Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A household-level parcel dataset of hurricane adaptation and recovery

Yanling Sang
Johns Hopkins University
Ruxiao Chen
Johns Hopkins University
Graham Hults
Johns Hopkins University
Maurizio Porfiri
New York University
Cristina Dragomir
New York University
Yongjun Zhang
Stony Brook University
Luis Ceferino
University of California, Berkeley
Susu Xu
susuxu@jhu.edu

Johns Hopkins University

data-descriptor

Keywords:

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11049701/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. Yongjun Zhang is a member of the Editorial Board of Scientific Data. The other authors declare no competing interests.

A household-level parcel dataset of hurricane adaptation and recovery

Yanling Sang[1], Ruxiao Chen[1], Graham Hults[1], Maurizio Porfiri[2], Cristina Dragomir[3], Yongjun Zhang[4,5], Luis Ceferino[6], and Susu Xu[1,7]

[1]Department of Civil and Systems Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA

[2]Center for Urban Science and Progress, New York University, New York, NY 10012, USA

[3]School of Liberal Studies, New York University, New York, NY 10012, USA

[4]Institute for Advanced Computational Science, Stony Brook University, Stony Brook, NY 11794, USA

[5]Department of Sociology, Stony Brook, NY 11790, USA

[6]Department of Civil and Environmental Engineering, UC Berkeley, Berkeley, CA 94720, USA

[7]Data Science and AI Institute, The Johns Hopkins University, Baltimore, MD 21218, USA

Post-hurricane housing recovery is a longitudinal process through which household repair and sale decisions accumulate to shape community recovery trajectories. Yet these decisions have not been systematically traced at parcel-level, longitudinal, and cross-event scales because doing so requires linking property damage to subsequent repair and sale events across disasters. We present the Hurricane Adaptation and Recovery Parcel (HARP) dataset for Hurricanes Harvey (2017), Ida (2021), and Ian (2022). HARP links geocoded single-family residential parcels to Census Block Groups and integrates parcel attributes, damage assessments, repair records, property transactions, and community-level socioeconomic covariates into monthly recovery sequences. The dataset covers 1,016,927 parcels for Harvey, 121,130 for Ida, and 202,926 for Ian. HARP also supplements official damage records with aerial-imagery-derived assessments and provides buyer-type labels and neighborhood influence networks. This dataset enables large-scale, cross-event, and parcel-level understanding of how individual housing conditions, household conditions, building damage, community resources, and neighborhood influences jointly reshape household-level and community-level housing recovery trajectories, supporting future research and policy designs to enhance coastal resilience.

[1]Corresponding author: Susu Xu, susuxu@jhu.edu

Background & Summary

Hurricanes are among the most severe and costly natural disasters that repeatedly affect populated regions in the United States [1, 2]. They cause widespread damage to housing, infrastructure, and local economies, with effects that persist for months to years following landfall [3, 4]. How affected communities recover from such disruptions has therefore become a central question in disaster research. Post-hurricane community recovery refers to the restoration of the societal functioning of the affected community to its normal or steady-state level [5]. It is a long-term process encompassing multiple interdependent dimensions, including the recovery of housing, the restoration of infrastructure and essential services, the recovery of businesses and the local economy, and the rebuilding of social networks [6–8].

Among these dimensions, housing recovery is widely regarded as a central component of community recovery, because how quickly damaged homes are repaired and how many are ultimately restored shape population retention and return, and subsequently condition household physical and psychological recovery [9–11]. Housing recovery is driven by the adaptation decisions of individual households [12]. After a hurricane, each owner of a damaged home may repair it, sell it or remain inactive [13, 14]. Prolonged inactivity can also end in abandonment, once repair costs exceed what a household can feasibly invest [15]. Repair reflects a willingness to stay and continued investment in the house, while sell transfers ownership and often accompanies relocation [14, 15].

How housing recovery unfolds through housing adaptation decisions is a complex process shaped by the interplay of various factors of household characteristics, building damage, community characteristics, and neighborhood activities [1, 12]. These determinants have been studied across engineering, economics, and the social sciences, and a recent multidisciplinary review synthesizes this dispersed evidence on the factors shaping whether households remain in or leave hazard-exposed areas [16]. At the household level, household characteristics set the baseline for recovery capacity. A household's financial capacity and insurance can affect whether repair is affordable and whether relocation is feasible, and family ties and place attachment can strengthen households' preferences to remain in place [16, 17].

Building damage is another primary driver of housing adaptation decisions, with both damage severity and damage type shaping recovery pathways. Greater damage can increase repair needs and reduce habitability. Following Hurricane Harvey, households exposed to $10,000 of flood damage were 40% more likely to move out within six months and 21% less likely to sell during the same period [18].

Community and market conditions shape housing adaptation decisions by influencing the opportunities and constraints households face. Socioeconomically vulnerable communities often have fewer private and public resources for recovery and experience slower repair and rebuilding [19]. Government buyout programs can further reshape households' recovery options by providing a pathway for relocation from hazard-prone areas [20, 21]. Investors may have more financial resources to restore damaged properties and return them to the market more quickly, potentially accelerating recovery. At the same time, these transactions may shift ownership toward wealthier buyers and contribute to post-disaster gentrification [22, 23].

Neighborhood activity is another important impact factor as housing adaptation decisions are not made in isolation: nearby repair activities and property listings can prompt similar decisions among surrounding households [16, 24, 25]. For example, in the Baltimore metropolitan housing market, a household whose 15 nearest neighbors included one additional home newly listed for sale was 11–15% more likely to list its own home than an otherwise comparable household with no such neighboring listing, an effect that held across alternative definitions of neighborhood size [26]. Similarly, after Hurricane Ike, the correlation between household return and nearby business reopening was $r = 0.189$ within 0.8 km, declined to $r = 0.103$ within 2.4 km, and became insignificant at 3.2 km [27]. As these locally interdependent decisions accumulate over time, they may push communities onto divergent recovery trajectories, ranging from stabilization to prolonged decline [10, 28, 29].

Together, these determinants have direct practical relevance for post-disaster recovery planning. For example, post-disaster recovery planning can combine measures of disaster damage with social vulnerability to identify areas with greater recovery needs and prioritize the allocation of recovery resources [30]. Information on household financial capacity, housing damage, community conditions, and neighborhood activity can further help characterize the constraints and opportunities households face during recovery. For policymakers, understanding where and when households repair or leave can support more targeted and timely recovery interventions and resource allocation [20, 21, 31]. For residents, such information can provide context for housing adaptation decisions as neighborhood recovery conditions evolve [32, 33].

Existing research has identified a wide range of determinants of post-disaster housing recovery. However, these factors have largely been examined in isolation. Empirical studies have quantified how financial resources shape individual recovery outcomes [34], how damage severity conditions housing re-

building [31], how social vulnerability produces uneven recovery across neighborhoods [19], how policy interventions alter household choices [20, 21], and how neighbors' choices spill over onto one another [24, 26, 33]. However, systematic longitudinal evidence remains limited on how these factors interact over time to jointly shape the timing and sequence of household repair and sale decisions, and how these dynamics vary across hurricanes. This limited understanding can hinder effective and targeted post-disaster recovery planning and interventions and, more fundamentally, lead to misleading interpretations of the mechanisms underlying housing recovery. For example, when interdependent neighborhood decisions are not accounted for, variation generated by neighborhood spillovers may instead be attributed to household characteristics, damage conditions, or financial resources, obscuring the mechanisms that actually drive recovery and limiting the generalizability of models to future disaster contexts [33, 35].

A central reason for this research gap is the limitations of existing data sources. One commonly used data source is household surveys, which directly capture residents' reported housing conditions and decisions [14, 36]. However, surveys are typically constrained by small samples and limited spatial coverage, making them less suitable for large-scale, cross-event studies. Mobile location data such as GPS data can serve as a useful proxy because they track people's traces, but their penetration rates are often limited, ranging from a few percent to approximately one-third of the local population [37–39], and observation periods are often limited to the immediate disaster period and short-term aftermath, ranging from about 11 days to 2 months in prior hurricane studies [38, 40, 41], making them less suitable for studying long-term housing recovery. Another approach is to use housing damage states to estimate whether houses are occupied [42, 43], but it is challenging to extend such inference to housing adaptation decisions as occupancy status does not directly correspond to whether it was sold or not. Together, these limitations make it difficult to consistently observe the timing and sequence of post-hurricane housing decisions over the long term. **This reveals the need for longitudinal data that links parcel-level (1) household characteristics, (2) housing damage, (3) repair activities, and (4) property transactions, and community-level (5) socioeconomic attributes and (6) neighborhood interdependencies**.

To enable various comprehensive housing recovery studies at large scale, the construction of a multi-event, parcel-level longitudinal dataset is non-trivial owing to three main challenges. First, cross-source data retrieval and harmonization are difficult because repair, sale, and other recovery records are dis-

tributed across different administrative systems and agencies, with varying access procedures, field definitions, identifiers, and conventions for recording event dates across jurisdictions [44]. For example, property transaction records may document the completion of a sale rather than the earlier listing of the property, while permit systems may record different stages of the repair process. Second, damage observation is fragmented. Official assessments sometimes cover only a subset of affected structures, and assigned damage levels vary with the personnel conducting the assessment [45]. For instance, the FEMA Visual Damage Assessment available for Hurricane Ida leaves a substantial fraction of parcels without an explicit damage label [46]. Damage records must therefore be supplemented and calibrated before they can be compared across events. Third, neighborhood interdependencies are not directly recorded in standard administrative data. Prior work has therefore approached them through behavioral models of interdependent household decisions [47], and in a longitudinal dataset they must instead be inferred from the timing and location of nearby repair and sale events.

To address these challenges, we present the Hurricane Adaptation and Recovery Parcel (HARP) dataset, a multi-event, parcel-level dataset for studying post-hurricane housing recovery dynamics. HARP is built around geocoded single-family residential parcels linked to Census Block Groups (CBGs), enabling analysis at both household and community scales. The dataset covers three major U.S. hurricanes, Harvey (Texas, 2017), Ida (Louisiana, 2021), and Ian (Florida, 2022), selected for their impact scale, data availability, and complementary damage mechanisms (flood-dominant, wind-dominant, and combined wind and surge, respectively), supporting cross-event comparison and assessment of whether observed recovery patterns generalize across distinct disaster contexts. For each event, parcel-level damage, repair activity, and property transactions are harmonized into monthly recovery sequences. HARP supplements damage assessments with aerial-imagery-derived damage labels for Ida and Ian, identifying 9,148 and 11,736 damaged parcels, respectively (Table 6). Beyond integrating heterogeneous administrative and geospatial sources, HARP also includes model-derived neighborhood influence networks inferred from repair and sale event histories using the decision diffusion framework developed in the companion study [48]. In total, HARP covers 1,340,983 geocoded single-family parcels (1,016,927 for Harvey, 121,130 for Ida, and 202,926 for Ian), grouped into 2,796 CBG communities (1,992, 382, and 422 respectively), and organized into monthly longitudinal panels spanning 31, 23, and 34 months (Table 2).

The large scale and fine-grained resolution of HARP support various research and analyses that were

previously difficult to conduct. Because HARP includes every single-family parcel in each study region rather than a sample of households, each household can be observed in full context, with its own conditions and its actual neighborhood, allowing the influencing factors discussed above to be studied jointly rather than in isolation. For example, whether households facing similar damage but different community characteristics repair at different speeds, or whether household decision timing is associated with nearby repair and sale activity. Because the panels extend for months after each landfall, the temporal evolution of recovery, such as early repairs, delayed sales, and prolonged inactivity can be traced, rather than observed only at a single point in time. Moreover, since attributes like damage, repair, and sale records are harmonized into a common format across the three storms, patterns estimated on one event can be tested on the others, such as whether the community characteristics associated with faster recovery in one storm show similar associations in the others. For Harvey, derived buyer type labels further distinguish whether a sold home was purchased by an individual or an investor, so ownership change can be examined alongside recovery. For example, the data can be used to assess whether damaged homes are disproportionately purchased by investors and which communities experience higher concentrations of investor purchases. Beyond these analyses, the released event sequences and influence networks allow recovery to be studied as a spatially interdependent diffusion process and enable examination of how decision dependencies emerge and propagate over time [49–51]. Taken together, HARP offers a harmonized, longitudinal, parcel-level dataset across three major hurricanes, providing a reproducible foundation for comparative studies of post-hurricane recovery.

These capabilities make HARP relevant to a broad range of coastal resilience researchers and practitioners working on post-hurricane housing and community recovery and resilience. In particular, the fine-grained spatio-temporal resolution (parcel-level spatial resolution and monthly temporal resolution), large spatial and temporal scales (more than one million parcels and 30-month time duration along gulf coast), and well-aligned detailed information (ranging from socio-demographic background and housing damage to sale and repair records) make it possible for researchers in civil engineering, sociology, urban planning, and economics to conduct longitudinal cross-event studies and uncover the highly complex dynamics of housing and community recovery trajectories. Beyond academic research, potential users also include federal and state agencies, emergency management organizations, and housing and planning departments, which may use the dataset to better understand recovery patterns and support post-disaster

policy design, planning, and resource allocation.

The remaining part of this paper is organized as follows. The Methods section details the construction pipeline of HARP, followed by the Data Record section describing the contents of the released datasets. The Technical Validation section evaluates data consistency and integrity, and the Example Analysis section illustrates representative research applications enabled by HARP. The Data Record section gives the public download location of the released dataset, and the Code Availability section provides links to the code used to construct it. The Usage Notes document the storm-specific variable definitions.

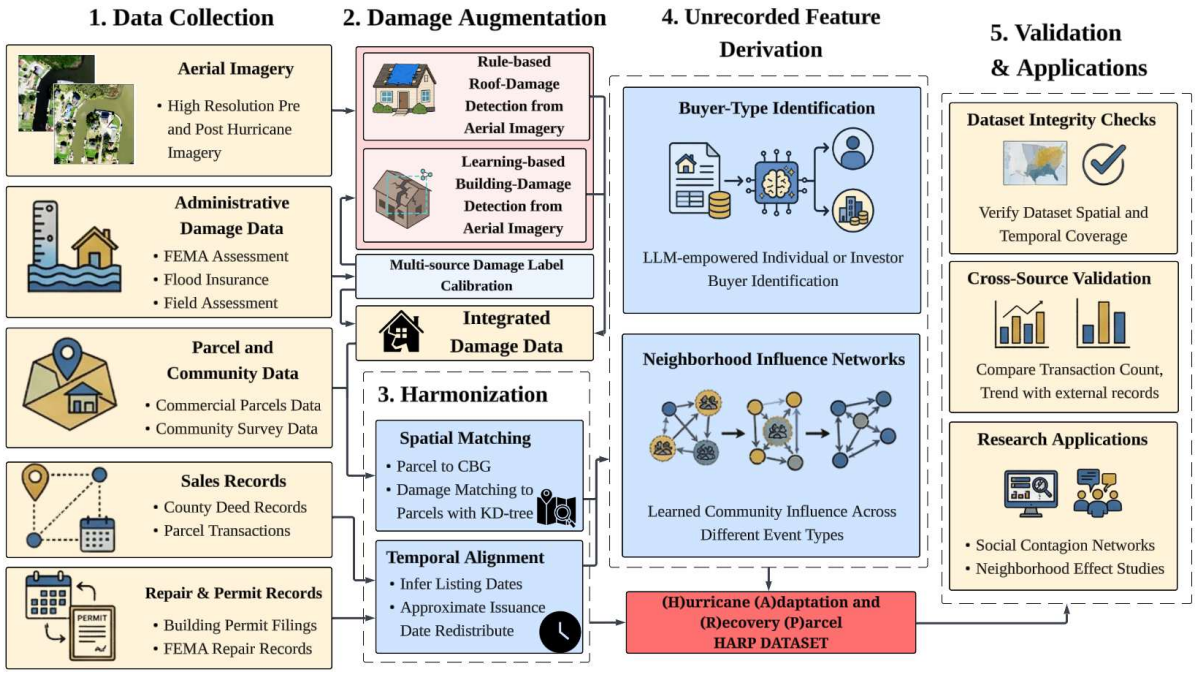


Figure 1: Construction pipeline of the HARP dataset, comprising five stages: *data collection* from administrative, geospatial, and Census sources; *damage augmentation*, which supplements official assessments with aerial-imagery-derived damage labels; *data harmonization*, which aligns event timing, damage severity, and spatial units across sources into monthly parcel-level panels; *derived feature construction*, including buyer type and neighborhood influence networks; and *validation and applications*, presented in the Technical Validation and Example Analysis sections.

Methods

HARP is constructed through a multi-stage pipeline that transforms heterogeneous source records into harmonized longitudinal parcel-level data. As illustrated in Fig. 1, the pipeline proceeds in five stages: *data collection*, *damage augmentation*, *data harmonization*, *derived feature construction*, and *validation and applications*. The first four stages are detailed in the subsections below, and the fifth is presented in the Technical Validation and Example Analysis sections.

Data collection

No single archive records post-hurricane household decisions. HARP's raw material must therefore be assembled piece by piece from whichever authority maintains each one. Property characteristics and transactions come from county appraisal districts and commercial parcel providers. Repair activity comes from municipal permit systems and federal assistance programs. Damage assessments come from FEMA and academic field-survey teams, post-event observations from aerial imagery programs, and socioeconomic context from the Census Bureau. Collection then proceeds in two steps: we first select the hurricanes to cover, then fix the parcel universe of each event and retrieve the records that populate it.

Hurricane selection

The empirical basis of HARP comprises three major U.S. hurricanes, Harvey (2017), Ida (2021), and Ian (2022), selected for their impact scale, data coverage, and complementarity in geographic setting and primary damage mechanism. All three storms rank among the costliest hurricanes in U.S. history, with combined economic losses exceeding $300 billion. Hurricane Harvey caused an estimated $125 billion in total damage, the second-costliest on record in the United States. Its catastrophic rainfall-induced flooding inundated approximately 70% of Harris County, Texas, and flooded over 136,000 structures [52, 53]. Hurricane Ida made landfall as a Category 4 storm near Port Fourchon, Louisiana in August 2021 with estimated damages exceeding $75 billion [54, 55]. Hurricane Ian struck Lee County, Florida in September 2022 with sustained winds of approximately 150 mph and a catastrophic storm surge, causing over $112 billion in damage, which is the costliest hurricane in Florida history and the third-costliest in U.S. history [56, 57].

Beyond their scale and data availability, the selected hurricanes differ substantially in primary damage mechanism, a dimension that is likely to condition recovery trajectories. Harvey was overwhelmingly flood-driven, with insured losses predominantly attributed to National Flood Insurance Program (NFIP) flood claims rather than wind [58]. Ida was wind-dominant with surge flooding limited to unprotected areas [54, 59]. Ian combined wind and storm surge, producing both widespread roofing losses and surge-driven structural damage [60]. Together, the three events span a spectrum from flood-dominant to wind-dominant to combined hazard, across distinct geographic and regulatory contexts (Harris County, New Orleans, and Lee County), providing complementary settings for analyzing how damage type shapes housing adaptation decisions.

Parcel universe and raw source integration

For each hurricane, we defined a parcel universe of geocoded single-family residential parcels. Parcel-level characteristics and transaction records were drawn from Harris Central Appraisal District (HCAD) public data for Harvey [61] and from Regrid administrative data for Ian and Ida [62]. Community-level socioeconomic attributes were obtained from the American Community Survey (ACS) at the Census Block Group (CBG) level [63], with CBG boundaries obtained from the U.S. Census Bureau TIGER/Line products [64]. Damage assessments were sourced from FEMA, Lee County, and StEER for Ian [65–67], from FEMA Visual Damage Assessment products for Ida [46], and from FEMA HydroShare and DesignSafe for Harvey [68, 69].

Property transactions and repair records were processed separately before being merged into the final panel. Sales were extracted from HCAD deed records for Harvey [61] and from Regrid parcel transaction files for Ian and Ida [62]. Repair events were extracted from building permit filings: HCAD permit records for Harvey [61], Lee County permits issued via the Accela platform for Ian [70, 71], and the City of New Orleans permit system for Ida [72]. As the City of New Orleans waived roofing permit requirements following Ida [72], repair activity for Ida was supplemented with Individuals and Households Program (IHP) habitability repair records [73]. After source-specific cleaning, sale and repair records were linked to parcels and treated as observable proxies for household housing decisions.

Damage augmentation

The FEMA Visual Damage Assessment record available for Hurricane Ida covers only a subset of structures within the impacted region [46], leaving a substantial fraction of parcels without an explicit damage label. We supplemented these records with two complementary detectors applied to pre and post Ida National Agriculture Imagery Program (NAIP) imagery [74], a fine-tuned existing deep change-detection model that captures visible structural damage, and a color-based detector targeting the blue tarps that homeowners install on damaged roofs after a hurricane. The two detectors operate on different backbones and together form the aerial-imagery-derived damage layer that we release as a new column alongside FEMA's record.

Because the aerial-imagery-derived damage layer is interpreted alongside FEMA's damage labels, we first clarify how those labels are assigned and how they relate to damage visible from overhead imagery. The FEMA Visual Damage Assessment is itself an imagery-based product: assessors compare pre- and post-event imagery structure by structure, and the assessment protocol instructs them to assess only what is visible, to assign the lower category when uncertain, and to record a structure whose view is obstructed as *Affected* [75]. For hurricane events the five levels are defined by the visible condition of the building. *No Visible Damage* requires no visible damage, no standing water, and no blue tarps. *Affected* is superficial or cosmetic damage, with shingles, siding, or gutters missing but no holes in the roof. *Minor* is non-structural damage that leaves the residence temporarily inhabitable and repairable, with a quarter to a half of the roof missing and the framing undamaged. *Major* is unsafe damage requiring extensive repairs, with partial or complete failure of roof, walls, or foundation, up to three quarters of the roof covering missing, and the interior exposed. *Destroyed* is two or more structural failures, beyond repair.

The detectors are trained as a binary problem, so these five levels are collapsed to two. Footprints labeled *Minor*, *Major*, or *Destroyed* form the positive class, and those labeled *No Visible Damage* form the negative class. *Affected* footprints are excluded from training altogether.

The cut is placed where damage becomes visible from overhead rather than where it becomes structurally severe, and FEMA's own criteria separate the two classes at that point: *Affected* stipulates no holes in the roof, whereas *Minor* begins at a quarter of the roof missing, a scale of roof damage that is generally visible at 0.4 m ground sampling distance. Excluding *Affected*, whose defining evidence is cosmetic and whose protocol also absorbs structures too obstructed to judge, keeps the target restricted

to damage the imagery can carry. The same boundary matches the backbone's own taxonomy, whose mildest damage class is *partial roof damage*, defined as up to half the roof covering removed [76], so FEMA *Minor* corresponds to a category DamageCAT was pretrained to recognize. *Minor* also carries most of the available supervision: among the damaged footprints on the 122 training tiles, 287 are *Minor* against 38 *Major* and 6 *Destroyed*, so excluding it would remove 87% of the positive examples and leave too few to fine-tune on. A manual review of the *Minor* footprints on those tiles found most of them showing blue tarps or exposed roof decking in the post-event imagery.

Model	Year	Precision	Recall	F1
DamageCAT [76]	2025	**91.38%**	**87.19%**	**89.24%**
ChangeOS [77]	2021	31.68%	9.27%	14.35%
Microsoft Siamese [78]	2021	2.89%	34.09%	5.33%
ChangeFormer [79]	2022	6.91%	0.50%	0.92%

Table 1: Binary any-damage detection on the DamageCAT (Hurricane Ida) test set. Precision, recall, and F1 are computed pixel-wise against the binarized DamageCAT ground truth.

Fine-tuned DamageCAT detector. Building damage extraction from pre and post event aerial imagery has been actively studied in recent years, with several deep change-detection models published as open-source baselines. To select a backbone for our pipeline, we benchmarked four such publicly available models on the DamageCAT test set, which is built on Ida imagery [76] (Table 1), and DamageCAT substantially outperformed the alternatives.

We adopted DamageCAT as the backbone. Because the released weights produced inconsistent predictions on NAIP coverage outside the original training footprint, we fine-tuned on a region-specific training set derived from FEMA damage points. Each FEMA damage point was matched to the nearest OpenStreetMap (OSM) building footprint within a 5 m radius, and the resulting footprint pixels were rasterized into per-tile ground truth masks.

The severe class imbalance in the underlying FEMA records shapes the rest of the training setup. Because *Minor* dominates the positive class and its damage is the mildest that the imagery resolves, recall is the quantity this design trades away; the operating point reported below reflects that.

Input imagery was sourced from NAIP at 0.4 m ground sampling distance (pre-Ida: 2019-07-07; post-Ida: 2021-11-16 to 2022-01-30) [74]. The network was initialized from the released DamageCAT checkpoint and trained with the original Dice-plus-focal loss. Because Major and Destroyed footprints

are far rarer than the milder grades, tiles containing them were oversampled during training. Drawing tiles uniformly would have exposed the model to severe damage too seldom to learn it, leaving it able to minimize its loss by predicting "no damage" almost everywhere.

At inference, geometric test-time augmentation was applied and the resulting softmax outputs were averaged. A parcel was then classified as damaged when the fraction of positively predicted pixels within its OSM footprint exceeded a fixed cutoff of 1%. Because these labels supplement the FEMA record rather than replace it, the cutoff was set to balance precision and recall, at this operating point, the detector attains 83% precision and 22.0% recall. Applied across the Ida study region, the detector flags 2,824 parcels as damaged (Fig. 2 (a1)–(c2)).

Blue-tarp detector. After Hurricane Ida, blue plastic tarps were widely installed as temporary covers on damaged roofs to keep homes habitable while permanent repairs were arranged. This practice is institutionalized in the U.S. Army Corps of Engineers' *Operation Blue Roof* program, which alone installed over 30,000 such covers following Hurricane Ida [80]. These tarps are visually distinctive and serve as a direct proxy for post-disaster damage and repair activity. We exploit this signal with a hand-crafted pixel-level rule combining three conditions. First, the pixel must appear noticeably more blue than the other channels in the post image. Second, the same location must have become bluer after Ida, a constraint that rules out swimming pools and naturally blue surfaces present in both pre- and post-Ida imagery. Third, an absolute-brightness threshold excludes the dark tree shadows that otherwise dominated false positives in an unconstrained variant. A pixel is flagged as a tarp candidate only when all three conditions hold. An OSM footprint is then flagged as damaged when the proportion of flagged pixels within it exceeds a cutoff. Thresholds were tuned by grid search against a manually annotated set of 200/7,253 NAIP tiles selected by FEMA-damaged point count per tile, in which every overlapping OSM building footprint was visually inspected and labeled as displaying a blue tarp or not. Validation on this annotated set yields $F1 = 88.4\%$ ($P = 86.9\%$, $R = 90.0\%$) at the chosen operating point. Fig. 2 (d1)–(f2) shows two illustrative tiles, each presented as a triplet of pre-Ida image, post-Ida image, and post-Ida image overlaid with damaged footprints (green fill): the top row (d1–f1) is a representative tile from the annotated set with manual ground-truth labels, and the bottom row (d2–f2) is a tile outside the annotation set with detector-flagged predictions. The predicted footprints align with the visible blue tarps on unseen imagery, confirming that the rule generalizes beyond the tuning set. Applied across the

Ida study region, the detector flags 7,072 parcels as damaged.

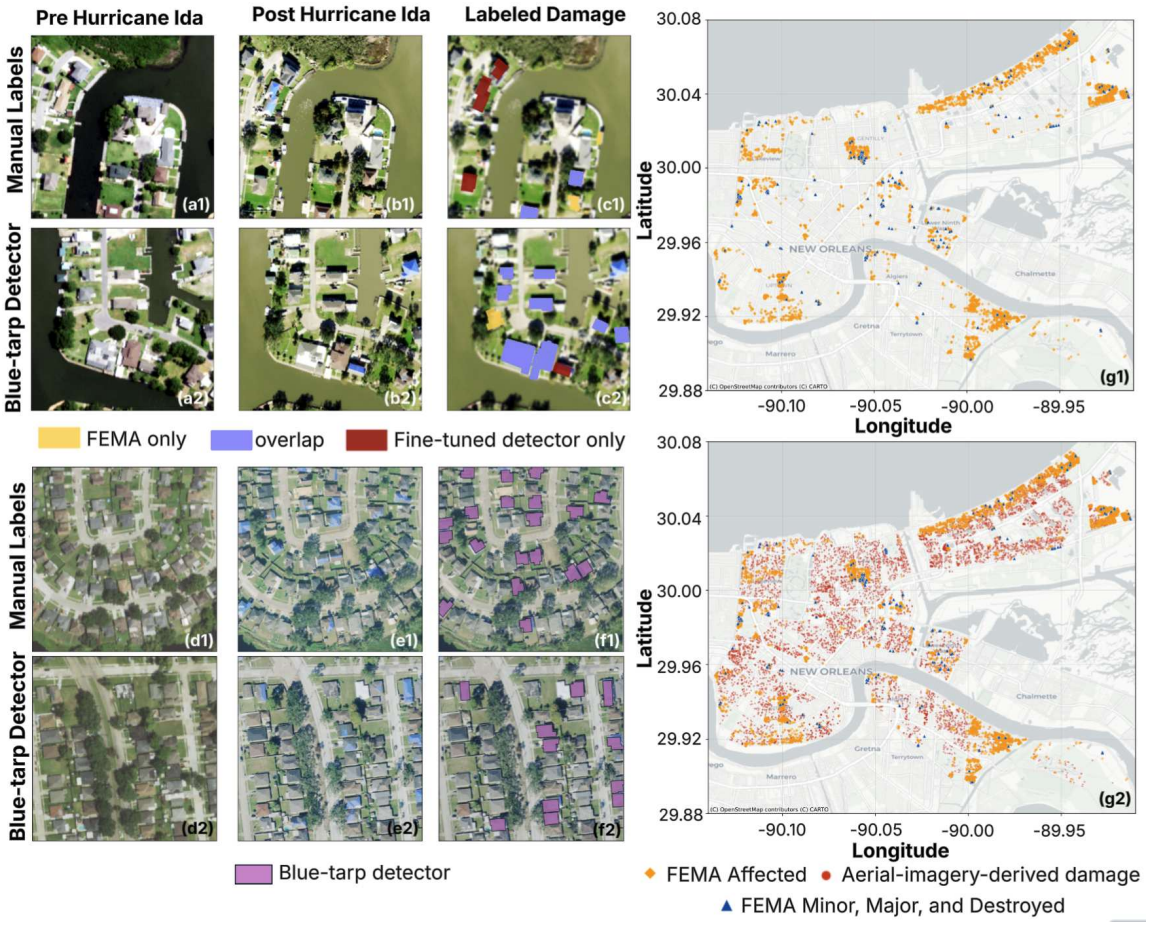


Figure 2: Aerial-imagery-derived damage layer for Hurricane Ida. **(a1)**–**(c1)** and **(a2)**–**(c2)**: two representative NAIP tiles for the fine-tuned DamageCAT detector, each shown as pre-Ida image, post-Ida image, and damage-label overlay on the post-Ida image (FEMA only in gold, overlap between FEMA and the fine-tuned detector in purple, fine-tuned detector only in red). **(d1)**–**(f1)** and **(d2)**–**(f2)**: two representative NAIP tiles for the blue-tarp detector, each shown as pre-Ida image, post-Ida image, and post-Ida image overlaid with damaged footprints (purple fill); the top row (d1–f1) is from the manually annotated set with ground-truth labels, and the bottom row (d2–f2) is a tile outside the annotation set with detector predictions. **(g1)**: spatial distribution of FEMA-labeled damaged parcels across the Hurricane Ida study region (Affected as orange diamonds; Minor, Major, and Destroyed as blue triangles). **(g2)**: the same region after augmentation, adding the aerial-imagery-derived damage layer released as `ida_AerialDamaged` (red circles) on top of the two FEMA layers. The legend below (g2) applies to both (g1) and (g2).

In total, these two detectors capture complementary signals and their union (9,148 parcels) is released as a new Aerial Damaged column alongside FEMA's record, an approximately 14× expansion over the FEMA Minor, Major, and Destroyed labels (Fig. 2 (g1)–(g2)).

Extension to Hurricane Ian. We applied the same blue-tarp detector to Hurricane Ian's NAIP coverage of Lee County, FL (pre-Ian: 2019-11-21 to 2022-01-30; post-Ian: 2023-01-16 to 2023-02-14) [74], with thresholds re-tuned against a manually annotated set of 57 tiles. Validation yields $F1 = 82.6\%$ ($P = 84.7\%$, $R = 80.6\%$). The detector flags 11,736 parcels as damaged, of which 8,976 (76.5%) lie outside FEMA's 22,274 Affected, Minor, Major, or Destroyed parcels and is released as the `ian_AerialDamaged` column alongside FEMA's record.

Data harmonization

The records assembled above are not directly comparable. Each agency collected its data for its own administrative purpose. The sources therefore disagree about when an event occurred, what a severity label means, and how a record maps to a place on the ground. Such discrepancies would otherwise be mistaken for real differences in recovery behavior. Harmonization resolves them in three steps. *Temporal alignment* converts each system's recorded timestamp into a proxy for the household decision behind it and bins every event to monthly resolution. *Damage calibration* maps the severity scales of the three events onto one ordinal taxonomy. *Spatial matching* uses geography to link records that share no identifier. It joins damage assessments and repair events to parcels, and each parcel to its Census Block Group.

Temporal alignment

Sale, repair, and assistance records originate from administrative systems that timestamp different stages of an event lifecycle, including deed closing dates, permit application or issuance dates, and FEMA registration dates. The recorded dates may occur days or even months after the household decision, with delays varying across data sources. We therefore provide both the transaction date and inferred listing date for sale event. For sale events, closing dates were converted to estimated listing dates, which serve as proxies for the household's decision to sell. For repair events, both permit record-opening (filing date) or issuance dates were provided as proxies for the repair decision. Placeholder and missing repair dates

were then reassigned based on estimated delay distributions. Following these adjustments, all resulting timestamps were binned to monthly resolution.

For sale events, each closing date $\boldsymbol{d_C}$ was converted to an estimated listing date $\boldsymbol{d_L}$ via $\boldsymbol{d_L} = \boldsymbol{d_C} - \boldsymbol{DOM}^{\boldsymbol{med}}_{\boldsymbol{z,m}}$, where $\boldsymbol{DOM}^{\boldsymbol{med}}_{\boldsymbol{z,m}}$ is the Redfin median days-on-market for the matching zip code and month. When zip-month information was unavailable, a fallback sequence of county-month, then county-year, and finally a 30-day default was used. For Ian and Ida, the resulting monthly listing counts were additionally calibrated against Redfin county-level new-listing totals: months exceeding the reference by more than 15% triggered a backward shift of 20–60 days applied to the excess listings, repeated iteratively for up to three passes. Harvey was not included because HCAD deed records cover a broader set of transactions than Redfin and are not directly comparable.

Repair events were timestamped using the permit record-opening date as a proxy for the household repair decision. When only the issuance date was available, the issuance date was used as a conservative approximation. Two source-specific adjustments were then applied. For Harvey, approximately 126,500 HCAD disaster-review records share a single placeholder issue date. To prevent these placeholders from artificially concentrating repair activity on a single day, their dates were redistributed by sampling from the empirical HCAD delay distribution stratified by FEMA damage level and blended 70/30 with the FEMA IHP rental-assistance end-date distribution. Sampling was deterministic under a fixed random seed. For Ida, FEMA IHP habitability repair records, used to supplement permit-based repair activity, lack precise completion dates. Repair dates were therefore sampled from the empirical distribution of the IHP rental-assistance end date, scaled by a uniformly sampled start fraction in $[0.15, 0.55]$, and clipped to $[7, 540]$ days post-landfall. All repair timestamps were then binned to monthly resolution; any timestamp falling beyond a storm's released panel is dropped at this binning step rather than clamped to the final month, which for Ida affects under 0.5% of IHP-supplemented records (all in the single month immediately following the panel).

Damage calibration

Damage assessments across the three events were produced by different agencies using incompatible severity classifications. We therefore harmonized these records into a common damage-level taxonomy. For Ian, records from FEMA, Lee County Emergency Management, and StEER FAST were merged [65–

67]. For Ida, the FEMA Visual Damage Assessment (VDA) product provides five damage classes [46]. Parcels with blank damage labels were reclassified as "Assessed", and parcels missing from the FEMA VDA but with a matching FEMA NFIP flood insurance claim [81] were labeled as flood-claim damage. For Harvey, damage levels were inferred from modeled flood inundation depth using fixed thresholds: Affected (0–2 ft), Minor (2–5 ft), Major (5–8 ft), and Destroyed (> 8 ft) [68, 69].

Flood depth was drawn from FEMA assessment fields, with sparse Ida values filled from a National Hurricane Center (NHC) storm surge raster. The limited FEMA VDA spatial coverage for Ida further motivated an aerial-imagery-derived supplemental damage layer, described in the Damage augmentation subsection above.

Three encoding details vary across storms and carry implications for downstream use. (i) The damage level is released as a categorical code rather than a severity ordinal, and its label set differs by storm. Harvey has five labels: *Affected*, *Minor*, *Major*, *Destroyed*, and *UNK* for parcels the assessment did not reach. Ian adds *Inaccessible*, for structures whose damage could not be visually verified. Ida has nine labels, adding *Assessed*, *No Visible Damage*, flood-insurance claim damage, and *Not Building/Remove*. The code-to-label mapping for each storm is given in the dataset documentation. (ii) The "truly damaged" set used for repair–damage matching excludes "Affected" for Harvey but includes it for Ian and Ida. (iii) The estimated-loss field carries a dollar loss estimate (in USD) for Ian (drawn from the FEMA assessment) but an ordinal severity score for Harvey and Ida, reflecting source-specific availability of monetary estimates.

Spatial matching

Parcel records and damage assessments originate from separate geospatial systems with no shared identifiers, requiring spatial join procedures to link them. Each parcel centroid was intersected with Census block-group boundaries to assign a CBG identifier, supporting both parcel-level analysis and aggregation into community-level recovery sequences [64]. Damage records, provided as geocoded point or polygon features, were then matched to parcels via a KDTree nearest-neighbor search within a maximum radius (0.01°, $\approx$1.1 km, for Ian; 0.02°, $\approx$2.2 km, for Ida and Harvey); when multiple records fell within the radius of the same parcel, the highest-severity assessment was retained. The assigned CBG identifiers were then used to attach CBG-level socioeconomic covariates from the American Community Survey to

each parcel [63].

For Harvey and Ian, an additional spatial propagation step addressed residual geocoding mismatches between FEMA damage point coordinates and permit-based repair addresses. In a forward, event-driven pass, each repair event was additionally duplicated to every damaged parcel within 200 m (Harvey) or 500 m (Ian) of the event's coordinates. In a reverse, parcel-driven pass, every damaged parcel still without a recorded repair received a copy of the repair history from the nearest repaired parcel within 200 m (both storms). Both passes preceded the monthly binning of repair timestamps.

Ida was excluded from this spatial propagation step because its permit coverage was too sparse. Instead, FEMA IHP habitability records, which supplement Ida's permit-based repair activity but lack parcel-level coordinates, were assigned to damaged parcels within the same zip code using a deterministic severity-based round-robin: parcels were processed in severity order (Destroyed, then Major, Minor, Affected, Assessed), and the chronologically earliest IHP records were routed to the highest-severity parcels.

Derived feature construction

Two quantities that HARP's intended analyses depend on appear in no source record. Deed files name the buyer of a property but never say whether that buyer is a household or an investor. Nor does any register document which parcels influence which others when owners decide to repair or sell. Both must be inferred from what the sources do record. *Buyer type* is recovered from free-text owner names by a two-level classifier that pairs keyword rules with a large language model. *Neighborhood influence networks* are estimated from the timing and location of the repair and sale events themselves, using the decision diffusion framework of the companion method paper [48]. Both are model outputs rather than observations, so we release them as clearly marked derived attributes that users may replace with their own estimates.

LLM-based buyer type identification

For analyses that distinguish the post-disaster behavior of owner-occupants and investors, the Harvey dataset was enriched with buyer type labels derived from administrative owner names. Sale parcels were linked to HCAD owner records by account number, and the primary owner name for each matched

account was assigned one of three labels: "individual", "investor", or "unknown".

Owner name classification from free-text administrative records is inherently ambiguous: names range from clearly personal (e.g., "JOHN A. SMITH") to clearly organizational (e.g., "ABC REALTY LLC"), but a large fraction occupy an intermediate category, including informal business names, family partnerships, and trust-related designations that cannot be reliably identified by keyword matching. Large language models are well suited to this task because they encode the variable surface forms of personal and organizational names [82]. Recent work has demonstrated that LLM-based classification of open-text categorical fields achieves strong accuracy on domain-specific labeling tasks where rule-based approaches are brittle [83], which motivates their deployment as a second-level classifier for residual ambiguous cases.

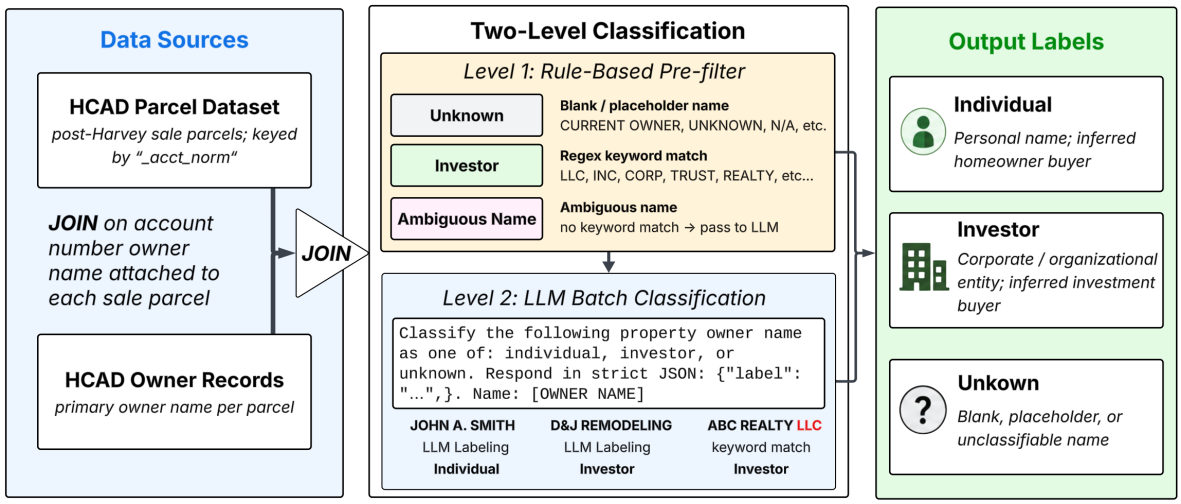


Figure 3: Two-level buyer type classification pipeline applied to Harvey post-hurricane sale parcels. The pipeline assigns each sale parcel one of three labels: Individual, Investor, or Unknown.

As illustrated in Fig. 3, classification proceeded in two levels. A rule-based pre-filter first resolved clear cases: blank or placeholder names (e.g., "CURRENT OWNER", "N/A") were labeled "unknown", and names matching organizational keywords via regular expression, including tokens such as LLC, INC, CORP, and REALTY, were labeled "investor". Remaining names with no keyword match were submitted in batches to an LLM (GPT-4o) with a prompt restricted to the three allowed labels. These labels support the released example analyses of investor versus individual buyer behavior, including repair timing relative to sale and damage-stratified buyer composition.

Latent neighborhood influence network

For each CBG community, a latent neighborhood influence network was inferred from the calibrated monthly parcel panel using the Household-level Decision Diffusion (HDD) framework developed in the companion method paper [48]. Briefly, HDD decomposes each household's decision propensity into three additive components: a damage-driven covariate baseline, a multi-scale self-excitation term encoding the household's own event history, and a neighborhood excitation signal propagated over a jointly learned, community-specific directed graph. Separate graph structures are inferred for sale and repair decisions. Each network is provided as a pair of directed, weighted adjacency matrices $\mathbf{A}^{(g,k)}$, where g indexes the CBG and $k \in \{\text{sale, repair}\}$ denotes the event type, with edge weights encoding the estimated pairwise influence strength between residential parcels within a community.

Data Record

The complete HARP dataset is publicly available for direct download, with no registration or access request required, at `https://github.com/SusuXu-s-Lab/HARP-Dataset`, under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

HARP is released as a collection of monthly parcel-level panels, one per hurricane event. Table 2 summarizes the three hurricanes covered. The first panel reports the temporal extent, number of Census Block Group communities, and total residential parcels. Full descriptions of all parcel-level attributes are provided in the Usage Notes section (Tables 4–13). For Harvey, each community record additionally carries a per-parcel `buyer_type` label (`individual`, `investor`, or `unknown`), derived from HCAD owner names as described in the Methods. The label is empty where no classification is available, which includes every parcel without a sale in the observation window.

Table 2: Per-hurricane dataset overview.

Hurricane	Landfall date	Temporal coverage	Months	Communities	Parcels
Ian	09/28/2022	01/01/2021–10/01/2023	34	422	202,926
Ida	08/29/2021	09/01/2020–07/01/2022	23	382	121,130
Harvey	08/25/2017	05/01/2016–11/01/2018	31	1,992	1,016,927

Hurricane	Permit fee waiver	IHP deadline	EIDL deadline
Ian	10/05/2022	01/12/2023	06/29/2023
Ida	10/07/2021	11/29/2021	06/07/2022
Harvey	—	11/30/2017	05/25/2018

To support analyses that anchor recovery behavior to policy timelines, for example, comparing decisions before and after a deadline or within the window between two milestones, the second panel of Table 2 reports three event-specific administrative milestones per hurricane: the local permit-fee-waiver date [84, 85], the FEMA Individuals and Households Program (IHP) application deadline [86–89], and the U.S. Small Business Administration Economic Injury Disaster Loan (EIDL) application deadline [90–92].

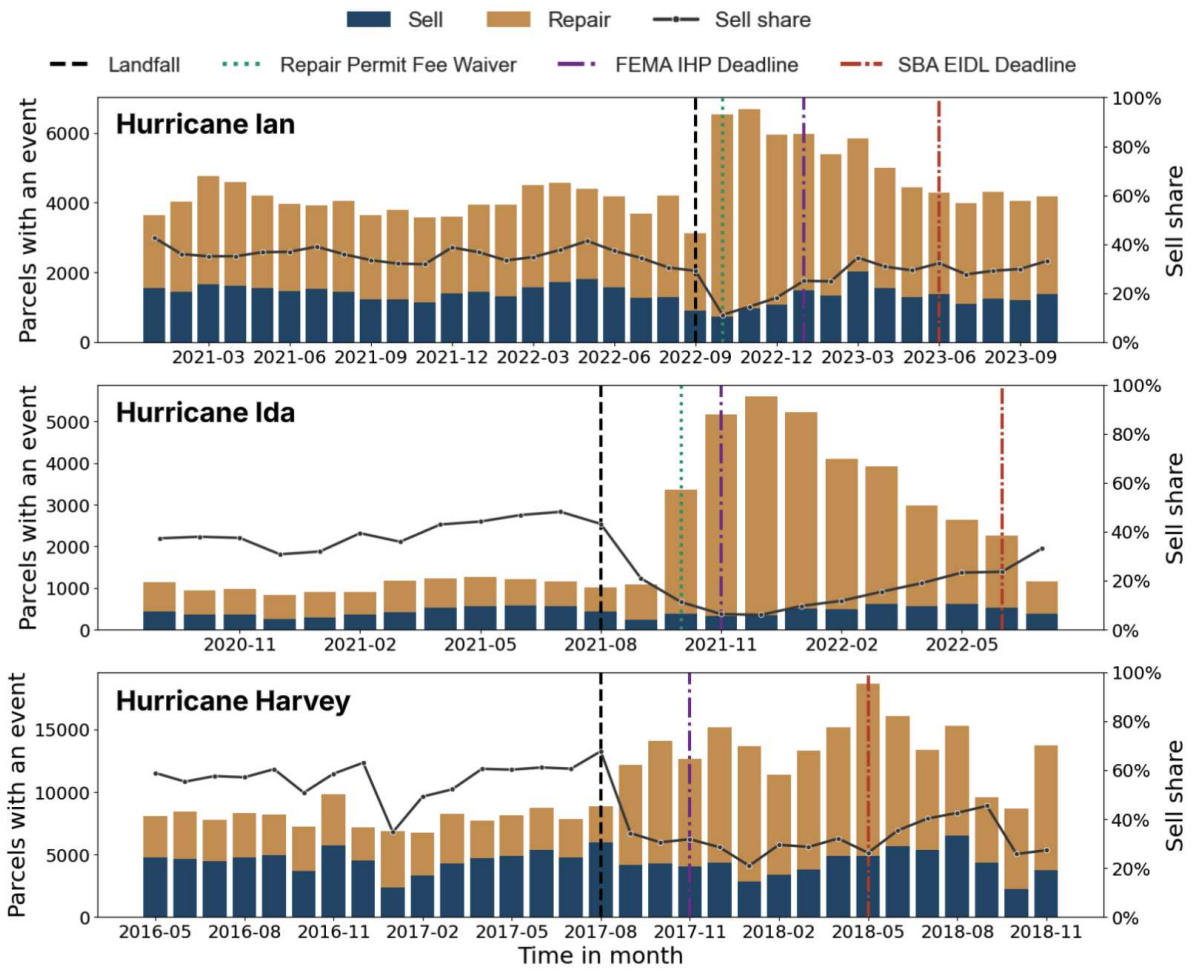


Figure 4: Monthly housing recovery decision records across three hurricane events. Bars give the monthly number of parcels with a recorded sale or repair, stacked with sales below and repairs above, for Hurricane Ian (Lee County, Florida, 2022), Hurricane Ida (New Orleans, Louisiana, 2021), and Hurricane Harvey (Harris County, Texas, 2017). A parcel is counted once per month per event type, so a month with several repairs on one parcel contributes one. The line shows the monthly share of sales among those parcels, read against the right-hand axis. Each panel spans that storm's full released window, so the three differ in length. Vertical lines mark hurricane landfall and the storm-specific administrative milestones listed in Table 2.

Figure 4 presents the monthly number of parcels with a sale or repair across all three hurricanes, over each storm's released window. Repair activity rises sharply after landfall across all three events.

Figure 5 provides a complementary spatial overview of cumulative sale and repair activity at the CBG level, with magnified insets of the inferred parcel-level influence networks for representative communities.

460 Sale and repair activity is geographically concentrated near the primary damage footprint of each storm.

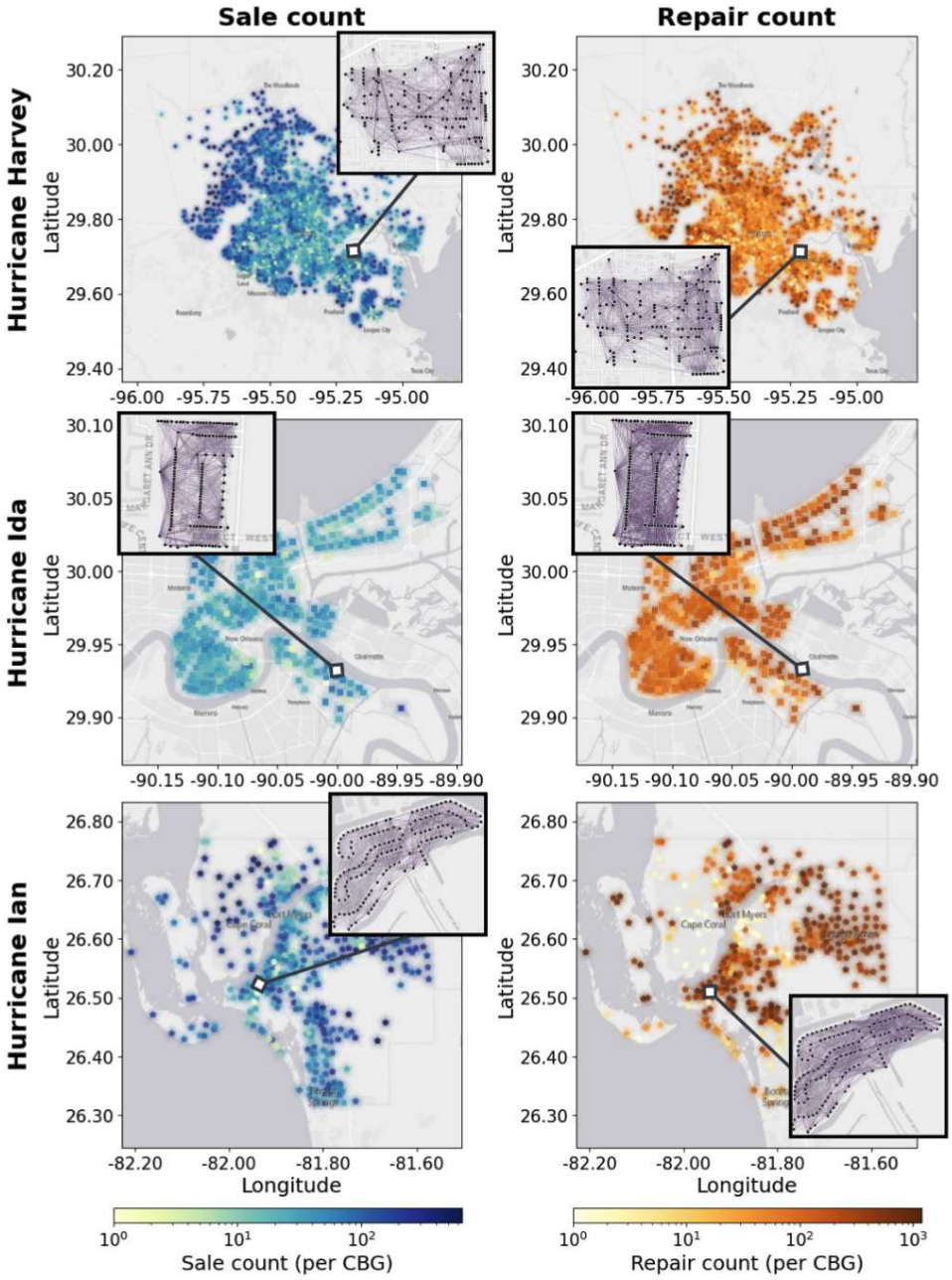


Figure 5: Spatial distribution of sale and repair events and post-hurricane learned neighborhood influence networks for three hurricanes. Each dot indicates a Census Block Group (CBG), colored by the cumulative number of parcels with a sale (left column, blue) or a repair (right column, orange-brown) over the storm's released window; one shared color scale is used per column. Insets show the learned neighborhood influence network for a representative CBG in each panel: vertices represent parcels, and edges indicate non-zero inferred influence weights with opacity reflecting strength. Left and right insets correspond to sale- and repair-type edges, respectively.

Technical Validation

HARP combines records that were never meant to be used together. The released panels therefore need evidence that the integration preserved the behavior they describe. Validation runs along two axes. The external axis compares HARP's aggregate sale and damage counts against independent sources. No parcel-level ground truth exists for post-hurricane repair and sale behavior, so aggregates are the only external check available. The internal axis confirms that the released files are structurally sound. The three subsections below report sale-record consistency, damage-count consistency, and data integrity.

Sale-record cross-source consistency

Monthly sale events are validated against two independent commercial real estate aggregators (Fig. 6): Redfin, restricted to single-family residential transactions to match HARP's scope [93], and Realtor.com, which aggregates single-family, condominium, and townhouse sales [94]. The two references differ from HARP in unit of observation as well as in scope: HARP counts geocoded single-family parcels, whereas the Realtor.com aggregate counts listed condominium and townhouse *units*, several of which may sit on a single parcel. Raw monthly volumes are therefore not directly comparable in level, and the Realtor.com series sits systematically above HARP in panels (a1), (b1), and (c1) for this reason rather than because of missing records. Validation accordingly reports both raw monthly sale volumes and month-to-month percentage changes; the latter cancels the fixed level offset induced by these unit and scope differences and serves as the primary basis for assessing temporal agreement. Across all three events, HARP sale sequences are consistent with both reference sources in capturing post-landfall transaction declines and subsequent market rebounds, supporting the temporal fidelity of the integrated sale records.

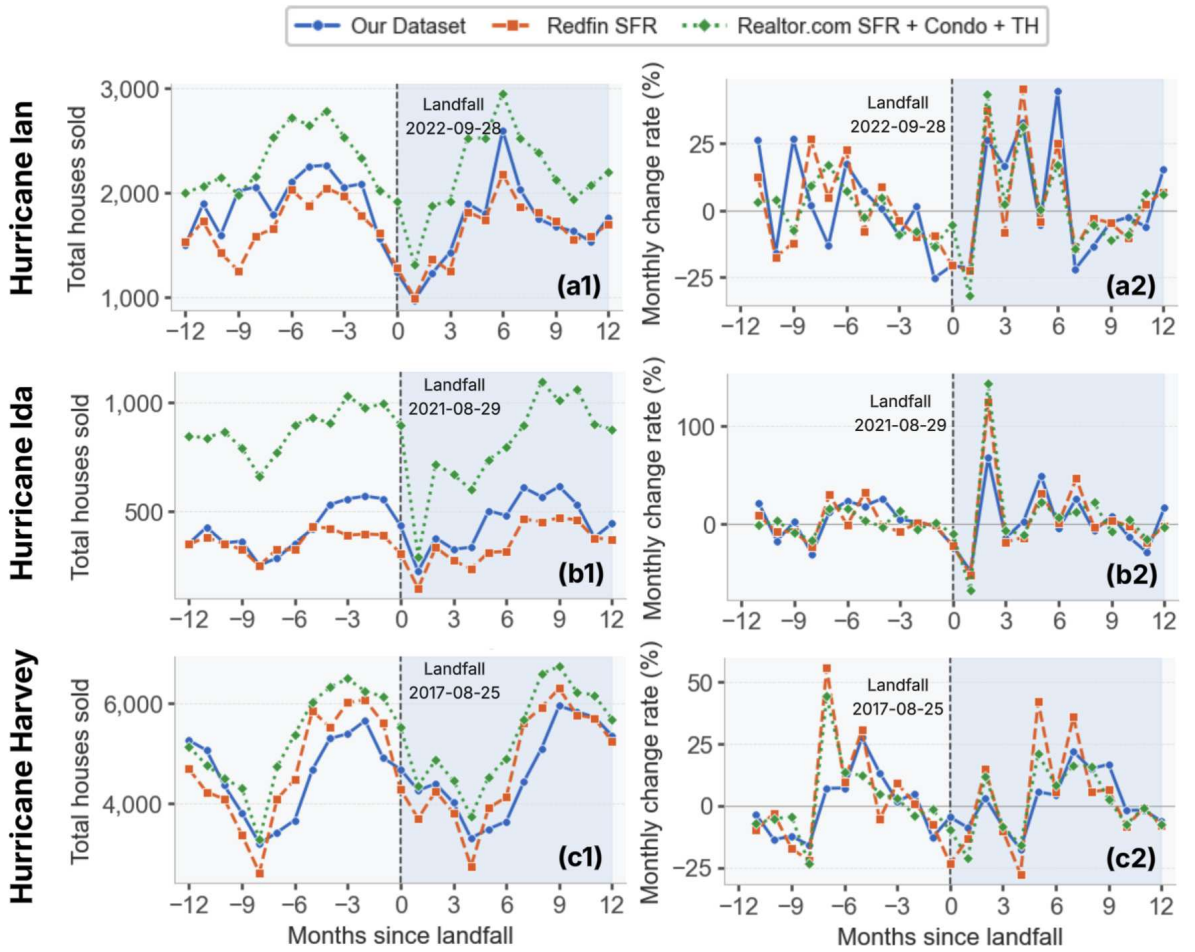


Figure 6: Cross-source validation of monthly sale records. Rows correspond to Hurricane Ian (a), Ida (b), and Harvey (c). Left column (a1, b1, c1): raw monthly sale counts; right column (a2, b2, c2): month-to-month percentage change. HARP (blue) is compared against Redfin single-family residential (orange) and Realtor.com single-family, condominium, and townhouse (green). Realtor.com releases these three property types only as a combined total, so its series cannot be restricted to single-family homes. The x-axis shows months relative to landfall; dashed vertical lines mark landfall.

Damage-count cross-source consistency

To further validate the harmonized damage layer, the HARP damaged-parcel counts are compared against independently reported third-party totals for each event (Table 3). The HARP totals combine FEMA-labeled damaged parcels (Affected, Minor, Major, Destroyed) with the aerial-imagery-derived augmentation released alongside FEMA's record. Across all three events, HARP recovers the same order of

magnitude as the reference counts. HARP is restricted to geocoded single-family residential parcels by construction, whereas the third-party references report broader populations, so the HARP counts are consistently lower in magnitude.

Event	HARP damaged parcels	Reference count
Hurricane Harvey	124,604	154,170 [95]
Hurricane Ida	13,545	53,000 [96]
Hurricane Ian	31,250	52,514 [56]

Table 3: Damage-count cross-source consistency.

Data integrity

We ran automated checks on the released files to confirm that all required fields were present, array dimensions were consistent, and records were correctly aligned across arrays. All three hurricane datasets passed these checks.

Example Analysis

We close with three short analyses that existing resources would make difficult. Each exercises a different part of the release rather than settling a substantive question. The first uses a derived attribute available for one storm only and examines repair pathways by buyer type after Harvey, asking whether investor and individual purchases differ in whether repair came before or after the sale. The second relies on the cross-event harmonization and compares how sale and repair events are sequenced across all three storms. The third uses the released influence networks to relate neighborhood influence intensity to community attributes at the Census Block Group level.

Buyer type and repair pathways.

Figure 7 shows post-hurricane transaction and repair pathways for damaged parcels stratified by buyer type (individual versus investor). Repair timing is classified into three categories: seller-repaired before sale, buyer-repaired after purchase, and sold as-is with no recorded repair. The plurality of damaged parcels were sold as-is, accounting for 37.6–49.7% across panels, indicating widespread ownership transfer without observed repair. Investor purchases involved a higher share of seller-repaired properties than individual purchases, with an overall difference of 6.0 percentage points (33.5% versus 27.5%). The

difference was larger among parcels with Major damage. The pattern is consistent with investor buyers preferentially acquiring properties whose owners completed repairs before listing.

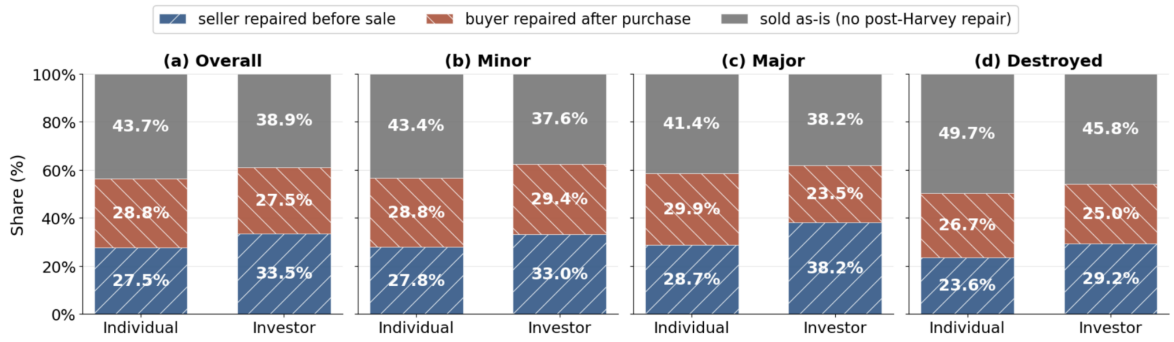


Figure 7: Repair timing in post-Harvey transactions by buyer type and damage level. Distribution of repair timing among damaged parcels that were sold after Hurricane Harvey. Bars show the percentage of post-Harvey transactions associated with individual and investor buyers, stratified by damage level.

Sale and repair sequencing.

Figure 8 characterizes the months between sale and repair events for parcels where both occurred within the twelve-month post-landfall window (the landfall month and the following eleven months), separately for each hurricane. The top row reports the within-group distribution of time gaps for repair-first parcels (those where repair preceded sale); the bottom row reports the analogous distribution for sale-first parcels. Across all three storms, sale-first parcels are more concentrated in the immediate post-landfall months (modal gap of 1–2 months) than repair-first parcels (modal gap of 2–3 months), and both distributions decay to near zero by month 11. The shapes are broadly similar across Ian, Ida, and Harvey, indicating a consistent post-disaster temporal structure between event types.

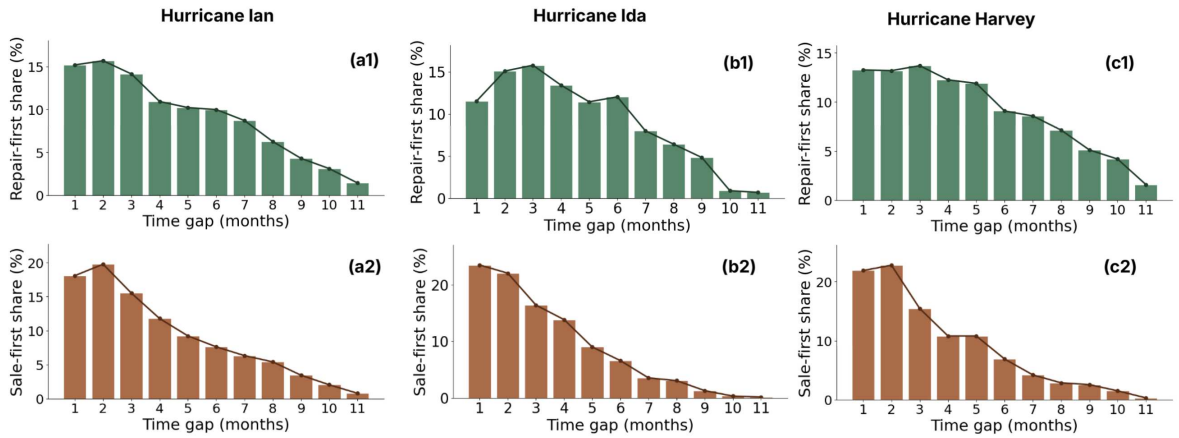


Figure 8: Distribution of time gaps between sale and repair events for parcels with both event types within the twelve-month post-landfall window (landfall month plus the following eleven months), for Hurricanes Ian (panels a1, a2), Ida (b1, b2), and Harvey (c1, c2). Top row (green): within-group distribution of repair-first gaps by months. Bottom row (orange): within-group distribution of sale-first gaps.

Neighborhood influence intensity across community attributes.

Figure 9 illustrates how the released influence networks can be combined with community attributes to summarize variation in neighborhood influence intensity across CBG communities. For each parcel, we compare the event-history signal propagated from its neighbors, computed as $\mathbf{R}_{\text{nei}} = \mathbf{A}\,\mathbf{R}_{\text{self}}$, with the parcel's own event-history signal $\mathbf{R}_{\text{self}}$. The neighborhood-to-self influence ratio is then averaged across active parcels within each CBG. Higher values indicate communities in which repair and sale dynamics are more strongly associated with neighboring parcels' prior activity than with each parcel's own prior activity.

Across all three storms, smaller communities show higher neighborhood-to-self influence ratios, suggesting stronger local decision dependence in more compact parcel networks. Median household income shows only weak variation across panels, while the association with population density differs by storm. The ratio is negative for Ian, approximately flat for Harvey, and slightly positive for Ida, likely reflecting differences in urban form across Lee County, Harris County, and New Orleans.

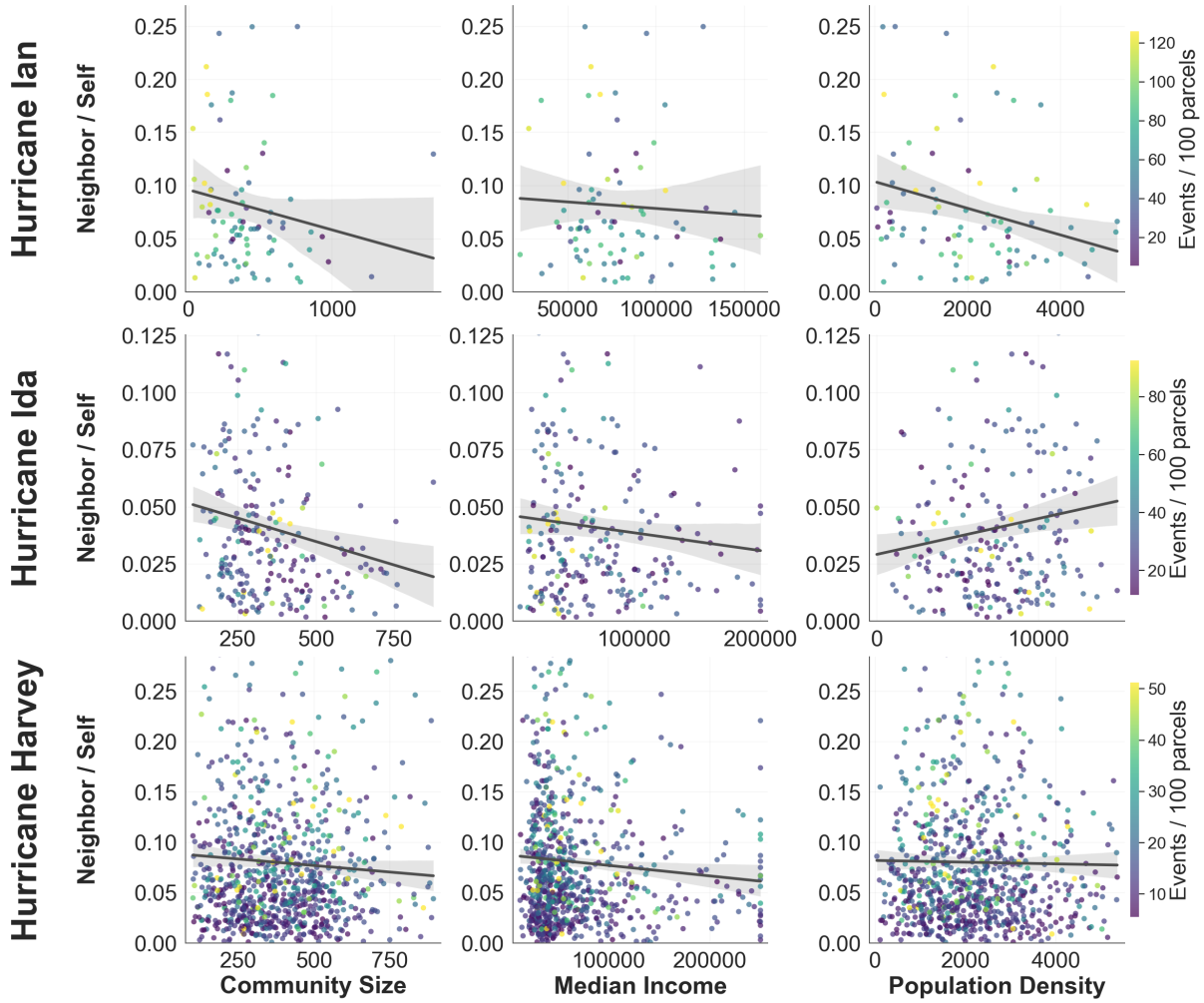


Figure 9: Inferred neighborhood-to-self influence ratio versus three CBG-level community attributes (community size, median household income, population density), for Hurricanes Ian (top), Ida (middle), and Harvey (bottom). Each point is a CBG colored by total post-hurricane events per 100 parcels; the solid line is a per-panel OLS fit with a 95% confidence band.

Code Availability

The code used to construct HARP, including the blue-tarp and fine-tuned DamageCAT damage detection and the buyer-type classification, together with the dataset usage guide, is publicly available in the same repository as the data, `https://github.com/SusuXu-s-Lab/HARP-Dataset` (see the Data Record section). The neighborhood influence networks were inferred with the HDD framework described in the method paper [48].

Usage Notes

HARP attributes are organized into four categories: (i) parcel-level household characteristics, documented separately for the Regrid-derived (Ian, Ida) and HCAD-derived (Harvey) releases (Tables 4 and 5), (ii) Census Block Group community-level attributes joined to each parcel (Table 7), (iii) damage assessment fields (Table 6), and (iv) sale and repair event records (Tables 8–13). The two household-characteristics tables are split because the Regrid and HCAD source schemas surface different underlying fields; Ian and Ida share the Regrid schema (with a small number of Ida-only fields marked †), while Harvey uses the HCAD schema, which includes a broader set of appraisal-district attributes not available from Regrid (identified in the tables by the `hcad_` prefix). Descriptions of these appraisal-district fields in Table 5 follow the official HCAD PDATA Codebook [97]. Sale and repair records use storm-specific field names and are documented separately per hurricane. Community-level attributes are joined to each parcel via its Census Block Group identifier (`census_blockgroup`), so parcels within the same CBG share identical values for these fields.

To protect the privacy of individual households while preserving the analytical value of their attributes, we release household features as learned embeddings rather than raw values. For each household we provide a 128-dimensional embedding that encodes its attributes in a compact latent form. The embedding is produced by standardizing the household's feature vector, projecting it to 128 dimensions, and passing it through two residual blocks. This embedding supports clustering, similarity search, and downstream prediction. Because it is a fixed, non-invertible transformation and the encoder weights are not distributed, the embeddings cannot be mapped back to individual attribute values. Tables 4 and 5 are therefore the one exception to the categories listed above: they document the source attributes the embedding encodes, not columns of the released files. Counting the damage fields of Table 6, these source attributes span 39 columns for Ian, 46 for Ida, and 48 for Harvey. Preprocessing transforms, merges, or drops some of them, so the encoder receives 34, 32, and 34 columns respectively.

Variable	Description
parcelnumb	Regrid parcel identification number.
census_blockgroup	Census Block Group identifier (12-digit FIPS).
lat	Latitude of parcel centroid (WGS-84).
lon	Longitude of parcel centroid (WGS-84).
usecode	Land use code assigned by the county assessor.
usedesc	Text description of the land use type (e.g. single family residential).
zoning	Local zoning code assigned by the jurisdiction (possible-use category).
lbcs_activity	Land-Based Classification Standards (LBCS) Activity code: the actual activity taking place on the land.
lbcs_function	LBCS Function code: the economic function served by the parcel.
lbcs_structure	LBCS Structure code: the type of structure on the parcel.
yearbuilt	Structure year built.
numstories	Number of stories.
sqft†	Building square footage (Ida only; Ian's release omits this field).
totalarea	Total building area (sq ft) of the existing structure on the parcel.
existing_heatedarea	Heated living area (sq ft) of the current structure on the parcel.
existing_bathrooms	Number of bathrooms in the current structure on the parcel.
existing_bedrooms	Number of bedrooms in the current structure on the parcel.
new_heatedarea	Heated area (sq ft) of new construction permitted for the parcel.
new_totalarea	Total area (sq ft) of new construction permitted for the parcel.
new_bathrooms	Number of bathrooms in new permitted construction for the parcel.
new_bedrooms	Number of bedrooms in new permitted construction for the parcel.
improvval	Improvement value: assessed value of improvements (structures) on the parcel.
landval	Land value: assessed value of the land.
parval	Total parcel value.
agval	Agricultural value.
owntype†	Ownership type code (Ida only).
highest_parcel_elevation	Highest elevation value (meters) intersecting the parcel, computed from the USGS 10 m DEM.
lowest_parcel_elevation	Lowest elevation value (meters) intersecting the parcel, computed from the USGS 10 m DEM.
roughness_rating	Within-parcel elevation variability class computed from the USGS 10 m DEM (Level / Nearly Level / Slightly / Intermediately / Moderately / Highly / Extremely Rugged).
transmission_line_distance	Distance (meters) from the parcel boundary to the nearest 69–765 kV electric transmission line.
fema_nri_risk_rating	FEMA National Risk Index (NRI) composite risk rating for the parcel's location.
fema_flood_zone†	FEMA flood-zone designation (Ida only).
nhc_in_tidal_zone†	Whether the parcel falls within the NHC tidal zone (0/1) (Ida only).
population_density	Population density (per square mile) at the Census Block Group level.
population_growth_past_5_years	CAGR of population over the past 5 years at the CBG level.
population_growth_next_5_years	Projected CAGR of population over the next 5 years at the CBG level.
housing_growth_past_5_years	CAGR of housing units over the past 5 years at the CBG level.
housing_growth_next_5_years	Projected CAGR of housing units over the next 5 years at the CBG level.
household_income_growth_next_5_years	Projected CAGR of median household income over the next 5 years at the CBG level.
median_household_income	Median household income (previous year), at the CBG level.
housing_affordability_index	Housing Affordability Index (HAI), at the CBG level.
pre_hurricane_repair_count	Number of repair permits recorded for the parcel before landfall.

† Encoded for Ida only; not part of Ian's feature set. Ian's attributes are otherwise a subset of Ida's.

Table 4: Household characteristics for Hurricane Ian and Hurricane Ida.

Variable	Description
parcelnumb	HCAD account identification number.
census_blockgroup	Census Block Group identifier (12-digit FIPS).
lat	Latitude of parcel centroid (WGS-84).
lon	Longitude of parcel centroid (WGS-84).
yearbuilt	Year the structure was originally built.
usecode	Land use code assigned by the county assessor.
sqft	Building square footage.
totalarea	Total area of the parcel (land and improvements) in square feet.
existing_heatedarea	Heated living area (sq ft) of the existing structure.
existing_bathrooms	Number of bathrooms in the existing structure.
existing_bedrooms	Number of bedrooms in the existing structure.
improvval	Improvement value: assessed value of improvements (structures) on the parcel.
landval	Land value: assessed value of the land.
parval	Total parcel value.
agval	Agricultural value.
hcad_improvval	Building Value.
hcad_landval	Market Land Value.
hcad_assessed_val	Assessed Value.
hcad_tot_appr_val	Appraised Value.
hcad_tot_mkt_val	Market Value.
hcad_tot_rcn_val	Total Building Replace Cost Value.
hcad_x_features_val	Extra Feature Value (value of other features such as a detached frame garage or a pool).
hcad_taxable_val	Taxable Value.
hcad_bld_ar	Total Building Area.
hcad_land_ar	Total Land Area Square Footage.
hcad_acreage	Total Land Area Acreage.
hcad_im_sqft	Improvement Area.
hcad_heated_area	Heated Area.
hcad_actual_area	Actual Area.
hcad_gross_area	Gross Area.
hcad_yr_impr	Parcel Year Improved.
hcad_year_built	Actual Year Built.
hcad_yr_remodel	Year Remodeled.
hcad_bedrooms	Number of bedrooms.
hcad_full_baths	Number of full bathrooms.
hcad_half_baths	Number of half bathrooms.
hcad_total_rooms	Total number of rooms.
hcad_story_height	Number of stories.
hcad_fireplaces	Number of fireplaces.
hcad_state_class	HCAD state / property class code.
hcad_neighborhood	Neighborhood Code.
hcad_market_area	HCAD primary market area code (Market_Area_1, valuation grouping).
hcad_school_dist	Independent School District (ISD) code.
population_density	Population density (per square mile) at the CBG level.
median_household_income	Median household income (previous year), at the CBG level.
pre_hurricane_repair_count	Number of repair permits recorded for the parcel before landfall.

Table 5: Household characteristics for Hurricane Harvey. The unprefixed valuation and area fields and the hcad_-prefixed appraisal-district fields are two parallel families in the HCAD source; both are encoded.

Variable	Description
FloodDepth	Flood depth at the parcel during the hurricane.
BldgValue	Building replacement value at time of damage assessment (Ian and Harvey only).
EstLoss	Estimated loss from the hurricane. **Ian: real dollar value (USD). Harvey and Ida: ordinal severity code.**
nhc_surge_category	NHC storm-surge raster category at the parcel (0–5; 7 = levee-protected) (Ida only).
nhc_surge_depth_ft	NHC modeled storm-surge depth (feet); 0 where none or levee-protected (Ida only).
nhc_tidalmask_category	NHC tidal-mask raster category at the parcel (0–5; 7 = levee, 15 = high-tide/intertidal zone) (Ida only).
Occupancy	Occupancy type of the assessed structure (Ian and Ida only), as a categorical code. Ian: single_family / multi_family / manufactured_home. Ida: the county assessor's land-use code of the matched assessment point.
DamageLevel	Damage label as a categorical code. Label set is event-specific (see Damage Calibration).
AerialDamaged	Damage flag derived from aerial imagery. Available for Ida and Ian.

Table 6: Damage assessment attributes.

Variable	Description
total_population	Total resident population.
total_households	Total number of households.
pct_age_under_18	Share of residents under 18 years of age (%).
pct_age_65_plus	Share of residents aged 65 and over (%).
median_age_total	Median age of all residents (years).
median_age_male	Median age of male residents (years).
median_age_female	Median age of female residents (years).
pct_white_alone	Share of residents reporting White alone (%).
pct_black_alone	Share of residents reporting Black or African American alone (%).
pct_asian_alone	Share of residents reporting Asian alone (%).
pct_two_or_more_race	Share of residents reporting two or more races (%).
pct_family_households	Share of households classified as family households (%).
pct_single_person_hh	Share of households consisting of a householder living alone (%).
pct_hh_with_children	Share of households with one or more children under 18 (%).
median_household_income	Median household income in the previous year (USD).
pct_hh_income_under_25k	Share of households with annual income below $25,000 (%).
pct_in_labor_force	Share of the population aged 16 and over in the labor force (%).
unemployment_rate	Unemployed as a share of the civilian labor force (%).
pct_hs_or_higher	Share of adults aged 25 and over with at least a high-school diploma (%).
pct_bachelor_or_higher	Share of adults aged 25 and over with a bachelor's degree or higher (%).
pct_owner_occupied	Share of occupied housing units that are owner-occupied (%).
pct_renter_occupied	Share of occupied housing units that are renter-occupied (%).
median_home_value	Median value of owner-occupied housing units (USD).
median_gross_rent	Median gross rent for renter-occupied units (USD per month).
median_year_built	Median year the structure was built.
total_housing_units	Total number of housing units, occupied and vacant.
pct_vacant_units	Share of housing units classified as vacant (%).
pct_single_family	Share of housing units in 1-unit detached or attached structures (%).
pct_mobile_home	Share of housing units classified as mobile homes (%).
pct_commute_60min_plus	Share of commuters with a one-way travel time of 60 minutes or more (%).
pct_65plus_living_alone	Share of residents aged 65 and over who are a householder living alone (%).

Table 7: Community-level (Census Block Group) attributes from the American Community Survey.

Variable	Description
first_sale_date	Parcel's sale close date; one transaction recorded per parcel.
listing_date	Listing date corresponding to sale date for the parcel (see Temporal Alignment).
sale_date_history	Semicolon-joined chronological sequence of deed-recorded close dates for every sale in the parcel record.

Table 8: Ian — Sale variables per parcel.

Variable	Description
lee_first_completion_date	Earliest permit completion date for the parcel, defined as the smallest record date among permits whose Status closed with a Certificate of Compliance, Certificate of Occupancy, or Certificate of Use, or that closed as Complete or Effective.
lee_statuses	Semicolon-joined sequence of permit Status labels for every Lee County permit matched to the parcel, sorted earliest first. The full Lee County Accela vocabulary contains 42 values across all record types; the most common ones are `Closed-CC Issued`, `Closed-CO Issued`, `Closed-Revision Approved`, `Permit Expired`, `Closed-Withdrawn`, `Permit Issued`, `Closed-Voided`, `Waiting on Applicant`, and `Closed-Cert of Use Issued`.
lee_record_dates	Semicolon-joined sequence of record dates corresponding to lee_statuses for those permits, in the same order.
fema_applied_dates	Semicolon-joined sequence of FEMA IHP application dates assigned to the parcel via zip-level mapping, ordered earliest first.

Table 9: Ian — Repair variables per parcel.

Variable	Description
first_sale_date	Parcel's sale close date; one transaction recorded per parcel.
listing_date	Listing date corresponding to sale date for the parcel.

Table 10: Ida — Sale variables per parcel.

Variable	Description
nola_app_submit_dates	Semicolon-joined sequence of application-submission dates for every New Orleans permit matched to the parcel, ordered earliest first.
nola_app_review_dates	Semicolon-joined sequence of application-review dates for those same permits, in the same order.
nola_plan_review_dates	Semicolon-joined sequence of plan-review dates for those same permits, in the same order.
nola_zoning_review_dates	Semicolon-joined sequence of zoning-review dates for those same permits, in the same order.
nola_approved_dates	Semicolon-joined sequence of permit-approval dates for those same permits, in the same order.
nola_issued_dates	Semicolon-joined sequence of permit-issuance dates for those same permits, in the same order.
nola_inspection_dates	Semicolon-joined sequence of first-inspection dates for those same permits, in the same order.
nola_co_dates	Semicolon-joined sequence of Certificate of Occupancy dates (final building-completion sign-off) for those same permits, in the same order.
fema_applied_dates	Semicolon-joined sequence of FEMA IHP application dates assigned to the parcel via zip-level mapping, ordered earliest first.

Table 11: Ida — Repair variables per parcel.

Variable	Description
first_sale_date	Earliest HCAD deed-of-sale date across the parcel's recorded transaction history.
listing_date	Listing date corresponding to sale date for the parcel.
sale_date_history	Semicolon-joined chronological sequence of HCAD deed-of-sale dates.

Table 12: Harvey — Sale variables per parcel.

Variable	Description
hcad_issue_dates	Semicolon-joined sequence of HCAD permit issue dates for every permit matched to the parcel by account number, ordered earliest first.
fema_applied_dates	Semicolon-joined sequence of FEMA IHP application dates assigned to the parcel via zip-level mapping, ordered earliest first.

Table 13: Harvey — Repair variables per parcel.

References

1. Nejat, A., Cong, Z. & Liang, D. Family structures, relationships, and housing recovery decisions after Hurricane Sandy. *Buildings* **6,** 14 (2016).

2. Park, G. A comprehensive analysis of hurricane damage across the US gulf and atlantic coasts using geospatial big data. *ISPRS International Journal of Geo-Information* **10,** 781 (2021).

3. NOAA National Centers for Environmental Information. *Billion-Dollar Weather and Climate Disasters* `https://www.ncei.noaa.gov/access/billions/`. DOI: 10.25921/stkw-7w73 (accessed 2026-03-27).

4. Waddell, S. L., Jayaweera, D. T., Mirsaeidi, M., Beier, J. C. & Kumar, N. Perspectives on the health effects of hurricanes: a review and challenges. *International journal of environmental research and public health* **18,** 2756 (2021).

5. Coleman, N., Liu, C. & Mostafavi, A. Anatomizing societal recovery at the microscale: Heterogeneity in household lifestyle activities rebounding after disasters. *arXiv preprint arXiv:2406.17993* (2024).

6. Kates, R. W., Colten, C. E., Laska, S. & Leatherman, S. P. Reconstruction of New Orleans after Hurricane Katrina: a research perspective. *Proceedings of the national Academy of Sciences* **103,** 14653–14660 (2006).

7. Miles, S. B., Burton, H. V. & Kang, H. Community of practice for modeling disaster recovery. *Natural Hazards Review* **20,** 04018023 (2019).

8. Aldrich, D. P. *Building resilience: Social capital in post-disaster recovery* (University of Chicago Press, 2012).

9. Zhang, Y. & Peacock, W. G. Planning for housing recovery? Lessons learned from Hurricane Andrew. *Journal of the American Planning Association* **76,** 5–24 (2009).

10. Hamideh, S., Peacock, W. G. & Van Zandt, S. Housing type matters for pace of recovery: Evidence from Hurricane Ike. *International journal of disaster risk reduction* **57,** 102149 (2021).

11. Peacock, W. G., Dash, N. & Zhang, Y. in *Handbook of Disaster Research* 258–274 (Springer, New York, NY, 2007).

12. Burton, H., Kang, H., Miles, S., Nejat, A. & Yi, Z. A framework and case study for integrating household decision-making into post-earthquake recovery models. *International Journal of Disaster Risk Reduction* **37,** 101167 (2019).

13. Moradi, S. & Nejat, A. RecovUS: An agent-based model of post-disaster household recovery. *Journal of Artificial Societies and Social Simulation* **23,** 13 (2020).

14. Binder, S. B., Baker, C. K. & Barile, J. P. Rebuild or relocate? Resilience and postdisaster decision-making after Hurricane Sandy. *American journal of community psychology* **56,** 180–196 (2015).

15. Negri, R. *et al.* When Recovery Becomes Infeasible: A Markov Model of Housing Abandonment Risk in Flood-Prone Areas. Preprint (2025).

16. Negri, R., Dragomir, C., Xu, S., Porfiri, M. & Ceferino, L. Permanent relocation into and out of areas exposed to natural hazards: A multidisciplinary review of the literature (2025).

17. Bukvic, A., Whittemore, A., Gonzales, J. & Wilhelmi, O. Understanding relocation in flood-prone coastal communities through the lens of place attachment. *Applied Geography* **146,** 102758 (2022).

18. Miller, W. A. Leveraging the Microgeography of Neighborhoods to Estimate the Causal Effects of Flood Damage: Evidence from Hurricane Harvey. *Available at SSRN 5376813* (2025).

19. Van Zandt, S. *et al.* Mapping social vulnerability to enhance housing and neighborhood resilience. *Housing Policy Debate* **22,** 29–55 (2012).

20. Binder, S. B. *et al.* Limbo: the unintended consequences of home buyout programmes on peripheral communities. *Environmental Hazards* **19** (2020).

21. Hashida, Y. & Dundas, S. J. The effects of a voluntary property buyout and acquisition program on coastal housing markets: Evidence from New York. *Journal of Environmental Economics and Management* **121** (2023).

22. Schwaller, N. L., Jaramillo, A. & El-Khattabi, A. R. Postdisaster Investment in Single-Family Housing: Insights for Harris County, Texas. *Journal of the American Planning Association,* 1–14 (2026).

23. Van Holm, E. J. & Wyczalkowski, C. K. Gentrification in the wake of a hurricane: New Orleans after Katrina. *Urban Studies* **56** (2019).

24. Hu, D. & Nejat, A. Role of spatial effect in postdisaster housing recovery: Case study of Hurricane Katrina. *Journal of Infrastructure Systems* **27,** 05020009 (2021).

25. Airriess, C. A., Li, W., Leong, K. J., Chen, A. C.-C. & Keith, V. M. Church-based social capital, networks and geographical scale: Katrina evacuation, relocation, and recovery in a New Orleans Vietnamese American community. *Geoforum* **39,** 1333–1346 (2008).

26. Irwin, N. B. & Livy, M. R. Close competition: The contagion effect of neighborhood listings on seller motivation (2023).

27. Xiao, Y. & Van Zandt, S. Building community resiliency: Spatial links between household and business post-disaster return. *Urban Studies* **49,** 2523–2542 (2012).

28. Zhang, Y. Will natural disasters accelerate neighborhood decline? A discrete-time hazard analysis of residential property vacancy and abandonment before and after Hurricane Andrew in Miami-Dade County (1991–2000). *Environment and Planning B: Planning and Design* **39,** 1084–1104 (2012).

29. Liu, C.-F. & Mostafavi, A. Network diffusion model reveals recovery multipliers and heterogeneous spatial effects in post-disaster community recovery. *Scientific Reports* **13,** 19032 (2023).

30. Blackwood, L. & Cutter, S. L. The application of the Social Vulnerability Index (SoVI) for geo-targeting of post-disaster recovery resources. *International Journal of Disaster Risk Reduction* **92,** 103722 (2023).

31. Peacock, W. G., Van Zandt, S., Zhang, Y. & Highfield, W. E. Inequities in long-term housing recovery after disasters. *Journal of the American Planning Association* **80,** 356–371 (2014).

32. Dwyer, C. & Horney, J. Validating indicators of disaster recovery with qualitative research. *PLoS currents* **6,** ecurrents–dis (2014).

33. Fu, C. & Gregory, J. Estimation of an equilibrium model with externalities: Post-disaster neighborhood rebuilding. *Econometrica* **87,** 387–421 (2019).

34. Nejat, A. & Ghosh, S. LASSO model of postdisaster housing recovery: Case study of Hurricane Sandy. *Natural Hazards Review* **17,** 04016007 (2016).

35. Manski, C. F. Identification of endogenous social effects: The reflection problem. *The review of economic studies* **60,** 531–542 (1993).

36. Nejat, A., Javid, R. J., Ghosh, S. & Moradi, S. A spatially explicit model of postdisaster housing recovery. *Computer-Aided Civil and Infrastructure Engineering* **35,** 150–161 (2020).

37. Deng, H. *et al.* High-resolution human mobility data reveal race and wealth disparities in disaster evacuation patterns. *Humanities and Social Sciences Communications* **8,** 144 (2021).

38. Hong, B., Bonczak, B. J., Gupta, A. & Kontokosta, C. E. Measuring inequality in community resilience to natural disasters using large-scale mobility data. *Nature Communications* **12,** 1870 (2021).

649 39. Li, Z., Ning, H., Jing, F. & Lessani, M. N. Understanding the bias of mobile location data across
650 spatial scales and over time: A comprehensive analysis of SafeGraph data in the United States.
651 *PLOS ONE* **19,** e0294430 (2024).

652 40. Guan, X. & Chen, C. A behaviorally-integrated individual-level state-transition model that can
653 predict rapid changes in evacuation demand days earlier. *Transportation research part E: logistics*
654 *and transportation review* **152,** 102381 (2021).

655 41. Long, E. F., Chen, M. K. & Rohla, R. Political storms: Emergent partisan skepticism of hurricane
656 risks. *Science advances* **6,** eabb7906 (2020).

657 42. Smith, S. K. & McCarty, C. Housing Damage and Population Displacement During Florida's 2004
658 Hurricane Season. *Journal of Florida Studies* **1** (2011).

659 43. Plyer, A., Bonaguro, J. & Hodges, K. Using administrative data to estimate population displacement
660 and resettlement following a catastrophic US disaster. *Population and Environment* **31,** 150–175
661 (2010).

662 44. Rathfon, D., Davidson, R., Bevington, J., Vicini, A. & Hill, A. Quantitative assessment of post-
663 disaster housing recovery: A case study of Punta Gorda, Florida, after hurricane charley. *Disasters*
664 **37,** 333–355 (2013).

665 45. Lozano, J.-M. & Tien, I. Data collection tools for post-disaster damage assessment of building and
666 lifeline infrastructure systems. *International journal of disaster risk reduction* **94,** 103819 (2023).

667 46. Federal Emergency Management Agency. *Hurricane Ida Final Visual Damage Assessment* FEMA
668 Hub (ArcGIS Feature Server). 2021. `https://services.arcgis.com/XG15cJAlne2vxtgt/arcgis/`
669 `rest/services/Hurricane_Ida_Final_Damage_Assessment_WFL1/FeatureServer/0`.

670 47. Nejat, A. & Damnjanovic, I. Agent-based modeling of behavioral housing recovery following disas-
671 ters. *Computer-Aided Civil and Infrastructure Engineering* **27,** 748–763 (2012).

672 48. Li, X. *et al. Repair and Sale Diffusion Drives Neighborhood Recovery Divergence after Hurricanes*
673 2025.

674 49. Gomez-Rodriguez, M., Leskovec, J. & Krause, A. Inferring Networks of Diffusion and Influence. In
675 *KDD* (2010).

50. Gomez-Rodriguez, M., Balduzzi, D. & Sch"olkopf, B. Uncovering the Temporal Dynamics of Diffusion Networks. In *ICML* (2011).

51. Gomez Rodriguez, M., Leskovec, J. & Schölkopf, B. Structure and dynamics of information pathways in online media. In *Proceedings of the sixth ACM international conference on Web search and data mining* 23–32 (2013).

52. NOAA National Centers for Environmental Information. *Costliest U.S. Tropical Cyclones* https://www.ncei.noaa.gov/access/billions/dcmi.pdf. Accessed: 2026-05-19. 2025.

53. Harris County Flood Control District. *Hurricane Harvey: Preliminary Report* https://www.bellairetx.gov/DocumentCenter/View/20559. Accessed: 2026-05-19. 2017.

54. Beven, J. L., Hagen, A. & Berg, R. *Tropical Cyclone Report: Hurricane Ida* tech. rep. AL092021. Accessed: 2026-05-19 (National Hurricane Center, Apr. 2022). https://www.nhc.noaa.gov/data/tcr/AL092021_Ida.pdf.

55. National Weather Service. *Service Assessment: 2021 Hurricane Ida* tech. rep. Accessed: 2026-05-19 (National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD, Apr. 2023). https://www.weather.gov/media/publications/assessments/Hurricane_Ida_Service_Assessment.pdf.

56. Bucci, L., Alaka, L., Hagen, A., Delgado, S. & Beven, J. *Tropical Cyclone Report: Hurricane Ian* tech. rep. AL092022. Original report dated 2 April 2023; revised 12 March 2026. Accessed: 2026-05-19 (National Hurricane Center, Apr. 2023). https://www.nhc.noaa.gov/data/tcr/AL092022_Ian.pdf.

57. EBSCO Research Starters. *Hurricane Ian* https://www.ebsco.com/research-starters/earth-and-atmospheric-sciences/hurricane-ian. Accessed: 2026-05-19. 2022.

58. Texas Department of Insurance. *Hurricane Harvey Data Call: Final Report* https://www.tdi.texas.gov/reports/documents/harvey-dc-final-06302019.pdf. Accessed May 2026. 2019.

59. Cape Analytics. *Hurricane Ida Damage Tracking* https://capeanalytics.com/resources/hurricane-ida-damage/. Accessed May 2026. 2021.

60. Florida Office of Insurance Regulation. *Hurricane Ian Claims Data* https://floir.gov/home/ian. Accessed May 2026. 2025.

61. Harris County Appraisal District. *Harris County Appraisal District Public Data Extracts* 2023. https://hcad.org/pdata/.

62. Regrid. *Regrid Parcel Data Platform* https://app.regrid.com. 2022.

63. U.S. Census Bureau. *American Community Survey 5-Year Data (2009–2024)* https://www.census.gov/data/developers/data-sets/acs-5year.html. Accessed 2026-03-27.

64. U.S. Census Bureau. *TIGER/Line Shapefiles* https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html. Accessed 2026-03-27.

65. Federal Emergency Management Agency, Response Geospatial Office. *FEMA Historical Geospatial Damage Assessments: Hurricane Ian* 2023. https://gis-fema.hub.arcgis.com/datasets/fema-historical-geospatial-damage-assessments.

66. Lee County Emergency Management. *Hurricane Ian Initial Damage Assessment* 2022. https://experience.arcgis.com/experience/27c1613132ac4bbf91c17f203b66ff0b.

67. Ceferino, L. *et al. StEER: Hurricane Ian Field Assessment Structural Team (FAST) Dataset* 2025. https://www.designsafe-ci.org/data/browser/public/designsafe.storage.published/PRJ-3709.

68. Federal Emergency Management Agency. *Hurricane Harvey Building Damage Assessment* 2018.

69. Roueche, D. B., Lombardo, F. T., Krupar III, R. J. & Smith, D. J. *Collection of Perishable Data on Wind- and Surge-Induced Residential Building Damage During Hurricane Harvey (TX)* 2018. https://www.designsafe-ci.org/data/browser/public/designsafe.storage.published/PRJ-1668.

70. Lee County Government. *Administrative Building Permit Records for Lee County* https://aca-prod.accela.com/LEECO/Cap/CapHome.aspx. Accessed via the Accela Civic Platform. 2022.

71. Office of the Governor of Florida. *Executive Order 22-218: Emergency Management — Hurricane Ian* 2022. https://www.flgov.com/2022/09/24/governor-ron-desantis-issues-executive-order-22-218/.

72. City of New Orleans. *Department of Safety and Permits: Storm Repair Permit Requirements* 2025. https://www.nola.gov/safety-and-permits/.

73. Federal Emergency Management Agency. *OpenFEMA Dataset: Individual Assistance Housing Registrants — Large Disasters* 2023. https://catalog.data.gov/dataset/individual-assistance-housing-registrants-large-disasters.

74. U.S. Department of Agriculture, Farm Service Agency. *National Agriculture Imagery Program (NAIP) DOQQ* Google Earth Engine ImageCollection USDA/NAIP/DOQQ. 2024. https://developers.google.com/earth-engine/datasets/catalog/USDA_NAIP_DOQQ.

75. Federal Emergency Management Agency. *Damage Assessment Quick Guide* https://gis-fema.hub.arcgis.com. Assessment criteria used for FEMA geospatial damage assessment products. Sept. 2023.

76. Xiao, Y. & Mostafavi, A. DamageCAT: A deep learning transformer framework for typology-based post-disaster building damage categorization. *International Journal of Disaster Risk Reduction,* 105704 (2025).

77. Zheng, Z., Zhong, Y., Wang, J., Ma, A. & Zhang, L. Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters. *Remote Sensing of Environment* **265,** 112636 (2021).

78. Microsoft AI for Good Research Lab. *Building Damage Assessment (CNN-Siamese) Reference Implementation* https://github.com/microsoft/building-damage-assessment-cnn-siamese. Accessed May 2026. 2021.

79. Bandara, W. G. C. & Patel, V. M. A Transformer-Based Siamese Network for Change Detection. In *IGARSS 2022 – 2022 IEEE International Geoscience and Remote Sensing Symposium* 207–210 (2022).

80. U.S. Army Corps of Engineers. *Operation Blue Roof* https://www.usace.army.mil/Missions/Emergency-Operations/Blue-Roof-Info/. Federal disaster-response program. 2021.

81. Federal Emergency Management Agency. *FEMA National Flood Insurance Program Claims Data* 2022. https://www.fema.gov/api/open/v2/FimaNfipClaims.

82. Achiam, J. *et al.* Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

83. Liu, Y. *et al.* Semantic trajectory data mining with llm-informed poi classification. In *2024 IEEE 27th International Conference on Intelligent Transportation Systems (ITSC)* 207–213 (2024).

84. New Orleans City Council. *Council Waives Certain Building and Demolition Permit Fees for Properties Damaged by Hurricane Ida* https://council.nola.gov/news/archive/2021/october-2021/council-waives-certain-building-and-demolition-per/. 2021.

85. Lee County Department of Community Development. *Lee County Sets Up Emergency Building and Well Repair Permitting Operation Post-Hurricane Ian* https://www.leegov.com/dcd/News%20Documents/Lee%20County%20Sets%20up%20Emergency%20Building%20and%20Well%20Repair%20Permitting%20Operation%20Post-Hurricane%20Ian.pdf. 2022.

86. Federal Emergency Management Agency. *Individuals and Households Program* https://www.fema.gov/assistance/individual/program.

87. KFDM News. *FEMA Registration Deadline for Harvey Assistance Extended to November 30* https://kfdm.com/news/local/fema-registration-deadline-for-harvey-assistance-extended-to-nov-30. 2017.

88. FEMA Region 6. *DR-4611-LA NR-029: Deadline to Apply for FEMA Assistance Today, But FEMA Remains in Louisiana* https://www.jedco.org/wp-content/uploads/2021/11/DR-4611-LA-NR-029-Deadline-to-Apply-for-FEMA-Assistance-Today-But-FEMA-Remains-in-Louisiana.pdf. FEMA news release. 2021.

89. Town of Fort Myers Beach. *FEMA Assistance Application Deadline Extended for Hurricane Ian* https://www.fortmyersbeachfl.gov/CivicAlerts.aspx?AID=940&ARC=2719. Civic alert. 2022.

90. U.S. Small Business Administration. *SBA Tops $3 Billion in Disaster Assistance Loans to Businesses and Residents Impacted by Hurricane Harvey* https://www.sba.gov/article/2018/jan/03/tx-15274-20-sba-tops-3-billion-disaster-assistance-loans-businesses-residents-impacted-hurricane. Press release. 2018.

91. U.S. Small Business Administration. *Deadline Approaching for Louisiana SBA Disaster Loans for Property Damage Due to Hurricane Ida* https://www.sba.gov/article/2021/oct/08/deadline-approaching-louisiana-sba-disaster-loans-property-damage-due-hurricane-ida. Press release. 2021.

92. U.S. Small Business Administration. *June 29 SBA Working Capital Disaster Loan Deadline for Florida Small Businesses Affected by Hurricane Ian* https://www.sba.gov/article/2023/05/

30/june-29-sba-working-capital-disaster-loan-deadline-florida-small-businesses-affected-hurricane-ian. Press release. 2023.

93. Redfin. *Redfin Housing Market Data Center* 2024. https://www.redfin.com/news/data-center/.

94. Realtor.com Economic Research. *Residential Real Estate Data Library* https://www.realtor.com/research/data/. Accessed April 2026. 2026.

95. Community Impact Newspaper. *Hurricane Harvey Damage Report: Harris County* 2018. https://communityimpact.com/news/2018/06/08/hurricane-harvey-damage-report-harris-county.

96. City of New Orleans Mayor's Office. *City of New Orleans Provides Hurricane Ida Response and Recovery Updates* https://nola.gov/next/mayors-office/news/articles/2021/september-2021/city-of-new-orleans-provides-hurricane-ida-response-and-recovery-updates-en/. 2021.

97. Harris Central Appraisal District. *PDATA Codebook: Public Data Field Names and Record Layout* https://hcad.org/assets/uploads/pdf/pdataCodebook.pdf. Accessed 2026-06-24.

Funding

This work was supported by the National Science Foundation through the CLIMA/Collaborative Research project *Riders on the Storm: Climate Adaptation of Housing Infrastructure for Resilient Communities in Coastal Cities* .