

「メッシュ統計の国際利活用検討会」第 4 期
報告書

一般社団法人世界メッシュ研究所

Research Institute for World Grid Squares

2025 年 5 月

【目次】

1 はじめに

2 検討作業工程

3 検討結果

(1) メッシュ統計間の類似性の測定方法

(2) メッシュ統計間の変換時の誤差評価方法

(3) メッシュ統計の表示方法

(4) グリッド統計のメッシュデータ変換における誤差評価の要点

4 まとめ

参考文献

※本報告書内容の著作権は一般社団法人世界メッシュ研究所が保持しています。利用に当たっては、次のとおり出典を明記してください。

一般社団法人世界メッシュ研究所, 「メッシュ統計の国際利活用検討会」第4期報告書 (2025)

本件に関するお問い合わせは、次のフォームにてご連絡ください。

https://www.fttsus.org/worldgrids/ja/contact_us/

1 はじめに

一般社団法人世界メッシュ研究所は、世界規模でのメッシュ統計の作成とその流通、利活用促進を支援することを主たる目的として、2020年11月に設立された非営利組織である。

2022年度、メッシュ統計の利活用を世界規模で検討する目的で一般社団法人世界メッシュ研究所「メッシュ統計の国際利活用検討会」を設置して以来、これまで継続してメッシュ統計の国際利活用に関する検討会を開催し議論を行ってきた。

- ・ メッシュ統計の国際利活用検討会：第1期（2022年8月～2022年10月）
- ・ メッシュ統計の国際利活用検討会：第2期（2023年1月～2023年3月）
- ・ メッシュ統計の国際利活用検討会：第3期（2023年9月～2023年11月）

本報告書は2024年9月～11月まで開催された、メッシュ統計の国際利活用検討会（第4期）で議論された検討結果をまとめたものである。第4期検討会は主として、メッシュ統計の分析手法や、応用方法について検討された。

表1には、メッシュ統計の国際利活用検討会（第4期）の構成員および陪席員を示す。産学からなる構成員により本検討作業が行われた。

表1 メッシュ統計の国際利活用検討会（第4期）委員構成および陪席者（所属と職名は検討会当時）

役割	氏名	所属・役職
委員長	西文彦氏	横浜市立大学・客員教授
委員	小西純氏	一般財団法人統計情報研究開発センター・主任研究員
委員	亀山正廣氏	横浜市立大学大学院データサイエンス研究科・客員講師
陪席者	佐藤 彰洋 氏	一般社団法人世界メッシュ研究所・代表理事
陪席者	加藤 茂博 氏	一般社団法人世界メッシュ研究所事・務局長

2 検討作業工程

表2にメッシュ統計の全3回の国際利活用検討会で取り扱った内容について列挙する。本検討会は1ヶ月に1度の割合で委員長、委員、陪席者が参加し、メッシュ統計の国際利活用に必要なと考えられるトピックについて調査し、その調査結果を委員会で共有することにより議論を深めた。

表2 メッシュ統計の国際利活用検討会（第4期）で議論した内容

開催回	開催日時	検討内容
第1回	2024年9月30日 (月)	メッシュ統計の国際利活用検討会（第4期）委員長、委員の挨拶につづき、議論のするめ方、対象について選定を行った。ISO/TC69/WG3 NP24108-2で想定される分析方法について主として選定を行った。
第2回	2024年10月29日 (火)	メッシュ統計間の類似性の測定方法について検討を行った。
第3回	2024年11月27日 (水)	メッシュ統計間の変換時の誤差評価方法について検討するとともに、一般社団法人世界メッシュ研究所公開用報告書の下書きについて検討した。

3 検討結果

本検討会では、メッシュ統計間の類似性を測定する方法について検討するとともに、メッシュ間の変換時の誤差評価方法の検討を主として行った。メッシュ間の類似性評価の方法とは、地域メッシュコード[1]や世界メッシュコード[2]を集計単位として集計した異なるデータ元のデータから作成された同じコード上のメッシュ統計を比較する場合である。また、メッシュ間の変換時の誤差評価方法とは、世界メッシュコードを集計単位として作成されたメッシュ統計をEU Grids[3]のような異なるグリッド体系へ変換する場合の変換時の誤差を評価する場合を想定している。これらの評価方法を検討する中で、メッシュ統計間の類似性および誤差の測定方法について量的メッシュ統計と質的メッシュ統計における評価方法の違いを議論した。更に、メッシュ統計のデータの供給方式について異なる行列データ表現系が共存していることを認識し、それらの整理を行った。グリッド統計のメッシュデータ変換における誤差評価を日本における国勢調査メッシュ統計を利用して実施し、面積按分による変換方法のほかが、面積按分+グリッド人口比の重み付けよりも誤差が小さくなることを確認した。本検討会における検討結果について、（1）～（4）に示す。

（1）メッシュ統計間の類似性の測定方法

対象となるデータを量的データと質的データの2種類に分けることができるが、まず、量的データの類似性の測定には、相関係数の利用が挙げられる。

・相関係数

相関係数を利用することにより、メッシュ統計間の相関係数が求められる。量的データの組 (x_i, y_i) ($i = 1, \dots, n$)が与えられたとき、相関係数 r は次の式で定義される。

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

ここで、 $\bar{x}$ と $\bar{y}$ はそれぞれデータ x_i と y_i の標本平均を意味する。(1)式で計算される相関係数 r は、 -1 から 1 までの間の値を取り、 r が -1 に近い場合負の相関、 0 に近い場合無相関、 1 に近い場合正の相関を示す。なお、相関係数を計算する場合はサンプルサイズ n が 100 以上となることが推奨されている。

相関係数は、JIS Z 8101-1: 2015 (ISO 3534-1: 2006)で定義されている。ISO についても、ISO 3534-1:2006 で定義が存在する。

次に、質的データの類似性の測定には、適合度検定の利用が挙げられる。

・適合度検定

適合度検定には、カイ二乗 (χ^2) 検定が使われる。この検定は、A群とB群の比率（メッシュ統計間の比率）に差があるかどうか検証する統計手法である。

A群とB群に対する2組の質的データがサンプルサイズを N で与えられているとし、A群に c 種類の質的変数が、B群に r 種類の質的変数が存在するとする。このとき、A群の c 種類の質的変数の周辺度数 n_i ($i = 1, \dots, c$)、B群の r 種類の質的変数の周辺度数が m_j ($j = 1, \dots, r$)であるとする。

このとき、これらのクロス集計により得られる結合度数 n_{ij} ($i = 1, \dots, c; j = 1, \dots, r$)とする。適合度検定を実行するために算出されるカイ二乗検定統計量は次式で定義される。

$$\chi^2 = \sum_{i=1}^c \sum_{j=1}^r \frac{(n_{ij} - E_{ij})^2}{E_{ij}}$$

ここで、 E_{ij} は各属性項目に対して独立性を仮定した場合の期待度数であり、

$$E_{ij} = \frac{n_i m_j}{N}$$

と計算される。カイ二乗検定統計量の値をもとに自由度 $k = (r - 1)(c - 1)$ のカイ二乗分布

$$P(x; k) = \begin{cases} \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} \exp\left(-\frac{x}{2}\right) x^{\frac{k}{2}-1} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

を基にその累積確率分布 $\Pr[X \geq x]$ を使って

$$p = \Pr[X \geq \chi^2] = 1 - \frac{\gamma(\frac{k}{2}, \frac{\chi^2}{2})}{\Gamma(k/2)} \quad (2)$$

により、 p 値を算出して検定を行う。ここで、 $\Gamma(a)$ と $\gamma(a, x)$ はそれぞれガンマ関数

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt$$

と不完全ガンマ関数

$$\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$$

である。有意水準 $\alpha = 0.05$ または 0.01 を設定し、(2)式で算出される p 値が有意水準 α 以下の値である場合、帰無仮説 H_0 (A群とB群は独立である) が棄却され対立仮説 H_1 (A群とB群は独立ではない) が採択される。結果、 $p \leq \alpha$ であれば、有意水準 α でA群とB群との間には相関性があると判断する。

適合度検定は、JIS や ISO では定義されていない質的データの検定評価方式である。

3 番目に、1 対 1 のメッシュ統計間の測定 (量的データ) には、偏差率が挙げられる。

・偏差率

前述の相関係数及び適合度検定は、群全体でのメッシュ統計間の相関性の測定 (例えば、A群とB群の類似性の測定) に用いる手法であり、群に対して類似性を示す評価値が1つ決まる性質のものであった。

これに対して、異なる推計方式における同一メッシュでの統計値の類似性を評価するなど、メッシュごとの値の評価方法として偏差率がある。偏差率は、例えば、メッシュ w におけるモバイル空間統計の推計人口を $P_m(w)$ 、国勢調査の人口を $P_c(w)$ としたとき、

$$d(w) = \frac{P_m(w) - P_c(w)}{P_m(w) + P_c(w)}$$

として定義され、メッシュ毎に偏差率 $d(w)$ 値を得る。偏差率 $d(w)$ は、相関係数と同様に、-1 から+1 の間で示され、0 であれば類似性が高く、-1 または+1 であれば、大きく2つの値は乖離していることを意味する。偏差率 $d(w)$ は測定する2者の類似性が高いほど、0に近い値を取ることに注意が必要である。

偏差率は、JIS や ISO では定義されていない。

(2) メッシュ統計間の変換時の誤差評価方法

量的データの誤差評価方法として、以下の指標を挙げることができる。

・誤差率

1対1のメッシュ統計間を評価する指標として、誤差率があり、次の式で表すことができる。なお、Aは変換前の値で、Bは変換後の値である（以下同様）。

$$(\text{誤差率}) = \frac{A - B}{B}$$

- 平均絶対誤差率

一方、メッシュ統計全体（多対多）を評価する指標として、平均絶対誤差率（MAPE）：Mean Absolute Percentage Error があり、次の式で表すことができる。

$$M = \frac{1}{N} \sum_{i=1}^N \left| \frac{A_i - B_i}{B_i} \right|$$

そのほか、メッシュ統計全体（多対多）の誤差を評価する方法として、以下の方法を挙げることができる。

- 基本統計量

変換前後のメッシュ統計全体の基本統計量（値が0のメッシュを除く）を比較する。すなわち、度数、平均値、標準偏差、最大値、最小値、中央値、最頻値等を比較する。

- 散布図及び相関係数

変換前後のメッシュ統計全体の散布図を作成し、相関係数（外れ値を除く）を求める。

（3）メッシュ統計の表示方法

メッシュ統計を表示する形式は、次に挙げる3つの形式で表示することができる^[4]。これら形式及び名称を定義することは、メッシュ統計利用者間のデータ形式の相互理解に資する。

- 地理行列（geographic matrix）

行方向にメッシュコード、列方向に属性情報（2列目以降）を持つ行列である。n個のメッシュコード、m個の属性情報を持つとき、m列n行の行列として、以下の様に表すことができる。

$$\begin{array}{c}
 \text{メッシュコード} \\
 \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ i \\ \vdots \\ m \end{matrix}
 \end{array}
 \begin{pmatrix}
 & \begin{matrix} 1 & 2 & 3 & \cdots & j & \cdots & n \end{matrix} \\
 \begin{matrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1j} & \cdots & x_{1n} \\
 x_{21} & x_{22} & x_{23} & \cdots & x_{2j} & \cdots & x_{2n} \\
 x_{31} & x_{32} & x_{33} & \cdots & x_{3j} & \cdots & x_{3n} \\
 \vdots & \vdots & \vdots & & \vdots & & \vdots \\
 x_{i1} & x_{i2} & x_{i3} & \cdots & x_{ij} & \cdots & x_{in} \\
 \vdots & \vdots & \vdots & & \vdots & & \vdots \\
 x_{m1} & x_{m2} & x_{m3} & \cdots & x_{mj} & \cdots & x_{mn}
 \end{matrix}
 \end{pmatrix}
 \end{pmatrix}$$

図1 地理行列

・取引型行列 (transaction matrix)

行方向、列方向ともにメッシュコードを取り、ペアとなったメッシュコード間の一つの属性値を持つ行列である。0-D 表がこの行列の例で、方向の区別がない場合対称行列となり、方向を持つ場合対角要素が異なる値となる場合がある。n 個のメッシュコードを持つとき、以下の正方行列として表すことができる。

$$\begin{array}{c}
 \text{メッシュコード} \\
 \begin{array}{c}
 1 \\ 2 \\ 3 \\ \vdots \\ i \\ \vdots \\ n
 \end{array}
 \end{array}
 \begin{array}{c}
 \text{メッシュコード} \\
 \begin{array}{c}
 1 \quad 2 \quad 3 \quad \cdots \quad i \quad \cdots \quad n
 \end{array}
 \end{array}
 \begin{pmatrix}
 x_{11} & x_{12} & x_{13} & \cdots & x_{1i} & \cdots & x_{1n} \\
 x_{21} & x_{22} & x_{23} & \cdots & x_{2i} & \cdots & x_{2n} \\
 x_{31} & x_{32} & x_{33} & \cdots & x_{3i} & \cdots & x_{3n} \\
 \vdots & \vdots & \vdots & & \vdots & & \vdots \\
 x_{i1} & x_{i2} & x_{i3} & \cdots & x_{ii} & \cdots & x_{in} \\
 \vdots & \vdots & \vdots & & \vdots & & \vdots \\
 x_{n1} & x_{n2} & x_{n3} & \cdots & x_{ni} & \cdots & x_{nn}
 \end{pmatrix}$$

図2 取引型行列

・相互交流型行列 (interaction matrix)

前項の取引型行列は、単一の属性値を扱う場合は、その簡便さから取り扱いやすい。一方でメッシュコード間に複数の属性値を持つ場合は属性値の数だけ取引型行列が必要になる。この点を考慮した行列が相互交流型行列である。

行方向にメッシュコードのペア、列方向に属性情報を持つ行列である。n 個のメッシュコード、m 個の属性情報を持つとき、m 列 n*n 行の行列として、以下の様に表すことができる。

$$\begin{array}{c}
 \text{メッシュコード} \\
 \begin{array}{c}
 1 \cdot 1 \\ 1 \cdot 2 \\ \vdots \\ 1 \cdot n \\ 2 \cdot 1 \\ \vdots \\ 2 \cdot n \\ \vdots \\ i \cdot n \\ \vdots \\ n \cdot n
 \end{array}
 \end{array}
 \begin{array}{c}
 \text{属性情報} \\
 \begin{array}{c}
 1 \quad 2 \quad \cdots \quad j \quad \cdots \quad m
 \end{array}
 \end{array}
 \begin{pmatrix}
 x_{111} & x_{211} & \cdots & x_{j11} & \cdots & x_{m11} \\
 x_{112} & x_{212} & \cdots & x_{j12} & \cdots & x_{m12} \\
 \vdots & \vdots & & \vdots & & \vdots \\
 x_{11n} & x_{21n} & \cdots & x_{j1n} & \cdots & x_{m1n} \\
 x_{121} & x_{221} & & x_{j21} & & x_{m22} \\
 \vdots & \vdots & & \vdots & & \vdots \\
 x_{12n} & x_{22n} & & x_{j2n} & & x_{m2n} \\
 \vdots & \vdots & & \vdots & & \vdots \\
 x_{1in} & x_{2in} & & x_{jin} & & x_{min} \\
 \vdots & \vdots & & \vdots & & \vdots \\
 x_{1nn} & x_{2nn} & & x_{jnn} & & x_{mnn}
 \end{pmatrix}$$

（４）グリッド統計のメッシュデータ変換における誤差評価の要点

・背景と目的

詳細地域単位（1 kmメッシュ）の変換精度については、課題が多く、「目的に応じた評価基準」が重要である。今回の分析では、主に統計データと地理データの変換を対象として精度の良い方法を検討した。国勢調査データ（日本）を利用して誤差評価方法について検討した。今後日本のデータで精度が向上する方法について検討した後、ニュージーランドの 250m グリッドデータを世界メッシュデータに変換する。

・変換方法の比較

変換は、①面積按分による変換、②面積按分+グリッド人口比の重み付け の2つの方法について行い、変換結果の誤差について比較した。

① 面積按分による変換

メッシュとグリッドを重ね、面積比に基づいて人口を配分した。誤差率が最も小さい傾向を示した。特に A~C ランクで±1%未満と良い精度を示している。

② 面積按分+グリッド人口比の重み付け

面積比に、グリッド内の人口比を掛けて配分した。都市部では過大推計が生じるケースが見られた。

・評価方法と結果

1000 組の標本抽出で誤差率の母平均・母標準偏差を推定した。情報量に基づき、メッシュをランク A~E に分類した。A は情報量が最も多く E は情報量が少ない。A の精度が高く、E では誤差が大きく参考値扱いとなる結果となった。

山間部では誤差が大きくなりやすく、幾何的手法（例：中心点）による補正が有効である。

・技術的実装の概要

Python により以下の処理を行った。

1. グリッド・メッシュデータの作成
2. 重ね合わせの情報で分割ポリゴンを作成し、面積計算や重心点の結合
3. 按分計算と誤差率の評価

・結果と今後の課題

面積按分の方が安定して精度が高いことが明らかとなった。グリッドの解像度を上げることで精度向上が可能となる。精度評価の A~E は地理的まとまりに基づいているので、変換メッシュの取捨選択が可能である。グローバル展開を行うためには、精度・コスト・実装容易性のバランスが必要である。

まとめと今後の課題

メッシュ統計の国際利活用検討会（第4期）では、メッシュ統計間の類似性の測定方法について、相関係数、適合度検定及び偏差率にかかる検討を行った。次に、メッ

シュ統計間の変換時の誤差評価方法について、誤差率、平均絶対誤差率、基本統計量、散布図・相関係数にかかる検討、メッシュ統計の表示について検討を行った。最後に、グリッドから地域メッシュ統計の変換の具体的な誤差評価について、日本の国勢調査データを利用して、地理的な誤差と統計的な誤差の2つの視点から評価した。

今後の課題としては、質的データの誤差評価方法について、具体的な方法を例示することが挙げられる。

参考文献

- [1] 総務省統計局, 地域メッシュコード, URL:
<https://www.stat.go.jp/english/data/mesh/05.html>
- [2] 一般社団法人世界メッシュ研究所, 世界メッシュコード, URL:
<https://www.fttsus.org/worldgrids/en/top-en/>
- [3] EU Commission, EU reference grids, URL: <https://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-2>
- [4] 奥野隆史, 計量地理学の基礎, 大明堂, 1977