

日本全土の商業統計ポイントデータの開発と
商業集積地域ポリゴンデータの信頼性検証
秋山祐樹・仙石裕明・田村賢哉・柴崎亮介

**Development of the Commercial Statistics Point data All Over Japan and
Reliability Verification of Commercial Accumulation Polygon Data**

Yuki AKIYAMA, Hiroaki SENGOKU, Kenya TAMURA and Ryosuke SHIBASAKI

Abstract: We have been developing new spatial dataset which can monitor shapes and locations of commercial accumulation (CA) throughout Japan. It has already accomplished to develop polygon data of CAs using digital telephone directory by our original spatial processing. However reliability of our data have not verified adequately. In this paper, reliability of our data has verified to cross-check our data with the Commercial Statistics of Japan. Supervised data to monitor distributions of CA is the list of commercial area in the Report by Characteristics of Location in the Commercial Statistics. The supervised data were converted into point data to add address and to geocode. In addition reliability of our data has verified to cross-check our data with the point data of the Commercial Statistics in terms of their locations and the number of shops. As a result, our data can monitor CAs with high reliability in not only prefectures in capital region but also prefectures without large cities.

Keywords: 商業集積地域 (commercial accumulation), 商業統計 (commercial statistics), 信頼性検証 (reliability verification), データ開発 (data development)

1. はじめに

地域の魅力、賑わいの大きな部分が商業地域によって支えられている。我が国では近年多くの地域においてロードサイド店舗群や大規模ショッピングモールなどの新しい商業集積が形成され、便利で魅力的な購買機会を提供している。一方、伝統的な中心市街地の商店街は流通競争に敗れて商業地域と

しての活力や魅力が失われつつある地域も多い。

そのため地域活性化、中心市街地活性化の目標の下に、国や地方自治体では様々な政策やインフラ整備事業等を行っている。しかし商業地域の成長・衰退、分布の変化等の実態を詳細に捉え、かつ全国で利用可能なデータは限られている。このことは商業地域の現状を把握し、全国規模で同じ基準で評価すること、政策効果の把握・評価すること等に大きな困難をきたしている。またこうした課題が実際に存在することも、著者らによる国の上記業務の担当者へのヒアリングにより明らかになっている。

なお多くの既存研究における商業地域・商店街の

秋山祐樹 〒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

Phone: 03-5452-6417

E-mail: aki@iis.u-tokyo.ac.jp

実態把握では、商業統計や事業所・企業統計調査等の既存統計を用いる場合が多いが（室町ほか（1994）；富永（2002））、多くの既存統計には集計単位の粗さや、回収率の低下・ばらつきの問題がある。また具体的な商店街の位置や形状に関する情報は公開されていない場合が多く、市区町村といったマクロ単位での情報しか収集することが出来ない。

こうした背景を受け、日本全国をカバーする商業地域の分布に関するデータの整備が望まれている。我々はこの課題に対してこれまでにデジタル電話帳データ（株式会社ゼンリン「テレデータシリーズ テレデータ Pack! 法人・個人電話帳データベース」）を用いて日本全国の商業集積地域1つ1つの分布・形状・規模（店舗数）をポリゴンで観察できるデータセットの開発を進めてきた（Akiyama et al. (2011)・本論文では「商業集積地域ポリゴンデータ」と呼ぶ。）。技術的には既に日本全土の任意の地域のデータ整備が可能になっている（図1）。

1.1 これまでに行われた信頼性検証

上記の課題解決に我々のデータを用いるためには、本データの信頼性の検証が必要である。ここでいう信頼性の検証とは本データで得られる商店街・商業地域の位置と規模（店舗数）が、他の統計情報とどの程度合致するか、という意味である。これまでに東京都全域における商店街名鑑との突き合わせ（Akiyama et al. (2010)・場所と店舗数の比較）、全国の都道府県ごとにおける商業統計との突き合わせ（Akiyama et al. (2011)・商業地域の数の比較）は行われており、何れも良好な結果が得られている。しかし全国規模での商業地域の位置と規模の検証はまだ行われていない。

1.2 目的

前述した全国規模の検証を行うために、商業統計から得られる日本全国の商業地域の分布データを作成する。平成19年度商業統計表立地環境特性別

統計編第10表（経済産業省）には、日本全国の商業地域・商店街の名称、店舗数、従業員数等が収録されている。本データを元に日本全国の商業地域のポイントデータ（商業統計ポイントデータ）を開発する。続いて我々が開発した商業集積地域ポリゴンデータと、商業統計ポイントデータを場所と規模（店舗数）に基づいて突き合わせを行い、我々のデータの信頼性を明らかにする。

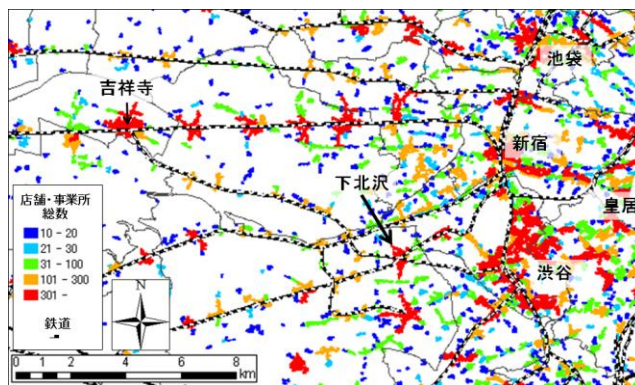


図-1 東京都における商業集積地域ポリゴンデータ

2. 商業統計ポイントデータの開発

商業統計表立地環境特性別統計編には、日本全国の商業集積地区・商店街の名称、店舗数、従業員数等が収録されている。このデータを用いることで、全国12,568箇所の商店街の情報を収集出来る。しかし商店街の位置情報は商店街の名称のみで住所や経緯度は不明である。

そこで商業統計に記載されている商店街の名前を手掛かりに、GoogleやYahoo!等の検索サイトと平成16年版全国商店街名鑑（全国商店街振興組合連合出版）を主に用いて各商店街に住所を付与し、「商業統計ポイントデータ」を作成した。住所収集の手がかりは商店街の名称のみであったため、当初は地方の商店街を中心に住所が発見出来ない等の問題があった。そこで調査員を25名に増強し約2週間に渡り調査を行うことでこの問題を解決し、12,568件中11,625件（92.4%）の商店街に住所情

報を与えることができた。また精度面での向上を図るために、商店街名と参照した住所との照らし合わせを複数グループで実施し、それらの結果のクロスチェックを行った。最後に東京大学空間情報科学研究センターの号レベルアドレスマッチングシステムを使用し、11,625 件中 11,603 件 (99.8%) の商店街に位置座標を与えることができた (図 2)。

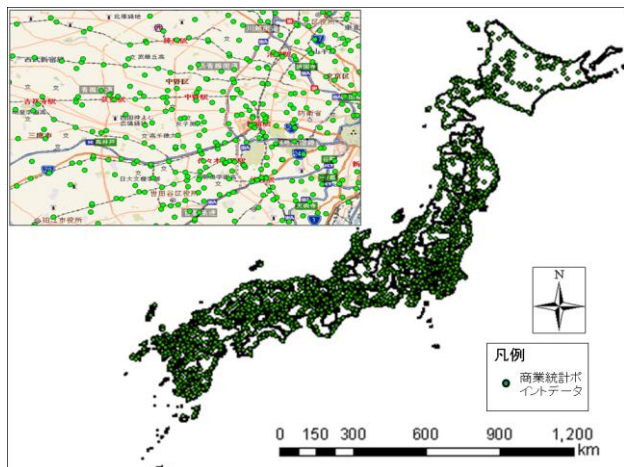


図-2 商業統計ポイントデータ

3. 信頼性の検証方法

信頼性の検証には2つの段階がある。第1段階として商業統計ポイントデータと商業集積地域ポリゴンデータの位置がどの程度合致するかを検証する。さらに第2段階として、それぞれが保有する商業地域の小売店舗数の比較を行う。商業集積地域ポリゴンデータは商業地域を構成する様々な事業所店舗を含むが、商業統計は小売店舗数しか含まない。そこで商業集積地域ポリゴンデータ側も電話帳の業種情報に基づき、小売店のみを抽出し、商業集積毎に小売店舗数を与えておく。なお商業統計の調査時期(2007年)に近い時期の商業集積地域ポリゴンデータを得るために、2008年の電話帳データを用いて商業集積地域ポリゴンデータを作成した。

3.1 本研究における信頼検証での問題

商業統計ポイントデータはポイントデータであり、一方商業集積地域ポリゴンデータはポリゴンデ

ータである。ポイントデータとポリゴンデータの位置の突き合わせ、またそれぞれの属性情報の突き合わせには幾つかの課題がある。

まずポイントの位置がポリゴンと微小にずれる場合がある(図3)。商業統計に与えた住所が不正確であったり、アドレスマッチングの精度が低かったりする場合に発生する。なお東京都の場合、882箇所の商業統計ポイントが存在するが、ポリゴンと重複するものは500件(56.69%)であった。こうしたポイントを近隣のポリゴンと関連付ける手法が必要である。また同時にポリゴンから距離が離れているポイントは関連付けから除外する必要もある。

次にポイントとポリゴンの関係が一对多、あるいは多対一の場合がある(図4)。図4左側の例では1つのポイントがその南北に分布するポリゴンの情報を包含している。この場合、1つの商業統計ポイントに対して、重複あるいは隣接する複数のポリゴンの小売店舗数を合計してポイントと比較する必要がある。一方図4右側の場合、1つのポリゴンが複数のポイントを包含している。この場合はポリゴンに重複する複数のポイントの小売店舗数を合計してポリゴンの小売店舗数と比較する必要がある。

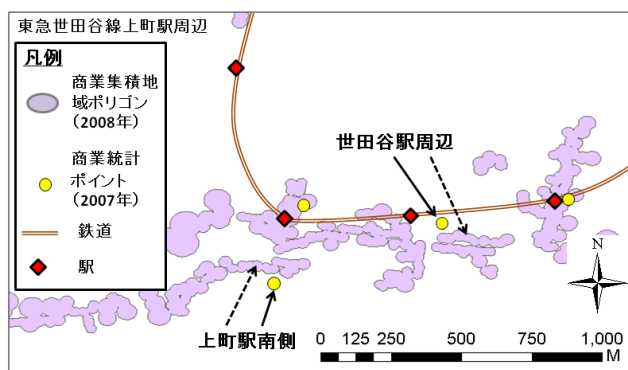


図-3 ポリゴンがポイントと微小にずれる場合

3.2 信頼性検証のためのデータ処理方法

上記の課題を解決するために、まずポイントが持つ小売店舗数とポイント周辺のポリゴンの店舗密

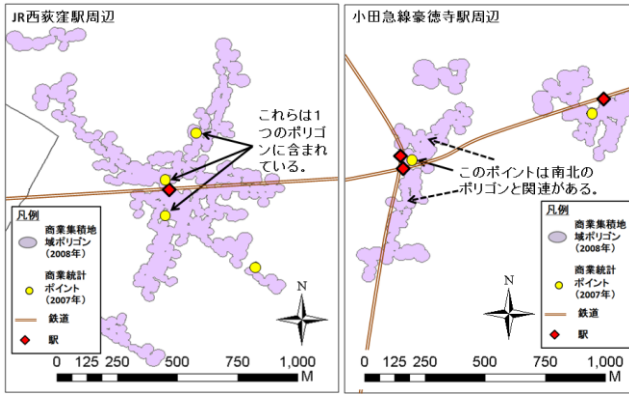


図-4 ポリゴンとポイントが一对多、あるいは多対一となる場合

度に応じてバッファ領域を発生させる。ポイントからの検索圏を半径 1km とし、検索圏内に存在するポリゴンの店舗数と集積面積を合算し、各ポイントが分布する地域における店舗 1 件が占める平均面積を算出する。続いて 1 店舗当りの平均面積と、各ポイントが持つ店舗数（商業統計の小売店舗数）を元にバッファリング距離を決定し、バッファ領域を発生させる。最後にそのバッファ領域と重複する商業集積地域ポリゴンを結合することで、新しい集計単位を生成し、その単位を元に場所と規模（店舗数）の合致の度合いを評価する。

図 5 を用いて上記の手法を具体的に説明する。ある商業統計ポイント n において、検索圏内に存在する商業集積地域ポリゴンの数を s 、ポリゴン 1 の面積を S_1 、店舗数を N_1 とする時、ポイント n の店舗 1 件当りの平均面積 A_n は式 1 で与えられる。

$$A_n = \frac{\sum_{k=1}^s S_k}{\sum_{k=1}^s N_k} \quad (1)$$

図 5 の場合、 A_n は 100 m^2 となる。

続いてポイント n の小売店舗数を用いて、ポイント n のバッファリング距離を決定する。ポイント n の店舗数を T_n とすると、ポイント n が保有出来る面積 M_n は式 2 で与えられる。

$$M_n = T_n A_n \quad (2)$$

そして M_n を用いてポイント n のバッファリング距離 r_n を式 3 で定義する。

$$r_n = \sqrt{\frac{M_n}{\pi}} = \sqrt{\frac{T_n A_n}{\pi}} \quad (3)$$

図 5 の場合、 r_n は 85.59m となる。

最後にバッファポリゴンとそれに重複する商業集積ポリゴンを結合して 1 つの領域を生成する。この処理で生まれた新しいポリゴン領域に含まれる商業集積（ポリゴン）の小売店件数と、商業統計（ポイント）の小売店件数を合計して比較する。バッファポリゴンと重複しないポリゴンおよびポイントは隣接データ無しとして処理する。

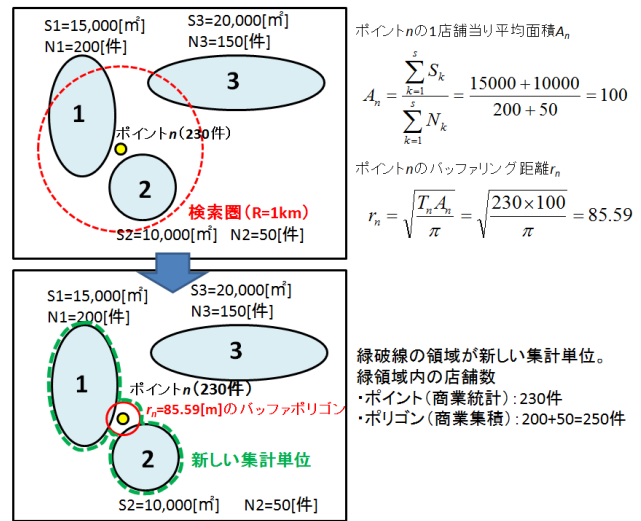


図-5 信頼性検証のためのデータ処理方法

4. 検証結果

4.1 東京都における検証結果

図 5 の集計方法を用いてまず両データが新規集計領域（図 5 の緑破線領域）にどの程度含まれるかを明らかにした。商業統計ポリゴンデータは 881 個中 758 個（86.04%）含まれており、商業集積地域ポリゴンデータが商業統計に記載された商業地域の場所の多くを把握出来ることが分かる。一方商業集

積地域ポリゴンデータは4373個中1033個(23.62%)に留まる。これはこのデータが商業統計に記載されていない小規模な商業地域・ロードサイド型店舗群・大規模小売店舗も含むためである。また店舗の空間的連続性の喪失により、本来1つの商店街であった地区が複数の小規模な集積に分断されてしまったことも原因である。これは表1の結果からも明らかである。表1は東京都の商業集積地域ポリゴンデータのうち、新規集計領域に含まれるものと含まれないものの規模(小売店舗数)を比較したものである。新規集計領域に含まれない集積の多くは小売店舗数が10件以下の小規模な集積となっていることが分かる。

続いて新規集計領域ごとに両データの小売店舗数を比較した。図6に両データの小売店舗数を新規集計領域ごとに取得し、相関を取った結果を示す。両データ間には強い相関が見られることが分かる。この結果から商業地域の規模(小売店舗数)の観点から見ても商業集積地域ポリゴンデータが商業統計を良好に説明出来ていることが分かる。また図7は新規集計領域ごとの小売店舗数の誤差率の分布を示している。この値が正の場合、商業集積地域ポリゴンデータで得られる値の方が大きく、その逆の場合は商業統計ポイントデータで得られる値の方が大きくなることを意味する。値が正の領域の多くは東京都心部に分布している。都心部では大規模な商業集積が連続的に分布するため、商店街組織をベースに集計された商業統計の件数が、実態の商業集積の規模より小さくなっていることが原因と考えられる。一方、値が負の地域の多くは郊外(特に郊外の駅周辺)に分布している。郊外地域の商店街を商業集積地域ポリゴンデータで見た場合、商業統計に登録された店舗数よりも、店舗数の少ない集積、あるいは分断されている集積が数多く分布している。商店街組織単位の商業地域(商業統計ベース)

表-1 新規集計領域に含まれる集積と含まれない集積の規模比較(東京都の場合)

小売店舗数	包含	割合(%)	包含されず	割合(%)
0～10	384	37.17	2417	72.37
11～20	195	18.88	540	16.17
21～30	100	9.68	142	4.25
31～50	93	9.00	119	3.56
51～100	140	13.55	86	2.57
101～200	86	8.33	28	0.84
201～300	22	2.13	5	0.15
301～	13	1.26	3	0.09
合計	1033		3340	

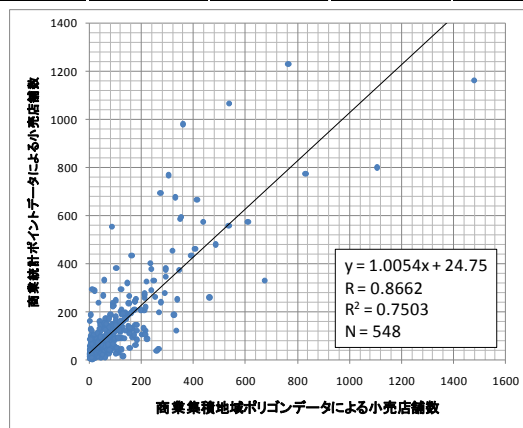
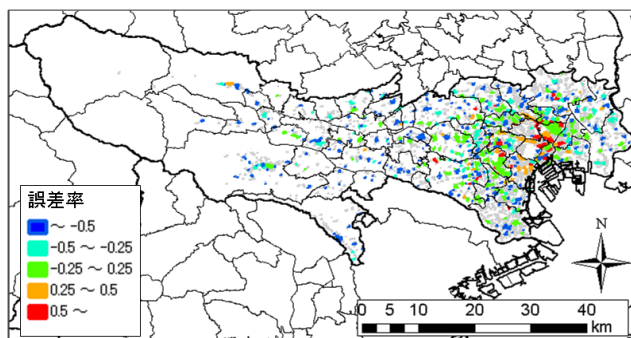


図-6 両データの小売店舗数比較



※誤差率=図6の残差 / 商業集積統計ポイントデータの小売店舗数

図-7 東京都における新規集計領域ごとの誤差率

が、実態の店舗分布状況(商業集積地域ポリゴンベース)に必ずしも一致しない事が原因と考えられる。

以上より商業集積地域ポリゴンデータは、商業統計に登録されている商業地域・商店街の場所の多くを把握出来るだけでなく、その規模についても高い精度で把握出来ることが分かった。同時に商業統計が実際の店舗集積状況を把握しきれていない可能性があることも明らかになった。

4.2 その他の都道府県における検証結果

同様の検証を東京都以外の南関東3県（埼玉県・千葉県・神奈川県）、比較的大規模な都市と郊外型の小規模商業地域を数多く持つ福岡県と鹿児島県、大都市を持たない鳥取県と島根県、そして古くからの伝統的な商店街が少ない北海道で実施した。表2に各道県の両データの新規集計領域との重複件数を、表3に図6と同様の集計による各道県の相関分析の結果を示す。何れの結果も東京都の場合と同程度の結果を得た。表2で北海道が他の県よりも補足できる商業地域の数が少なくなっているが、これは北海道の商業地域は他の地域と比較して店舗の連坦・密度の度合いが低いことが原因である。また何れの道県でも表1、図7と同様の傾向が確認された。

以上より東京都以外の地域でも東京都と同程度の信頼性が見られる事が分かった。また同時に前節と同様の問題が発生することも明らかになった。

表-2 分析対象の道県における両データの新規集計領域との重複件数

都道府県名	商業集積地域ポリゴンデータ			商業統計ポイントデータ		
	総件数	うち重複	割合(%)	総件数	うち重複	割合(%)
埼玉県	1724	367	21.29	484	383	79.13
千葉県	1615	411	25.45	436	344	78.90
神奈川県	2334	546	23.39	513	419	81.68
福岡県	1887	455	24.11	362	299	82.60
鹿児島県	605	159	26.28	184	150	81.52
鳥取県	216	50	23.15	64	51	79.69
島根県	292	80	27.40	112	84	75.00
北海道	2039	365	17.90	482	326	67.63

表-3 分析対象の道県における相関分析の結果

都道府県名	集計領域数(N)	回帰係数(a)	相関係数(R)	決定係数(R ²)
埼玉県	216	1.4282	0.9024	0.8143
千葉県	244	1.5108	0.8461	0.7159
神奈川県	313	1.1680	0.8513	0.7247
福岡県	208	1.9919	0.9421	0.8875
鹿児島県	101	0.9424	0.8253	0.6811
鳥取県	27	0.8710	0.8631	0.7450
島根県	56	1.0146	0.8254	0.6813
北海道	235	1.1816	0.8506	0.7236

5. おわりに

商業統計ポイントデータの実現と、そのデータとの突き合わせ検証により、我々が開発してきた商業集積地域ポリゴンデータは、商業統計に記載された

商業地域・商店街の位置と規模を高い精度で捕捉出来ていることが分かった。1章でも紹介した既にこれまでに行われた他の統計情報との突き合わせ結果も合わせると、我々のデータが様々な既存統計を説明出来る信頼性を持っていると言えよう。今後は他の府県でも同様の検証を進めるとともに、既存統計では突き合わせ検証が出来ない小規模な集積の信頼性検証方法（例えば現地調査等）についての検討も進めていく予定である。

謝辞

本研究は株式会社ゼンリンとの共同研究によって実現した。また本研究に関連して内閣府、経済産業省、国土交通省の方々にもお時間を頂いた。ここに記して謝意を表したい。

参考文献

室町泰徳・原田昇・太田勝敏（1994）：都心商業地域の衰退状況と大規模小売店舗の立地動向に関する研究，日本都市計画学会学術研究論文集，29，529-534.

富永滋（2002）：市街地商店街の活性化に関する一考察：商店街実態調査とGIS活用による分析を通して，地理情報システム学会研究発表大会講演論文集，11，121-125.

Akiyama, Y., Sengoku, H., Takada, H. and Shibasaki, R., 2011. *Development of Commercial Accumulation Polygon Data Throughout Japan Based on the Digital Classified Telephone Directory, CUPUM2011*, F-TC-3(1).

Akiyama, Y., Sengoku, H. and Shibasaki, R., 2010. *Automatic Detection and Spatio-temporal Analysis of Commercial Accumulations Using Digital Yellow Page Data. World Academy of Science, Engineering and Technology* **66**, 122-127.