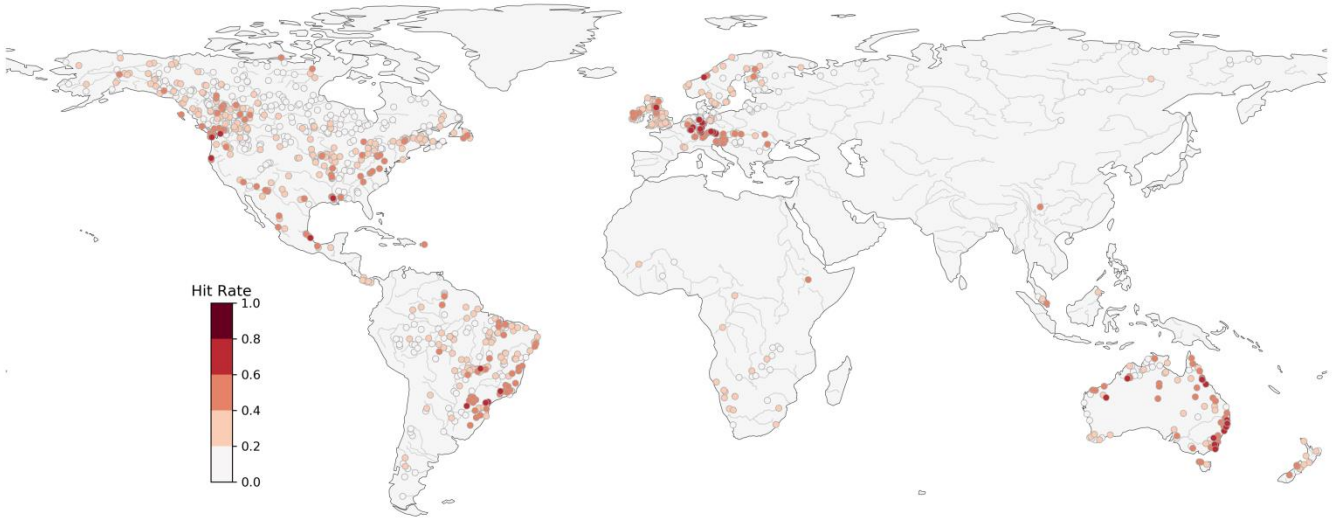

Figure S1: Probability of correctly detecting the date of the discharge annual maxima within ± 3 days.

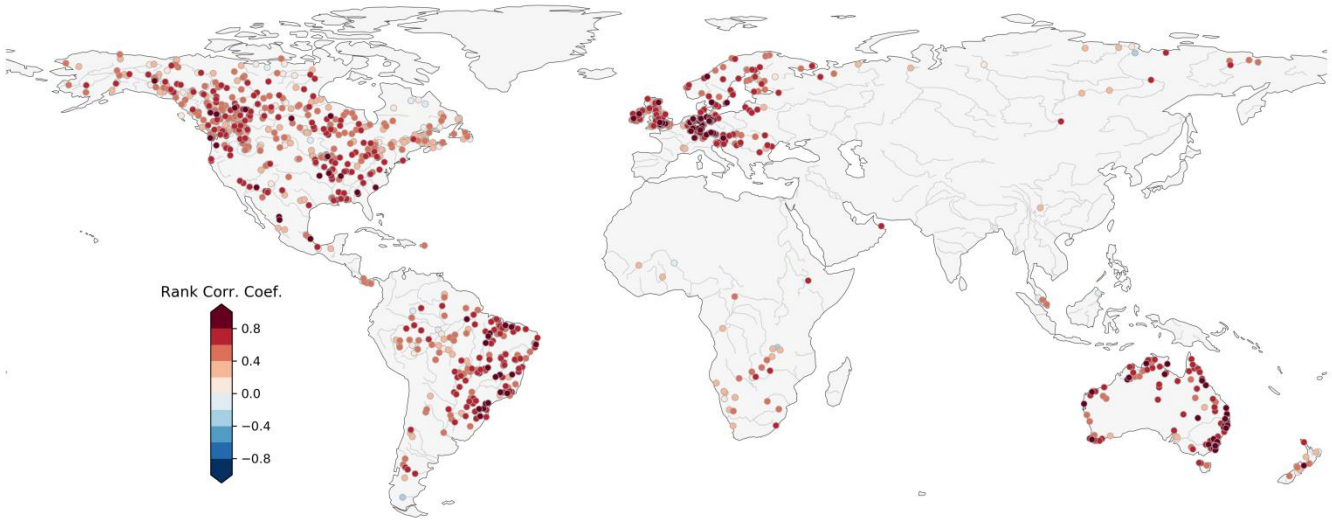

Figure S2: Spearman's rank correlation between the discharge annual maxima obtained from the model and from the observations.

S1.2 Annual maxima of storm surge

In order to compare the simulated storm surge variable with observations, we extract the equivalent of the storm surge from the sea levels observations of the Global Extreme Sea-level Analysis Version 2 database (GESLA-2) database (Woodworth et al., 2017). We select coastal stations if they have at least 20 years of data and a minimum completeness of 75% per year and compare it with the closest GTSM output location within a maximum radius of 20 km. This leads to the selection of 165

stations, shown in Figure S3 and S4. The timing of the simulated storm surge annual maxima compared with observations varies greatly globally. We find a median hit rate of 0.34 (min: 0, max: 0.70, s.d.: 0.22) and a median rank correlation coefficient of 0.37 (min: -0.45, max: 0.81, s.d.:0.31).

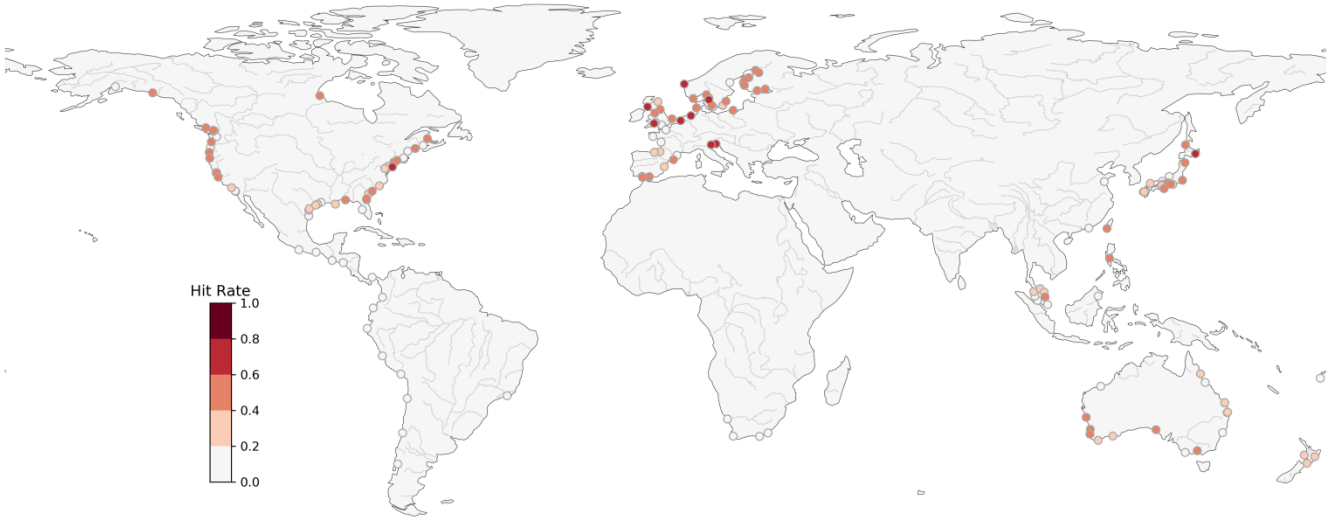

Figure S3: Probability of correctly detecting the date of the storm surge annual maxima.

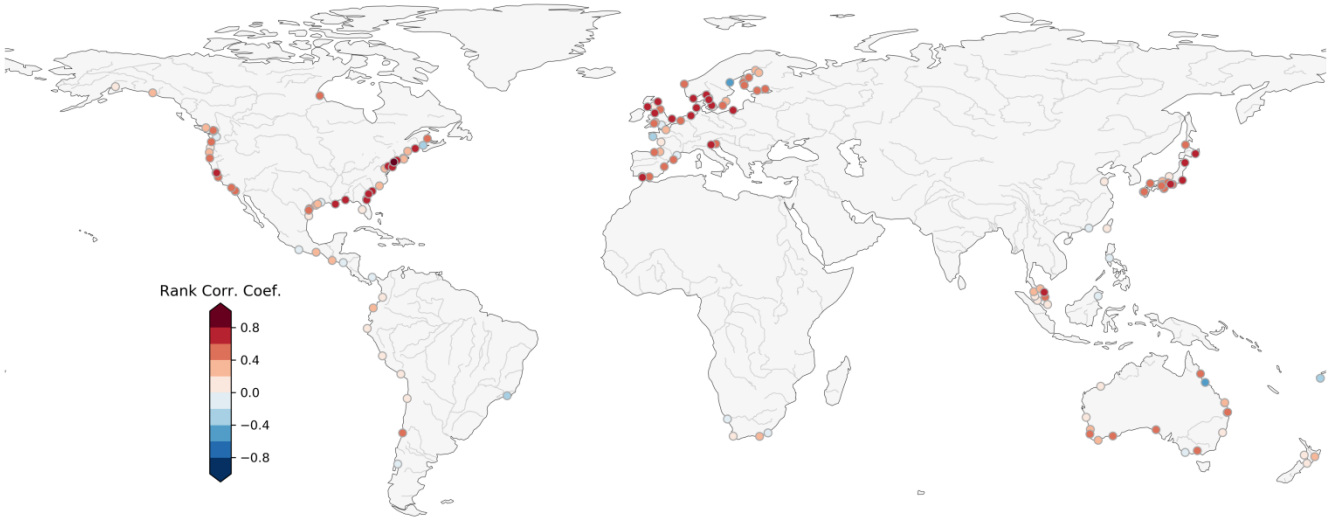

Figure S4: Spearman's rank correlation between the storm surge annual maxima obtained from the model and from the observations.

S1.3 Compound flood potential measure from simulated and observed time series

In order to assess how the performance of the global models presented in section S1.1 and S1.2 might affect the results presented in this study, we compare the covariability of the discharge and storm surge. To do so, we start with the combinations of discharge and tide stations presented in Ward et al. (2018) and filter this list to keep combinations that have at least 75% of overlapping data per year and a minimum of 20 years within the period 1980-2014. We only keep GESLA-2 tidal stations for which a GTSM output location is within a 20 km radius from the latter and GRDB discharge stations in near-natural areas as described in Zhao et al. (2017). This leads to the selection of 25 combinations of stations with the majority being located in Europe. Figure S5 presents the conditional dependence strength (Fig. S5a and S5b) and the total number of co-occurring annual maxima (Fig. S5c) obtained from the paired observations stations and the corresponding model output locations.

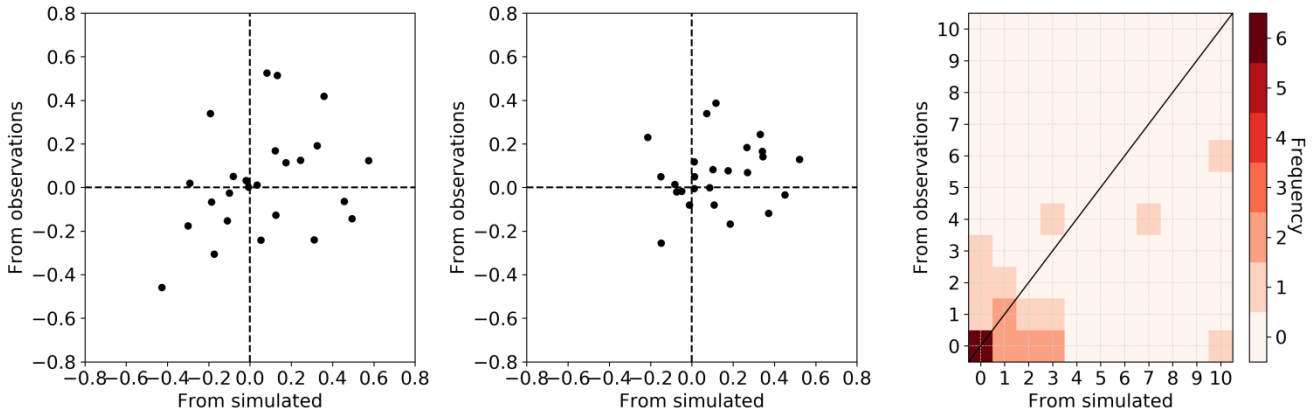

Figure S5: Comparison of the spearman's rank correlation between the storm surge conditional on annual maxima discharge (a), the discharge conditional on annual maxima storm surge (b) and the number of co-occurring annual maxima within a 3-day time window (c) obtained from the simulated discharge and storm surge variables (x-axis) and from observations (y-axis) .

Concerning the conditional dependence, we observe an overall positive agreement but with a large spread between the model outputs and the observations. This is particularly noticeable in Figures S6 and S7, which show the corresponding results at the paired stations. Some locations, such as in the south of England or the southwest of Australia, exhibit a similar dependence behavior as with the observations, albeit with some bias, while other locations, such as in Italy or the northwest coast of Australia, show opposite results. Nevertheless, we note that at most locations the model can capture the sign of the correlation correctly. Figure S8 indicates that, in general, the selected models tend to correctly identify the locations with the highest co-occurrences but overestimate the number of co-occurring annual maxima. An outlier is the northwest of Australia, where no co-occurring annual maxima are measured based on observations and 10 based on the simulated variables. These discrepancies could be due to the fact that the selected global models fail to capture important small-scale features and processes driving extreme discharge and storm surge at these locations.

(a) From observation stations

(b) From the simulated variables

Figure S6: Spearman's rank correlation between the storm surge conditional on annual maxima discharge obtained from observation stations (a), and from the simulated variables at these locations (b) .

(a) From observation stations

(b) From the simulated variables

Figure S7: Spearman's rank correlation between discharge conditional on annual maxima storm surge obtained from observation stations (a), and from the simulated variables at these locations (b) .

(a) From observation stations

(b) From the simulated variables

Figure S8: Number of co-occuring annual maxima of discharge and storm surge from observation stations (a), and from the simulated variables (b) .

Supplement S2. Co-occurrences of joint annual maxima at selected locations

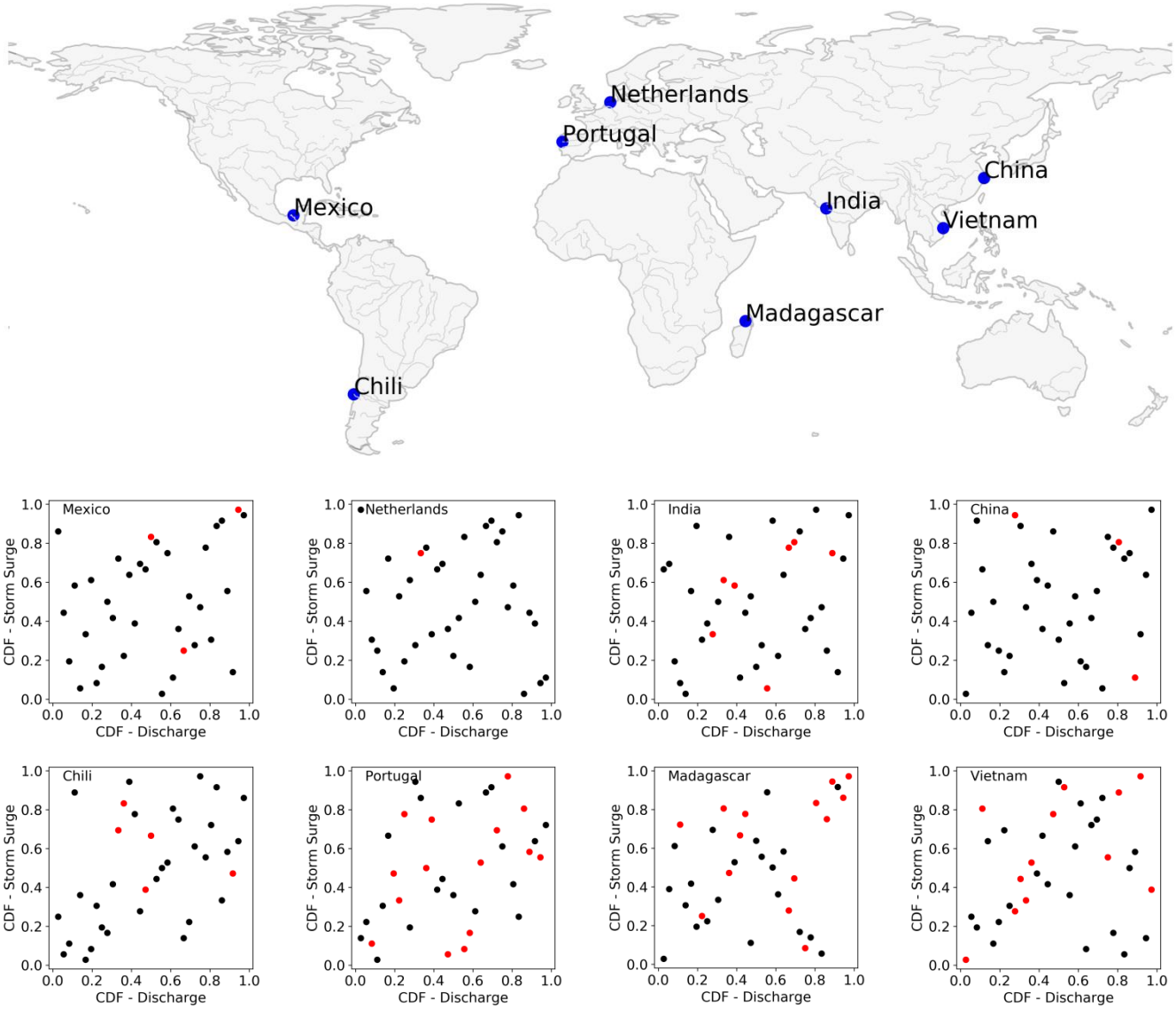

Figure S9: Examples of pseudo-observations from simulated annual maxima of discharge Q and storm surge S at selected locations. Red dots indicate a co-occurrence of Q and S, (Q^*, S^*), within a time lag of 3 days.

5

Supplement S3. Sensitivity of time window on Spearman's r_s correlation coefficient

Time window Δ (days)	Storm surge conditional on discharge annual maxima (Q_n, s_n)		Discharge conditional on storm surge annual maxima (S_n, q_n)	
	α = 0.05	α = 0.10	α = 0.05	α = 0.10
0	11	17	19	25
1	14	19	19	24
2	14	20	19	25
3	14	20	20	25
4	15	20	19	24
5	15	21	19	24
6	15	21	19	24
7	16	22	18	24

Table S1: Total percentage (%) of paired locations along the global coastline with a positive and statistically significant Spearman's rank correlation coefficient both for (Q_n, s_n) and (S_n, q_n) pairs and significance levels of α = 0.05 and α = 0.10.

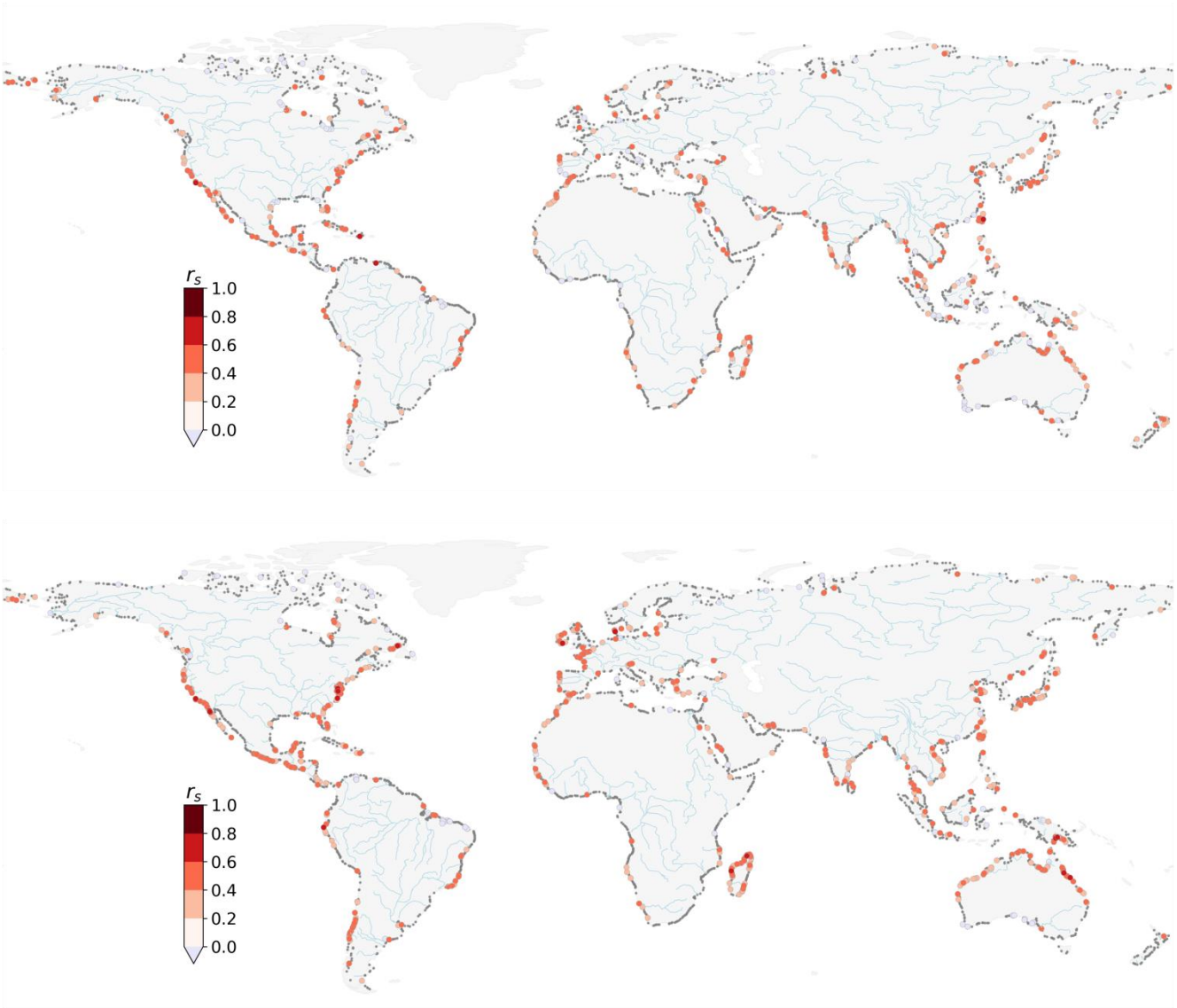

Figure S10: Spearman's r_s correlation coefficient between storm surge conditional on discharge annual maxima (Q_n, s_n) for a time window of Δ = 0 days (top) and Δ = 7 days (bottom). Black dots denote locations with no significant dependence (α = 0.05).

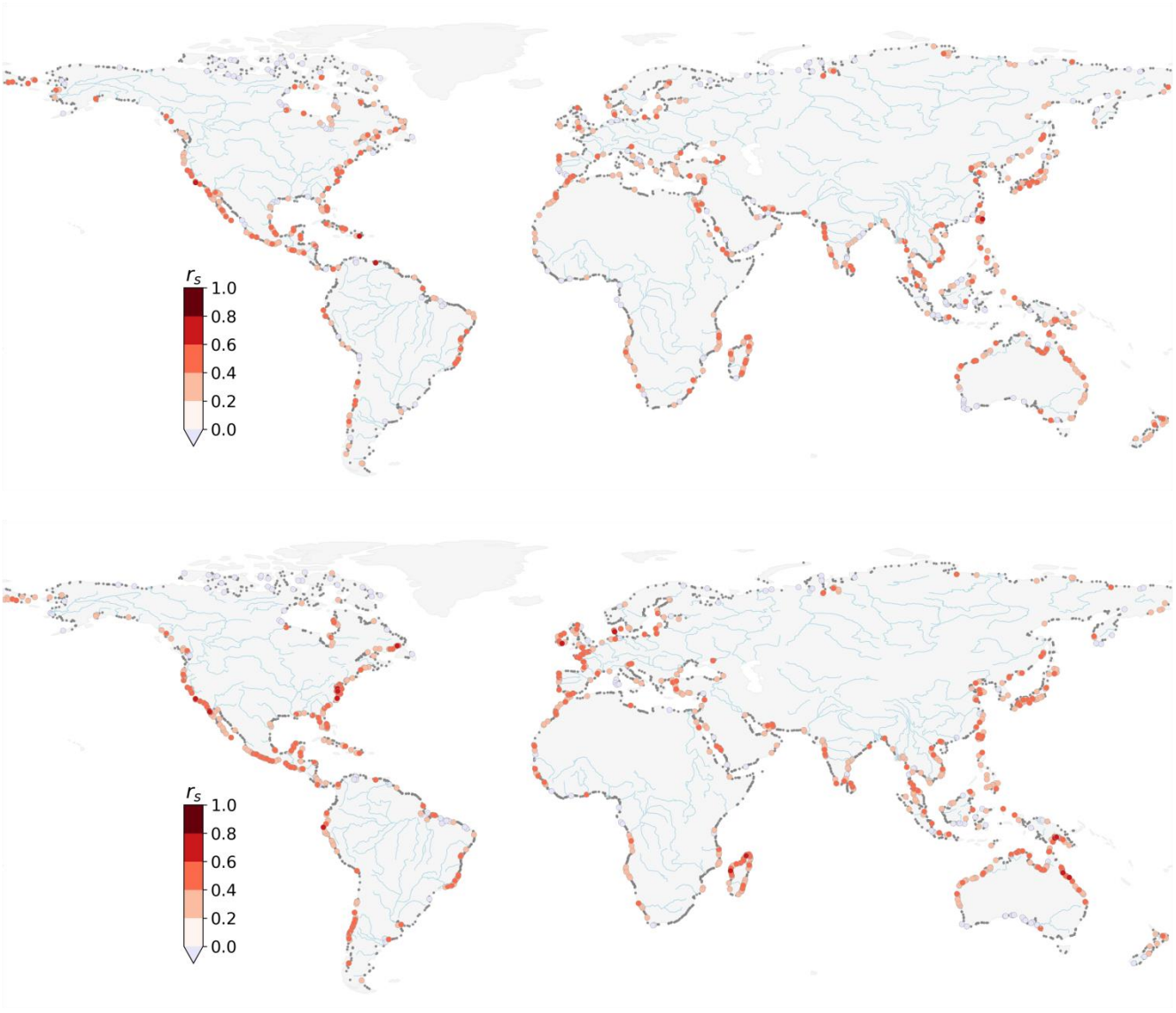

Figure S11: Spearman's r_s correlation coefficient between storm surge conditional on discharge annual maxima (Q_n, s_n) for a time window of Δ = 0 days (top) and Δ = 7 days (bottom). Black dots denote locations with no significant dependence (α = 0.10).

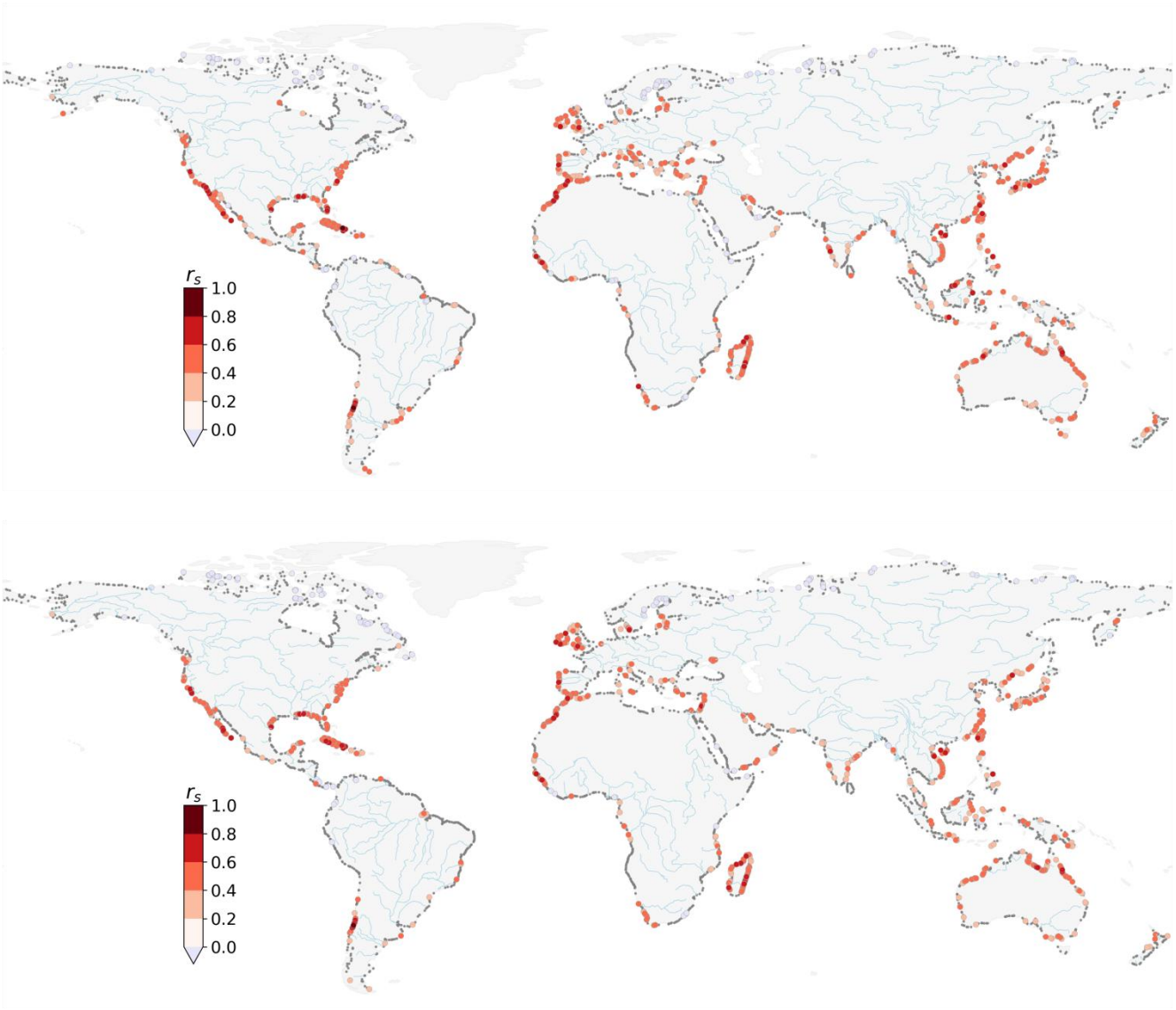

Figure S12: Spearman's r_s correlation coefficient between discharge conditional on storm surge annual maxima (S_n, q_n) for a time window of $\Delta = 0$ days (top) and $\Delta = 7$ days (bottom). Black dots denote locations with no significant dependence ($\alpha = 0.05$).

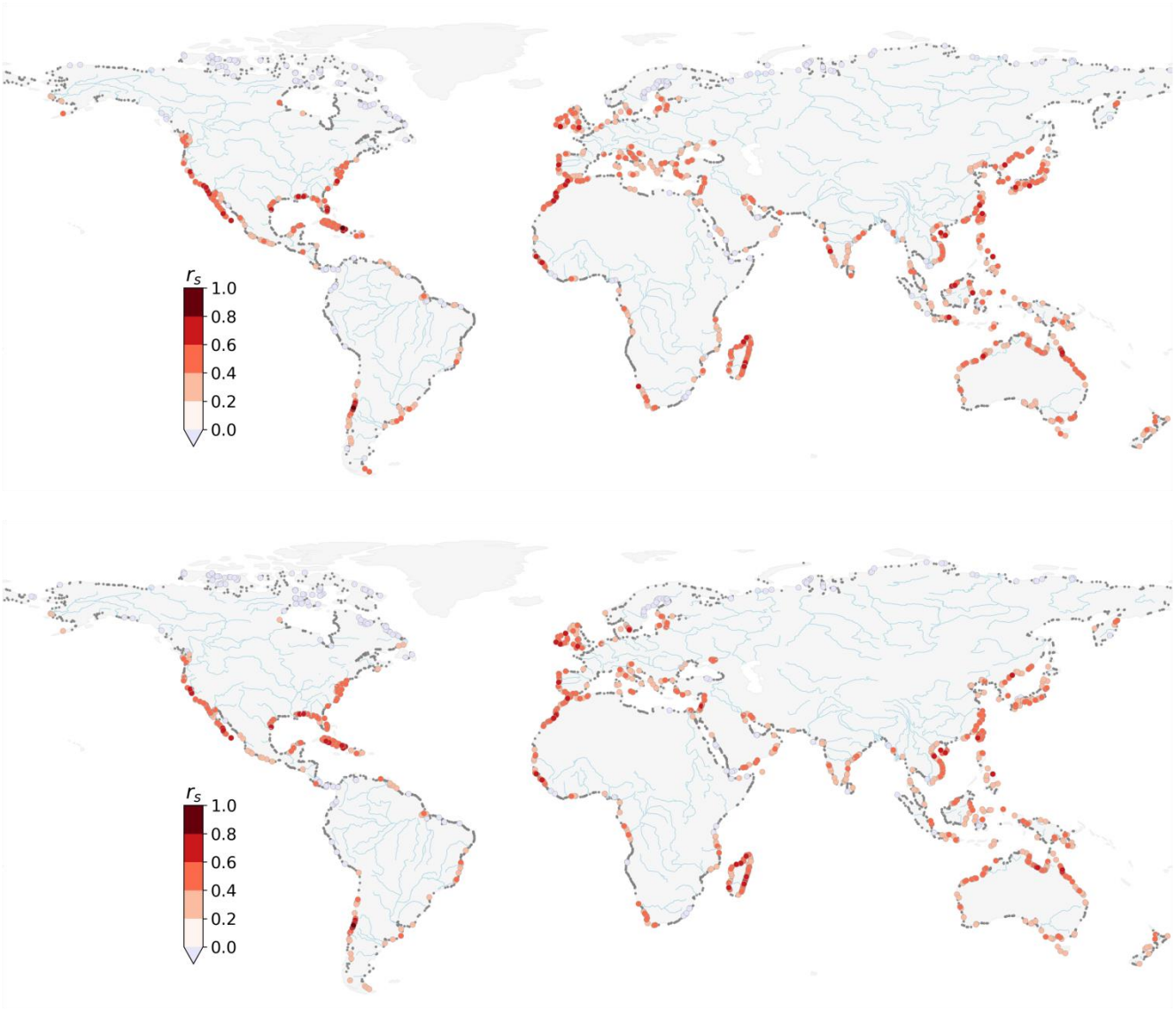

Figure S13: Spearman's r_s correlation coefficient between discharge conditional on storm surge annual maxima (S_n, q_n) for a time window of $\Delta = 0$ days (top) and $\Delta = 7$ days (bottom). Black dots denote locations with no significant dependence ($\alpha = 0.10$).

Supplement S4. Influence of the time window on the number of co-occurring annual maxima

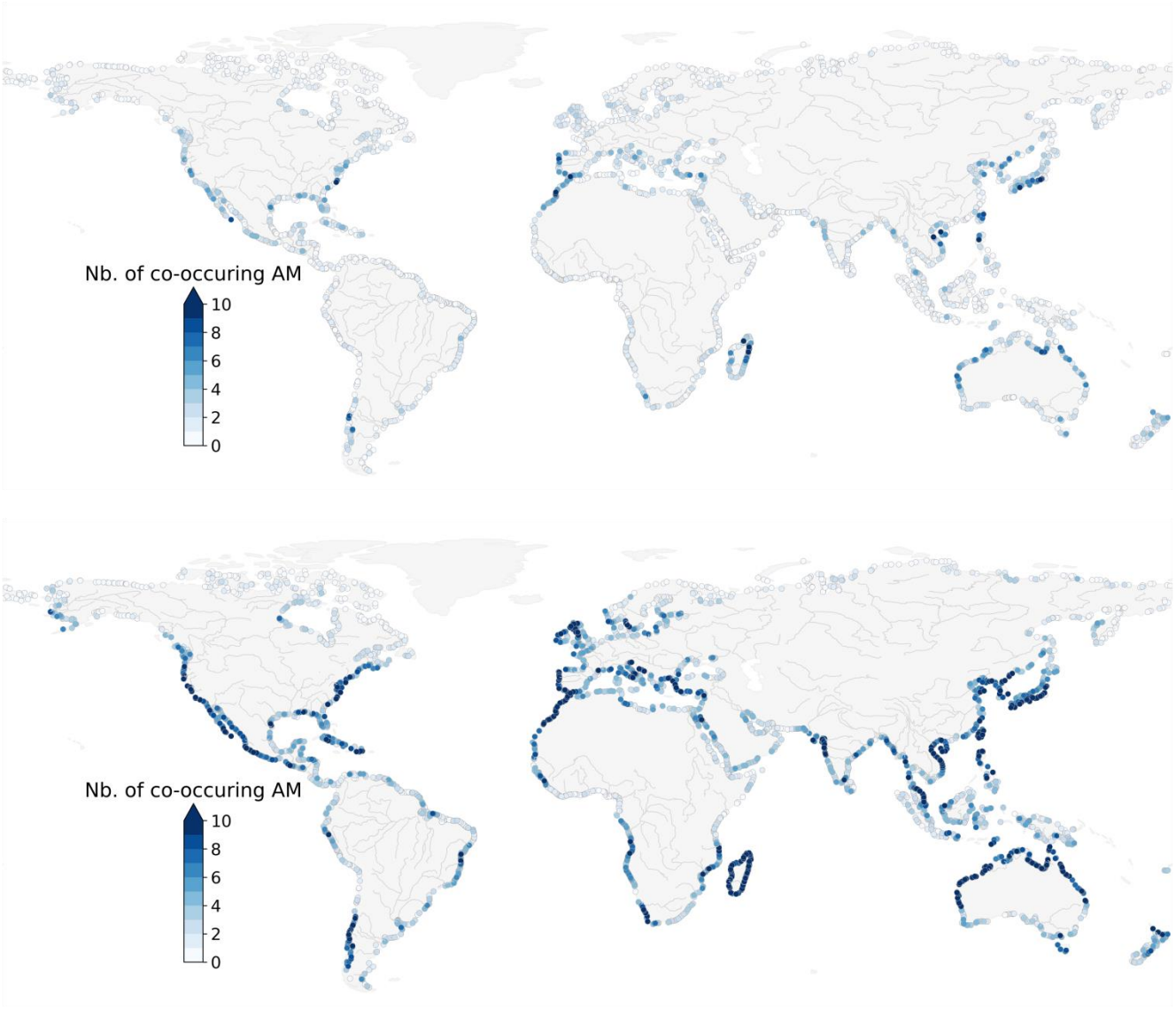

Figure S14: Number of co-occurring yearly maxima of storm surge and discharge obtained between 1980-2014 using a time window of Δ = 0 days (top) and Δ = 7 days (bottom).

5

References

Beck, H. E., Van Dijk, A. I. J. M., De Roo, A., Dutra, E., Fink, G., Orth, R. and Schellekens, J.: Global evaluation of runoff from ten state - of - the - art hydrological models, , 2881–2903, doi:10.5194/hess-2016-124, 2017.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., van Dijk, A., Sperna Weiland, F., Minvielle, M., Calvet, J.-C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W. and Weedon, G. P.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, Earth Syst. Sci. Data, 9(2), 389–413, doi:10.5194/essd-9-389-2017, 2017.

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, Environ. Res. Lett., 13(8), 084012, doi: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, Geosci. Data J., 3(2), doi:10.1002/gdj3.42, 2017.

Zhao, F., Veldkamp, T. I. E., Frieler, K., Schewe, J., Ostberg, S., Willner, S., Schauberger, B., Gosling, S. N., Schmied, H. M., Portmann, F. T., Leng, G., Huang, M., Liu, X., Tang, Q., Hanasaki, N., Biemans, H., Gerten, D., Satoh, Y., Pokhrel, Y., Stacke, T., Ciais, P., Chang, J., Ducharne, A., Guimberteau, M., Wada, Y., Kim, H. and Yamazaki, D.: The critical role of the routing scheme in simulating peak river discharge in global hydrological models, Environ. Res. Lett., 12(7), 075003, doi:10.1088/1748-9326/aa7250, 2017.