Socio-Environmental Data Analytics for Sustainable Communities

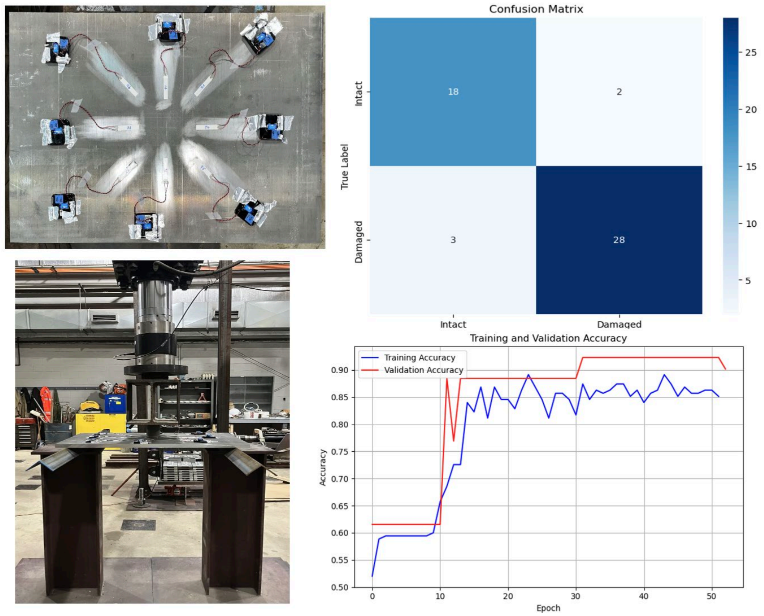
Hadi Salehi
Assistant Professor of Civil Engineering
Louisiana Tech University

November 7, 2025

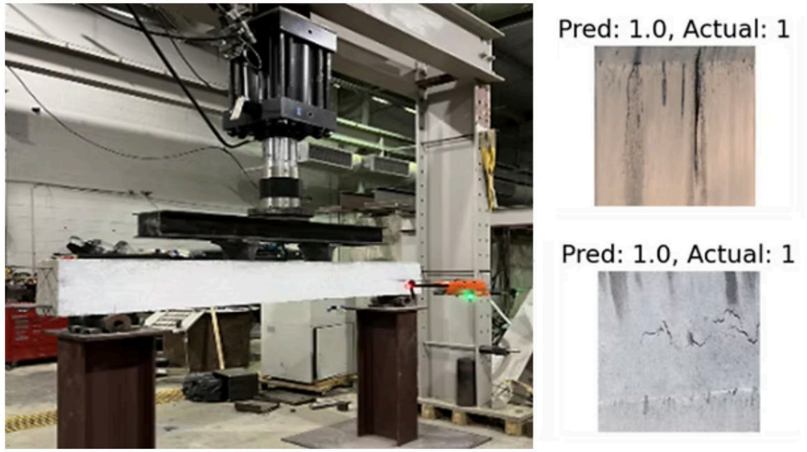
Intelligent, Data-driven, Emerging, & Adaptive Systems Technology (IDEASTech) Lab

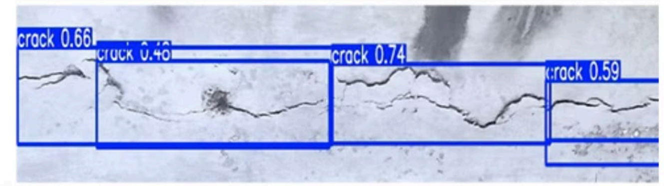
Development of adaptive technology solutions to enhance the resilience of the built environment based on the integration of *energy harvesting*, *multifunctional materials*, *real-time monitoring*, *robotics*, and *data analytics*

(a) Experimental Tests (b) Classification Results



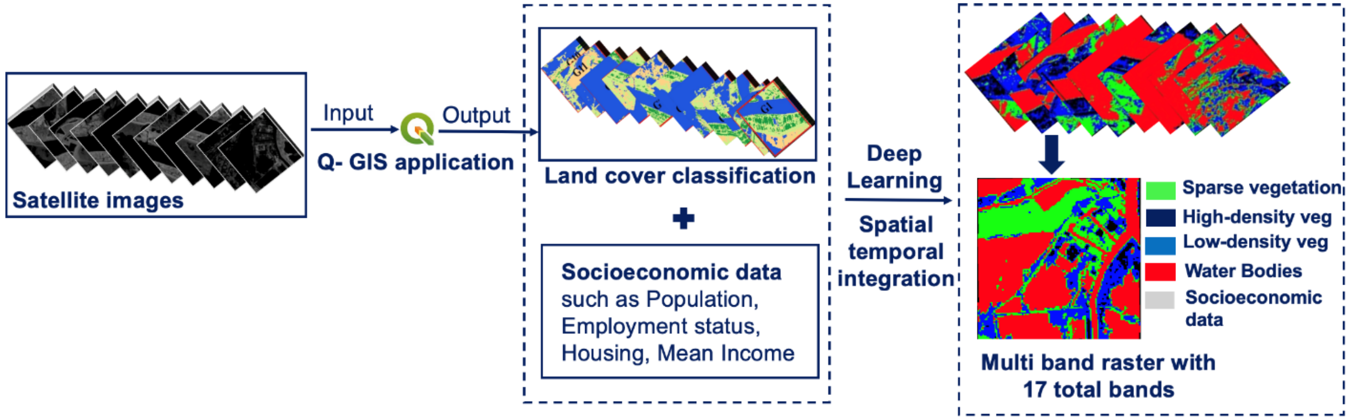
Self-powered sensing for structural monitoring of space habitats





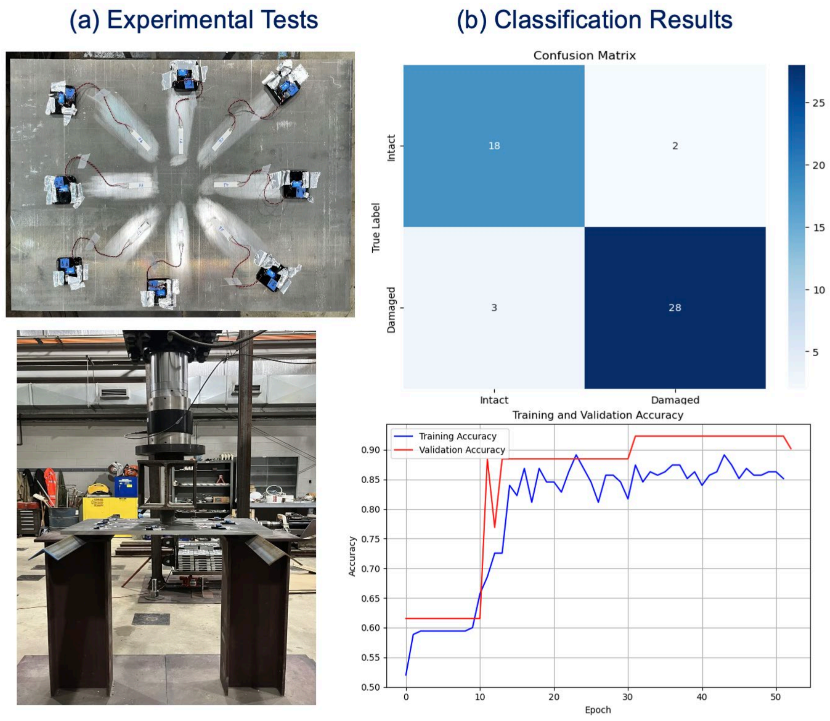
Bridge monitoring utilizing computer vision and robotics

Advanced data analytics for mitigating the impacts of human-induced changes on coastal communities



LaSPACE REA Program

Autonomous Monitoring of Space Habitats Utilizing Self-Powered Sensing and Advanced Data Analytics Towards Deep Space Exploration - LaSPACE REA Program



Self-powered sensing for structural monitoring in space habitats

NASA EPSCoR

Towards mitigating human-induced changes on coastal communities employing advanced data analytics

This project aligns with **NASA's Science Mission Directorate (SMD**) vision of promoting scientific knowledge to tackle coastal resilience and equity justice.

This work is classified under the research focus area "Impacts of human activity on coastal physical, geomorphological and ecological Variability."

Background and Motivation: Coastal Vulnerability in Southern Louisiana

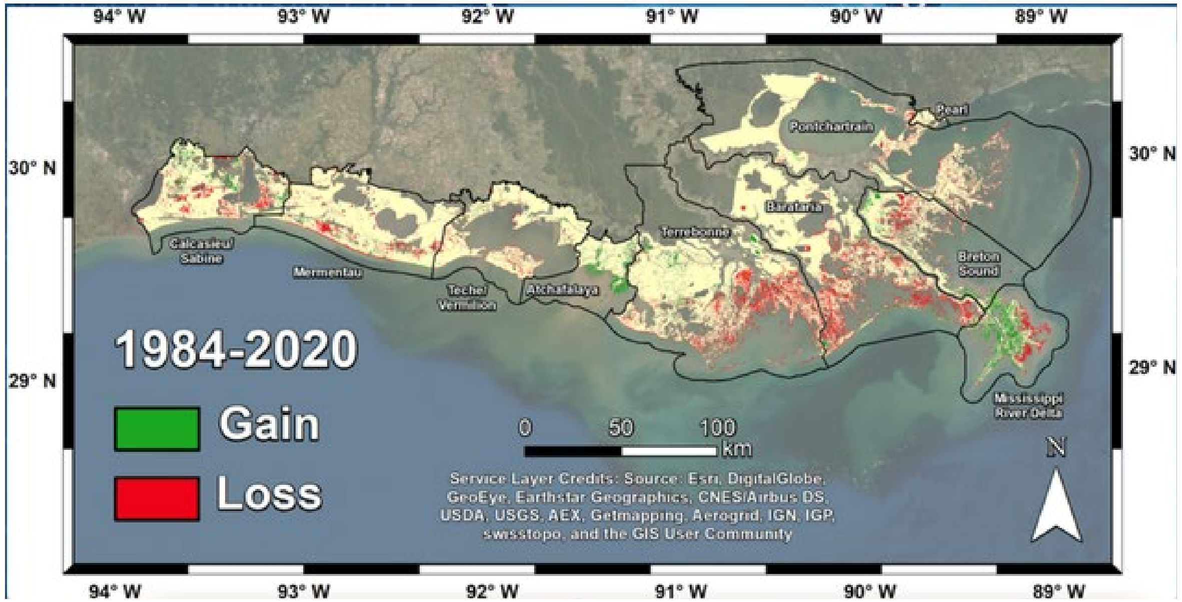
- **Severe Land Loss:** Over **1.2 million acres** of Louisiana's coastal land have vanished since the 1930s due to erosion, subsidence, and sea-level rise

- **Human-Induced Stress:** Urban expansion, industrial activities, and agriculture continue to degrade wetlands and accelerate habitat loss

- **Frequent Natural Disasters:** Intensifying **hurricanes, floods, and storm surges** repeatedly damage infrastructure and disrupt livelihoods

- **Socio-Economic Fragility:** Communities in southern Louisiana, particularly **Plaquemines Parish**, face low-income levels, aging infrastructure, and limited adaptive capacity

- **Critical Research Gap:** Existing studies focus *either* on environmental or socio-economic aspects, **few integrate both spatially and temporally** to capture real vulnerability dynamics



Land loss (red) and gain (green) across coastal Louisiana from 1984–2020 (Source: Esri, USGS, CNES/Airbus DS)

Knowledge Gap

- **Existing Limitation:** Most previous studies analyze *either* environmental hazards (e.g., land loss, flooding) *or* socio-economic vulnerability (e.g., income, housing, education) in isolation

- **Missing Integration:** Very few models **combine environmental and socio-economic factors** to understand how they interact *spatially* and *temporally* in shaping community vulnerability

- **Modeling Gap:** Traditional GIS or statistical methods fail to capture **non-linear patterns**, **temporal dependencies**, and **multi-scale interactions** that drive vulnerability dynamics

- **Need for Innovation:** A data-driven framework is required to **link remote-sensing data with socio-economic indicators** and learn evolving vulnerability behavior

Research Objectives

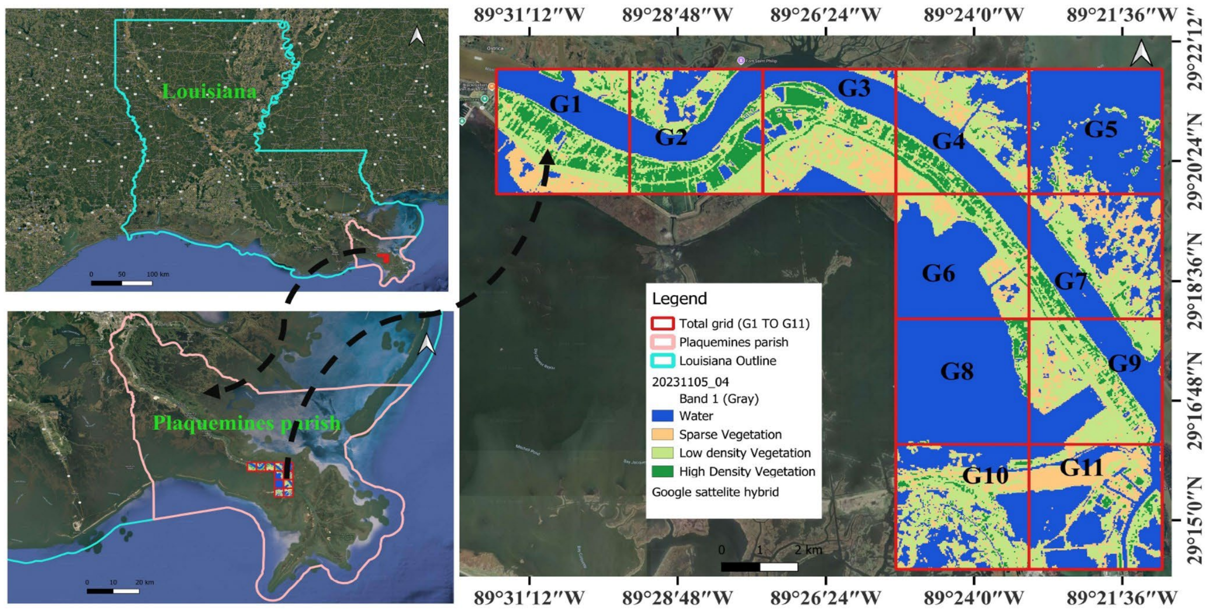
• Develop a data-driven model that rasterizes and co-registers socioeconomic indicators with multi-sensor remote sensing, enabling grid-level spatial-temporal learning

• Integrate environmental features (e.g., vegetation, water bodies, land use) with socioeconomic factors (e.g., income, population density) using multi-temporal raster data collected from Plaquemines Parish, Louisiana, spanning 2017 to 2023

• Deploy a convolutional neural network (CNN) architecture with a and long short-term memory (LSTM) network to make quantitative, grid-level predictions of vulnerability

Data Collection and Processing – Area of Study

- Located in southeastern Louisiana, forming part of the Mississippi River Delta
- A highly vulnerable coastal region experiencing severe land loss, flooding, and erosion
- Faces socio-economic challenges such as low income, aging infrastructure, and limited resilience capacity
- The study area is divided into 11 grids (G1–G11), each covering roughly 12 km², for detailed spatial analysis
- Each grid was used to monitor land use and land cover changes from 2017 to 2023 using satellite and socio-economic data
- The region serves as a representative case of coastal vulnerability in southern Louisiana



Study area map showing Louisiana, Plaquemines Parish, and grid divisions (G1 to G12) with classification of land cover

Data Collection and Processing – Data Acquisition Source

- **Satellite Data:**
  - Landsat-8 OLI imagery acquired from USGS Earth Explorer for years 2017 and 2023.
  - A six-year bi-temporal window (2017–2023) was selected to capture measurable, long-term human-induced changes while minimizing the influence of short-term environmental fluctuations

- **Socioeconomic Data:**
  - Extracted from the U.S. Census Bureau (ACS 5-year dataset) at block-group level.
  - Variables include income, poverty rate, education level, housing, and employment indicators.
  - Each variable was normalized and aggregated to match the 12 km^2 grid scale (G1–G11)

- **Integration Process:**
  - All datasets projected to WGS-84 (EPSG:4326) and aligned using QGIS
  - Create a multiband raster stack combining environmental and socio-economic features (17 total bands)
  - This enables the CNN–LSTM model to learn complex relationships between, for example, vegetation loss (from NDVI) and socioeconomic stress (from income or housing data)

Summary of satellite and socio-economic datasets used for the study

Data	Data types	Resolution
Satellite images	Landsat sensor: 8 OLI Band 3 (Green: 0.53–0.60 µm), Band 4 (Red: 0.64–0.67 µm), Band 5 (NIR: 0.85–0.88 µm)	30 m pixel size
Socio-economic data	Population, Employment status, housing units, mean income, poverty rate, male, female, bachelor's degree, graduate degree, high school degree, less than high school, rented houses and owned houses	12 km^2

Data Collection and Processing – Socioeconomic Data

- The satellite images were categorized into four bands: *water, sparse vegetation, low-density vegetation,* and *high-density vegetation*

Socio-economic indicators across 11 grids (G1–G11) for 2017 and 2023 in Plaquemines Parish, Louisiana

Owned Houses–Band 17		Rental Houses–Band 16		less than High School–Band 15		High School Degree–Band 14		Graduate Degree–Band 13		Bachelor Degree–Band 12		Female Population–Band 11	
2023	2017	2023	2017	2023	2017	2023	2017	2023	2017	2023	2017	2023	2017
186	210	10	35	121	92	358	375	0	28	29	44	254	283
20	77	6	7	10	32	51	142	0	12	0	54	41	120
58	104	12	10	20	48	184	181	0	6	0	34	88	164
56	78	11	7	18	40	188	131	0	0	0	9	80	125
0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	63	8	6	14	32	152	105	0	0	0	8	64	101
37	50	7	5	11	26	122	85	0	0	0	6	51	80
0	0	0	0	0	0	0	0	0	0	0	0	0	0
72	65	10	35	53	35	213	158	0	15	0	46	125	123
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0

Male Population–Band 10		Poverty rate (%)–Band 9		Mean Income ($)–Band 8		Employment status (Pop 16 yr and)		Housing Unit–Band 7		Population–Band 5	
2023	2017	2023	2017	2023	2017	2023	2017	2023	2017	2023	2017
220	266	18.6	12.3	40856	32600	369	391	336	317	474	549
15	135	42	8.28	22728	30952	31	142	56	108	56	255
63	140	29.1	13.2	23724	34281	123	223	100	144	151	304
65	86	9.96	10.7	14743	22397	127	179	85	106	145	211
0	0	0	0	0	0	0	0	0	0	0	0
53	70	8.08	8.71	11951	18156	103	146	69	86	117	171
43	57	6.49	7	9600	14584	83	117	56	69	94	137
0	0	0	0	0	0	0	0	0	0	0	0
117	128	14.1	17.4	69949	55418	179	205	107	123	242	251
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0

Data Processing – Spatial-Temporal Integration

• This study integrates multi-temporal Landsat imagery (2017–2023) with socioeconomic indicators to analyze environmental change and community risk

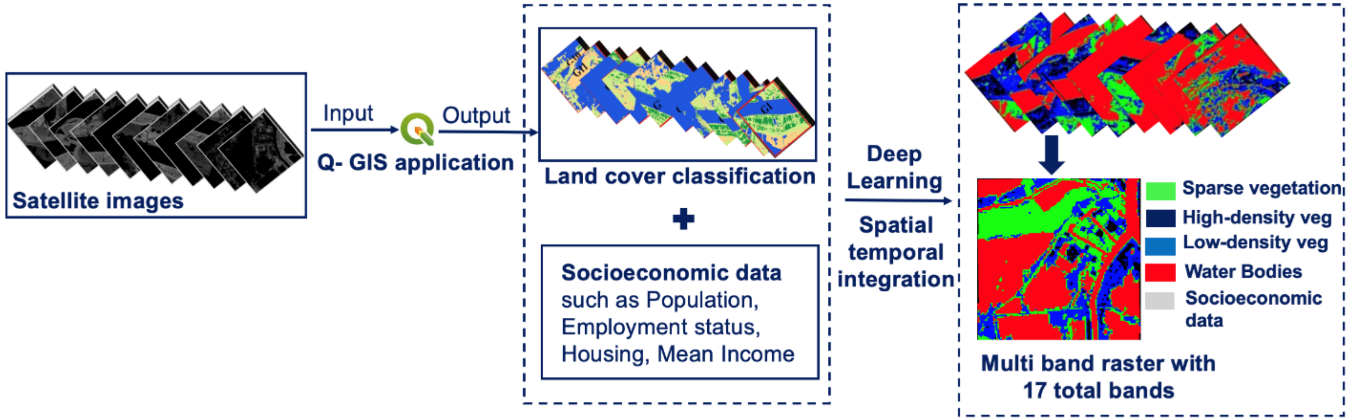
• Combine QGIS-based satellite image processing with socioeconomic data to construct 17-band raster as model inputs

•To unify data formats for spatial analysis we need rasterization

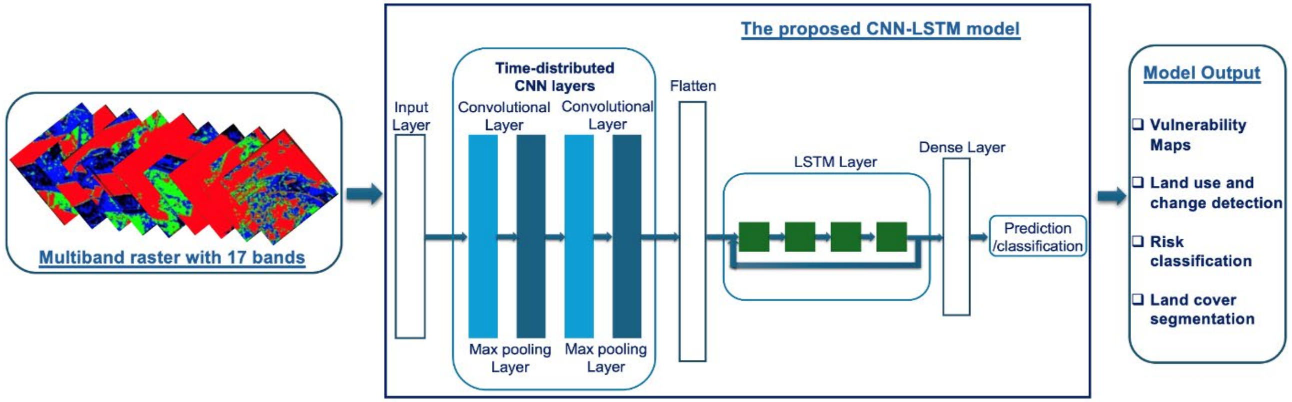
•Satellite and aerial images (e.g., NDVI, NDBI, land use maps) are stored as *rasters*, grids of pixels, each with a value representing a physical quantity (like vegetation index

•Socioeconomic data (population, income, housing, etc.) typically come as *vector or tabular data* tied to irregular polygons

• Rasterization converts those vector polygons into the same grid-based format as the environmental rasters

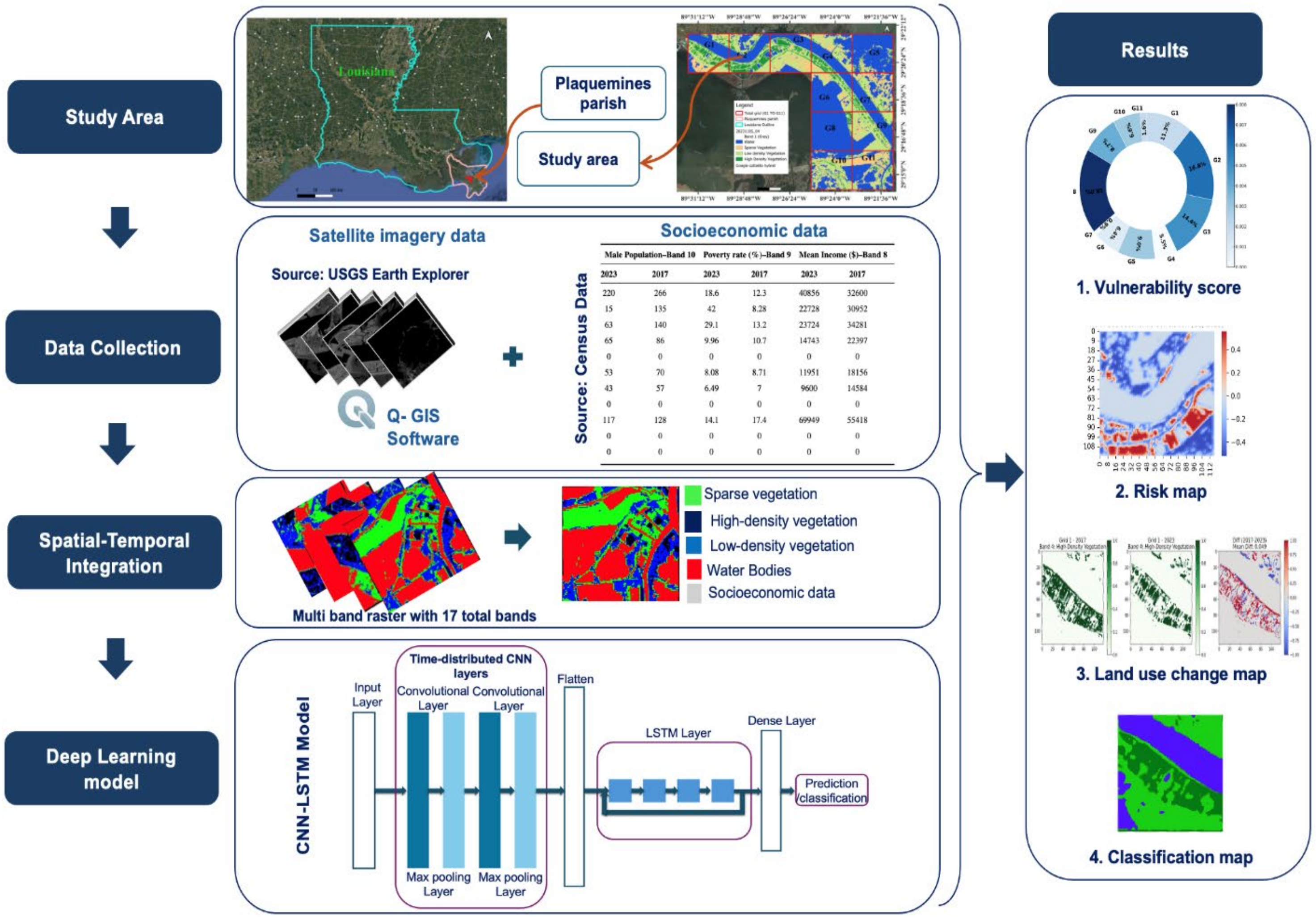


Data Processing Workflow for Multi-Band Raster Creation



Hybrid CNN–LSTM architecture integrating environmental and socio-economic data for vulnerability prediction

Methodology



Proposed CNN–LSTM model integrating environmental and socio-economic data for vulnerability mapping

- **Data Collection and Processing:** Collect satellite images (Landsat-8, NDVI, land cover) and socio-economic data (population, income, housing, education); align all datasets to a 1 km × 1 km grid for southern Louisiana; cleaned, normalized, and pre-processed in QGIS and Python
- **CNN – Spatial Learning:** Extract spatial patterns such as vegetation cover, water expansion, and land-use variation using convolutional layers
- **LSTM – Temporal Learning:** Model temporal trends from 2017 → 2023 to capture environmental and socio-economic changes over time
- **Prediction and Mapping:** Integrate CNN and LSTM outputs to classify vulnerability levels (Low / Medium / High) and generate vulnerability and land-use change maps for Plaquemines Parish
- **Output:** Create final vulnerability maps highlighting high-risk coastal zones to support data-driven resilience planning

Model Optimization and Hyperparameter Tuning

Key parameters tuned:

• *Learning rate:* 0.001 – 0.0001 (stability–speed balance)

• *Batch size:* 16, 32, 64 (trade-off between memory and convergence)

• Epochs: 100 – 300 (monitored validation loss to avoid overtraining)

• Dropout rate: 0.2 – 0.4 (regularization to improve generalization)

• Optimizer: Adam (best convergence for CNN–LSTM hybrid)

Model Optimization and Hyperparameter Tuning

• The model was tested with three train–test splits: 75/25, 80/20, and 85/15

➢ The vulnerability heatmap shows that the 75/25 and 80/20 splits produced stable, consistent patterns, while the 85/15 split caused overfitting and instability due to excessive training bias

•The 80/20 split provided the best performance, balancing accuracy and generalization while avoiding overfitting

➢ **For Land use change detection: MSE (Mean Squared Error):**

•80/20: 0.01895 → most accurate predictions
•75/25: 0.01976 → close performance
•85/15: 0.02444 → higher error

➢ For classification tasks, 80/20 achieved:

•Training Accuracy: 88.06%
•Validation Accuracy: 87.04%

•The 85/15 split, though showing high accuracy (≈90%), was less generalizable

•The 80/20 configuration ensured **model stability, reliable learning**, and minimal variance between training and testing

Results

1.Vulnerability Mapping

• The CNN–LSTM model combined 17 satellite bands with socio-economic data to predict grid-level vulnerability (G1–G11)

• The predicted vulnerability is visualized using multi-temporal satellite and socio-economic inputs

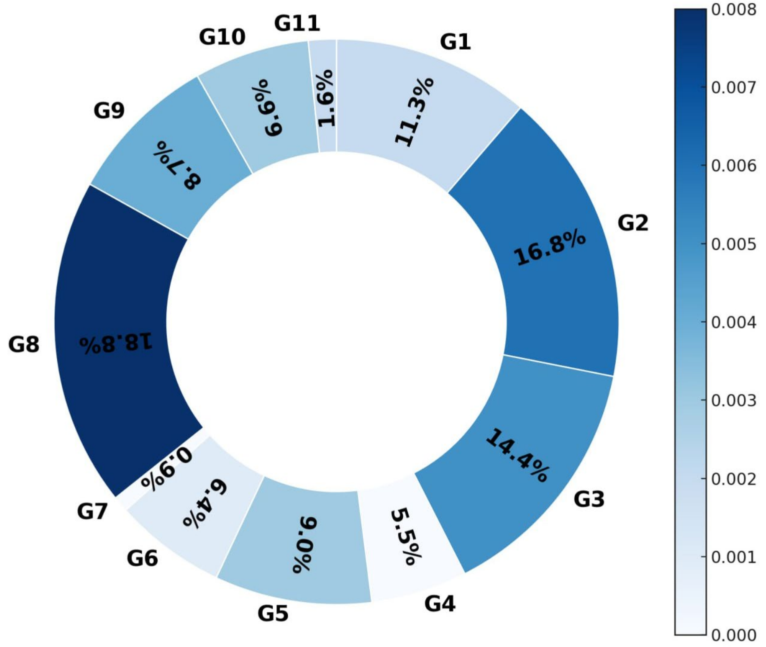
•The color bar shows vulnerability scores, **darker colors indicate higher vulnerability**.

• **Grid G8 (18.8%)** shows the highest vulnerability; **Grid G7 (0.9%)** shows the lowest

• High-risk zones occur around **Buras and Venice** in Plaquemines Parish

• Areas with better vegetation and income exhibit **lower vulnerability**

• The model captures **spatial variation** missed by traditional static analyses



Predicted grid-level vulnerability distribution (G1–G11) using CNN–LSTM model

Results

2. Land-Use and Change Detection

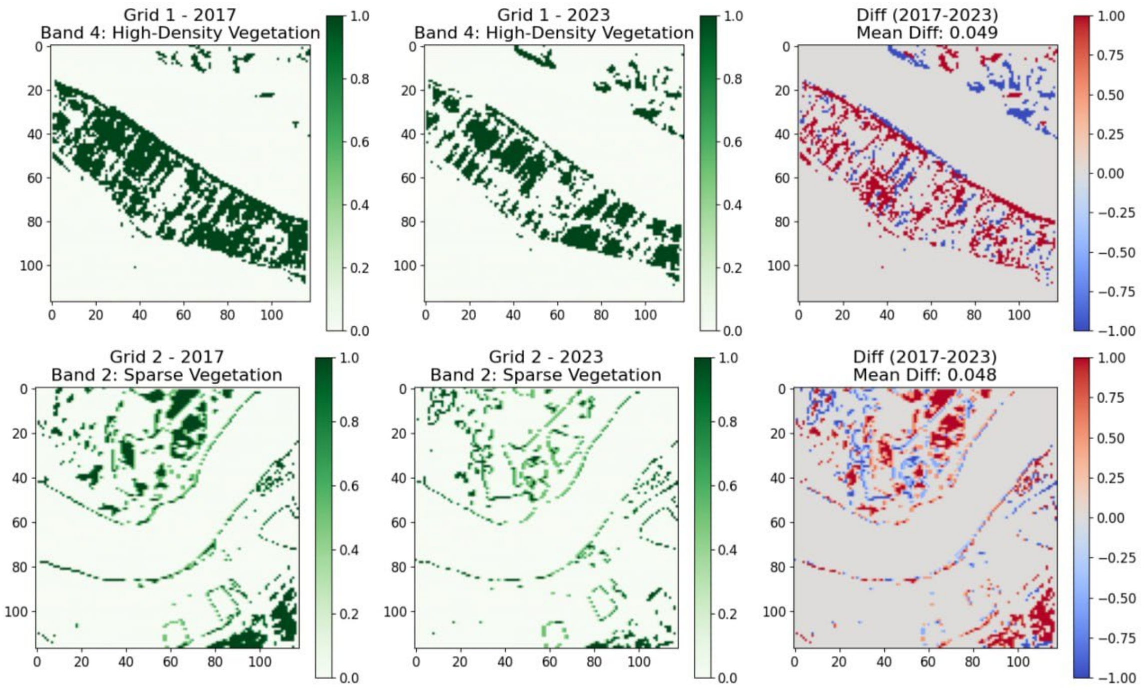
- The **CNN–LSTM model** analyzed raster data from **two time points (2017 → 2023)** to detect key spectral changes and major land-use transitions

- The model assessed **land-use change across 11 grids (G1–G11)** using 17-band satellite imagery

- Results for **Grid 1 and Grid 2 highlight the most changed bands** between 2017 and 2023

- The **area of each spectral band** (vegetation, water, urban, soil, etc.) was calculated for all 11 grids

- **Change charts** validated predictions, confirming the model's ability to capture vegetation decline and water-body expansion



Land Use Change Detection and Analysis for Environmental and Urban Planning

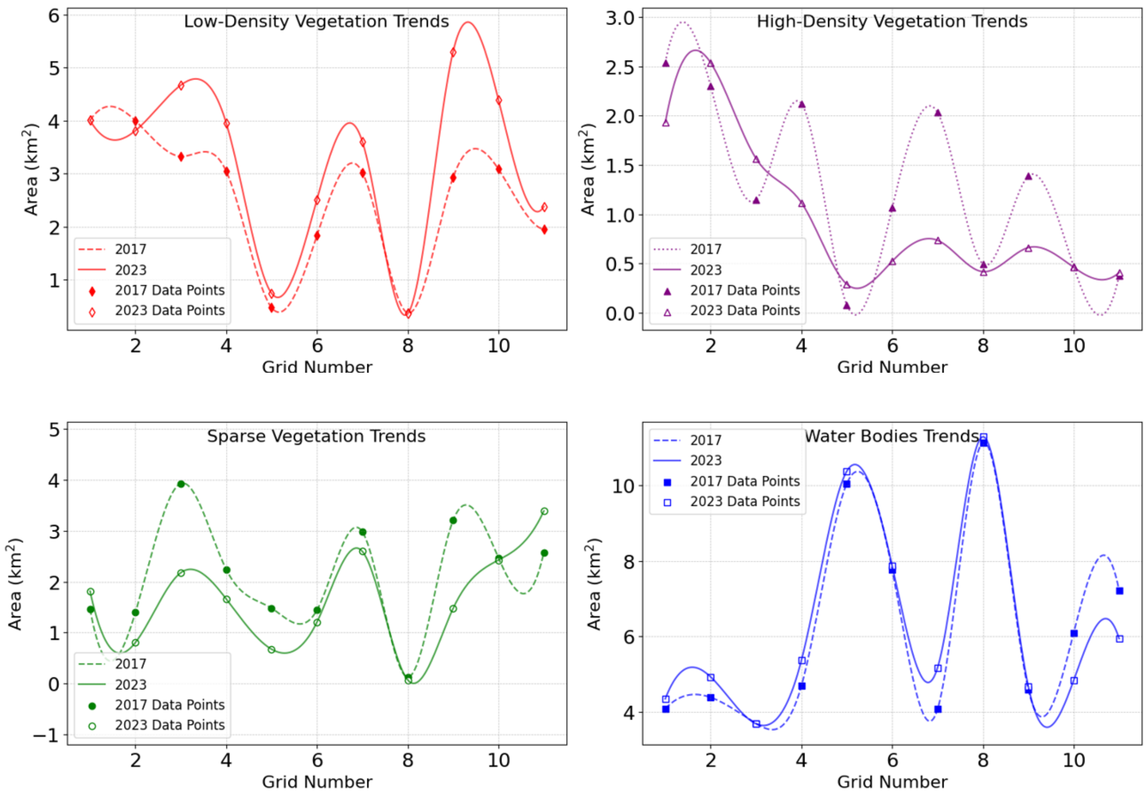
Results

2. Land-Use and Change Detection

• This study integrates QGIS to validate model outputs and ensure the accuracy of vulnerability predictions in coastal areas

• Model predictions are compared against spatial data derived from QGIS to visualize trends in low-density vegetation, high-density vegetation, sparse vegetation, and water bodies for 2017 and 2023

• By reliably identifying areas where vegetation and land cover are most affected, the model can help detect early warning signs of human-induced pressures such as urban expansion or deforestation



Trends in Land Cover and Vegetation Across Coastal Grids (2017-2023)

Results

3. Risk Classification

• The model used **NDVI and socio-economic data (2017–2023)** to study how vegetation relates to community conditions

• **Red areas**: Show **high NDVI linked with better socio-economic status** in Grids 1–4 and 8, while in Grids 5–11 they mainly represent water bodies or mixed transitional zones

• **Dark blue areas**: High vegetation but **low socio-economic activity**, typical of **rural or undeveloped areas**

• **Light blue zones**: **Urban regions** with less greenery due to development or land degradation

• **Gray/orange areas**: **Water bodies or mixed zones** with weak correlation between vegetation and people

• **Grids 1, 2, 3, 4, and 8**: Show **positive correlation**, healthy vegetation with higher income and infrastructure

• **Grids 5–11**: Show **inverse or complex patterns**, sparse vegetation and lower economic activity

• Overall, the results show how **urbanization and land management** drive environmental and social risk in southern Louisiana

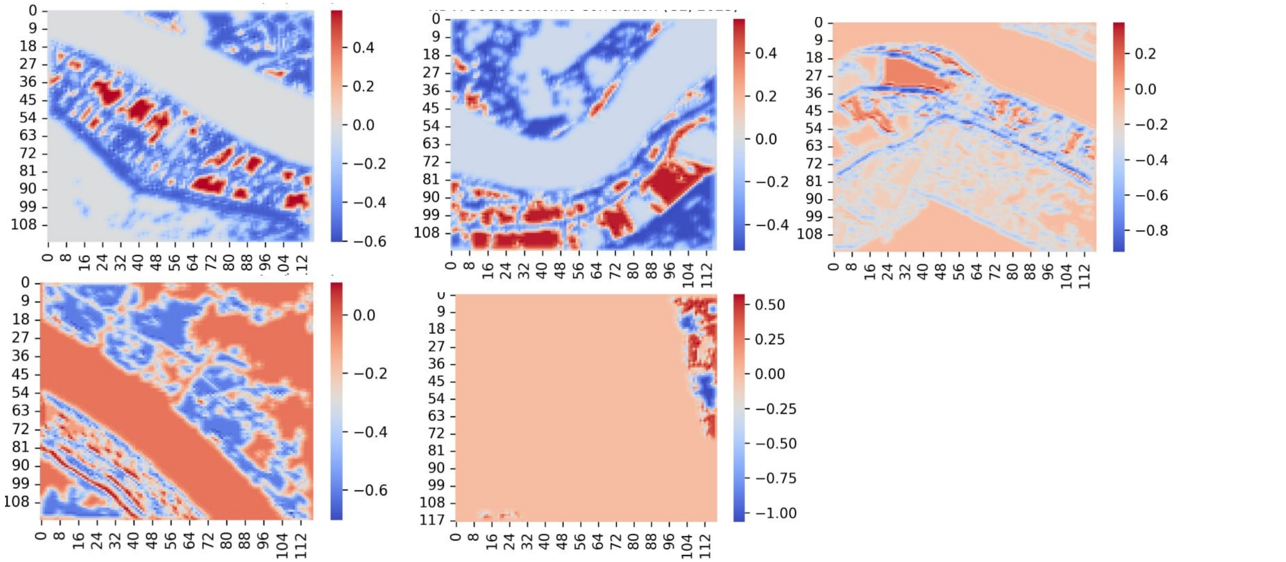
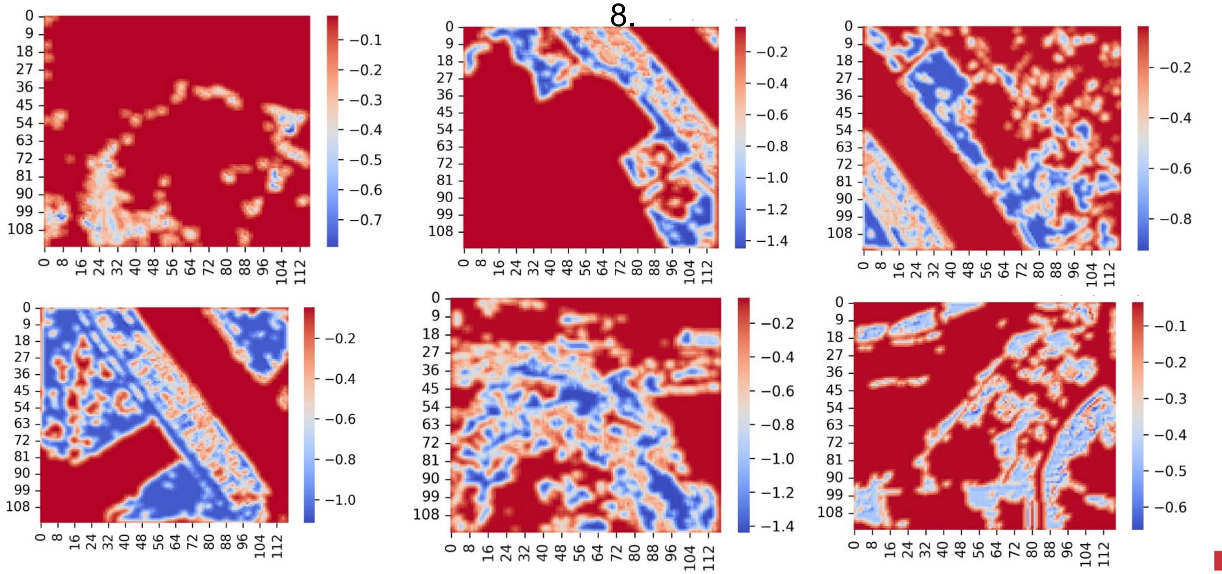


Figure 8: Risk maps for vegetation index and socioeconomic factors for Grids 1, 2, 3, 4, and 8.



Risk maps for vegetation index and socioeconomic factors for Grids 5, 6, 7, 9, 1 and 11

Results

4. Land Cover Classification and Urban Segmentation

• The model segmented land cover using **raster data (2017–2023)** into three classes:
Water (blue), Low Vegetation (lime green), and High Vegetation (green)

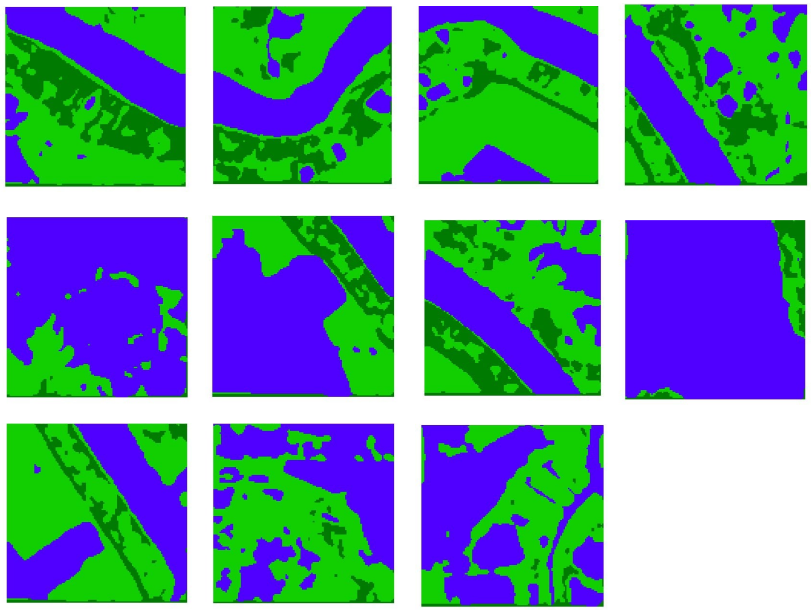
• **Grids 1–8:** Dominated by **vegetation**, with mixed zones of low and high vegetation and smaller water patches

• **Grids 9–11:** Show **more complex distributions** with higher water coverage and transitional landscapes

• **Grid 9** shows a significant mix of high and low vegetation, suggesting **combined urban and natural influences**

• The segmentation maps clearly highlight **vegetation–water balance** and reveal **land-use transitions** across the coastal region

• These classified maps provide valuable insights for **environmental monitoring, land management, and coastal planning**



Spatial Distribution and Segmentation of Land Cover Across Grids 1 to 11

Discussion

- The CNN–LSTM framework captured the spatial and temporal dynamics of environmental and socio-economic change across southern Louisiana
- Vulnerability mapping revealed that areas with high vegetation loss and low income face the greatest risk, aligning with observed coastal degradation zones
- Land-use change detection confirmed that urban growth and wetland reduction directly contribute to increased vulnerability
- Risk classification demonstrated clear patterns between vegetation health and socio-economic status, showing how community well-being is tied to environmental stability
- Segmentation results provided precise visualization of land-cover transitions, essential for targeted coastal restoration and flood-risk planning
- Together, these findings highlight the interdependence of environmental and human systems, reinforcing the need for data-driven, spatially adaptive management strategies

Conclusion

- The CNN–LSTM model effectively integrates environmental and socio-economic data for spatial vulnerability prediction in southern Louisiana

- The approach captures nonlinear spatial–temporal interactions, improving accuracy over traditional GIS-based assessments

- Results confirm that land loss, vegetation decline, and low income jointly increase vulnerability in Plaquemines Parish and nearby coastal zones

- The developed framework can guide coastal restoration, infrastructure planning, and disaster preparedness

Future Work

- Incorporate real-time satellite data for dynamic vulnerability monitoring
- Conduct household-level interviews and field surveys to collect ground-truth socio-economic data for model calibration and precision
- Apply Bayesian uncertainty modeling to improve reliability of predictions
- Integrate with policy and resilience planning tools for statewide application

Acknowledgment

- This research is supported by NASA grant #80NSSC24M0125
- PhD students Rashmi Bhaila, Nazila Jamshidighadikolaei, Ishwor Saud have worked on this project
- NASA Ames Center

Thank you!