

2011年東日本大震災に基づく 津波被害モデルの構築と被害予測

小川 芳樹¹・秋山 祐樹²・柴崎 亮介³

¹非会員 大学院生 東京大学大学院新領域創成科学研究科
(〒153-8505 東京都目黒区駒場4-6-1 東京大学生産技術研究所Cw-503号室)
E-mail:ogawa@csis.u-tokyo.ac.jp

²非会員 東京大学助教 東京大学地球観測データ統融合連携研究機構
(〒153-8505 東京都目黒区駒場4-6-1 東京大学生産技術研究所Cw-503号室)
E-mail:aki@iis.u-tokyo.ac.jp

³正会員 東京大学教授 東京大学空間情報科学研究センター
(〒277-8568 千葉県柏市柏の葉5-1-5 総合研究棟4階414号室)
E-mail:shiba@csis.u-tokyo.ac.jp

将来予測される巨大地震における津波被害軽減のためには、現在予測される地震津波における被害を定量的に評価・比較しておく必要がある。そこで本研究では、東日本大震災時における東日本大震災津波被災市街地復興支援調査(国土交通省都市局、東京大学CSIS)から得られる1棟1棟の建物被害データ(青森・岩手・宮城・福島・茨城・千葉)を統計的に解析し、建物構造や津波の浸水深などとの被害関係を評価することで、建物被害の津波被害モデルの構築を試みると共に考察を行った。その際、本研究では機械学習アルゴリズムの一つであるランダムフォレスト及び累積正規分布を用いてそれぞれモデルの構築をした。こうした大規模な調査データを用いた津波被害モデルの構築は前例がなく、同モデルを用いることで津波被害予測をこれまでよりもより高精度に実施可能なモデルとなりうる。また開発した津波被害関数と全国の建物データを用いることで日本全国において任意の浸水データを入力し、被害予測を行った。

Key Words : machine learning, cumulative normal distribution, damage estimation, micro geo data

1. はじめに

近い将来、日本では南海トラフ沖を震源とした東海・東南海・南海地震による巨大津波が予測されている。加えて2014年8月には、日本政府の調査員会が日本海における津波の想定高さを初めて発表をした。¹⁾このように津波リスクの高い日本で予測されている津波に対して、建物被害を定量的に推定し、物理的なリスクを事前に評価することは、国土や地域防災政策を考える上で重要である。また企業においてもBCP(Business Continuity Plan)を策定する際にどの程度の被害が予測されるかを把握しておくことは必要不可欠である。このようなことからハード面とソフト面を含めた施策のためには、より高い信頼

性をもった津波被害推定が求められる。

津波による家屋被害を推計するには津波被害関数と呼ばれる浸水深や流速などの外力と被害率の関係を定式化したものを用いるのや指標化したものがある。多くの被害想定^{2,3)}で用いられている被害評価手法に、首藤⁴⁾の津波強度指標があり、例えば浸水深が2m以上で木造家屋については全壊という家屋がどの程度持ちこたえられるかを評価している。また越村ほか⁵⁾は、2004年スマトラ島沖地震津波による家屋被害を浸水深や流速を説明変数とし、確率的に倒壊率と死者率を評価する津波被害関数の構築をしているが、日本と建物の耐震性の違いなどの観点から日本における適用は困難である。

一方、2011年東北地方太平洋沖地震津波においては、

被災後に国土交通省により浸水域を対象に現地調査が行われた。浸水域の全棟調査であるため建物被災状況や浸水深などのデータが豊富であり、従来の衛星画像による判読した被災データ等に比べれば信頼性が高く、被害関数や被害モデルの構築をするに検証を含めて適したデータであるといえよう。

そこで本研究では、2011年東北地方太平洋沖地震津波で浸水域全域の建物を対象として、津波による建物被害推定手法について検討を行う。本研究では(i)回帰分析による津波被害関数(モデル1)と(ii)機械学習による津波被害推定モデル(モデル2)を構築する。(i)については、最大浸水深と家屋被害関係を構造別に回帰分析を行い簡易的な津波被害関数の構築をする。(ii)については、機械学習手法の一つであるランダムフォレスト⁹⁾を用いて津波被害モデルを提案する。ランダムフォレストによる分類は、大きなデータに対しても計算速度が早く学習データのノイズに強いと現地調査データによるノイズにも強いと考えられる。また、分類精度も他の機械学習手法に比べ相対的に高いことが知られている。⁷⁾最後に、構築した津波被害モデルを用いて日本全国を対象として任意の浸水深を与えることで日本全国の建物1棟1棟の建物被害推定をすることで、地域間の相対的なリスク評価が可能なデータベースを作成する。本研究において1棟1棟の構造や築年代を考慮して被害推定することで、任意集計単位に対応でき、自治体のみならず地域コミュニティ単位でのリスク評価が可能となる。

2. 津波被害推定モデルの作成

(1) モデル作成の流れ

本研究で用いる2011年の東北地方太平洋沖地震津波による建物被害状況は国土交通省都市局の「東日本大震災津波被災市街地復興支援調査」をアーカイブ化し東京大学空間情報科学センターが管理運営している「復興支援調査アーカイブ」⁸⁾内のデータを利用する。このデータは浸水域全域(青森県、岩手県、宮城県、福島県、茨城県、千葉県)の各自治体の罹災証明、国土交通省の現地調査と衛星画像により判別した建物ポリゴンデータがシェープファイル形式で限定公開されており、個別家屋の被災状況、浸水深や構造などが含まれる。表-1に本研究で利用した個別建物の属性を示す。図-1に本研究の流れを示す。まず、モデル評価のためにデータを学習データと検証用データに分ける。次いでモデル1:最大浸水深と家屋被害データを用いて建物構造別の津波被害関数とモデル2:機械学習によるモデルを構築する。次に構築したそれぞれのモデルに検証用データを適用し、被害の真値と比較することで検証を行う。

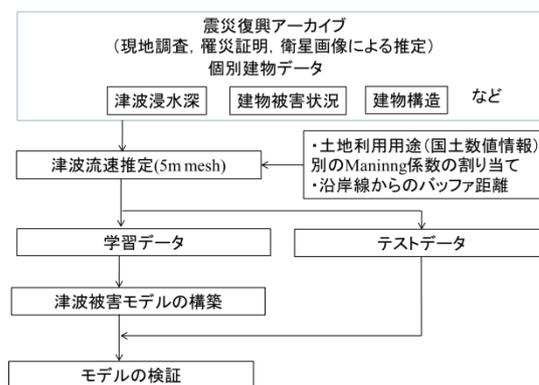


図-1 本研究における被害モデル構築の全体像

(2) 東北地方太平洋沖地震津波における建物被害状況

「復興支援調査アーカイブ」から得られる個別建物の被災状況と浸水深を図-2と3に示す。このように個別建物単位でのデータとなっている。また浸水域と全壊・半壊した棟数を市区町村別に集計した結果を図-4に示す。宮城県石巻市で被害が一番大きくなっており、そこを中心に南北方向にいく連れて被害棟数が小さくなっている。被害関数の構築に際して本研究では、衛星画像により被害の判読を行ったものを除いた信頼性の高い罹災証明と現地調査によるデータかつ表-1にある各属性が満たされているデータを用いるとする。そのため被災した全家屋約22万棟(全壊約12万棟)の内の約10万棟(全壊約5万棟)を本研究の対象建物とする。家屋の構造別の最大浸水深と被害率の関係は、図-5に示す通りである。最大浸水深2mにおいて全壊率は木造で50%、鉄骨造(以下S造)で40%、鉄筋コンクリート造(以下RC造)で25%となっている。

表-1 「復興支援調査アーカイブ」のシェープファイルに含まれる属性について

フィールド名	フィールド内容
市区町村id	id
建物id	id
建物構造	RC造、S造、木造、その他
建物階数	階数
築年数	年代に変換し利用
建物用途	住宅、商業、工業、公益、その他
被災区分	全壊、大規模半壊、半壊、一部損壊、被災なし
浸水深	現地調査によるもの
浸水高	現地調査によるもの
標高	現地調査によるもの

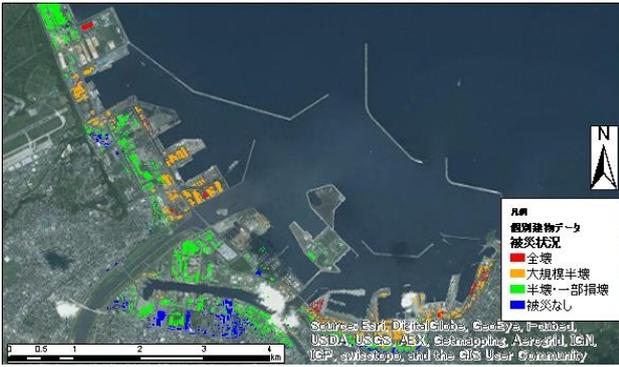


図-2 個別建物の被災状況の例(青森県八戸市)

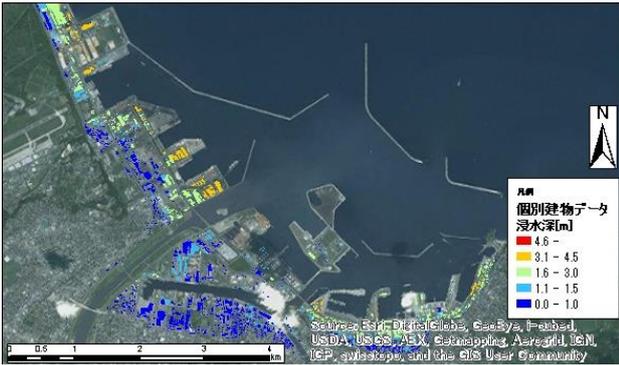


図-3 個別建物の浸水深の例(青森県八戸市)

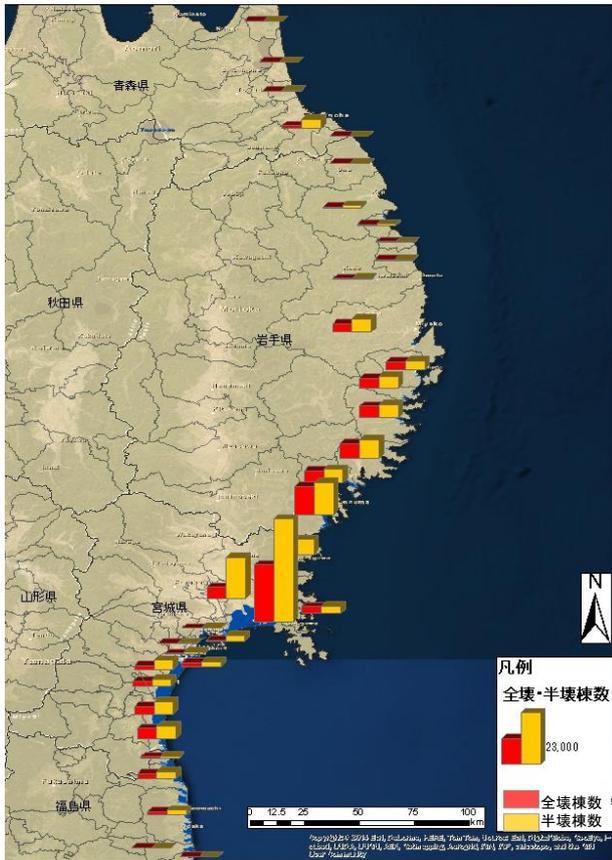


図-4 市区町村別の家屋被災状況

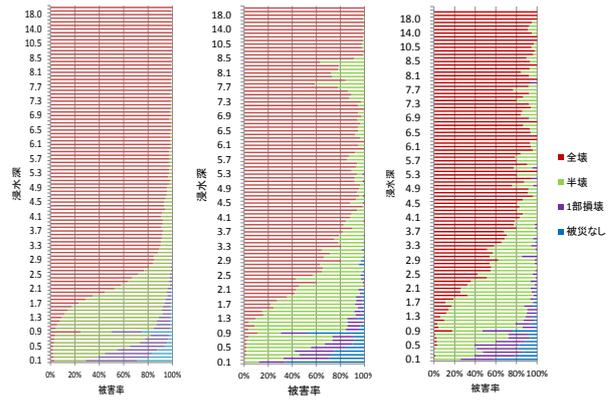


図-5 家屋の構造別被害率のヒストグラム

(3) 津波の流速推定

松富・飯塚⁹⁾の手法を用いて津波の陸上流速を 5m メッシュ(市区町村により 100m メッシュ)で推定し各建物に空間結合する。式(1)に示すように流速 u はフルード数 Fr と浸水深 h_r の関数である。

$$Fr = \frac{u}{(gh_r)^{0.5}} \quad (1)$$

g は重力加速度。いま、浸水深は既知であるからフルード数が推定出来れば流速 $u = u_r$ が推定できる。最大の沿岸津波高が生起したときの波峰から陸上部を考え、この部分の流れは定常的と仮定しエネルギーと質量の保存則を用いることで算出する(式(2)と(3))。

$$z + hr + \frac{u^2}{2g} + f \left(\frac{\delta}{hr} \right) \frac{u^2}{2g} \cong H_0 + \frac{u_0^2}{2g} \quad (2)$$

$$h_r u \cong H_0 u_0 = (h_0 + a_0) u_0 \quad (3)$$

z は標高, f は摩擦損失係数, n は Manning の粗度係数であり, 小谷ら¹⁰⁾を参考に表-1 の示すように土地利用別に設定をする。土地利用データは国土数値情報¹¹⁾からダウンロードできる細密メッシュデータ(100m メッシュ)を用いるとする。 δ は沿岸から対象地点までの距離で沿岸線からのバッファを生成することで各メッシュに割り当てる(図-6)。また h_0 は静水深, u_0 は沿岸津波高生起時の流速である。式(2)と(3)を解くことによりフルード数の評価式として式(4)を得る。フルード数を式(1)に代入することで流速の推定値を得る。

$$Fr \cong \left\{ 2(H_0 - z - h_r) / h_r \left(1 + \frac{2gn^2\delta}{h_r^{3/4}} - \frac{h_0^2}{H_0^2} \right) \right\}^{0.5} \quad (4)$$

図-7 に流速の計算結果を建物に空間結合したものを示す。図-3 の浸水深が同じ場所においても場所によって土地利用や沿岸距離を考慮することで流速が変化しているのがわかる。

表-2 Manningの粗度係数の土地利用別の設定値

土地利用	粗度係数
建物用地 80%-100%	0.08
40-80%	0.06
10-40%	0.04
田・その他農地	0.02
森林	0.03
河川地(水域)	0.025
その他(荒地,公園など)	0.025

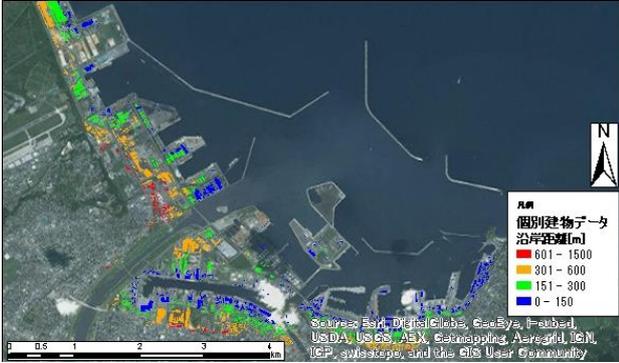


図-6 個別建物の沿岸線からの距離の例(青森県八戸市)

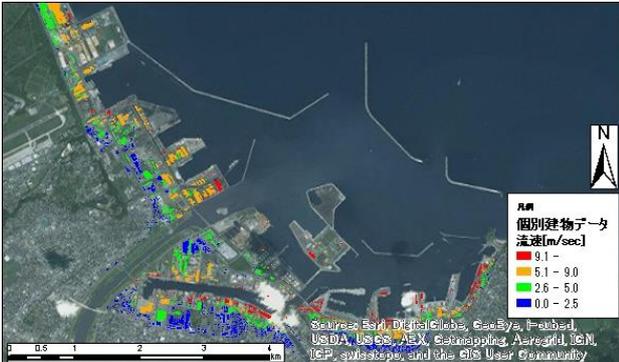


図-7 個別建物の流速計算結果の例(青森県八戸市)

(4) 被害モデル

モデル 1: 累積正規分布とモデル 2: 機械学習による 2つのモデルを提案する。

a) 累積正規分布によるモデル

地震や津波による被害関数は一般的に累積対数正規分布または累積正規分布として定義される(式(5)と(6)).⁵⁾

$$P(x) = \Phi \left[\frac{\ln x - \lambda}{\xi} \right] \quad (5)$$

$$P(x) = \Phi \left[\frac{x - \lambda}{\xi} \right] \quad (6)$$

ここで x は地震動や浸水深等の外力, λ は平均値, ξ は標準偏差である. 村尾・山崎¹²⁾による 1995 年の阪神淡路大震災における地震動による建物被害関数を築年代や構造別に構築したものは現在の被害想定においても広く利用されている. 本研究では全壊と全壊+半壊の 2つの被害関数を構築する. 考慮する変数は既存研究^{4,5)}を参考に浸水深と建物構造とする. 図-5のように浸水深別に被害棟数を集計し, 図-8に示すように累積正規分布の逆関数 ϕ^{-1} を縦軸に最大浸

水深を取り回帰分析することで, パラメータを求める. その際の級数は, ステージェスの公式により決定し, 級数幅はデータ数が同じになるようにする. 累積正規分布か累積対数正規分布かは適合度の高い方を選択するとするが本研究では, 累積正規分布も採用する. それぞれ回帰分析した結果のパラメータを表-3に示す. また, 構築した津波被害関数を図-9に示す. 全壊曲線の S 造と RC 造は最大浸水深が 2.5m 付近から急激に被害率が大きくなり似た傾向を示すのに対し, 木造は 1.5m 付近から被害率が大きくなっている.

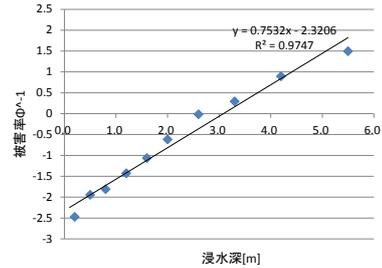


図-8 RC造における最大浸水深と家屋被害率

表-3 家屋被害関数のパラメータ

関数種別	データ	平均 λ	標準偏差 σ	決定係数R ²
全壊	木造	1.8255	0.7595	0.889
	S造	2.1408	0.9243	0.8851
	RC造	2.3206	0.7532	0.9747
全半壊	木造	-0.1248	0.7665	0.9274
	S造	0.1179	0.763	0.8723
	RC造	0.2318	0.9599	0.9161

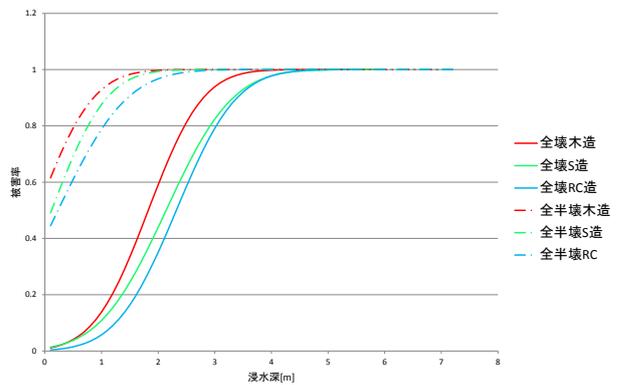


図-9 建物構造別の被害関数

本研究で求めた津波被害関数の検証をするために, 被害関数を用いた推定結果と実被害結果を集計(市区町村)比較することで信頼性検証をする. 全半壊についての結果を図-10に示す. 全壊・全半壊のいずれの構造においてもばらつきはあるものの決定係数が 0.95 以上 ($p < 0.05$)であり高い信頼性であるといえよう.

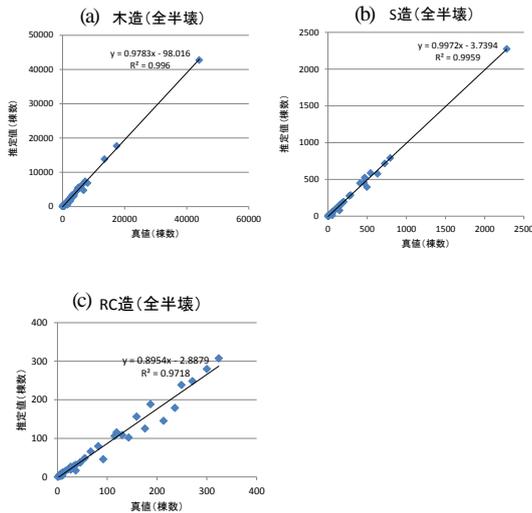


図-10 (a)木造における全半壊の推定値と真値の集計値の比較, (b)S造における全半壊の推定値と真値の集計値の比較, (c)RC造における全半壊の推定値と真値の集計値の比較

b) 機械学習を用いたモデル

次に機械学習の1つであるランダムフォレスト⁹⁾を利用した建物被害推定モデルも構築する。ランダムフォレストを用いるメリットとしては計算速度がSVM¹³⁾などに比べ早いため大規模データに対応でき、ノイズに強い点が挙げられる。ランダムフォレストの詳しいアルゴリズムは引用文献に委ねるが、アルゴリズムの概念イメージは図-11の示す通りである。ランダムフォレストは、以下の手順で学習・判別を行う。

- (i) 重複可能な抽出により n 個のブートストラップ標本を作成する。
- (ii) 標本からランダムに選択された m 個の変数を用いて決定木を作る。
- (iii) 作成した決定木に対して予測データを入れる。
- (iv) 全ての決定木から得られた予測結果の多数決をとり、予測値を決定する。

様々な機械学習手法があるがランダムフォレストは相対的に高い予測精度を示すという報告もある。¹⁴⁾ また本研究では、ランダムフォレストのパラメータとして各決定木の特徴量は、特徴量数の平方根を用いるとする。¹⁵⁾ 決定木の数は、500とし各枝の最小データ数は1とする。一方で、ランダムフォレストはモデル構築に際して、変数の重要度としてGini係数をOOB(out of bag)データによる予測精度から算出し、同時に誤差率も算出される。OOBデータはブートストラップサンプリングに含まれないデータであり、OOBによる誤差率はクロスバリデーションの代わりになる。図-12に学習データより算出した特徴量ごとのGini係数を示す。津波の流速が一番に寄与しており、続いて築年代、浸水深、沿岸からの距離

離の順に大きく影響していることがわかる。しかしながら階数や構造は低いGini係数になっている。これは同じ構造や階数でも被災状況にはばらつきが大きいためであると考えられる。本モデルの精度は、図-12の特徴量全てを利用した場合でOOB誤差率が0.82%であり、特徴量数を5にした場合(流速, 築年代, 浸水深, 沿岸距離, 浸水高)で1.05%である。また築年代データは一般に得られにくいことから、特徴量から築年代を除いた場合(特徴量数を4)においてもOOB誤差率は4.17%である。加えて、特徴量全てを利用した場合にクロスバリデーション($k=10$)を行なった結果の平均誤差率は、2.00%であった(表4)。さらにモデル1(累積正規分布モデル)と同様に集計し、被災別に棟数の比較をした結果、どの被災状況においても $R2 = 1.00$ であり本モデルの方が高い信頼性を示していることが示された。これは数多くの特徴量(変数)を考慮できる点が大いと思われる。

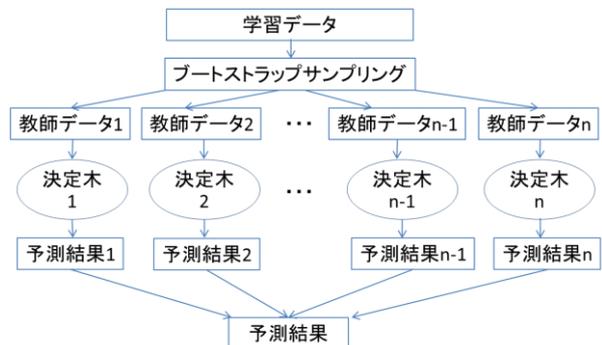


図-11 ランダムフォレストのアルゴリズム

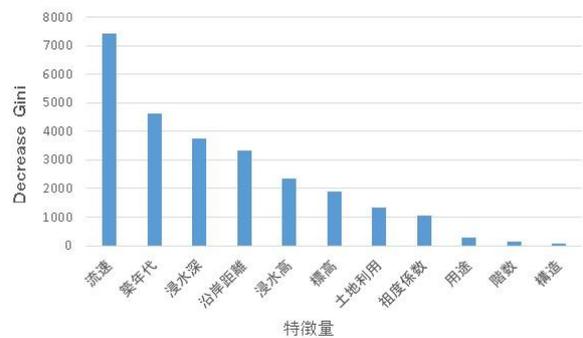


図-12 各特徴量の重要度(Gini係数)

表4 クロスバリデーションによる本モデルの精度

真値	被害状況	推定値				精度
		被災なし	半壊・一部損壊	大規模半壊	全壊	
	被災なし	21692	190	183	96	0.98
	半壊・一部損壊	61	8499	0	3	0.99
	大規模半壊	41	2	12574	8	1.00
	全壊	37	42	32	1540	0.93
総合精度: 0.98						

3. 本モデルの適用

本研究で提案したモデルを将来予測される津波に対して他の地域で適用するには、建物構造や標高が個別建物データで整備されている必要がある。個別建物の建物構造や築年数などは各自治体が保有しており公開している自治体は数少ない。例として世田谷区はWebで建物耐火性能を公開しているが築年数を公開している自治体は見られない。¹⁶⁾そのため本研究ではマイクロジオデータ¹⁷⁾¹⁸⁾を用いるとする。日本全国約6000万棟を対象として建物構造・築年代を住宅地図ベースで整備したマイクロジオデータである。マイクロジオデータは統計値や住宅地図の属性などを組み合わせ、建物単位の構造や築年数などを整備したものである。建物構造や築年代の推定手法・精度についての詳細はOgawa et al.¹⁹⁾に委ねる。この推定データは、最終的に集計することを考えているため1棟1棟か必ずしも正しく推定されていない点に注意する必要があるが集計すれば十分な信頼性があることは確認されている。本研究では、日本全国を対象にこのマイクロジオデータに対して津波被害推定をするために標高データ(10mメッシュDEM(国土地理院))を各建物に空間結合し、そこに任意の浸水データもしくは浸水高を入力することで被害推定を行うことが可能になる。被害推定においては、信頼性の高いランダムフォレストを利用したモデルを採用する。一律に津波高さ3mを与えた場合の愛知県豊橋市における推定結果を図-13に示す。また図-14に西日本における一律に津波高さ3mを与えた場合の4次メッシュ(500m四方)集計における全壊・全半壊率を示す。一律に浸水高を与えたため標高の低い地域で被害が大きくなっている。本来であれば日本全国における津波浸水域のデータが公開されていれば用いることが望ましいがほとんど公開されていないため今後の課題とする。

推定結果は個別建物というマイクロなスケールから都道府県や市区町村という広域なスケールで整備されている。そのため防災政策を議論する際の基礎データとして利用できるだけでなく、津波避難シュミレーションをする際には、どこの建物を避難ビルに選定するのか、どの建物が倒壊するのかという今まで多くは恣意的に決めていたものを定量的に把握することが可能になる。また任意の集計単位でその地域の被害状況を把握することが出来るため、例えば町丁目や100mメッシュなどの小地域単位で現状の把握が出来る。そのため、人口データ等と重ね合わせて現状の避難ビル容量や避難経路幅の評価をすることで避難ビルやタワーの増設などの計画を小地域単位で出来る。

以上のように本研究で推定した成果は、町丁目などの小地域単位で、その地域のリスクを議論し地域防災政策やBCPなどを考える際にも有意義であると考えられる。

今後はデータの見せ方として色々なパラメータを変えて住民に誤解なく分かりやすい見せ方を検討し個人の問題意識として捉えられるような見せ方を提案することで、本研究で整備したデータは住民や自治体におけるリスクコミュニケーションの促進と共に防災意識の向上に寄与するものと期待できる。

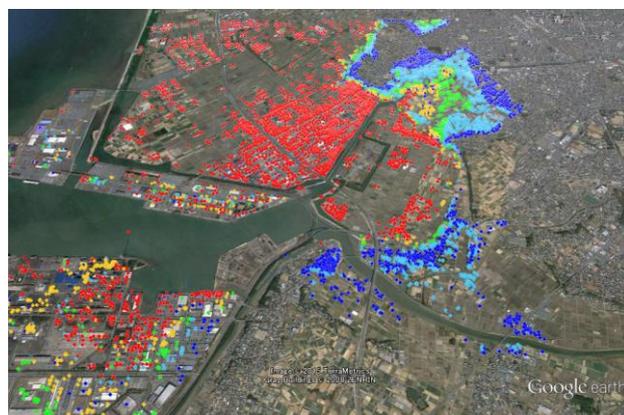


図-13 豊橋における3mの津波高を与えた時の推定建物被害結果(豊橋市、赤色：全壊、橙色：大規模半壊、黄緑色：半壊、水色：一部損壊)

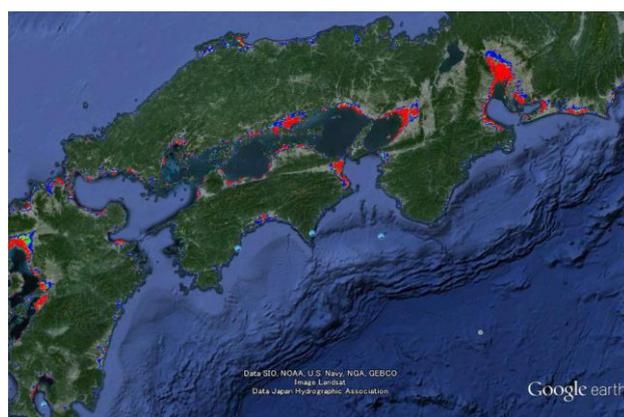


図-14 西日本における3mの津波高を与えた時の建物被害率(全壊・全半壊の割合)の推定結果(4次メッシュ集計)

4. 結論

2011年東北地方太平洋沖地震津波における建物被害データを用いて被害状況の把握をすると共に新しい津波被害評価モデルの構築・検証をした。津波被害モデルは被害関数による確率的に表現するものと全壊・半壊・被災なしというように表現をする識別モデルがあり、外力は津波流速、浸水深などとし、建物構造、築年数別に構築(ランダムフォレストの場合は決定木の枝として)するものである。

本稿では東北地方太平洋沖地震の現地調査データに土地利用や住宅地図を統合することで、津波の陸上流速計算をし、累積正規分布による確率的モデルと機械学習ア

ルゴリズムの1つであるランダムフォレストを用いて、新たな津波被害モデルの構築をした。累積正規分布モデルは簡易的に被害推定し、集計して被害棟数を把握する分には適用可能なモデルだと思われる。より高精度に被害推定する場合はランダムフォレストによる推定が誤差率2%程度(正答率98%)なので望ましいと思われる。

また、ランダムフォレストを用いて各特徴量の Gini 係数を算出することで津波被害にどの特徴量が大きく影響を与えるかを明らかにした。既存研究においては建物構造の影響が大きいと考えられてきたが本研究により築年数の影響が大きいことが明らかになった。

さらにマイクロジオデータと呼ばれるデジタル住宅地図等の個別建物の階数はもちろん推定構造や築年数などを用いることで日本全土を対象に、任意の浸水深を入力することで建物の倒壊被害を評価するための基盤データを整備した。それらを用いて任意の浸水深に対する建物被害を、任意の集計単位で定量的に評価する手法を提案した。ただし本研究で整備したデータは推定値であるため、集計した巨視的な被害量のレンジとして考える必要がある。一方で日本全土を対象としたデータ整備と被害推定から、我が国ではどの地域が相対的に津波に脆弱な状態にあるのかという情報が日本全土で明らかになった。

本研究で整備したデータには依然として数多くの課題が残されている。例えば津波の流速計算は5mメッシュを基本としたものの簡易的な手法であるため検証の必要がある。また任意の浸水深を入力できるため南海トラフ地震における想定浸水域を入力し推定する予定である。結果の公開・共有方法についても適切な方法を議論すると共に、集計データの元になっている建物単位の非集計データの流出、一人歩きが起こらないような管理が必要である。国土スケールで東北地方太平洋沖地震津波の被害を基にした任意の浸水深に対する地震津波の被害を評価出来るデータ基盤が整ったことは、防災政策を考える材料に大いに貢献出来る可能性がある。

謝辞: 文部科学省のグリーン・ネットワーク・オブ・エクセレンス(GRENE)事業 環境情報分野「環境情報技術を用いたレジリエントな国土のデザイン」の一環として実施したものである。ここに記して謝意を表したい。

参考文献

- 1) 国土交通省, 内閣府, 文部科学省: 日本海における大規模地震に関する調査検討会報告, 2014
- 2) 中央防災会議: 南海トラフの巨大地震 建物被害・人的被害の被害想定項目及び手法の概要, 2012.
- 3) 愛知県: 愛知県被害想定資料 2003.
- 4) 首藤伸夫, 1992年, 津波強度と被害: 津波工学研究報告, 第9号 pp.101-36
- 5) 越村俊一, 行谷祐一, 柳沢英明: 津波被害関数の構築, 土木学会論文集 B, Vol.65 No. 4, pp. 320-331, 2009
- 6) Breimman L, Random forests, Mach.Learn.2001, Vol.45, pp. 5-32
- 7) 杉本 知之, 下川 敏雄, 後藤 昌司: 樹木接近傍法と近年の発展, 計算機統計学 Vol. 18(2), pp. 123-164, 2007
- 8) 国土交通省: 復興支援アーカイブ: URL: <http://fukukou.csis.u-tokyo.ac.jp/>
- 9) 枝松富英夫, 飯塚秀則: 津波の陸上流速とその簡易推定法, 海岸工学論文集 Vol. 45, pp. 361-365, 1998
- 10) 小谷美佐, 今村文彦, 首藤伸夫: GIS を利用した津波遡上計算と被害推定法, 海岸工学論文集 Vol. 45, pp. 356-360, 1988
- 11) 国土数値情報ダウンロードサービス: <http://nlftp.mlit.go.jp/ksj/>
- 12) 村尾修・山崎文雄: 自治体の被害調査結果に基づく兵庫県南部地震の建物被害関数, 日本建築学会構造系論文集, 527, pp.189-196, 2000.
- 13) Boser, B. E.; Guyon, I. M.; Vapnik, V. N., "A training algorithm for optimal margin classifiers". Proceedings of the fifth annual workshop on Computational learning theory - COLT '92. p. 144, 1992
- 14) 金明哲編, 下川敏雄, 杉本知之, 後藤昌司: Rで学ぶデータサイエンス9, 共立出版, 2013
- 15) R Díaz-Uriarte, SA De Andres: , Gene selection and classification of microarray data using random forest, BMC bioinformatics 7 (1), 3
- 16) 世田谷 iMap: <http://www.sonicweb-asp.jp/setagaya/>
- 17) Akiyama Y, Sengoku H, Hiroyuki T, Shibasaki R (2011) Development of Commercial Accumulation Polygon Data Throughout Japan Based on the Digital Classified Telephone Directory. Computers in Urban Planning and Urban Management, F-TC-3(1)
- 18) Akiyama, Y., Takada, H., and Shibasaki, R. : Development of Micro Population Census through Disaggregation of National Population Census, CUPUM2013, conference papers, 110.
- 19) Ogawa, Y., Akiyama, Y. and Shibasaki, R., 2013, "Evaluation of Catastrophic Earthquake Damage throughout Japan using Estimated Micro Data", CUPUM2013 conference papers, 103.

(?受付)