Supplement of

Impact-based flood forecasting in the Greater Horn of Africa

Lorenzo Alfieri et al.

Correspondence to: Lorenzo Alfieri (Lorenzo.Alfieri@cimafoundation.org)

Sensitivity of calibration parameters

We performed a global sensitivity analysis (GSA) on 8 Continuum parameters to investigate their sensitivity and the most influential parameters for each output variable. GSA was based on the SAFE Toolbox (Pianosi et al., 2015), using the Elementary Effects Test (EET) (or method of Morris) and One-at-a-time (OAT) sampling using Latin Hyper-cube. Continuum's parameters to perturb are summarized in Table S1.

Table S1: Sampled parameters of the Continuum model.

uc	Friction coefficient in channels
uh	Flow motion coefficient in hillslopes
cf	Infiltration capacity at saturation
ct	Field capacity
CN	Curve Number
ws	Water Sources
WTableHbr	Maximum water capacity of the aquifer
Fr	Fracturation

We run 900 model simulations, corresponding to 100 runs for each parameter, plus 100 needed to enable bootstrap analysis (based on 1000 samplings with 5% significance level). Sensitivity was assessed towards soil moisture, evaporation, and discharge (the latter both considering sensitivity to the model output and to the model performance). Results were diagnosed by assessing 1) the sensitivity index of the perturbed parameters, 2) mean versus standard deviation of each parameter, 3) scatter plot of the sampled parameters, 4) convergence and 5) behavioral runs. A sample of diagnostic plots is shown in Figure S1.
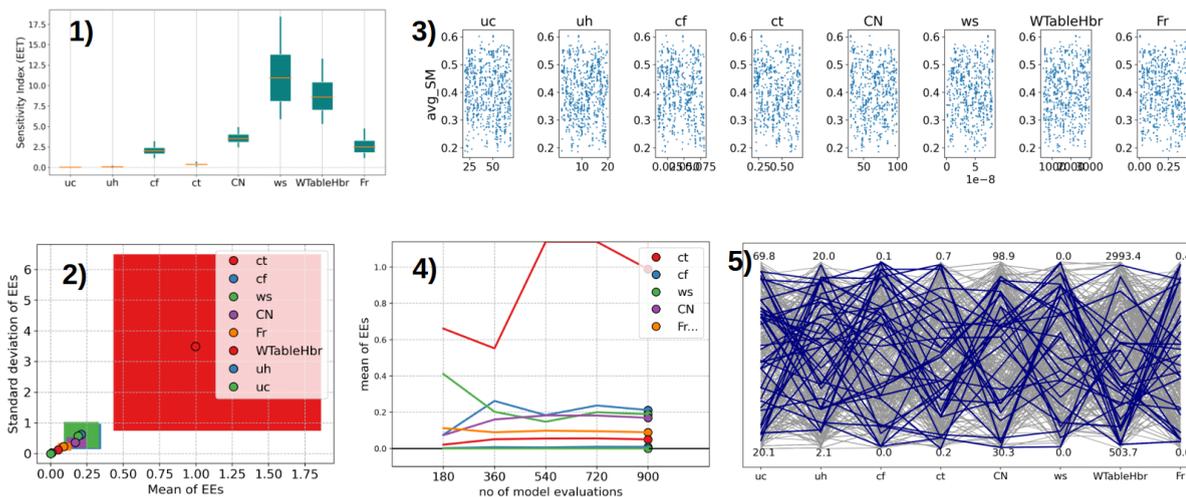
Figure S1: Diagnostic plots of the GSA: 1) sensitivity index of the perturbed parameters, 2) mean versus standard deviation of each parameter, 3) scatter plot of the sampled parameters, 4) convergence and 5) behavioral runs (here based on the threshold KGE >0.5).

Results of the GSA on the Continuum model can be summarized in the following key messages:

- Soil moisture is mainly influenced by the field capacity parameter (ct).

- Evaporation is mainly influenced by the Curve Number (CN) and by strong cross-parameter interactions with ct.

- The simulated flow metrics (with respect to observed values) show maximum sensitivity to the parameters ws and WtableHbr, followed by Fr and CN.

- CN has (usually) the greatest sensitivity in acting on advances or delays in the hydrogram, with influence also of other parameters.

- The analyses show a minimum sensitivity of the parameters uc and uh, followed by cf.

- Based on the analyses carried out, it is recommended to calibrate the parameters ct, CN, WtableHbr, Fr. This strategy was applied in the calibration of the hydrological domains in the GHA region, additionally by constraining such parameters within a physically meaningful range. Some minor modifications were then included in the choice of calibration parameters for selected cases following results of initial tests. For instance, in the Nile River basin we included the friction coefficient in channels as calibration parameter, given the larger extent of the river network and the increasing weight of river routing as compared to the case used for the sensitivity analysis. In Figure S2, key graphs on the sensitivity index of the perturbed parameters versus the three variables are shown.
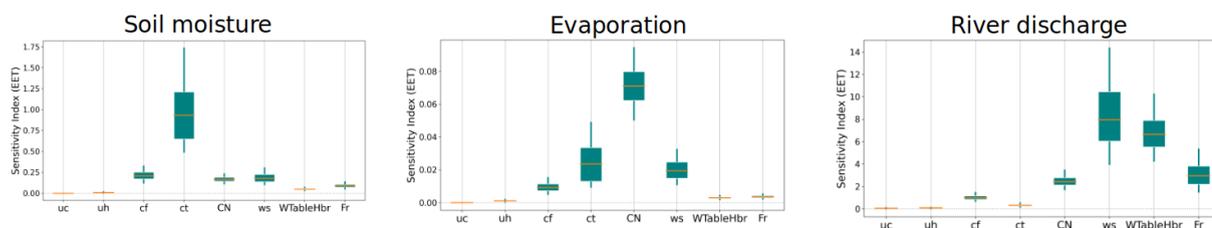
Figure S2: Boxplots indicate the most sensitive parameters for different output variables: soil moisture, evaporation, and river discharge.

Perturbation method

We have worked to improve the perturbation method for those parameters which are not kept constant across the selected domain but are defined as the product of a default map times a scaling factor. Such an approach has the advantage of preserving the spatial distribution of selected parameters, linking it to a physically based quantity, and defining the relation and ranking among such quantities across the domain. However, such an approach is complicated by the need for constraining the calibrated parameter map to physically based constraints derived from the literature, which forces the procedure to a non-linear scaling between the calibration factor and the final (i.e., calibrated) parameter map. Among the chosen calibration parameters of Continuum, such considerations apply to two parameters: the Curve Number (CN) and the Field Capacity (Ct).

The trigonometric arc-tangent function is suitable for applying a scaling of the values in a map within a predefined range. However, an analysis of past applications showed that such function tends to select perturbed parameter maps at the edges of the range (hence with little physical meaning), leaving wide ranges of realistic values undersampled. Our work was focused on addressing such limitations and producing perturbed parameter maps following a more uniform distribution, hence enabling more efficient search of the best values. Results obtained were successful and the current algorithm produces quasi uniform sampling which narrows the sampling range at each iteration and speeds up the convergence to optimal values. Sample results are shown in Figure S3. In addition, we have adapted the sampling method so that the number of perturbations is set higher in the first iteration (default n=50 samples), which is then

reduced by 20% at each subsequent iteration (i.e., 50, 40, 32, 26, 21). Such addition enables a thorough sampling at the initial iterations, yet an efficient use of the computing resources by reducing the number of runs in subsequent iterations. The calibration algorithm now performs a minimum of two iterations (50+40, i.e., a minimum of 90 model runs), where at the end of each iteration after the first evaluates the improvement in the objective function. The calibration stops when the improvement in the objective function is smaller than a predefined threshold, which default value is set to 1%. In addition, a maximum of 5 iterations was imposed, leading to a maximum of 169 model runs for each calibrated domain.
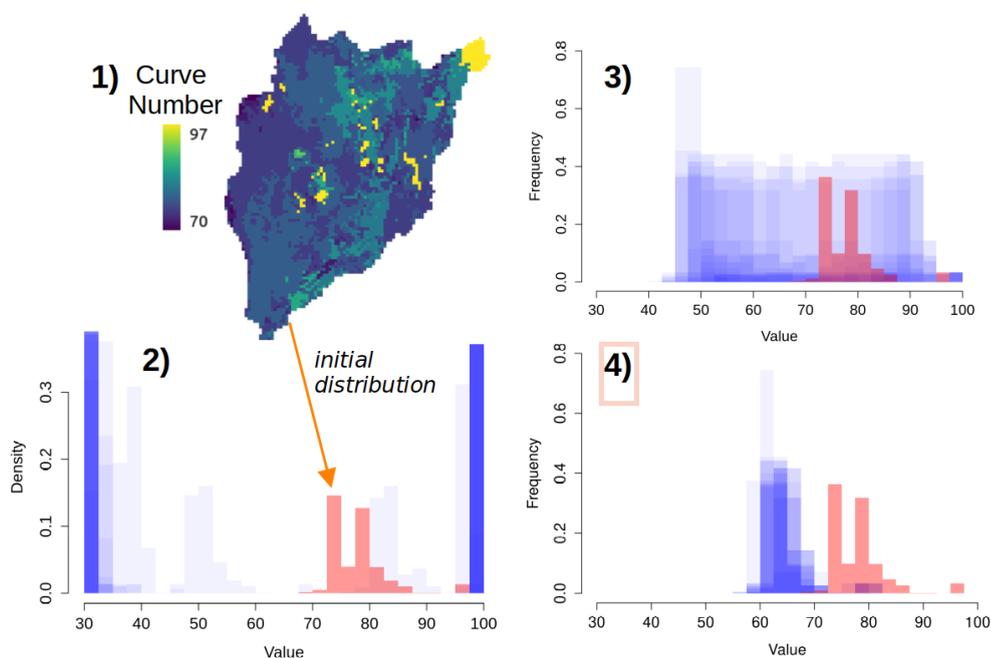
Figure S3: 1) Default map of the Curve Number for a subbasin of the White Nile River, which pdf is shown in red in panels 2) to 4). 2) Unbalanced sampling leading to extreme values (in blue). 3) Improved balanced sampling around the initial distribution (iteration #1). 4) Narrower sampling towards the best parameter distribution (iteration #5).

Objective function

We have tested different objective functions (KGE, nRMSE, correlation, NSE) to evaluate the optimal choice for the model parameterization, keeping in mind the priority of this implementation which is operational flood forecasting and early warning. Ideally, such a system must be capable not only to capture adequately well periods of high and low flows keeping a moderate bias, but also to preserve a skilful ranking between flow peaks and the warning thresholds derived statistically from historical long-term simulations.

The Kling-Gupta Efficiency (KGE) used in previous applications was dismissed from the calibration procedure as it often gave unsatisfactory results. In addition, it has a number of limitations, including its subjective attribution of equal weights to the error components, the non-linear behaviour with the Nash-Sutcliffe Efficiency (NSE), and its varying performance which heavily depends on the hydrological regimes and on the coefficient of Variation (CV) of simulated flows, as reported by Knoben et al. (2019).
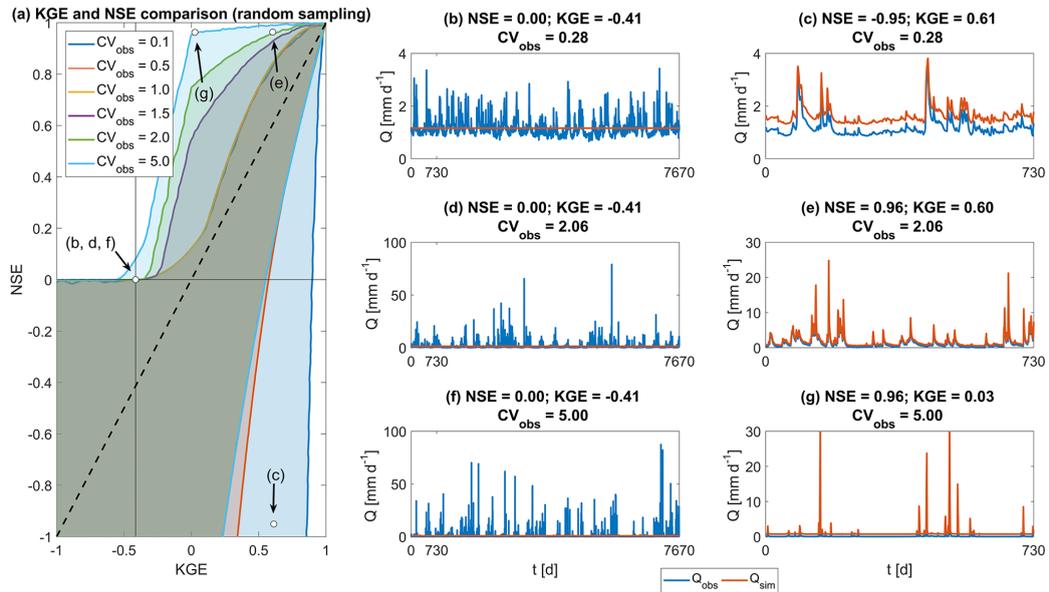
Figure S4: The non-linear relation between KGE and NSE and their influence versus the hydrological regime (from Knoben et al., 2019).

After comparing results of calibrations with different objective functions, we opted for the use of the normalized Root Mean Square Error (nRMSE), where the RMSE of each calibration station is normalized by its average flow obtained from long term records. nRMSE preserves a linear scaling of performance and enables a good trade-off in achieving low bias and good correlation. The optimization of the objective function is performed on the entire time series.

We performed a multi-site calibration where all stations within a model domain contribute at the same time to the evaluation of the objective function, where the nRMSE at each calibration station is weighted by the logarithm of their upstream area, to give a comparable but higher weight to stations located downstream. Multi-site calibrations are known to give on average lower performance than cascading calibrations in the calibration period, but they give higher performance in validation, especially in river sections where no calibration station is available. Therefore, such an approach proves particularly useful in river basins with a limited number of stations.
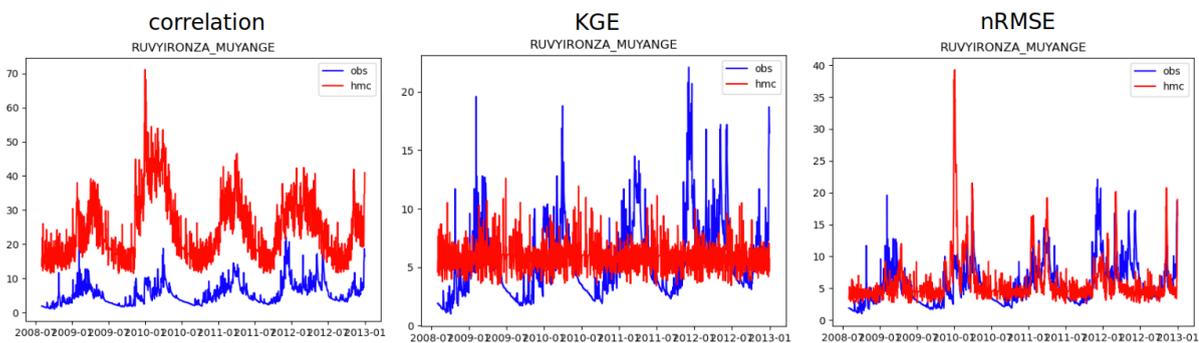
Figure S5: Comparison of observed vs. simulated discharges at a station in Burundi, using model parameters obtained with different objective functions (correlation, KGE, nRMSE).

Exposure layers

Table S2: Exposure data used for the operational impact-based forecasts.

DATA	DESCRIPTION
Population (2020)	Population distribution at 90 meter resolution (2020). This layer contains the number of people per pixel and it is based on the population distribution data from WorldPop top-down modeling methods (https://www.worldpop.org/methods/populations) adjusted to match United Nations national population estimates (UN 2020). The layer has been corrected with reference to the official Census data, when available.
Crop land (2019)	Crop land map at 90 meter resolution (reference year: 2019). Each pixel represents the crop land area in hectares. These data derive from the ASAP crop mask (Version 03, Anomaly Hotspots of Agricultural Production, JRC)
Grazing (2019)	Grazing land map at 90 meter resolution (reference year: 2019). Each pixel represents the grazing land area in hectares. These data derive from ASAP rangeland mask (Version 03, Anomaly Hotspots of Agricultural Production, JRC)
GDP (2019)	Gross Domestic Product (GDP) map at 90 meter resolution (reference year: 2019). Each pixel contains the amount of GDP in USD produced in that pixel. These data derive from the exposure data developed for the GAR 2015 risk atlas (A. de Bono, B. Chatenoux, UNEP/GRID-Geneva. A global exposure model for GAR 2015) adjusted to match 2019 GDP estimates from the World Bank.
Livestock units (2010)	Cattle population in the GHA Region at 90 meter resolution (reference year: 2010). This layer contains the number of cattle per pixel and it is based on data derived from the Harvard dataverse, provided by the ICPAC Geoportal (https://geoportal.icpac.net/layers/geonode%3Acattle_gha).
Road network (2021)	Road Network based on OpenStreetMap shapefile of roads (© OpenStreetMap contributors, March 2021). The length of each road branch is calculated in a GIS environment. Two types of roads are classified (primary and secondary) based on the original class of OpenStreetMap.

Ensemble forecasts
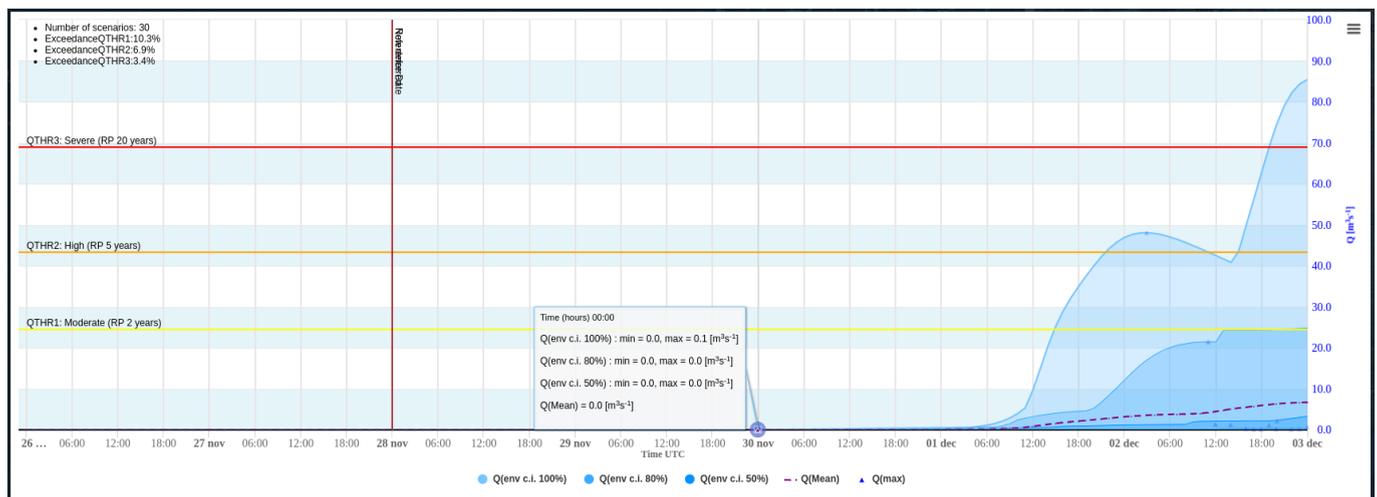
Figure S6: Example of 5-day ensemble discharge forecast visualized in the myDewetra platform for a reporting point in FloodPROOFS East Africa.

Case study - the Nile floods in summer 2020
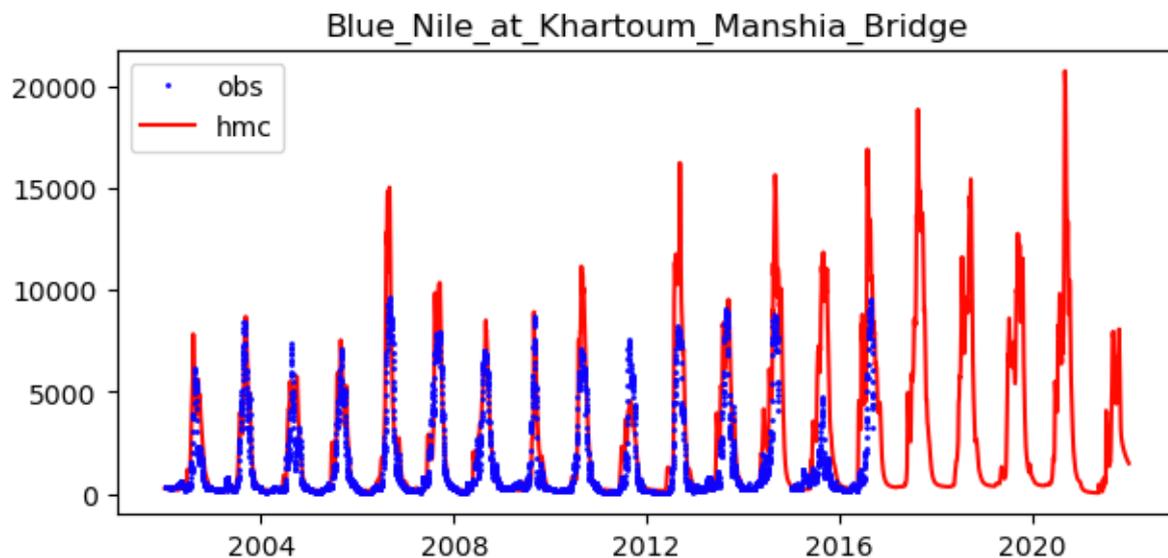
Figure S7: Comparison of simulated versus available observed discharges in the Blue Nile at Khartoum Manshia Bridge. The September 2020 flow peak is the largest in the long term simulation.
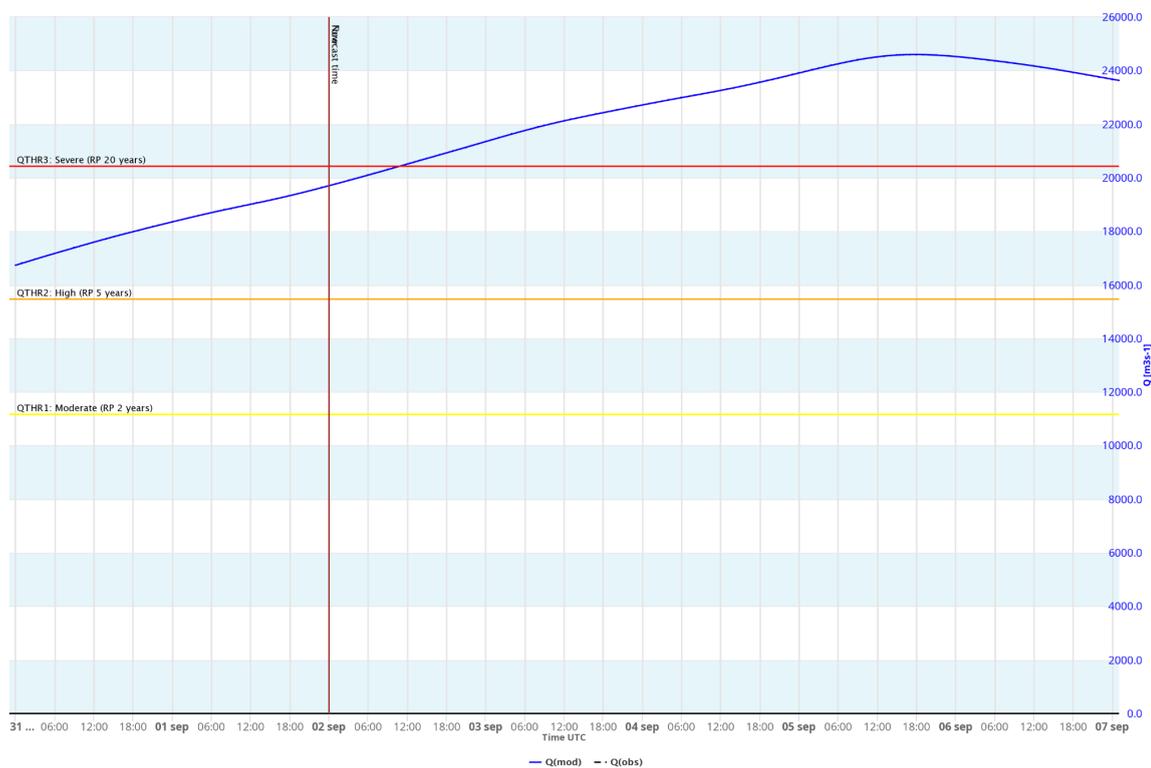
Figure S8: FloodPROOFS East Africa forecast run of 2 September 2020, 00 UTC. Reporting point in the Blue Nile at Khartoum Manshia Bridge. Peak flow is forecast in the evening of the 5 September. Maximum water level was observed on 7 September 2020.
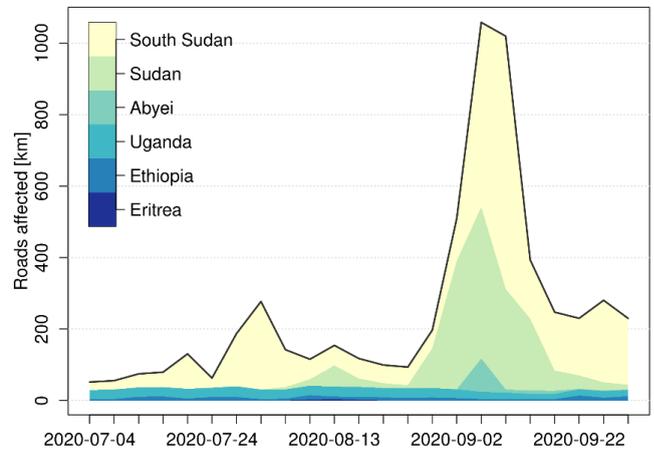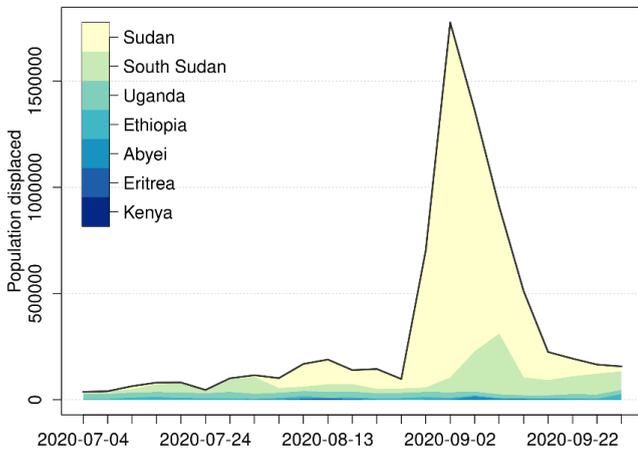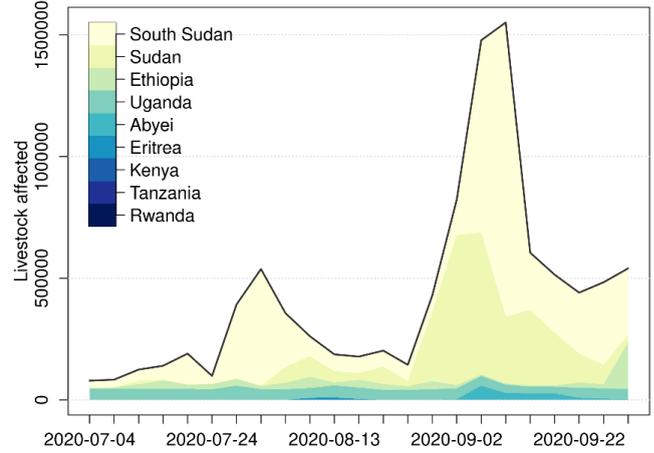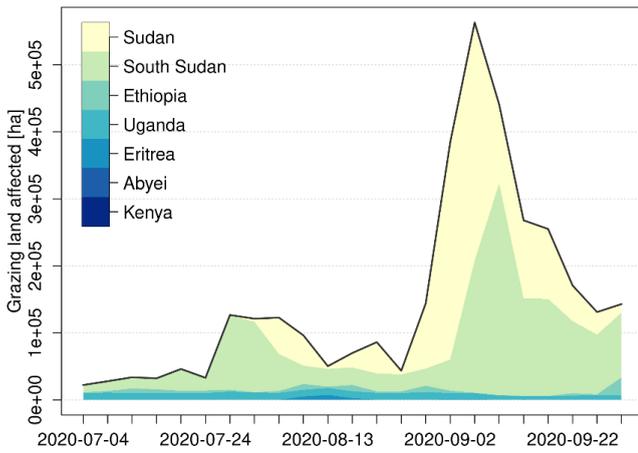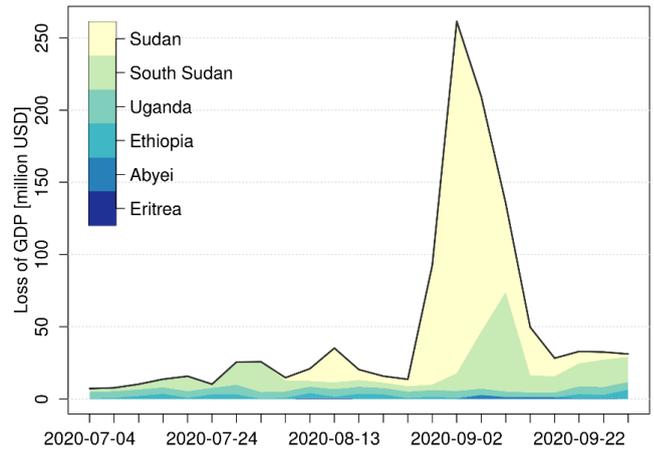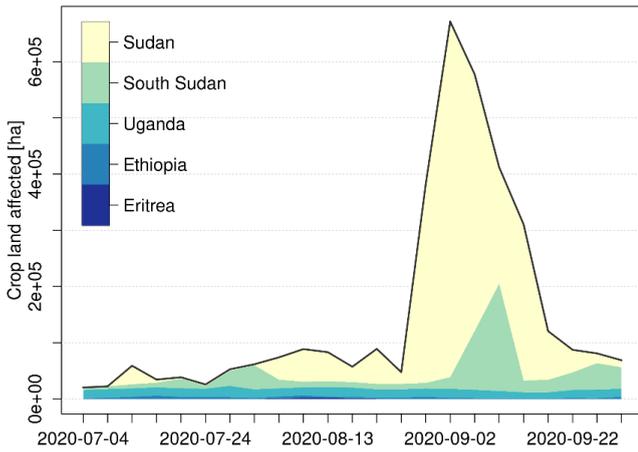
Figure S9: FPEA 5-day impact forecasts for six exposure categories for the period July-September 2020 in the Nile Basin. Aggregations by country.
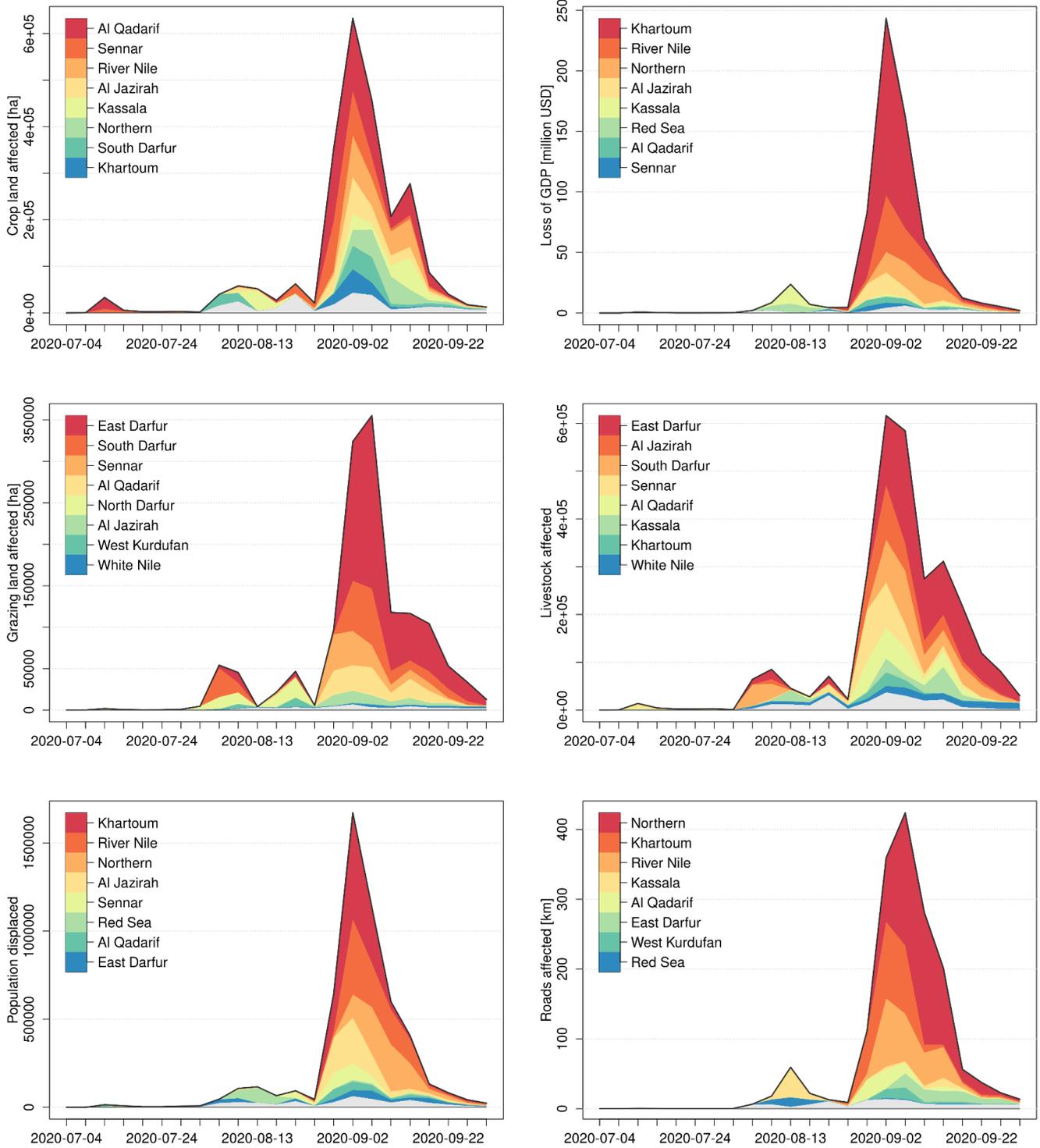
Figure S10: FPEA 5-day impact forecasts for six exposure categories for the period July-September 2020 in Sudan. Aggregations by state (i.e., 1ˢᵗ level sub-national administrative regions). Only the top 8 states for each category are shown. Others are grouped and drawn in light gray.

Hydrological model evaluation

Table S3: *Performance of simulated versus observed discharge at 78 validation stations for the entire period of data availability, ranked by decreasing correlation values.*

ID	Name	nRMSE	KGE	Correlation (r)	Bias rate	Variability rate	longitude	latitude
1	Blue Nile River at GERD	0.51	0.84	0.92	1.09	0.9	34.92	11.24
2	Nile River at Ad Damar	1.13	0.16	0.87	0.79	1.81	33.97	17.65
3	Awash River at Hombole	1.06	0.67	0.86	1.29	0.94	38.78	8.38
4	Blue Nile at El Roseires Dam	0.77	0.75	0.85	1.17	1.12	34.39	11.8
5	Blue Nile at Khartoum Manshia Bridge	1.26	0.5	0.85	0.97	1.47	32.53	15.61
6	Nile at Khartoum Al Halfaia Bridge	0.88	0.44	0.84	0.87	1.52	32.56	15.98
7	Awash River at Melka Kuntire	1.08	0.69	0.84	1.11	0.76	38.6	8.7
8	Blue Nile River at Sennar	1.07	0.57	0.83	0.91	1.38	33.64	13.55
9	Nile River at Dongola	1.52	0.06	0.83	0.96	1.92	30.49	19.18
10	Ruvubu River at Muyiga	1.23	-0.11	0.77	1.96	1.51	30.47	-2.98
11	Shabelle River at Beledweyne	0.71	0.72	0.75	0.98	1.13	45.2	4.74
12	Tana RIver at Garsen	0.88	-0.11	0.73	1.97	0.54	40.12	-2.28
13	Awash River at Haro Adi	1.95	-0.18	0.72	2.09	1.34	39.85	8.85
14	White Nile River at Mongalla	0.17	0.5	0.67	0.63	0.91	31.77	5.2
15	Shabelle River at Bulobarde	1.15	0.36	0.67	0.94	1.54	45.57	3.86
16	Kachwamba River at Sofia	1.02	-3.24	0.59	5.12	0.06	30.46	-1.05
17	Jubba River at Doolow	2.04	-0.21	0.58	2.13	1.08	42.22	4.16
18	Nzoia River at Ruambwa Ferry	0.96	-0.21	0.58	1.88	0.29	34.09	0.12
19	Awash River at Melka Sedi	1.82	-0.99	0.56	2.94	1.06	40.13	9.39
20	Jubba River at Luuq	0.99	0.31	0.54	1.33	0.61	42.56	3.79
21	White Nile River at Malakal	0.71	0.23	0.53	1.08	1.61	31.67	9.55
22	Athi River at Munyu	2.36	-0.65	0.52	2.4	0.27	37.19	-1.09
23	Ewaso Narok River at Sosian	2.75	0.39	0.51	0.99	1.36	36.72	0.44
24	Jubba River at Bardere	0.95	-1.75	0.51	3.56	0.14	42.27	2.34
25	Tana River at Sagana	1.44	-3.35	0.48	5.24	0.19	37.21	-0.67
26	Songwe River at Galula	1.46	-0.99	0.47	2.72	0.13	33.06	-8.73
27	Mbarali River at Igawa	1.24	-1.2	0.47	2.99	0.22	34.38	-8.78
28	Awash River at Adaitu	1.12	0.21	0.45	1.45	0.66	40.78	11.13

29	Lukuledi River at Mkwaya	9.05	-1.18	0.44	2.41	2.57	39.66	-10.14
30	Great Ruaha River at Msembe	1.97	-0.01	0.44	1.48	0.31	34.9	-7.71
31	Thiba RIver at Nyamindi confluence	4.67	-0.21	0.43	1.43	0.01	37.51	-0.73
32	Gucha River at Migori	1.16	0.22	0.41	1.41	0.69	34.21	-0.95
33	Ruvyronza River at Kibaya	1.13	-0.36	0.41	2.21	1.14	29.92	-3.32
34	Mayanja River at Kapeka	1.17	0.03	0.41	1.67	0.62	32.17	0.68
35	Ruhuhu River at Mavanga	1.04	-3.29	0.4	5.17	0.2	35.19	-9.94
36	Rovuma River at Ruvumachini	1.24	-1.59	0.39	3.42	0.28	35.34	-11.2
37	Tana River at Garissa	1.15	-2.82	0.36	4.68	0.18	39.7	-0.45
38	Kizigo River at Chinugulu	1.89	0.04	0.36	1.38	0.39	35.4	-6.93
39	Little Ruaha River at Mawande	1.24	-1.73	0.35	3.5	0.13	35.46	-7.5
40	Wami River at Dakawa	1.8	-1.22	0.34	2.93	0.12	37.53	-6.45
41	Lach Dera River at ElKarama	6.95	-0.57	0.34	0.92	2.42	36.91	0.09
42	Rumpungwe River at Gisuru	1.83	-0.34	0.34	1.67	0.04	30.49	-3.44
43	Great Ruaha River at Salimwani	1.78	-0.81	0.33	2.4	0.06	34.11	-8.9
44	Ruvubu River at Kanabusoro	1.11	0.05	0.33	1.63	0.77	29.71	-3.03
45	Ruaha River Tributary at Kilolo	1.13	-3.28	0.32	5.14	0.15	35.78	-8.02
46	Ruaha River at Iringa	1.18	-2.13	0.32	3.93	0.12	35.72	-7.79
47	Malagarasi River at Mbelagule	1.13	-0.42	0.31	2.19	0.66	30.08	-5.18
48	Wami River at Mandera T2	1.55	-1.67	0.3	3.39	0.03	38.39	-6.25
49	Basse Mulembwe at Mutambara	1.08	-3.91	0.3	5.76	0.03	29.44	-4
50	Ruvu River at Kibaha	1.62	-1.33	0.3	3.06	0.16	38.71	-6.69
51	Mutonga River at Kierera	1.55	-2.27	0.29	4.07	0.13	37.9	-0.38
52	Tsavo River at Tsavo	1.2	-14.78	0.29	16.74	0.07	38.47	-3
53	Kigogo Ruaha River at Lugema	1.14	-3.29	0.29	5.14	0.14	35.34	-8.77
54	Kagera River at Masangano	10.95	-6.53	0.27	1	8.5	31.76	-0.94
55	Tana River at Hola	0.9	-0.7	0.27	2.32	0.21	40.03	-1.5
56	Luichi River at Isangwa	1.34	-1.26	0.26	3.03	0.34	31.6	-7.74
57	Ruaha River at Mafinga	1.25	-1.47	0.26	3.19	0.12	35.33	-8.37
58	Kimani River at GreatNorthRoad	1.71	-1.49	0.26	3.17	0.05	34.17	-8.85
59	Ndembera River at Ilongo	1.62	-1.18	0.25	2.84	0.11	35.16	-8.26
60	Ruvu River at Kifaru	1.42	-5.01	0.24	6.87	0.01	37.56	-3.53

61	Lufilyo River at Lufilyo	1.07	-2.21	0.23	4	0.16	33.92	-9.32
62	Kafu River at Nile confluence	2	0.16	0.23	1.21	1.27	32.04	1.55
63	Kagogo River at Gatore	1.04	-8.07	0.22	9.99	0.07	30.56	-2.27
64	Ruizi River at Mbarara	1.28	-1.36	0.22	3.01	0.05	30.65	-0.62
65	Nyabarongo River at Mugesera	0.99	-6.43	0.2	8.33	0.04	30.28	-2.2
66	Kagera River at Resumo	1.01	-7.1	0.18	9	0.03	30.78	-2.38
67	Pangani River at Mnyuzi	1.45	-3.01	0.18	4.8	0.04	38.56	-5.23
68	Victoria Nile River at Karuma Falls	0.5	-0.58	0.18	2.34	1.17	32.27	2.27
69	White Nile River at Jebel Aulia Dam	2.34	-1.17	0.16	1.76	2.85	32.49	15.25
70	Buzimba River at Gatete	1.4	-1.15	0.16	2.76	0.09	29.54	-4.17
71	White Nile at Panyango	0.29	0.14	0.15	0.96	1.07	31.43	2.57
72	Mbaka River at Mwaya	1.25	-1.18	0.13	2.83	0.19	33.96	-9.52
73	Awash River at Tendaho	1.61	0.01	0.11	1.27	0.65	40.95	11.68
74	Victoria Nile River at Masindi Port	0.63	-0.46	0.11	2.12	1.28	32.09	1.69
75	Mvuha River at Tulo School	1.31	-4.79	0.1	6.63	0.03	37.92	-7.24
76	Ruvu River at Kibangile	1.41	-3.22	0.09	5	0.03	37.8	-7.02
77	Victoria Nile at Paraa	0.59	-0.14	0.07	1.61	1.23	31.56	2.28
78	Victoria Nile River at Mbulamuti	0.67	-1.08	0	2.81	1.18	33.03	0.84

References

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores, Hydrology and Earth System Sciences Discussions, 1–7, https://doi.org/10.5194/hess-2019-327, 2019.

Pianosi, F., Sarrazin, F., and Wagener, T.: A Matlab toolbox for Global Sensitivity Analysis, Environmental Modelling & Software, 70, 80–85, https://doi.org/10.1016/j.envsoft.2015.04.009, 2015.