

リアルタイムな人口変動を利用した 災害規模推定における特徴量の検討

落合 桂一^{1,a)} 赤塚 裕人¹ 山田 渉¹ 寺田 雅之¹

概要: 台風や大雨などの自然災害が激甚化している。災害規模が大きくなるほど災害規模の把握に時間を要し、初動対応が遅れる可能性がある。そこで本研究では、リアルタイムな人口変動のデータを利用し災害規模を推定することを提案し、推定に有用な特徴量の検討を行う。

1. はじめに

気候変動により台風やゲリラ豪雨などの気象災害が世界各地で以前より頻繁に発生し、重大な被害が発生している [1], [2]. 日本では 2019 年に台風 15 号および 19 号が東日本に上陸し、甚大な被害をもたらした。台風 15 号では千葉県を中心に大規模な停電が発生したり^{*1}, 台風 19 号では多摩川や千曲川など少なくとも 71 の河川で堤防が決壊し浸水被害が発生するなど大きな被害があった^{*2}. このような状況下では、国や自治体は早急に災害への対応を行うことが求められる。しかしながら、被害規模が大きくなるほど被害状況の把握が困難となり、結果として対応が遅れる可能性がある。そのため、自然災害の発生初期において被害規模を推定することが重要である。

災害発生時の被害状況の把握の困難さについて、2019 年の台風 15 号を実例として説明する。図 1 (a) は台風 15 号の進路を示す。図中の赤い点が台風の中心を示す。台風 15 号は 2019 年 9 月 9 日午前 3 時前に三浦半島に上陸し、東京湾を通過し、午前 5 時頃に千葉県付近に上陸した。図 1 (b) と (c) は内閣府が発表^{*3}した東京都、神奈川県、千葉県における住家の被害件数を時系列でプロットした推移である。図 1 (b) では 9 月 9 日から 9 月 16 日のみをプロットし、図 1 (c) は 9 月 9 日から 10 月 12 日までをプロットしている。図 1 (b) において、台風上陸の 3 日後 (赤い実線) では神奈川県の被害が大きく見えるが 1 週間後 (赤い点線) では千葉県の被害が多いことがわかる。さらに台風上陸の 1 ヶ月後のデータ (図 1 (c) の赤い実線) では、千葉県で大

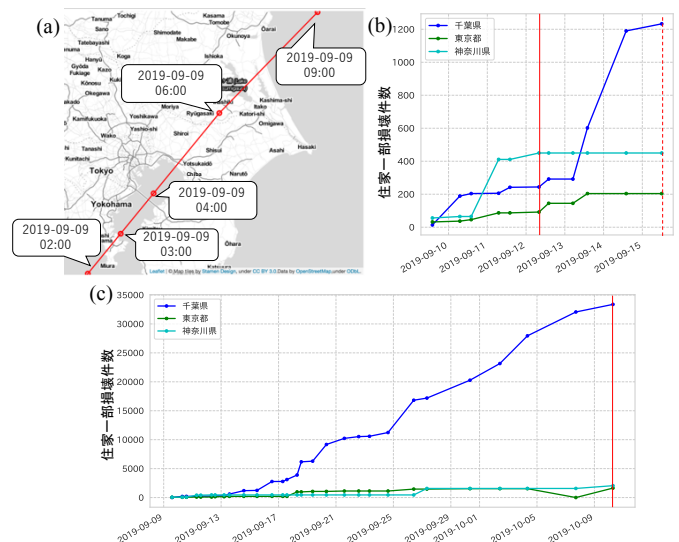


図 1 2019 年台風 15 号における住家の一部損壊件数の推移

規模な被害があったことがわかる。この例は、大規模な被害が発生した場合、災害発生直後に被害規模を把握することの困難さを示している。

災害発生直後に被害規模を把握する方法として、衛星写真を活用する方法が考えられる [3]. しかしながら、衛星写真では時間解像度が限られたり、雲による遮蔽などの影響で正確に撮影できない場合がある。別の方法として、ソーシャルメディアや位置情報などの人の行動データを利用する方法が考えられる。いくつかの研究でこれらの可能性が研究されている [1], [4], [5] しかしながら、ソーシャルメディアを活用する場合は情報の信頼性が課題になったり [5], [6], GPS の位置情報を活用する場合は、特に郊外においてエリアのカバレッジが課題となる [4].

そこで本研究では、携帯電話ネットワークの運用データ

¹ 株式会社 NTT ドコモ, NTT DOCOMO, INC.

^{a)} ochiaike@nttdocomo.com

^{*1} <https://www.nikkei.com/article/DGXMZO49558350Y9A900C1000000/>

^{*2} <https://jp.weathernews.com/news/29409/>

^{*3} <http://www.bousai.go.jp/updates/r1typhoon15/index.html>

から生成した人口データであるモバイル空間統計^④[7]を活用することを検討する。特にリアルタイムな人口を反映した国内人口分布統計（リアルタイム版）（以下、リアルタイム人口統計）を利用して、災害の被害規模を推定するときに有用な特徴量の検討を行う。本論文で扱うリアルタイム人口統計データは、1時間前の人口分布を10分ごとに、500mメッシュ単位に推定した人口分布のデータである。携帯電話ネットワークを元に行っているため、地理的なカバレッジが広く、また10分ごとにデータがあるため時間解像度も高いため、本研究で取り組むタスクに適したデータであると考えられる。本研究では、災害発生時の人々の行動に関していくつかの仮説に基づき特徴量を設計する。そして、市区町村ごとの被害規模の大小を推定するモデルを構築し特徴量の有効性を評価する。

本研究の貢献は以下の通りである。

- 災害発生時の人口データを利用することで各地域の被害状況を推定する災害被害規模推定問題を提案した。
- 被害規模を推定するための特徴量を検討し、各特徴量の有効性を評価した。

本稿の構成は以下の通りである。次章で災害被害規模推定問題と利用するデータについて説明する。次に、3章で推定に利用する特徴量を説明し、4章では実データを用いた評価を行い、特徴量の有効性を検証する。5章で関連研究を概観し、最後に6章にて本研究のまとめおよび今後の課題について述べる。

2. 問題設定とデータ

2.1 定義

本節では、用語定義と問題設定について説明する。

定義1（リアルタイム人口統計） 市区町村 c のタイムスタンプ t におけるリアルタイム人口統計を $pop_{c,t,a}$ とする。ここで人口種別 a は対象とする市区町村 c の居住者、非居住者、その両者の和である総人口のいずれかを示す値である。また、 c におけるリアルタイム人口統計系列を $x_c = \{pop_{c,t_0,a}, pop_{c,t_1,a}, \dots, pop_{c,t_s,a}\}$ とし、 s は x_c の長さを示す。

定義2（災害被害レベル） 市区町村 c の被害規模は住家の被害件数により決定する。災害救助法^{*5}や被災者生活再建支援法^{*6}などの災害関連の法律では、全壊、半壊などの件数から法律の適用が行われる場合がある。そこで、これらの法律の住家被害件数のカウント方法を参考に以下の式で被害件数 $total_damage_c$ をカウントする。

$$total_damage_c = num_{complete} + \frac{1}{2}num_{half} + \frac{1}{3}num_{flood}$$

ここで、 $num_{complete}$ は全壊、 num_{half} は半壊、 num_{flood}

*4 「モバイル空間統計」は株式会社 NTT ドコモの登録商標です。

*5 <http://www.bousai.go.jp/taisaku/kyuujo/kyuujo.html>

*6 <http://www.bousai.go.jp/taisaku/seikatsusaiken/shiensya.html>

表1 リアルタイム人口統計データの例

Timestamp	Grid code	Residential city	Estimated Population
2019-09-09 00:00	533946113	千代田区	20,000
2019-09-09 00:00	533946113	千葉市中央区	2,000
2019-09-09 00:00	533946113	横浜市西区	5,000

は床上浸水の件数を示す。本研究では被害レベル y_c を大小を2値で推定することとし、 $total_damage_c$ が閾値以上の場合に $y_c = 1$ 、そうではない場合 $y_c = 0$ とする。

（問題定義） リアルタイム人口統計系列 x_c が与えられたとき、市区町村 c の被害レベル y_c を推定する問題を災害被害規模推定問題と定義する。

2.2 データ

本研究では2種類のデータ（リアルタイム人口統計データおよび災害被害データ）を利用する。対象とする災害は2019年の台風15号および19号の2つ、分析対象エリアは東京都（島嶼部は大島のみ含む）、神奈川県、千葉県とし、対象の市区町村数は171である。

リアルタイム人口統計データ：リアルタイム人口統計データはNTTドコモの携帯電話ネットワークの運用データを利用して推定したモバイル空間統計と呼ばれる人口データであり、500mメッシュごとに10分間隔で1時間前までの人口を推定している。詳細は文献[7]を参照されたい。リアルタイム人口統計データでは5歳刻みの年齢、性別、居住地という属性情報が利用できる。本研究では年齢、性別は集約し、居住地に関して居住者と非居住者という2つの区分で利用する。ユーザのプライバシー保護のため、NTTドコモではガイドラインを公開している^{*7}。モバイル空間統計の人口データは集計データであり、個人の位置の軌跡は利用できないため、特徴量の検討においても集計した人口の変動から作成できる特徴量を考える必要がある。表1にリアルタイム人口統計データの例を示す。地域メッシュコード533946113は千代田区を示すため、居住者人口は20,000、非居住者人口は7,000と計算できる。リアルタイム人口統計データは時系列データのため、カルマンフィルタによる成分分解が可能である。トレンド成分および季節成分は定常的な要素とみなすことができ、この定常成分で説明できない成分をイベント成分とする。本研究では、リアルタイム人口統計を定常成分と非定常成分（イベント成分）に分解したデータも利用する。成分分解の手法については文献[8]を参照されたい。

災害被害データ：住家被害件数のデータは分析対象の各都府県の公式ページから取得した^{*8*9*10}。このデータから定義

*7 https://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/guideline/index.html

*8 <https://www.bousai.metro.tokyo.lg.jp/taisaku/saigai/>

*9 <http://www.bousai.pref.chiba.lg.jp/>

*10 <https://www.bousai.pref.kanagawa.jp/>

表 2 被害レベルの人口別の閾値

人口	閾値
100,000 以上	10
50,000 以上 100,000 未満	5
50,000 未満	2

2に記載の式で被害件数 $total_damage_e$ を計算し被害レベルを判定する。被害レベルの閾値は定義2に記載の法律を参考に表2の通りとした。表2の人口はリアルタイム人口統計ではなく国勢調査の人口である。本研究では2015年の国勢調査のデータを利用した。

3. 災害被害レベル推定

3.1 特徴量の検討

本研究では、災害発生後の人々の動きとして以下の2つの仮説を立てる。そして、これらの仮説を表す特徴量を生成し推定精度を算出することで特徴量の有効性を検証する。
仮説1: 被害が大きいエリアでは、自宅の補修や交通機関が止まる影響で自宅付近に留まる人が多く居住者人口が増加し、反対に非居住者は減少する。

仮説2: 居住者/非居住者問わず、定常的にある場所を訪問する人数が減少する一方、非定常的な人数が増加する。

図2に特徴量生成の流れを示す。仮説1に対応する特徴量として、平常時の人口と災害後の人口の比率（特徴量1）および居住者と非居住者の人口比率（特徴量2）が考えられる。仮説2については、カルマンフィルタで成分分解した定常成分とリアルタイム人口統計の比率（特徴量3）が考えられる。また、特徴量1~3について、平常時の値からの異常度合いも特徴量として考えられる。本研究ではホテリング理論[9]に基づき、以下の手順で異常度スコアを計算する。

Step 1: リアルタイム人口統計または人口比率の平均 $\hat{\mu}_{c,a,d,h}$ と分散 $\hat{\sigma}_{c,a,d,h}$ を各市区町村 c 、人口種別 a 、曜日種別 d 、時間帯 h ごとに計算する。ここで曜日種別 d はタイムスタンプ t が平日か休日かを示す値、時間帯 h は t の1時間ごとの時間帯を示す。

Step 2: 異常度スコア $anom(pop_{c,a,t}) = \frac{(pop_{c,a,t} - \hat{\mu}_{c,a,d,h})^2}{\hat{\sigma}_{c,a,d,h}}$ を計算する。

加えて、上記の特徴量を所定期間の所定のタイムスロットごとに計算するため、人口比率および異常度スコアの基本統計量（最大、最小、平均、中央値、標準偏差）を特徴量とする。最終的に、これらの特徴量をすべて連結したベクトルを各市区町村に対する特徴量として利用する。

3.2 推定モデル

本研究では市区町村のつながりをグラフ構造として捉えた Graph Convolutional Networks (GCN) モデルと、グラフ構造は考慮しないモデル (XGBoost[10] および Light-

特徴量1: 平常時の人口と災害後の人口の比率
特徴量2: 居住者と非居住者の人口比率
特徴量3: 定常成分とリアルタイム人口の比率

X

方法1: 人口比率
方法2: 異常度スコア
方法3: 方法1および2の基本統計量

市区町村別の
特徴量 (98次元)

図 2 特徴量生成の流れ

GBM[11]) で推定を行う。2つのモデルを利用する理由は、GCN モデルでは隣接市町村の情報も加味できるため推定精度の向上が期待できるが、特徴量の重要度に関する情報を取得できないため、XGBoost および LightGBM も利用し有用な特徴量の検討を行う。なお、XGBoost および LightGBM については、特徴量の工夫を行なったため、モデルの詳細は文献 [10], [11] を参照されたい。ここでは GCN モデルについて説明する。

ノード特徴量 x_v を持つ市区町村ノード $v \in V$ および市区町村間の移動を示すエッジ $e \in E$ に基づく市区町村グラフ $G = (V, E)$ を考える。GCN モデルでは一般にグラフを無向グラフとして扱い、各ノードの特徴表現を学習する。GCN モデルでは反復的に周辺ノードの特徴量を集約し対象ノードの特徴表現を更新する。 k 層の GCN モデルでは k 回の反復計算を行い、対象ノードの k ホップ先までのノード特徴量を集約する。

$$a_v^k = AGGREGATE_k(\{h_u^{k-1}, u \in N(v)\}) \quad (1)$$

$$h_v^k = COMBINE_k(h_v^{k-1}, a_v^k) \quad (2)$$

ここで $h_v^0 = x_v$ 、 h_v^k は k 層の集約後のノード v の特徴表現を示し、 $N(v)$ はノード v の隣接ノードを示す。表記は文献 [12], [13] に従う。Kipf らの GCN モデル [14] では AGGREGATE および COMBINE の操作を統合し以下の式で計算する。

$$h_v^k = ReLU(W \cdot MEAN\{h_u^{k-1}, u \in N(v) \cup \{v\}\}) \quad (3)$$

ここで W は学習可能なパラメータ行列、MEAN は要素ごとの平均を示す。本研究では、人の移動には向きがあるためグラフを有向グラフとして扱い、ノード間の重みに移動人数を設定グラフを構築する。エッジの重みを考慮するため、式 (3) の MEAN をエッジの重みを加味した重み付き平均とする。

4. 評価

本章では、提案する特徴量の有効性を評価するため以下の観点で検証を行う。

- 評価1: 特徴量1~3のそれぞれの有効性を評価するため特徴量を1つ抜いた状態で精度を比較する。
- 評価2: 災害発生後の経過時間と精度の関係を調査するため、災害発生からの経過時間を変えて精度を比較する。

表 3 特徴量を 1 つ抜きで評価した結果 (適合率)

モデル	特徴量 1+2+3	特徴量 1+2	特徴量 1+3	特徴量 2+3
XGBoost	0.6814	0.6774	0.692	0.6367
LightGBM	0.7060	0.6800	0.7437	0.7432
GCN	0.7693	0.6211	0.6774	0.7639

表 4 特徴量を 1 つ抜きで評価した結果 (再現率)

モデル	特徴量 1+2+3	特徴量 1+2	特徴量 1+3	特徴量 2+3
XGBoost	0.6173	0.5913	0.6112	0.5766
LightGBM	0.5956	0.5974	0.6298	0.6004
GCN	0.7026	0.6086	0.6401	0.7421

4.1 データ

対象とする災害は 2019 年の台風 15 号および 19 号とし、ラベル 1 (被害大) が 27.2%, ラベル 0 (被害小) が 72.8% である。個別の台風におけるラベルの比率は、ラベル 1 (被害大) が台風 15 号では 31.6%, 台風 19 号では 22.8% である。平常時との人口比率や異常スコアに利用する平均と分散は各台風が通過する直前 1 ヶ月分のデータから計算した。特徴量の計算には 24 時間分の人口データを利用し、計算量削減のため 3 時間ごとのデータを利用した。

4.2 評価条件

評価は 5 分割交差検証で行い、評価指標の平均値を計算した。具体的には 171 の市区町村を 5 分割し、4 つのグループで学習を行い残りの 1 グループで評価した。XGBoost および LightGBM については Optuna[15] を利用してハイパーパラメータのチューニングを行なった。

4.3 評価結果

評価 1: 表 3 および表 4 に特徴量を全て利用した場合と 1 つ抜いて学習した場合の適合率と再現率を示す。XGBoost および LightGBM では全ての特徴量を利用するより特徴量 1 と 3 を利用する場合の方が精度が高い結果となった。一方 GCN モデルでは、適合率は全ての特徴量を利用する場合が最も高く、再現率は特徴量 2 と 3 のみを利用した場合の方が高かった。表 5 にすべての特徴量を利用した場合の XGBoost および LightGBM の特徴量重要度の上位 10 件を示す。表 5 の結果から 0 時や 24 時の特徴量が多く上位に入っていることから、深夜帯の人口変動が災害規模の推定に有効であると考えられる。また、人口比率よりも異常度スコアに基づく特徴量が多いことから、異常度スコアが有効であると考えられる。さらに、表 3 および表 4 の結果から特徴量 2 を除いた場合に精度が高くなっていることと、表 5 の結果から、特徴量 2 (居住者と非居住者の比率) はあまり有効な特徴量ではないことがわかる。

評価 2: 評価 2 では推定に利用するデータを台風通過直後、1 日後、2 日後と変化させたときの推定精度を比較する。表 6 および表 7 に適合率、再現率のそれぞれの結果を示す。全体的な傾向としては、時間経過とともに適合率は上昇し、

再現率は減少している。この要因としては、災害の被害規模が大きいほど平常状態に戻るために時間を要するため、時間が経過しても特異な特徴量を示す市区町村は被害規模が大きいと判定されている可能性が考えられる。

5. 関連研究

文献 [16] のサーベイ論文によると、災害対応には衛星写真 [3] やソーシャルメディア [5], [6] のデータが頻繁に利用される。加えて、スマートフォンの GPS[4], [17] や携帯電話ネットワークのデータに基づくモビリティデータも利用される [18], [19], [20]。ここでは、本研究に関連するモビリティデータについて関連研究を述べる。Song らは東日本大震災における人々の移動の分析を行い、隠れマルコフモデルに基づく移動の予測を行なった [4]。Yabe らは Web 検索ログから避難行動の意思決定を予測するモデルを提案している。このモデルでは、スマートフォンの位置情報に対して異常検知を行い避難を実施したかどうかという教師ラベルを自動で生成している [21]。また、Yabe らは東日本大震災や熊本地震を対象に災害後の人口減少率のモデル化を行い、長期的な人口減少に影響している要素を分析した [22]。

6. おわりに

本研究では、リアルタイム人口統計データを利用して災害発生時の各地域の被害状況を推定する災害被害規模推定問題を提案した。被害規模を推定するための特徴量を複数パターン検討し、各特徴量の有効性を評価した。その結果、平常時との人口比率、時系列の成分分解を行い定常成分とリアルタイム人口統計の比率などが特徴量として有用であることを確認した。また、人口比率よりも人口比率の異常度スコアが有用であること、深夜帯の人口の特徴量が有効であることがわかった。今後の課題として、災害事例を増やし手法や特徴量の一般性を検証すること、地理的な解像度を向上させ詳細な被害規模を推定することなどが挙げられる。

参考文献

- [1] Kryvasheyev, Y., Chen, H., Obradovich, N., Moro, E., Van Hentenryck, P., Fowler, J. and Cebrian, M.: Rapid assessment of disaster damage using social media activity, *Science advances*, Vol. 2, No. 3, p. e1500779 (2016).
- [2] First, P. J.: Global Warming of 1.5 °C An IPCC Special Report on the Impacts of Global Warming of 1.5 °C Above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, *Sustainable Development, and Efforts to Eradicate Poverty*. <https://www.ipcc.ch/sr15/>. Accessed, Vol. 1 (2019).
- [3] Panteras, G. and Cervone, G.: Enhancing the temporal resolution of satellite-based flood extent generation

表 5 特徴量重要度

順位	XGBoost		LightGBM	
	特徴量	重要度	特徴量	重要度
1	異常度スコア/特徴量 1/総人口/24 時	0.127256	異常度スコア/特徴量 1/総人口/24 時	127.32943
2	異常度スコア/特徴量 1/総人口/0 時	0.066743	異常度スコア/特徴量 3/総人口/18 時	33.094078
3	異常度スコア/特徴量 3/居住者最小値	0.066516	異常度スコア/特徴量 1/非居住者/0 時	30.099055
4	異常度スコア/特徴量 3/総人口/18 時	0.031215	異常度スコア/特徴量 1/総人口最小値	28.582007
5	異常度スコア/特徴量 1/総人口最小値	0.023499	異常度スコア/特徴量 3/非居住者/24 時	16.68695
6	異常度スコア/特徴量 3/居住者/3 時	0.018997	人口比率/特徴量 3/総人口平均値	15.856616
7	異常度スコア/特徴量 1/非居住者/0 時	0.016271	異常度スコア/特徴量 1/総人口/6 時	14.76719
8	人口比率/特徴量 1/居住者/21 時	0.016065	異常度スコア/特徴量 3/総人口/21 時	13.448712
9	異常度スコア/特徴量 3/総人口/15 時	0.015973	人口比率/特徴量 1/居住者標準偏差	11.65774
10	異常度スコア/特徴量 1/居住者/3 時	0.01494	異常度スコア/特徴量 3/非居住者最小値	11.57418

表 6 入力データの時間を変化させた評価結果 (適合率)

モデル	直後	1 日後	2 日後
XGBoost	0.6579	0.6814	0.6548
LightGBM	0.6777	0.706	0.7297
GCN	0.7286	0.7693	0.7739

表 7 入力データの時間を変化させた評価結果 (再現率)

モデル	直後	1 日後	2 日後
XGBoost	0.6373	0.6173	0.6005
LightGBM	0.602	0.5956	0.595
GCN	0.7158	0.7026	0.6448

using crowdsourced data for disaster monitoring, *International Journal of Remote Sensing*, Vol. 39, No. 5, pp. 1459–1474 (2018).

[4] Song, X., Zhang, Q., Sekimoto, Y. and Shibasaki, R.: Prediction of Human Emergency Behavior and Their Mobility Following Large-Scale Disaster, *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pp. 5–14 (2014).

[5] Thomson, R., Ito, N., Suda, H., Lin, F., Liu, Y., Hayasaka, R., Isochi, R. and Wang, Z.: Trusting tweets: The Fukushima disaster and information source credibility on Twitter, *Proceedings of the 9th International ISCRAM Conference*, pp. 1–10 (2012).

[6] Nadamoto, A., Miyabe, M. and Aramaki, E.: Analysis of Microblog Rumors and Correction Texts for Disaster Situations, *Proceedings of International Conference on Information Integration and Web-Based Applications and Services*, IIWAS '13, Association for Computing Machinery, pp. 44–52 (2013).

[7] Terada, M., Nagata, T. and Kobayashi, M.: Population estimation technology for mobile spatial statistics, *NTT DOCOMO Technical Journal*, Vol. 14, No. 3 (2013).

[8] Akatsuka, H. and Terada, M.: Application of Kalman Filter to Large-Scale Geospatial Data, *International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL 2020)* (2020).

[9] Hotelling, H. et al.: A generalized T test and measure of multivariate dispersion, *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, The Regents of the University of California (1951).

[10] Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Dis-*

covery and Data Mining, KDD '16, pp. 785–794 (2016).

[11] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T.-Y.: Lightgbm: A highly efficient gradient boosting decision tree, *Advances in neural information processing systems*, pp. 3146–3154 (2017).

[12] Hamilton, W. L., Ying, R. and Leskovec, J.: Inductive Representation Learning on Large Graphs, *NIPS* (2017).

[13] Xu, K., Hu, W., Leskovec, J. and Jegelka, S.: How Powerful are Graph Neural Networks?, *International Conference on Learning Representations (ICLR)* (2019).

[14] Kipf, T. N. and Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks, *International Conference on Learning Representations (ICLR)* (2017).

[15] Akiba, T., Sano, S., Yanase, T., Ohta, T. and Koyama, M.: Optuna: A Next-Generation Hyperparameter Optimization Framework, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '19, pp. 2623–2631 (2019).

[16] Yu, M., Yang, C. and Li, Y.: Big data in natural disaster management: a review, *Geosciences*, Vol. 8, No. 5, p. 165 (2018).

[17] Yabe, T., Ukkusuri, S. V. and Rao, P. S. C.: Mobile phone data reveals the importance of pre-disaster inter-city social ties for recovery after Hurricane Maria, *Applied Network Science*, Vol. 4, No. 1, p. 98 (2019).

[18] Bagrow, J. P., Wang, D. and Barabasi, A.-L.: Collective response of human populations to large-scale emergencies, *PLoS one*, Vol. 6, No. 3, p. e17680 (2011).

[19] Lu, X., Bengtsson, L. and Holme, P.: Predictability of population displacement after the 2010 Haiti earthquake, *Proceedings of the National Academy of Sciences*, Vol. 109, No. 29, pp. 11576–11581 (2012).

[20] Becker, R., Cáceres, R., Hanson, K., Isaacman, S., Loh, J. M., Martonosi, M., Rowland, J., Urbanek, S., Varshavsky, A. and Volinsky, C.: Human mobility characterization from cellular network data, *Communications of the ACM*, Vol. 56, No. 1, pp. 74–82 (2013).

[21] Yabe, T., Tsubouchi, K., Shimizu, T., Sekimoto, Y. and Ukkusuri, S. V.: Predicting Evacuation Decisions Using Representations of Individuals' Pre-Disaster Web Search Behavior, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '19, pp. 2707–2717 (2019).

[22] Yabe, T., Tsubouchi, K., Fujiwara, N., Sekimoto, Y. and Ukkusuri, S. V.: Understanding post-disaster population recovery patterns, *Journal of the Royal Society Interface*, Vol. 17, No. 163, p. 20190532 (2020).