Supplementary Information for 'Partial Flood Defenses Shift Risks and Amplify Inequality in Core–Periphery City'

Ashish S. Kumar[1], Rajarshi Majumder[2], Vivek P. Kapadia[1], Udit Bhatia[1,3*]

[1]Department of Civil Engineering, Indian Institute of Technology Gandhinagar, Palaj, Gandhinagar, 382055, Gujarat, India.
[2]Department of Economics, University of Burdwan,, Burdwan, 713104, West Bengal, India.
[3]Department of Computer Science and Engineering, Indian Institute of Technology Gandhinagar, Palaj, Gandhinagar, 382055, Gujarat, India.

*Corresponding author(s). E-mail(s): bhatia.u@iitgn.ac.in;
Contributing authors: ashishkumar@iitgn.ac.in;
rmajumder@eco.buruniv.ac.in; vivek.kapadia@iitgn.ac.in;

Text S1: Preprocessing for flood modeling

Dataset

The data required for the flood modeling are displayed in Table S1.

Roughness Coeffiecient of Tapi River and 2D Domain

We adopted a Manning's roughness coefficient of 0.03 for the Tapi River, aligning with values reported in existing literature [1]. To validate this parameter, we simulated the 2013 flood event in the lower Tapi River for the period from September 22, 12:00 PM, to September 26, 6:00 PM (103 hours). Model performance was evaluated by comparing the simulated water level time series with observed data at the Nehru Bridge location during the 2013 flood. The comparison yielded a root mean square error (RMSE) of 0.57 meters and a coefficient of determination (R^2) of 0.96, as illustrated in Supplementary (Fig. S1 (a,b)), indicating satisfactory model accuracy.

Table S1 Dataset Used for Flood Modeling.

Data	Resolution	Source	Remarks
Streamflow	Hourly, Ukai reservoir station	Surat Irrigation Circle (SIC)	Available through request
Tapi river stage data	Hourly, Nehru bridge and Singanpur Wier station	SIC, Central Water Commission (CWC) and SMC	Available through request
Tide level data	Hourly	Literture [1]	Digitized from literature
Land Cover Data	10 m	NRSC Bhuvan Land Cover Map[2] and World Cover Annual Composites[3]	Open source data
Digital Elevation Model (DEM)	30 m	SRTM DEM	Inland elevation data, opensource
Coastal DEM	30 m	NCEI DEM, NOAA	Coastal bathymetry
River cross section	Available at each 200m distance along the river	SMC	Available through request
River Defenses Data	-	Drainage Division Surat	Available through request (Includes location, levee cross-section, and elevation)

Floodplain roughness was derived from the existing LULC map of the study area and classified into five major land cover types: built-up (n = 0.150), agriculture (n = 0.035), forest (n = 0.150), water (n = 0.030), and wasteland (n = 0.055). The roughness coefficient ('n') values for each land class were obtained from the literature [1, 4, 5]. The same roughness values were used for the 2006 flood model to validate the 2D flood depth and for subsequent flood event simulations.

Text S2: Validation of Flood Model

The validation of the flood model is performed for the 2006 flood scenario. We simulate the flood event from August 3, 2006 to August 11, 2006 at an hourly time scale for the entire flood domain. The simulated output of flood extent and depth was then validated against the observed flood depth and extent.

- To validate the simulated flood depth, we utilized observed flood depth data from 80 locations within the study region from the flood-affected area map of 2006 obtained from the Surat Municpal Coorporation. The flood depth map was georeferenced to align with the study region, and the observed flood depth points were digitized. Corresponding points between the observed and simulated flood depths were identified to assess the model's accuracy in replicating flood depths across various spatial locations.
- To validate the simulated flood extent, we utilized available SMC 2006 flood extent map data to extract the water-submerged area during the flood event. The map diplayed the classification of submerged area in range of flood depths which area was used as the observed flood extent for validation. Validation was conducted through

an accuracy assessment using 1,000 stratified random sampling points generated across the study region [6]. Each point was classified into flooded and non-flooded categories based on both the observed and simulated flood extents. A confusion matrix was constructed to calculate the overall accuracy of the validation process. Additionally, Cohen's kappa coefficient [7] and the weighted F1 score [8] were estimated to further assess model performance. These metrics help address potential biases arising from class imbalances.

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{1}$$

where κ is cohen's kappa coefficient, P_o is observed agreement and P_e is expected chance agreement.

$$\mathrm{F1} = \frac{2 \times (\mathrm{Precision} \times \mathrm{Recall})}{\mathrm{Precision} + \mathrm{Recall}} \tag{2}$$

where *Precision* is

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{3}$$

and *Recall* is

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{4}$$

where True Positives (TP) refers to the number of correctly predicted positive instances (i.e., the cases where the model correctly identifies a flood area). False Positives (FP) represents the number of instances where the model incorrectly predicted a positive class (i.e., the model falsely identified a flood area when it wasn't flooded). False Negatives (FN) represents the number of instances where the model failed to predict an actual flood area (i.e., the model missed some flooded regions, falsely classifying them as non-flooded).

$$\mathrm{F1}_{\mathrm{weighted}} = \frac{\sum_{i=1}^{n} w_i \cdot \mathrm{F1}_i}{\sum_{i=1}^{n} w_i} \tag{5}$$

where w_i is the weight for class i, typically the number of true instances for class i, $\mathrm{F1}_i$ is the F1 score for class i, and n is the total number of classes.

Table S2 Error matrix of accuracy assessment for 1D-2D hydrodynamic model with Cohen's Kappa value and F1 score

Actual \Predicted	**Non-Flooded**	**Flooded**	**Total**	**Producer's Accuracy**
Non-Flooded	364	32	396	0.9192
Flooded	222	344	566	0.6078
Total	586	376	962	
User's Accuracy	0.6212	0.9149		Overall Accuracy = 0.74
Kappa Coefficient		**0.49**		
Weighted F1-Score		**0.73**		

Text S3: Flood Frequency analysis

We employed the Generalized Extreme Value (GEV) distribution [9, 10] to analyze streamflow return levels for 5-,25-,50-,100- and 250-year return periods. The analysis followed the block maxima approach[11], where annual maximum streamflow values were extracted from the dataset. The GEV distribution, characterized by the location parameter (μ), scale parameter (σ), and shape parameter (ξ), was fitted to the block maxima data. The return levels (x_T) corresponding to a return period (T) were calculated using the formula:

$$x_T = \begin{cases} \mu + \frac{\sigma}{\xi} \left[\left(-\ln\left(1 - \frac{1}{T}\right) \right)^{-\xi} - 1 \right], & \text{if } \xi \neq 0, \\ \mu - \sigma \ln\left(-\ln\left(1 - \frac{1}{T}\right) \right), & \text{if } \xi = 0. \end{cases} \tag{6}$$

To ensure the suitability of the GEV distribution for modeling the streamflow extremes, we performed a Kolmogorov-Smirnov (K-S) goodness-of-fit test. The test yielded a K-S statistic of 0.10 and a p-value of 0.68, indicating an excellent fit of the GEV distribution to the observed data, as the high p-value confirms the null hypothesis that the data follows the specified distribution.

Text S4: Cost-Benefit Analysis

Expected Annual Loss

An Expected Annual Loss (EAL) approach was adopted to integrate damages from floods of varying magnitudes, weighted by their probabilities of occurrence[12]. We considered return periods of 5, 25, 50, 100, and 250 years. EAL was computed as

$$\text{EAL} = \sum_i P_{(i)} \times L_{(i)} \tag{7}$$

where $P_{(i)}$ is the probability of the event with return period i and $L_{(i)}$ is the associated economic loss. Including a range of events allowed us to assess not only moderate but also extreme floods that could exacerbate residual risks under partial adaptation.

Cost of Levee Construction and Maintenance

The cost of levee construction was estimated based on a global meta-analysis dataset, which provides unit costs in Million Euros per kilometer and per meter (reference year: 2014)[13]. We specifically extracted the unit cost for India and followed these steps for cost estimation: Extracting Baseline Cost: The levee construction cost for India was obtained from the global dataset (reference year: 2014). Currency Conversion: Costs were converted from Euros to Indian Rupees using the exchange rate of the reference year. Inflation Adjustment: The cost was adjusted to 2022 prices using the Consumer Price Index (CPI) to account for inflation[14]. Maintenance Cost Estimation: A 1% cost of construction was assumed based on standard engineering practices for levee structures.

Estimation of Net Avoided Damage and Benefit-Cost Ratio

To assess the long-term economic feasibility of levee construction, we considered a service life of 50 years and computed the Net Present Value (NPV) of expected annual flood damages over this period.

Net Present Worth of Expected Annual Damage

The total flood damage over 50 years was discounted using a 6% discount rate, which represents the time value of money in economic assessments. The Net Present Worth (NPW) of expected annual damages (EAD_{NPW}) was estimated as:

$$\mathrm{EAD_{NPW}} = \sum_{t=1}^{50} \frac{\mathrm{EAL}}{(1+r)^t} \tag{8}$$

where:r is the discount rate (6%) and t is the year in the assessment period (1 to 50).

Net Avoided Damage Over 50 Years

The Net Avoided Damage due to levee construction was computed as:

$$\text{Net Avoided Damage} = \mathrm{EAD_{NPW}} - (\text{Construction Cost} + \text{Maintenance Cost over 50 years}) \tag{9}$$

Benefit-Cost Ratio (BCR)

The Benefit-Cost Ratio (BCR) determines the economic feasibility of adaptation measures and is given by:

$$\mathrm{BCR} = \frac{\text{Net Avoided Damage}}{\text{Construction Cost} + \text{Maintenance Cost over 50 years}} \tag{10}$$

A $BCR > 1$ indicates that the levee investment is economically justified, whereas a $BCR < 1$ suggests that the costs outweigh the benefits.

Table S3 Cost-Benefit Analysis Results

Parameter	**Lower Estimate**	**Absolute Estimate**	**Upper Estimate**
Avoided Expected Annual Damage (EAL) [Billion ₹]	0.94	2.02	3.15
Avoided EAL Net Present Worth (50 years) [Billion ₹]	20.02	54.54	88.90
Cost of Construction + Maintenance [Billion ₹]	37.55	37.55	37.55
Benefit-Cost Ratio	0.53	1.45	2.36

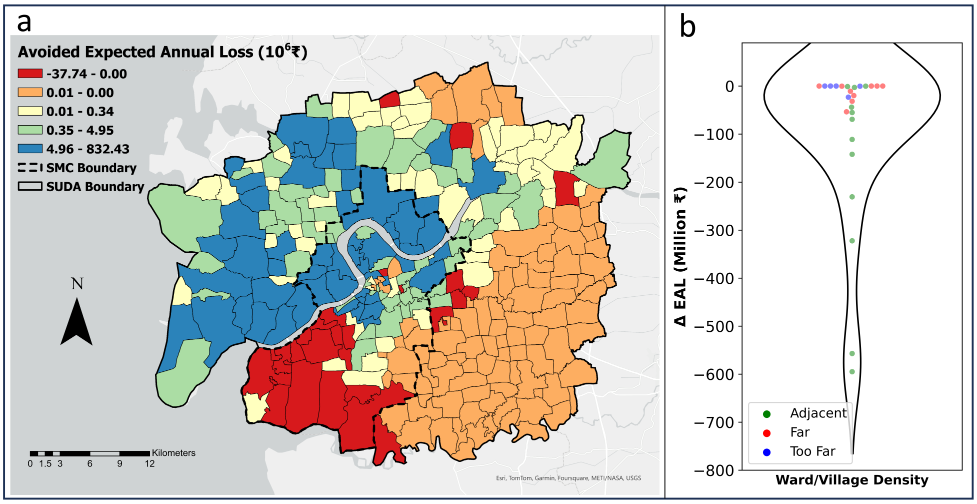

Fig. S1 (a) Spatial distribution of avoided expected annual loss (Base-Protected EAL). (b) Distribution of avoided expected annual loss in downstream wards, expressed in net present worth over a 50-year service life, shown as a swarm-violin plot. The maximum negative ΔEAL is 800 million ₹.

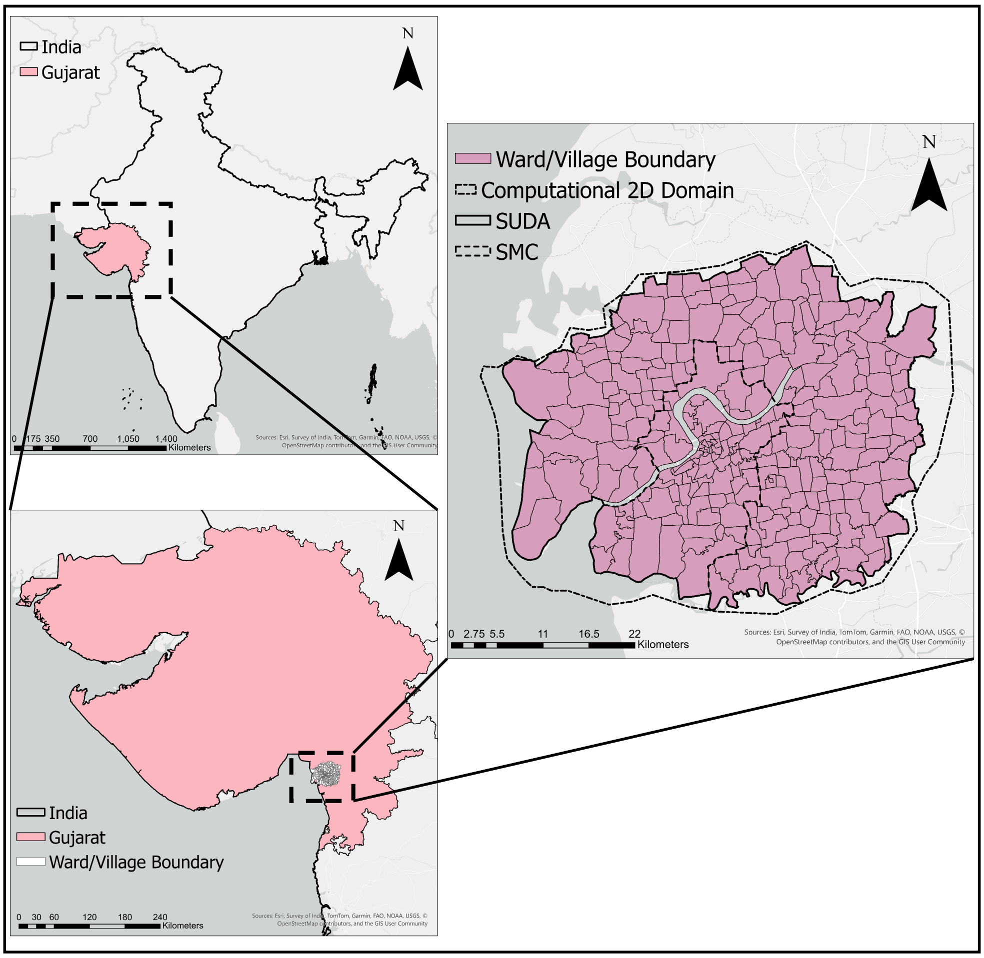

Fig. S2 Study area of Surat (Both SMC and SUDA), located in the state of Gujarat, Western India. The shaded wards/villages represent the areas considered for flood risk assessment, as they overlap with the computational 2D flood domain.

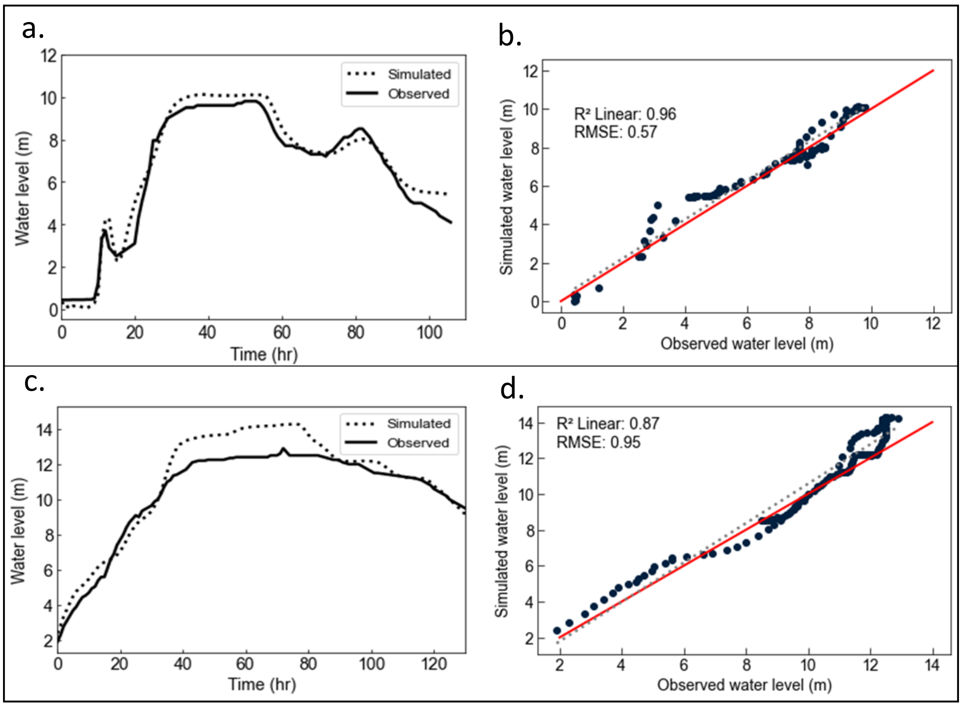

Fig. S3 Calibration and Validation: (a) Time series comparison of observed and simulated flood levels (Location: Nehru Bridge) for the 2013 flood, used for model calibration. (b) Scatter plot of simulated vs. observed water levels for accuracy assessment (2013 flood) . (c)Time series comparison of observed and simulated flood levels (Location: Nehru Bridge) for the 2006 flood, used for model validation. (d) Scatter plot of simulated vs. observed water levels for accuracy assessment (2006 flood).

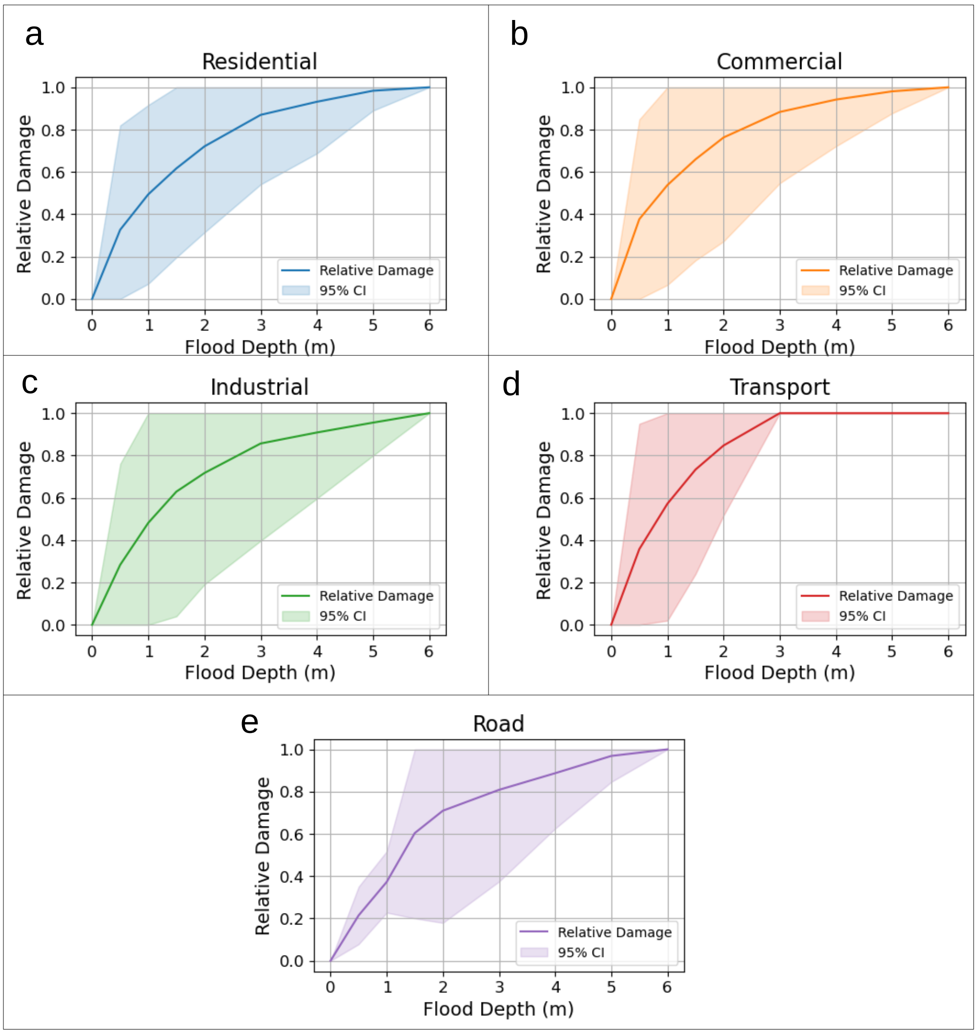

Fig. S4 Relative Depth-damage functions for different land use types. Fig.(a–e) depict the depth-damage relationships for Residential, Commercial, Industrial, Transport, and Road, respectively. The shaded region in all subfigures represents the 95% confidence interval.

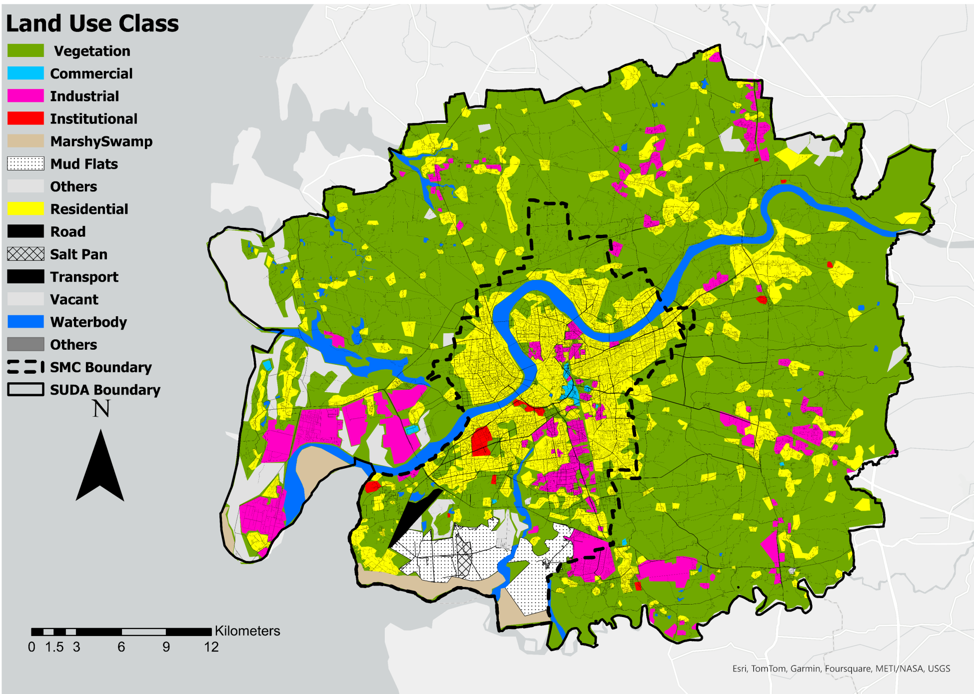

Fig. S5 Land use map of Study Area.

References

[1] Timbadiya, P., Patel, P., Porey, P.: A 1d–2d coupled hydrodynamic model for river flood prediction in a coastal urban floodplain. Journal of Hydrologic Engineering **20**(2), 05014017 (2015)

[2] National Remote Sensing Centre (NRSC): Bhuvan - Indian Geo-Platform of ISRO. Accessed: 2025-03-08 (2025). https://bhuvan.nrsc.gov.in/

[3] European Space Agency: WorldCover: Global Land Cover Mapping at 10 m Resolution. https://esa-worldcover.org/en. Accessed: 2025-03-08 (2021)

[4] Chow, V.t.: Open-channel hydraulics, classical textbook reissue. MC Graw Hill Seattle, WA (1988)

[5] Patel, D.P., Ramirez, J.A., Srivastava, P.K., Bray, M., Han, D.: Assessment of flood inundation mapping of surat city by coupled 1d/2d hydrodynamic modeling: a case application of the new hec-ras 5. Natural Hazards **89**, 93–130 (2017)

[6] Story, M., Congalton, R.G.: Accuracy assessment: a user's perspective. Photogrammetric Engineering and remote sensing **52**(3), 397–399 (1986)

[7] Landis, J.: The measurement of observer agreement for categorical data. Biometrics (1977)

[8] Goutte, C., Gaussier, E.: A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In: European Conference on Information Retrieval, pp. 345–359 (2005). Springer

[9] Hamed, K., Rao, A.R.: Flood Frequency Analysis. CRC press, ??? (2019)

[10] Lettenmaier, D.P., Wallis, J.R., Wood, E.F.: Effect of regional heterogeneity on flood frequency estimation. Water resources research **23**(2), 313–323 (1987)

[11] Coles, S.: An Introduction to Statistical Modeling of Extreme Values. Springer, London (2001). https://doi.org/10.1007/978-1-4471-3675-0

[12] Wobus, C., Porter, J., Lorie, M., Martinich, J., Bash, R.: Climate change, riverine flood risk and adaptation for the conterminous united states. Environmental Research Letters **16**(9), 094034 (2021)

[13] Nicholls, R., Hinkel, J., Lincke, D., Pol, T.: Global investment costs for coastal defense through the 21st century (2019)

[14] World Bank: Consumer Price Index (CPI). Accessed: 2025-01-10 (2023). https://data.worldbank.org/indicator/FP.CPI.TOTL