The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume L-4/W1-2026
Free and Open Source Software for Geospatial (FOSS4G) 2026 – Academic Track "Bridging Geospatial Technology and Humanity",
30 August–5 September 2026, Hiroshima, Japan

Spatiotemporal Analysis of OpenStreetMap Editing Activities in Japan Using the OSMCha

Toshikazu Seto [1,2]

[1] Department of Geography, Komazawa University, Tokyo, Japan – tosseto@komazawa-u.ac.jp
[2] 3D Geoinformation group, Department of Urbanism, Faculty of Architecture and Built Environment,
Delft University of Technology, Delft, The Netherlands

Keywords: OpenStreetMap, quality assessment, OSM changeset analyzer, spatiotemporal analysis, contributor behaviour

Abstract

OpenStreetMap (OSM) is a prominent platform for Volunteered Geographic Information (VGI), wherein geographic data are updated daily by a global community of contributors. The OpenStreetMap Changeset Analyzer (OSMCha) serves as a quality assurance tool that automatically flags suspicious edits based on detection rules covering geometric and tag plausibility, edit scale, and contributor behavioral patterns. In this study, we developed Python scripts to systematically collect 740,038 changesets spanning four years (2022–2025) for Japan via the OSMCha API, consolidizing them into a FlatGeobuf database with 60 reason_id mappings embedded as attributes and prefecture-level spatial joins applied.
Our analysis revealed a 63.3% increase in annual changesets and a 72.5% expansion in unique contributors, alongside an 11.9-percentage-point decline in the suspicious edit rate. Mobile editors and survey-based edits grew rapidly, consistently demonstrating lower suspicion rates than non-survey edits. A structural shift in the OSMCha detection logic was empirically identified in 2024. The KDE analysis confirmed editing hotspots in three major metropolitan areas, while the January 2024 Noto Peninsula earthquake triggered concentrated crisis mapping, engaging 1,536 contributors.

1. Introduction

Since its establishment in 2004, OpenStreetMap (OSM) has grown to encompass more than ten million registered contributor accounts worldwide, and its use has expanded across diverse fields, including international disaster response and humanitarian support, municipal open data initiatives, and geospatial infrastructure development. In Japan, more than 35,000 contributors have participated (Seto, 2022), and a distinctive mapping culture has emerged that draws on a variety of data sources, including base maps and aerial imagery provided by the Japanese government and the 3D city model of Project PLATEAU (Seto et al., 2023). In addition, the recent proliferation of lightweight editors for smartphones and the rise of AI-assisted mapping have led to rapid diversification of contributor profiles and editing methods.

As the social importance of VGI continues to grow, the question of how to evaluate, maintain, and improve data quality remains a central concern in the OSM research. Previous studies have proposed a wide range of approaches, including the development of OSHDB and ohsome-API as platforms for spatiotemporal analysis of edit histories (Raifer et al., 2019), freshness evaluation based on update frequency (Minghini and Frassinelli, 2019), analysis of corporate contributors' activities (Anderson et al., 2019), and investigation of spatial biases in editing areas based on contributor discussions (Seto et al., 2020). These studies demonstrate that OSM quality assessment has moved beyond simple comparisons of positional accuracy or building geometry, entering a stage where it is discussed from multiple perspectives, encompassing contributor behavior, the use of AI technology, and the detection of malicious edits.

In a preceding study (Seto, 2024), the author analyzed 172,143 changesets recorded in Japan over one year in 2023, examining trends in suspicious editing detected using the OSM Changeset Analyzer (OSMCha). However, several issues could not be fully addressed within such a limited timeframe: (1) the extent to which changes in OSMCha detection rules influenced observed trends; (2) how the community balances an influx of new contributors with continued editing by existing ones and what effect AI-assisted editing has had; and (3) the scale of VGI community mobilization in regions outside major cities, particularly in disaster-affected areas. To address these questions, the present study extends the analytical period to four years (2022–2025), the full range for which data are available, and conducts a comprehensive analysis of all 740,038 changesets recorded in Japan.

2. Related Work

Early OSM quality research focused primarily on comparing positional accuracy and attribute completeness with commercial maps and administrative data. Kanasugi et al. (2019) evaluated the positional accuracy and completeness of OSM road data in Japan by comparing it with a digital road map. However, since the latter half of the 2010s, research interest has shifted from the data themselves to contributor behavior and the editing process. Quattrone et al. (2015) distinguished two groups in OSM—"power users" (fewer than 20% of contributors responsible for more than 80% of edits) and "the crowd" (the majority of occasional contributors)—and proposed a method for quantitatively evaluating the divergence (bias) between these groups in terms of mapping content along three axes: what is edited, where it is edited, and how meticulously it is edited, using data from 40 countries over three years. Their analysis revealed that while there is no notable divergence between the two groups in the types of features edited (what), there is a strong mismatch in the spatial distribution of edits (where), with power users tending to concentrate their activity in metropolitan areas and the crowd tending to be more widely distributed. Furthermore, they showed that editing meticulousness (how meticulous) correlates with national cultural indicators such as the Power Distance Index and Individualism. Anderson et al. (2019) quantified the growing share of corporate contributors in OSM using OSM-QA tiles, highlighting the diversification of actors in VGI. Li et al. (2021) attempted to improve vandalism detection accuracy using user embeddings, demonstrating the effectiveness of contributor-level behavioural analysis.

Tools for evaluating editing behavior have evolved. Raifer et al. (2019) proposed the OSHDB framework and the ohsome-API for spatio-temporal aggregation of all OSM historical data, providing a foundation for international comparative research. Minghini and Frassinelli (2019) developed the "Is OSM up-to-date?" tool for assessing the freshness of OSM. These represent mechanisms for the aggregative evaluation of edit histories and are positioned as a distinct approach from OSMCha, which is used in the present study.

Two particularly important recent developments in the OSM landscape deserve to be mentioned. The first is the full-scale introduction of AI-assisted mapping (Andorful et al., 2026), wherein semi-automatic mapping workflows, where contributors approve and modify automatically estimated features from satellite imagery, have become widespread. The RapiD editor and MapwithAI services, as well as fAIr, present contributors with candidate buildings and roads estimated primarily from satellite imagery. While enabling large-scale mapping, especially in developing countries, the accuracy of automatically estimated features is not uniform. The second development is the diversification of the contributor community and the evaluation of sustained participation by new contributors in the community. The proliferation of lightweight editors such as StreetComplete and EveryDoor, which run on smartphones, has rapidly increased micromapping based on field surveys. Tracking the long-term impact of these changes on editing behavior and quality is an indispensable challenge for considering VGI sustainability. Choe et al. (2023) analyzed conflict and its management among contributors in such peer-production environments, providing a perspective complementary to the quantitative approach of the present study.

OSMCha, used in this study, is a quality verification tool with a widely used web UI and API in the OSM community. It is characterised by its per-changeset approach, applying primarily algorithm-based judgements to flag edits as "suspicious." Specific detection conditions include "possible import" when created elements exceed 70% of all edits, "mass deletion" when more than 1,000 elements are deleted, "New mapper" for users with fewer than five edits or fewer than five days of mapping, "User has multiple blocks" for edits by users who have been blocked by the OSM Foundation multiple times, and "suspect_word" for changesets containing specific terms (e.g., "import," "google," "pokemon") in the comment field. Although these detection logics are publicly available, the operational rules have changed over time, and analyses comparing detection results across years require awareness of these changes.

3. Methodology and Data

3.1 Overview of OSMCha

OSMCha is an open-source Python package and web interface (Figure 1) for aggregating and visualizing, on a per-changeset basis, the editing activity and spatial extent of OSM changesets (https://github.com/OSMCha/osmcha). A changeset is a logical unit comprising a series of edits (creation, modification, deletion, and tag assignment) made by a single user within a short period; it is automatically closed either when the user explicitly ends editing or after one hour of inactivity. The maximum number of edits per changeset was 10,000 elements, and the maximum editing duration was 24 h. OSMCha applies a multidimensional set of rules to these changesets—based on edit counts, element counts, tag content, comment strings, and feature geometry—and assigns a "suspicious" flag (is_suspect) together with the detection reasons (reasons). The detection rules compiled for this study total 60 categories, which are broadly divided into: (i) rules concerning the plausibility of feature geometry and tags (e.g., crossing ways, disconnected way, invalid tag modification); (ii) rules concerning edit scale (e.g., mass deletion, mass modification, possible import); (iii) rules concerning contributor attributes and behavior (e.g., new mapper, user has multiple blocks, suspect_word); and (iv) rules concerning edits to specific feature types (e.g., edited a major road, motorway/trunk geometry modified, park added by new user).

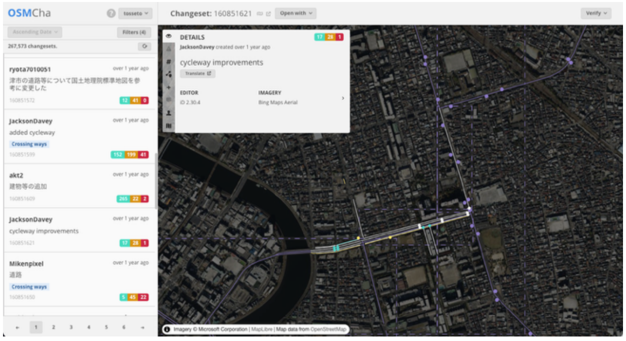

Figure 1. OSMCha interface. (https://osmcha.org/)

3.2 Data Collection and Analysis Items

In this study, changesets targeting areas within Japan were collected in GeoJSON format via the OSMCha API at monthly intervals spanning four years from January 1, 2022, to December 31, 2025. Because the API imposes constraints on the number of records retrievable per request and on timeouts, the developed Python scripts combined monthly batch processing with GeoJSON storage, a checkpoint function enabling resumption from interruption points, and a validation logic that issues a warning when the number of records acquired falls below 90% of the expected total for a given month, thereby achieving stable large-scale data collection. Although the raw data were provided as polygon data representing the editing extent, they contained numerous geometry errors; accordingly, the data were expressed as centroid points incorporating the area value as an attribute and integrated into a single FlatGeobuf database.

The database contains the following primary attributes: contributor name, contributor UID, editor used, comment, data source, imagery used, edit timestamp, creation count, modification count, deletion count, editing area, suspect flag, detection reason IDs, harmful flag, review status, and tag information. Regarding annual data availability, detection reason text (reasons) is available only for 2022; from 2023 onward, only reason IDs (reason_ids) are stored. Considering this difference, the present study cross-referenced the reason field for 2022 with the definitions in the OSMCha public repository to compile mappings for all 60 reason_ids and their corresponding names. The total number of detection occurrences over the four years was 339,114.

The following analyses were conducted using the integrated database. First, the basic monthly trends were characterized. Second, the annual frequency of detection of reason IDs was compared, aggregating both sole and co-occurring detections. Third, changes in the composition of the editors and data sources were tracked. Fourth, contributor continuity was analyzed by tabulating the number of users by years of activity across the four-year period. Fifth, the spatial distribution was evaluated using kernel density estimation (KDE) in QGIS to extract hotspots at the national scale and to compare the editing volume and suspect rates by prefecture.

https://doi.org/10.5194/isprs-archives-L-4-W1-2026-275-2026 |

4. Results

4.1 Growth and Monthly Variation in Editing Activity

The total number of changesets recorded in Japan over the four years was 740,038, representing an increase of approximately 63.3% from 144,625 in 2022 to 236,203 in 2025. The number of unique contributors also increased by 72.5%, from 3,660 in 2022 to 6,315 in 2025, indicating steady expansion in contributor participation. The total number of edits—the sum of created, modified, and deleted elements—reached approximately 142.77 million over four years, growing from approximately 28.59 million in 2022 to approximately 42.99 million in 2025.

Examining the monthly trend (Figure 2), the number of creations expanded approximately 1.9-fold, from approximately 2.15 million in January 2022 to approximately 4.15 million in December 2025, indicating a steady growth in the volume of geographic data generated. January 2024 recorded a four-year maximum of approximately 5.39 million creations, strongly reflecting the crisis mapping associated with the Noto Peninsula Earthquake, which is described later. The suspect count in the same month reached 15,385, the highest in the four-year period, confirming that the expansion in overall editing activity and the increase in suspect-flagged edits are linked along the time axis. The number of modifications peaked at approximately 1.16 million in January 2023 and subsequently remained in the range of 600,000–900,000, showing a more gradual variation compared with the growth rate of creations. Deletion counts generally remained stable in the range of 150,000–300,000; however, a notable outlier of approximately 1.16 million was recorded in March 2022. This is presumed to have been caused by large-scale batch deletions (including revert operations) by a specific contributor and is designated as a subject for future analyses.

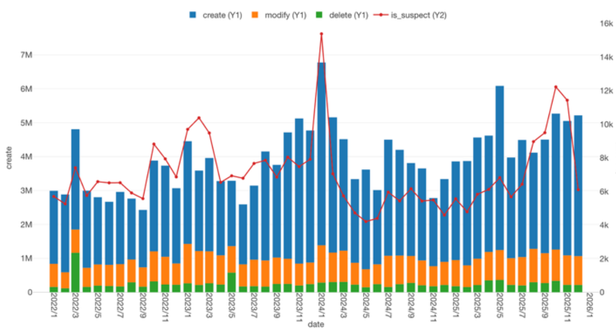

Figure 2. Monthly OSM edits in Japan from 2022 to 2025.

The suspect rate, which stood at 40.7% in both 2022 and 2023, fell to 30.8% in 2024 and 28.8% in 2025, a decline of approximately 11.9 percentage points. However, in the fourth quarter of 2025, monthly figures rose again—12,224 in October and 11,430 in November—suggesting that the annual average decline does not simply signify "quality improvement." There are two possible interpretations of the declining suspect rate. The first is genuine quality improvement through the accumulation of contributor experience and the diffusion of community norms. The other is an apparent change attributable to revisions in the OSMCha detection rules. The analysis of the detection reason composition presented below indicates that the latter influence is dominant in this study.

4.2 Annual Changes in Detection Reason Composition

Tabulating the frequency of detection reason ID occurrences (total occurrences, including both sole and co-occurring detections), a major structural change was confirmed around 2024 (Table 1). Of the total 339,114 detection reason occurrences across the four years, the top five—"New mapper", "Crossing ways”, "possible import”, "Review requested" and "Invalid tag modification"— accounted for approximately 70% of the total. The two categories dominant throughout the entire period were new mapper and Crossing ways. In particular, the number of new mappers increased by approximately 52%, from 17,356 in 2022 to 26,414 in 2025, constituting approximately 24.4% of the total detections across the four years and becoming the largest segment. Crossing ways also recorded 20,144 occurrences in 2025, establishing itself as a major component, second only to new mapper. Detection reasons outside the top categories (Others) remained generally in the range of 10,000–13,000 per year, indicating that diverse factors not concentrated in specific reason IDs continue to exist at a certain scale.

In contrast, "Invalid tag modification" and "Motorway/trunk geometry modified," which had been detected frequently in 2022 and 2023, virtually disappeared from 2024. The former occurred 10,574 times in 2022 and 11,148 times in 2023, declining to 863 times in 2024 and zero times in 2025. The latter showed a similar pattern, falling from 10,836 in 2023 to 482 in 2024, and zero in 2025. Conversely, "User has multiple blocks " surged from 2024 onward: from 541 in 2022 and 177 in 2023 to 3,069 in 2024 and 3,710 in 2025. In addition, "suspect_word" increased by approximately 10.5-fold, from 387 in 2022 to 4,069 in 2025. These changes strongly suggest that OSMCha's detection logic shifted around 2024 from checking the quality of geographic data itself—"the plausibility of feature geometry and tags"—toward monitoring user-side behavioural patterns—"user behaviour and comment strings." Furthermore, "Review requested " increased from 3,151 in 2022 to 8,600 in 2025, indicating the growth of mutual review within the community.

reasons_id	reasons	2022	2023	2024	2025	Total
40	New mapper	17,356	16,607	22,397	26,414	82,774
517	Crossing ways	15,275	18,409	15,865	20,144	69,693
2	possible import	6,958	8,353	9,411	9,389	34,111
86	Review requested	3,151	7,973	5,770	8,600	25,494
42	Invalid tag modification	10,574	11,148	863		22,585
91	Motorway/trunk geometry modified	7,612	10,836	482		18,930
526	Outdated tags	2,537	3,559	3,055	4,411	13,562
521	Disconnected way	2,488	2,570	2,838	2,251	10,147
4	mass modification	1,879	1,771	1,807	2,383	7,840
1	suspect_word	387	1,149	2,190	4,069	7,795
83	User has multiple blocks	541	177	3,069	3,710	7,497
516	Almost junction	1,676	1,414	1,610	1,823	6,523
520	Very close points	536	832	1,352	1,533	4,253
3	mass deletion	1,084	1,157	918	1,068	4,227
43	Invalid key value combination	1,046	2,302	214		3,562
57	Feature overlaps with existing features	1,422	1,199	131		2,752
531	Mapbox: Overlapping features	258	422	548	612	1,840
528	Impossible oneway	496	453	435	438	1,822
530	Mapbox: Incorrect mapping	66	133	252	895	1,346
527	Mismatched geometry	254	289	388	356	1,287
31	Deleted a wikidata/wikipedia tag	513	657	37		1,207
44	Park added by new user	672	484	30		1,186
489	Mapbox: Spam text	197	246	228	398	1,069
—	Others	2,038	3,589	921	1,064	7,612
	Total	79,016	95,729	74,811	89,558	339,114

Table 1. Occurrence counts by major detection reason IDs. (only reason IDs with ≥1,000 occurrences in aggregate)

4.3 Diversification of Editors and Data Sources

Examining the composition by editor (Table 2), web-based iD accounted for 63.1% (467,257 changesets) across the entire period, maintaining its position as the dominant editor, while its suspect rate also remained high at 41.25% against the overall average of 34.39%, exerting an extremely large influence on the platform's overall quality indicator. Desktop-based JOSM accounted for the second largest share at 19.8% (146,288 changesets); notably, its suspect rates improved markedly from

https://doi.org/10.5194/isprs-archives-L-4-W1-2026-275-2026

41.0% in 2022 to 28.2% by 2025. Another major characteristic is the substantial expansion in the use of mobile lightweight editors. For example, StreetComplete grew approximately 4.5-fold from 5,709 changesets in 2022 to 25,576 in 2025, and EveryDoor exhibited a 10.6-fold increase from 539 to 5,692 changesets. StreetComplete's suspect rate was extremely low at 4.0%. Regarding AI-assisted editing, only RapiD appeared in the upper rankings, recording a peak of 2,558 changesets in 2023 before declining slightly to approximately 1,700 annually thereafter. Although the overall average suspect rate for RapiD is 37.54%, it fell from 59.11% in 2023 to 15.97% in 2025, indicating a certain degree of quality improvement in mapping.

reasons	2022	2023	2024	2025	Total	suspect (%)
iD	99,753	106,601	112,286	148,610	467,250	41.25
JOSM	28,033	34,088	40,453	43,714	146,288	33.08
StreetComplete	5,709	13,153	18,015	25,576	62,453	5.23
Go Map!!	5,416	7,875	6,355	5,161	24,807	11.11
Every Door	539	2,616	4,539	5,692	13,386	9.78
RapiD	209	2,670	1,659	1,739	6,277	37.54
Automated	670	2,983	1,199	987	5,839	3.78
Vespucci	992	726	729	1,122	3,569	17.12
Organic Maps	214	279	439	1,181	2,113	27.17
Level0	260	171	262	1,168	1,861	6.13
MapRoulette	1,265	4		2	1,271	47.21
OsmAnd	313	185	229	223	950	14.11
Merkaartor	159	154	181	200	694	17.87
reverter_plugin	131	195	206	98	630	53.02
MAPS.ME	182	119	62	12	375	42.40
osmapi	19	94	98	158	369	7.59
Others	761	230	355	560	1,906	39.27
Total	144,625	172,143	187,067	236,203	740,038	34.39

Table 2. Annual changes in the OSM editors and suspect rates.

4.4 Quality Characteristics of Edits and Contributor Trends

Regarding data source usage, "survey" was the most frequently used term, appearing 97,919 times over the entire period; an increase from 16,952 in 2022 to 47,619 in 2025. PLATEAU-derived edits also expanded 4.3-fold from 927 in 2022 to 3,970 in 2025, indicating growing activity in reflecting administratively derived open data in OSM.

A clear difference in the suspect rates was observed between the survey-based edits and all other edits (Figure 3). The suspect rate for survey-derived edits (124,255 changes sets) was 15.4%, substantially lower than the 38.2% for non-survey edits (615,783 changes sets). The annual trend shows that the suspect rate for survey-derived edits fell from 25.3% in 2022 to 10.6% in 2025, which is less than half the starting value, confirming that field-survey-based edits are consistently sound in quality. The average number of edits per changeset for survey-derived edits was also small and stable at 39–71, suggesting that a micro-mapping-oriented editing style has been established.

A trend toward quality improvement is also evident in AI-assisted editing of manuscripts. The suspect rate for RapiD-based editing fell from 59.11% in 2023 to 15.97% in 2025, indicating a degree of quality improvement. This is thought to reflect the fact that, whereas in the early trial phase, large-scale, wide-ranging automated estimates were easily detected as "possible import" or "mass modification," from 2024 onward, a workflow in which contributors more carefully scrutinize and correct AI-estimated results has become established.

Examining the composition of users by years of activity (Figure 4), of the 14,978 contributors in total, 84.7% (12, 689 contributors) were active for only one year, and only 433 contributors (2.9%) sustained activity across the entire four-year period. In contrast, the top 15 contributors (all four-year continuers) accounted for 120,308 changesets, representing 16.3% of the overall total (Table 3), confirming a structure in which a small number of continuing contributors make an oligopolistic contribution.

The suspect rates of the top contributors varied widely across individuals, ranging from 0.9% (contributor G) to 96.5% (contributor N). N's changesets suggests that this is an account created exclusively for importing; while accounting for approximately 13.1% of the total created elements, it has the structural characteristic of being automatically flagged as suspicious. The top contributors accounted for about 19% of overall edits, not necessarily an oligopolistic situation, but in 2023 alone, they accounted for 25% of the total, indicating a dependent relationship on a small number of core contributors.

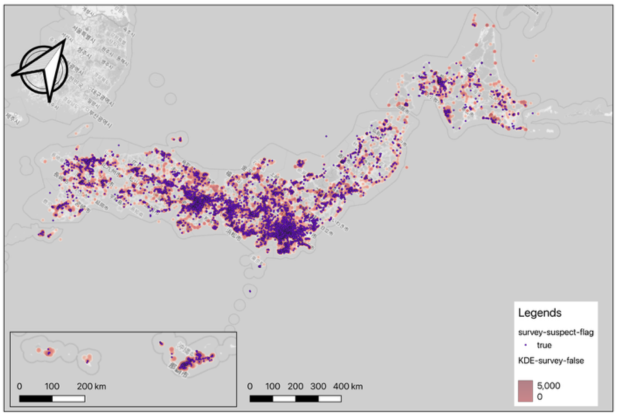

Figure 3: Distribution of survey-based changeset.
(KDE representation of changesets flagged as suspect: 19,156, and those flagged as false: 105,099)

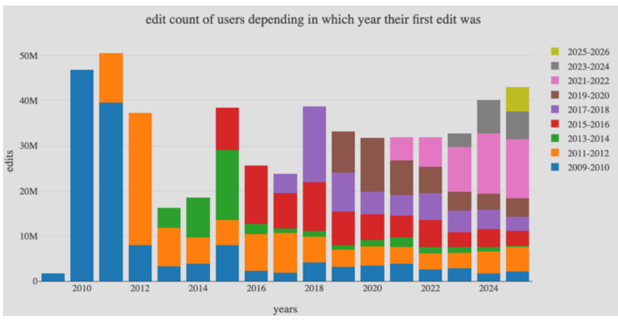

Figure 4. Distribution of contributors by years of activity.

contributors	2022	2023	2024	2025	Total	suspect (%)
A	2,719	4,212	4,694	3,299	14,924	73.52
B	997	3,418	4,130	3,831	12,376	8.10
C	2,750	2,868	2,737	3,430	11,785	54.81
D	819	2,626	2,851	2,362	8,658	9.25
E	1,250	2,064	3,146	2,027	8,487	33.50
F	1,671	2,686	1,410	1,635	7,402	36.31
G	1,439	1,631	2,422	1,878	7,370	0.90
H	3,600	1,815	1,010	815	7,240	24.43
I	1,368	2,091	1,332	2,297	7,088	35.09
J	363	4,089	878	1,206	6,536	35.31
K	327	2,522	1,691	1,320	5,860	3.16
L	330	1,548	1,090	2,889	5,857	57.44
M	352	788	2,070	2,500	5,710	8.51
N	90	851	1,291	3,332	5,564	96.53
O	1,381	2,101	760	1,209	5,451	60.83
Others	120,600	136,184	155,555	202,173	614,512	33.87
Total	144,625	172,143	187,067	236,203	740,038	34.39

Table 3. Transition in changeset counts for top 15 contributors.

4.5 Spatial Distribution of Editing Activity

KDE analysis (Figure 5) confirmed that editing activity hotspots are concentrated in the three major metropolitan areas (Tokyo, Osaka, and Nagoya), consistent with the findings of prior research. At the prefecture level, Tokyo and Aichi show outstanding volumes of creative activity. Tokyo recorded approximately 3.38 million creations in 2022, with the suspect count reaching 8,901, among the highest volumes throughout the entire period. In the same year, deletions in Tokyo reached approximately 778,000, suggesting the existence of batch deletion processes under specific conditions. In Aichi, creation counts reached approximately 6.60 million in 2024, an extreme spike of approximately 4.6-fold compared with the approximately 1.42 million in the previous year. This suggests the presence of a specific large-scale import or organized building data improvement initiative, which has also influenced national trends.

In contrast, the geographic distribution of the suspect rates presents a different picture from the edit volumes. In Iwate, 59.3% of edits are flagged, and in Kumamoto, 55.4%—with high proportions of suspicious edits in areas outside the major cities. For example, in Kumamoto, approximately 1.72 million creations in 2025 were accompanied by 13,711 suspect flags, indicating a rapid increase in the anomaly detection rate. In contrast, Kyoto (20.3%) and Oita (20.6%) showed low suspect rates. These differences are thought to stem from structural factors, including the degree of dependence on a small number of contributors in rural areas and the spatial concentration of AI-assisted editing and imports.

Furthermore, in the crisis mapping associated with the January 2024 Noto Peninsula earthquake (Figure 6), edits in Ishikawa and Toyama Prefectures accounted for approximately 45% of all national edits in that month, with nationwide changeset counts reaching 30,228—nearly double those of adjacent months—and 1,536 unique contributors participating within Japan. A layer showing the KDE distribution (search radius: 1 km) of changesets throughout the full year 2024 is also overlaid on this figure; it is evident that activity was centered on the coastal areas of Ishikawa, where the earthquake damage was greatest, before expanding into Toyama, which experienced the second-greatest damage from March onward. Suspect-flagged changesets (true) totalled 13,633 over the two months, slightly fewer than the false-flagged total (16,341), and the average number of features created per edit (approximately 160) was relatively low compared with the surrounding months (average 250), indicating that editing was conducted with care. Regarding suspect flags, more than 7,000 changesets were assigned the new mapper flag, confirming new participation beyond the regular contributor. This is thought to reflect task allocation functioning through the "HOT Tasking Manager" and demonstrates the high mobilization capacity of the VGI community during disasters.

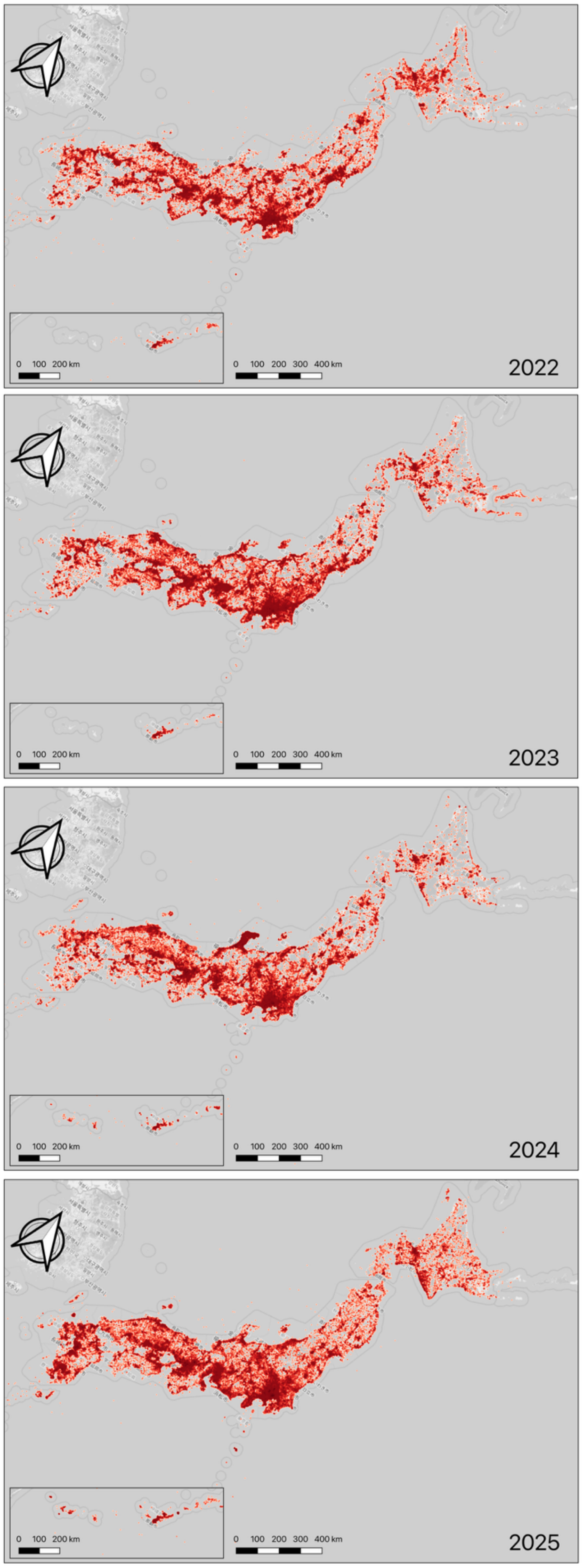

Figure 5. Hotspot analysis of OSM edits at the national scale using kernel density estimation (KDE).

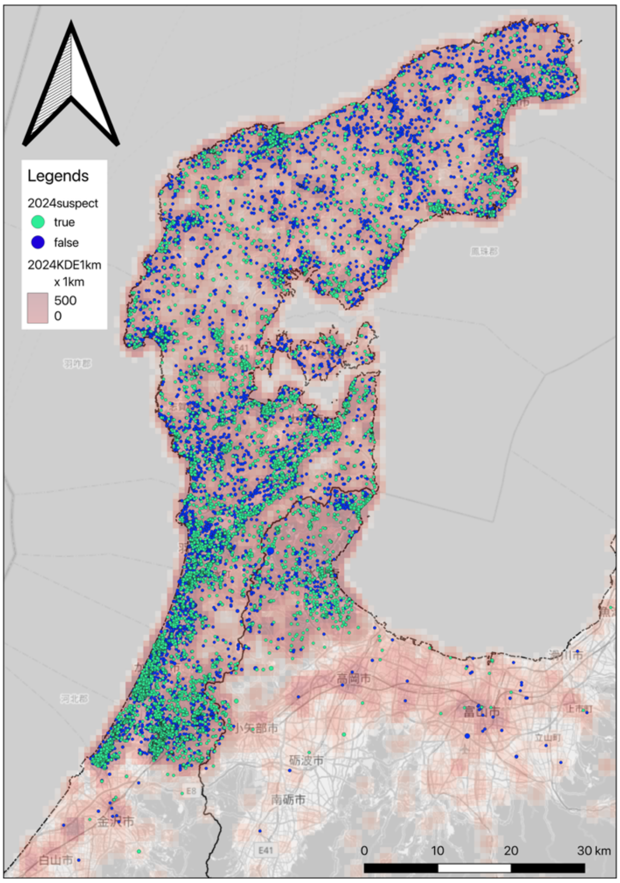

Figure 6. Spatial distribution of crisis mapping during the two months following the Noto Peninsula earthquake in Jan. 2024.

5. Discussion

5.1 VGI Quality Based on OSMCha Detection Rules

The most noteworthy finding of this study is that the analysis spanning 2022–2025 revealed how changes in OSMCha's detection rules have altered the quality indicators used for OSM. As shown in Section 4.2, reason_id=42 (Invalid tag modification) and reason_id=91 (Motorway/trunk geometry modified), which had fired at high frequency in 2022–2023, together accounted for approximately 20,000 detections per year; these decreased sharply in 2024 and fell to zero in 2025. The detection reasons that increased in their place were reason_id=83 (User has multiple blocks) and reason_id=1 (suspect_word), conditions based not on the edited feature itself but on the contributor's history of blocks and the text of the comment field. Mapping onto the four-category taxonomy of detection rules organized in Section 3.1; (i) feature geometry/tag plausibility, (ii) edit scale, (iii) contributor attributes/behaviour, and (iv) specific feature editing, the shift in the detection system's center of gravity from (i) to (iii) is clearly discernible.

What this transition implies is that the very definition of VGI quality is expanding from "the correctness of the data itself" to "the correctness of the process by which the data are produced." Viewed in the light of Li et al. (2021), who showed that user embeddings are effective for vandalism detection, the post-2024 indicator change in OSMCha can be interpreted as incorporating the contributor's behavior as an evaluation axis to address grey zones. Such as edits that are externally normal but systematically inappropriate that cannot be fully captured by feature-based automated checking alone. As Andorful et al. (2026) raised concerns about transparency and trust in AI-assisted editing, automatically estimated features may appear plausible in shape and tags alone; however, their provenance and the validity of the editing process require separate verification. The change in OSMCha's design philosophy is consistent with this broader shift in the concerns of VGI research.

However, the change in detection logic introduces discontinuity in the long-term comparison of VGI quality indicators. The decline in the suspect rate over four years observed in this study (from 40.7% to 28.8%) cannot be interpreted as a straightforward quality improvement. Of the 60 confirmed detection rules, those that appear more than 1,000 times per year are concentrated in approximately a dozen, with the remaining rules existing as "latent detection criteria" (Table 1). Whereas existing evaluation tools such as OSHDB and ohsome-API (Raifer et al., 2019) and "Is OSM up-to-date?" (Minghini and Frassinelli, 2019) aim for cumulative aggregation of edit histories, OSMCha operates as a collection of event-driven detection rules and has the peculiar characteristic that VGI quality cannot be correctly evaluated unless the year-on-year changes in the detection rules are included as an object of analysis. For long-term VGI quality monitoring, a mechanism for continuously recording the operational reality of the rules, that is, which rules are detected at what frequency, is indispensable.

In addition, alongside this paradigm shift in automated detection, it is also noteworthy that a culture of mutual peer review is growing within the community. The occurrence of review requests (reason_id=86) increased approximately 2.7-fold, from 3,151 in 2022 to 8,600 in 2025 (Table 1). This represents contributors proactively registering their edits as objects of review, indicating that detection logic is not completed by algorithms alone but is moving toward integration with a human feedback loop. The conflict management practices in peer-production environments that Choe et al. (2023) analyzed qualitatively are consistent with the quantitative increase in Review requested observed in this study; the combination of automated systems and human review is expected to become the basic configuration of future VGI governance. The FlatGeobuf dataset planned for release in this study contains the structural changes of detection rules in a form traceable through reason_id attributes and can also function as a foundation enabling meta-research on VGI quality monitoring.

5.2 What is the Sustainability of VGI through OSM Contribution?

Another major contribution of this study is that it demonstrates that OSM editing activity in Japan is diversifying simultaneously along three axes—contributors, editing methods, and geographic distribution—and that these are not independent phenomena but are closely intertwined in determining the sustainability of the VGI community.

Along the contributor axis, the bipolar structure shown in Section 4.5, wherein 84.7% are single-year contributors and only 2.9% continue for all four years, is confirmed as the concrete manifestation in Japan of the "1% rule" in VGI (Figure 5). This structure has dimensions of both vulnerability and robustness. This vulnerability stems from the dependency risk arising from the oligopolistic nature in which the top 15 contributors (all four-year continuers) account for 16.3% of changesets (Table 3). In particular, the fact that an import-type account with a 96.5% suspect rate accounts for 13.1% of the total creation counts means that the activities of a small number of large-scale importers can affect the interpretation of national trends. In conjunction with

Anderson et al. (2019), who noted the growing presence of corporate contributors, the inequality of actors in VGI is a globally common challenge; in Japan, this problem manifests in a different form through import-type accounts. On the other hand, the robustness is illustrated by the mobilization example during the Noto Peninsula earthquake shown in Section 4.5 and Figure 7 (30,228 changesets in that month, 1,536 unique contributors, and more than 7,000 new mapper flags). The influx of new participants during disasters through the HOT Tasking Manager temporarily far exceeded the scale of the peacetime editing community and served to meet the demand for emergency data development. The layer of single-year participants, which in peacetime appears as a weak tie to VGI, has a dual character as a decisive contributor group during an emergency. Accordingly, two fundamentally different policies are simultaneously required: support for retaining new participants and load distribution among continuing contributors.

Along the editing methods axis, as observed in Section 4.3, the rapid growth of mobile lightweight editors (StreetComplete 4.5-fold, EveryDoor 10.6-fold) and field survey-derived edits (survey 2.8-fold, local knowledge 2.5-fold) is forming a new style of VGI practice (Figure 4). As shown in Section 4.4, these edits maintain extremely low suspect rates (15.4% across the survey period and 10.6% in 2025), confirming that micro-mapping based on field observation is a method capable of reconciling quantitative expansion with quality maintenance. AI-assisted editing, on the other hand, saw a dramatic improvement in the suspect rate from 64.0% to 16.0%, yet, in absolute terms, it remained limited at 1,700–1,800 changesets per year, with no explosive expansion. This indicates that in response to transparency and trust concerns regarding AI-assisted editing raised by Andorful et al. (2026), the Japanese community is establishing a cautious workflow (scrutiny and correction of AI-estimated results by contributors). The convergence of survey (field-based, small-scale) and AI-assisted (remote, large-scale) methodologically contrasting approaches toward quality improvement can be interpreted as the result of contributors understanding the characteristics of each method and making appropriate choices about scope and workflow. The expansion of PLATEAU-derived edits (Section 4.3) also indicates, within this diversification, that a third methodological lineage unique to Japan—OSM editing originating from administratively provided 3D city models—is becoming established (Seto et al., 2023), suggesting the formation of a new field of practice in the mutual reference between digital twins and VGI.

Along the geographic axis, the contrast between the metropolitan concentration shown in Section 4.5 (Figure 6) and the high suspect rates in rural areas highlights the geographic heterogeneity of the VGI community into sharp relief. The gap of approximately 40 percentage points between the high suspect rates of Iwate (59.3%) and Kumamoto (55.4%) and the low rates of Kyoto (20.3%) and Oita (20.6%) makes it clear that edit volume and quality characteristics are not necessarily correlated. The tendency for the proportion of suspicious flags to be higher in rural areas with fewer edits stems from structural factors, including the degree of dependence on a small number of contributors in rural areas and the spatial concentration of AI-assisted editing and large-scale imports. The extreme spike of approximately 6.60 million creations in Aichi, about 4.6 times the previous year's figure (Section 4.5), is a concrete example of how a small number of activities can exert an asymmetric influence on national trends. This suggests that both region-specific development strategies and prior consent processes for large-scale editing activities (the conflict management in peer-production environments discussed by Choe et al., 2023) are necessary for the sustainability of VGI.

These three axes of diversification are interrelated and complementary. The proliferation of mobile editors lowers the barriers to entry for new contributors, resulting in field-survey-derived edits spreading to provincial cities. Conversely, a small number of large-scale import activities asymmetrically amplify a particular contributor's activity volume, manifesting as a spike in specific prefectures, such as Aichi. These spatial heterogeneities are consistent with the spatial bias structure observed by Quattrone et al. (2015) and this study has shown that this globally common tendency is also reproduced at the intra-national prefecture scale. The extremely low suspect rate of survey-derived edits observed in this study (10.6% in 2025) and the high activity volume of continuing contributors provide a clue for reconsidering the positioning of Japanese VGI practices within such cultural contexts. Future VGI research needs to move toward multivariate analysis by combining all three axes, rather than the long-term tracking of a single indicator.

The dataset of 740,038 changesets spanning four years constructed in this study serves as a starting point for such a multidimensional analysis. Going forward, we plan to conduct international comparative research—acquiring and integrating OSMCha data from other countries under the same schema—to clarify how the proportions of field-survey culture, administrative data integration, and AI-assisted editing differ across countries and regions, and to elucidate cultural and institutional differences in VGI editing behavior. OSM editing activity in Japan is characterized by a high proportion of survey-based edits compared with other countries and by advanced use of administratively derived data such as PLATEAU (Seto, 2022; Seto et al., 2023); comparing this characteristic with similarly situated countries will enable the international positioning of a "Japanese model" in VGI. The Python scripts planned for open release in this study and the FlatGeobuf-format dataset with 60 reason_id mappings embedded as attributes are expected to serve as a shared foundation for such comparative research.

6. Conclusions

This study conducted a longitudinal analysis of 740,038 OSM changesets recorded in Japan over four years (2022–2025) using OSMCha. The principal findings are as follows: (1) steady expansion in editing activity and unique contributor counts; (2) a structural change in detection reason composition around 2024 and a paradigm shift in detection logic; (3) rapid growth of mobile editors and survey-derived edits; (4) a marked improvement in the quality of AI-assisted editing; (5) a bipolar structure of a small number of continuing contributors and a large number of single-year participants; and (6) geographic disparities manifesting as outstanding edit volumes in Tokyo and Aichi and high-suspect rates in rural areas.

Future tasks include qualitative analysis of cases labelled as harmful, international comparative analysis, and panel-data analysis tracking individual contributors' activity trajectories. In addition, the mass data acquisition scripts from the OSMCha API developed in this study will be published as open source to ensure reproducibility, and the four-year dataset used in the analysis will be made available via Zenodo in the FlatGeobuf format with 60 reason_id mappings embedded as attributes. This will enable researchers in other countries to conduct equivalent analyses, and we anticipate that it will contribute to the formation of an international baseline for VGI quality monitoring.

Acknowledgements

This research was supported by JSPS KAKENHI Grant Number 24K15662 and the overseas research fellowship program in Komazawa university. Generative AI tools were utilized to assist in the development of the OSMCha data extraction scripts used in this study and were also consulted during the data aggregation process. Specifically, Claude Code was employed to support aspects of code design, implementation, debugging, and data handling procedures. The author critically reviewed, revised, and validated all AI-assisted outputs to ensure their correctness, methodological consistency, and reproducibility.

References

Anderson, J., Sarkar, D., Palen, L., 2019. Corporate editors in the evolving landscape of OpenStreetMap. *ISPRS Int. J. Geo-Inf.*, 8(5), 232. doi.org/10.3390/ijgi8050232.

Andorful, F., Herfort, B., Melanda, E. A., Antonio, N. D., Zipf, A., Camboim, S. P., 2026. Transparency and trust in collaborative mapping: concerns and dilemmas in AI-assisted road integration within OpenStreetMap. *Annals of the American Association of Geographers*, 116(3), 1–22. doi.org/10.1080/24694452.2025.2589286.

Choe, Y., Tomko, M., Kalantari, M., 2023. Producer conflict management approaches in online peer production communities: a case study of OpenStreetMap. *Proc. of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, Article 667, 1–12. doi.org/10.1145/3544548.3581036.

Kanasugi, H., Seto, T., Sekimoto, Y., Shibasaki, R., 2019. Comparison of OpenStreetMap road data and Digital Road Map: focusing on positional accuracy and completeness. *Theory and Applications of GIS*, 27(1), 43–48. (in Japanese)

Li, Y., Anderson, J., Niu, Y., 2021. Vandalism detection in OpenStreetMap via user embeddings. *Proc. of the 30th ACM International Conference on Information and Knowledge Management (CIKM '21)*, 3232–3236. doi.org/10.1145/3459637.3482213.

Minghini, M., Frassinelli, F., 2019. OpenStreetMap history for intrinsic quality assessment: is OSM up-to-date?. *Open Geospatial Data, Software and Standards*, 4, 9. doi.org/10.1186/s40965-019-0067-x.

Raifer, M., Troilo, R., Kowatsch, F., Auer, M., Loos, L., Marx, S., Przybill, K., Fendrich, S., Mocnik, F.-B., Zipf, A., 2019. OSHDB: a framework for spatio-temporal analysis of OpenStreetMap history data. *Open Geospatial Data, Software and Standards*, 4, 3. doi.org/10.1186/s40965-019-0061-3.

Quattrone, G., Capra, L., De Meo, P., 2015. There's no such thing as the perfect map: quantifying bias in spatial crowd-sourcing datasets. *Proc. of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*, 1021–1032. doi.org/10.1145/2675133.2675235.

Seto, T., Kanasugi, H., Nishimura, Y., 2020. Quality verification of volunteered geographic information using OSM notes data in a global context. *ISPRS Int. J. Geo-Inf.*, 9(6), 372. doi.org/10.3390/ijgi9060372.

Seto, T., 2022. Development of OpenStreetMap data in Japan. In: Wakabayashi, Y., Morita, T. (eds.) *Ubiquitous Mapping. Advances in Geographical and Environmental Sciences*, 97–115. Springer: Singapore. doi.org/10.1007/978-981-19-1536-9_7.

Seto, T., Furuhashi, T., Uchiyama, Y., 2023. The role of 3D city model data as an open digital commons: a case study of openness in Japan's digital twin "Project PLATEAU". *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-4/W7-2023, 201–208. doi.org/10.5194/isprs-archives-XLVIII-4-W7-2023-201-2023.

Seto, T., 2024. Editing trends and characteristics of OpenStreetMap in Japan using a data quality verification tool. *Proc. of the 33rd Annual Conference of GIS Association of Japan*. (in Japanese)

Appendix

The data extraction scripts used in this analysis are publicly available on GitHub: https://github.com/tossetolab/osmcha-pipeline.

The extracted nationwide dataset for Japan has been deposited on Zenodo: https://doi.org/10.5281/zenodo.20407813.