Buildings Information & Land-use Dataset (BILD)
Technical Working Group

March 3, 2025
10:00-11:00

Attendees: Julia McWest, Sanjaya kumar, Catherine Adler, Richard Bell, Christine Zampa, Jayme Breschard, Juan Junca, Beth O'Reilly, Christina Croll, Stephanie Notar, Dr. Catherine Lawson, Michael Saliani, Thomas LaLiberte Jr, James Samons, Daniel Rozell, Jeffrey Langella, Peter Lauridsen, Bethany Greenaway, Shelby Meyers, Matt Miller, Gwen LaSelva, Chris Leo, Christa Hay, Joshua Clyburn, Brian Wright, Karis Manning, John Lee, Elizabeth Clohessy, Shaun Brennan, Matthew Wiley, Tiffany Campo, Cherry Mui, Sam Horowitz, Brandon Ostiguy, Andrew Blancero, Kristen Vacca, Adam Tobey, Eric Krans, Adam Greenwade, Nathaniel opare darko Asare, Alex Muro, Amar Kakirde, Lauren Steinberg, Sean O'Connell, Jennifer Singh, Gabriella CebadaMora, Kenneth Kemp, Erika Jozwiak, Nilus Klingel, Carol Lamb-Lafay, Ted Fisch, Greg Yetman, Lynn Bogan, Gerald Farrell, Alex Pennington, Kisha Santiago, Dean Chuang

Objective: Present BILD v3.5. Review data clarifications, privacy considerations, and example analyses in preparation for final acceptance.

Introduction
- This is the second of three final acceptance meetings, the last of which will be held in about one month.

Data Availability and Clarification
- BILD is a large dataset that typically cannot be opened in Excel. Computers typically need Geographic Information Systems (GIS) capability in order to open and leverage the data.
    - The dataset is only available statewide, not pre-filtered by county/jurisdiction.
- BILD attributes are still being updated.
    - The Office of Real Property Tax Services (ORPTS) property class crosswalk with NYC building codes was successfully integrated; statewide property class analyses will now include NYC buildings.
    - Geometry matches to ORPTS records were improved; more improvements and column streamlining are expected before final acceptance.
    - Critical infrastructure from multiple public Homeland Infrastructure Foundation-Level Data (HIFLD) datasets is almost completely included.
- Final acceptance will include longitudinal datasets from 2018, methodology updates, and data documentation (metadata/data dictionary).
- Missing data should be expected - data coverage by attribute is not expected to be 100% complete.
    - ORPTS data is specific to building type; therefore, certain attributes are only mapped to certain groups of buildings (e.g., residential buildings will have no values in industrial/commercial columns).

- State ownership information is coming from ORPTS data, not Statewide Financial System (SFS) data.

Use cases

- Example analyses are available at https://buildings.mitigateny.org/use_cases.
- Case Study 1: Identify state-owned buildings in Utica by year built and heat type.
- Case Study 2: Devise a heat vulnerability index by calculating the percentage of buildings with central air in residential buildings (suggested by TWG members from the Department of Health)
- Case Study 3: Identify critical infrastructure in a town/jurisdiction.
- One could use critical infrastructure codes or property class codes to identify colleges and universities or other building types (schools, fire or police stations, grocery stores, hospitals, etc.).
  - Multiple buildings on the same campus parcel will not be identical – many columns will be the same, but others will differ (e.g., building valuations, differences from ORPTS matches)
- Causes for missing data:
  - Data does not apply to the building
  - Information isn't available/is missing
  - BILD's data dictionary will clarify this where possible.

Privacy Concerns

- NYS Information Technology Services (ITS) informed AVAIL that all ORPTS data included in BILD is publicly available.
  - Parcel geometries are private in some counties, but these are not included.

Next Steps

- Assess additional datasets for possible inclusion (e.g., historical buildings)
- Quality assurance checks, e.g., name standardization, removing buildings that are no longer standing
- Longitudinal datasets
- Risk file with overlays and centroids for floodplains and National Risk Index (NRI) hazard risk, respectively
  - NRI data will come with a disclaimer on its limitations.
- Data dictionary
- The final acceptance meeting will convene in approximately one month.