

令和 3 年 5 月 29 日現在

機関番号：12601

研究種目：基盤研究(C)（一般）

研究期間：2018～2020

課題番号：18K11987

研究課題名（和文）オープンストリートマップにおける道路データの品質推定手法の構築

研究課題名（英文）Development of quality assessment methodologies for road data of OpenStreetMap

研究代表者

金杉 洋（Kanasugi, Hiroshi）

東京大学・空間情報科学研究センター・協力研究員

研究者番号：00526907

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：OpenStreetMap(OSM)はオープンな地理空間データとしてグローバルに整備・更新され利用機会が増加している一方で、そのデータ品質の評価が課題とされている。

本研究ではOSMの道路データを対象に、他の道路データとの比較やOSM自体から得られるデータの特徴からOSM道路データの品質を検証する手法について検討した。先行研究で提案された指標を援用し、道路データの比較指標として交差率・網羅率を定義し、日本全国でデータ比較を実施した。結果としてOSM道路データは、高速道路ではトンネルなどの一部地域を除いて概ねDRMと同程度に位置し、DRMが含まない幅員の狭い道路をより網羅していることが示された。

研究成果の学術的意義や社会的意義

OSMは参加するボランティアが任意に地域のデータ編集を行うため、データの鮮度や精度は保証されていない。グローバルに整備されるオープンな地理空間データとして途上国や大規模災害発生地域などでも活用機会が増加しているOSMデータの品質評価は、データを利用する上でも重要な指標となるだけでなく、OSMデータを編集するボランティアやコミュニティにとっても、データの品質が地域ごとに定量的に示されることはデータ編集を行う優先順位や編集内容を定める上でも重要となる。

研究成果の概要（英文）：OpenStreetMap (OSM) is an open geospatial data and has been developed and updated globally. In recent years, the opportunities of OSM utilization have been increasing. However, due to the data characteristics being compiled by volunteers, the evaluation of data quality has been discussed.

In this research, we examined a method for data quality assessment for OSM road data by comparing it with other road data and by extracting the features of OSM road data itself. Based on the indices proposed by the previous study, we defined the intersection ration and coverage ratio as the comparison indices of road data, and conducted data comparison for entire Japan. The results show that OSM road data generally exist at the same area as DRM highways except for tunnels, and that it covers narrow roads which DRM does not include.

研究分野：図書館情報学

キーワード：OpenStreetMap データ品質評価 デジタル道路地図 VGI

1. 研究開始当初の背景

オープンな地理空間データの構築を目的にイギリスで 2004 年から始まった OpenStreetMap(OSM)は、現在も日本を含めた世界中の多くのデータ編集ボランティアによってデータの編集・更新が進められている。世界の多くの地域をカバーするオープンな地理空間データの蓄積は、国連の掲げる Sustainable Development Goals(SDGs)の達成へ向けた指標検討でも重要なデータとして貢献する可能性が示されている(Ibiseh.2016)。日本国内でも 2016 年 12 月の官民データ活用推進基本法の施行により、地理空間情報を含めたオープンデータの整備・利活用の推進が加速することが予想され、日本国内でも OSM データの利活用が進むことが期待される。

しかし、OSM に代表される VGI(Volunteered Geographic Information)では、参加するボランティアが任意に地域のデータ編集を行うため、データの鮮度や精度は保証されていない。そのため作成されたデータを公的或いは商用目的に本格的に使用する際は、利用に先立って OSM データの品質を評価・検証することが必要になる。他方で、OSM データを編集するボランティアやコミュニティにとっても、データの品質が地域ごとに定量的に示されることは、データ編集を行う優先順位や編集内容を定める上でも重要である。

地理空間データの品質に関する指標の算出方法は対象となるデータの特徴に依存するため、OSM データについても定量的な評価指標の検討が望まれる。しかし、道路データの場合、国ごとの道路種別や交通規則の差異に加えて、品質評価の基準となる道路データとして一定の精度を持ちながら国全体を網羅するデータの入手の困難さもあり、国土レベルで OSM 道路データの品質評価を行った事例は少ない。また日本をはじめとした先進国においては、公共測量等により定期的に地理空間データが更新されるため、精度や鮮度の担保されたデータを品質評価の比較対象として利用することができるが、そうしたデータが十分に整備されていない途上国においては同様の品質評価指標が採用できるとは限らない。そのため、他の道路データとの比較に依らず、一定の精度で品質評価するための新たな指標も望まれる。更に、OSM のデータは日々世界中のボランティアによって編集されるため、品質評価も一定期間毎に実施し結果を更新することが必要となる。

2. 研究の目的

本研究では、多種多様な地理空間データを含む OSM から、特に道路データを対象に先行研究で提案された品質評価指標を踏襲・援用しながらデータの品質評価指標を検討する。主に Haklay(2010)の提案した位置精度と完全性の指標を援用しながら、日本全国スケールで商用に提供されている道路データ(デジタル道路地図(DRM)や基盤地図道路縁データ)と OSM 道路データの比較・評価を試みる。具体的な指標として交差率と網羅率を定義し、市区町村及び 1 km メッシュの単位で両道路データと比較し、差異を考察する。

OSM データは日々世界各地域で更新されるため、一定期間毎に品質評価の結果も更新されることが望ましい。そのため、作成した品質評価を継続的に実施し、結果をアーカイブするための仕組みを検討する。

3. 研究の方法

(1) 道路データの概要

デジタル道路地図(DRM)

デジタル道路地図(DRM)は主にカーナビでの利用を中心整備されている日本全国の道路データであり、日本における一般的な道路データのひとつである。DRM は全国デジタル道路地図データベース標準に基づいて 1/25,000 縮尺相当で整備され、道路の地理空間形状だけでなく道路経路探索用に、交差点をノード、交差点間の道路区間をリンクとしたグラフ構造のネットワークデータとして提供される。また高速道路や一般都道府県道などの基本道路網に加え、街路など詳細な道路網を表す全道路が、道路種別や幅員などの詳細な属性情報と合わせて含まれている。本研究では、東京大学空間情報科学研究センターの共同利用データで提供されている、拡張版全国デジタル道路地図データベース 2017 年版(住友電工)を使用した。

OpenStreetMap(OSM)

OSM は道路に限らず多様な地理空間データを含むため、key と value からなる「タグ」によりラベルを付けることで地物の属性情報を分類している。道路全般を表す highway タグは、自動車道路に限らず歩道や競技場まで様々な道路を包含しており総延長は DRM よりも長い。そこで、OSM に含まれる主な自動車道路を抽出するため、highway タグのうち主要な道路 8 種類と接続道路 5 種類の計 13 種類 (motorway, trunk, primary, secondary, tertiary, residential, unclassified, motorway_link, road, trunk_link, primary_link, tertiary_link, secondary_link)を抽出し比較に使用する。ただし、ここで採用した 13 種類の highway タグは、

交通規則や道路種別の定義の違いから、必ずしも DRM の道路種別と明確に対応づかないことを留意されたい。本研究では OSM の履歴データを提供している Geofabrik から取得した 2017 年 7 月 26 日時点の OSM データを使用した。

データの分割

DRM 及び OSM 道路データを、比較の単位領域となる 1 km メッシュ（三次メッシュ）と市区町村界でそれぞれ分割する。両道路データが交差する 1 km メッシュは、DRM では 265,028 件、OSM では 306,668 件であり、共通するメッシュは 262,270 件であった。一方、市区町村の行政区界は国土数値情報の平成 29 年度行政区界データを採用し、政令指定都市の区分けも含めて市区町村コードの有効な 1,892 件を単位領域として使用する。

(2) データ比較の指標

Haklay(2010)の提案した位置精度と完全性の指標を援用しながら、道路データ比較の指標として交差率と網羅率を定義し、市区町村及び 1 km メッシュの単位で両道路道路データを比較し、差異を考察する

交差率

道路種別の定義が異なる両道路データに対し、相互に路線名称や路線番号の振られていない細道路までを正確に対応付けるのは難しい。そこで本研究では OSM 道路データ上の motorway タグデータ(約 24,257 km)のみを比較対象とする。OSM motorway は「入口の少ない中央分離帯付きの 2 車線以上の道路」で「アメリカのフリーウェイやドイツのアウトバーンが相当する」道路と定義される。これに対応する DRM の道路として、一般に「高速道路」とされる道路を対象に、DRM motorway(約 26,017 km)を抽出し比較対象とした。

更に抽出した DRM motorway の路線名称 211 件に対して、OSM motorway の道路名称は 2,371 件で冗長性が高く、路線名称からの正確な対応付けは難しい。そこで、Haklay(2010)では対応する道路同士で位置精度を算出することで位置を比較したのに対し、本研究では道路単位ではなく DRM・OSM 共に 1 km メッシュ単位に motorway を区切り、各メッシュの道路の交差率として次式から算出する。

$$F_{int}(m) = \frac{\sum_{l_i \in L'_m} Len(l_i \cap B_{drm})}{\sum_{l_i \in L'_m} Len(l_i)}$$

F_{int} は DRM の道路幅員情報をもとに路線形状(道路中心線)からのバッファ B_{drm} を生成し、そのバッファに対して同一メッシュ(m)に含まれる OSM motorway の道路ライン L'_m が交差する道路延長の割合として算出する。また、前述の通り、DRM は 1/25,000 相当の縮尺であるため、位置の標準偏差が 17.5 m 以内であることに留意し、道路幅員ごとにバッファ半径を 17.5 m 加算した場合の交差率も合わせて算出し結果を比較する。

網羅率

Haklay(2010)が示した OSM 道路データの完全性(Completeness)は、任意の 1 km メッシュ内において、対象の OSM 道路データ(13 種類の highway タグ)の総延長と、同メッシュ内における基準道路(Ordinance survey 道路)の総延長との差分で定義されている。しかし、対象範囲内における道路延長の絶対差もあるため、道路延長の差分のみでは網羅度合いの直感的な比較は難しい。そこで、比較単位領域 m (1 km メッシュと市区町村界)における OSM 道路データの網羅率 F_{cvr} を、DRM の総延長を基準とした割合として次式の通り定義する。

$$F_{cvr}(m) = \frac{\sum_{l_i \in L_m} Len(l_i)}{\sum_{r_i \in R_m} Len(r_i)}$$

L_m, R_m は、それぞれ比較単位領域 m 内に区切られた OSM 道路データと DRM の集合を表す。なお、領域 m 内に DRM が存在しない場合は、網羅率 F_{cvr} は定義しない。

4. 研究成果

図 1 に DRM motorway と OSM motorway の両データが通過する 1 km メッシュ全 13,774 件での交差率の算出結果を示す。DRM の標準誤差を考慮しない場合(図 1 a)でも、交差率が 0.8 以上のメッシュが約 84.1%、0.9 以上が約 79.2%であり、motorway に限られた比較結果ではあるが、全体として OSM 道路データは DRM と概ね同程度の位置に整備されている。

一方で DRM の標準誤差を考慮した場合でも、交差率 0.9 以上のメッシュが約 95.3%に留まり、完全には交差しない。その原因は、OSM motorway がメッシュ内に存在しない場合や、一部のエリアで両道路データの位置が大きくずれていることにある。特に交差率の低いメッシュは、トンネル区間や山間部などに多く分布している。これは衛星画像のトレースが一般的な編集方法である OSM において、これらのエリアで道路を衛星画像から判読困難であることが、位置のずれに

起因していることが伺える。

図 2 に 1 km メッシュ 262,270 件, 市区町村 1,896 件での網羅率の算出結果を示す。網羅率が 1 よりも大きい場合に, 比較単位領域内で OSM 道路データが DRM よりも道路延長が長いことを表す。網羅率の算出結果では, 91,133 メッシュ(約 34.7%), 804 市区町村(約 42.4%)で, OSM 道路データが DRM よりも道路延長が短い(網羅率が 1 未満)ことが示された。図 2a から, 特に東京, 名古屋, 大阪などの大都市近郊で網羅率が低い傾向が見られるが, 必ずしも中心部の網羅率は低くない。例えば, 東京都 23 区の網羅率の平均値は 1.01 であり, 最大値の新宿区が 1.103, 最小値の江戸川区も 0.911 であり, 両道路データは同程度の道路延長を有している。他方で, 23 区周辺の郊外地域に, 網羅率の低い自治体が分布している。これは新規道路の開通が, OSM 道路データに反映されるまでの時間遅れが, 可能性としてあげられる。つまり, 衛星画像のトレースが一般的な編集方法である OSM において, 衛星画像の更新までの時間遅れと, 実際に編集されるまでの時間遅れが影響していると推測する。これは新規道路の開通状況, OSM 道路データの更新日時, OSM 編集者数の分布などを加味して比較することで, 今後明らかにしていきたい。

一方で, 北海道・東北や内陸山間部には網羅率の高い結果が分布している。こちらは, DRM に登録のない細かな道路(幅員 3 m に満たない市町村道以下の道路)の多くが, OSM に登録されていることが影響している。特に, 網羅率の算出時に対象外とした, OSM 道路データがあり DRM のない 1km メッシュは, 44,398 件(総延長は約 49,960 km)に上る(図 2c)。これらのメッシュは, 2015 年国勢調査の平均人口で 0.3 人(最大 256 名)であることから, OSM 道路データは DRM に比べ, 人口の少ない地域を広く網羅しているといえる。一方で, こうした地域の OSM 道路データは長らく更新がされていない。今後はデータの更新日時も考慮した比較が必要であろう。

引用文献

1. Haklay, M. (2010) How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets, *Environment and Planning B: Planning and Design*, 37, pp.682-703.
2. Pierre L. Ibiseh, et.al, (2016), A global map of roadless areas and their conservation status, *Science*, 354(6318), pp.1423-1427

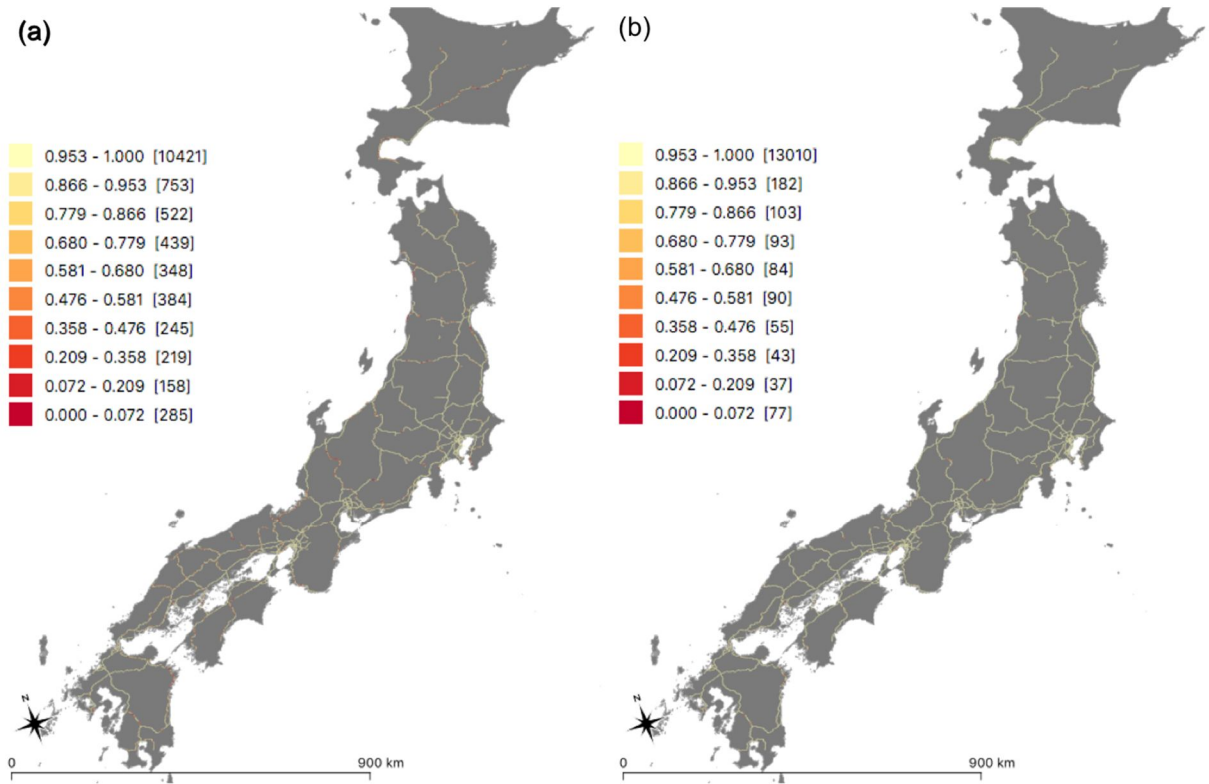


図 1. 1 km メッシュ単位での DRM motorway と OSM motorway の交差率 (N=13,774)
 (a) DRM 標準誤差の考慮なしの結果 (b) DRM 標準誤差を考慮した結果

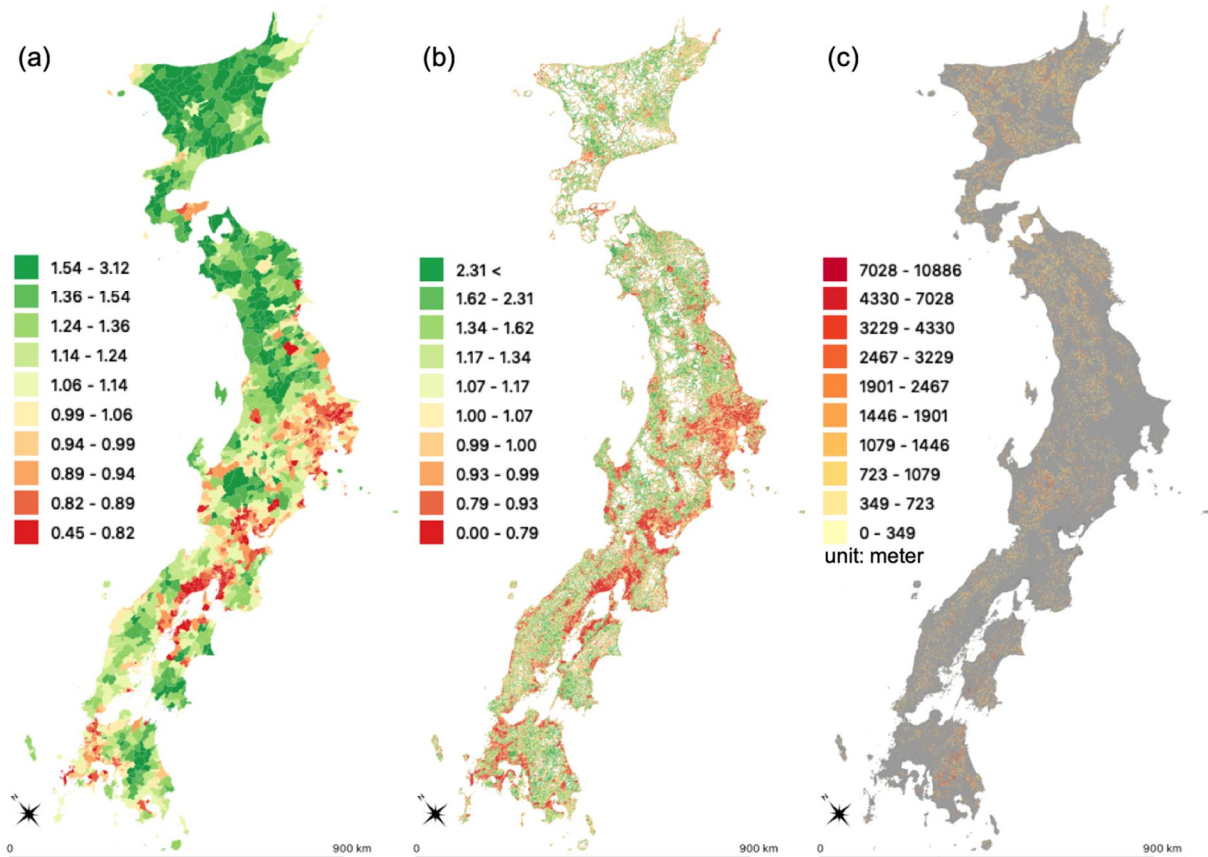


図 2. DRM を基準とした OSM 道路データの網羅率 (a) 市区町村(N=1,896) (b) 1 km メッシュ (N=262,270) (c) 網羅率適用範囲外(DRM のない)メッシュの OSM 道路データの総延長 (N=44,398)

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 金杉洋, 瀬戸寿一, 関本義秀, 柴崎亮介	4. 巻 27
2. 論文標題 オープンストリートマップ道路データとデジタル道路地図の比較 - 位置と完全性に注目して -	5. 発行年 2018年
3. 雑誌名 GIS-理論と応用	6. 最初と最後の頁 43-48
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 金杉洋, 瀬戸寿一, 関本義秀, 柴崎亮介
2. 発表標題 オープンストリートマップ道路データとデジタル道路地図の比較
3. 学会等名 CSIS DAYS 2019
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	瀬戸 寿一 (Seto Toshikazu) (80454502)	東京大学・空間情報科学研究センター・特任講師 (12601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------