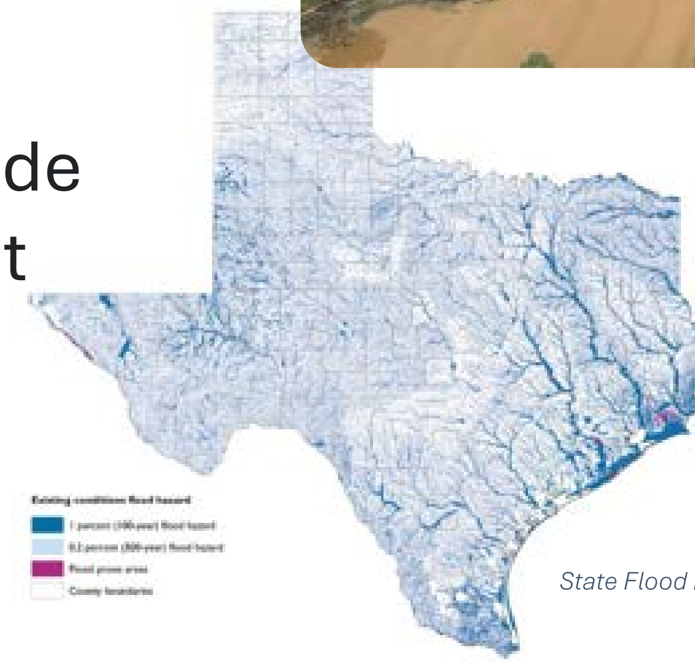
Shingles to Pixels: Texas Building Footprint Project

November 6th, 2025

DianaJoyce Ojeda &
Nick Wellbrock

Introduction

- Texas growth + disaster declarations → urgent need for updated data
- **Goal:** Create publicly, robust, available statewide building footprint dataset
- **Benefits:** Flood risk management + multi-sector planning

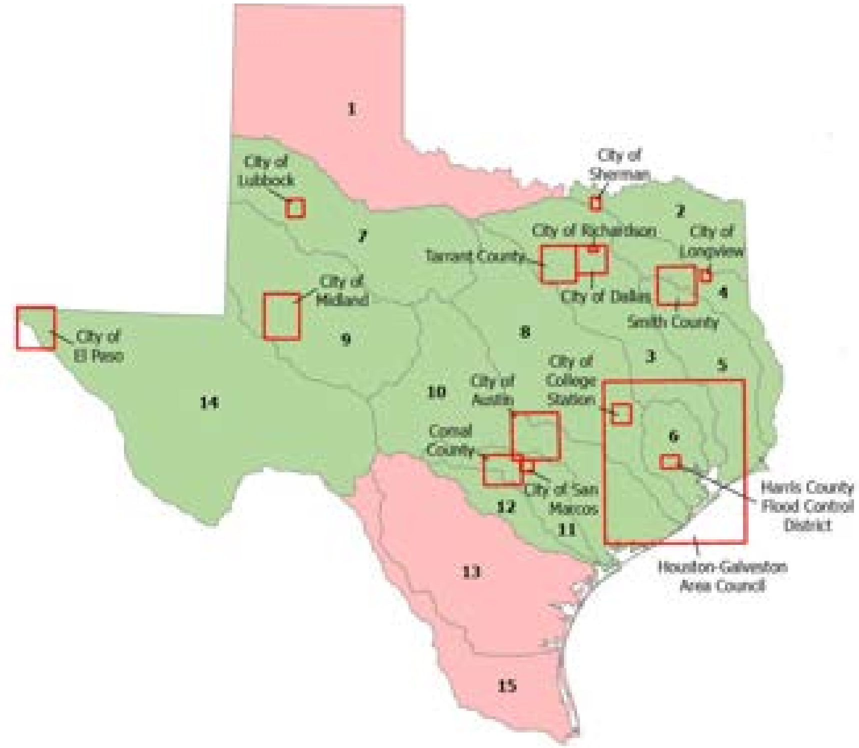


State Flood Plan 2024, TWDB

Driving Factors

- Current statewide dataset:
  - Outdated (2010-2021)
  - Gaps in coverage
  - Better tech/solutions for collection
- 200k+ new permits/year means many missing structures.
- Important for emergency response, infrastructure, and insurance.

Total Spatial Coverage of Acquired Regional Building Footprint Datasets

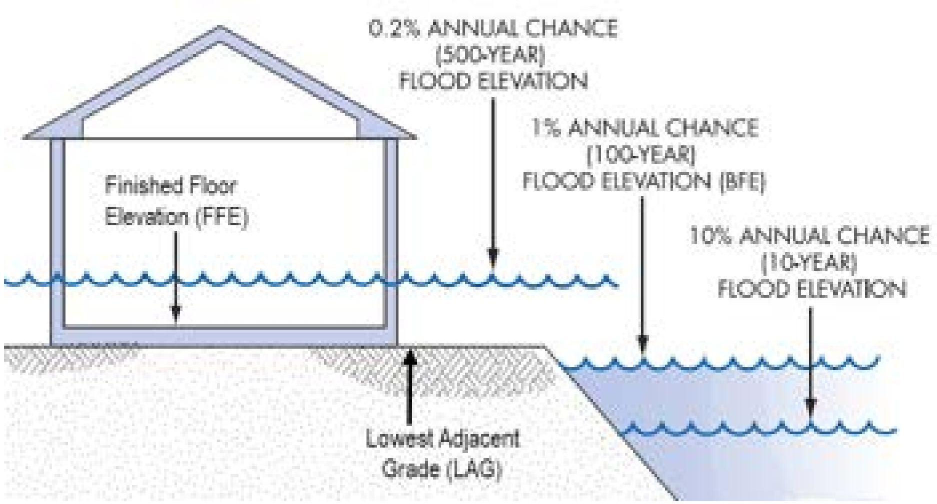


Building footprint dataset available

Building footprint dataset unavailable

Stakeholder Input

- Interview and Surveys in and out of Texas
- Private companies, academic institutions, governing bodies, existing BF stewards for other organizations
- Flood related analysis overwhelming the greatest use
- FFE (finished floor elevation) most requested attribute
- Structure/occupancy type followed closely



Flood Insurance Solutions, 2020

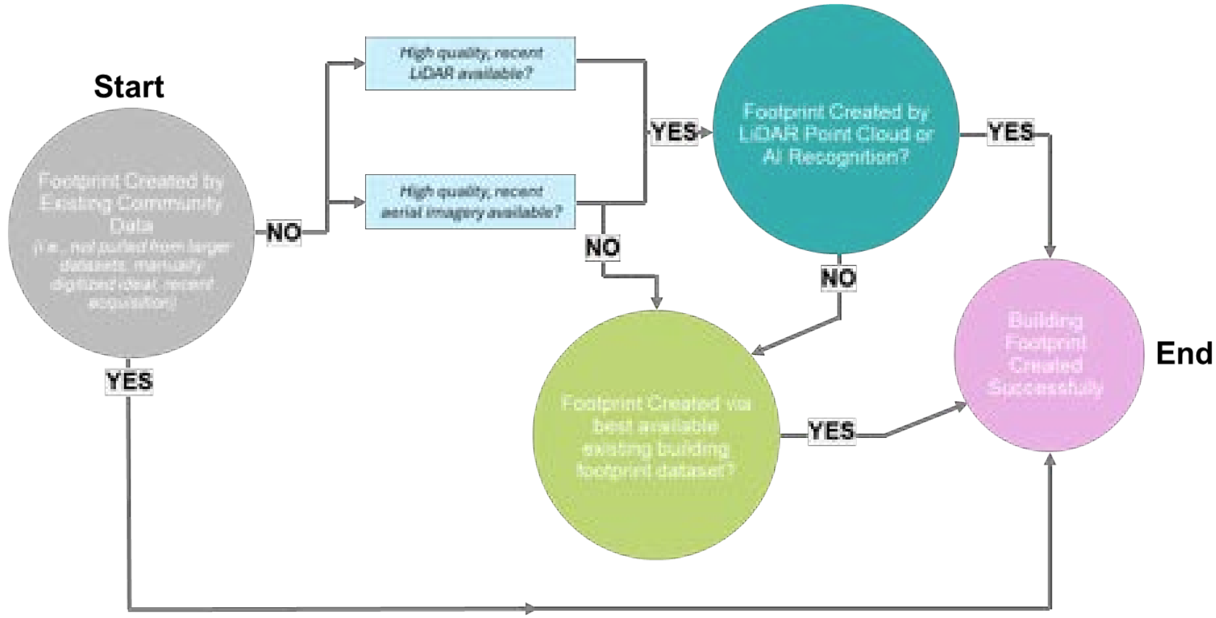
Current Schema & Domains

Group	Field	Description	Role in Flood Risk
Structure Characteristics	BuildingID	Numerical unique ID	Reference to polygon/row
	Building Type	HAZUS building construction material type	Structure durability and resilience
	Square Footage	Building living area in square feet	Structure size
	Number of Stories	Number of floors	Structure size
	Occupancy Type	HAZUS building and facilities occupancy type	Economic activity at risk
	FFE	First Floor Elevation	Minimum flood elevation
	FFE Type	Type of survey used to obtain First Floor Elevation	Measurement source/accuracy
	HAG	DEM-derived Highest Adjacent Grade	Maximum ground elevation at structure
	LAG	DEM-derived Lowest Adjacent Grade	Minimum ground elevation at structure
	LiDAR Source	LiDAR collection used	Measurement source/accuracy
	Foundation Type	Structure foundation type	Structure durability and resilience
	Max Building Height	Maximum building height	Structure size
	Min Building Height	Minimum building height	Structure size
	Year Built	Year structure was built	Structure durability and resilience
Data Sources	Year Built Source	Year Built attribute source	Measurement source/accuracy
	Average Building Height	Average building height	Structure size
	BestAvaila	Is the LiDAR collection used the best available	Measurement source/accuracy
	Height Source	Type of survey used to obtain building heights	Measurement source/accuracy
	Tax Year	Reported tax year for land parcel	Measurement source/accuracy
Characteristics of Underlying Area	SVI	Social Vulnerability Index	Additional risk type
	POL_NAME1	Political name of community	Reference to geography
	Postal Code	Structure zip code	Reference to geography
	DFIRM_ID	County FIPS code	Reference to flood study
	Flood Zone	Flood hazard zone	Reference to flood hazard type
	Source	Building footprint origin	Measurement source/accuracy

TEXAS WATER DEVELOPMENT BOARD

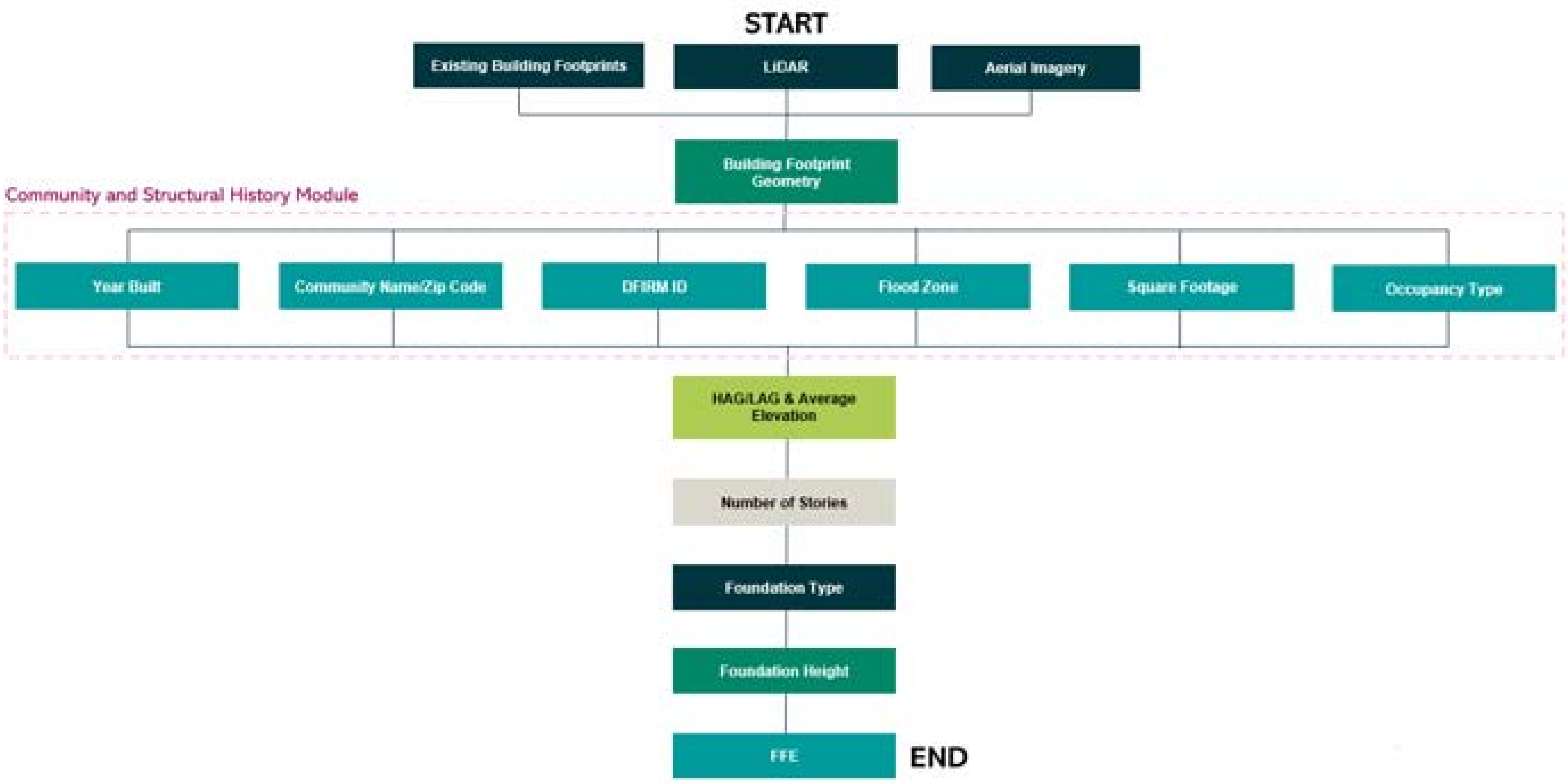
Technical Approach

- **Workflow Development** - Series of modules covering building footprints and structure attributes



- **Decisions Matrix**
  - What should be tested? Informed by literature review
  - Scored on suitability by project team -Highest prioritized attributes used

Methodology



Pilot Study Objectives

- **Pilot Study Geographies:** Selected diverse regions across Texas to test building footprint methods.
- **Key Factors Assessed:** Building footprint extraction (Structure density, environmental conditions, structure types, imagery/LiDAR availability), and general data availability for necessary attributes (FFE, HAG/LAG, structure characteristics, etc.)
- **Data Prioritization:** Evaluated quality and usefulness of available data to guide dataset selection and necessary effort.

Pilot Results - Prosper

Selection Criteria: typical suburban setting with dense housing and low tree cover.

Method Changes

- Split building footprints by parcel to distinguish closely spaced homes.
- Process was effective but computationally intensive (took several hours – 205 acres).

Building Identification and Attribute Results

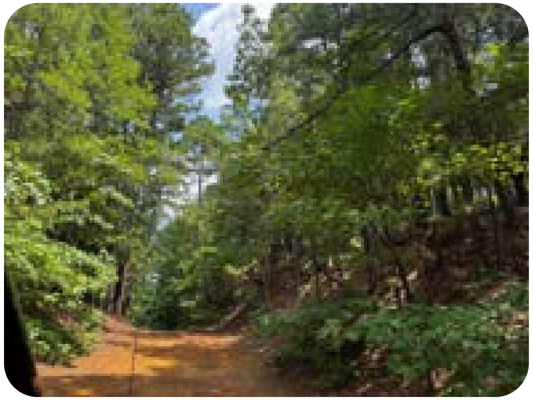
- 153 total structures identified.
- Good LiDAR and parcel coverage/vintage made for easy extraction of many required attributes

Limitations

- Clustered starter homes may merge into single outlines.
- Large, irregular homes (e.g., "McMansions") are harder to delineate accurately.
- Some parcel data from local CAD proved inaccurate

Pilot Results - Carthage

Selection Criteria: Rural area with low-density housing and high tree cover.



Method Changes

- Aerial model retrained to detect tree-covered buildings and trailers.
- Default 1.0 cell resolution used (homes not closely spaced).
- Footprints not split by parcels to preserve shape.

Building Identification and Attribute Results

- 193 total structures identified.
- Rural location led to older LiDAR and imagery coverage – complicating BF and attrribute collection results
- Parcel data for the area lacks details of more developed pilot areas

Limitations

- Tree shadows reduced aerial detection accuracy.
- Most features identified via LiDAR, not imagery.

Pilot Results - Seabrook

Selection Criteria: Coastal area with moderate housing density and many stilted structures.

Method Changes

- Improved building number assignment using overlap metrics.
- Substituted missing data (e.g., year built, ZIP, square footage).
- Estimated stories and foundation height where data was unavailable - LiDAR

Building Identification Results

- 124 total structures identified.
- Elevation Certificates (EC's) allowed some interpretation of foundation heights
- Occupany estimated based on local CAD/Parcels

Limitations

- Shadows from elevated buildings distorted aerial imagery.
- Foundation height estimation was difficult due to limited EC availability and manual integration requirements.

Pilot Results - El Paso

Selection Criteria: Mountainous, desert region with steep terrain and sparse vegetation.

Method Changes
- Aerial model retrained for desert terrain and to exclude carports.
- LAS points filtered: only structures >3m height and >41.8m^2 area included.
- Footprints not split by parcels to preserve shape.

Building Identification and Attribute Results
- 161 total structures identified.
- Structure stories derived from LiDAR
- Occupany type derived from HAZUS generalized products
- Census block group median year built used wherelocal data was missing.

Limitations
- Parcel-based splitting caused footprint loss in multi-parcel buildings.
- Tree cover minimal, but still affected aerial detection accuracy.

Final Products

- **Pilot Study Results** – GDBs including building footprints, attributes, and QC documentation
- **Schema** – blueprint of how the future BF dataset attributes should be laid out
- **Project Report** – documents literature review, stakeholder interviews, technical approach development (including assessed data and methods), and pilot study results, lessons learned, and recommendations.
- **User Guide -** step by step of how to replicate the pilot studies, necessary changes, complications, future development and maintenance considerations.

Main Takeaways

- LiDAR + imagery combo increases accuracy.
- Attribute gaps remain (FFE hardest to scale) → may require more investment in AI tech to scale
- Combined data approach is required since statewide coverage and consistency of data does not yet exist
- Training and process development for different regional characteristics is necessary
- Scalable approach for statewide implementation

Final Thoughts

- Robust building footprint dataset = better flood planning and resilience.
- Potential future steps: statewide scaling, automation, stakeholder collaboration, and long-term maintenance.

Questions?

DianaJoyce Ojeda
Dianajoyce.Ojeda@twdb.texas.gov

TEXAS WATER DEVELOPMENT BOARD

Nick Wellbrock
nicholas.Wellbrock@aecom.com

AECOM