A Flood Risk Analytics and Visualisation Platform for Socio-Economic Scenarios

Isha Dev[1], Sabirul Sk[2], Nirul Ranjan Patra[1] and Subhankar Karmakar[1,3]

Environmental Science and Engineering Department, Indian Institute of Technology Bombay, Mumbai 400076, India [1]
E-mail: ishadev.iitb@gmail.com; sabirul@iitb.ac.in; nirul23@iitb.ac.in
Centre for Climate Studies, Indian Institute of Technology Bombay, Mumbai 400076, India[2]
E-mail of corresponding author: skarmakar@iitb.ac.in

ABSTRACT

Flood risk assessments increasingly require integration of hazard impacts with socioeconomic vulnerability (SEV) to support equitable resilience planning, as emphasised by the Intergovernmental Panel on Climate Change (IPCC) in their last two cycles of Assessment Reports (AR5 and AR6). This study presents a Flood Risk Analytics and Visualisation Platform, an interactive tool that links flood hazard information with diverse socioeconomic scenarios to support decision-making. Developed using PowerBI, an interactive data visualisation tool that enables rapid exploration of flood-prone areas, associated socioeconomic conditions, and localised vulnerability drivers, providing actionable insights for policymakers and practitioners. Maharashtra, India's second-most populous state and a region experiencing recurrent floods every two to three years, is used as a case study to demonstrate the platform's applicability. While previous national and regional assessments have evaluated vulnerability, they often lack sub-district (talukas/tehsils) level resolution with limited SEV indicators. To address these gaps, the study quantifies SEV for 357 sub-districts using demographic and census-based indicators, with indicator weighting optimised through a non-parametric Data Envelopment Analysis (DEA) approach to minimise subjective bias. The results reveal high vulnerability clusters in the Central and Eastern Vidarbha Zone, moderate vulnerability across the Central Maharashtra Plateau, and relatively lower vulnerability in the North Konkan Coastal and Western Maharashtra zones. By integrating these analytical outputs into an interactive visual platform, the study provides a scalable and adaptable decision-support tool capable of informing targeted vulnerability reduction and resilience strategies. The platform's capacity to evaluate fine-resolution SEV alongside flood hazard information strengthens its utility for locally grounded adaptation planning and policy formulation.

KEYWORDS: Adaptation planning, decision-support platform, disaster risk reduction, flood risk, Socioeconomic vulnerability, vulnerability indicators

1 INTRODUCTION

Natural catastrophes have become both more frequent and more financially devastating since the 1970s (Yamamura, 2015), a trajectory expected to intensify as greenhouse gas concentrations continue to rise (IPCC, 2018). The uneven spatial and social distribution of disaster impacts underscores the interplay between climate change, hazard recurrence, and socioeconomic vulnerability (Groeschl & Noy, 2020). Marginalised and resource-constrained populations consistently bear a disproportionate share of losses (Benson et al., 2001; Sakai et al., 2017), highlighting the critical need for inclusive risk-reduction policies that prioritise vulnerability reduction and resilience building. In developing nations such as India, where more than 95% of global disaster-related deaths between 1970 and 2008 were recorded (IPCC, 2012), the growing recurrence of floods, droughts, and heatwaves continues to threaten sustainable development trajectories.

Floods are among the most recurrent natural hazards worldwide, primarily driven by extreme rainfall events that disrupt livelihoods and cause substantial socio-economic losses (IFRC, 2011; Dottori et al., 2018).

While accurate flood forecasting and preparedness are essential for risk reduction, they remain challenging due to the complexity of atmospheric and land–surface interactions (Fritsch et al., 1988). The IPCC AR6 (2023) conceptualises disaster risk as a function of hazard, exposure, and vulnerability, with hazards escalating into disasters when they intersect with populations lacking adaptive capacity (Birkmann, 2006). In Maharashtra, existing vulnerability assessments are commonly aggregated at the district level, obscuring intra-district heterogeneity that is crucial for targeted interventions. To address this limitation, the present study develops a high-resolution socioeconomic vulnerability (SEV) assessment at the sub-district (tehsil) scale using 23 indicators derived from the 2011 Census of India. Indicator weighting subjectivity is reduced via a non-parametric Data Envelopment Analysis (DEA) optimisation approach. Complementarily, a probabilistic rainfall hazard model combining IMD observations with GEFSv12 forecasts captures event-specific hydroclimatic extremes, validated using two intense rainfall events in Raigad during 2015 and 2016.

Building on these analytical outputs, the study operationalises high-resolution SEV and hazard results through a Flood Risk Visualisation Platform developed in PowerBI (Bhargava et al., 2018). Integrating both components within an interactive dashboard environment enables users to explore spatial patterns of flood susceptibility, interrogate the socioeconomic drivers of vulnerability, and compare risk profiles across sub-districts. By combining analytical rigour with intuitive geo-visualisation, the platform strengthens the accessibility and usability of spatial risk evidence for policymakers, disaster managers, and development practitioners supporting informed prioritisation, early-warning preparedness, and locally grounded adaptation planning.

2 DATA AND METHODOLOGY

2.1 Study Area

Maharashtra, located in western–central India, is the country's second-most populous and third-largest state, spanning ~0.31 million km² and hosting 112.4 million people (Census of India, 2011). It is highly urbanised, with major metropolitan centres such as Mumbai, Pune, Nagpur, and Nashik, and serves as India's leading state economy with a GSDP of ~USD 435 billion in 2022–23 (Maharashtra State Disaster Management Plan, 2023).

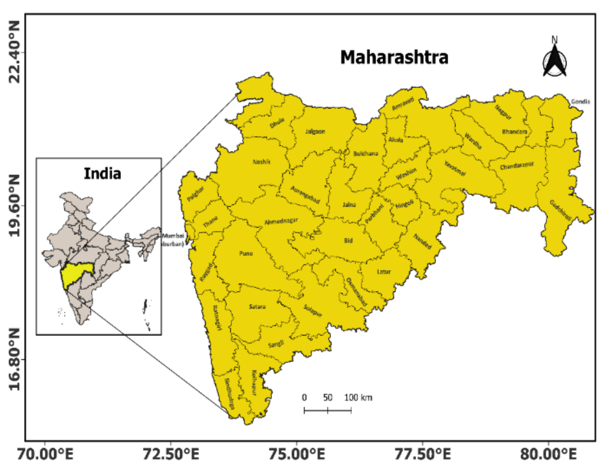

Figure 1. A map developed in QGIS showing the study area with a subdistrict-level map of Maharashtra showing the area.

2.2 Socioeconomic Vulnerability

The methodology comprises two core components: (i) high-resolution socioeconomic vulnerability (SEV) assessment at the sub-district scale, and (ii) probabilistic rainfall-based hazard estimation. The integrated workflow adopted in this study is summarised in Figures 2 and 3. First, SEV indicators derived from demographic and census datasets were pre-processed, normalised, dimension-reduced using Principal Component Analysis (PCA), and weighted through a Data Envelopment Analysis (DEA) framework to produce sub-district-level vulnerability scores (Figure 2). Second, observed IMD rainfall records and GEFSv12 forecasts (2000–2019) were jointly modelled using kernel-based bivariate functions to estimate conditional exceedance probabilities representing rainfall-induced hazard values (Figure 3). The resulting SEV and hazard layers form the basis for spatial flood risk estimation and visualisation.

2.2.1. Normalisation of Indicators

The socioeconomic data for the 357 sub-districts in Maharashtra were extracted from the recent Indian Census of 2011 from the Census of India (CoI) official website. Indicator selection was based on a comprehensive review of relevant literature. The selected indicators should be relevant, highly feasible, better representative, and justifiable with respect to the particular study (Bogardi & Birkmann, 2004; Sherly et al., 2015). Initially, 35 indicators were chosen for their relevance and availability at the sub-district level. These indicators were classified as sensitive and adaptive, depending on whether they increase or decrease SEV. The selected indicators were then merged into 23 composite indicators. The indicators were then normalised to make them dimensionless (Wu et al., 2002; Karmakar et al., 2010). The normalisation process followed the equations mentioned in the study –

$$V_{ij}^{std} = \frac{V_{ij} - V_i^{min}}{V_i^{max} - V_i^{min}} \qquad (\forall i \subseteq S) \tag{1}$$

$$V_{ij}^{std} = \frac{V_i^{max} - V_{ij}}{V_i^{max} - V_i^{min}} \qquad (\forall i \subseteq S^C) \tag{2}$$

where, V_{ij}^{std} represents the standardised indicator value and V_{ij} is the actual value of the i^{th} indicator for the jth decision-making unit (DMU). The DMUs are the sub-districts selected in the study domain, i.e., the entire Maharashtra state. The S and S^C represent the set of sensitive and adaptive indicators for all the DMUs. In the above equation, the V_i^{max} and V_i^{min} are the maximum and minimum values of the indicator for the ith indicator, considering all sub-districts.

2.2.2 Principal Component Analysis and Data Envelopment Analysis

Principal Component Analysis (PCA) was used to identify the key components and to decorrelate and reduce the indicators' dimensions after normalisation. The Kaiser criterion was used to select the number of significant PCs. According to the Kaiser criterion, six components (i.e., the first 6 PCs) were identified as significant. The Kaiser rule suggests discarding any components with eigenvalues less than 1.0, as an eigenvalue of 1.0 represents the amount of information explained by an average individual item. PCA was performed using MATLAB Software. Six significant PCs were selected as per the analysis, which explains 84.86% of the total variance, and the resulting components were then input into the Data Envelopment Analysis (DEA) model in R Software. This model was used to determine the efficiency of the decision-making units (DMUs), the sub-districts in this case. This technique is a non-parametric approach, which optimises each sub-district to calculate a discrete piecewise frontier, as per the estimation by the complete set of Pareto-efficient DMUs (Vittal et al., 2020).

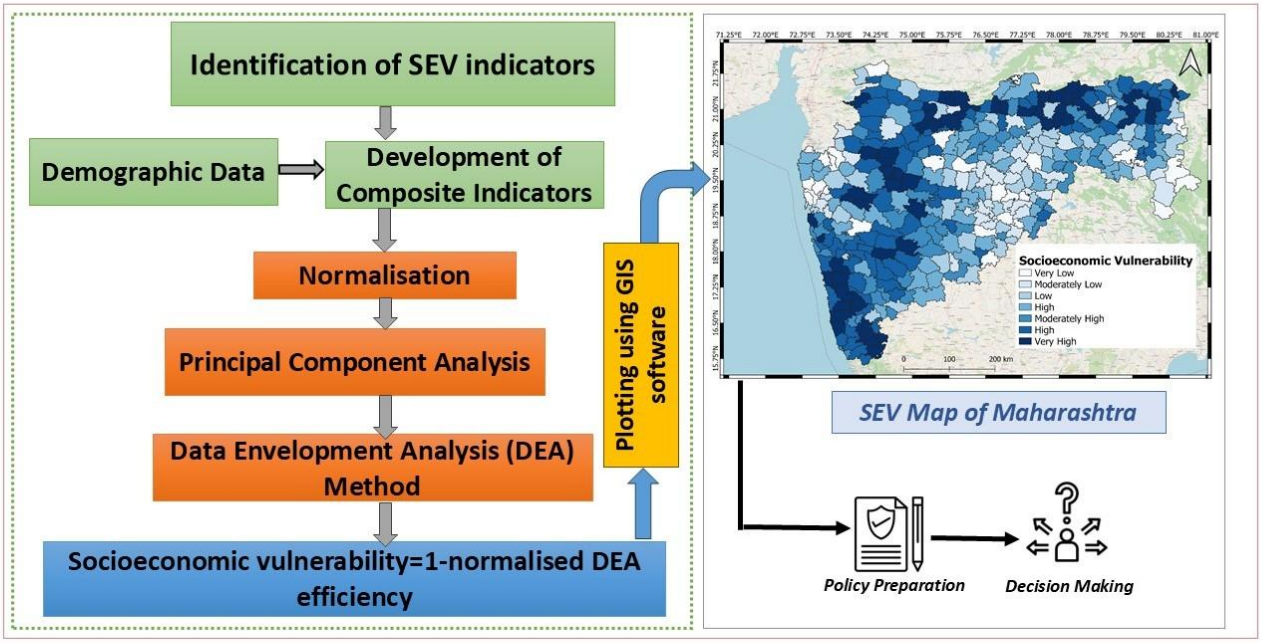

Figure 2. Methodological framework for constructing the Socio-Economic Vulnerability (SEV) map using a hybrid approach integrating Principal Component Analysis (PCA) and Data Envelopment Analysis (DEA). It also represents the sequential workflow from indicator selection and data normalisation to dimensionality reduction via PCA and efficiency score estimation using DEA

2.2.3 Finding major drivers of vulnerability

To identify the major drivers of socioeconomic vulnerability, the variance-based factor analysis was done to find the contribution of the important indicators in the significant PCs. The significance of an observation for a component is determined by the ratio of the observation's squared factor score to the eigenvalue corresponding to that component. This ratio is referred to as the observation's contribution to the component (Abdi & Williams, 2010; Dhakal et al., 2020; Shlens, 2014; Uddin et al., 2019). This method is based on variance; the variance of individual indicators is evaluated only on the significant PCs (here, 6 PCs).

2.3 Hazard Calculation

For this study, year-wise precipitation forecasts during the Indian summer monsoon months—June to September (122 days/year) were utilised. The final dataset has dimensions of 122 × 11 × 40 × 33 × 26, representing days, ensemble members, 6-hourly forecast steps across 10 days, and spatial grids (latitude × longitude). GEFSv12 provides 5 ensemble members per day (c00, p01–p04) and 11 members on Wednesdays (c00, p01–p10). From these files, only the 6-hourly forecast steps were retained. Time series models were applied separately to observed and forecasted datasets to capture temporal dependencies. A bivariate kernel density estimation method was used to construct a joint distribution, enabling the computation of conditional probabilities of extreme rainfall events. The methodology was applied to two heavy rainfall events—3rd August 2016 and 22nd June 2015—in the Raigad district of Maharashtra.

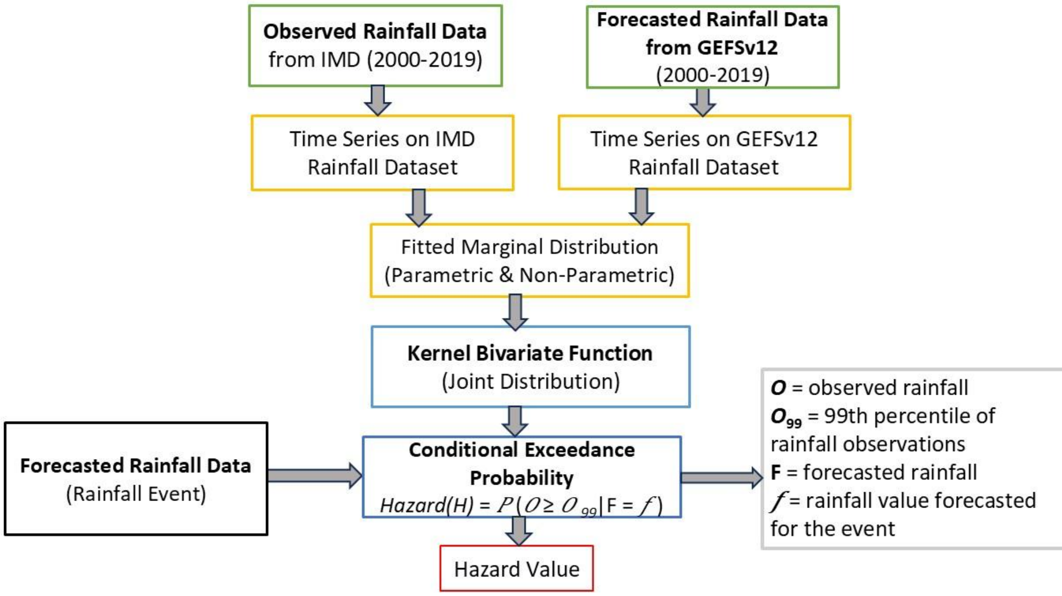

Figure 3. Methodological framework for deriving hazard values using bivariate modelling of observed and forecasted rainfall. The observed (IMD) and forecasted (GEFSv12) rainfall datasets for the period 2020–2019 are individually fitted with marginal distributions and time series models. These are then combined using a Kernel Bivariate Function to estimate the joint distribution. This joint probability is used as the hazard value.

2.3.1 Extreme Precipitation Risk Mapping

Conventionally, risk associated with extreme events is conceptualised as the product of hazard, vulnerability, and exposure, a framework widely adopted in disaster risk science (Chen et al., 2015; Gusain et al., 2020; IPCC, 2012; Karmakar et al., 2010; Kron, 2005; Sahani et al., 2019). This formulation emphasises that the overall risk is not solely dependent on the likelihood of an extreme event (hazard), but also on how exposed a system or population is, and how susceptible they are to harm.

$$\text{Risk} = \text{Vulnerability} \times \text{Hazard} \tag{3}$$

3 RESULTS

3.1 Socioeconomic Vulnerability Assessment

Figure 4 illustrates the spatial distribution of socioeconomic vulnerability at the sub-district level across Maharashtra, derived using a composite index developed through Data Envelopment Analysis (DEA). DEA, a non-parametric linear programming technique, was employed to assess the relative vulnerability of sub-districts by benchmarking multiple Census of India (2011) indicators representing demographic pressure, housing conditions, access to basic services (drinking water, electricity, sanitation), employment characteristics, and social marginalisation. Vulnerability is quantified as deviation from an efficiency frontier, enabling the identification of sub-districts performing poorly relative to comparable administrative units under similar constraints. The resulting vulnerability scores were classified into seven ordinal categories—Very Low (VL), Low (L), Moderately Low (ML), Medium (M), Moderately High (MH), High (H), and Very High (VH)—using the Jenks natural breaks algorithm. This approach minimises within-class variance while maximising inter-class differences, thereby enhancing the representation of spatial heterogeneity.

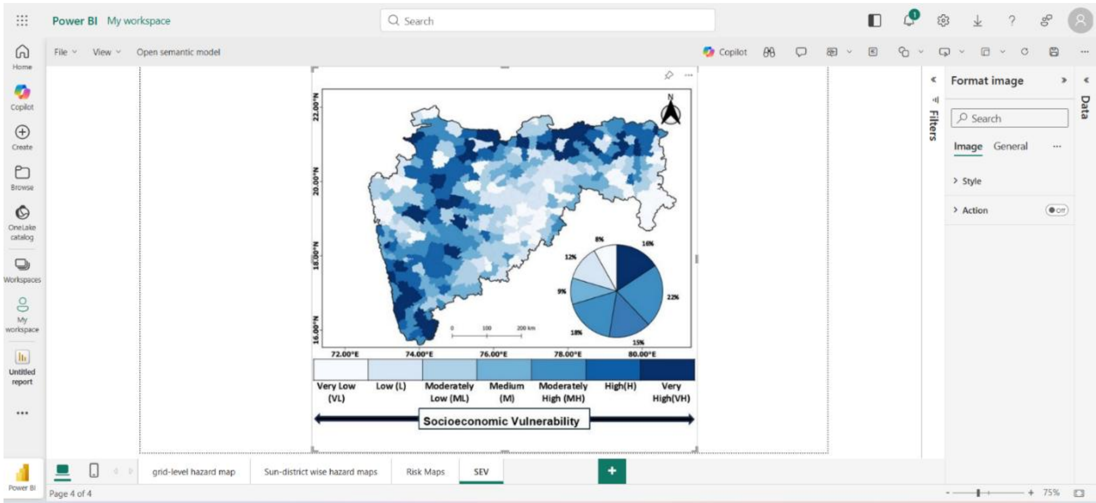

Figure 4. A PowerBI snippet showing the spatial distribution of the Socio-Economic Vulnerability (SEV) Index across sub-districts in Maharashtra, India. The SEV scores are classified into seven ordinal categories, ranging from "Very Low" to "Very High," with progressively darker shades indicating higher vulnerability levels. An inset pie chart displays the proportion of the total geographic area falling within each SEV class across the state.

The map reveals pronounced clustering of higher vulnerability classes (H and VH) in the southern, southeastern, and northeastern regions of Maharashtra. These areas are predominantly characterised by rural settlements, greater dependence on climate-sensitive livelihoods such as agriculture, and relatively limited access to public infrastructure, increasing their susceptibility to hazard-induced disruptions.

3.2 Hazard

To demonstrate and validate the proposed hydroclimatic hazard assessment framework, two major extreme rainfall events in Maharashtra were selected as case studies. The first occurred in June 2015 during the early monsoon, with a maximum daily rainfall of 292.21 mm centred near Poladpur. The second event took place in August 2016 in the Konkan region and recorded one of the highest single-day rainfall intensities in the dataset (338.52 mm), again focused around Poladpur in Raigad district (Fig.5). These events were chosen due to their high spatial rainfall variability, severe impacts on infrastructure and livelihoods, and their representativeness of monsoon-driven hazards along the Western Ghats.

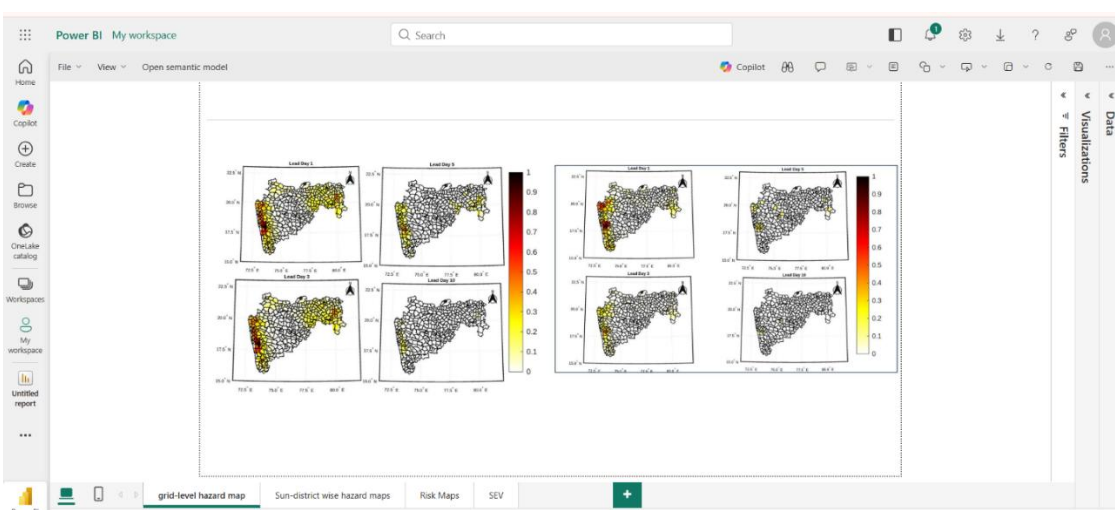

Figure 5. A PowerBI snippet showing Grid-level hazard probability maps for 19 June 2015 and 3rd August 2016 across increasing lead times

For both events, grid-wise hazard values were computed for forecast lead times of 1, 3, 5, and 10 days, representing the joint probability of observed rainfall exceeding the 99th percentile given forecasted rainfall. For the June 19, 2015 event, lead day 1 exhibits high hazard values (>0.9) concentrated along the Konkan coast, particularly in southern Raigad, indicating strong forecast–observation agreement (Fig.6). The hazard zone expands modestly at lead day 3, while probabilities weaken at lead day 5 and become spatially diffuse by lead day 10. A similar pattern is observed for the August 3, 2016 event, with sharply localised high hazard at short lead times and rapid degradation of predictability at longer leads. Overall, the Konkan coast, especially Raigad, consistently emerges as a high-risk zone, with maximum predictive skill at shorter forecast horizons.

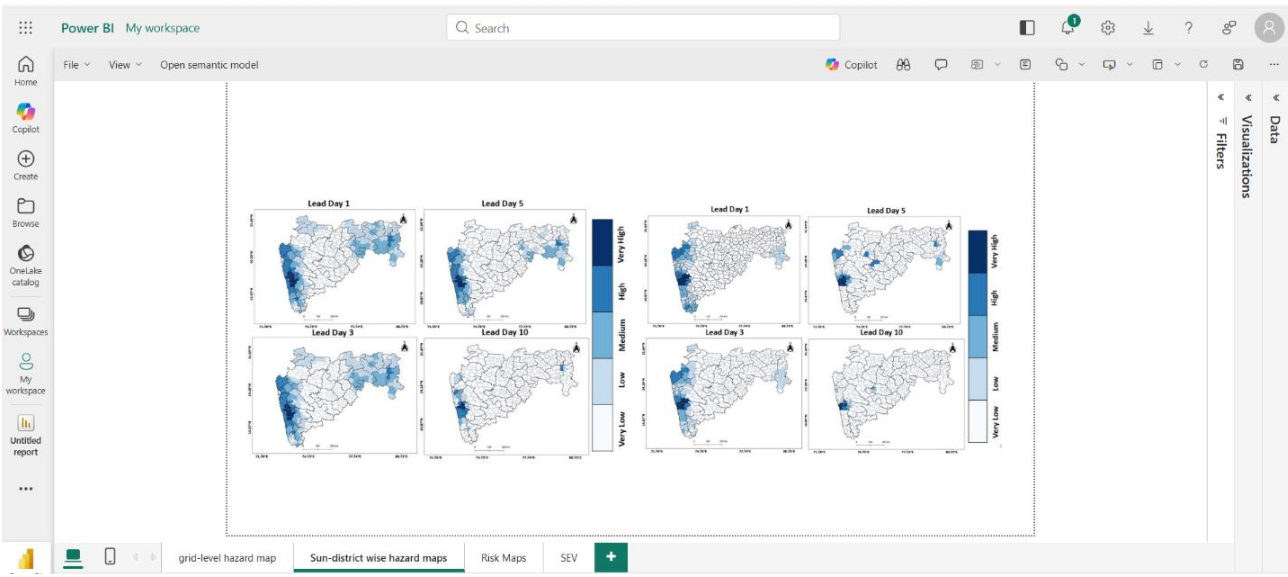

Figure 6. A PowerBI snippet showing the sub-district level hazard classification maps for 19 June 2015 and 3rd August for lead times

3.3 Risks Mapping

The spatial distribution of flood risk for the June 2015 event was analysed across forecast lead times of 1, 3, 5, and 10 days. The risk maps indicate that sub-districts in the Konkan region, particularly Raigad, Ratnagiri, and Sindhudurg, consistently exhibit High to Very High-risk levels at shorter lead times (Lead Days 1 and 3) (Fig.7). This pronounced clustering reflects the convergence of elevated hazard probabilities with high socioeconomic vulnerability. At Lead Day 1, risk is most intense, forming large contiguous Very High-risk zones along the western coast, where extreme rainfall coincided with vulnerable populations and infrastructure. With increasing lead time, particularly by Lead Day 10, both the spatial extent and intensity of risk decline markedly, indicating reduced forecast reliability and weaker alignment between hazard and vulnerability patterns. Notably, parts of the Vidarbha region display Moderate to High risk despite lower hazard levels, underscoring the amplifying role of vulnerability in shaping overall risk outcomes. For the August 2016 event, risk patterns show both similarities and contrasts. At Lead Day 1, the Konkan coast again emerges as the dominant high-risk zone, especially in southern districts such as Ratnagiri and Sindhudurg. By Lead Day 3, although Very High-risk areas contract, much of the western belt remains in High to Medium risk categories. At Lead Day 5, the focus of elevated risk shifts inland towards the Western Ghats and central Maharashtra, notably parts of Satara and Pune. By Lead Day 10, risk becomes sparse and spatially fragmented, reaffirming the importance of short-lead forecasts for effective flood risk preparedness.

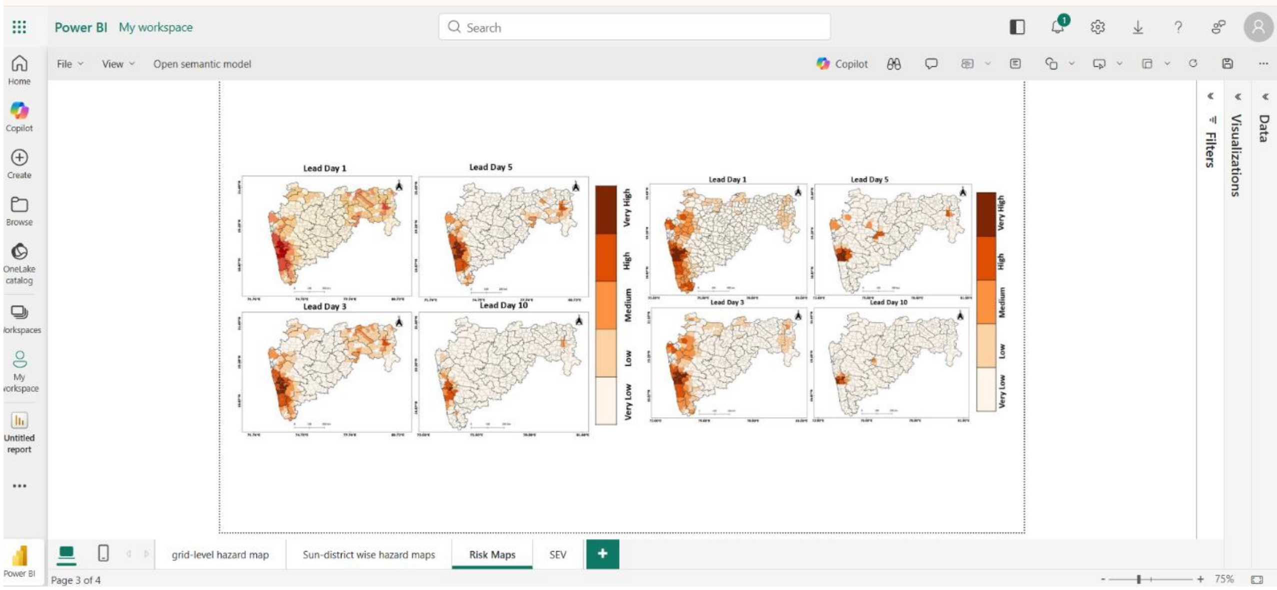

Figure 7. A PowerBI snippet showing sub-district level risk classification maps for 19 June 2015 and 3rd August

4 CONCLUSION

This study develops a high-resolution flood risk assessment framework for Maharashtra by combining sub-district-scale socioeconomic vulnerability (SEV) mapping with probabilistic rainfall hazard modelling, and translating these outputs into an interactive geospatial visualisation and decision-support platform. The resulting SEV surface highlights pronounced vulnerability in Central and Eastern Vidarbha and comparatively lower levels in Konkan and Western Maharashtra. Rainfall-induced hazard was estimated using a bivariate kernel density model of IMD observations and GEFSv12 forecasts (2000–2019), enabling robust characterisation of extreme rainfall probabilities.

Integrating SEV and hazard fields produced spatially explicit flood risk maps that were operationalised through a PowerBI dashboard to improve accessibility, interpretation, and scenario exploration, which are important prerequisites to sustainable adaptation planning. The platform enables policymakers and practitioners to interactively compare regions, analyse risk drivers, and evaluate targeted adaptation options, thereby strengthening evidence-based decision-making.The proposed framework is scalable, adaptable to additional hazard types, and demonstrates how modern visual analytics platforms can bridge the gap between research-oriented risk assessment and operational disaster risk management.

5 ACKNOWLEDGEMENTS

This research was supported by the Department of Science and Technology (DST), Government of India, through the SPLICE–Climate Change Programme (Sponsored Project: DST/CCP/CoE/140/2018), and by the Green Energy and Sustainability Research Hub (GESH) at the Indian Institute of Technology Bombay (Sponsored Project: DO/2024-RHUB004-003; Grant Number: 00000000000050004495). IIT Bombay has provided support with computational resources. The corresponding author gratefully acknowledges the generous support of the donors, Dr Vinaya Kapoor and Dr Samir Kapoor, for endowing the Vinaya & Samir Kapoor Chair Professorship in Climate Studies at the Centre for Climate Studies. The authors also gratefully acknowledge Dr Sohom Mondal for his valuable guidance and insights provided through telephonic and virtual discussions. The authors extend their gratitude to the Census of India and the Government of India for providing the detailed demographic dataset.

REFERENCES

Abdi, H. and Williams, L.J. (2010) Principal Component Analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2, 433-459.

Benson, C., Twigg, J., & Myers, M. (2001). NGO initiatives in risk reduction: an overview. *Disasters*, *25*(3), 199-215.

Bhargava, M. G., Kiran, K. T. P. S., & Rao, D. R. (2018). Analysis and design of the visualisation of the educational institution database using PowerBI tool. *Global Journal of Computer Science and Technology*, *18*(C4), 1-8.

Birkmann, J. (2006) Measuring Vulnerability to Promote Disaster-Resilient Societies: Conceptual Frameworks and Definitions. In: Birkmann, J., Ed., Measuring Vulnerability to Natural Hazards: Towards Disaster Resilient Societies, United Nations University Press, Tokyo, 9-54.

Dottori, F., Szewczyk, W., Ciscar, JC. *et al.* Increased human and economic losses from river flooding with anthropogenic warming. *Nature Clim Change* **8**, 781–786 (2018).

Fritsch, J. M., Houze, R. A., Adler, R., et al. (1998). Quantitative Precipitation Forecasting. *Bulletin of the American Meteorological Society, 79*(2), 285–299.

Groeschl, J., & Noy, I. (2020). *Poverty, Inequality, and Disasters – An Introduction to the Special Issue*. Economics of Disasters and Climate Change, Springer.

IFRC. (2018). *World Disasters Report 2018: Leaving no one behind.*

IPCC. (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation: A special report of Working Groups I and II of the Intergovernmental Panel on Climate Change* (C. B. Field, V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, ... P. M. Midgley, Eds.). Cambridge University Press.

Sakai, Y., Estudillo, J. P., Fuwa, N., Higuchi, Y., & Sawada, Y. (2017). Do natural disasters affect the poor disproportionately? Price change and welfare impact in the aftermath of Typhoon Milenyo in the rural Philippines. *World Development*, *94*, 16-26.

Vittal, H., Karmakar, S., Ghosh, S., & Murtugudde, R. (2020). A comprehensive India-wide social vulnerability analysis: highlighting its influence on hydro-climatic risk. *Environmental Research Letters*, *15*(1), 014005.

Yamamura. (2015). *The impact of natural disasters on income inequality: Analysis using panel data from 1970 to 2004*. International Economic Journal, 29(3), 359–374.

Web References

Asian Development Bank < https://www.adb.org/what-we-do/topics/climate-change/overview#disaster-resilience> (accessed on 28.09.2024)

Census of India (CoI) <https://censusindia.gov.in/census.website/data/census-tables> (accessed on 5.08.2023)

Organisation for Economic Co-operation and Development (OECD) < https://web-archive.oecd.org/temp/201708/374943approachtowardsdisasterriskreductionandresilience.htm> (accessed on 2.08.2023)

United Nations Disaster Risk Reduction (UNDRR) <https://www.undrr.org/gar> (accessed on 15.09.2024)