

2022/05/17

理論談話会#7

Measuring inequality in community resilience to natural disasters using large-scale mobility data

Boyeong Hong, Bartosz J. Bonczak, Arpit Gupta, Constantine E. Kontokosta

Nature Communications, Vol. 12, 1870, (2021)

M1 近藤愛子

課題

特性の違う地域の災害反応・復旧の時空間的動態を考慮した
一般化・計量可能な、「レジリエンス」の客観的な尺度の開発

自然災害に限らず、極度の/予期せぬ
出来事への反応

- Houston, Texasの人口約35%にあたる800,000人の位置情報データを用いて, Hurricane Harvey (Aug, 2017)への応答を分析
- **発災前→発災→被災後の行動の変化**に着目し, Community resilience capacity を「災害の影響の大きさ」と「復旧までの時間」の関数として定義

▶ resilience capacityや避難パターンは地区居住者の社会経済的特性や人種の傾向により
明確な格差

災害時の行動反応に関する新たな洞察

▶ 脆弱性の高い地域に優先的に資源配分するような, 公共部門のデータ駆動型
の意思決定に貢献

Main Article

- Introduction
- Results
- Discussion
- Methods

Supplementary Information

*図表の半分ほどが補足資料に掲載

本日の流れ

0. Introduction
1. データ概要
2. 地区の分類と resilience capacity
 1. Methods
 2. Results
3. 移動パターン
 1. Methods
 2. Results
4. Discussion

問題意識

気候変動による気象被害の拡大

政策立案者が避難, その影響, 復旧を **時空間的に高解像度**で理解するためのツール不足

- ▶ 解像度の高い情報があれば, **地区ごと**の避難の効率性や格差を評価でき, 資源配分や長期的な政策立案につながる

位置情報が付加された大規模なモバイルデータ, ビッグデータ

より頑健で, 検証可能な都市モデルの開発. 特に交通需要推定に役立ってきた
Urban Resilienceに適用された例はない

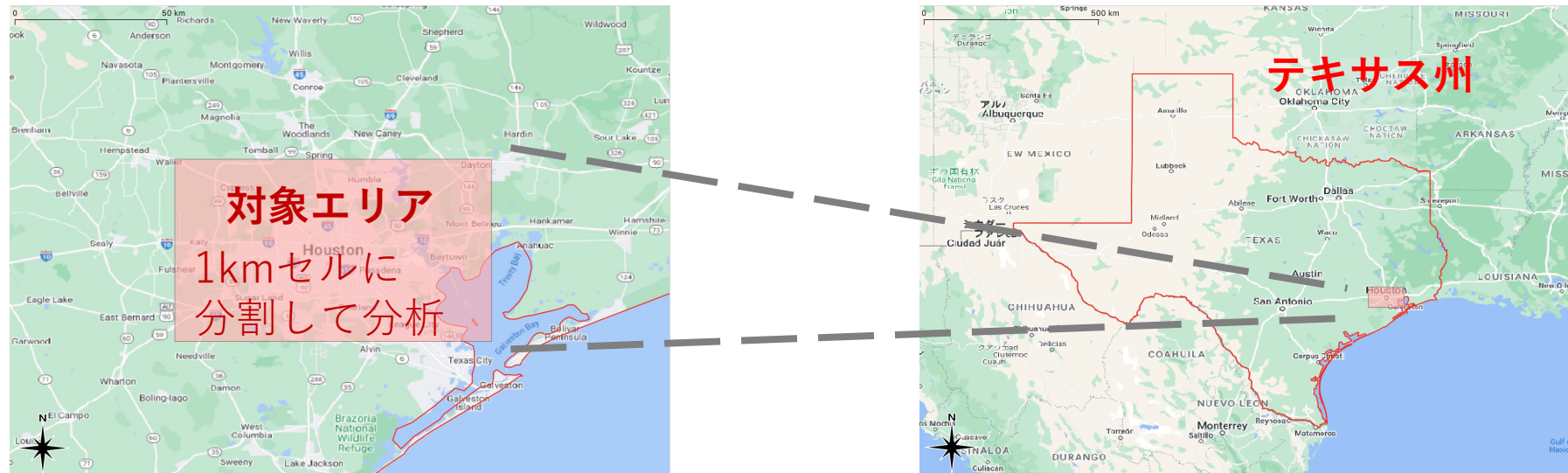
災害に関する既往研究は, 少サンプルの質的調査に基づくことが多かったが, バイアスや規模の小ささ

- ▶ 大規模なモバイルデータを活用し, 大規模でありながら, 地区ごとの特性に着目した分析を行う

1. データ概要

内容	対象日時	対象エリア	サンプル数
Venpath. Inc. から取得した匿名化された携帯電話の位置情報	2017年8月1日～9月30日*	Harris County (Greater Houston Area)	829,350機器

*Hurricane Harveyは8月23日に熱帯低気圧に，25日～30日にTexas上陸¹⁾



2.地区の分類と resilience capacity

■ 地区ごとの避難・復旧パターンの定量化

1. user u , day n の活動拠点（家や避難先など） H_{u,d_n} の特定

$$H_{(u,d_n)} = \operatorname{argmax}(\forall p \in P_u | t_p \in R_u : \sum R_u(i,p))$$

平日は午後8時から午前7時までのみを考慮
最も観測される回数が多いセル p を活動拠点とする

P_u : ユーザ u が観測された全セル集合

t_p : セル p にいた時刻

R_u : 時刻集合?

$\sum R_u(i,p)$: メッシュ p が観測された回数

2. 近隣に滞在するユーザ数の変化割合

$$D_{N,d_n} = \frac{U_{Nd_n} - U_{NA}}{U_{NA}}$$

U_{NA} : 8/1~8/16のセル N の平均ユーザ数

U_{Nd_n} : day n におけるセル N の平均ユーザ数

対象期間中の活動レベルのばらつきを考慮する指標

■ 地区ごとの避難・復旧パターンの定量化

3. 災害応答・復旧パターンを基にした地区のクラス分け

Agglomerative hierarchical clustering algorithm

凝集型

階層的

クラスタリング

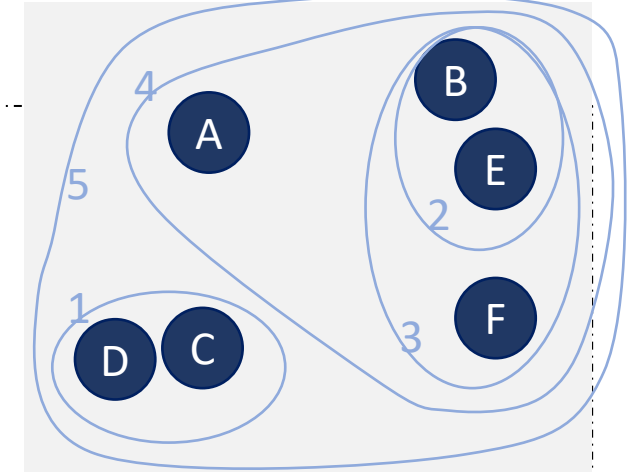
クラスター C_i, C_j 間の距離 $D_w(C_i, C_j)$ は Ward 距離を採用：

$$D_w(C_i, C_j) = \sum_{x \in C_i} (x - r_i)^2 + \sum_{x \in C_j} (x - r_j)^2 + \sum_{x \in C_{ij}} (x - r_{ij})^2$$

C_{ij} : C_i と C_j の和クラスター
 r_k : C_k のセントロイド

各セルの入力は、 D_{N,d_n} の3日間平均ベクトル(8/4~9/26)

「クラスター内での分散を最小化」かつ「クラスター間の分散を最大化」
するような個数に分類する



■ 地区ごとの避難・復旧パターンの定量化

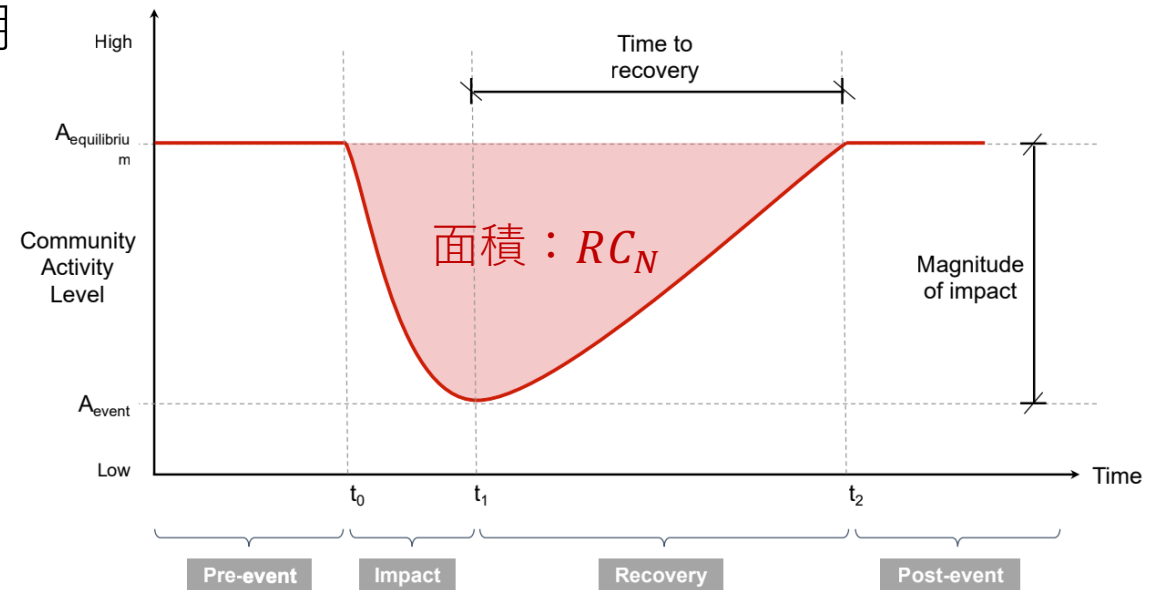
4. Community resilience capacityの定式化

地区クラスごとの D_{N,d_n} の3日間平均を時間軸上にプロットした
conceptual resilience curve²⁾

pre-event, impact, recovery, post-eventの4期
復旧にかかる時間と, 影響の大きさ

- ▶ 両方を考慮した指標が**必要**
- ▶ クラス N のresilience capacity RC_N を,
Area under/over the curve(AUC)として定義

$$RC_N = \int_{t_0}^{t_2} A_{equilibrium} - A_N(t) dt$$



2.2 Results

■ 地区クラス分け結果

Methods 2,3により4つに分類

Group 4: 活動が増加, “Shelter-in-place” (=一時避難場所)

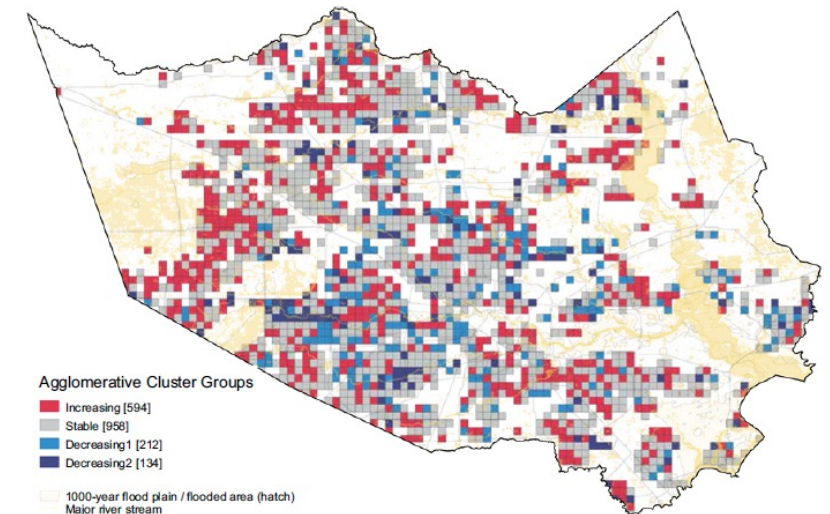
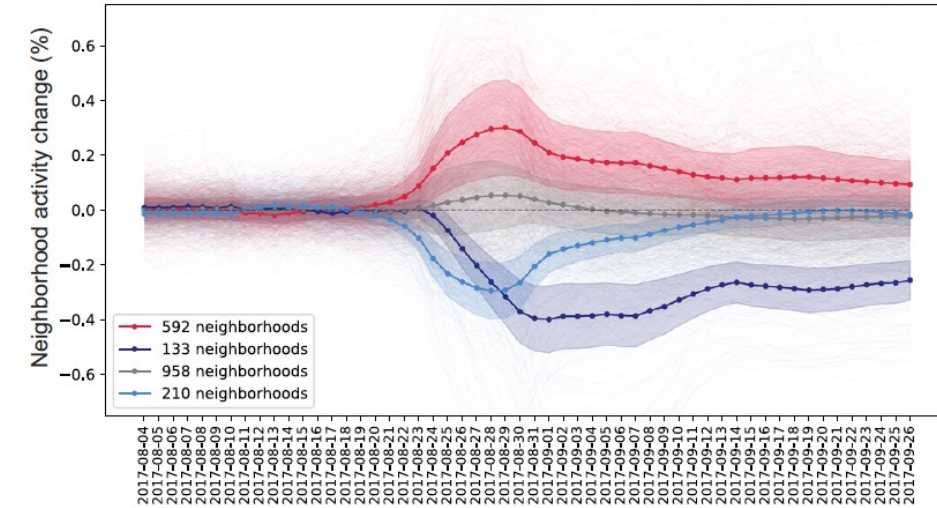
Group 3: 活動レベルは一定, “Stable”

Group 2: 活動が減少, “Distressed”

Group 1: 活動が著しく減少, “Abandoned”

グループ毎の「影響の大きさ」 「復旧にかかる時間」 「Resilience capacity」

Neighborhood cluster	# Of neighborhoods	Impact	Time to recovery	Resilience capacity
Group 4—Shelter-in-place (red)	958	0.37	2 weeks	5.67
Group 3—Stable (gray)	592	0.10	5 days	0.09
Group 2—Distressed (sky blue)	212	0.34	2 weeks	-3.65
Group 1—Abandoned (dark blue)	133	0.48	>3 weeks	-9.86



Groupごとにどのような特徴があるのか？

Table 2 Neighborhood cluster characteristics.

Feature	4—Shelter-in-place n = 958	3—Stable n = 592	2—Distressed n = 212	1—Abandoned n = 133
Demographic and socioeconomic features				
Black (%)	0.16 (0.18)	0.17 (0.18)	0.20 (0.23)	0.12 (0.14)
Hispanic (%)	0.38 (0.24)	0.39 (0.23)	0.44 (0.26)	0.33 (0.23)
Limited English speakers (%)	0.10 (0.11)	0.11 (0.11)	0.15 (0.15)	0.09 (0.12)
Educational attainment (College degree)	0.10 (0.09)	0.11 (0.10)	0.10 (0.11)	0.15 (0.13)
Educational attainment (High school degree)	0.23 (0.08)	0.23 (0.09)	0.25 (0.10)	0.19 (0.10)
Median income (USD)	\$75,545 (\$57,244)	\$71,718 (\$36,951)	\$61,157 (\$35,631)	\$82,126 (\$42,390)
Unemployment rate	0.06 (0.03)	0.06 (0.04)	0.07 (0.04)	0.05 (0.03)
Households without health insurance (%)	0.18 (0.11)	0.19 (0.11)	0.23 (0.12)	0.16 (0.11)
Households with food stamps (%)	0.11 (0.09)	0.12 (0.09)	0.14 (0.10)	0.09 (0.09)
Households without internet (%)	0.15 (0.12)	0.17 (0.13)	0.22 (0.15)	0.14 (0.13)
Homeowners (%)	0.68 (0.20)	0.62 (0.22)	0.54 (0.24)	0.63 (0.23)
Households living in mobile homes (%)	0.03 (0.07)	0.03 (0.06)	0.05 (0.10)	0.03 (0.07)
Land use and topographical features				
Median building age (years)	32 (16)	35 (56)	39 (14)	35 (14)
Median number of rooms	6.01 (1.16)	5.71 (1.22)	5.24 (1.11)	5.84 (1.27)
Vacancy rate (%)	0.07 (0.05)	0.08 (0.05)	0.10 (0.05)	0.09 (0.06)
Multifamily housing (%)	0.02 (0.03)	0.03 (0.05)	0.04 (0.08)	0.03 (0.05)
Lower elevation (% of land area)	0.07 (0.16)	0.08 (0.17)	0.09 (0.17)	0.13 (0.20)
Floodplain (% of land area)	0.30 (0.26)	0.30 (0.28)	0.34 (0.32)	0.45 (0.31)

Statistically significant differences between groups are tested using one-way ANOVA (analysis of variance) and Tukey's multi-comparison method. Mean values with standard deviation in parentheses; highest values across the groups in bold.

1の次に裕福
避難場所は4に立地

マイノリティ
所得低

収入や
教育レベル高

地形リスクは最も低い

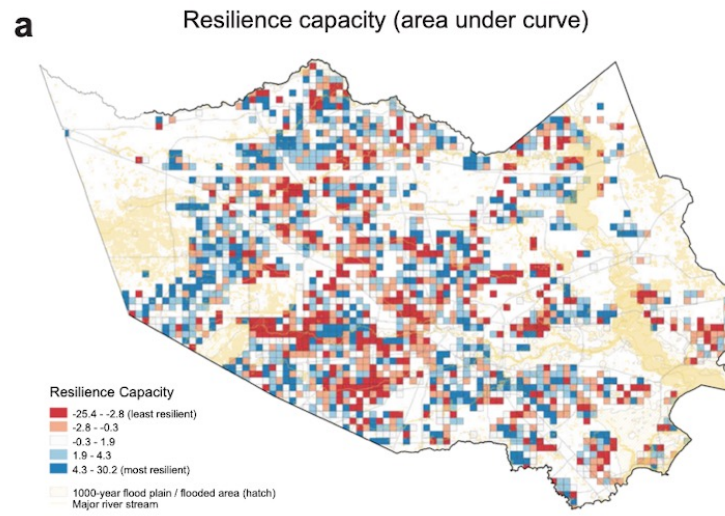
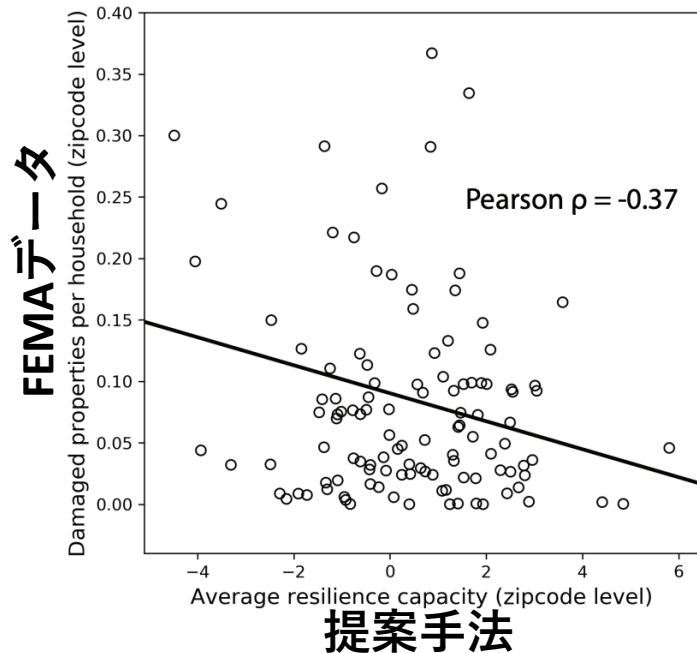
浸水リスク高

地形的リスクは2と
それほど変わらない

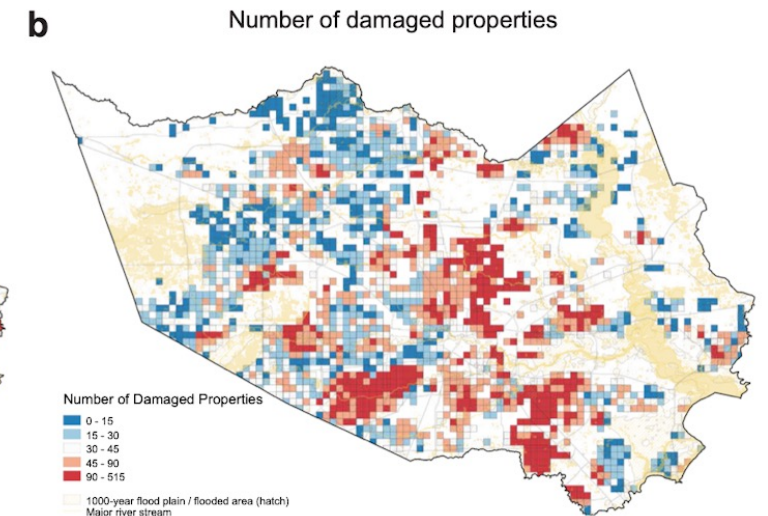
Validation: 損壊被害データとの比較

Federal Emergency Management Agencyの損壊被害のデータと、AUC値を比較
統計的に有意な負の相関がある（左プロット）

But...低収入世帯の方が(今回の場合Group2など)、政府の援助を欲するために
報告しやすい可能性 = self-selection bias. FEMAデータを盲目的に信じれば言い訳でもない。



提案手法



FEMAデータ

*赤がleast resilient, 青がmost resilientなのでこれまでと色が逆転していることには注意

3. 移動パターン分析

■ 行動パターンの把握

- 対象期間中を通してHoustonに在住
- ハリケーン襲来前に「自宅」セルで生活

するサンプルのうち、最もアクティブな25%を抽出 (51,020機器)

→これらの機器に対し、対象地域を米国全域に広げ分析

DBSCANの利点：

クラスターの数を与えなくて良い
ノイズに頑健 (クラスターに所属しない点)

DBSCAN clustering algorithm³⁾

活動場所を特定

p の近隣は次のように定める

$$N_\varepsilon(p) = \{q \in C \mid \text{dist}(p, q) \leq \varepsilon\} \wedge |C| \geq n$$

C : クラスター

n : クラスターの最小ポイント数

$\text{dist}(p, q)$: p と q のユークリッド距離

指標は空間距離と時間

$\varepsilon = 0.25$ つまり、250m, 6時間以内の点群をクラスターとみなす

■ 行動パターンの把握

「家」の設定：k-means clustering

クラスターの平均を使って、 k 個に分類する手法
サンプル $x_i (= x_1, \dots, x_n)$ を以下を解くようにクラスターに割り当てる

$$\sum_{i=0} \min_{\mu_j \in C} (\|x_i - \mu_j\|^2)$$

μ_j : j 番目のクラスターの平均

DBSCANで得た点群を時間ごとの活動頻度で4つに分類

住居(朝夜に観測多), 職場(昼間に観測多), CBD, ミックス

*これらの分類は、GitHubに公開されている
本研究のコードに記載されていました

住居に分類された点群の中からハリケーン襲来前の最頻点を0,
同様にハリケーン後の居住点を求めDとし, OD表を作って分析する

■ 前提

Houston 政府は避難指示を出さなかった

- ▶ 各世帯はそれぞれのリスク許容度, 経済力, 住居事情に応じて行動を決定
- ▶ 避難世帯は
 - i. 独自の判断で事前に避難
 - ii. 被害状況からやむを得ず避難

Table 3 Disparities in disaster response and evacuation patterns across the four cluster groups.

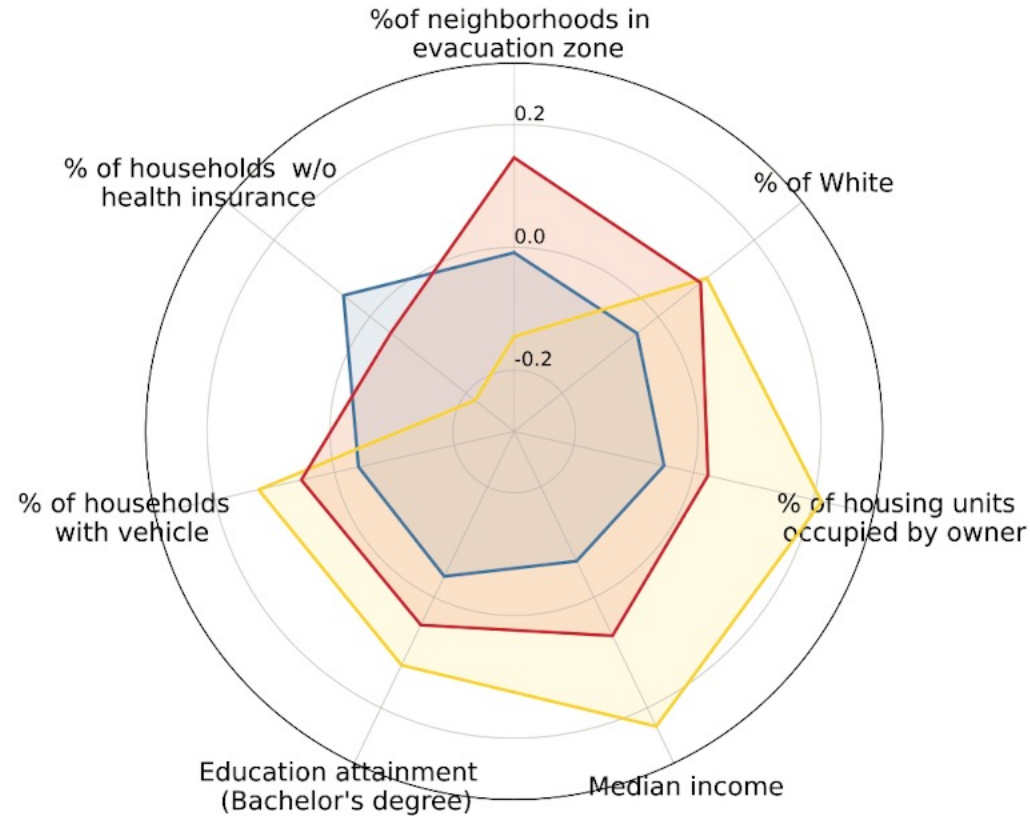
Origin group	Sample size	Disaster response			Where evacuees go		
		Stay home	Evacuate	Inactive (No ping)	Harris County	Texas (outside of Harris County)	US (outside of Texas)
Shelter-in-place	9953	8493 (85.33%)	837 (8.41%)	623 (6.26%)	552 (65.95%)	133 (15.89%)	152 (18.16%)
Stable	16,506	13,936 (84.43%)	1506 (9.12%)	1064 (6.45%)	1008 (66.93%)	220 (14.61%)	278 (18.46%)
Distressed	4187	3338 (79.72%)	559 (13.35%)	290 (6.93%)	381 (68.16%)	57 (10.20%)	121 (21.65%)
Abandoned	1960	1320 (67.35%)	438 (22.35%)	2020 (10.31%)	343 (78.31%)	49 (11.19%)	46 (10.50%)

*202の誤植

This table summarizes evacuation destinations based on residents' origin (home neighborhood) associated with neighborhood activity clusters.

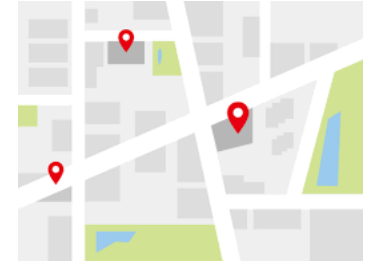
避難者のうちでは、避難者のうちでは、
脆弱な地区に住む人の方が、Resilientな地区に住む人の方が
Harris Countyにとどまり外に出ている

3.2 Results



- 位置データの活用による分析

従来手法は小規模なエリアに対してはリッチな情報を得られるが、**低コスト**で**大規模**な情報を得られる位置データ活用の可能性



- 一般性

ほぼ同じデータ・分析手法により、他の災害に対しても分析ができる

- 地区間の**格差**の存在を確認

同等の地形リスクでも、**人種マイノリティ**が多い地域2では**活動レベル**が大きく低下
地域2は1/4が健康保険に未加入



- 地区ごとのレジリエンスを考慮した意思決定が可能になる

より脆弱な地区にフォーカスした、**公平**な資源配分、避難所配置、避難路整備につながる



Boyeong Hong, Bartosz J. Bonczak, Arpit Gupta and Constantine E. Kontokosta. Measuring inequality in community resilience to natural disasters using large-scale mobility data, *Nature Communications*, Vol. 12, 1870, (2021)

1. <https://www.weather.gov/hgx/hurricaneharvey>
2. Constantine E. Kontokosta and Awais Malik. The Resilience to Emergencies and Disasters Index: Applying big data to benchmark and validate neighborhood resilience capacity, *Sustainable Cities and Society*, Vol. 36, 272-285 (2018).
3. Ester, M., Kriegel, H. P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD*, Vol. 96, 226-231 (1996).

- 基礎分析
- Pre-disaster: spatial distributionが道路にheavily concentrated
During-disaster: significantly lower levels of activity
Post-disaster: まだ完全には戻らない

