AGILE: GIScience Series, 5, 14, 2024. https://doi.org/10.5194/agile-giss-5-14-2024
Proceedings of the 27th AGILE Conference on Geographic Information Science, 4–7 Sept. 2024.
Editors: Alison Heppenstall, Mingshu Wang, Urska Demsar, Rob Lemmens, and Jing Yao.
This contribution underwent peer review based on a full paper submission.

Evaluating geotemporal behaviours of OpenStreetMap contributors

Guy Solomon [1], Dominick Sutton [1], Merve Polat Kayali [1], Xinyi Yuan [1], Zoe Gardner [2], and Ana Basiri [1]

[1]School of Geographical and Earth Sciences, University of Glasgow, Glasgow, UK
[2]School of Geography, Geology and the Environment, University of Leicester, Leicester, UK

Correspondence: Guy Solomon (guy.solomon@glasgow.ac.uk)

Abstract. Volunteered Geographic Information ('VGI') and crowdsourcing are integral for projects such as OpenStreetMap ('OSM'). However, despite the wide use of OSM as one of the most successful crowdsourcing platforms, the under-representation of certain demographic groups amongst those who contribute information may ultimately mean this information favours the interests of some groups over others. This can result in misleading conclusions for analyses conducted on the basis of these data. This paper connects OSM user contributions to demographic data collected via a survey. It shows that, in relation to geographic diversity of contributions, men and women demonstrate distinct trends over time. It then considers the extent to which this observed pattern can be seen as influenced by the COVID-19 pandemic. In this regard, it concludes that there does not appear to be a distinct 'pandemic effect' divergent from longer-term trends.

Keywords. Volunteered Geographic Information, crowdsourcing, OpenStreetMap, bias, user behaviour, gendered participation, disproportionate effect of COVID-19

1 Introduction

Previous research has highlighted the potential biases present within crowdsourcing platforms such as OSM, with particular socio-demographic groups being more likely to contribute (Haklay, 2010; Wood, 2014). This may lead to the information contained within such platforms being skewed towards the interests of certain groups and the omission of others (Lin, 2015; Haklay, 2016; Basiri et al., 2019). This can have implications for analyses undertaken on the basis of these data.

A number of studies have attempted to evaluate this concept, in relation to OSM and other crowdsourcing platforms (Stephens, 2013; Das et al., 2019; Gardner et al., 2020). However, a key issue which sometimes receives less attention is the change in users' behaviour over time. This paper focuses on studying changes in the behaviour of OSM contributors over time, with a specific focus on the extent to which trends in user contributions might differ by gender. Previous work has demonstrated that gender effects may be prominent in contributor behaviours, and as the OSM contributor community demonstrates such a heavy gender skew, this means it is particularly important to establish patterns in the behaviour of these contributor groups (Stephens, 2013; Gardner et al., 2017; Gardner and Mooney, 2018; Das et al., 2019).

Employing geographic diversity of edits as a metric for one type of contributor behaviour demonstrates that users (self-identifying) as male and female exhibit divergent trends in the geographic breadth of their contributions since 2017. These trends appear to have continued through the period of disruption accompanying the COVID-19 pandemic; and whilst they may have been aggravated by the effects of the pandemic, there does not appear to be a 'pandemic effect' divergent from longer-term trends.

2 Motivations and Background

OSM is one of the most successful examples of VGI to emerge from a new technological landscape which increasingly draws on crowdsourced data (Yan et al., 2020; Biljecki et al., 2023). It now contains over 28 billion GPS points (as reported by PlanetOSM), as well as other geospatial and descriptive information. All contributors to OSM are registered as users, but a comparatively small proportion of users are consistently active contributors. This means that the information collected by OSM is used by a large number of people worldwide, including over 10 million registered users, and further users who access OSM on an 'ad hoc' basis.

As research such as Cohen (2011) and Comber et al. (2016) has highlighted, just because crowdsourcing projects (such as OSM) are typically *open* to all individuals to contribute, this does not mean that contributions are *actually* drawn from a balanced cross-section of the population. Whilst OSM users are not required to provide demographic information when they register, various authors have estimated that a significant majority (around 96 percent) of contributors are male (Budhathoki, 2010; Haklay, 2010; Schmidt and Klettner, 2013). This is important because of the potential for 'self focus' biases (Hecht and Gergle, 2009; Das et al., 2019). Bégin et al. (2013), for example, suggest that VGI contributors make choices that are in their own interests (such as the mapping of roads, hiking trails, etc.), depending on their own preferences and motivations, in choosing the areas and features to which they will contribute. If there are similar effects on contributions as a consequence of socio-demographic profiles then biases in *contributors* may be translated into biases in *contributions*.

There is already evidence that men and women contribute to OSM in different ways. Gardner et al. (2017), Gardner and Mooney (2018), and Gardner et al. (2020) found that men and women were likely to demonstrate different behaviours in terms of the types of contributions they made to OSM (such as tendencies to create, modify, or delete information; or to focus on specific types of features, such as 'nodes', 'ways', or 'relations', as described by the OpenStreetMap Wiki). Das et al. (2019) also argue for gender biases in contributions to urbanised and rural areas. More recently, Sutton et al. (2023) have suggested that contributors with different socio-demographic characteristics (gender, age, and education) contribute different types of changesets to OSM as well as offering different geographic dispersions of edits.

However, one element of Sutton et al.'s analysis which has the potential to influence their conclusions is the issue of change over time. As the method of identifying users employed by Sutton et al. involves linking survey data to OSM usernames, this results in all contributions by a single user being identified, regardless of when the survey was carried out. Considering all the contributions made by a single user is important for understanding the types of information they provide, but it is also possible that their behaviour may evolve throughout the period they contribute. For example, a user may have different capacities to contribute changes to OSM at different stages of their life. This can even vary among or within demographic groups. As different demographic groups were affected by the COVID-19 pandemic disproportionately, it might also be worth studying if contributing to OSM was influenced by other external factors, such as national 'lockdowns'.

The geographic scope of a user's contributions may also differ over time as an outcome of a change in confidence in their own abilities. Generally, crowdsourcing projects have a low expertise barrier for contributors, in that they provide various tools to aid users without the need for extensive technical experience. However, Hacar et al. (2018) argue that more detailed contributions are made by the contributors with more experience, the types of contributions also change depending on the level of experience, and participants with less experience contribute to places with fewer details.

In response to life-stage, external stimuli and learning effects (which are not mutually exclusive and may occur in combination), it is likely that contributor behaviour is not static and evolves with the user. Therefore, studies of user contributions must address two related but distinct time-series issues: have the types of contribution made by different groups of contributors *in general* changed over time? And do the contributions of *specific users* evolve over their period of activity?

3 Data

The dataset for this paper is created from the combination of two sources: Gardner et al. (2017)'s OSM demographic survey, and OSM 'changesets' accessed via the OpenStreetMap API. The former asked OSM users to volunteer basic demographic information: gender, age, country of residence, nationality, and highest level of education. They were also asked to provide their OSM username, which was then used to link their demographic information to the latter source, which provides details of their OSM contributions.

Once duplication, irreconcilable OSM usernames, and lack of OSM contributions associated with the identified user account are considered, the survey yields 267 OSM contributors with demographic information – although given the focus on gender in the following analysis, a further three users who responded 'Prefer not to say' to this question are not utilised. It is possible that these individuals would have preferred to select a non-binary gender category, but this was not given as an option in the original survey – and even if this were to be the case, this category would have been too small for meaningful statistical analysis. Of the remainder, 230 gave their gender as 'Male' and 34 as 'Female'. Therefore, whilst more men responded than women, the latter are over-sampled relative to the proportion of OSM contributors estimated to be female. This approach is necessary to evaluate potential gender-based differences in contributor behaviour.

The information relating to the OSM contributions of each identified individual is obtained by analysing their 'changesets' – groups of edits contributed by a single user, typically within a limited geographic scope and relatively short time period. Each changeset contains a 'bounding box', which provides the geographic extend of the changes, and the date and time at which it was closed. Using these two pieces of information, each changeset can be classified by the country in which it took place – by using world country boundaries obtained from ArcGIS Hub (Esri, 2022) and taking the centroid of the bounding box as the point of intersection – and the quarter of the year (taken as January–March, April–June, July–September, and October–December).

The location of each changeset for all users in the sample is shown in Figure 1. The number of changesets contributed by each user in the sample varies significantly, from occasional contributors to 'super-users'.

4 Analysis

The following analysis is divided into three sections. The first addresses the temporal pattern in OSM contributions by gender. The second considers gender distinctions in contribution patterns pre-, during, and post- COVID-19. In both cases, men and women demonstrate different trends in contribution levels and geographic dispersion of edits. The third evaluates changes in the contributions over time of the same users.

4.1 Gendered trends in behaviour over time

A major question regarding contributor behaviour is the extent to which it might change over time. In addition to the level of activity, the metric adopted for analysis is the diversity of countries to which the user contributed, by time period. This is measured by calculating the Gini-Simpson's Index, as discussed by Hunter and Gaston (1988). The geographic spread of changesets is only one aspect of user contributions, but it provides a good proxy for engagement. Contributors with more diverse changesets do not necessarily have higher expertise, but this does suggest a willingness to undertake the mapping of areas in which they are unlikely to be personally familiar, which may be indicative of user confidence and experience.

The Gini-Simpson's Index is calculated as:

$$GSI = 1 - \frac{\Sigma_{i=1}^{249} n_i(n_i - 1)}{N(N-1)}$$

where n_i is the number of changesets by a user in a single country, the summation index, i, ranges across the 249 countries for which OSM edits were made, and N is the total changesets by a user summed across all countries. The value of this index must fall between 0 and 1. If a user (for a given time period) possesses a high GSI score, near 1, this can be interpreted as they made contributions across multiple different countries in relatively even proportions. Alternatively, if a user has a low GSI score, near 0, this would indicate that their changesets were heavily concentrated in a few countries. Given that the goal is to compare the diversity of the contributions of men and women, this is the most appropriate diversity measure for comparing between types.

Figure 2 shows the trends in GSI for each quarter (from Q1 2017, up to and including Q3 2023) by gender. Despite the volatility in the values for female contributors, likely a consequence of their smaller sample size, the diversity of contributions at the start of the period of study for men and women are fairly comparable. Unsurprisingly, given that the vast majority of contributors are male, the male data is showing a pattern closer to that of the overall OSM average. However, from 2017, these trends seem to diverge: with the mean GSI for women increasing, particularly from 2021 onwards; and the mean GSI for men undergoing a gradual decline and plateau. This indicates that female users in the sample gradually began to contribute changesets across a more diverse range of countries than their male counterparts.

It is possible that using change*sets* as a unit of measurement could obscure some of the trends. As the name suggests, changesets are made up of a number of smaller changes. If the average size of a changeset were to vary with time, then the actual contributions made in different countries might display a different trend. Utilising individual *changes* is not necessarily a superior measure, as changesets are arguably a better overview of general activity, but can potentially provide a different perspective on the volume of contributions. The mean GSI for geographic diversity of changes is therefore shown in Figure 3. As is evident from a comparison with Figure 2, although there are some small differences, the overall trends persist.

It is important to note that diversity in contributions is not directly related to overall activity. Considering the number of changesets by year, is it evident from Figure 4 that – for the most part – female users in the sample made fewer contributions than their male counterparts, although this was notably more comparable in 2021 and 2022. In addition, since 2019 the upper quartile of female contributors has been close to (or exceeding) the upper quartile of male contributors; and the lower quartiles for both groups have generally converged. This appears to have somewhat reverted on the basis of 2023 data, which is discussed in the next section.

There are also a high number of what would typically be termed 'outliers'– however, this is to be expected, given the 'long tail' effects present within OSM data (Wood, 2014). What this does suggest is that mean contributions – at least, in terms of activity – will be sensitive to these large values.

The patterns observed in Figure 4 appear to be related, as demonstrated by Figure 5, to a dramatic rise in mean quarterly changesets for female contributors observed between Q1 2019 and Q3 2022. This is discussed in more detail in the following section, however the main peak in Figure 5 – during Q1 and Q2 2021 – saw a particularly prominent number of higher volume users, relative to the other female contributors in the sample.

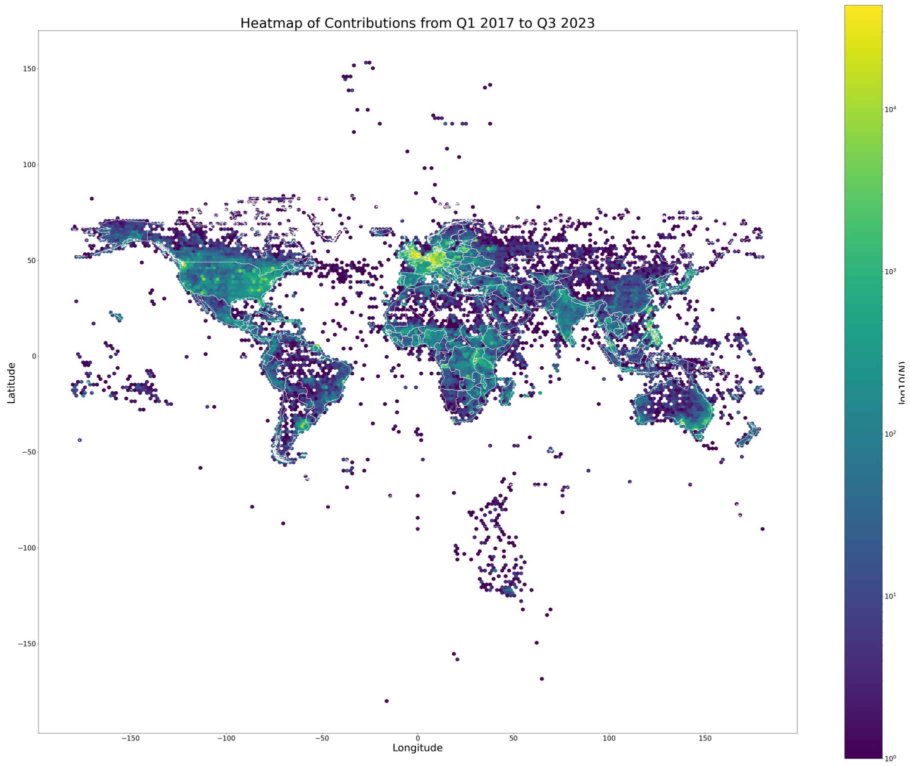

Figure 1. The hexagonally binned locations of all changesets contributed by users in the sample, with national borders: the colour corresponds to the logarithm of the number of changesets.

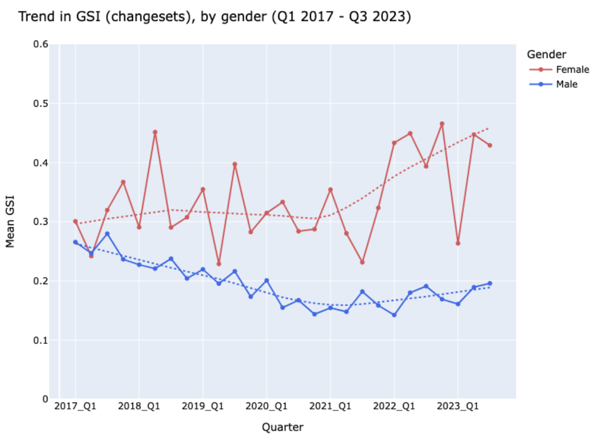

Figure 2. Trend in GSI (changesets), by gender (Q1 2017 – Q3 2023) (inclusive). Trendline (dashed) calculated by Plotly's 'Locally Weighted Scatterplot Smoothing (LOWESS)' function.

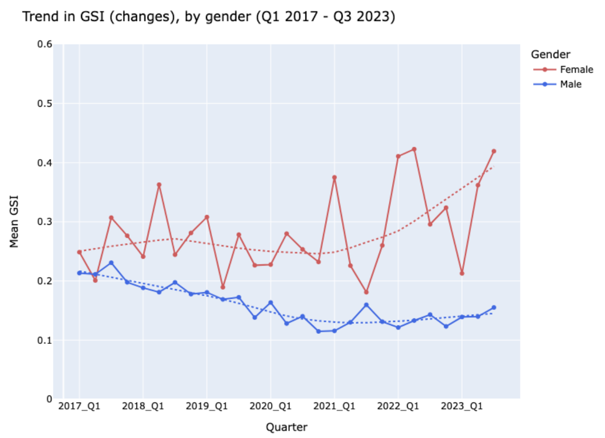

Figure 3. Trend in GSI (changes), by gender (Q1 2017 – Q3 2023) (inclusive). Trendline (dashed) calculated by Plotly's 'Locally Weighted Scatterplot Smoothing (LOWESS)' function.

4.2 The disproportionate effect of the COVID-19 pandemic

Although the previous section goes some way to explaining the trend in female activity levels, it does not indicate the extent to which this might be a process driven (or exacerbated) by the COVID-19 pandemic; nor does it examine how the trend in geographic diversity of contributions might be related to this. There is evidence that the engagement of some contributors may be stimulated by humanitarian events and crises (Roche et al., 2013). Soden and Palen (2014), for example, highlight the catalyzing process of the 2010 Haitian Earthquake. Imi et al. (2012) also observe an increase in OSM activity following the Great East Japan Earthquake of 2011. More recently, the coordination of OSM mapping activities for humanitarian causes, through organisations such as the Humanitarian OpenStreetMap Team (HOT), has led to the involvement

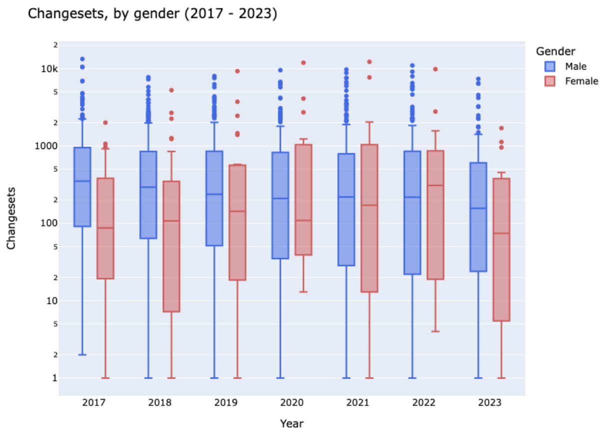

Figure 4. Changesets, by gender (2017 – 2023, inclusive). Note that: a) data for 2023 is incomplete at time of writing, and b) the y-axis is presented on a log scale due to the high number of large values.

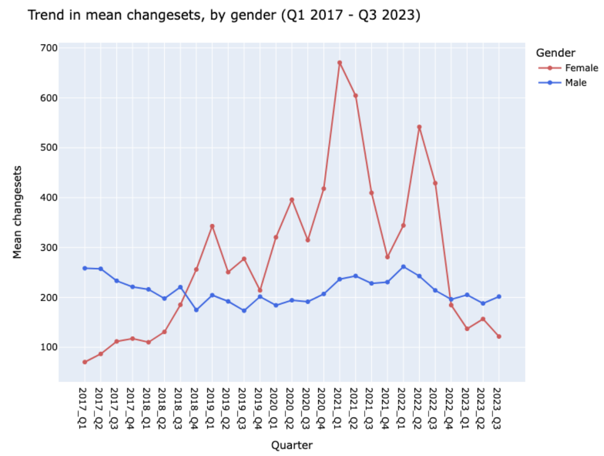

Figure 5. Trend in mean changesets, by gender (Q1 2017 – Q3 2023, inclusive).

of thousands of volunteers (Herfort et al., 2021). However, during participation in such initiatives, as Gardner and Mooney (2018) highlight, tasks are often 'highly prescriptive' and may therefore diverge from a user's typical behaviour.

The COVID-19 pandemic provides perhaps the most broadly applicable example of mapping being influenced by external factors, but changes in user engagement may not be a consequence of precisely the same mechanism. Whereas individuals may well be motivated to contribute to OSM in response to earthquakes, floods, wars, or other similar events, these are not universal motivations for all mappers, and the impact of these events are experienced by a relatively small proportion of people, compared to the global population as a whole. By contrast, it seems plausible that even if the COVID-19 pandemic may not have directly motivated some individuals to change their mapping habits *per se*, the influence of the pandemic (felt through lockdowns, social distancing, and other social restrictions) may still have had an influence (particularly given the biases in the locations of OSM contributors).

This is not to entirely exclude the possibility that the COVID-19 pandemic had a direct influence on information contributed OSM. VGI and spatial data are important for use in COVID-19 or other disasters, as volunteer data allows for the utilisation of real-time contributions for crisis response, especially in the initial and interference stages of such situations (Tzavella et al., 2022). OSM information was used by a variety of different organisations and governments to determine the locations of pharmacies, health institutions, road maps and places with high epidemic risk during the COVID-19 pandemic (Minghini et al., 2020). For example, Minghini et al. (2022), report an increase in Italian pharmacy data of 9.5 percent in the first two months of the pandemic. HOT report that their volunteers provided 4.6 million map features to OSM in support of COVID-19 responses (Deffner et al., 2021). And Mooney et al. (2021) highlight the widespread use of tools like healthsites.io.

However, the softer organising effects of the COVID-19 pandemic on OSM contributors are yet to be studied. Contributors may have had more time (as a consequence of social restrictions limiting their ability to undertake other activities) or less time (for example, due to increased caring responsibilities) which might have influenced their engagement. They may also have felt more urgency and need (due to the state of disaster), or have changed their lifestyles due to restrictions (such as limited capacity to leave their homes resulting in a shift from 'surveying' towards 'armchair' mapping; or alternatively, the inverse, motivated by a desire to undertake physical activities outdoors). Notably, these effects may have disproportionately influenced particular demographic groups. Research suggests, for example, that women became responsible for a greater proportion of caring needs within family units (see, for example, Hupkau and Petrongolo, 2020; Wenham et al., 2020; Herten-Crabb and Wenham, 2022).

As has been seen in the preceding section, the divergence in trends relating to GSI between men and women appears to pre-date the COVID-19 pandemic. Nonetheless, they seemingly become exacerbated around the pandemic period. It is therefore worthwhile examining this change in greater depth, in order to establish the extent to which this might be a temporary break in the previous pattern, or transition to a new model of contributor behaviour.

The data are therefore split into three time periods: Q1 2017 – Q4 2019 (prior to COVID-19), Q1 2020 – Q2 2021 (COVID-19), and Q3 2021 – Q3 2023 (post COVID-19). This is an imperfect categorisation, as the pandemic and its effects were not constant across different countries, and it is also unclear how long any adjustment from COVID to post-COVID behaviour would take. However, it does provide a useful coarse measure for examining the potential influence of the COVID-19 pandemic.

In addition, not every contributor was active in the COVID period but excluding their data may skew the overall picture. It is not possible to tell if this altered activity level was a result of a contribution pattern change, the imposition of different lockdown criteria in different locales, changing levels of responsibility (e.g., additional caring duties or working on the COVID-19 front line as a key worker), or having contracted COVID-19 itself. Finally, analysis based on individual contributors is seriously impacted by the uneven number of quarters in which they were active. Owing to these difficulties in using summary statistics of individual contributor data to compare the pre-COVID, COVID and post-COVID periods the following analysis pools all contributor data, segregated by gender, into the appropriate time periods.

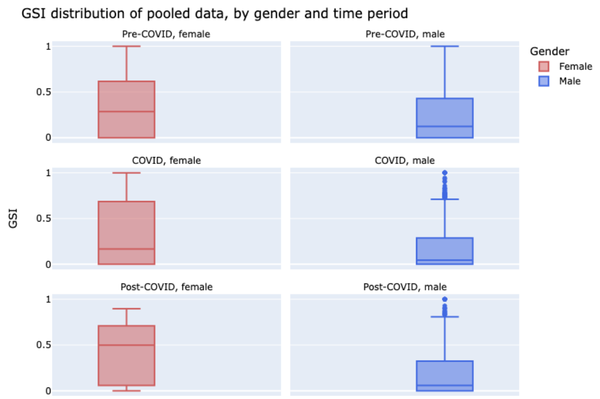

Figure 6. GSI distribution of pooled female and male data for the Pre-COVID, COVID and Post-COVID periods.

Clearly, all of these factors are also likely to influence any analysis on the pooled data. Nonetheless, by pooling the data some insight into contributor activity in the three different time periods can be gleaned. As can be seen in Figure 6, female diversity was higher than male in the pre-COVID period – and this gap persisted in the COVID period. Interestingly, male contributions during COVID (as measured by the GSI) became less diverse, with a more compact inter-quartile range ('IQR') on lower GSI values, and the skew (as measured by position of the median line) increased. Outlier values also began to appear. This contraction does not appear to have been significantly reversed in the post-COVID period.

In contrast, the female IQR widened slightly during COVID in comparison to the pre-COVID period, indicating an overall increase in diversity of contributions. There was also an increase in skew. Post-COVID the female IQR increased slightly again but, more strikingly, there was rise in the median diversity of contributions, as measured by the GSI from 0.1667 during COVID to 0.4985 post-COVID.

However, mean GSI for female contributors is broadly comparable between the pre-COVID and COVID periods, and both are lower than the post-COVID period. This is shown in Figure 7. For male contributors, the mean GSI is lower for the active COVID period than for the preceding period, and largely plateaus into the post-COVID period. This appears to be an extension of the existing trend, as shown by Figure 2.

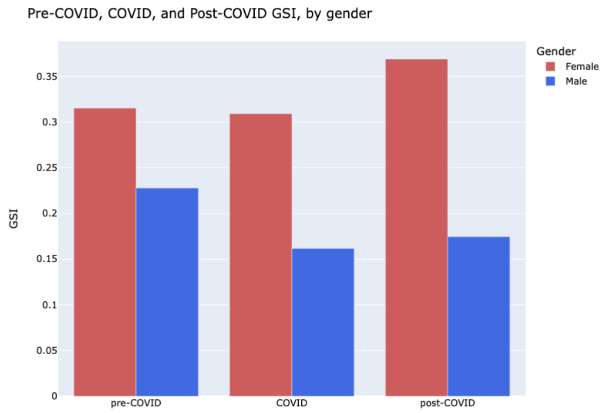

Figure 7. Mean GSI by gender, for Pre-COVID, COVID, and Post-COVID time periods.

As the previous section (and Figure 5) suggested, mean quarterly changesets provided by female users increased dramatically over the period associated with the pandemic. At its peak, mean changesets for female contributors was over double that of the first quarter in the COVID period (321 versus 671). It is therefore valid to question the extent to which this, or more specifically, the social arrangements associated with this period, might have played a role in this behaviour. As evidenced by Figure 8, it can be observed that a decrease in the number of female contributors coincides with the spike in mean changesets. This suggests that the trend is driven by female lower-volume users ceasing to contribute, and the mean is therefore reflecting the group of highly-active female contributors. In addition, whilst the timing of the peak in Figure 5 coincides with the beginning of the pandemic, the rise to this point begins noticeably earlier. Therefore, whilst the COVID-19 pandemic may have been a contributing factor in lower-volume female users choosing not to contribute to OSM, it does not appear that the pandemic drove a higher volume of contributions amongst female users. Nonetheless, the data demonstrates that the female heavy-mappers in our sample continued to contribute to OSM throughout the pandemic, whereas the lighter female mappers did not.

4.3 Change in individual user behaviour over time

If these trends are rooted in the long-term evolution of contributor behaviour, rather than a short-term adaptation during COVID-19, then a fruitful line of inquiry is the extent to which this is a change in the *activities* of particular users; versus a change in the *composition* of the user-base. Thebault-Spieker et al. (2018) argue that user con-

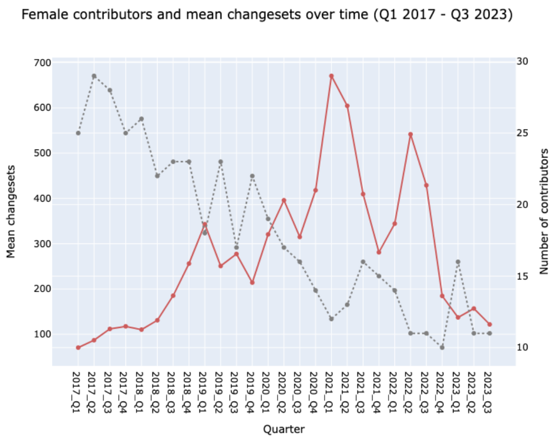

Figure 8. Female contributors (dashed grey) and mean changesets (red) over time, Q1 2017 – Q3 2023

tributions are relatively constant over time, but they do not break this down into demographic groups – instead favouring user classes based on the number of edits. They also restrict their analysis to the continental United States, which may demonstrate different patterns to that of a global perspective.

The combination of a relatively small sample size for female users, the number of contributors who have extensive quarters without activity, and the exogenous impact of COVID – as well as the relatively short period post-COVID – makes assessing the underlying activity trend for individuals difficult. Nonetheless, some insight into the stability of contributor activity over time may be gleaned by looking at the mean standard deviation of individual contributions over time. Table 1 shows the grand mean and mean standard deviation of GSI, as calculated on individual contributor activity. It can be seen that the mean and standard deviations are of similar magnitude, with the standard deviation somewhat smaller than the mean. The standard deviation for males is smaller than for females, which is consistent with the greater volatility for females observed in the previous figures. The full distribution of these GSI standard deviations calculated on individual contributor data, segregated by gender, is given in Figure 9. The relatively compact range for male data is suggestive of relatively stable GSI activity levels. Although the range for female data is wider and has a higher mean it may also be associated with a relatively stable GSI contribution. In both cases the Coefficient of Variation ('CV', see Reed et al., 2002) is less than 100%, with the female CV at 78.7% and male at 84.1%. Together, this suggests that the diversity in contributions of both male and female contributors is relatively stable over time.

4.4 Discussion

There are several elements of the work discussed here which require careful consideration. First, although it is possible to link surveyed users to the changesets prior to 2017, it is not clear if this would provide a representative sample of OSM contributors in those years. These changesets are therefore excluded, but it is therefore not possible to say with confidence how far the trends observed after 2017 are a break with those in previous years, or the continuation of a longer-term process.

Table 1. Standard deviation of GSI by gender, calculated by individual

	Mean	Standard Deviation	Coefficient of Variation
Female	0.341	0.268	78.7 %
Male	0.214	0.180	84.1 %

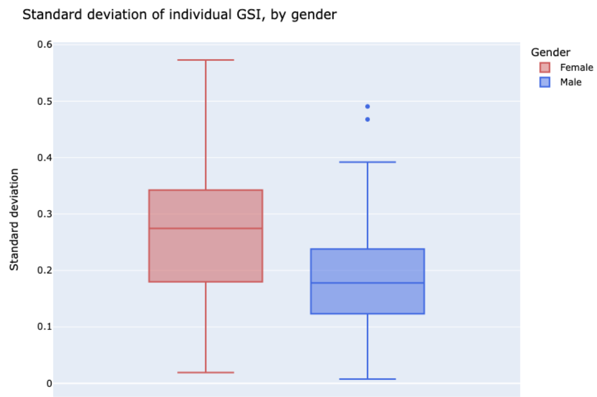

Figure 9. A boxplot of the standard deviations of individual contributor GSI by quarter.

Second, the fairly coarse definition of the COVID-19 pandemic likely has the potential to blur some modified patterns of behaviour as a consequence of social restrictions. OSM is a global project, and contributors would have seen different restrictions at different times, depending on their geographic location. The results presented here provide a good indication of how far social restrictions can be observed in contributor's behaviours, but a fine-grained analysis may provide further insight.

Third, although there is a noticeable drop-out of lower volume female contributors around the period of the COVID-19 pandemic, it is not possible on the basis of these data to say with certainty that there is a link between the two. Additional data collection may allow for greater examination of this aspect.

5 Conclusions and future work

In conclusion, the data suggest that the diversity of contributions of male and female OSM contributors have been diverging over time – at least, since 2017; but particularly, since 2021. Although these trends have persisted through the disruption associated with the COVID-19 pandemic, there is only limited evidence of a 'pandemic effect' – at least on this element of contributions – as these trends appear to largely have been in motion prior to the period associated with COVID-19. Geographic diversity in contributor activity also appears to be relatively stable over time, suggesting that users do not substantially modify this aspect of their behaviour.

Lower-volume female users in the sample ceased to contribute in the time period associated with the COVID-19 pandemic. This leads to a particularly high mean for female contributors in certain quarters (as discussed in relation to Figure 8), as the remainder appear to be skewed towards 'uber' mappers, who consistently contribute a high number of changesets. Further work is required to establish the breadth of this trend, how closely it might have been related to conditions caused by the COVID-19 pandemic, and the extent to which it has persisted in the post-COVID period.

There are several other elements which provide promising prospects for future work in this area. First, re-surveying OSM contributors will allow for the expansion of the sample beyond the 267 participants utilised here. This will also facilitate broadening the scope of the data collected, including explicit consideration of contributor participation in humanitarian, special interest, and other mapping schemes. Second, expanding the analysis to consider other socio-demographic factors (such as age, country of residence, and education) and other influences on the long term trajectory of OSM contributions will allow for a more rounded approach to the assessment of contributor behaviour. This includes the potential combinations of socio-demographic factors, touched upon by Sutton et al. (2023). Finally, whilst utilising GSI as a measure of geographic diversity yields valuable insights, adopting measures of spatial entropy for future analysis will facilitate a more direct engagement with the spatial location of contributions in relation to each other, without the need to classify contributions into nations. This has the potential to offer a powerful analysis of both how individual users contribute, and how these contributions relate to patterns present within OSM as a whole.

Data and Software Availability

The data cannot be made publicly available due to their containing personally identifiable information that could compromise the privacy of research participants. Code utilised in the creation and analysis of these data has previously been made available at https://doi.org/10.5281/zenodo.10817832. Additional code to support the analysis and visualisations for this paper is available at https://doi.org/10.5281/zenodo.10988945.

Author contributions. Author contributions are as follows:

GS: conceptualisation and scoping, data processing, methodology, analysis, visualisation, writing, editing and reviewing, organisation and supervision.
DS: conceptualisation and scoping, data processing, methodology, analysis, visualisation, writing, editing and reviewing.
MPK: conceptualisation and scoping, literature review, data processing, methodology, writing.
XY: conceptualisation and scoping, data processing, methodology, visualisation, writing.
ZG: conceptualisation and scoping, data collection, editing and reviewing.
AB: conceptualisation and scoping, methodology, writing, editing and reviewing, organisation and supervision, funding acquisition.

Competing interests. The authors declare none.

Acknowledgements. The authors received support from the UK Research and Innovation Future Leaders Fellowship "Indicative Data: Extracting 3D Models of Cities from Unavailability and Degradation of Global Navigation Satellite Systems (GNSS)", grant number MR/S01795X/2, the Republic of Turkey Ministry of National Education Scholarship Program, and China Scholarship Council.
The authors would like to thank three anonymous AGILE referees for their comments on the original manuscript.

References

Basiri, A., Haklay, M., Foody, G., and Mooney, P.: Crowdsourced geospatial data quality: challenges and future directions, International Journal of Geographical Information Science, 33, 1588–1593, https://doi.org/10.1080/13658816.2019.1593422, 2019.

Bégin, D., Devillers, R., and Roche, S.: ASSESSING VOLUNTEERED GEOGRAPHIC INFORMATION (VGI) QUALITY BASED ON CONTRIBUTORS' MAPPING BEHAVIOURS, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-2/W1, 149–154, https://doi.org/10.5194/isprsarchives-XL-2-W1-149-2013, 2013.

Biljecki, F., Chow, Y. S., and Lee, K.: Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes, Building and Environment, 237, 110 295, https://doi.org/10.1016/j.buildenv.2023.110295, 2023.

Budhathoki, N. R.: Participants' motivations to contribute geographic information in an online community, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2010.

Cohen, N.: Define Gender Gap? Look Up Wikipedia's Contributor List, The New York Times, https://www.nytimes.com/2011/01/31/business/media/31link.html?smid=url-share, 2011.

Comber, A., Mooney, P., Purves, R. S., Rocchini, D., and Walz, A.: Crowdsourcing: It Matters Who the Crowd Are. The Impacts of between Group Variations in Recording Land Cover, PloS one, 11, e0158 329, https://doi.org/10.1371/journal.pone.0158329, 2016.

Das, M., Hecht, B., and Gergle, D.: The gendered geography of contributions to OpenStreetMap: Complexities in self-focus bias, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–14, https://doi.org/10.1145/3290605.3300793, 2019.

Deffner, R., Martinez, A. R., and Reyes, C. d. l.: Building a global network to map for COVID-19 response, https://www.hotosm.org/updates/building-a-global-network-to-map-for-covid-19-response/, 2021.

Esri: World Countries Generalized, https://hub.arcgis.com/datasets/esri::world-countries-generalized/, 2022.

Gardner, Z. and Mooney, P.: Investigating gender differences in OpenStreetMap activities in Malawi: a small case-study, in: AGILE, Lund, Sweden, 2018.

Gardner, Z., Mooney, P., Dowthwaite, L., and Foody, G.: Gender differences in OpenStreetMap contributor activity, editing and tagging behaviour, 25th Annual Geographical Information Science Research UK (GISRUK), Manchester, United Kingdom, p. 6, 2017.

Gardner, Z., Mooney, P., De Sabbata, S., and Dowthwaite, L.: Quantifying gendered participation in OpenStreetMap: responding to theories of female (under) representation in crowdsourced mapping, GeoJournal, 85, 1603–1620, https://doi.org/10.1007/s10708-019-10035-z, 2020.

Hacar, M., Kılıç, B., and Şahbaz, K.: Analyzing OpenStreetMap Road Data and Characterizing the Behavior of Contributors in Ankara, Turkey, ISPRS International Journal of Geo-Information, 7, https://doi.org/10.3390/ijgi7100400, 2018.

Haklay, M.: How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets, Environment and planning B: Planning and design, 37, 682–703, https://doi.org/10.1068/b35097, 2010.

Haklay, M. M.: Why is participation inequality important?, in: European Handbook of Crowdsourced Geographic Information, Ubiquity Press, https://doi.org/10.5334/bax.c, 2016.

healthsites.io: https://healthsites.io/.

Hecht, B. and Gergle, D.: Measuring self-focus bias in community-maintained knowledge repositories, in: C and T 2009 - Proceedings of the 4th International Conference on Communities and Technologies, C and T 2009 - Proceedings of the 4th International Conference on Communities and Technologies, pp. 11–19, Association for Computing Machinery, https://doi.org/10.1145/1556460.1556463, 2009.

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., and Zipf, A.: The evolution of humanitarian mapping within the OpenStreetMap community, Scientific reports, 11, 3037, https://doi.org/10.1038/s41598-021-82404-z, 2021.

Herten-Crabb, A. and Wenham, C.: “I Was Facilitating Everybody Else’s Life. And Mine Had Just Ground to a Halt”: The COVID-19 Pandemic and its Impact on Women in the United Kingdom, Soc Polit, 29, 1213–1235, https://doi.org/10.1093/sp/jxac006, 2022.

Hunter, P. R. and Gaston, M. A.: Numerical index of the discriminatory ability of typing systems: an application of Simpson’s index of diversity, Journal of Clinical Microbiology, 26, 2465–2466, https://doi.org/10.1128/JCM.26.11.2465-2466.1988, 1988.

Hupkau, C. and Petrongolo, B.: Work, Care and Gender during the COVID-19 Crisis, Fiscal Studies, 41, 623–651, https://doi.org/https://doi.org/10.1111/1475-5890.12245, 2020.

Imi, Y., Hayakawa, T., and Ito, T.: Analyzing the Effect of Open Street Map During Crises: The Great East Japan Earthquake, in: 2012 IEEE 14th International Conference on Commerce and Enterprise Computing, pp. 126–130, https://doi.org/10.1109/CEC.2012.28, 2012.

Lin, W.: Revealing the making of OpenStreetMap: A limited account, The Canadian Geographer/Le Géographe canadien, 59, 69–81, https://doi.org/https://doi.org/10.1111/cag.12137, 2015.

Minghini, M., Coetzee, S., Grinberger, A. Y., Yeboah, G., Juhász, L., and Mooney, P.: Editorial: OpenStreetMap research in the COVID-19 era, Proceedings of the Academic Track at the State of the Map 2020 Online Conference, https://doi.org/10.5281/zenodo.3922054, 2020.

Minghini, M., Sarretta, A., and Napolitano, M.: OpenStreetMap Contribution to Local Data Ecosystems in COVID-19 Times: Experiences and Reflections from the Italian Case, Data, 7, 39, https://doi.org/10.3390/data7040039, 2022.

Mooney, P., Grinberger, A. Y., Minghini, M., Coetzee, S., Juhasz, L., and Yeboah, G.: OpenStreetMap Data Use Cases During the Early Months of the COVID-19 Pandemic, pp. 171–186, COVID-19 Pandemic, Geospatial Information, and Community Resilience, CRC Press, 2021.

OpenStreetMap API: API v0.6 — OpenStreetMap Wiki, https://wiki.openstreetmap.org/w/index.php?title=API_v0.6&oldid=2589236, online; accessed 29-August-2023, 2023.

OpenStreetMap Wiki: Elements — OpenStreetMap Wiki, https://wiki.openstreetmap.org/w/index.php?title=Elements&oldid=2582681, online; accessed 28-August-2023, 2023.

PlanetOSM: https://planet.openstreetmap.org/statistics/data_stats.html.

Plotly: Linear and Non-Linear Trendlines in Python, https://plotly.com/python/linear-fits/.

Reed, G. F., Lynn, F., and Meade, B. D.: Use of Coefficient of Variation in Assessing Variability of Quantitative Assays, Clinical and Vaccine Immunology, 9, 1235–1239, https://doi.org/10.1128/CDLI.9.6.1235-1239.2002, 2002.

Roche, S., Propeck-Zimmermann, E., and Mericskay, B.: GeoWeb and crisis management: issues and perspectives of volunteered geographic information, GeoJournal, 78, 21–40, https://doi.org/10.1007/s10708-011-9423-9, 2013.

Schmidt, M. and Klettner, S.: Gender and experience-related motivators for contributing to OpenStreetMap, in: AGILE2013: International workshop on action and interaction in volunteered geographic information (ACTIVITY), pp. 13–18, 2013.

Soden, R. and Palen, L.: From Crowdsourced Mapping to Community Mapping: The Post-earthquake Work of OpenStreetMap Haiti, in: COOP 2014 - Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France), pp. 311–326, https://doi.org/10.1007/978-3-319-06498-7_19, 2014.

Stephens, M.: Gender and the GeoWeb: divisions in the production of user-generated cartographic information, GeoJournal, 78, 981–996, https://doi.org/10.1007/s10708-013-9492-z, 2013.

Sutton, D., Solomon, G., Yuan, X., Polat Kayali, M., Gardner, Z., and Basiri, A.: Assessing the Relationship between Socio-Demographic Characteristics and OpenStreetMap Contributor Behaviours, in: Proceedings of the 1st ACM SIGSPATIAL International Workshop on Geocomputational Analysis of Socio-Economic Data, GeoSocial '23, p. 5–11, Association for Computing Machinery, New York, NY, USA, https://doi.org/10.1145/3615892.3628477, 2023.

Thebault-Spieker, J., Hecht, B., and Terveen, L.: Geographic Biases Are 'Born, Not Made': Exploring Contributors' Spatiotemporal Behavior in OpenStreetMap, in: Proceedings of the 2018 ACM International Conference on Supporting Group Work, GROUP '18, p. 71–82, Association for Computing Machinery, New York, NY, USA, https://doi.org/10.1145/3148330.3148350, 2018.

Tzavella, K., Skopeliti, A., and Fekete, A.: Volunteered geographic information use in crisis, emergency and disaster management: a scoping review and a web atlas, Geo-spatial Information Science, 27, 423–454, https://doi.org/10.1080/10095020.2022.2139642, 2022.

Wenham, C., Smith, J., and Morgan, R.: COVID-19: the gendered impacts of the outbreak, The Lancet, 395, 846–848, https://doi.org/10.1016/S0140-6736(20)30526-2, 2020.

Wood, H.: The Long Tail of OpenStreetMap, https://harrywood.co.uk/blog/2014/11/17/the-long-tail-of-openstreetmap/, 2014.

Yan, Y., Feng, C.-C., Huang, W., Fan, H., Wang, Y.-C., and Zipf, A.: Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience, International Journal of Geographical Information Science, 34, 1765–1791, https://doi.org/10.1080/13658816.2020.1730848, 2020.