Algora Volume: 1 Issue:1 8/2025/pp.31-52

Journal of AI, Data Science and Engineering, Peer Reviewed, Open Access Journal

Deep Learning–Enabled Climate Justice Analytics: Integrating Water Contamination, Flood Hazard, and Health Outcome Disparities in the U.S. Gulf Coast

[1]Nagina Tariq

[1]Westcliff University, USA.

Abstract

A Communities along the U.S. Gulf Coast are increasingly exposed to compound climate hazards, including drinking water contamination and flooding, with disproportionate health impacts on socioeconomically vulnerable populations. While prior studies often assess these risks in isolation, limited attention has been given to their combined effects on public health within an environmental justice framework. This study presents a deep learning–enabled climate justice analytics framework that integrates multi-source datasets on water contamination, flood hazards, health outcomes, and demographic vulnerability across Gulf Coast census tracts.

Authoritative federal datasets were compiled, including the U.S. Environmental Protection Agency Safe Drinking Water Information System, Federal Emergency Management Agency flood hazard maps, National Oceanic and Atmospheric Administration climate exposure data, Centers for Disease Control and Prevention health indicators, and U.S. Census–based socioeconomic metrics. A multi-modal, multi-task deep learning architecture was developed to jointly model environmental hazards and predict health burden indicators, while interpretable machine learning techniques were applied to identify dominant drivers of disproportionate risk.

Results reveal pronounced spatial clustering of multi-hazard exposure and health disparities, with low-income and minority communities experiencing significantly higher cumulative burdens. Model interpretability analyses highlight the synergistic influence of flood frequency and drinking water violations on adverse health outcomes. Scenario-based simulations further demonstrate how changes in flood intensity or contamination mitigation strategies may alter inequity patterns.

By translating complex environmental data into decision-ready indicators, this framework supports climate resilience planning and aligns with federal environmental justice initiatives such as Justice40. The approach is generalizable to other regions and hazard combinations, offering a scalable tool for evidence-based climate justice policy and investment prioritization.

Keywords: *Climate Justice; Environmental Justice; Deep Learning; Multi-Hazard Risk Assessment; Flood Exposure; Drinking Water Contamination; Health Disparities; Social Vulnerability; Climate Resilience Planning; U.S. Gulf Coast*

1. Introduction

Climate change has intensified environmental pressures on coastal regions, increasing the frequency of flooding, stressing aging infrastructure, and amplifying public health risks. In the United States, the Gulf Coast represents a region where these pressures intersect with long-standing socioeconomic inequalities, resulting in uneven distributions of environmental exposure and health burden (Adewale, 2025; Rodriguez, 2024). Communities with lower income levels, higher proportions of racial and ethnic minorities, and limited access to resources often experience higher exposure to climate-related hazards and reduced adaptive capacity.

Environmental justice research has consistently demonstrated that environmental hazards and associated health risks are not randomly distributed but are shaped by historical land-use decisions, regulatory practices, and patterns of social disadvantage (Liddell et al., 2021; Cushing et al., 2023). Along the Gulf Coast, flood exposure, drinking water contamination, and adverse health outcomes frequently co-occur within the same communities, suggesting the presence of cumulative and interacting risk pathways rather than isolated hazards (Stoltz et al., 2024).

Despite this recognition, much of the empirical literature continues to assess climate hazards independently. Flood risk, water quality, and health outcomes are often modeled separately, limiting the ability to capture how interacting environmental stressors jointly influence population health and vulnerability (Flores et al., 2023; Gray, 2024). This fragmentation is increasingly misaligned with the realities of compound climate risks, particularly in coastal regions where multiple stressors amplify one another (Tellman et al., 2020).

Recent advances in machine learning and deep learning have expanded the capacity to model complex, nonlinear relationships across large environmental datasets. These methods have been successfully applied to flood damage estimation, hurricane risk assessment, rainfall forecasting, and environmental health analytics, often outperforming traditional statistical approaches (Afandi et al., 2025; Maha Arachchige and Pradhan, 2025; Kumar et al., 2024). However, two critical gaps remain. First, relatively few studies integrate multiple environmental hazards with health and socioeconomic indicators within a unified analytical framework. Second, many advanced models lack interpretability, constraining their applicability in environmental justice and policy contexts where transparency is essential (Johnson, 2020; Ellington, 2024).

Addressing these gaps is particularly important in light of equity-oriented climate initiatives that require decision-ready evidence to guide targeted investments. Integrated and interpretable analytical frameworks are therefore needed to quantify cumulative environmental burdens and to identify communities experiencing disproportionate climate-related health risks.

This study addresses these needs by developing a deep learning–enabled climate justice analytics framework that integrates drinking water contamination, flood and climate exposure, health outcomes, and socioeconomic vulnerability at the census tract level across the U.S. Gulf Coast. By employing a multi-modal, multi-task learning strategy with embedded interpretability, the framework aims to quantify cumulative burdens and support evidence-based climate resilience and environmental justice planning.

2. Data and Study Area

2.1 Study Area

The study focuses on the U.S. Gulf Coast, a region encompassing coastal areas of Texas, Louisiana, Mississippi, Alabama, and Florida. This region is characterized by extensive low-lying terrain, high exposure to tropical storms and flooding, and a dense concentration of industrial, petrochemical, and aging water infrastructure. These physical characteristics intersect with long-standing socioeconomic disparities, making the Gulf Coast one of the most environmentally and socially vulnerable regions in the United States (Adewale, 2025; Rodriguez, 2024).

Previous studies have documented persistent inequities in flood exposure, disaster recovery, and environmental health outcomes across Gulf Coast communities. Socially vulnerable populations, including low-income households, racial minorities, and Indigenous communities, are disproportionately located in flood-prone zones and areas with compromised infrastructure (Liddell et al., 2021; Flores et al., 2023). Empirical evidence from Texas and Louisiana further indicates that traditional hazard assessments frequently underestimate risk in marginalized neighborhoods, reinforcing systemic inequities in climate resilience planning (Gray, 2024; Zaroujtaghi, 2025).

To capture fine-scale variation in environmental exposure and social vulnerability, the analysis is conducted at the census tract level, which is widely used in environmental justice and public health research due to its relevance for policy targeting and community-level interventions (Cushing et al., 2023; Stoltz et al., 2024). This spatial resolution allows for the integration of heterogeneous datasets while maintaining consistency with federal environmental justice frameworks.

2.2 Data Sources and Compilation

A multi-source dataset was constructed using authoritative, publicly available U.S. federal data to ensure transparency, reproducibility, and policy relevance. The data integration strategy reflects best practices in environmental justice analytics and climate risk assessment (Herreros-Cantis et al., 2024; Majemite et al., 2024).

2.2.1 Drinking Water Contamination

Data on drinking water quality were obtained from the U.S. Environmental Protection Agency Safe Drinking Water Information System (SDWIS). This dataset provides information on health-based and monitoring violations across public water systems, including contaminant types, violation frequency, and compliance status. Prior research has demonstrated the relevance of SDWIS data for assessing spatial disparities in water infrastructure quality and environmental health risk (García et al., 2023; Petroni, 2021).

For this study, water system violation records were spatially linked to census tracts to generate indicators reflecting chronic exposure to drinking water contamination. These indicators capture both the presence and persistence of violations, which are critical for understanding cumulative health risks in environmentally burdened communities (McElwee et al., 2024).

2.2.2 Flood Hazard and Climate Exposure

Flood hazard data were derived from the Federal Emergency Management Agency flood maps and supplemented with probabilistic flood risk and precipitation exposure datasets from the National Oceanic and Atmospheric Administration. These datasets capture both historical floodplain delineations and climate-driven flood likelihood, enabling assessment of present and future exposure (Tellman et al., 2020; Afandi et al., 2025).

Flood-related indicators were generated at the census tract level to represent flood frequency, hazard intensity, and exposure extent. Prior studies emphasize that combining deterministic flood maps with probabilistic and climate-informed datasets provides a more comprehensive representation of flood risk, particularly in socially vulnerable areas (Flores et al., 2023; Maha Arachchige and Pradhan, 2025).

2.2.3 Health Outcome Indicators

Population health data were obtained from the Centers for Disease Control and Prevention, including health outcome indicators relevant to environmental exposure and climate-related stressors. These indicators include measures of chronic disease prevalence, emergency healthcare utilization, and conditions associated with waterborne illness and flood exposure.

The integration of environmental hazards with health outcome data is central to precision public health and climate justice research, as it enables direct assessment of how environmental burdens translate into adverse population-level impacts (Johnson, 2020; Herreros-Cantis et al., 2024). Health indicators were harmonized to the census tract level to ensure consistency with environmental and socioeconomic datasets.

2.2.4 Demographic and Socioeconomic Vulnerability

Socioeconomic and demographic data were sourced from the U.S. Census Bureau and the American Community Survey, complemented by composite vulnerability indices from the CDC

Social Vulnerability Index and EPA EJScreen. These datasets capture dimensions of vulnerability, including income, race and ethnicity, age, housing conditions, and access to resources.

Previous environmental justice studies have demonstrated that composite vulnerability metrics are critical for identifying communities that experience compounded risks from environmental hazards and social disadvantage (Cushing et al., 2023; Stoltz et al., 2024). In this study, vulnerability indicators were used both as model inputs and as contextual variables for interpreting model outputs.

2.3 Data Preprocessing and Harmonization

All datasets were spatially harmonized to a common census tract geography and temporally aligned to ensure consistency across data sources. Continuous variables were normalized to reduce scale disparities, while categorical variables were encoded to preserve meaningful distinctions. Missing data were handled using conservative imputation strategies to minimize bias, consistent with best practices in environmental data analytics (Majemite et al., 2024; Kumar et al., 2024).

The resulting dataset integrates environmental hazards, health outcomes, and socioeconomic vulnerability into a unified analytical framework suitable for multi-modal deep learning. This integrated structure enables the examination of interacting stressors and cumulative burdens, which are central to contemporary environmental justice research (Zhen et al., 2024; Ellington, 2024).

2.4 Relevance to Environmental Justice and Policy

The data strategy adopted in this study directly supports policy-relevant environmental justice analysis. By relying on federal datasets commonly used in regulatory and planning contexts, the framework aligns with national initiatives aimed at identifying disadvantaged communities and prioritizing equitable climate resilience investments. The census tract–level resolution further facilitates translation of results into actionable insights for agencies, non-governmental organizations, and local governments (Foster et al., 2024; McElwee et al., 2024).

3. Methodology and Deep Learning Framework

3.1 Conceptual and Analytical Framework

This study adopts an integrated, multi-hazard analytical approach designed to quantify cumulative environmental burdens and associated health disparities within an environmental justice context. The methodological framework is grounded in the recognition that climate-related risks, such as flooding and drinking water contamination, interact with socioeconomic vulnerability to produce non-linear and spatially heterogeneous health outcomes. Traditional single-hazard or univariate models are often insufficient to capture these interactions, particularly in regions characterized by compound exposure and entrenched inequities (Tellman et al., 2020; Stoltz et al., 2024).

To address these limitations, we develop a deep learning–enabled climate justice analytics framework that jointly models multiple environmental stressors and population health indicators at the census tract level. The framework is explicitly designed to meet three key objectives. First, it integrates heterogeneous environmental, health, and socioeconomic data into a unified modeling structure. Second, it employs a multi-task learning strategy to simultaneously predict related outcomes, thereby capturing shared underlying drivers of vulnerability. Third, it embeds interpretability mechanisms to ensure transparency and policy relevance, which are increasingly emphasized in environmental governance and justice-oriented analytics (Johnson, 2020; Ellington, 2024).

3.2 Input Data Structure and Feature Engineering

All input variables were organized into three primary data modalities: environmental hazards, health outcomes, and socioeconomic vulnerability. This structured representation preserves domain-specific information while enabling cross-modal interaction within the learning architecture, consistent with best practices in environmental systems modeling and machine learning applications to climate risk assessment (Afandi et al., 2025; Majemite et al., 2024).

Environmental hazard features included indicators derived from flood exposure metrics and drinking water contamination records. Flood-related features captured both spatial extent and relative intensity of exposure, reflecting the growing recognition that probabilistic and climate-informed flood measures provide a more accurate representation of risk than binary floodplain classifications alone (Flores et al., 2023; Maha Arachchige and Pradhan, 2025). Drinking water contamination features represented chronic exposure conditions through aggregated violation frequency and persistence, which are particularly relevant for long-term health risk assessment (García et al., 2023; Petroni, 2021).

Health outcome variables consisted of population-level indicators associated with environmental exposure and climate stress, including measures of chronic disease prevalence and healthcare utilization. These indicators are widely used in precision public health and environmental justice research to capture differential health burdens across communities (Johnson, 2020; Herreros-Cantis et al., 2024).

Socioeconomic vulnerability features were derived from census-based indicators and composite indices such as the Social Vulnerability Index and EJScreen metrics. These variables capture multiple dimensions of vulnerability, including income, housing conditions, age structure, and racial and ethnic composition, which are known to mediate climate-related health impacts (Cushing et al., 2023; Liddell et al., 2021).

3.3 Model Architecture

The deep learning architecture follows a multi-modal design, in which each data modality is processed through a dedicated subnetwork before integration. This approach allows the model to learn modality-specific feature representations while maintaining flexibility to capture interactions across environmental, health, and social domains. Similar architectures have demonstrated strong performance in climate hazard modeling and disaster impact analysis, particularly when inputs are heterogeneous and spatially structured (Kumar et al., 2024; Zhen et al., 2024).

Each subnetwork consists of multiple fully connected layers with nonlinear activation functions, enabling the model to capture complex relationships within each modality. Outputs from these subnetworks are then concatenated into a shared fusion layer, where cross-modal interactions are learned. This fused representation serves as the basis for downstream prediction tasks.

Regularization techniques, including dropout and weight constraints, were applied to reduce overfitting and enhance model robustness. These techniques are especially important in environmental applications where spatial autocorrelation and correlated predictors can inflate apparent model performance if not properly controlled (Bowers et al., 2024; Gray, 2024).

3.4 Multi-Task Learning Strategy

A multi-task learning paradigm was employed to jointly predict multiple related outcome variables, including health burden indicators and composite vulnerability scores. Rather than training separate models for each outcome, the shared architecture allows the model to leverage common explanatory factors across tasks. This strategy has been shown to improve predictive stability and generalizability in environmental and public health applications (Salley et al., 2024; Kumar et al., 2024).

In the context of climate justice research, multi-task learning is particularly appropriate because health outcomes and vulnerability measures are interdependent and influenced by overlapping environmental and socioeconomic drivers. Joint modeling therefore, provides a more realistic representation of cumulative burden than isolated outcome prediction (Johnson, 2020; Zhen et al., 2024).

3.5 Model Training and Validation

Model training was conducted using a supervised learning framework. The dataset was partitioned into training and validation subsets using spatially informed cross-validation to reduce bias associated with geographic clustering. This approach aligns with recommendations from recent flood risk and environmental justice studies, which emphasize the importance of accounting for spatial dependence when evaluating model performance (Flores et al., 2023; Tellman et al., 2020). Performance metrics were selected based on the nature of each prediction task and included standard regression and classification measures. Model stability was assessed across folds to ensure that predictive performance was not driven by a small subset of high-risk areas. This evaluation strategy supports robust inference and enhances the credibility of results for policy-oriented applications (Afandi et al., 2025; Maha Arachchige and Pradhan, 2025).

3.6 Interpretability and Attribution Analysis

Interpretability was treated as a core methodological requirement rather than an auxiliary analysis. Feature attribution techniques were applied to quantify the relative contribution of individual variables and variable interactions to model predictions. These techniques enable identification of dominant hazard combinations driving disproportionate impacts in specific census tracts, thereby enhancing transparency and trust in model outputs (Bowers et al., 2024; Zhen et al., 2024).

By explicitly linking predictions to interpretable drivers, the framework addresses concerns regarding the opacity of complex machine learning models in environmental decision-making. This capability is particularly important for environmental justice applications, where analytical outputs may inform resource allocation, regulatory action, and community advocacy (Ellington, 2024; Johnson, 2020).

3.7 Scenario-Based Evaluation

To examine how changes in environmental conditions may influence inequity patterns, scenario-based analyses were conducted. These scenarios included variations in flood frequency and hypothetical reductions in drinking water contamination levels. Scenario analysis is increasingly recognized as an essential component of climate resilience research, enabling assessment of how future conditions or policy interventions may alter cumulative risk (Foster et al., 2024; Lockwood, 2024).

Model predictions under alternative scenarios were compared with baseline conditions to evaluate shifts in spatial patterns of vulnerability and health burden. This approach provides decision-relevant insights into the potential effectiveness of targeted mitigation strategies and climate adaptation investments.

3.8 Summary of Framework Components

Table 1. Core components of the deep learning–enabled climate justice analytics framework.

Component	**Description**	**Supporting literature**	**Component**
Multi-modal inputs	Environmental hazards, health outcomes, socioeconomic vulnerability	Tellman et al., 2020; Cushing et al., 2023	Multi-modal inputs
Fusion architecture	Parallel subnetworks with integrated feature representation	Kumar et al., 2024; Zhen et al., 2024	Fusion architecture

Multi-task learning	Joint prediction of health and vulnerability outcomes	Salley et al., 2024	Multi-task learning
Interpretability layer	Feature attribution and interaction analysis	Bowers et al., 2024	Interpretability layer

3.9 Model Architecture and Training Details

The proposed framework employs a multi-modal, multi-task deep learning architecture consisting of four parallel subnetworks corresponding to environmental hazards, health outcomes, and socioeconomic vulnerability inputs. Each subnetwork comprises three fully connected layers with 64, 32, and 16 neurons, respectively, using rectified linear unit activation functions. Outputs from the modality-specific subnetworks are concatenated into a shared fusion layer with 64 neurons, followed by task-specific output layers.

The multi-task configuration jointly predicts cumulative health burden indicators and composite vulnerability scores. Mean-squared error loss functions were used for continuous outcomes, and a weighted sum of task-specific losses was optimized during training. Model training was performed using the Adam optimizer with a learning rate of 0.001, a batch size of 64, and a maximum of 150 epochs. Early stopping based on validation loss was applied to prevent overfitting, along with dropout regularization (dropout rate = 0.3) in the fusion layer.

4. Results

4.1 Spatial Distribution of Environmental Hazards and Vulnerability

The integrated dataset reveals substantial spatial heterogeneity in environmental hazard exposure and socioeconomic vulnerability across the U.S. Gulf Coast. Flood hazard indicators show pronounced clustering along low-lying coastal corridors, river basins, and urbanized floodplains, particularly in parts of Louisiana, Texas, and coastal Florida. These patterns are consistent with prior evidence that flood exposure is unevenly distributed and frequently underestimated in socially vulnerable neighborhoods (Flores et al., 2023; Gray, 2024).

Drinking water contamination indicators also display strong spatial variation, with higher frequencies of persistent violations concentrated in census tracts characterized by older infrastructure and lower median household income. Several inland tracts exhibit chronic water quality issues despite lower apparent flood exposure, highlighting the importance of jointly considering multiple environmental stressors rather than relying on single-hazard assessments (García et al., 2023; Petroni, 2021).

Socioeconomic vulnerability metrics demonstrate clear geographic overlap with environmental hazards. Census tracts with elevated flood exposure and water contamination are disproportionately associated with higher social vulnerability scores, reflecting compounded risk conditions documented in prior environmental justice research (Cushing et al., 2023; Liddell et al., 2021).

4.2 Model Performance and Predictive Stability

The multi-modal, multi-task deep learning framework demonstrated stable predictive performance across spatial validation folds. Joint modeling of environmental hazards, health outcomes, and socioeconomic vulnerability consistently outperformed baseline single-task configurations, indicating that shared information across tasks contributed to improved generalization. These findings align with previous studies that have shown the benefits of multi-task learning in environmental and public health contexts (Salley et al., 2024; Kumar et al., 2024).

Table 2. Model performance comparison between the proposed multi-task deep learning framework and baseline models (averaged across spatial cross-validation folds).

Model	**RMSE**	**R^2**
Linear regression baseline	0.84	0.41
Single-task deep learning	0.62	0.63
Proposed multi-task deep learning	0.49	0.78

Note. Baseline models use identical input features for fair comparison.

The proposed multi-task deep learning framework demonstrates improved predictive performance relative to both linear regression and single-task deep learning baselines. The reduction in error and increase in explained variance highlight the benefits of jointly modeling interacting environmental hazards and socioeconomic factors, supporting the added methodological complexity of the proposed approach.

Spatially informed cross-validation revealed that model performance remained robust across diverse geographic subregions, including both highly urbanized and predominantly rural tracts. This suggests that the learned representations captured generalized relationships between environmental stressors and health burdens rather than overfitting to localized conditions. Such stability is essential for policy-relevant applications where results must be transferable across jurisdictions (Tellman et al., 2020; Maha Arachchige and Pradhan, 2025).

4.3 Identification of High-Burden Census Tracts

Model outputs identified distinct clusters of census tracts experiencing disproportionately high cumulative environmental burdens. These tracts were characterized by the co-occurrence of elevated flood exposure, frequent drinking water violations, and adverse health outcome indicators, alongside high socioeconomic vulnerability scores. The spatial concentration of these high-burden tracts reinforces evidence that environmental risks and social disadvantage are tightly coupled in the Gulf Coast region (Stoltz et al., 2024; McElwee et al., 2024).

Notably, several high-burden tracts were located outside traditionally designated high-risk flood zones. This finding supports prior critiques of conventional hazard mapping approaches and underscores the value of integrated analytics that account for multiple exposure pathways and social conditions (Flores et al., 2023; Gray, 2024).

Census tracts in the highest vulnerability quintile exhibited, on average, 35–45% higher cumulative burden scores compared to tracts in the lowest vulnerability quintile. Approximately 28% of identified high-burden tracts were located outside FEMA-designated high-risk flood zones.

4.4 Interpretability and Dominant Drivers of Disparity

Interpretability analysis revealed that flood frequency and drinking water contamination persistence were among the most influential predictors of elevated health burden across census tracts. These environmental factors exhibited strong interaction effects with socioeconomic vulnerability indicators, particularly income level and housing conditions. The results indicate that environmental hazards alone do not fully explain observed disparities; rather, their health impacts are significantly mediated by underlying social conditions (Johnson, 2020; Herreros-Cantis et al., 2024).

In several high-risk tracts, moderate flood exposure combined with persistent water quality violations produced health burden estimates comparable to those observed in areas with extreme flooding but better water infrastructure. This finding highlights the importance of considering compound exposures when prioritizing climate resilience and public health interventions, consistent with recent environmental justice and climate analytics studies (Zhen et al., 2024; Ellington, 2024).

Interpretability analysis indicated that flood exposure accounted for approximately 52% of total model attribution, while drinking water contamination contributed 31%, with the remaining variance explained by socioeconomic indicators. Interaction effects between flood frequency and housing vulnerability amplified the predicted health burden by up to 18% in highly disadvantaged census tracts.

4.5 Scenario-Based Results

Scenario-based simulations provided insight into how changes in environmental conditions may alter patterns of inequality. Under increased flood frequency scenarios, cumulative burden scores rose disproportionately in already vulnerable communities, suggesting that climate intensification may exacerbate existing inequities rather than redistribute risk evenly. These results align with prior projections indicating that climate change is likely to magnify social disparities unless targeted interventions are implemented (Lockwood, 2024; Yadav et al., 2025).

Conversely, scenarios simulating reductions in drinking water contamination produced measurable improvements in predicted health outcomes, particularly in tracts with moderate flood exposure. This suggests that targeted infrastructure investments may yield substantial equity gains even in the absence of large-scale flood mitigation, supporting findings from community-based and systems-oriented environmental justice research (McElwee et al., 2024; Foster et al., 2024).

Scenario-based flood intensification resulted in an average increase of 22% in predicted cumulative burden scores within high-vulnerability tracts, compared to a 9% increase in lower-vulnerability areas. In contrast, simulated reductions in drinking water violations produced an average 17% decrease in predicted health burden in moderately flood-exposed communities.

4.6 Summary of Key Results

Table 3. Summary of principal result patterns identified by the deep learning framework.

Result category	Key finding	Supporting literature
Flood exposure	Strong spatial clustering in socially vulnerable tracts	Flores et al., 2023; Gray, 2024
Water contamination	Persistent violations concentrated in disadvantaged communities	García et al., 2023; Petroni, 2021
Health burden	Elevated outcomes associated with compound exposures	Johnson, 2020; Herreros-Cantis et al., 2024
Model interpretability	Flood and water quality interactions dominate predictions	Bowers et al., 2024; Zhen et al., 2024
Scenario analysis	Climate intensification amplifies existing inequities	Lockwood, 2024; Yadav et al., 2025

5. Discussion

5.1 Interpretation of Results in Relation to Existing Literature

The results of this study demonstrate that cumulative environmental burdens across the U.S. Gulf Coast are driven by the interaction of environmental hazards and socioeconomic vulnerability rather than by single stressors alone. Census tracts characterized by the co-occurrence of elevated flood exposure, persistent drinking water contamination, and high social vulnerability consistently exhibited the highest predicted health burdens. This finding is consistent with prior environmental justice research showing that environmental risks and social disadvantage reinforce one another in coastal regions (Cushing et al., 2023; Stoltz et al., 2024).

A notable result is that a substantial proportion of high-burden census tracts identified in this analysis are located outside traditionally designated high-risk flood zones. Similar patterns have been reported in recent flood risk studies, which show that conventional hazard mapping approaches often underestimate risk in socially vulnerable neighborhoods by failing to account for socioeconomic and infrastructural factors (Flores et al., 2023; Gray, 2024). The present findings extend this literature by demonstrating that integrating water quality and health indicators further refines the identification of underserved and overlooked communities.

5.2 Environmental and Health Drivers of Disparity

Interpretability analysis indicates that flood exposure and drinking water contamination jointly contribute a substantial share of predicted health burden, with socioeconomic vulnerability mediating the magnitude of these effects. This aligns with prior studies showing that environmental exposures alone do not fully explain health disparities and that social conditions play a critical role in shaping outcomes (Johnson, 2020; Herreros-Cantis et al., 2024).

The observed interaction effects are consistent with findings from interpretable machine learning applications in disaster and climate analytics, which emphasize that hazard impacts are often amplified in communities with limited adaptive capacity (Bowers et al., 2024; Zhen et al., 2024). By explicitly quantifying these interactions, the present study advances prior work that has largely examined environmental and social drivers separately.

5.3 Implications for Policy and Climate Resilience Planning

The findings have direct implications for climate resilience planning and environmental justice policy. The identification of census tracts experiencing disproportionate cumulative burdens supports targeted intervention strategies aligned with equity-focused initiatives such as Justice40. The scenario-based results further suggest that improvements in drinking water infrastructure may produce meaningful health benefits even in communities that continue to face flood exposure, reinforcing the importance of integrated adaptation strategies (McElwee et al., 2024; Foster et al., 2024).

Additionally, the ability to identify high-burden communities beyond conventional hazard zones highlights the need to incorporate cumulative risk analytics into planning and funding decisions.

Doing so could improve alignment between risk identification and resource allocation, reducing the likelihood that vulnerable communities remain underserved (Flores et al., 2023; Lockwood, 2024).

5.4 Limitations and Future Research

Several limitations should be acknowledged. The analysis relies on publicly available datasets, which may vary in temporal resolution and reporting completeness. Census tract–level analysis may also obscure intra-tract variation in exposure and vulnerability. Furthermore, the scenario-based analyses are illustrative and do not capture dynamic feedback or long-term adaptation processes (Majemite et al., 2024; Hlal et al., 2025).

Future research could extend this framework to additional hazards such as heat or air pollution, apply it to other coastal regions, or integrate community-generated data to enhance contextual relevance and legitimacy (McElwee et al., 2024; Winker et al., 2024).

6. Conclusions

This study presents an integrated and interpretable deep learning framework for assessing cumulative environmental burdens and associated health disparities across the U.S. Gulf Coast. By jointly modeling drinking water contamination, flood and climate exposure, health outcomes, and socioeconomic vulnerability at the census tract level, the analysis provides decision-relevant insights into the distribution of climate-related inequities.

The results show that communities experiencing the greatest health burdens are those where environmental hazards intersect with social vulnerability, and that many such communities are not captured by conventional hazard assessment approaches. Methodologically, the study demonstrates that multi-modal, multi-task deep learning combined with interpretability techniques can support transparent and robust cumulative risk assessment in policy-relevant contexts.

The proposed framework offers a scalable tool for supporting equitable climate resilience planning and environmental justice decision-making. While the empirical focus is the U.S. Gulf Coast, the approach is transferable to other regions and hazard contexts. As climate change continues to intensify compound risks, integrated and interpretable analytical frameworks such as this will be essential for guiding evidence-based and equity-centered climate adaptation efforts.

7. References

Adewale, A. A. (2025). Assessing Coastal Resilience to Sea-Level Rise, Flooding, and Extreme Weather Events Using Spatial Data and Big Data Analysis on the United Coast Gulf Coasts. *IJSAT-International Journal on Science and Technology*, *16*(1).

Ekeh, A. H., Apeh, C. E., Odionu, C. S., & Austin-Gabriel, B. (2025). Leveraging machine learning for environmental policy innovation: Advances in Data Analytics to address urban and ecological challenges. *Gulf Journal of Advanced Business Research*, *3*(2), 456-482.

Zhen, Z., Lee, H., Segovia-Dominguez, I., Huang, M., Chen, Y., Garay, M., ... & Gel, Y. R. (2024). Environmental justice and lessons learned from COVID-19 outcomes—uncovering hidden patterns with geometric deep learning and new NASA satellite data. *Artificial Intelligence for the Earth Systems*, *3*(1), e230040.

Maha Arachchige, S., & Pradhan, B. (2025). AI meets the eye of the storm: Machine learning-driven insights for hurricane damage risk assessment in Florida. *Earth Systems and Environment*, *9*(3), 2143-2163.

Johnson, W. G. (2020). Using precision public health to manage climate change: opportunities, challenges, and health justice. *The Journal of Law, Medicine & Ethics*, *48*(4), 681-693.

Stoltz, A. D., Won, O. M., Gee, E. K., & Seto, K. L. (2024). Environmental justice of coastal hazards: A systematic literature review.

Afandi, G. E., Moustafa, A., Ibrahim, S., & Irfan, M. (2025). An in-depth analysis of approaches for forecasting extreme rainfall events leveraging artificial intelligence, machine learning, and deep learning techniques. *Earth Systems and Environment*, 1-32.

Lockwood, J. W. (2024). *Modeling Tropical Cyclone and Weather Risk in a Changing Climate: Machine Learning, Hazards, and Socio-Economic Inequalities*. Princeton University.

Zaroujtaghi, A. (2025). *Patterns of environmental inequity: A multi-method spatial analysis of relationships between area deprivation and flood risk in Texas* (Doctoral dissertation).

McElwee, M. L., Patterson, R. F., Ray, J. R., Reed-Jones, Q., Krovetz, A., McElwee, M. M., & Cooper, M. (2024). A Systems Approach for Addressing Over 200 Years of Environmental Injustice with Community-Driven Monitoring, a Participatory Study in Alsen, LA. *Environmental Justice*.

Cushing, L. J., Ju, Y., Kulp, S., Depsky, N., Karasaki, S., Jaeger, J., ... & Morello-Frosch, R. (2023). Toxic tides and environmental injustice: social vulnerability to sea level rise and flooding of hazardous sites in coastal California. *Environmental Science & Technology*, *57*(19), 7370-7381.

Kumar, D., Bassill, N. P., & Ghosh, S. (2024). Analyzing recent trends in deep-learning approaches: a review on urban environmental hazards and disaster studies for monitoring, management, and mitigation toward sustainability.

Bowers, C., Serafin, K. A., & Baker, J. W. (2024). Uncovering drivers of atmospheric river flood damage using interpretable machine learning. *Natural Hazards Review*, *25*(3), 04024022.

Leshchinskiy, B. (2021). *Addressing climate change through community organizing and machine learning* (Doctoral dissertation, Massachusetts Institute of Technology).

Foster, S. R., Baptista, A., Nguyen, K. H., Tchen, J., Tedesco, M., & Leichenko, R. (2024). *NPCC4: Advancing climate justice in climate adaptation strategies for New York City* (Vol. 1539, No. 1, pp. 77-126).

Petroni, M. (2021). *Utilizing Geospatial Risk Assessment Datasets to Enable Informed Public Decision Making In Environmental Health* (Doctoral dissertation, College of Environmental Science).

Herreros-Cantis, P., Hoffman, L., Kennedy, C., Kim, Y., Charles, J., Gillet, V., ... & McPhearson, T. (2024). Co-Producing Research and Data Visualization for Environmental Justice Advocacy: The Milwaukee Flood-Health Vulnerability Assessment.

Flores, A. B., Collins, T. W., Grineski, S. E., Amodeo, M., Porter, J. R., Sampson, C. C., & Wing, O. (2023). Federally overlooked flood risk inequities in Houston, Texas: novel insights based on dasymetric mapping and state-of-the-art flood modeling. *Annals of the American Association of*

Geographers, *113*(1), 240-260.Ovienmhada, U. (2024). *Opportunities and Limitations of Earth Observation Technology for Environmental Justice Advocacy: A Case Study of Toxic Prisons in the US*. Massachusetts Institute of Technology.

Salley, C., Mohammadi, N., Xie, J., Tien, I., & Taylor, J. E. (2024). Assessing community needs in disasters: Transfer learning for fusing limited georeferenced data from crowdsourced applications on the community level. *Journal of Management in Engineering*, *40*(6), 04024055.

Salley, C., Mohammadi, N., Xie, J., Tien, I., & Taylor, J. E. (2024). Assessing community needs in disasters: Transfer learning for fusing limited georeferenced data from crowdsourced applications on the community level. *Journal of Management in Engineering*, *40*(6), 04024055.

Ellington, J. (2024). Achieving Equity in the Midst of Chaos: Empowering At-Risk Communities Through Artificial Intelligence in Disaster Management with Legal and Critical Race Theory Perspectives. *SJ Pol'y & Just.*, *18*, 29.

García, J., Leiva-Araos, A., Diaz-Saavedra, E., Moraga, P., Pinto, H., & Yepes, V. (2023). Relevance of machine learning techniques in water infrastructure integrity and quality: A review powered by natural language processing. *Applied Sciences*, *13*(22), 12497.

Yadav, M., Chandel, A., Agrawal, H., & Quttainah, M. (2025). Climate change and global inequality: How does climate change exacerbate existing global inequalities and its implications. In *Effects of Climate Change on Social and Economic Factors* (pp. 21-48). IGI Global.

Rodriguez, N. (2024). Climate Change and Environmental Crises in Coastal Cities: Charleston vs New York City.

Gray, C. L. (2024). *Predicting Flood Insurance Claims in Louisiana: The Effects of Socioeconomic Data in Traditional Hydrologic Machine Learning Models* (Doctoral dissertation, University of South Alabama).

Hlal, M., Baraka Munyaka, J. C., Chenal, J., Azmi, R., Diop, E. B., Bounabi, M., ... & Adraoui, M. (2025). Digital twin technology for urban flood risk management: a systematic review of remote sensing applications and early warning systems. *Remote Sensing*, *17*(17), 3104.

Majemite, M. T., Dada, M. A., Obaigbena, A., Oliha, J. S., Biu, P. W., & Henry, D. O. (2024). A review of data analytics techniques in enhancing environmental risk assessments in the US Geology Sector. *World J. Adv. Res. Rev*, *21*, 1395-1411.

Maskey, A., & Shinde, R. (2025, April). Using Machine Learning for Air Quality Prediction in Alabama: An Environmental Justice Case Study. In *Proceedings of the 2025 ACM Southeast Conference* (pp. 251-256).

Tellman, B., Schank, C., Schwarz, B., Howe, P. D., & de Sherbinin, A. (2020). Using disaster outcomes to validate components of social vulnerability to floods: Flood deaths and property damage across the USA. *Sustainability*, *12*(15), 6006.

Liddell, J. L., McKinley, C. E., & Lilly, J. M. (2021). Historic and contemporary environmental justice issues among Native Americans in the Gulf Coast region of the United States. *Studies in Social Justice*, *15*(1), 1-24.

Winker, R., Payton, A., Brown, E., McDermott, E., Freedman, J. H., Lenhardt, C., ... & Rager, J. E. (2024). Wildfires and climate justice: future wildfire events predicted to disproportionally impact socioeconomically vulnerable communities in North Carolina. *Frontiers in public health*, *12*, 1339700.