

世界メッシュコードの提案と世界メッシュ統計の作成および利活用

佐藤彰洋（京都大学、JST さきがけ）

aki@i.kyoto-u.ac.jp

椿広計（独立行政法人統計センター、統計数理研究所）

概要

本稿では日本工業規格地域メッシュコード(JIS X0410)を独自に世界拡張した世界メッシュコードとそれを用いた世界メッシュ統計の作成方法ならびに世界メッシュ統計を用いた異なるデータ間の連結分析の事例を紹介する。特に, JAXA だいち(ALOS)の 30m 標高データから作成した標高統計 3 次メッシュデータを紹介し, この標高メッシュ統計を用いることによりフィリピンにおける津波ハザード 3 次メッシュデータを作成した。更に, NASA が提供する夜間光データから作成した 3 次メッシュデータと総務省統計局国勢調査 3 次メッシュデータとの相関関係を調べた。最後に世界メッシュ統計・データが社会科学分野における利活用方法について議論する。

1 はじめに

地域メッシュコード[1] は現在日本国内において, 地域メッシュ統計を計算する上でなくてはならないものとなっている。総務省統計局は e-Stat[2] を通じて, 地図で見る統計(統計 GIS) [3] から総務省統計局国勢調査やセンサスなどの政府統計をメッシュデータ化し政府オープンデータとして公開している。更に, 国土交通省国土政策局国土情報課は国土情報に関するメッシュ統計データが公開されている[4]。このように地域メッシュ統計が我が国で広く作成され, 利用されるようになった背景には, 我が国における地域メッシュ統計制度の策定から利用に至る長い歴史が存在している。

著者らは統計 2015 年 9 月号で発表した「ビッグデータ時代に必要な標準化」[5]で, 地域メッシュコードの世界メッシュコードへの拡張方法を示した。この方式は, 我が国で利用されてきた, 地域メッシュコードの上位互換性を有する世界を覆うことができるグリッド定義と, 固有番号の割当に対する方法である。

本稿では, この世界メッシュコードに関する説明と, 世界メッシュコードを利用した計算とデータに基づく社会経済の定量的な把握方法について提案する。このデータ源として, 政府統計, インターネットから収集されるポイントデータならびに衛星データが挙げられる。本稿では, 衛星データを用いた世界メッシュデータの作成方法および, 日本の政府統計調査と衛星データから生成される世界メッシュデータとの相関性ならびに, 著者のひとり(佐藤)が日本の地域メッシュ統計を用いて開発した津波ハザード計算方式が世界メッシュコード体系を用いることにより他国においても計算可能であることを実証する。

2 地域メッシュコード(JIS X0410)とその拡張

本論に入る前に、地域メッシュ統計(JIS X0410)について概観する。地域メッシュ統計とは「グリッド法」と呼ばれる統計集計方法のひとつである。これは、空間をある区画（グリッドまたはメッシュと呼ぶ）に分割し、個票など含まれる位置属性を利用して、個票データの各項目を集計することで、空間に依存した統計を作り出す方法として開発されてきた。我が国においては、このグリッド法を日本標準の下で作成できるよう、地域メッシュ統計を作成する場合、共通のコード体系に基づいた区画定義に従うことが推奨されている。この規格を地域メッシュコードとよび、日本工業規格(JIS)の手法 X0410 として計算方法が定義されている。

我が国では、経済成長、市町村合併、社会基盤の共同運用など社会の状況変化に伴い、国、都道府県、自治体といった行政区単位の統計情報だけでなく、より小さい区域の情報が求められるようになってきた。そうしたなか、行政機関は、政府統計調査の情報を、経緯度で区分した区域を用いて地域メッシュ統計として公表する方法について、1960年代後半から調査を開始した。1970年代に入り、総理府統計局よりこの地域メッシュコードが提唱され、行政管理庁の投信として統計諮問委員会において審議を経て、正式に国の空間統計標準としての承認がなされた。その結果、国の行政機関が作成する地域メッシュ統計では、日本工業規格で制定されている「地域メッシュコード」を利用している。

この「地域メッシュコード」は、1次(約 80km)から 6次(約 125m)までの粒度があり、用途により使い分けられている。商業的に作成されている例として、ドコモモバイル空間統計[6]のように携帯電話と基地局との通信データを用いた空間統計が作成されている事例がある。また、衛星画像や航空写真に基づくメッシュ統計データが国土交通省国土地理院により作成されている。その他、インターネットから収集される位置情報付きデータを用いることにより地域メッシュ統計データを作成することが可能である。例えば、佐藤らは[7]、求人広告に含まれる位置情報を用いることで、求人数および業種別求人広告に関する地域メッシュ統計を作成している。

著者らは日本国内でしか利用されていなかったこの地域メッシュコードの上位互換性を有する世界メッシュコードを提案し[5]、全世界を覆う世界メッシュコードの定義方法地域メッシュコード同様に 6階層に渡り提案した。この方法は JIS X0410 での定義域を緯度に対して -90° から 90° まで、経度に対して -180° から 180° まで拡張した方法である。

図 1 は世界メッシュコードの 6階層の模式図を示している。世界メッシュコードでは世界を 8区画に分割し、それぞれ 1 から 8 までの 0次メッシュコードを割り付け、それぞれの 0次メッシュ内に JIS X0410 の 6階層と同じ方法で 1次メッシュから 6次メッシュまでを定義することを行っている。1次メッシュコード(緯度方向 40分、経度方向 1度)は 6桁の整数で表現される。2次メッシュコード(緯度方向 5分、経度方向 7分 30秒)は 8桁、3次メッシュコード(緯度方向 30秒、経度方向 45秒)は 10桁の整数列により表記される。更に、4次メッシュコード(緯度方向 15秒、経度方向 22.5秒)は 11桁、5次メッシュコード(緯

度方向 7.5 秒, 経度方向 11.25 秒)は 12 桁, 6 次メッシュコード(緯度方向 3.25 秒, 経度方向 5.625 秒)は 13 桁の整数列により表示される. 1 次メッシュから 6 次メッシュコードを緯度と経度から計算する関数および 1 次メッシュから 6 次メッシュコードからそのメッシュの角の緯度と経度を計算するためのライブラリを世界メッシュ研究所[8]のページから公開している. 2017 年 2 月執筆現在, ライブラリは R, Javascript, PHP の 3 種類の言語により提供を行っている. 更に, いくつかの世界メッシュ統計・データの例をホームページより公開している.

3 世界メッシュデータの生成方法およびその利活用の可能性

政府統計が整備され先進国では日本と同じようにグリッド統計が作成されて公開されている事例が存在している. しかしながら, 発展途上国ではそもそも統計制度が十分に整備されておらず, 国の様子に関する理解を深めるためには基礎データの整備から必要である場合がほとんどである. そこで, 衛星データを用いた世界メッシュデータの事例と我が国における衛星データから作成したメッシュデータとの相関関係について論じたい.

図 2 は JAXA だいち(ALOS)による 30m 角グリッド標高データ[9]を用いて作成したフィリピンの標高に関する世界メッシュデータである. 3 次メッシュごとの最小標高, 平均標高, メディアン標高, 最大標高の統計としてまとめた. この標高メッシュ統計の作成方法は, 以下ようになる.

(1) JAXA だいち(ALOS)による 30m 角グリッド標高データは緯度方向 1 度, 経度方向 1 度のタイルごとに緯度方向 1 秒, 経度方向 1 秒の 3600×3600 の geotiff データとして提供されている.

(2) タイル内に含まれる 3600×3600 の標高ごとに, 3 次メッシュコードを付与して同じ 3 次メッシュコードを有する標高データ集合について最小値, 平均値, メディアン, 最大値を計算する.

(3) タイルごとの 3 次標高メッシュ統計を各国の行政界 3 次メッシュデータを用いて連結し, 国ごとに標高 3 次メッシュ統計としてまとめる.

標高 3 次メッシュ統計とした. 海上では標高 0 となっているため, 今回海上のデータは標高計算に用いないこととした. 3 次メッシュは緯度方向 30 分, 経度方向 45 分であるので 1350 点が最大標本数となる. 実際はデータ欠損や海上の標高 0 点である理由から 3 次メッシュ内の標本数は 1350 点以下となる. 衛星データは一般に緯度と経度で区切られる画像としてデータの収集と公開が行われているが, 社会科学の文脈で衛星データを分析しようとする場合, 国ごとに集計結果を取りまとめて公開するほうが取り扱いしやすいため, (3) の国ごとにまとめる操作を入れている.

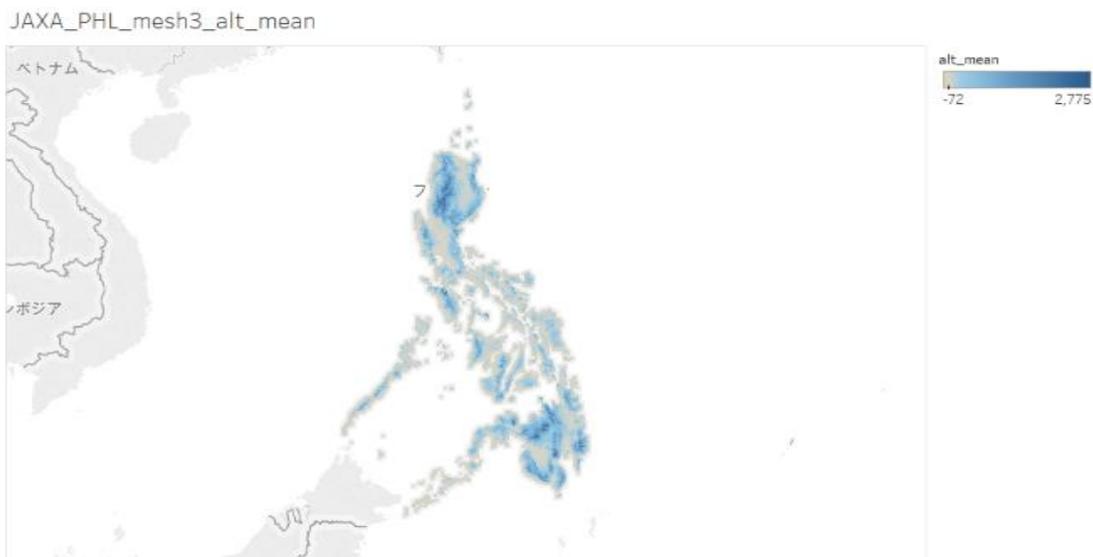


図2 平均標高3次メッシュデータ (フィリピン)

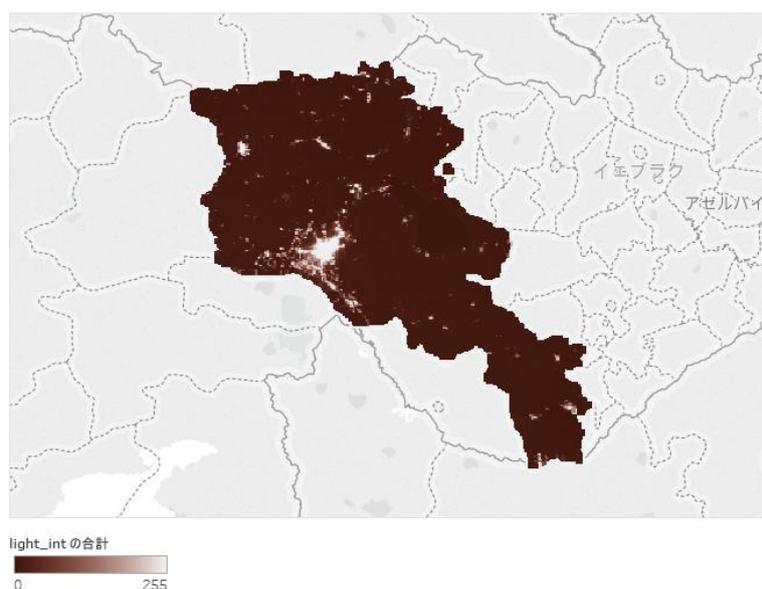


図3 NASA 夜間光データの3次メッシュデータ(アルメニア)

図3はアルメニアにおけるNASAの夜間光データ[10]を世界メッシュデータに変換したものを可視化したものである。NASA夜間光データは緯度方向12秒、経度方向12秒各内の人工衛星から撮影した夜間光強度が0から255の強度値として記録されている。3次メッシュ上では緯度方向 $45/12=3.75$ 個、経度方向 $30/12=2.5$ 個が含まれる。3次メッシュ内中に含まれる光強度の代表値を1つ選び世界メッシュデータとして変換した。

これらのメッシュ統計を作成するために利用した衛星データは、緯度と経度に基づくグリッド的メッシュ構造を有しているが、グリッドの緯度と経度方向に対する幅がまちまちであり、これらのデータを連結して比較、分析するには、共通の枠組みを必要とする。

我々が提案する,世界メッシュコードを用いてデータの変換を行うことにより,一部データの劣化する部分は存在するが,同じメッシュ番号上のデータを連結して分析することが可能である.

社会科学を豊富な計算とデータを用いて研究を行うためには,異なる組織が異なる目的で収集,公開するデータを相互に共有,結合して分析を行う必要がある. 更に,ポイントデータや衛星画像などのデータは個人や企業のプライバシーに抵触することが多いデータであるため,一般にそのままの形式でデータを公開,共有しようとするとは様々な問題が露呈する. この問題を解決するひとつの方法として,各組織は個人情報をも特定できない統計の形で公表を行い,データの共有や結合を行う前に基本的な統合分析ができる環境を整備すべきである. この目的で世界メッシュコードを用いた世界メッシュ統計を作成して,公開し,共有結合して分析することは,グローバル規模で社会活動のデータを結合的に分析できる環境構築につながると考える.

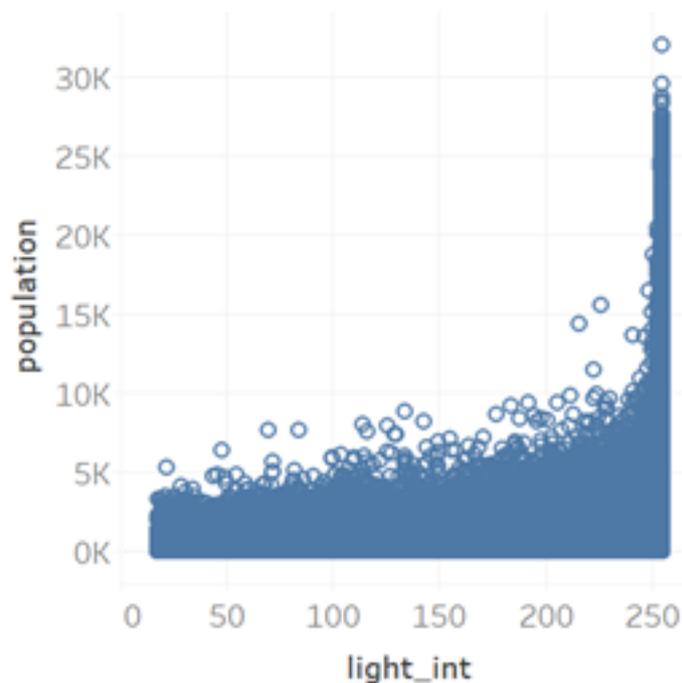


図4 日本における3次メッシュデータ(2010年国勢調査人口と2012年NASA夜間光強度)間の相関関係

図4は2010年総務省統計局国勢調査人口メッシュデータと2012年NASA夜間光強度データ3次メッシュデータの散布図を示している. これらのデータは異なる目的で作成され公開されているデータではあるが,連結分析を行うことにより人口集中地域における人工的に発生する明かりの光量がどの程度であるかの相関関係を調べることが可能である. 人口10,000人以上の2,120個の3次メッシュについて調べたところ,夜間光強度は210以上の値を示すことが分かった. すなわち,人口が10,000人以上居住する3次メッシュで

は夜間光強度は 210 以上の値を日本では必ず示すことを意味する。

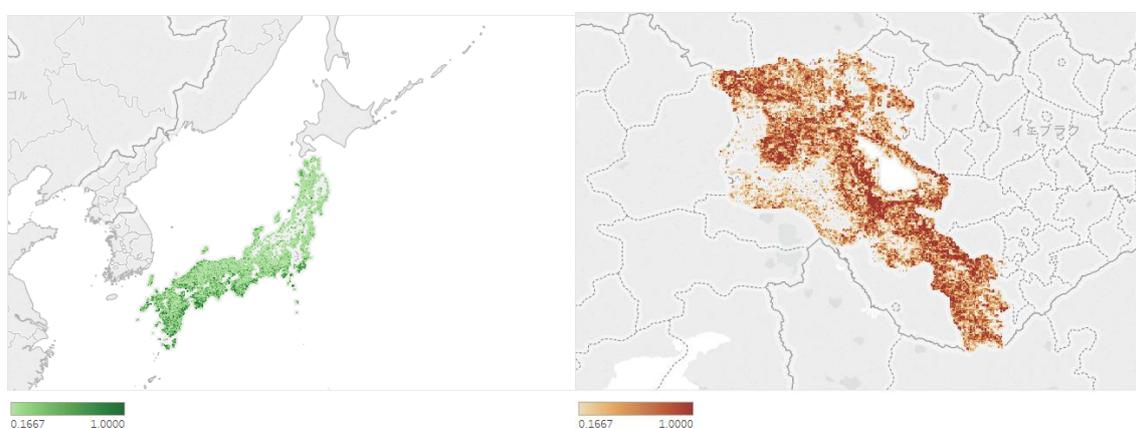


図 5 日本とアルメニアにおける土地利用状況 3 次メッシュ統計データ。(a)日本における常緑広葉樹の割合(b)アルメニアにおける農地の割合

図 5 は衛星画像に基づく土地被覆分類データ[11]を元に 3 次世界メッシュ内の土地利用比率として作成した土地被覆分類 3 次メッシュ統計の例である。[11]で提供される土地被覆分類データには 20 分類(常緑広葉樹, 落葉広葉樹, 常緑針葉樹, 落葉針葉樹, マングローブ, 都市, 水域など)が存在しており、最小分類区画は 15 秒度となっている。そのため、3 次世界メッシュ内では 6 セル分の分類が存在している。

この土地被覆分類 3 次メッシュ統計では、それぞれ 20 種類の分類が各 3 次メッシュでどの程度含まれているかを 1/6 刻みで 0 から 1 の間の割合で表記している。図 4(a)は日本における常緑広葉樹による被覆割合を示している。色の濃いほどその 3 次メッシュ内に常緑広葉樹が存在する比率が高いことを示している。また、図 4(b)はアルメニアにおける農地による被覆割合を示している。異なるデータ源(インターネットから収集されるポイントデータなど)から生成されるメッシュデータと連結することにより、土地利用と社会経済活動との関係の分析に利用することが可能である。

更に、図 2 で示した標高 3 次メッシュ統計データと津波上陸カタログデータを用いることで、佐藤ら[12]の方法で提案した一般化パレート分布を用いた津波上陸ハザードの計算に世界メッシュコードを用いることにより、JIS X0410 で定義されている範囲以外の国であっても、津波ハザードの計算が可能である。図 6 は NOAA が公開する過去 1000 年分の津波上陸カタログデータ[13]を用いて、フィリピンにおける津波上陸のハザードを計算した結果である。最低標高に到達するハザード、最大標高まで到達するハザードをそれぞれ、JAXA だいち(ALOS)の 30m 標高データから作成した標高 3 次メッシュ統計を用いて算出した。

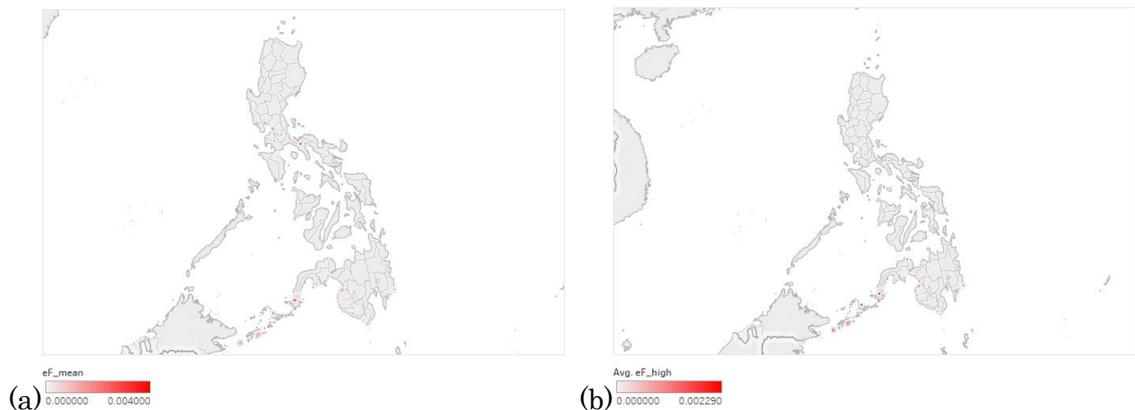


図 6 フィリピンにおける津波上陸ハザード 3 次メッシュデータ. (a)最低標高到達ハザード, (b)最高標高到達ハザード.

4 まとめ

本稿では、世界メッシュコード提案の経緯と世界メッシュ統計について概説するとともに、この世界メッシュコードを利用するためのライブラリを紹介した。世界メッシュコードを利用することにより、インターネット上で収集される位置情報付きデータから世界メッシュ統計を作成できるのみならず、衛星画像や衛星計測データをもとに世界のあらゆる場所における特徴を空間統計として共通の枠組みにより集出し公開することが可能である。

共通の世界メッシュコードを用いることで、異なる組織、異なる目的で収集されるデータを統計化（劣化）させることで、秘匿化しデータ量を減らしつつ空間の特徴を捉え、比較・結合を通じてデータマイニングや知識発見の目的で利用できる。

本稿では、特に JAXA が提供するだいち(ALOS)標高データ、NASA が提供する夜間光データ [9]、International steering Committee for Global Mapping が提供する土地利用データをデータ源として世界メッシュ統計が作成できることを示した。さらに、我が国の 2010 年国勢調査人口 3 次メッシュデータと NASA が提供する夜間光データから作成した 3 次メッシュデータを連結することにより、人口と夜間光強度との間の相関関係について調べた。その結果、3 次メッシュ内に人口 1 万人以上が確認される場所の夜間光強度は日本の全ての場所において 210 以上の値を示すことを確認した。

今後の課題として、世界メッシュ統計を大量に作成した場合における、メタデータ付与に関する問題が挙げられる。この問題は、近年普及がすすむ LOD および経済産業省が行う共通語彙基盤(IMI)と連動することにより、メタデータの世界標準化の動きの中で解決していくべき課題である。さらに、作成した世界メッシュ統計を相互に連結することにより、空間の特徴を自動的に判定し、政策意思決定者は理解できていないが、社会的問題となっている課題の発見とその解決に貢献する社会設計への貢献方法が挙げられる。この社会設計への貢献ならびにそれを支援するためのアプリケーションの開発が必要である。筆者の一人（佐藤）は世界メッシュ統計データを利用するための Web アプリケーションとしてグロ

ーバル・システムの持続可能性評価基盤(MESHSTATS)を開発し、一部の機能を公開している[14]。このような世界メッシュ統計を利用するアプリケーションが今後発展していくことにより、匿名性の高い状態で個人のプライバシーを配慮した社会システムの分析とその社会実装が可能となると期待する。本研究は、科学技術振興機構（JST）平成 27 年度戦略的創造研究推進事業（さきがけ）の研究成果を含む。本研究は 2016 年度京を中核とする HPCI システム課題研究（hp160060）（統計数理研究所）の計算資源を用いて実施された。

参考文献

- [1] 総務省統計局地域メッシュ統計の概要[10 August 2015]
<http://www.stat.go.jp/data/mesh/gaiyou.htm>
- [2] e-Stat, <https://www.e-stat.go.jp/SG1/estat/eStatTopPortal.do>
- [3] 地図で見る統計（統計 GIS）
<https://www.e-stat.go.jp/SG1/estat/toukeiChiri.do?method=init>
- [4] 国土交通省国土政策局国土情報課
<http://www.mlit.go.jp/kokudoseisaku/kokudojoho.html>
- [5] 佐藤彰洋, 椿広計, ビッグデータ時代に必要な標準化, 統計, 2015 年 9 月号 (2015) pp.32-38.
- [6] ドコモモバイル空間統計,
https://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/
- [7] Aki-Hiro Sato, Chihiro Shimizu, Takayuki Mizuno, Takaaki Ohnishi, Tsutomu Watanabe, "Relationship between job opportunities and economic environments measured from data in internet job searching sites", *Procedia Computer Science*, Volume 60 (2015) pp. 1255—1262
- [8] 世界メッシュ研究所, <http://www.fttsus.jp/worldgrids/>
- [9] ALOS 全球数値地表面モデル (DSM) "ALOS World 3D - 30m" (AW3D30)
- [10] NASA, VISIBLE EARTH, <http://visibleearth.nasa.gov/view.php?id=55167>
- [11] Land Cover, International steering Committee for Global Mapping,
<https://www.iscgm.org/gmd/>
- [12] 佐藤彰洋, 榎嶋弘樹, Tae-Seok Jang, 澤井秀文, "経済社会データおよび環境データを用いた空間評価指標の大規模計算:地域メッシュ統計の利活用", 横幹, 第 10 巻第 2 号 (2016) pp. 76-83
- [13] NGDC/WDS Global Historical Tsunami Database, tsunami runup data,
doi:10.7289/V5PN93H7
- [14] グローバルシステムの持続可能性評価基盤（統計情報可視化システム MESHSTATS）,
<https://www.meshstats.xyz/meshstats/>