Fine-scale spatial distribution of biodiversity citizen science in a natural area depends on area accessibility and differs from other recreational area use

C.P. Mandeville[1], E.B. Nilsen[2,3], A.G. Finstad[4]

[1]Centre for Biodiversity Dynamics, Department of Natural History, Norwegian University of Science and Technology. Erling Skakkes Gate 47B, 7021 Trondheim, Norway.

caitlin.mandeville@ntnu.no

[2]Norwegian Institute for Nature Research, Postboks 5685 Torgarden, 7485 Trondheim, Norway.

[3]Nord University, Kongens gate 42, 7713 Steinkjer, Norway.

[4]Centre for Biodiversity Dynamics, Department of Natural History, Norwegian University of Science and Technology. Erling Skakkes Gate 47B, 7021 Trondheim, Norway.

Abstract

1. Opportunistic citizen science produces large amounts of primary biodiversity data but is underutilized in the conservation and management of protected areas despite these areas' status as citizen science hotspots. Application of these data may be limited by the challenge of understanding sampling patterns associated with opportunistic data at a scale relevant to local area management. An improved understanding of citizen science activity patterns within protected areas could strengthen both data analysis and the local promotion and guidance of citizen science activity.
2. We investigated local-scale patterns of citizen science distribution, using a case study approach to examine citizen science activity in a recreationally popular natural area that serves as a regional citizen science hotspot. We modeled the relationship between local citizen science activity and ten spatial covariates broadly related to ease of access and natural interest, which have been shown to drive citizen science activity at regional scales in previous studies. We further compared the distribution of citizen science activity with that of professional data collection, and with data on recreational visitor activity in the study area.
3. We found that citizen science largely complements professional data collection in space. Citizen science participation was primarily driven by ease of access, especially the presence of trails. However, citizen science use of the trail network differed from other types of recreational trail use, including a weaker preference for well-established trails and a stronger association with developed areas.
4. This improved understanding of patterns in citizen science participation may be used to better account for spatial biases in citizen science data and to manage natural areas in a way that supports and guides future citizen science activity.

Keywords

biodiversity; biodiversity data; citizen science; community science; protected areas; recreation ecology

Introduction

Public participation in biodiversity research, often termed biodiversity citizen science, continues to grow in popularity. Citizen science engages millions of individuals in biodiversity science, produces massive amounts of data, and contributes extensively to research in biodiversity, conservation, and related fields (Cooper et al. 2014, Kays et al. 2020, Bonney 2021, Callaghan et al. 2021). Much of this contribution comes from mass participation citizen science, in which participants opportunistically upload species observations to digital platforms that are often national to international in scope, due largely to the accessibility of these data in open digital repositories (Ball-Damerow et al. 2019, Callaghan et al. 2021, Mandeville et al. 2021). But despite the mainstream recognition and broad application of mass participation citizen science, it is generally underutilized in the conservation and management of protected areas (Danielsen et al. 2010, Callaghan and Gawlik 2015, Binley et al. 2021, 2021, Mandeville and Finstad 2021, Rapacciuolo et al. 2021, Salmon et al. 2021, Cheung et al. 2022).

Biodiversity data from mass participation citizen science could contribute to filling a critical gap for small protected areas and multiple-use areas that contribute to other effective area-based conservation measures (OECMs [IUCN 2019]), which are increasingly recognized as crucial for meeting biodiversity conservation targets (Kendal et al. 2017, Baldwin and Fouch 2018, Bonnet et al. 2020, Häkkilä et al. 2021, Rodríguez-Rodríguez et al. 2021, Riva and Fahrig 2022). Such areas enhance connectivity, support ecosystem services, and play a key role in addressing environmental threats that manifest at a local scale (Oldekop et al. 2015, Volenec and Dobson 2019, Wintle et al. 2019, Hlasny et al. 2021, Gaget et al. 2021, Dreiss and Malcolm 2022). Still, small protected areas often have limited resources for biodiversity conservation, management, and monitoring, despite their high conservation value (Armsworth et al. 2011, Maxwell et al. 2020, Jansujwicz et al. 2020).

Mass participation citizen science data are already regularly collected in protected areas, which tend to be hotspots for citizen science activity (Tulloch et al. 2013). At broad spatial resolutions, citizen science activity is largely associated with two main types of predictors: accessibility (e.g., population density, road access, regional trail availability) and natural interest (e.g., aesthetic and recreational value, high biodiversity, and threatened ecosystems) (Tulloch et al. 2013, Geldmann et al. 2016, Boakes et al. 2016, Mair & Ruete 2016, Tiago et al. 2017, Millar et al. 2018, Petersen et al. 2021). Because they are accessible areas of local natural interest, small protected areas and OECMs are popular destinations for citizen science participants. Nevertheless, data collected in these areas remain underutilized in local area-based conservation and are instead more commonly applied in studies at broad spatial scales (Danielsen et al. 2010, Callaghan and Gawlik 2015, Rapacciuolo et al. 2021, Mandeville et al. 2021). A key reason for their underutilization at local scales may be that spatial patterns of citizen science activity are not well understood at a scale relevant to local area management (Callaghan and Gawlik 2015).

This may inhibit local applications of citizen science data in multiple ways. First, a limited understanding of the citizen science sampling process, coupled with the rarity of species non-detection data, means that spatial and temporal biases in the data are hard to quantify and species absences are hard to infer, limiting the potential for statistical inference (Welvaert et al. 2016, Johnston et al. 2020, Di Cecco et al. 2021). At a broader spatial scale, this challenge is sometimes addressed by using trends in the spatial distribution of citizen science activity to approximate the sampling process (Mueller et al. 2019, Sicacha Parada et al. 2020, Cretois et al. 2020, Johnston et al. 2020, Di Cecco et al. 2021, Zulian et al. 2021). But covariates commonly used to model the citizen science sampling process at broader spatial scales are often not well suited to characterize the sampling process at scales relevant to local management. As such, little is known about how the fine-scale distribution of citizen science activity varies within regional citizen science hotspots (Callaghan and Gawlik 2015, Dobson et al. 2020).

Further, anmproved understanding of citizen science activity within natural areas is required for area managers to utilize citizen science more effectively (Feldman et al. 2021). Such areas are commonly managed for both conservation and recreation objectives, both of which can be furthered by citizen science (Buta et al. 2014, Newman et al. 2017, Gurney et al. 2021, Vimal et al. 2021, Halliwell et al. 2021). Citizen science participants are increasingly recognized as an important category of protected area visitors, and a better understanding of spatial patterns in their activity would allow managers to actively promote and direct citizen science to meet local objectives. Such direction (e.g., interpretive signage, the use of customized settings on citizen science platforms, and promotional events such as bioblitzes) can effectively guide mass participation citizen science data collection (Callaghan et al. 2019, Knape et al. 2021, Koen & Newton 2021, Kays et al. 2021, Salmon et al. 2021). For these reasons, researchers and managers of protected areas have called for greater research into trends in citizen science participation within protected areas (Weaver and Lawton 2017, Leung et al. 2018, Miller et al. 2019, Binley et al. 2021, Gosal et al. 2021).

We aimed to respond to this challenge by investigating the spatial distribution of citizen science participation at a scale relevant to local area management. We took a case study approach, characterizing citizen science activity within a small, recreationally popular natural area in Central Norway. The site was selected because it meets many criteria as a citizen science hotspot. Our objectives were to 1) test the hypothesis that the main predictors of citizen science activity at a broad spatial resolution—accessibility and natural interest—also drive citizen science at a local scale; 2) test the hypothesis that citizen science activity would primarily occur within a short distance of trails and roads; and 3) compare the patterns of citizen science activity within the study area and along the area's trail network with patterns in both general recreational visitor activity, represented by activity tracking data from Strava Metro, and professional biodiversity data collection.

Methods

2.1 Study site

Our study site is an 86 km^2 natural area located on the periphery of Trondheim, Central Norway, a regionally dominant city with a population of around 190,000 (Figure 1; Trondheim Municipality 2020). The area consists of a diverse range of southern-boreal habitat types, including mires, mixed forest, lakes, and coastline (Moen 1999). Land management objectives vary within the study area; the entire area is designated as a natural area for public use, while three smaller subsets of the area comprising a total of 12 km^2 are designated as nature reserves with greater conservation protections. The area contains an extensive trail network that is used throughout the year for a range of activities including hiking, running, cycling, and skiing, as well as a small number of access roads. There are also a small number of private homes within the area, primarily concentrated near the access roads. The area is recognized as highly important for recreation, but visitor activity patterns are not well studied (Hagen et al. 2019).

2.2 Data

2.2.1 Citizen science and professional biodiversity data. All biodiversity data available on the Global Biodiversity Information Facility (GBIF) for the study area were downloaded on August 3, 2021 (GBIF 2021). The descriptions on GBIF of contributing data providers were used to classify all data as either opportunistic citizen science, structured citizen science, or professionally collected data. The single dataset classified as structured citizen science was excluded from analysis (Supporting Information). Data from before 2000 were excluded, as digital platforms for opportunistic citizen science largely grew in popularity after that year. Bacteria and freshwater-obligate species, including fish and aquatic invertebrates, were excluded because the citizen science observation process for these species is expected to differ fundamentally from that of terrestrial species. Finally, data points with a recorded coordinate uncertainty of greater than 150 meters were excluded.

The filtered citizen science data consisted of 44206 observations from seven citizen science platforms. The majority were contributed through the Norwegian Species Observation Service (https://www.biodiversity.no/), which is Norway's main biodiversity citizen science platform. Citizen science data were contributed by 560 participants in 8614 observation events (events being defined as unique combinations of observer, location coordinates, and date). As is typical of digital citizen science datasets, a small number of highly active participants contributed the majority of the data; the most active five percent of participants contributed 79% of the total data, while the median participant contributed just six observations. The filtered professionally collected data consisted of 2059 observations from 31 data providers, collected in 907 observation events (Supporting Information).

In total, the citizen science data contained reports of 1524 species and the professional data contained reports of 991 species (Figure 2). Both types of data collection took place year-round with a peak in intensity in the summer months, but annual variation in sampling intensity was more extreme in the professional data, with sampling intensity peaking later in the summer and falling to a much lower rate in the winter than in the citizen science data (Figure 2). Observations occurred in all available land cover types (Figure 2).

2.2.2 Recreational visitor data. Data on recreational trail use were accessed from Strava Metro (https://metro.strava.com). Strava Metro publishes public data from users of Strava, a mobile app used by recreationists to log running, cycling, skiing, and other recreational activities. Data were summarized as the number of recorded trips per Open Street Map (https://www.openstreetmap.org) segment, defined as sections of trail or road between intersections. Strava Metro data were available from 2016 through 2020.

2.2.3 Environmental covariates. We identified ten environmental covariates, broadly related to ease of area access and natural interest, that we expected to relate to citizen science activity. Five covariates were related to area access: access points, trail locations, recreational facilities (e.g., public tourist cabins, playgrounds, maintained swimming beaches, and similar), elevation, and longitude. Area access points were defined by intersections between a road or trail and the boundary of the natural area as well as public parking areas and public transit stops within or adjacent to the area. The locations of trails and recreational facilities were derived from maps provided by Trondheim Municipality (https://kart.trondheim.kommune.no). Elevation was accessed from the Norwegian Digital Elevation Model (https://www.kartverket.no). Longitude was used as a proxy for distance from the nearest population center; the study area lies to the west of Trondheim's population center, so it was expected that eastern longitudes would be positively associated with citizen science activity. Five covariates were related to natural interest: cultivated land cover and developed land cover were expected to relate negatively to citizen science activity, as these land cover types may be perceived as less natural and thus less interesting than other land cover types in the study area. Conversely, forest and wetland land cover and proximity to a freshwater lake or stream were expected to relate positively to citizen science activity. All natural interest variables were derived from the Norwegian Institute for Bioeconomics AR5 land cover data at a 1:5000 scale (Ahlstrøm et al. 2014).

2.3 Analysis

2.3.1 Environmental covariates of citizen science activity. To examine the relationship between our ease of access and natural interest covariates and the distribution of citizen science activity, we established a grid of 150 x 150 m^2 cells in the study area, resulting in 4130 cells. Using the number of citizen science observations in each grid cell as a response variable, we fit a negative binomial generalized linear model with the ten covariates as predictor variables. This approach follows other studies that have examined covariates of citizen science activity at a broader spatial scale (e.g., Romo et al. 2006, Tulloch et al. 2012, Tiago et al. 2017). There were a small number of outlier cells (n = 7) where the number of citizen science observations was between two and eight times greater than in any other cells with citizen science activity. Citizen science participation in these highly active cells was most likely driven by processes that differ from typical patterns of citizen science participation; for instance, three such cells were located in the vicinity of a birdwatching tower and two were adjacent to a school and a residential neighborhood at the edge of the study area. Because the drivers of sampling activity in these outlier cells might differ fundamentally from the typical sampling process within the study area, they were excluded from the analysis.

The ten covariates were summarized by grid cell in the following ways: access was summarized by distance from grid centroid to nearest access point; trail locations were summarized by the total length of trail within each grid cell; recreational facilities were summarized as a binary variable expressing whether or not the grid contained a facility; elevation was summarized as the maximum elevation per grid cell; longitude was summarized by the grid centroid; all land cover types were summarized as the area within the grid cell covered by the land cover type; and proximity to freshwater was summarized as a binary variable expressing whether or not the cell contained a freshwater body. We tested for spatial autocorrelation using Moran's I and included a distance-weighted autocovariate in the model, which reduced autocorrelation (Bardos et al. 2015). We used Akaike's information criterion for small sample sizes (AIC_c) to rank all possible models consisting of combinations of our covariates with no interactions. The ranked models were used to determine the relative importance of each covariate, and we used multi-model inference to obtain the model-averaged estimate and standard error for each covariate (Burnham and Anderson 2002).

To compare the distribution of citizen science activity with comparable professional data collection processes, we first used a Pearson rank correlation analysis to compare the distribution of the two activity types and then repeated the modeling analyses using the number of professional biodiversity data observations per grid cell as the response variable.

2.3.2 Relationship between citizen science and trail network. Because accessibility, including regional trail density, has been shown to predict citizen science activity at broad spatial scales (Tiago et al. 2017), we examined the relationship between citizen science activity and trails in our study area. We used a linear model to test the hypothesis that the average distance from citizen science observation to the nearest trail would be smaller than if the points were distributed randomly.

We further investigated whether citizen science participants who have likely accessed their observation sites via a trail tend to make observations from the trail or to leave the trail before making observations. We expected that if participants tend to make observations from the trail, then the distance between the recorded observation coordinates and the nearest trail would be greatest for taxonomic groups that are most often visible and identifiable from a distance (e.g., mammals, birds, some plants). If participants tend to leave the trail to make observations, then we would not expect this relationship. We used a linear model to examine the relationship between distance to trail and the observed taxonomic group as an indicator of off-trail observation activity.

Both analyses related to the relationship between observation sites and trail locations were repeated for the professional dataset.

2.3.3 Citizen science and other recreational trail use. Due to the previously documented relationship between citizen science activity and the presence of trails (Maire and Ruete 2016, Tiago et al. 2017), we hypothesized that the spatial distribution of citizen science activity along trail segments would be positively correlated with the intensity of activity by other recreational trail users. To test this, we first used a Pearson rank correlation test to compare the number of citizen science observations within a 300-meter-wide corridor along each trail segment in our study area (n = 7153), standardized by segment length, with the total number of Strava Metro activities reported on the segment.

Finally, we compared the relationship between our covariates and citizen science activity along trail segments with the relationship that those covariates have to Strava activity. We first used the number of citizen science observations per segment corridor, standardized by segment length, as a response variable and fit a negative binomial generalized linear model with a modified set of landscape covariates. We then used the number of Strava activities reported along each trail segment to fit a second model with the same structure and covariates.

The ten covariates were adapted to relate to trail segments rather than grid cells: all distance covariates, elevation, and longitude were summarized relative to the segment centroid, and all land cover covariates were summarized by percentage of area in the trail segment corridor covered by the land cover type. Rather than examining trail density as in the grid-based model, we added an additional covariate to examine the characteristics of the trail segment itself: the percentage of the trail segment characterized by the "transportation" land cover category was used to indicate the function of the segment as a main travel route. We adjusted for spatial autocorrelation, determined the relative variable importance, and conducted multi-model averaging using the same approaches as in the grid-based model.

All analyses were conducted in R version 4.1.2 (R Core Team 2021), and analysis scripts are available (Mandeville et al. 2022). Key R packages included tidyverse for data management

(Wickham et al. 2019), sf for spatial analyses (Pebesma 2018), and glmulti for multi-model inference (Calcagno et al. 2010).

Results

3.1 Environmental covariates of citizen science activity.

As predicted, ease of area access was positively correlated with citizen science activity among grid cells (Figure 3). The total trail length per grid cell was the most important covariate and had a large positive effect on citizen science activity. Grid cells nearer to an area access point and to the closest population center were also positively associated with citizen science activity, though the effect of these covariates was smaller. Neither elevation nor the presence of recreational facilities were related to citizen science activity. The relationship with natural interest covariates did not match expectations; it was expected that developed and cultivated areas would be negatively associated with citizen science activity while land cover types often perceived to be more "natural" would be positively associated. Instead, the developed and cultivated land cover types had a large positive association with citizen science, while the wetland and forest land cover types were unimportant (Figure 3). The presence of freshwater had a small but important positive relationship to citizen science activity. These results were consistent among the six models that had a substantial level of support ($\Delta AIC_c < 2$), in total accounting for 65.9% of the weight of evidence (Supporting Information).

We found no evidence that citizen science activity was correlated with professional data collection activity among grid cells (Figure 1; Pearson correlation $r = 0.035$, $p = 0.232$). Three access covariates—distance to access points, longitude, and trail length—were important, but the effect sizes were much smaller than in the models of citizen science activity (Figure 4). As with citizen science, the presence of water and cultivated land had a small positive relationship to professional data collection. Unlike with citizen science, the presence of wetland land cover had a very small negative relationship to professional data collection and the developed land cover type did not have an important effect. These results were consistent among the eight models with a substantial level of support ($\Delta AIC_c < 2$), in total accounting for 57.2% of the weight of evidence (Supporting Information)

3.2 Relationship between citizen science and trail network.

Citizen science observations in the study area were on average made 26 meters (SD 42 meters) from the nearest trail, which was closer than professional data collection points (mean 53 meters; SD 68 meters) as well as a random distribution of sites, which would be expected to have a mean distance from the nearest trail of 72 meters.

There was high variability in the distance between observation points and the nearest trail within taxonomic groups. Still, the variation in mean distance values between taxonomic groups for citizen science data was consistent with the trend expected if observations tended to be made from a trail: taxonomic groups that are difficult to see from a distance (fungi, reptiles and amphibians) were associated with the smallest mean distance from the trail, while taxonomic groups that are easiest to spot from a distance (birds, mammals) were associated with the greatest distance. There was greater within-taxa variability and little evidence for a trend between taxonomic groups in the professional dataset (Figure 5).

3.3 Citizen science and other recreational trail use.

The tested covariates had limited ability to explain variation in citizen science activity among trail segments; four models had a substantial level of support, totaling 23.8% of the weight of evidence (Figure 6; Supporting Information). All effect sizes were relatively small compared to the grid-level models (Figure 6). Notably, most covariates that were important at the grid scale were not important to describe variation between trail segments; distance from the nearest access point, longitude, and proximity to freshwater were important at the grid level but had only a small and uncertain relationship to citizen science activity along trail segments. The most important variable was forest cover, which had a small negative relationship to citizen science activity.

The number of citizen science observations per trail segment corridor had no relationship to the number of reported Strava activities (Figure 1; Pearson correlation test, $r = -0.01$, $p = 0.414$). The relationship between the covariates and Strava activity differed substantially from their relationship to citizen science activity. The degree to which a trail segment functioned as a main travel route was the most important covariate, with a large positive relationship to Strava activity (Figure 7). In contrast, this covariate had a very small, marginally important positive effect on citizen science activity (Figure 6). Elevation had a strong positive association with Strava activity but a small negative association with citizen science activity. Wetland land cover had a positive association, while the association with forest and developed land cover was negative. These results were consistent among the twelve models with a substantial level of support ($\Delta AIC_c < 2$), together accounting for 39.2% of the weight of evidence (Supporting Information)

Discussion

We responded to calls for research on citizen science within protected areas by examining citizen science activity in a small natural area that serves as a regional citizen science hotspot. Our results illustrate that citizen science participation is spatially heterogeneous on a local scale. Ease of area access was the dominant landscape characteristic driving the distribution of citizen science in our study area, and a key component of accessibility is the use of a trail network to access observation sites. However, the distribution of citizen science activity along the trail network differed from that of other trail users. In general, citizen science activity was more evenly dispersed over a wider range of trail characteristics than other trail use; for example, citizen science participants were more likely than other trail users to spend time both in more developed parts of the natural area and also on less well-established paths that do not function as main travel routes.

The importance of area access is a notable result of our study. It is known that accessibility and natural interest are major regional determinants of citizen science activity, but our results are among the first to show that, within a small natural area, accessibility has a stronger relationship to citizen science activity than particular landscapes perceived as the most natural. To the contrary, citizen science activity was positively associated with cultivated and developed land within the area. This may be partially explained by the increased accessibility afforded by infrastructure in these areas. But it may also stem from an affinity for these land cover types, as suggested by recent findings that the integration of biodiversity with cultural and agricultural heritage plays an important role in communities' relationship to natural areas (Cusens et al. 2021). Proximity to water was positively associated with citizen science activity, as has previously been shown at regional scales and within urban areas (Boakes et al. 2016, Tiago et al. 2017). This could be explained by trends in either participant behavior (e.g., participants might prefer spending time near water or observing species found near water) or in species availability (e.g., landscapes containing freshwater may be more species-rich or afford greater detectability for species that are present).

The strong association between accessibility and citizen science participation offers some possibilities for improving the analysis of citizen science data. First, it may be possible to coarsely model the citizen science sampling intensity in local-scale analyses by accounting for access opportunities, as has been done previously at broader scales (Johnston et al. 2020, Cretois et al. 2020, Sicacha Parada et al. 2020). Further, accounting for accessibility could allow for better estimation of the citizen science sampling process on a regional scale. Our results suggest that even areas with a regionally high density of citizen science have likely not been sampled evenly, and that the regional citizen science sampling process could be better estimated by accounting for fine scale access patterns within citizen science hotspots.

At the same time, our results emphasize that mass participation citizen science can be a valuable supplement or, where needed, surrogate for biodiversity data from structured sampling. The citizen science data on GBIF include a greater number of species from all taxonomic groups than the equivalent professional datasets, covering a similarly diverse range of land cover types. In some ways, citizen science expands the reach of professional data collection; for instance, citizen science far outpaced professional data collection in the winter months in our study area. Winter ecology is recognized as understudied yet critical to conservation in the face of climate change (Studd et al. 2021, Sutton et al. 2021), so the contribution to this research area by citizen science is noteworthy. When comparing citizen science and professionally collected data, is important to note that the professional data on GBIF for our study area is almost certainly not a complete record of professional biodiversity data that has been collected in the area; while the value of openly sharing data is increasingly recognized, barriers still prevent much biodiversity data from being shared (Mandeville et al. 2021). For this reason, citizen science data are particularly valuable for their relatively easy accessibility.

In addition to informing more effective analysis of existing citizen science data, knowledge of citizen science activity patterns can be used by area managers to promote and guide future data collection. First, managers could use knowledge about citizen science trends to reach out to current participants to prompt collection of data to meet specific monitoring needs, for example by posting signs in areas regularly frequented by citizen science participants or communicating through customization features offered by citizen science platforms (Loen & Newton 2021, Gosal et al. 2021). Second, managers could identify areas of low citizen science activity to target for recruiting new participants (Weaver and Lawton 2017). For instance, recreational facilities were not closely associated with citizen science participation in our study area, so promotional information placed near such facilities could engage area visitors who do not yet participate in citizen science. Finally, managers may be able to prioritize professional data collection to complement citizen science by emphasizing areas of low citizen science activity.

Knowledge of spatial trends in citizen science activity can further inform overall recreational area management strategies. The needs and preferences of recreational visitors are regularly used to make management decisions about protected areas, but because different subsets of visitors prioritize different types of area management, it is challenging to fully capture the diverse needs of area visitors (Muñoz et al. 2020, Komossa et al. 2021). Because recreational preferences often play a key role in justifying area protection, it is important to accurately understand the full range of visitor experiences (Hornigold et al. 2016, Mancini et al. 2018, Cambria et al. 2021). Our results show that citizen science participants tend to use the trail network differently from other visitors, so their needs may be overlooked if not explicitly considered. Citizen science participants may even serve as a useful proxy to represent a broader group of area visitors, characterized by a desire to experience and learn about nature, who may be otherwise difficult to account for when assessing overall visitation trends (Havinga et al. 2020, Cambria et al. 2021).

Moving forward, there is much left to learn about citizen science participation at a local scale. The knowledge gained from modeling spatial patterns in citizen science participation is critical for understanding the ways that participants' motivations and behaviors manifest in spatially heterogeneous data collection (Sisneros-Kidd et al. 2021). Our results demonstrate spatial trends in citizen science participants' behavior. They also suggest new directions that could be followed up with research to address the motivation for this behavior: for instance, it would be useful to survey citizen science participants about their selection of trail routes or their on- and off-trail activity. Importantly, our goal of understanding citizen science activity in local protected areas responds to a commonly documented motivation for citizen science participation: participants regularly indicate that they want their data to be used for the conservation and management of places that they value (Ganzevoort et al. 2017, Larson et al. 2020, Maund et al. 2020). Through facilitation of improved data analysis and citizen science program implementation, a stronger understanding of citizen science activity within protected areas can be a step towards increasing the local conservation impact of participants' contributions.

Acknowledgements

We are grateful to Francesco Frassinelli from the Norwegian Institute for Nature Research for assistance with the Strava Metro data and to Trondheim Municipality, especially Kristian Morin, for assistance with spatial data layers. We also wish to thank the Transforming Citizen Science for Biodiversity research group, of which the authors are a part, for many valuable conversations about biodiversity citizen science. Funding for CM was provided by the Digital Transformation Initiative of the Norwegian University of Science and Technology. Finally, we extend deep gratitude to all the citizen science participants and practitioners who have played a role in assembling the biodiversity data used in this paper.

Conflict of interest

The authors declare no conflicts of interest.

Author contributions

CM conceived of the idea, analyzed the data, and led the writing of the manuscript. EN and AF supported the conceptual development and writing of the manuscript. All authors contributed to the drafts and gave approval for publication.

Data availability

Biodiversity data are obtained from the Global Biodiversity Information Facility and are available here: https://doi.org/10.15468/dl.pd3tce (GBIF 2021). This report includes aggregated and de-identified data from Strava Metro. R scripts for the analyses are available from Open Science Framework: 10.17605/OSF.IO/BGF3D (Mandeville et al. 2022).

References

Ahlstrøm A., Bjørkelo K., Frydenlund J. 2014. AR5 klassifikasjonssystem: Klassifikasjon av Arealressurser. Rapport 6. Ås: Norsk institutt for skog og landskap.

Alfonso L, Gharesifard M, Wehn U. 2022. Analysing the value of environmental citizen generated data: Complementarity and cost per observation. Journal of Environmental Management 303: 114157.

Armsworth PR, Cantú-Salazar L, Parnell M, Davies ZG, Stoneman R. 2011. Management costs for small protected areas and economies of scale in habitat conservation. Biological Conservation 144: 423–429.

Baldwin RF, Fouch NT. 2018. Understanding the Biodiversity Contributions of Small Protected Areas Presents Many Challenges. Land 7: 123.

Ball-Damerow JE, Brenskelle L, Barve N, Soltis PS, Sierwald P, Bieler R, LaFrance R, Ariño AH, Guralnick RP. 2019. Research applications of primary biodiversity databases in the digital age. PLOS ONE 14: e0215794.

Bardos DC, Guillera-Arroita G, Wintle BA. 2015. Valid auto-models for spatially autocorrelated occupancy and abundance data. Methods in Ecology and Evolution 6: 1137–1149.

Binley AD, Proctor CA, Pither R, Davis SA, Bennett JR. 2021. The unrealized potential of dommunity science to support research on the resilience of protected areas. Conservation Science and Practice 3: e376.

Boakes EH, Gliozzo G, Seymour V, Harvey M, Smith C, Roy DB, Haklay M. 2016. Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. Scientific Reports 6: 33051.

Bonnet P, Joly A, Faton J-M, Brown S, Kimiti D, Deneu B, Servajean M, Affouard A, Lombardo J-C, Mary L, Vignau C, Munoz F. 2020. How citizen scientists contribute to monitor protected areas thanks to automatic plant identification tools. Ecological Solutions and Evidence 1: e12023.

Bonney R. 2021. Expanding the Impact of Citizen Science. BioScience 71: 448–451.
Burnham KP, Anderson DR. 2002. Model selection and multimodel inference: A practical information-theoretic approach. New York: Springer. 488 p.

Buta N, Holland SM, Kaplanidou K. 2014. Local communities and protected areas: The mediating role of place attachment for pro-environmental civic engagement. Journal of Outdoor Recreation and Tourism 5–6: 1–10.

Calcagno, V., & de Mazancourt, C. (2010). glmulti: An R Package for Easy Automated Model Selection with (Generalized) Linear Models. Journal of Statistical Software, 34(12), 1–29. https://doi.org/10.18637/jss.v034.i12

Callaghan C, Gawlik D. 2015. Efficacy of eBird data as an aid in conservation planning and monitoring. Journal of Field Ornithology 86: 298–304.

Callaghan CT, Poore AGB, Mesaglio T, Moles AT, Nakagawa S, Roberts C, Rowley JJL, Vergés A, Wilshire JH, Cornwell WK. 2021. Three Frontiers for the Future of Biodiversity Research Using Citizen Science Data. BioScience 71: 55–63.

Callaghan CT, Rowley JJL, Cornwell WK, Poore AGB, Major RE. 2019. Improving big citizen science data: Moving beyond haphazard sampling. PLOS Biology 17: e3000357.

Cambria VE, Campagnaro T, Trentanovi G, Testolin R, Attorre F, Sitzia T. 2021. Citizen Science Data to Measure Human Use of Green Areas and Forests in European Cities. Forests 12: 779.

Cheung SY, Leung Y-F, Larson LR. 2022. Citizen science as a tool for enhancing recreation research in protected areas: Applications and opportunities. Journal of Environmental Management 305: 114353.

Cooper CB, Shirk J, Zuckerberg B. 2014. The Invisible Prevalence of Citizen Science in Global Research: Migratory Birds and Climate Change. PLOS ONE 9: e106508.

Cretois B, Simmonds EG, Linnell JDC, van Moorter B, Rolandsen CM, Solberg EJ, Strand O, Gundersen V, Roer O, Rød JK. 2021. Identifying and correcting spatial bias in opportunistic citizen science data for wild ungulates in Norway. Ecology and Evolution 11: 15191–15204.

Cusens J, Barraclough AMD, Måren IE. 2021. Participatory mapping reveals biocultural and nature values in the shared landscape of a Nordic UNESCO Biosphere Reserve. People and Nature 2021.

Danielsen F, Burgess ND, Jensen PM, Pirhofer-Walzl K. 2010. Environmental monitoring: the scale and speed of implementation varies according to the degree of peoples involvement. Journal of Applied Ecology 47: 1166–1168.

Di Cecco GJ, Barve V, Belitz MW, Stucky BJ, Guralnick RP, Hurlbert AH. 2021. Observing the Observers: How Participants Contribute Data to iNaturalist and Implications for Biodiversity Science. BioScience 71: 1179–1188.

Dobson ADM, Milner-Gulland EJ, Aebischer NJ, Beale CM, Brozovic R, Coals P, Critchlow R, Dancer A, Greve M, Hinsley A, Ibbett H, Johnston A, Kuiper T, Le Comber S, Mahood SP, Moore JF, Nilsen EB, Pocock MJO, Quinn A, Travers H, Wilfred P, Wright J, Keane A. 2020. Making Messy Data Work for Conservation. One Earth 2: 455–465.

Dreiss LM, Malcom JW. Title identifying key federal, state, and private lands strategies for achieving 30 × 30 in the United States. Conservation Letters: e12849.

Feldman MJ, Imbeau L, Marchand P, Mazerolle MJ, Darveau M, Fenton NJ. 2021. Trends and gaps in the use of citizen science derived data as input for species distribution models: A quantitative review. PLOS ONE 16: e0234587.

Gaget E, Johnston A, Pavón-Jordán D, Lehikoinen AS, Sandercock BK, Soultan A, Božič L, Clausen P, Devos K, Domsa C, Encarnação V, Faragó S, Fitzgerald N, Frost T, Gaudard C, Gosztonyi L, Haas F, Hornman M, Langendoen T, Ieronymidou C, Luigujõe L, Meissner W, Mikuska T, Molina B, Musilová Z, Paquet J-Y, Petkov N, Portolou D, Ridzoň J, Sniauksta L, Stīpniece A, Teufelbauer N, Wahl J, Zenatello M, Brommer JE. Protected area characteristics that help waterbirds respond to climate warming. Conservation Biology: 10.1111/cobi.13877.

Ganzevoort W, van den Born RJG, Halffman W, Turnhout S. 2017. Sharing biodiversity data: citizen scientists' concerns and motivations. Biodiversity and Conservation 26: 2821–2837.

GBIF. 3 August 2021. GBIF Occurrence Download https://doi.org/10.15468/dl.pd3tce

Geldmann J, Heilmann-Clausen J, Holm TE, Levinsky I, Markussen B, Olsen K, Rahbek C, Tøttrup AP. 2016. What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. Diversity and Distributions 22: 1139–1149.

Gosal AS, McMahon JA, Bowgen KM, Hoppe CH, Ziv G. 2021. Identifying and Mapping Groups of Protected Area Visitors by Environmental Awareness. Land 10: 560.

Gurney GG, Darling ES, Ahmadia GN, Agostini VN, Ban NC, Blythe J, Claudet J, Epstein G, Estradivari, Himes-Cornell A, Jonas HD, Armitage D, Campbell SJ, Cox C, Friedman WR, Gill D, Lestari P, Mangubhai S, McLeod E, Muthiga NA, Naggea J, Ranaivoson R, Wenger A, Yulianto I, Jupiter SD. 2021. Biodiversity needs every tool in the box: use OECMs. Nature 595: 646–649.

Häkkilä M. 2021. Are small protected habitat patches within boreal production forests effective in conserving species richness, abundance and community composition? A systematic review. Environmental Evidence 10: 2.

Halliwell P, Whipple S, Bowser G. 2021. Learning to love protected areas: Citizen science projects inspire place attachment for diverse students in United States National Parks. Journal of Geoscience Education 0: 1–9.

Havinga I, Bogaart PW, Hein L, Tuia D. 2020. Defining and spatially modelling cultural ecosystem services using crowdsourced data. Ecosystem Services 43: 101091.

Hlásny T, Mokroš M, Dobor L, Merganičová K, Lukac M. 2021. Fine-scale variation in projected climate change presents opportunities for biodiversity conservation in Europe. Scientific Reports 11: 17242.

Hornigold K, Lake I, Dolman P. 2016. Recreational Use of the Countryside: No Evidence that High Nature Value Enhances a Key Ecosystem Service. PloS One 11: e0165043.

IUCN WCPA Task Force on OECMs. 2019. Recognising and reporting other effective area-based conservation measures. IUCN, International Union for Conservation of Nature.

Jansujwicz JS, Calhoun AJK, Bieluch KH, McGreavy B, Silka L, Sponarski C. 2021. Localism “Reimagined”: Building a Robust Localist Paradigm for Overcoming Emerging Conservation Challenges. Environmental Management 67: 91–108.

Johnston A, Moran N, Musgrove A, Fink D, Baillie SR. 2020. Estimating species distributions from spatially biased citizen science data. Ecological Modelling 422: 108927.

Kays R, McShea WJ, Wikelski M. 2020. Born-digital biodiversity data: Millions and billions. Diversity and Distributions 26: 644–648.

Kendal D, Zeeman BJ, Ikin K, Lunt ID, McDonnell MJ, Farrar A, Pearce LM, Morgan JW. 2017. The importance of small urban reserves for plant conservation. Biological Conservation 213: 146–153.

Knape J, Coulson SJ, van der Wal R, Arlt D. 2022. Temporal trends in opportunistic citizen science reports across multiple taxa. Ambio 51: 183–198.

Koen EL, Newton EJ. 2021. Outreach increases detections of an invasive species in a crowdsourced monitoring program. Biological Invasions 23: 2611–2620.

Komossa F, Wartmann FM, Verburg PH. 2021. Expanding the toolbox: Assessing methods for local outdoor recreation planning. Landscape and Urban Planning 212: 104105.

Larson LR, Cooper CB, Futch S, Singh D, Shipley NJ, Dale K, LeBaron GS, Takekawa JY.

2020. The diverse motivations of citizen scientists: Does conservation emphasis grow as volunteer participation progresses? Biological Conservation 242: 108428.

Mair L, Ruete A. 2016. Explaining Spatial Variation in the Recording Effort of Citizen Science Data across Multiple Taxa. PLOS ONE 11: e0147796.

Mancini F, Coghill GM, Lusseau D. 2019. Quantifying wildlife watchers' preferences to investigate the overlap between recreational and conservation value of natural areas. Journal of Applied Ecology 56: 387–397.

Mandeville CP, Finstad AG. 2021. Community science supports research on protected are resilience. Conservation Science and Practice 3: e442.

Mandeville CP, Koch W, Nilsen EB, Finstad AG. 2021. Open Data Practices among Users of Primary Biodiversity Data. BioScience 71: 1128–1147.

Mandeville CP, Nilsen EB, Finstad AG. 2022. Spatial distribution of biodiversity citizen science in a natural area. Open Science Framework 10.17605/OSF.IO/BGF3D.

Maund PR, Irvine KN, Lawson B, Steadman J, Risely K, Cunningham AA, Davies ZG. 2020. What motivates the masses: Understanding why people contribute to conservation citizen science projects. Biological Conservation 246: 108587.

Maxwell SL, Cazalis V, Dudley N, Hoffmann M, Rodrigues ASL, Stolton S, Visconti P, Woodley S, Kingston N, Lewis E, Maron M, Strassburg BBN, Wenger A, Jonas HD, Venter O, Watson JEM. 2020. Area-based conservation in the twenty-first century. Nature 586: 217–227.

Millar EE, Hazell EC, Melles SJ. 2019. The 'cottage effect' in citizen science? Spatial bias in aquatic monitoring programs. International Journal of Geographical Information Science 33: 1612–1632.

Mueller MA, Drake D, Allen ML. 2019. Using citizen science to inform urban canid management. Landscape and Urban Planning 189: 362–371.

Newman G, Chandler M, Clyde M, McGreavy B, Haklay M, Ballard H, Gray S, Scarpino R, Hauptfeld R, Mellor D, Gallo J. 2017. Leveraging the power of place in citizen science for effective conservation decision making. Biological Conservation 208: 55–64.

Muñoz L, Hausner VH, Runge C, Brown G, Diagle R. 2020. Using crowdsourced spatial data from Flickr vs. PPGIS for understanding nature's contribution to people in Southern Norway. People and Nature 2: 437-449.

Oldekop JA, Holmes G, Harris WE, Evans KL. 2016. A global assessment of the social and conservation outcomes of protected areas. Conservation Biology 30: 133–141.

Pebesma E (2018). "Simple Features for R: Standardized Support for Spatial Vector Data." The R Journal, 10(1), 439–446. doi: 10.32614/RJ-2018-009, https://doi.org/10.32614/RJ-2018-009.

Petersen TK, Speed JDM, Grøtan V, Austrheim G. 2021. Species data for understanding biodiversity dynamics: The what, where and when of species occurrence data collection. Ecological Solutions and Evidence 2: e12048.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rapacciuolo G, Young A, Johnson R. 2021. Deriving indicators of biodiversity change from unstructured community-contributed data. Oikos 130: 1225–1239.

Riva F, Fahrig L. 2022. The disproportionately high value of small patches for biodiversity conservation. Conservation Letters: e12881.

Rodríguez-Rodríguez D, Sánchez-Espinosa A, Abdul Malak D. 2021. Potential contribution of OECMs to international area-based conservation targets in a biodiversity rich country, Spain. Journal for Nature Conservation 62: 126019.

Romo H, García-Barros E, Lobo JM. 2006. Identifying recorder-induced geographic bias in an Iberian butterfly database. Ecography 29: 873–885. 37.

Salmon RA, Rammell S, Emeny MT, Hartley S. 2021. Citizens, Scientists, and Enablers: A Tripartite Model for Citizen Science Projects. Diversity 13: 309.

Sicacha-Parada J, Steinsland I, Cretois B, Borgelt J. 2021. Accounting for spatial varying sampling effort due to accessibility in Citizen Science data: A case study of moose in Norway. Spatial Statistics 42: 100446.

Sisneros-Kidd AM, D'Antonia A, Monz C, Mitrovich M. 2021. Improving understanding and management of the complex relationship between visitor motivations and spatial behaviors in parks and protected areas. Journal of Environmental Management 280: 111841.

Studd EK, Bates AE, Bramburger AJ, Fernandes T, Hayden B, Henry HAL, Humphries MM, Martin R, McMeans BC, Moise ERD, O'Sullivan AM, Sharma S, Sinclair BJ, Sutton AO, Templer PH, Cooke SJ. 2021. Nine Maxims for the Ecology of Cold-Climate Winters. BioScience 71: 820–830.

Sutton AO, Studd EK, Fernandes T, Bates AE, Bramburger AJ, Cooke SJ, Hayden B, Henry HAL, Humphries MM, Martin R, McMeans B, Moise E, O'Sullivan AM, Sharma S, Templer PH. 2021. Frozen out: unanswered questions about winter biology. Environmental Reviews.

Tiago P, Ceia-Hasse A, Marques TA, Capinha C, Pereira HM. 2017. Spatial distribution of citizen science casuistic observations for different taxonomic groups. Scientific Reports 7: 12832.

Tulloch AIT, Mustin K, Possingham HP, Szabo JK, Wilson KA. 2013. To boldly go where no volunteer has gone before: predicting volunteer activity to prioritize surveys at the landscape scale. Diversity and Distributions 19: 465–480.

Vimal R, Navarro LM, Jones Y, Wolf F, Le Moguédec G, Réjou-Méchain M. 2021. The global distribution of protected areas management strategies and their complementarity for biodiversity conservation. Biological Conservation 256: 109014.

Volenec ZM, Dobson AP. 2020. Conservation value of small reserves. Conservation Biology 34: 66-79.

Weaver DB, Lawton LJ. 2017. A new visitation paradigm for protected areas. Tourism Management 60: 140–146.

Welvaert M, Caley P. 2016. Citizen surveillance for environmental monitoring: combining the efforts of citizen science and crowdsourcing in a quantitative data framework. SpringerPlus 5: 1890.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Wintle BA, Kujala H, Whitehead A, Cameron A, Veloz S, Kukkala A, Moilanen A, Gordon A, Lentini PE, Cadenhead NCR, Bekessy SA. 2019. Global synthesis of conservation studies reveals the importance of small habitat patches for biodiversity. Proceedings of the National Academy of Sciences 116: 909–914.

Zulian V, Miller DAW, Ferraz G. 2021. Integrating citizen-science and planned-survey data improves species distribution estimates. Diversity and Distributions 27: 2498–2509.

Figures

Figure 1. Map of study area in Trondheim, Central Norway. (a) indicates the position of the study area in relation to the population center of Trondheim. (b) indicates the density of reported Strava activities per trail segment. (c) and (d) indicate the density of citizen science data and professional data per grid cell, respectively.

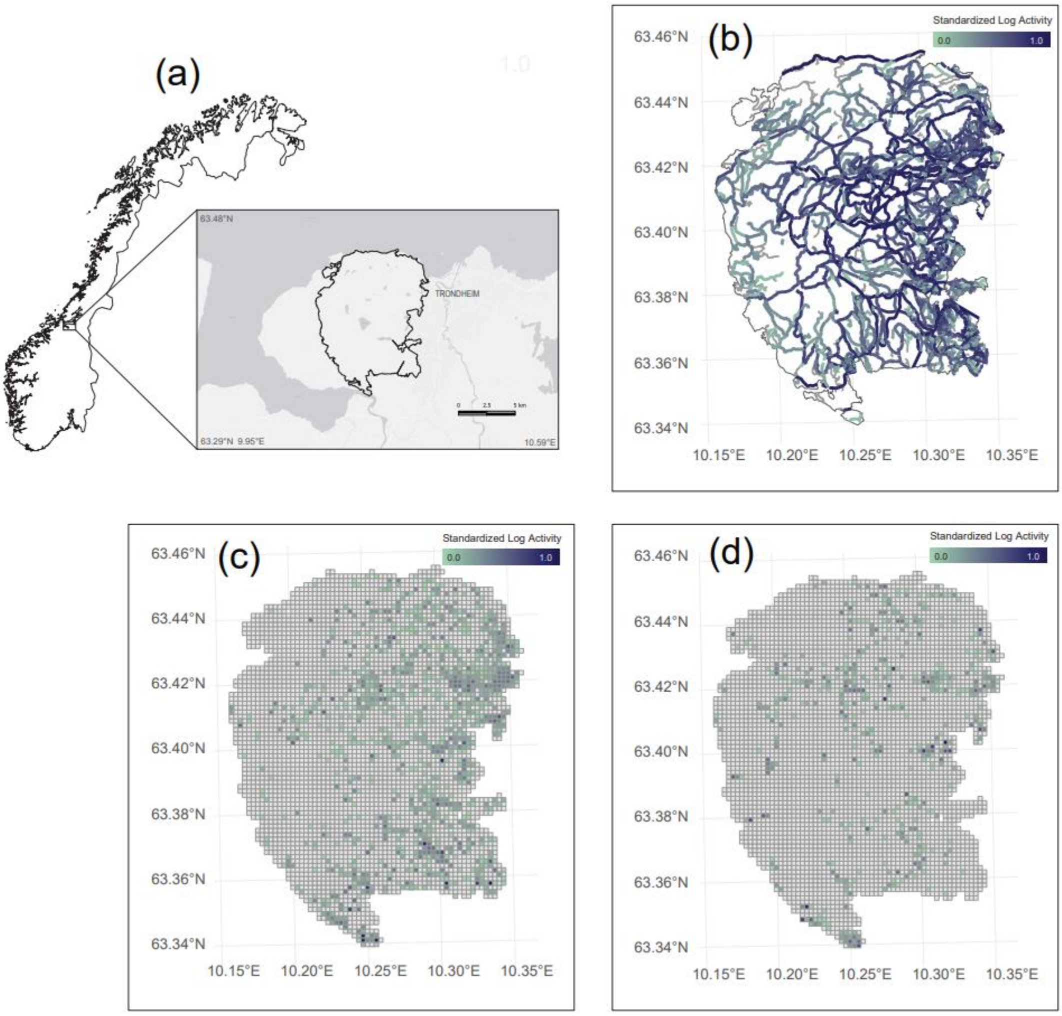


Figure 2. (a) Number of observations from each taxonomic group for citizen science and professional data; (b) number of species from each taxonomic group for citizen science and professional data; (c) month of observation for citizen science and professional data; (d) land cover type for citizen science and professional observations, shown relative to the availability of land cover types within the area.

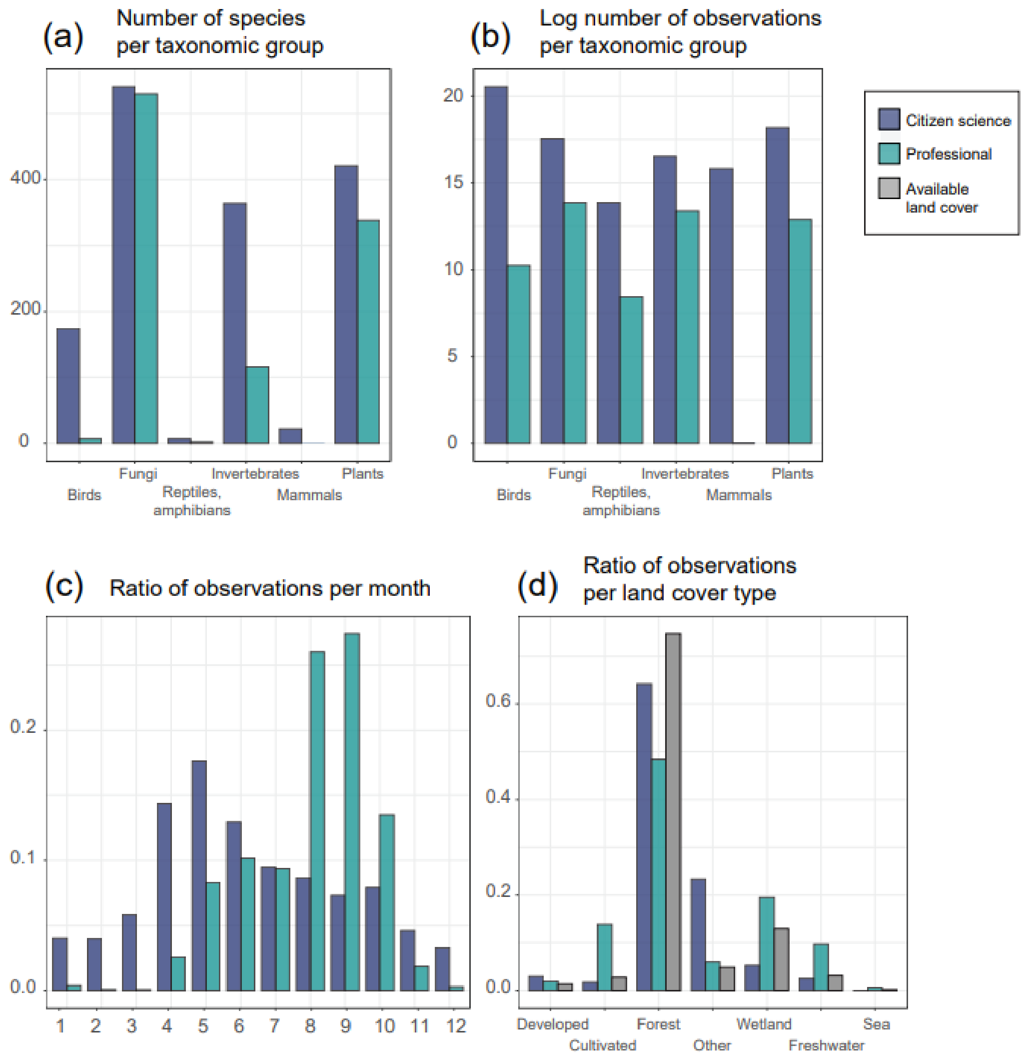


Figure 3. Modeled effect of all covariates on the number of citizen science observations per grid cell, modeled with a negative binomial generalized linear model structure. All six models with substantial support ($\Delta AIC_c < 2$) are shown. Ribbons indicate a 95% confidence interval. Relative variable importance, calculated with a weighted average of all models, is indicated for each covariate.

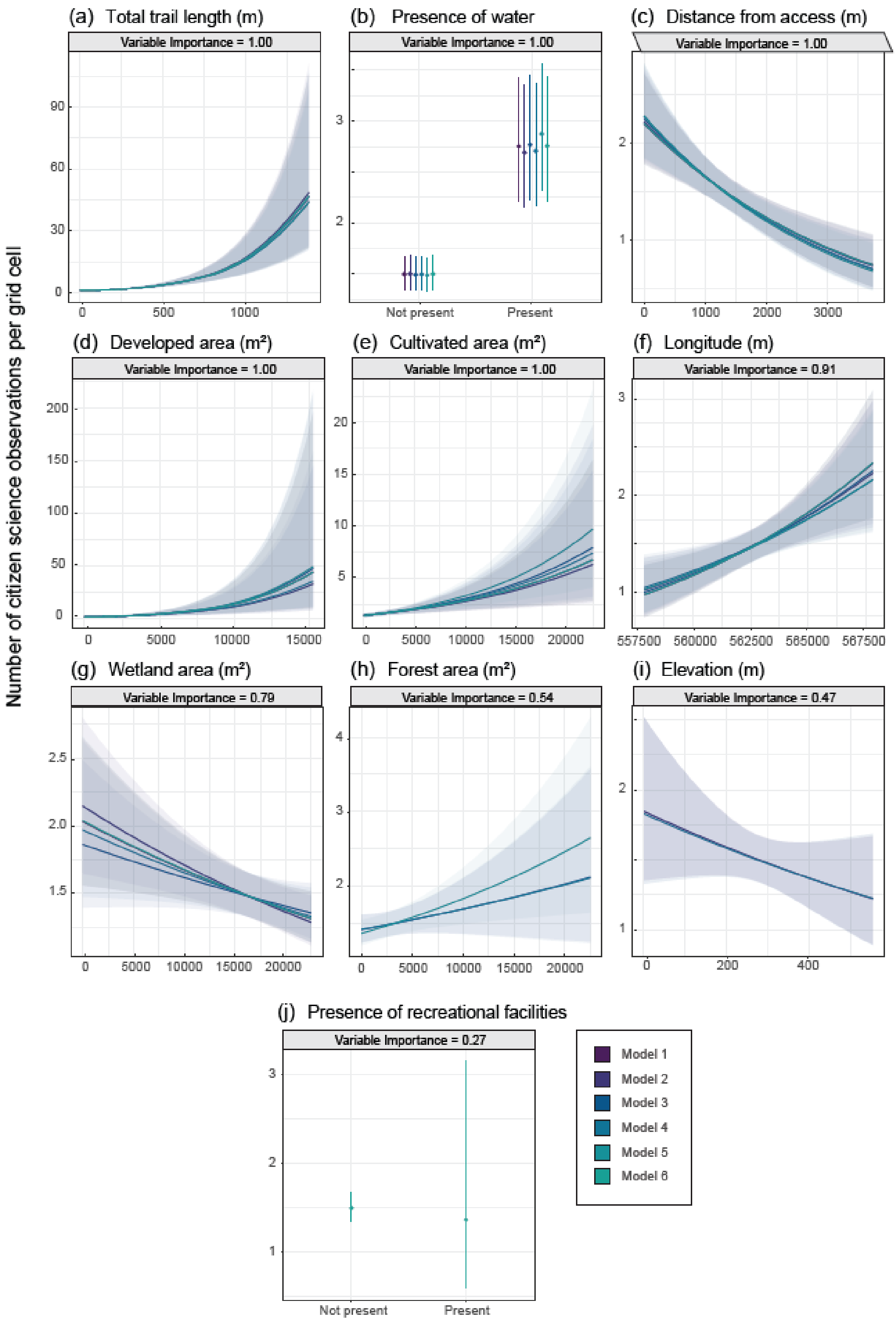


Figure 4. Modeled effect of all covariates on the number of professional biodiversity observations per grid cell, modeled with a negative binomial generalized linear model structure. All eight models with substantial support ($\Delta AIC_c < 2$) are shown. Ribbons indicate a 95% confidence interval. Relative variable importance, calculated with a weighted average of all models, is indicated for each covariate.

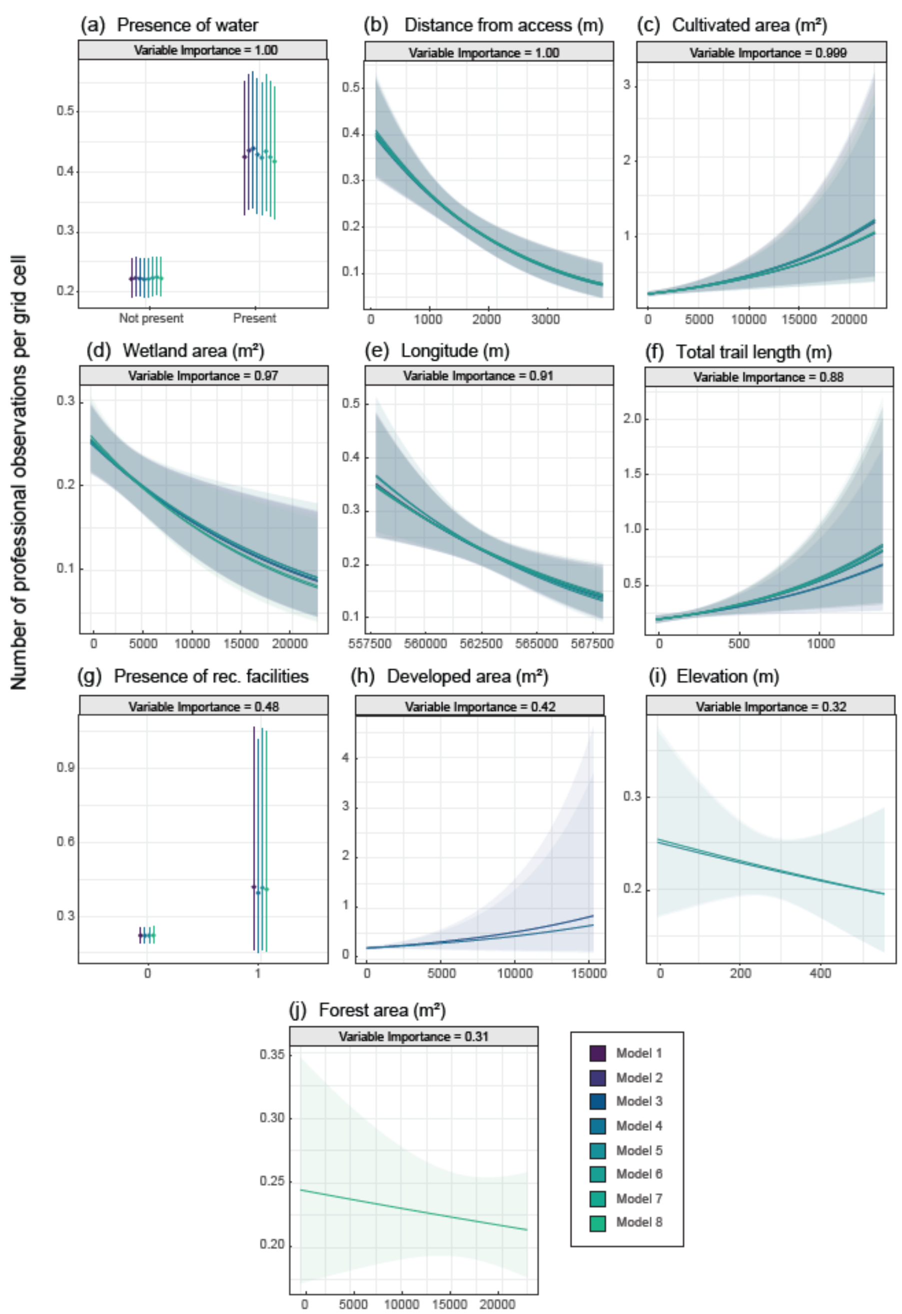


Figure 5. Distance between reported observation coordinates and the nearest trail for citizen science and professional data. Letters indicate groups of taxonomic classes that were identified as distinct from each other at a $p < 0.05$ level with a Tukey HSD test.

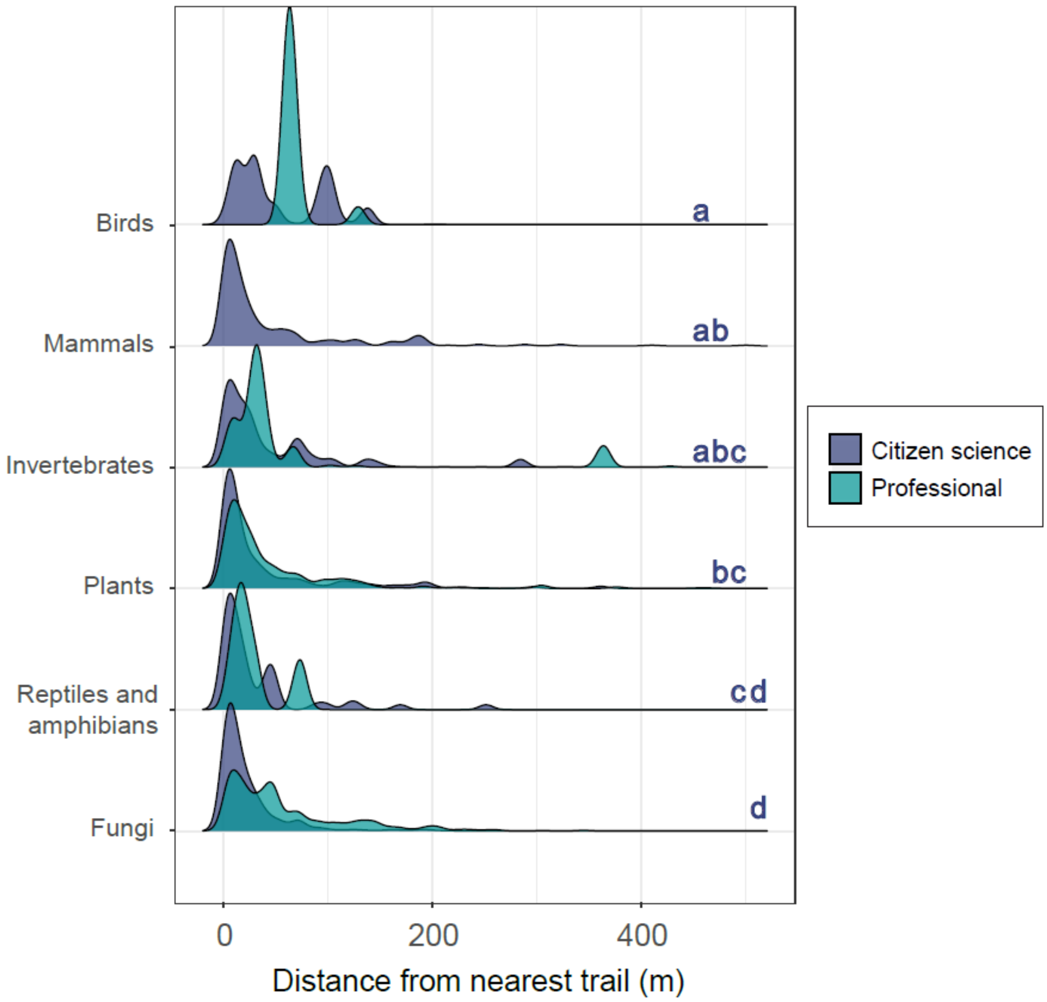


Figure 6. Modeled effect of all covariates on the number of citizen science observations per 300-meter-wide trail segment corridor, standardized by segment length, modeled with a negative binomial generalized linear model structure. All four models with substantial support ($\Delta AIC_c < 2$) are shown. Ribbons indicate a 95% confidence interval. Relative variable importance, calculated with a weighted average of all models, is indicated for each covariate.

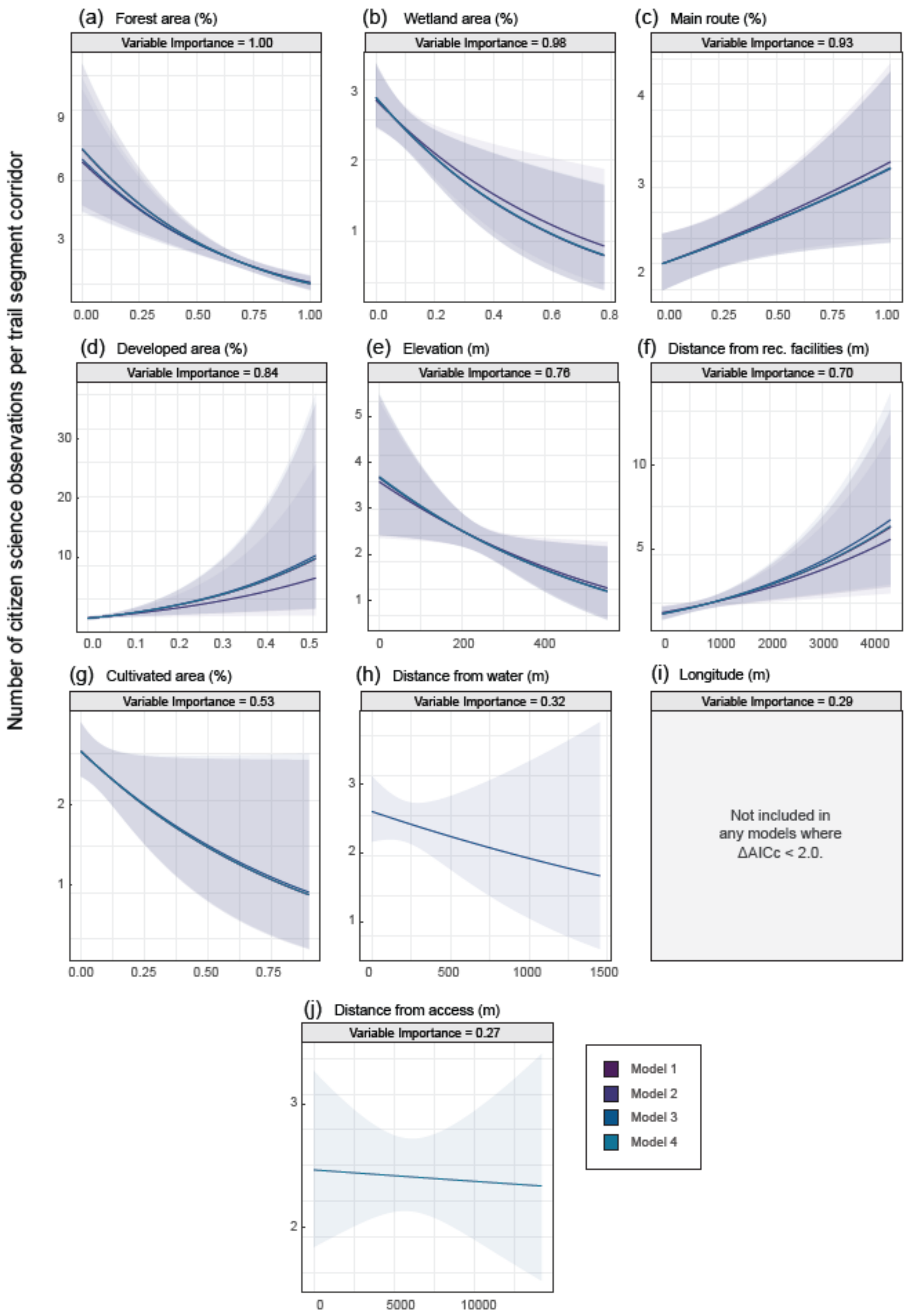
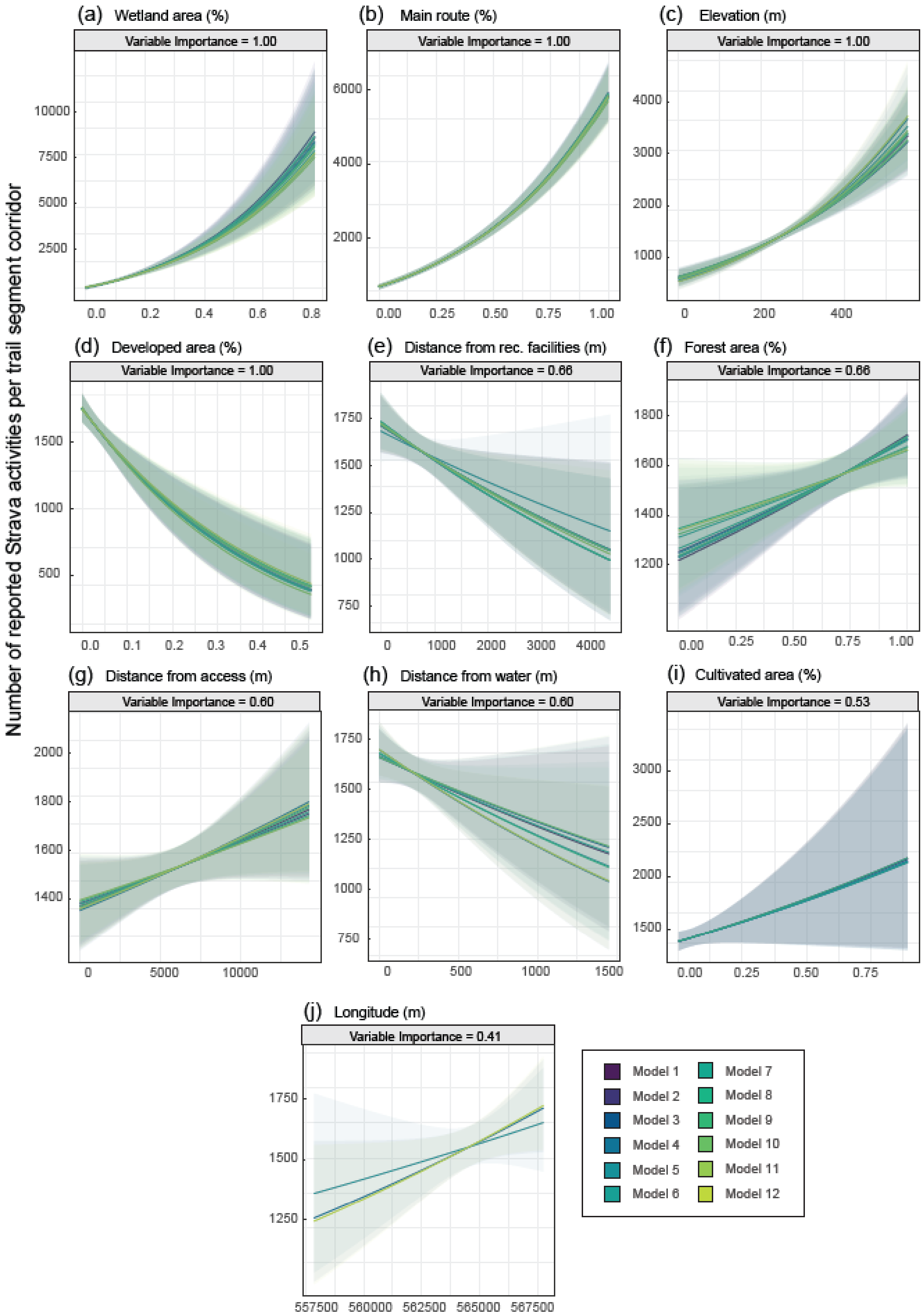


Figure 7. Modeled effect of all covariates on the number of reported Strava activities per trail segment corridor, modeled with a negative binomial generalized linear model structure. All twelve models with substantial support ($\Delta AIC_c < 2$) are shown. Ribbons indicate a 95% confidence interval. Relative variable importance, calculated with a weighted average of all models, is indicated for each covariate.



Supporting Information

Fine-scale spatial distribution of biodiversity citizen science in a natural area depends on area accessibility and differs from other recreational area use

C.P. Mandeville, E.B. Nilsen, A.G. Finstad

Contents

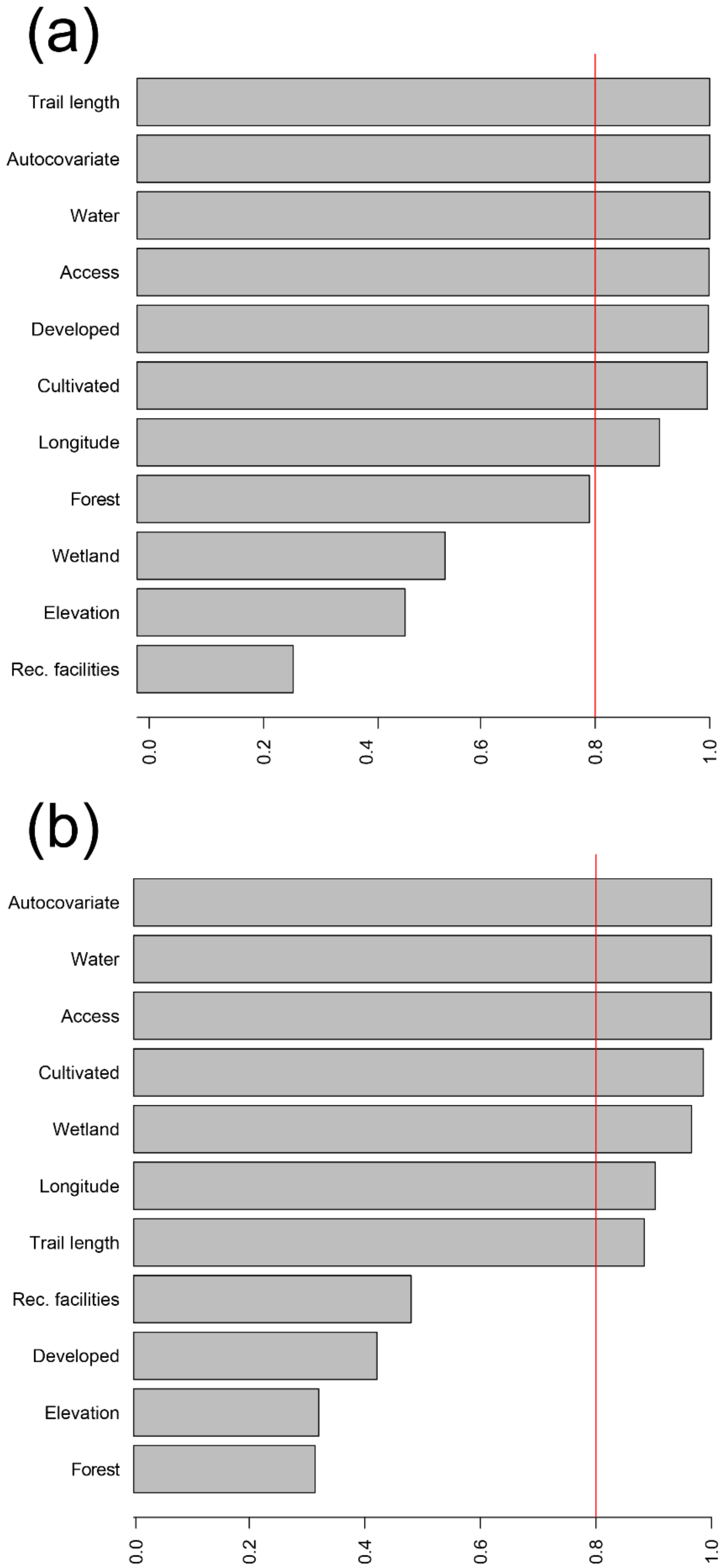
S1.1 Data contributors to the biodiversity data accessed from GBIF, after filtering for inclusion in this study. *n* indicates the number of included data points contributed by the indicated data source. Data sources where *n* = 0 were present in the study area on GBIF but all data from these sources were excluded through the filtering described in the Methods section.

Data source	**Type of source**	***n* (included)**
Norwegian Species Observation Service	Citizen science - opportunistic	40376
eBird Observation Dataset	Citizen science - opportunistic	3450
iNaturalist Research-grade Observations	Citizen science - opportunistic	299
Pl@ntNet	Citizen science - opportunistic	50
Skandobs	Citizen science - opportunistic	11
Naturgucker	Citizen science - opportunistic	11
Observation.org	Citizen science - opportunistic	6
Vascular plant herbarium TRH, NTNU University Museum	Professional	492
Lichen herbarium TRH, NTNU University Museum	Professional	291
Terrestrial and limnic invertebrates systematic collection, NTNU University Museum	Professional	246
Mycology herbarium TRH, NTNU University Museum	Professional	218
Fungi field notes, Oslo (O)	Professional	170
NINA insect database	Professional	120
International Barcode of Life project (iBOL)	Professional	100
BioFokus	Professional	99
Geographically tagged INSDC sequences	Professional	67
Bryophyte herbarium TRH, NTNU University Museum	Professional	53
Lichen field notes, Oslo (O)	Professional	29
Royal Botanic Garden Edinburgh Living Plant Collections (E)	Professional	28
Bird collection NTNU University Museum	Professional	27
Lichen herbarium, Oslo (O) UiO	Professional	25
Mycology herbarium, Oslo (O) UiO	Professional	23
NHMO DNA Bank Vascular plants collection	Professional	19
Vascular Plant Herbarium, Oslo (O) UiO	Professional	15
NHMO DNA Bank Fungi and Lichens collection	Professional	7

Danish Mycological Society, fungal records database	Professional	6
Artsprosjekt: hypogeous_macrofungi	Professional	4
Bryophyte Herbarium, Oslo (O) UiO	Professional	4
Herpetile collection NTNU University Museum	Professional	4
Entomological collections, UiB	Professional	3
Lichen herbarium, UiB	Professional	2
Algae herbarium TRH, NTNU University Museum	Professional	1
Mycology collection, Norwegian Forest and Landscape Institute	Professional	1
Reptilia notes, NTNU University Museum	Professional	1
Seabirds in Norway - Estimated population sizes	Professional	1
The cryptogamy collection (PC) at the Herbarium of the Muséum national d'Histoire Naturelle (MNHN - Paris)	Professional	1
Tropicos Specimen Data	Professional	1
Vascular plant herbarium (KMN) UiA	Professional	1
Norwegian Biodiversity Information Centre - Other datasets	Citizen science - structured	0
Algae collection, Oslo (O) UiO	Professional	0
Algae, Norwegian College of Fishery Science	Professional	0
Birds ringed with Norwegian rings 1914-1960	Professional	0
Birds ringed with Norwegian rings 1961-1990	Professional	0
Bryophyte herbarium, UiT Tromsø Museum	Professional	0
Collembola collection of Arne Fjellberg, Norway	Professional	0
Entomology collection, UiT Tromsø Museum	Professional	0
Entomology Division, Yale Peabody Museum	Professional	0
Entomology, Natural History Museum, University of Oslo	Professional	0
Fish collection NTNU University Museum	Professional	0
Herbarium GB, University of Gothenburg	Professional	0
Huitfeldt Kaas: Freswhater fish distribution in Norway 1918	Professional	0
Ims fish tag database	Professional	0
Lichen herbarium, UiT Tromsø Museum	Professional	0
Limnic freshwater benthic invertebrates biogeographical mapping/inventory NTNU University Museum	Professional	0

Limnic freshwater pelagic invertebrates biogeographical mapping/inventory NTNU University Museum	Professional	0
Lund Botanical Museum (LD	Professional	0
Mammal collection NTNU University Museum	Professional	0
Marine invertebrate collection NTNU University Museum	Professional	0
Mycology herbarium, UiT Tromsø Museum	Professional	0
National fish tag database	Professional	0
NHMO DNA Bank Fish and Herptile collection	Professional	0
NINA Vanndata fisk	Professional	0
NINA Vanndata øvrige arter	Professional	0
NMNH Extant Specimen Records	Professional	0
Notes from the Mycology Herbarium, Oslo (O)	Professional	0
NSW AVH data	Professional	0
Provincial Museum of Alberta, Edmonton, AB, Canada. Birds (Aves)	Professional	0
SEAPOP - Last observation per locality in breeding season	Professional	0
Thrips (Thysanoptera) in Norway	Professional	0
Vascular plant field notes, NTNU University Museum	Professional	0
Vascular plant herbarium, UiT Tromsø Museum	Professional	0
Vascular Plants, Field notes, Oslo (O)	Professional	0
Vascular Plants, Museum of Archaeology, University of Stavanger	Professional	0

S2.1 Model-averaged relative variable importance of each covariate for the models of (a) citizen science and (b) professional biodiversity data observations among grid cells in the study area.



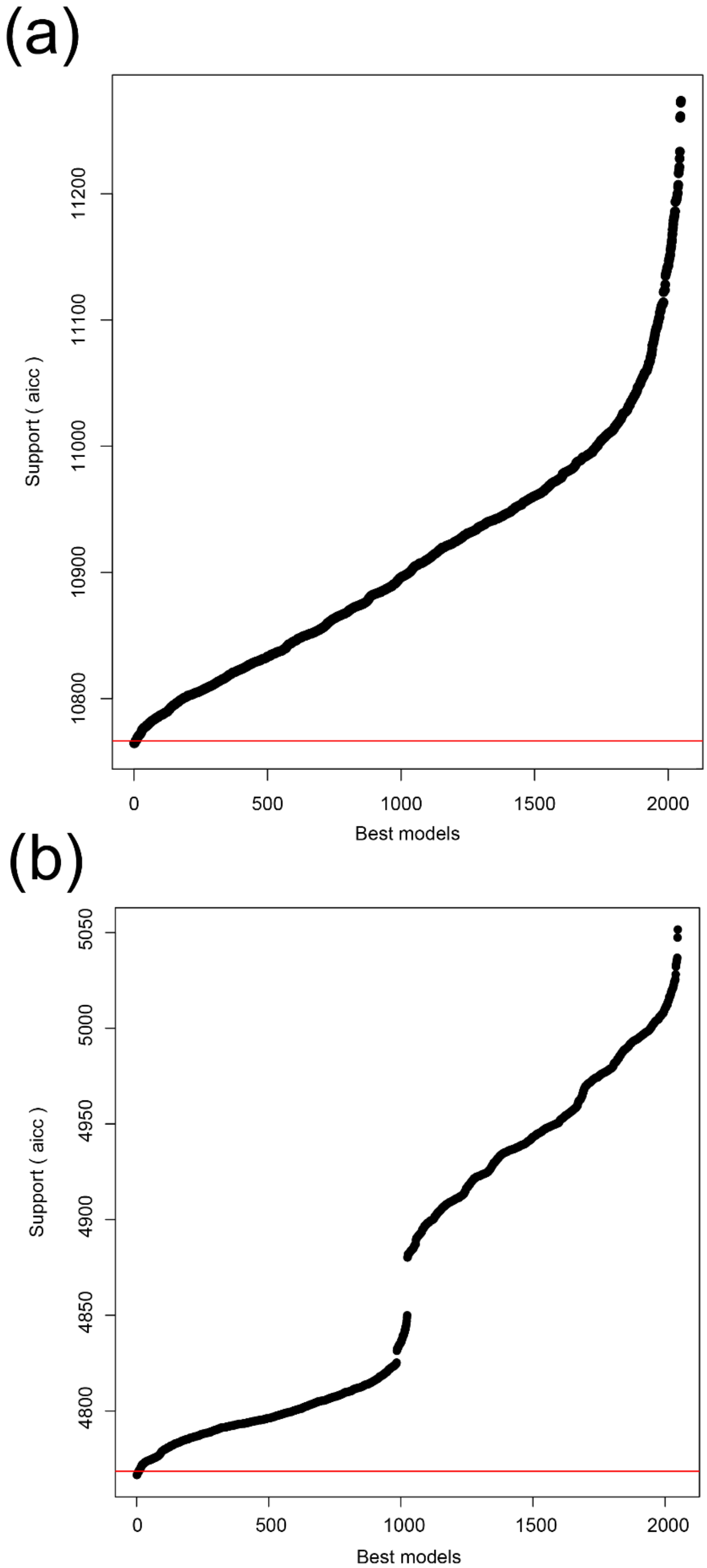
S2.2 Importance and model-averaged estimates and standard error for each covariate in the (a) citizen science and (b) professional model of biodiversity observations among grid cells in the study area.

(a) Citizen science			
	Importance	Estimate	Standard error
Intercept	1.0000	-40.8046	18.1530
Trails	1.0000	2.8719e-03	3.2799e-04
Water	1.0000	0.6120	0.1286
Access	0.9988	-3.1417e-04	7.6526e-05
Developed	0.9976	2.1905e-04	5.2257e-05
Cultivated	0.9959	7.3890e-05	2.1878e-05
Longitude	0.9126	7.3116e-05	3.2137e-05
Forest	0.7900	-1.4755e-05	1.4731e-05
Wetland	0.5379	-3.4152e-04	1.4731e-05
Elevation	0.4681	-3.4152e-04	5.0224e-04
Facilities	0.2726	-2.5224e-02	0.1365
(b) Professional			
	Importance	Estimate	Standard error
Intercept	1.0000	45.9149	22.0541
Water	0.9997	0.6563	0.1528
Access	0.9996	-4.3222e-04	9.7075e-05
Cultivated	0.9857	7.1324e-05	2.450e-05
Wetland	0.9655	-4.6951e-05	1.8090e-05
Longitude	0.9103	-8.4016e-05	3.9093e-05
Trails	0.8837	9.0124e-04	4.7133e-04
Facilities	0.4802	0.2920	0.4288
Developed	0.4212	4.01662e-05	5.7759e-05
Elevation	0.3205	-1.3417e-04	3.2941e-04
Forest	0.3139	-2.1098e-06	5.0725e-06

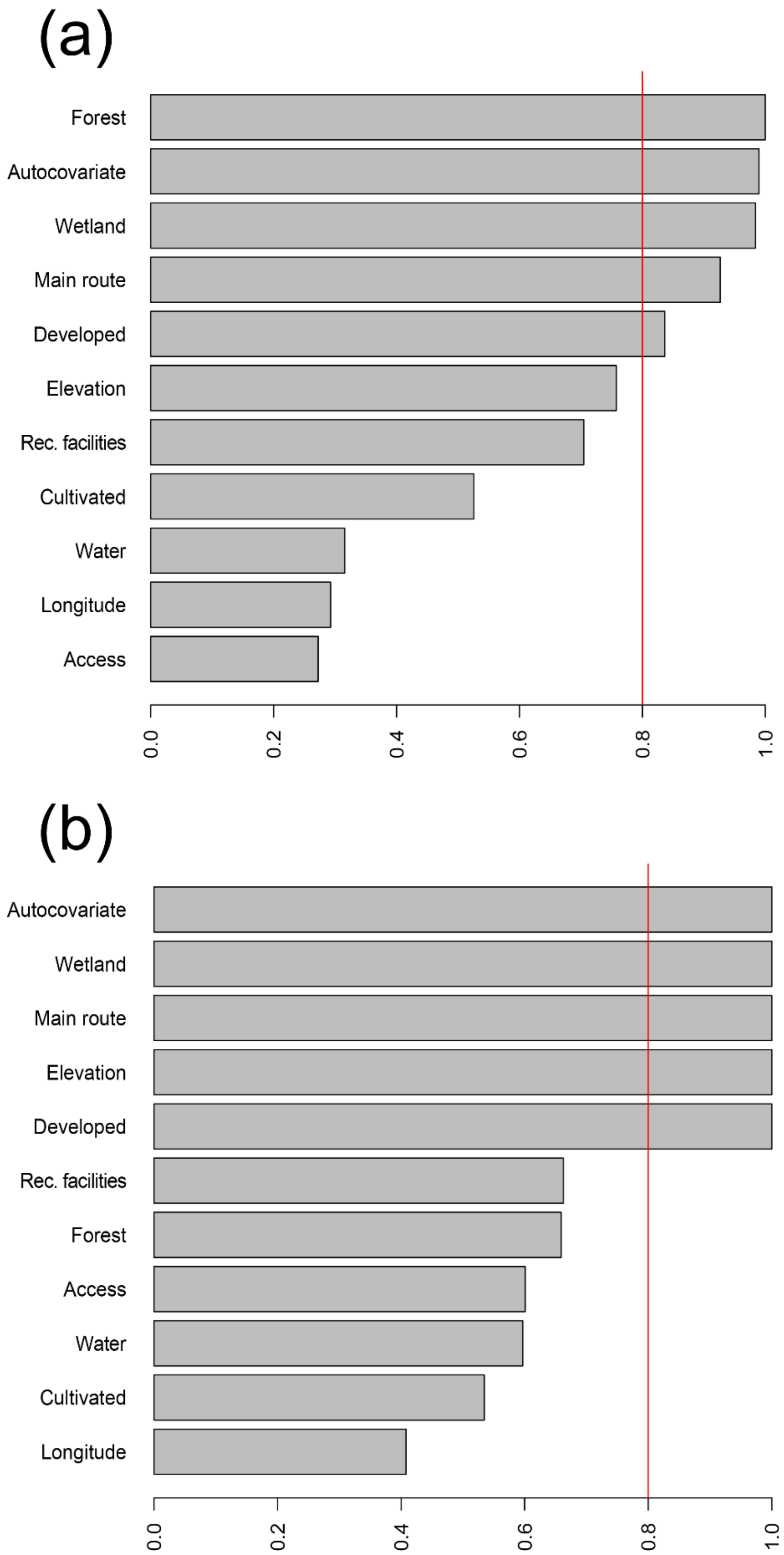
S2.3 All negative binomial generalized linear models of (a) citizen science and (b) professional biodiversity observations within grid cells with a substantial level of support ($\Delta AIC_c < 2$).

Model	AIC_c	ΔAIC_c	k	Evidence weight
(a) Citizen science				
water + access + trails + longitude + developed + cultivated + forest + ac	10764.62	0.00	9	0.152
water + access + trails + longitude + elevation + developed + cultivated + forest + ac	10764.76	0.14	10	0.142
water + access + trails + longitude + developed + cultivated + forest + wetlands + ac	10765.07	0.45	10	0.121
water + access + trails + longitude + elevation + developed + cultivated + forest + wetlands + ac	10765.30	0.68	11	0.109
water + access + trails + longitude + developed + cultivated + wetlands + ac	10765.95	1.33	9	0.078
facilities + water + access + trails + longitude + developed + cultivated + forest + ac	10766.58	1.96	10	0.057
(b) Professional				
facilities + water + access + trails + longitude + cultivated + wetlands + ac	4766.53	0.00	8	0.113
water + access + trails + longitude + cultivated + wetlands + ac	4766.62	0.09	7	0.109
water + access + trails + longitude + developed + cultivated + wetlands + ac	4767.32	0.70	8	0.076
facilities + water + access + trails + longitude + developed + cultivated + wetlands + ac	4767.65	1.12	9	0.065
facilities + water + access + trails + longitude + elevation + cultivated + wetlands + ac	4768.00	1.47	8	0.055
water + access + trails + longitude + elevation + cultivated + wetlands + ac	4768.03	1.50	7	0.054
water + access + trails + longitude + cultivated + forest + wetlands + ac	4768.10	1.57	7	0.052
facilities + water + access + trails + longitude + cultivated + forest + wetlands + ac	4768.24	1.71	8	0.048

S2.4 AIC_c weights of the 2000 highest rated negative binomial generalized linear models for the number of (a) citizen science and (b) professional observations per grid cell, out of a set consisting of all possible combinations of the ten covariates with no interactions. Models below the red line have substantial support ($\Delta AIC_c < 2$).



S3.1 Model-averaged relative variable importance of each covariate for the models of (a) citizen science observations, standardized by trail length, and (b) reported Strava activities among trail segment corridors in the study area.



S3.2 Importance and model-averaged estimates and standard error for each covariate in the (a) citizen science and (b) Strava model among trail segments in the study area.

(a) Citizen science			
	Importance	Estimate	Standard error
Intercept	1.0000	2.5276	6.7570
Forest	0.9999	-1.4894	0.2837
Wetland	0.9840	-1.5201	0.5525
Main route	0.9267	0.3813	0.1839
Developed	0.8364	2.2695	1.5327
Elevation	0.7575	-1.1050e-03	9.4727e-04
Facilities	0.7045	1.8293e-04	1.4776e-04
Cultivated	0.5258	-0.4971	0.6186
Water	0.3156	-7.3325e-05	1.5947e-04
Longitude	0.2928	-1.0377e-06	1.1786e-05
Access	0.2725	-3.434e-09	5.3803e-06
(b) Strava			
	Importance	Estimate	Standard error
Intercept	1.0000	-0.9757	7.4036
Wetland	1.0000	2.3556	0.2464
Main route	1.0000	1.6205	6.9936e-02
Elevation	1.0000	2.4697e-03	3.5730e-04
Developed	0.9999	-2.5075	0.5347
Facilities	0.6626	-6.4751e-05	6.0945e-05
Forest	0.6590	0.1587	0.1540
Access	0.6008	9.0454e-06	9.9301e-06
Water	0.5969	-1.3327e-04	1.5104e-04
Cultivated	0.5346	0.1924	0.2492
Longitude	0.4080	7.8845e-06	1.3055e-05

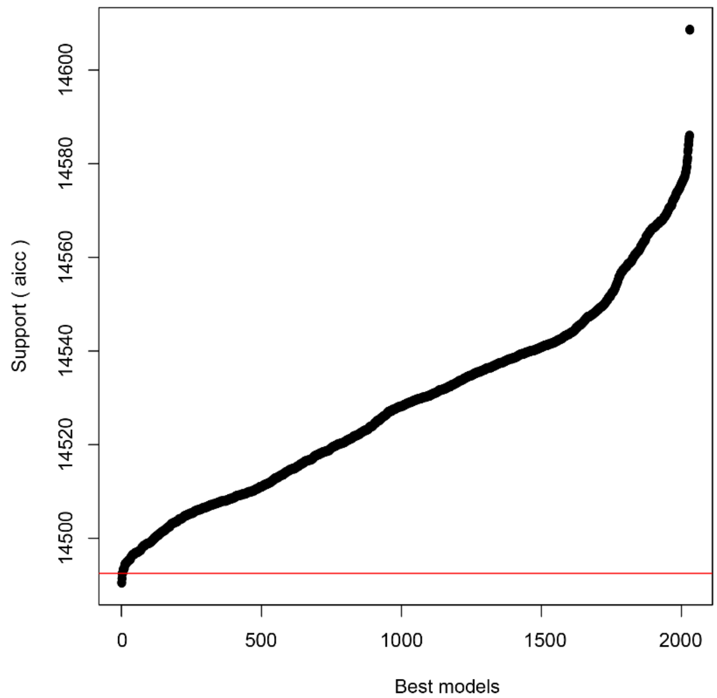
S3.3 All negative binomial generalized linear models of (a) citizen science observations, standardized by trail length, and (b) reported Strava activities within trail segment corridors with a substantial level of support ($\Delta AIC_c < 2$).

Model	AIC_c	ΔAIC_c	k	Evidence weight
(a) Citizen science				
mainroute + facilities + elevation + wetlands + forest + developed + cultivated + ac	14490.53	0.00	8	0.095
mainroute + facilities + elevation + wetlands + forest + developed + ac	14491.42	0.00	7	0.061
mainroute + facilities + elevation + wetlands + forest + developed + cultivated + water + ac	14491.99	0.00	9	0.046
access + mainroute + facilities + elevation + wetlands + forest + developed + cultivated + ac	14492.50	0.89	7	0.036
(b) Strava				
access + mainroute + facilities + elevation + wetlands + forest + developed + cultivated + water + ac	105747.2	0.0	10	0.065
access + mainroute + facilities + elevation + wetlands + forest + developed + cultivated + ac	105747.8	0.6	9	0.048
mainroute + facilities + elevation + wetlands + forest + developed + cultivated + water + ac	105748.2	1.0	9	0.039
access + mainroute + longitude + elevation + wetlands + forest + developed + cultivated + water + ac	105748.2	1.0	10	0.038
access + mainroute + longitude + facilities + elevation + wetlands + forest + developed + cultivated + water + ac	105748.3	1.1	11	0.037
access + mainroute + facilities + elevation + wetlands + forest + developed + water + ac	105748.3	1.1	9	0.036
mainroute + facilities + elevation + wetlands + forest + developed + cultivated + ac	105748.4	1.2	8	0.036
access + mainroute + facilities + elevation + wetlands + forest + developed + ac	105748.8	1.6	8	0.028
access + mainroute + elevation + wetlands + forest + developed + cultivated + water + ac	105748.9	1.7	9	0.027

mainroute + facilities + elevation + wetlands + forest + developed + water + ac	105749.1	1.9	8	0.025
access + mainroute + facilities + elevation + wetlands + developed + ac	105749.1	1.9	7	0.025
access + mainroute + longitude + elevation + wetlands + forest + developed + water + ac	105749.1	1.9	9	0.024

S3.4 AIC$_c$ weights of the 2000 highest rated negative binomial generalized linear models for the number of (a) citizen science observations, standardized by trail length, and (b) recorded Strava activities per trail segment, out of a set consisting of all possible combinations of the ten covariates with no interactions. Models below the red line have substantial support (ΔAIC$_c$ < 2).

(a)



(b)