

Article

A Bayesian Approach to Estimate the Spatial Distribution of Crowdsourced Radiation Measurements around Fukushima

Carolynne Hultquist ^{1,*}, Zita Oravecz ² and Guido Cervone ¹

¹ Department of Geography, Institute for Computational and Data Sciences (ICDS), Earth and Environmental Systems Institute (EESI), The Pennsylvania State University, University Park, PA 16802, USA; cervone@psu.edu

² Department of Human Development & Family Studies, The Pennsylvania State University, University Park, PA 16802, USA; zzo1@psu.edu

* Correspondence: hultquist@psu.edu

Abstract: Citizen-led movements producing spatio-temporal big data are potential sources of useful information during hazards. Yet, the sampling of crowdsourced data is often opportunistic and the statistical variations in the datasets are not typically assessed. There is a scientific need to understand the characteristics and geostatistical variability of big spatial data from these diverse sources if they are to be used for decision making. Crowdsourced radiation measurements can be visualized as raw, often overlapping, points or processed for an aggregated comparison with traditional sources to confirm patterns of elevated radiation levels. However, crowdsourced data from citizen-led projects do not typically use a spatial sampling method so classical geostatistical techniques may not seamlessly be applied. Standard aggregation and interpolation methods were adapted to represent variance, sampling patterns, and the reliability of modeled trends. Finally, a Bayesian approach was used to model the spatial distribution of crowdsourced radiation measurements around Fukushima and quantify uncertainty introduced by the spatial data characteristics. Bayesian kriging of the crowdsourced data captures hotspots and the probabilistic approach could provide timely contextualized information that can improve situational awareness during hazards. This paper calls for the development of methods and metrics to clearly communicate spatial uncertainty by evaluating data characteristics, representing observational gaps and model error, and providing probabilistic outputs for decision making.

Keywords: crowdsourcing; big spatial data; Bayesian statistics; uncertainty; representation



Citation: Hultquist, C.; Oravecz, Z.; Cervone, G. A Bayesian Approach to Estimate the Spatial Distribution of Crowdsourced Radiation Measurements around Fukushima. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 822. <https://doi.org/10.3390/ijgi10120822>

Academic Editors: A-Xing Zhu, Guiming Zhang and Wolfgang Kainz

Received: 1 September 2021

Accepted: 28 November 2021

Published: 6 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Crowdsourced environmental monitoring is of interest to researchers as a source of human relevant social data that may provide localized insights on environmental phenomena and human impacts. It is a form of **Volunteered Geographic Information (VGI)**; in that these data are associated with a spatial location and the contributors intentionally provide entries. **VGI** is particularly noted as having the potential to provide up-to-date information [1,2] and that “citizens as sensors” could measure their local environments [3]. A particular time when crowdsourced environmental data could be useful is to rapidly generate knowledge during hazards when real-time reliable information is needed to improve situational awareness [4]. However, these crowdsourced data collection projects would need to be implemented prior to disasters or quickly after to provide useful information in these contexts. Sometimes, the quickest environmental data during and immediately after disasters are pre-existing or crisis triggered citizen science projects initiated directly from citizens as grassroots initiatives [5].

Pre-existing crowdsourced environmental monitoring can maintain data collection during disasters and provide useful information on environmental changes in a human relevant context. Some projects may evolve from monitoring one aspect in the environment

to then incorporate hazard-specific monitoring. Other projects are initiated during a crisis to collect hazard-related data during an event and may continue collection after the event of interest has passed. Crisis projects that sustain collection after hazardous events are typically those that form a long-term motivating interest, gain resources, develop technical capacities, and maintain outreach to grow a community. Therefore, these projects may be tied to scientific objectives and the generation of knowledge.

Citizen science is a term that describes the activity of the general public contributing collectively to projects or collaborating for a scientific endeavor that often has a social outcome or educational component [6]. The term originated to signify those without formal training in science engaging in authoritative knowledge production [7]. Historically the involvement of citizens was limited mostly to contributions to institutional projects such as data collection. However, environmental monitoring is no longer an activity that only government and industry sectors initiate. The public is increasingly involved in the collection of environmental measurements through the implementation of citizen-led crowdsourcing projects. Communities can develop capabilities to both collect and make use of the data as a source of spatial information. The ubiquitous use of connected technologies spurs the growth of networked and technically enabled communities capable of building devices, maintaining projects, and sharing data that challenges traditional relationships between scientist, public, and data production [2].

Citizen science projects are often categorized by their participation structure through a framework that groups projects by the degree of participation of citizens. Models of participation include contributory, with citizens volunteering data by participating in a project designed by scientists to advance scientific research; collaborative, on a project designed by scientists with citizen input; co-created, as a partnership of active engagement in all part of the project; contractual, in which researchers investigate on the behalf of citizens; and collegial, which are independent projects by non-institutional, often non-credentialed citizens [8–12]. Collegial projects can be specified as citizens involved in knowledge production to answer scientific questions as opposed to institutional researchers [13]. Non-credentialed individuals may be completely in charge of all aspects of the practices of the scientific study [14]. In fact, collegial groups can form a core team that organizes activities and crowdsources data collection from the general public. The collegial participation model is a subset of citizen science that might particularly want institutional recognition of their work [15] and, as a result, are more likely to produce documentation of their project and even publish academically [9].

Projects began to collect data in 2011 around the Fukushima area of Japan after radioactive releases from the damaged [Fukushima Daiichi Nuclear Power Plant \(FDNPP\)](#). The idea was that crowdsourcing could inform populations of environmental risk in areas and times without the availability of government measurements. One example was by researchers from the Tokyo area that started an open-source platform [radiation-watch.org](#) and also developed the Pocket Geiger (POKEGA) to connect to smartphones and distributed 12,000 devices within 6 months of the release [16]. Another project is the Kyoto University Radiation MApping system (KURAMA) which was developed at the end of April 2011 by the Research Reactor Institute to monitor gamma radiation using mobile sensors on public vehicles with 20 to Fukushima and 125 sensors for the government [17]. This process saw rapid development with review by selected experts, partnerships with device distributors for closed hardware and software, and limits on data sharing. Meanwhile Safecast was started by a group in Tokyo to crowdsource ambient environmental radiation levels associated with publicly shared geographic coordinates and a timestamp. It eventually involved the participation of thousands of volunteers contributing open data and sustained itself by adapting its goals beyond the initial collection of crisis data to the scientific goal of providing global data on background radiation levels in order to have a global baseline and to detect elevated levels [18]. The group continues to perform environmental monitoring with a community of volunteers to contribute data including for the collection of other types of environmental data.

The Safecast project participation structure can be categorized as collegial [9] as it involves citizens performing independent data collection and sharing their knowledge. Volunteers serve with different capabilities and the core team provide expertise to meet certain needs of the project. The original core team developed the technology required to collect and share measurements with the intention of providing open data and produced documentation on collection standards [19–21]. Initially, the project crowdfunded to produce Geiger counters, and they have continued to attract financial or physical donations that encourage operations. The Safecast project has been successful over many years at attracting resources and has industry, academic, and government connections. There could also be arguments that this project could belong in other models of participation as the project has connections to people who have been associated with universities. Safecast was co-founded by Joi Ito, an entrepreneur who was the director of the MIT Media Lab at the time, and later joined by Azby Brown, an artist with university affiliations related to Japanese architecture [22]. However, these ties are anecdotal connections and not coming from the top down of academics driving the scientific nature of the project. The project maintained its volunteer participation and grew a community of volunteers through continuous outreach that includes educational benefits. The experiential educational aspects met a need of locals on understanding the risks of radiation levels in their environments.

For the purposes of this research, traditional data refers to typical government or industry data collections and products. These datasets typically have standardized characteristics such as spatially consistent sampling at specified times and places, clearly outlined and replicable collection standards, and statements of limitations at the time of data release. Systematic, spatially sampled datasets are often collected by sensor networks, remote or airborne surveys, and tasked in-person measurements. On the other hand, crowdsourced data is generally opportunistically collected where humans are located. This contributed data does not typically follow consistent sampling patterns, so the data may not completely represent the spatio-temporal patterns of the phenomenon of interest. In addition, crisis sparked projects may not initially report on replicable standards or evaluate limitations for the ongoing collection. Over time many projects add documentation to improve the potential of the data to be evaluated and used in academic and government settings.

Big spatio-temporally varying social data originating from citizen projects can be difficult to assess, as these environmental collections have a number of complex factors and are unstructured for traditional means of evaluation. In addition, evaluation of these data sources can involve big data challenges of volume, as the datasets contain a larger quantity of observations than what many systems can handle; velocity, as spatio-temporally varying real-time production; variety, in the multitude of types of data; veracity, as pertaining to the quality; and vinculation, given the interdependent nature and relationships of the data [23]. Vicinity addresses this nature in the spatial sense as the sampling, density, and resolution of data influences how it can be appropriately used in analysis [24]. Generalizing opportunistic crowdsourced data over time and space is fraught with challenges as the data are not sampled to be generalized.

Many environmental monitoring studies focus on veracity. For some tasks, non-experts have been found to perform just as well as experts and also collecting responses on levels of confidence for observations can produce more robust data [25]. VGI environmental monitoring work on air quality demonstrates a method for quality assurance (QA) protocols to handle the unstructured nature of crowdsourced data [26]. Another study recommends an approach of cross-comparison with a comparable data source [27]. Concurrent work notes that Safecast flags anomalous data for the manual verification of moderators and that study, which compares Safecast to KURAMA, approaches extreme values as anomalies [28]. However, how extreme values should be treated is complex, as the variance of observations in relation to measurement error related uncertainty can be difficult to assess for spatio-temporally varying phenomena. Even so, a dataset may still have low measurement error, but exhibit spatial patterns that are not spatially representative of the phenomenon. As our intention is to visualize patterns of elevated levels near the [FDNPP](#), we focus on the extreme

values collected immediately after the release and treat them as observations to identify elevated trends.

The ‘fitness of use’ to be included in analysis at the initial time of crisis is based on several factors including spatio-temporal completeness, as well as data quality in the sense of the precision and the accuracy of sensors [29,30]. Completeness is an essential aspect of data quality, as it is necessary for the actual patterns of the phenomenon to be observable. Complications can arise in interpretation of crowdsourced data due to the nature of opportunistic social data when using methods that are intended for data with a spatial sampling design. Representation and completeness of the data are essential aspects of data quality ‘fitness of use’ as it is necessary for the data to accurately represent the broad spatial pattern of the phenomenon for opportunistic data to be used for this purpose [31]. These factors contribute to uncertainty in a generalized representation as while even if the individual observation is accurate, a trend may not be representative. The most important aspect of the reliable use of citizen science data to represent spatial patterns is therefore the issue of inherently patchy data [32].

We evaluate crowdsourced data in terms of uncertainty in representation by considering the spatio-temporal completeness at the initial time of crisis [29,30]. Temporal and spatial patterns of collection of social environmental monitoring vary by the motivations and activities of human actors, but mostly occur in populated areas and along roads. The data are not consistently collected in a sampling pattern and collection can be biased by observer effects [33]. How and why the contributed data were collected determines the sorts of questions that are able to be answered through analysis. For example, patchy data over a broad area might still be used to consistently monitor temporal changes in urban areas, but it might not be appropriate to capture the full extent, distribution pattern, and hotspots of the phenomenon of interest. Such as the extreme values near the release which are in infrequently sampled areas. It can be difficult to predict exactly when and where opportunistic data will be collected and under what crisis circumstances there may be gaps given the uncertainties in human behavior. Data may be collected more frequently during a crisis as people are compelled to volunteer data during these disasters while extreme phenomena of interest are occurring. On the other hand, data collection may also decrease at times of crisis as a result of severe impacts, evacuations, or power outages [34]. Nonetheless, these fluctuations themselves can be useful indicators of human activities or barriers when evaluated in this context.

Issues that arise from mapping multiple observations in the same place require the use of geostatistical methods. Spatial point data can be represented through aggregation and interpolation methods used to visualize spatial patterns of the phenomenon. There can be significant measurement variance in the same or nearby locations so exploratory analysis and appropriate selection of measures of aggregation are important. We also use interpolation which is broad category of geostatistical methods that estimates the values of unknown places with known sample values. Kriging is an interpolation technique that weights the sample values based on their distance and degree of variation. Classical kriging uses a linear combination of unknown values at nearby locations based on a single variogram and estimates tend towards the mean value [35]. Therefore, extreme values will always be modeled as decreasing away from the observation, which can result in areas of underestimation. Overestimation can also occur in data gaps in which there is actually a rapid decline. Interpolation with classical kriging is only accurate when the data is spatially sampled at a sufficient density over a continuous surface with a normal distribution, no systematic trends or patterns, spatial autocorrelation at the same rate of change over distance, and stationarity as correlation by distance matters not the location. Clearly, these specifications do not hold true for opportunistic collection so we need to consider other methods to represent spatial patterns in the data.

Bayesian methods offer promising opportunities to provide robust estimates of uncertainty and quantify input error in spatial data [36,37]. Diggle et al. [35] applied Bayesian kriging methods to model the distribution of radiation from radionuclide sampling and

it has proven effective for estimating contamination [38]. Wainwright et al. [39] used a Bayesian model to integrate air dose rate datasets in the FDNPP area with the goal of improving the resolution and completeness. Bayesian kriging is used for capturing uncertainty in the spatial distribution of sparse [40] and non-Gaussian, non-stationary environmental data [41]. In this study, we analyze contrasting characteristics of crowdsourced and government data for the same phenomenon and demonstrate an approach to estimate the spatial distribution with uncertainty metrics. We demonstrate the use of an Empirical Bayesian Kriging approach which uses multiple semivariograms to model spatial variations in the variance of values for domains. This Bayesian approach is used to estimate the spatial distribution of crowdsourced data with uncertainty metrics.

2. Safecast and DOE Data

Following radioactive releases from the damaged FDNPP which spread across the Fukushima prefecture in March 2011, the Safecast project emerged by making use of voluntary “citizens as sensors” [1,3,42]. The Safecast volunteers carry Geiger counters to produce a publicly available collection of radiation levels with a record of coordinates and time collected [19]. The Geiger counters used by the Safecast project are scientifically calibrated, and the ambient environmental radiation is measured in counts per minute (CPM) of ionizing radiation particles [43]. In this way, the radiation level is measured in counts collected over a duration at certain locations and times. The count-based data in CPM records non-negative real value integers, which can be represented as a Poisson distribution.

Safecast has web-based data download options and a map collection which store specific coordinates, times, and radiation levels of millions of measurements from thousands of users of Safecast devices around the world [19]. The focus of Safecast is to provide open data, and the general movement of collecting radiation data was a reaction to government data not being made immediately available after the radioactive releases [22]. The founders of Safecast are committed to providing open access data to the public domain using a Creative Commons “No Rights Reserved” (CC0) designation [44]. The Safecast initiative has expanded its goals to provide global data on background radiation levels to detect elevated radiation levels as well as other environmental themes of interest. Details on the Safecast project can be found in the annual reports of the Safecast team [19–21].

Globally, the Safecast dataset contained 40 million point-based measurements by the end of 2015 and 100 million by 2018. However, the large quantities of data that are available today were not initially available to be used in the immediate aftermath of the nuclear accident. Figure 1 shows the gradual growth of the dataset globally. For our study, we will later subset the data to evaluate the suitability of using the data collected in the initial time of crisis within Fukushima. However, this figure shows how the project grew as it expanded beyond Tokyo and the Fukushima prefecture to gain a global following of volunteers interested in collecting background radiation levels. Even if these data were not initially available with broad and thorough coverage, an established project capable of replicating a government survey as demonstrated in Cervone and Hultquist [45] could be used as a ground-based observational input to models in the event of a future crisis.

We provide context for the spatial pattern of phenomenon from the crowdsourced data through comparison to a traditional, government collected dataset. The U.S. Department of Energy (DOE) / U.S. National Nuclear Security Administration (NNSA) data was collected by airborne surveys in consistent swath sampling patterns over the Fukushima prefecture [46]. The collection is an intentional designed study, and the “aircraft ground speed, altitude, and line spacing are chosen to optimize the detector system’s sensitivity and spatial resolution” [47]. The study used quality assurance techniques, and the documentation clearly states limitations that complicated the collection. These limitations including terrain considerations, plume modeling for scope, and the inability to collect nearby background levels. The DOE collection involved the use of calibration methods to deal with the complexity of the analysis and the technique has associated uncertainty as it is an airborne collection that is extrapolated to 1 meter above ground level by using

on-ground measurements [47]. This dataset has also been compared with other radiation surveys in the Fukushima area [48].

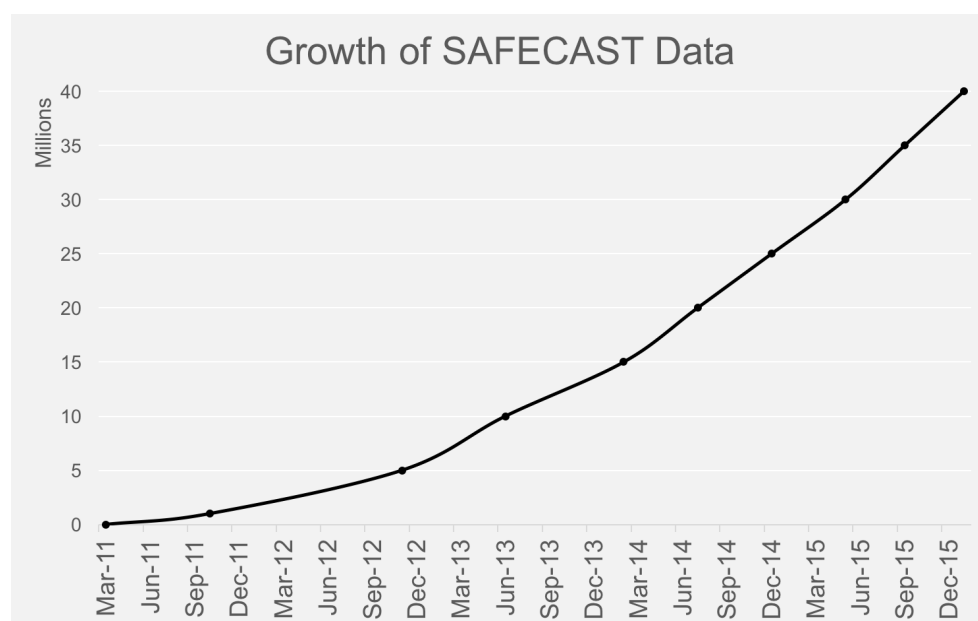


Figure 1. Growth of the quantity of Safecast data globally from 2011 to 2015 by millions of measurements.

The DOE dataset is used in this paper for comparison to Safecast because it has a broad spatial coverage of the Fukushima prefecture. The DOE data were previously aggregated in comparison to the Safecast data at the same location to show a high correlation of measurements over three months in 2011 by Coletti et al. [49] and over years with the development of a decay correction method by Hultquist and Cervone [50]. Finally, Cervone and Hultquist [45] demonstrated how Safecast data could be used to reconstruct the high concentrations of radiation levels of the plume from DOE and predict future radiation levels. This paper takes the evaluation further to analyze contrasting characteristics of the collection of the two datasets in 2011 and demonstrate an approach to estimate a spatial distribution with uncertainty metrics.

The Safecast data is voluntary, opportunistic contributions, and the DOE survey is used for comparison to the Safecast data as it is a broad collection with regular spatial sampling [47]. Table 1 provides a summary of both of these sources of data that were collected in the Fukushima prefecture during the specified time periods.

Table 1. Table summarizing data sources with the number of measurements within the spatial extent of the Fukushima prefecture.

Dataset	Year (Quantity)	Collected in Units	Type
DOE	2011 (107,147)	Microsieverts ($\mu\text{Sv/h}$)	Government
Safecast	2011–2015 (6,249,165)	Counts per minute (CPM)	Citizen Science

We visualize in a Geographic Information System (GIS) the Safecast and DOE datasets for the Fukushima prefecture in Japan. Figure 2 shows similar trends in the raw point values of the datasets in a common unit of $\mu\text{Sv/h}$ with all the data until November 2011 in the Fukushima prefecture. The Fukushima prefecture is on the east of the island of Honshu with the east coast bordering the Pacific Ocean and mountainous terrain in the west. FDNPP, the source of the release, is located midway on the coast in Figure 2. The DOE data clearly identifies the extent of the radioactive plume of elevated levels and the

sampling design of the airborne survey ensures that the data is spatially representative of the phenomenon.

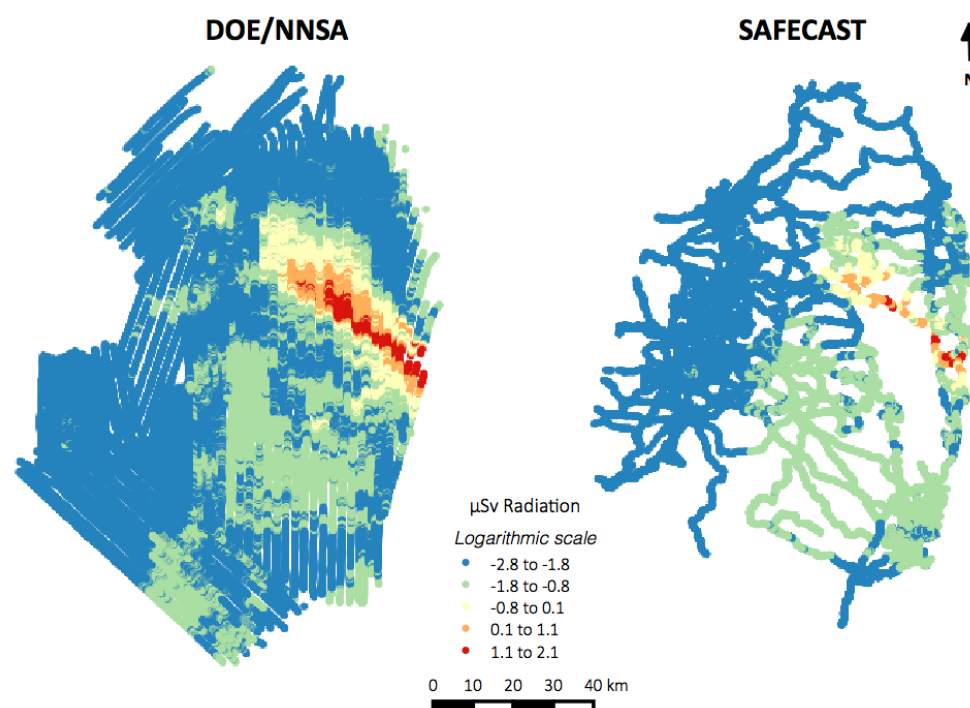


Figure 2. Safecast and DOE raw data points overplotted from 2011 [24].

In this paper, we evaluate techniques for aggregation and interpolation of crowd-sourced data, but first, we show the raw data here to demonstrate why these visualization techniques are necessary. Figure 2 is an example of overplotting as all of the measurements are displayed as points and unless the collection of the data is spaced far enough apart, the data will overlap in the visualization. This makes it impossible to see each individual measurement, especially for the Safecast data in which many observations are taken at the same locations. Using a simple plot of the points is ill-advised in many circumstances as the last value plotted in that location will end up being seen while other values are masked over. This is particularly an issue for data that are not sampled and that vary temporally. If the data are sorted by date then the last value collected in an area will obscure previous data and the data are collected at different times so the visualization will not be temporally standardized. These point-based mapping issues are why methods such as interpolation or aggregation are necessary. In this case, it is still possible to see that there are elevated levels of radiation simply by plotting the points. This plume area of elevated radiation levels is the focus of the paper to show how the representation of the spatial pattern varies as a result of the method chosen.

Representation and statistical outcomes will vary based on the temporal subset of Safecast data selected for analysis. This is because, unlike traditional data, the data are not evenly sampled spatially at regular temporal intervals to provide consistency. While the Safecast dataset is quite large as a whole, the selection of subsets of a big spatio-temporal varying dataset provide unique statistical outcomes. This is to say that each subset by time and space provides different results due to the distribution of the data collection. While this is less prominent if all the data from an area are within a normal background level range, capturing enough of the extreme values of elevated radiation levels in a time period is particularly important to ensure a representative result. We are actually most interested in the extreme values themselves as an indicator of a hazard, but these elevated areas tend to have fewer and infrequent observations, particularly early on after the event.

A visual depiction of the weekly quantities of data available over time from both of the sources is shown in Figure 3, which was produced in R. The DOE data was collected

from mid-March through May and the Safecast data was collected in larger quantities later that same year. While the DOE data was collected shortly after the nuclear accident, it should be kept in mind that government data might not initially be released to the public. We use this temporal subset of the data from 2011 for the following analysis.

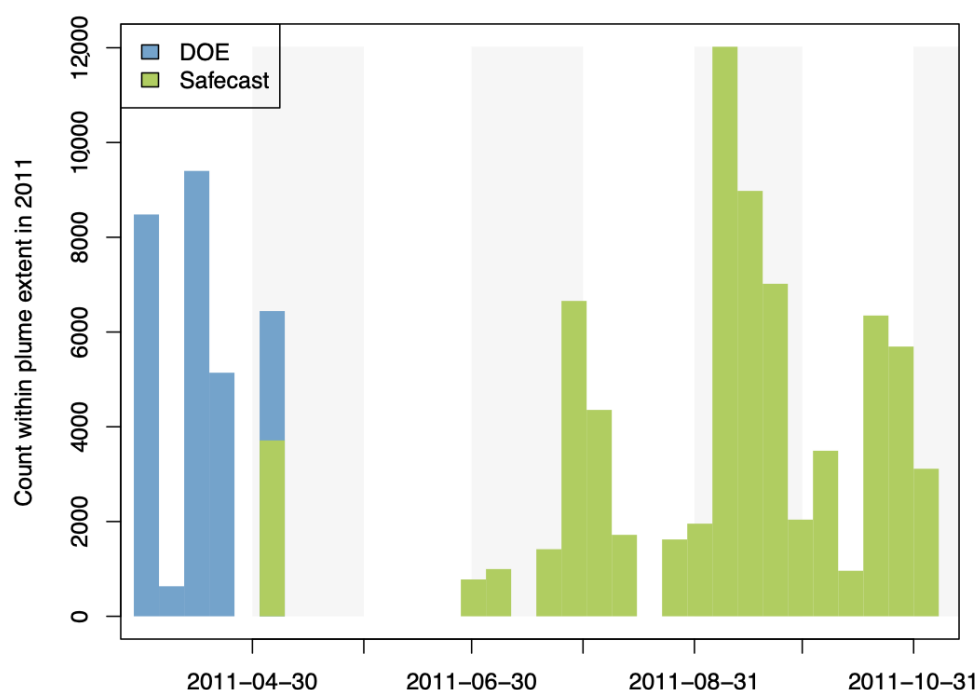


Figure 3. Temporal plot of DOE and Safecast data for 2011.

3. Methods for Crowdsourced Data

The research methodology of this paper is summarized here for the estimation of the spatial distribution of crowdsourced radiation measurements. This approach deals with analytic challenges that arise with the data analysis of unevenly-collected or opportunistic crowdsourced data. Many factors influence how patterns of spatial data are represented, including how the data is collected, and it becomes even more important to capture uncertainty. We demonstrate a methodology to use geostatistical techniques for the evaluation and representation of crowdsourced data. This section is divided into three overarching methodological categories to evaluate the spatial distribution of crowdsourced data in relation to sampling (1) Processing for Comparison, (2) Aggregation and Interpolation Techniques, and (3) Bayesian Methods.

3.1. Processing for Comparison

We recognise that a primary way of ensuring the reasonableness of datasets is to compare the data to other sources. The spatially sampled government survey from the DOE is used to check the patterns observed in the Safecast data when visualized under various constraints and representations. We select an area of interest and a time frame to make a subset of the datasets and evaluate the spatial sampling of the data. We use the entire dataset for this time period of 2011 in the Fukushima area.

A necessary aspect for comparison is converting the data to a common structure with an understanding of background information to be able to make the data comparable through standardization, aggregation, and statistical evaluation [50]. Since radiation has temporal properties this can be done by decay correcting values using a deposition survey ratio [51] given the decay rates of ^{134}Cs and ^{137}Cs radiation for areas above natural background radiation levels. Underlying questions about the spatial data dimension are addressed as the methodology of Hultquist and Cervone [50] provides a way to standardize the temporal continuum and compare nearby spatial data. Multiple datasets can be used

together in order to recognize the limitations that each dataset and model has in identifying hazards [52]. These methods are not a validation of each measurement, but they compare aggregations of measurements, and are intended to validate the derivation of the spatial pattern from unevenly sampled contributed data.

Initial processing is performed to standardize the datasets for comparison. Selection of a 0.5 kilometer grid of the contributed data shows a spatial pattern consistent with the government data, even with variations in space and across time periods. The unit of measure is converted from CPM to $\mu\text{Sv/h}$ using

$$1 \mu\text{Sv/h} = \frac{1}{334} \text{ cpm} \quad (1)$$

Refs. [49,50] and rasterized to the same extent. A levelplot using the R lattice package [53] is used for visualization with a formula of

$$z \sim x * y \quad (2)$$

in which x and y are longitude and latitude in a grid and z is the log of $\mu\text{Sv/h}$. Maps produced in R in this paper use RColorBrewer color schemes [54].

3.2. Aggregation and Interpolation Techniques

Statistical representations and models are used to aid in the interpretation of spatial patterns of crowdsourced data in light of their data characteristics. Crowdsourced collections are typically unstructured; not following a specific sampling method, and the statistical variations in the datasets are not typically assessed. Assessment of the representation is necessary especially when using methods intended for data with spatial sampling design as complications can arise in the interpretation of crowdsourced data results due to the nature of the opportunistic social data collection.

We explore how aggregation and interpolation techniques can be used to better visualize crowdsourced data. The methods that are commonly used for traditional data are adjusted for an understanding of the characteristics of crowdsourced data and the phenomena of interest. We use statistical measures of standard deviation and local deviation from global mean. Standard deviation is calculated from the R stats package as

$$s = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}} \quad (3)$$

while using the rasterize function in the R raster package [55] for each 0.5 km grid cell. Raster deviation is a function from the R SpatialEco package [56] which evaluates the local deviation from global statistic measure as

$$[\text{stat}(x) - \hat{y}] \quad (4)$$

in which $\text{stat}(x)$ is the global value of the specified statistic and $\hat{y}$ is the focal statistic. We use local deviation from the global mean for data aggregated within each 0.5 km grid cell. We display the results of these measures with the crowdsourced data range for both the same as the traditional data for comparison and its natural range.

Maps are abstractions, so the representation of values and features on maps involves generalization. We demonstrate the influence of how the data are visualized in comparison. In order to compare between datasets, the same numerical classification scheme is used for visualization; otherwise, the visual representation is not an equal comparison. However, as the crowdsourced data are not systematically sampled, it can be difficult to determine which scheme is appropriate to use for comparison if the best for one source is actually inappropriate for the other. Likewise, a comparison of maps using different spatial sampling schemes and data distributions can be inconclusive or misleading if a

common classification is not used. The optimal choice of numerical classification methods to visualize patterns can differ depending on the sampling scheme and the intention, such as showing all the data equally or emphasizing extreme values. The decision process of which interval to use depends on the motivation of the mapper, the type of phenomena, and the spatial distribution of the data.

Radiation levels indicated in cells with few observations and high variance are not typically differentiated from areas with many observations and low variance. Therefore, we use a mapping approach that visualizes that number of observations as well as the aggregated statistical result. We use the R hexbin package [57] to show the number of observations in each hexbin while the color scheme is used to show the statistical mean of the radiation value. Multivariate representations enable the visualization of the relationship between factors to consider characteristics of the collection.

Interpolation outputs typically mask the raw spatial pattern and do not reveal the inherent uncertainty introduced by the distribution of the data. An interpolation does not typically show where there is and is not data supporting the resulting surface that is visualized. Therefore, we demonstrate a method to visualize the weight of the raw data as an overlay on the interpolation to show how much the data is and is not supporting areas of the interpolation. We demonstrate an output with an overlay of the density of observations. The surface was produced in R using Multilevel B-spline Approximation (MBA) [58]. This method applies transparency to show the quantity of measurements based on the weight of the data using imagePlot from the R package fields [59]. Finally, we show the statistical distribution of data by itself and as context for the distribution of the interpolation. Histograms are shown alongside using the hist function from the R raster package [55]. The interpretation of spatial data can be improved when viewed alongside statistical measures and plots for context.

Here, we provided examples of representations from contributed data to capture uneven spatial sampling and variance as these are not generally captured by standard mapping practices. When there is an intention to produce these types of output, standard geostatistical methods and mapping techniques can be used to represent uncertainty.

3.3. Bayesian Methods

We demonstrate a Bayesian approach to estimate the spatial distribution of crowdsourced data with uncertainty metrics. A Bayesian approach is a rational statistical perspective that considers the weight of evidence and has mechanisms to evaluate uncertainty [60]. One primary benefit of the Bayesian approach is the ability to handle uncertainty when few observations are available [61]. Bayesian methods have proven useful for representing uncertainty in statistical estimates of spatial data [36]. We apply empirical Bayesian kriging to quantify uncertainty in the spatial representation of crowdsourced environmental hazard data. The full formulation of empirical Bayesian kriging method can be found in Krivoruchko and Gribov [62].

Bayesian kriging is a flexible tool to mathematically formulate spatial dependencies that increase the overall prediction accuracy of the interpolation [63]. We use Bayesian kriging to investigate how contributed radiation measurements vary over space, and if the opportunistic spatial distribution is sufficient to regularly and reliably be used for spatial interpolation. We map the Bayesian kriging model with an overlay of the crowdsourced data to provide context on the locations of the data that are used in the output. Because this method does not assume tendency towards the mean [38] it can predict extreme values over space. Semivariograms are used for interpolation within spatial domains to model consistent changes within that area in the spatial distribution. It automates parameterization over the spatial domains of the model by subsetting and simulating the parameters. The use of multiple semivariograms means that the spatial variance does not have to be modeled as a constant over the entire domain as in classical kriging.

We demonstrate the use of an empirical Bayesian kriging approach which uses multiple semivariograms to model spatial variations in the variance of values for domains.

The variance of the model is inferred while minimizing error within localized spatial domains which improves the accuracy with a spatially dependent model and semivariograms [64]. Allowing for different forms of spatial variance of values in the model is particularly important when there are extreme values in an area that have a sudden shift while the majority of the domain has many consistently low values. With this approach we spatially represent the uncertainty of the model by visualizing prediction error. We produce an error surface to help to distinguish between the observed characteristics of the phenomenon and gaps in the data at different locations.

Finally, we create probability maps to compare outputs at multiple thresholds. While aggregations of the maximum value or other extreme value measures may produce maps with inconsistent patterns. This technique can be used to compare thresholds and isolate areas where extreme values are more likely to occur. We use this method to calculate probabilities of radiation values at specific thresholds for the study area. The method can be used to identify areas of consistent extremes even from sparse and noisy data.

3.4. Methodological Summary

This study evaluates the spatial distribution of crowdsourced measurements. We use this approach due to the nature of opportunistic crowdsourced data that is unevenly-collected in order to better understand how the characteristics of the collection influences the representation. First, we process the crowdsourced data and select a spatial and temporal subset to compare to another data source with consistent sampling. Then we demonstrate how aggregations and interpolations are influenced by sampling. Finally, we use Bayesian methods to represent uncertainty and probabilistic representations of the data.

4. Results

Geostatistical methods include traditional techniques that assume spatial sampling, consistency, and completeness. These standard methods that are used to aggregate and interpolate spatial point data need to be evaluated in light of the different spatio-temporal characteristics of crowdsourced data. The use of novel data streams with unique spatio-temporal properties highlights the need to consider uncertainty in the representation when statistical techniques are applied to these datasets.

4.1. Processing for Comparison

To compare the data aggregated to grids, we create a common format, scale of comparison, aggregation, and ensure the grid method itself does not cause issue with the pattern. We compare crowdsourced with traditional data by subsetting and processing the data to the same time and place. Critical dimensions of time and space are determining factors in the distribution of radiation from a release.

Table 2 summarizes observations in the elevated plume region of the Fukushima prefecture from March until November 2011 which is the data shown in the following visualizations.

Table 2. Table summarizing data sources with the number of measurements within the selected spatial extent in 2011.

Dataset	Year (Quantity)	Collected in Units	Type
DOE	2011 (30,098 in extent)	Microsieverts ($\mu\text{Sv/h}$)	Government
Safecast	2011 (72,878 in extent)	Counts per minute (cpm)	Citizen Science

We recognize that sub-setting based on random, aspatial sampling would significantly decrease the potential spatial resolution of a surface representation. Locations with sparse contributed collection are unlikely to be randomly selected and, therefore, would not be represented in a subset. This is a result of the characteristics of the crowdsourced dataset

as some locations have many more observations than other places. The crowdsourced data collection patterns are constrained in this case by limited access areas because much smaller quantities of data are available for the region and at the time of interest. As more users participated, the spatial coverage of the contributed dataset expanded over the years within this area of interest.

In this example, the spatial points are averaged within a raster of a standard cell size. Figure 4, which was produced using the R lattice package [53], shows Safecast data from 2011 aggregated into 0.5 km spatial units for just the area of elevated values in $\log(\mu\text{Sv/h})$ as it aligns with the same in DOE.

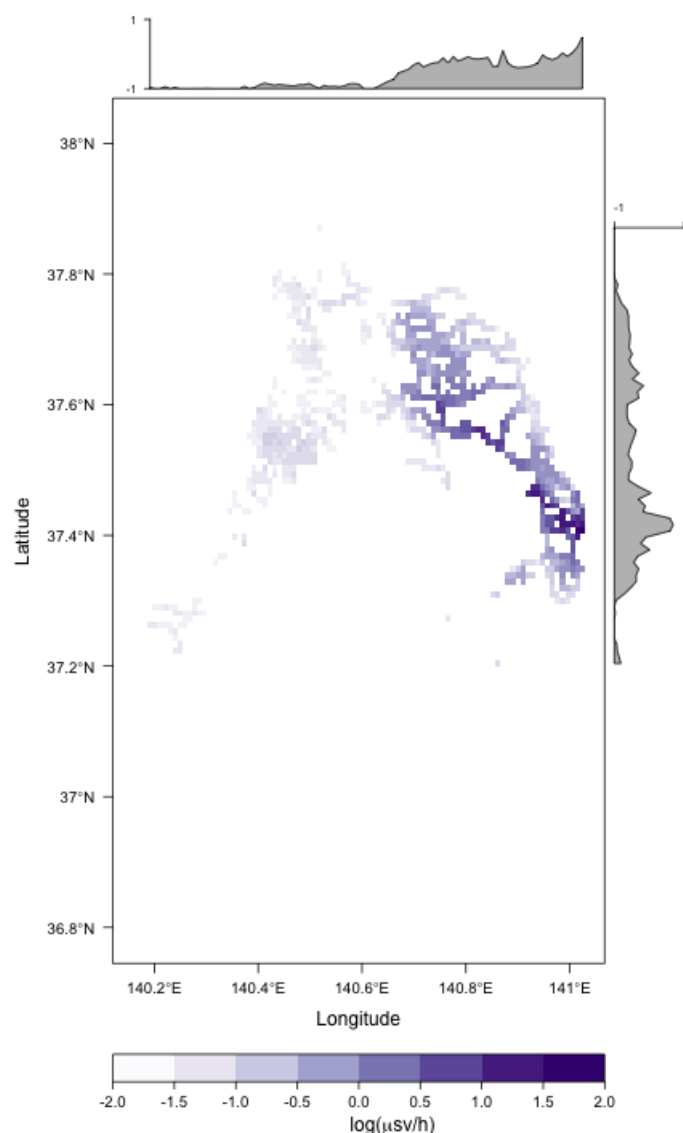


Figure 4. Average elevated levels from 2011 Safecast data aggregated at 0.5 km grid in $\log(\mu\text{Sv/h})$. The side panels are the average values projected onto the axis.

4.2. Aggregation and Interpolation Techniques

Evaluation of statistical aggregations can help to determine if the data are reasonably distributed to represent the phenomenon. We evaluate the appropriate type of numerical statistic for aggregated grid cells, such as the mean, median, minimum, or maximum value, which needs to be determined on a case by case basis. The selection of the measure of central tendency will not have a significant impact on the surface if there are few observations with little variance in a cell. Statistical averages may be more precise when there are many

measurements in a place, but then the dataset will likely have a higher variance, too. As the data are aggregated in this form of visualization, more repetitive data at the same location will not provide an informational benefit. Sampled datasets can have a consistent number of values in each area while opportunistic datasets might have many observations in one cell and few or no observations in other cells. It may then prove useful to evaluate within each cell how much the data vary from the local mean and how much the data vary from the global mean. In fact, the extreme values are themselves of interest as they could indicate an environmental hazard, noisiness in the data, or even measurement error that needs to be understood in context.

The following maps were produced in R to provide a comparison of the data using statistical measures and adjusting for matching ranges. The first column of Figure 5 shows the standard deviation of the data within each 0.5 km grid cell. The data are aggregated using the rasterize function in the R raster package [55]. There can be a lot of variation in observations within each cell, with the most variation occurring within the plume near the source of the release where there are more extreme values. The second column of Figure 5 shows the local deviation from the global mean produced using R spatialEco [56] of the subset of dataset shown. Note that the top row is DOE data while the last two are Safecast with the middle row constrained to the same range as DOE and the bottom row at the natural range of Safecast. The last two rows demonstrate how comparisons require data to have the same bounds in their representation to be meaningful. Without the middle row showing that there is variation in the Safecast data using the same range, the final row could make it appear that there is less variation due to the cartographic decision.

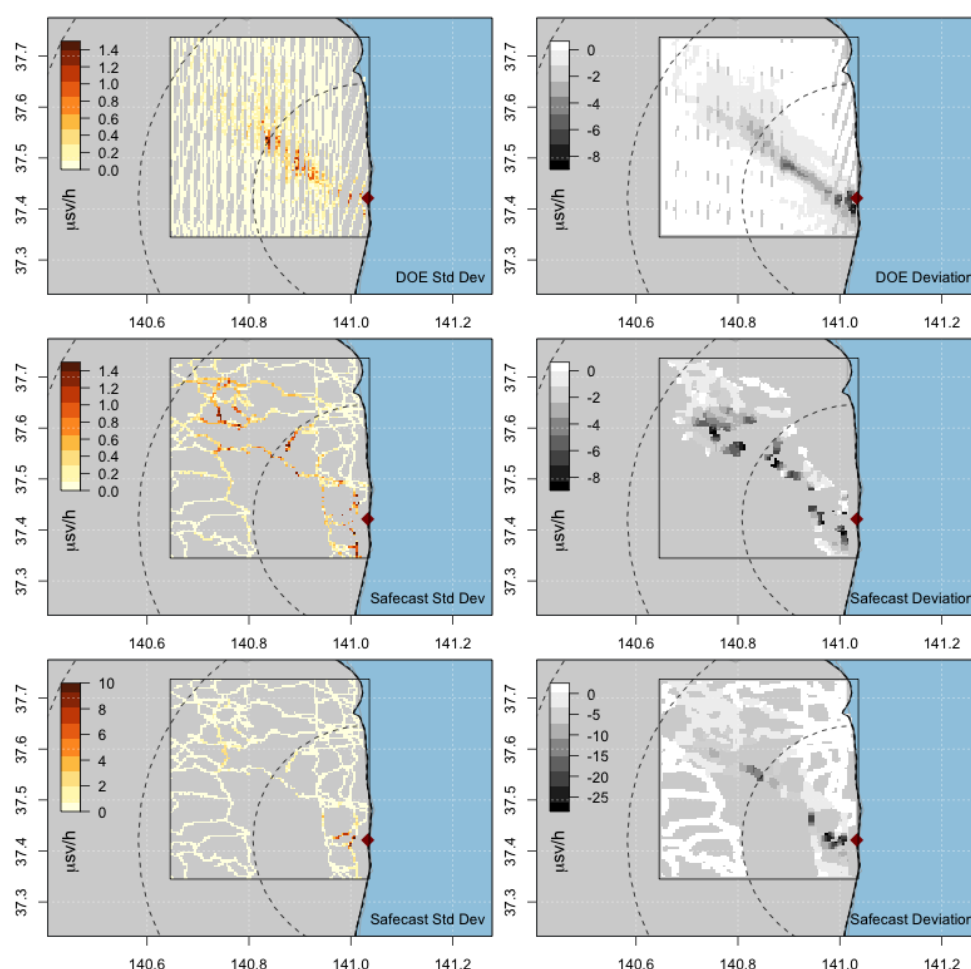


Figure 5. Statistical measures for 2011 data in $\mu\text{Sv/h}$: in column 1, Standard Deviation with 0.5 km cell and column 2, local deviation of cell from the global mean. (**Row 1**): DOE, (**Row 2**): Safecast data for the same range of values as DOE, and (**Row 3**): Safecast data for its own range.

Figure 6, which was produced using the R hexbin package [57], is zoomed in on the area with the highest observed radiation values around the damaged FDNPP with data from over the six months after the 2011 event. It shows a count of the number of observations in each hexbin. The color scheme is the statistical mean of the radiation value with most of the hexbin areas having an average below 4 $\mu\text{Sv/h}$, and with the highest few hexbin areas having an average of up to 24 $\mu\text{Sv/h}$. There are very few areas that have extreme values; in these areas there are relatively fewer observations, and many hexbins with extreme values have neighboring hexbins without any observations. The sparse sampling of the Safecast data in elevated areas, especially immediately after the accident, is a limitation of the dataset in regard to the observed spatial pattern as it introduces uncertainty related to the spatial distributions. We cannot be certain just from this subset of the dataset that the peak levels or areas are captured.

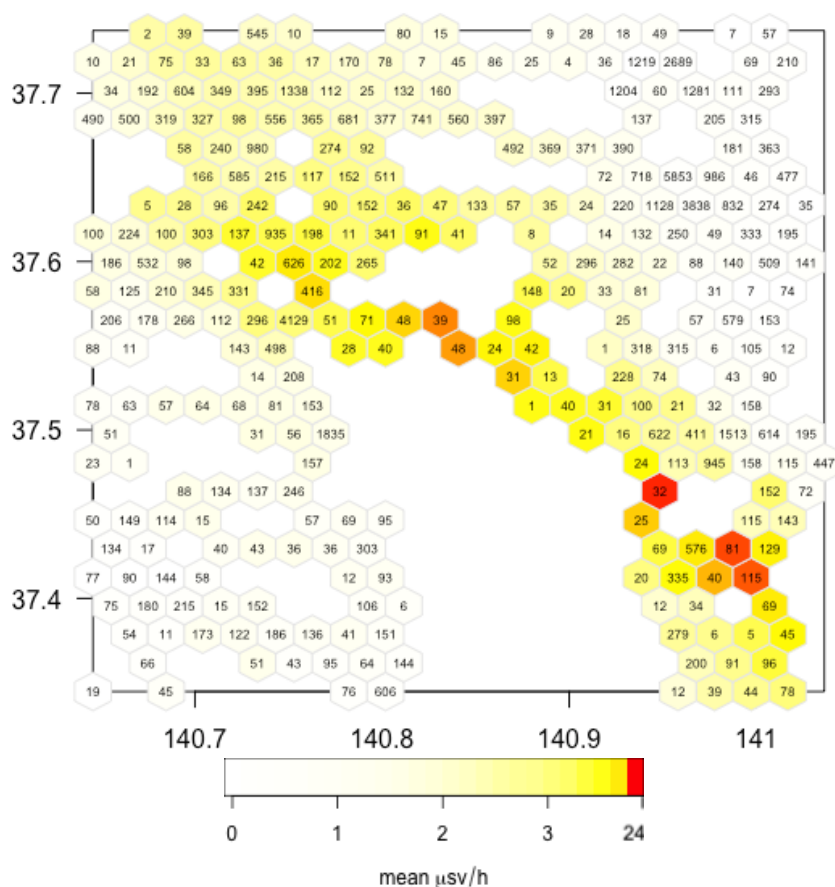


Figure 6. Counts binned by location with text of number of measurements and color showing mean $\mu\text{Sv/h}$.

Some places in the dataset have very few observations while other nearby locations have thousands of observations. Therefore, it is necessary to take care with the analysis of contributed data when subsetting the data and to evaluate an appropriate method of visual representation given the spatial distribution. Areas with the highest radiation values are unlikely to be as represented, for the most part, these locations with high values have relatively few total observations. Therefore, taking a subset of the dataset without considering spatial sampling constraints might result in the phenomenon of interest not being observable or accurately represented in the spatial pattern. When a small, consecutive sample of the dataset is taken, then the spatial distribution is also limited as consecutive observations are typically collected by the same device in nearby locations.

An interpolation of radiation values is shown in Figures 7 and 8, which were produced in R using Multilevel B-spline Approximation (MBA) [58] and imagePlot from fields [59],

with an overlay of the density of observations. The black traces have transparency applied to indicate the quantity of measurements so the brightness is not a representation of the radiation values themselves. A pattern of elevated values of radiation are observed from the [FDNPP](#) at the coastline on the east edge to the northwest inland. Spectral color schemes are used here as they can be appropriate at times to emphasize both low and high extremes of quantitative data [65]. These interpolations were visualized using different scales as the Safecast data contains some extreme values that we want to capture; the intention here is not to compare observational values, but to contrast the spatial patterns of data collection in light of a straightforward interpolation of each dataset. This is not a standardized comparison of the data, but an exercise to highlight the differences in the sampling of the raw data collection when displaying interpolations.

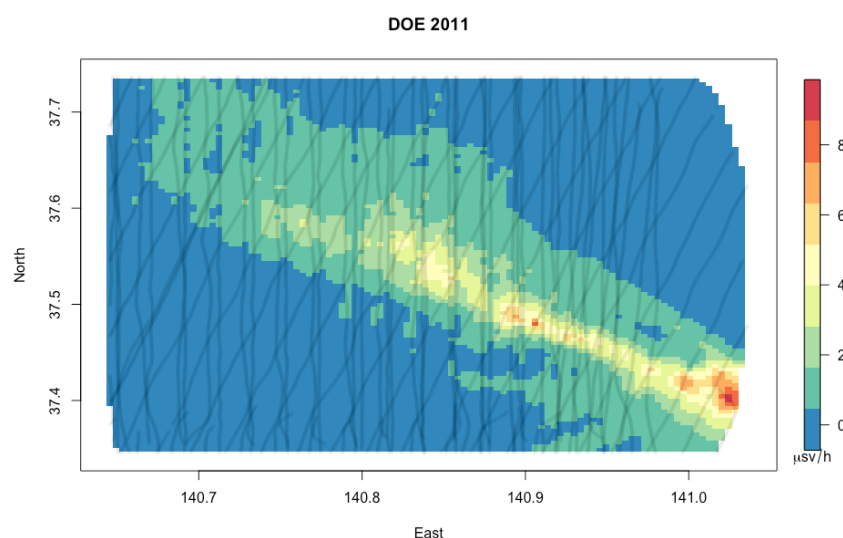


Figure 7. Interpolation and locations of the DOE data.

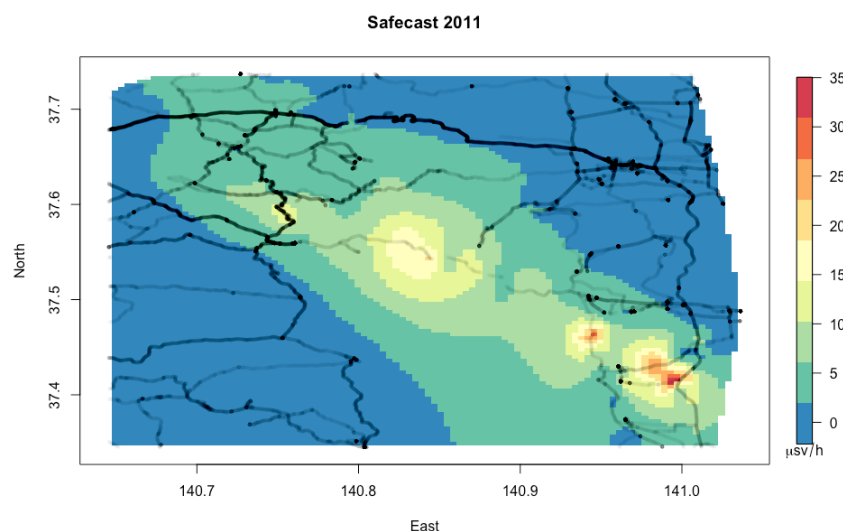


Figure 8. Interpolation and locations of the Safecast data.

The surface layer of Figure 8 by itself does not indicate what parts of the surface are interpolated to fill gaps in the data or how much uncertainty is present. The data are not spatially sampled so it is not clear if the places showing the highest values are actually where the peak of the phenomenon is or if the visualization is just the highest value recorded in the area. For example, it is possible that the collection did not record data in the location of the actual highest values. These questions can often be addressed by comparing a data source to other sources that have different limitations.

The DOE data were consistently collected in a swath pattern in order to identify the extent of the release. Whereas, in some areas, the Safecast data does not have a sufficient spatial distribution to support this interpolated surface, especially when temporal subsets are selected. Figure 8 shows how there could be an issue with interpolating over space in which there are no observations. For this time period, there are large areas in the interpolated surface without Safecast measurements. Only consistent spatial sampling would limit the horizontal spread of the interpolated distribution. A classical kriging model can erroneously widen the plume beyond what is observed in the DOE data if the spatial decay parameters do not limit the interpolation in that domain. When the extent of the release is unknown, caution needs to be taken when interpolating from crowdsourced data of a phenomenon as the spacing is based on the general domain instead of the spatial limit of observed extreme values.

Here in Figure 9 we see how the data from DOE and Safecast are used to interpolate a surface along with histograms in $\log(\mu\text{Sv/h})$. Even if the data are produced for the same phenomenon, crowdsourced data can have uniquely different characteristics when compared to traditional data collections, such as government surveys. While the spatial trends of the Safecast measurements appear visually similar to DOE measurements, the statistical representations can be used to analyze how reliable it is to infer information from sparse interpolated surfaces. In this case, it is clear that the crowdsourced interpolation surface involves the estimation of many more spatial locations.

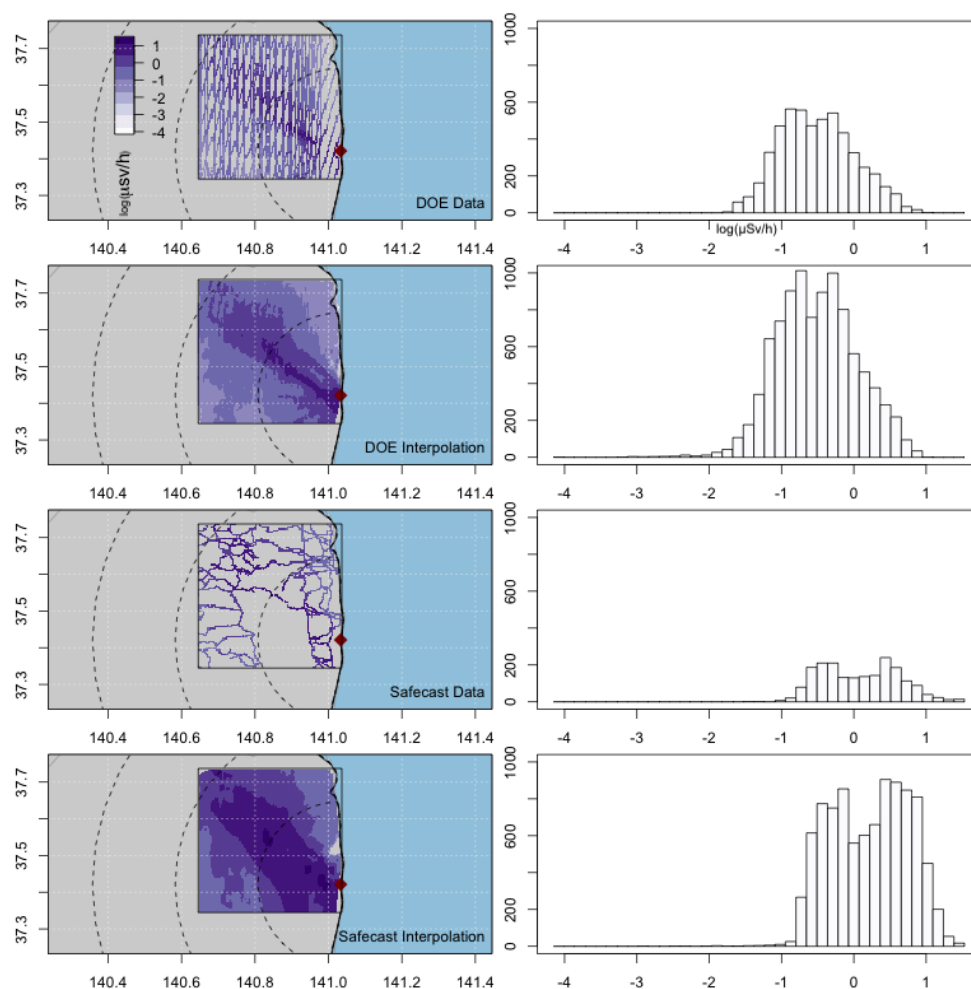


Figure 9. Maps and graphs of DOE and Safecast data for 2011 in $\log(\mu\text{Sv/h})$.

4.3. Bayesian Methods

A Bayesian approach is used to model the spatial distribution and assess uncertainty in spatial predictions. Bayesian interpolation methods are used to evaluate the uncertainty of the surface for specific information needs when modeling predictions over space in data gaps. Figure 10, which was produced in GIS, shows the location of Safecast data observations in relation to the derived surface using Bayesian interpolation methods. We can see that most of the Safecast measurements are collected in urban areas away from the release and along roads. This method of overlaying also shows us that unevenly sampled data are spatially interpolated to show elevated levels in the radiation unit of CPM. It is clear that there is associated uncertainty in the spatial patterns as excluding a few traces or subsets of the dataset in specific areas can result in different interpolated spatial patterns. Likewise, the spatial pattern could also change by adding more data in spatial gaps. This means that the visual representation of an interpolated surface, derived from a subset, or potentially even from all the data that happened to be collected, may not show an accurate depiction of the phenomenon even if the measurements themselves are accurate.

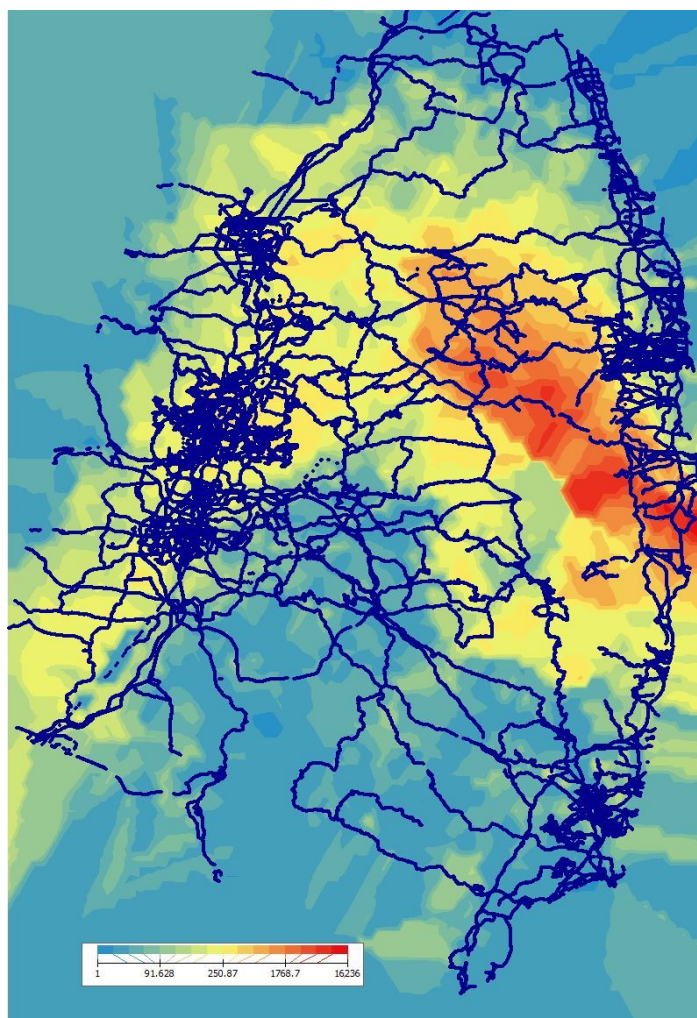


Figure 10. Point locations of 2011 Safecast data in Fukushima overlaid on a Bayesian kriging model in CPM.

Statistical representations and models are demonstrated, including an Empirical Bayesian Kriging approach to represent uncertainty and probabilistic output, that can aid in interpreting spatial patterns of the crowdsourced data. The use of this approach allows for interpolation that fills spatio-temporal gaps in the data based on how the values vary in each area by splitting the areas to be represented by semivariograms consistent with the

degree of change in observational value given distance. That is to say that, the Empirical Bayesian Kriging uses multiple semivariograms to model the degree of the shift over space within separate spatial domains to decrease overall error. A Bayesian kriging approach can be used to model the spatial distribution of datasets that vary in some places differently than in other spatial areas.

A Bayesian approach to kriging allows for the calculation and visualization of prediction errors. Figure 11, which was produced in GIS, shows contour areas with the predicted value and the other panel shows the standard error associated with each contour area prediction. For areas in which there are no observed radiation measurements, radiation levels are predicted using a domain dependent spatial interpolation function and the uncertainty is represented in the visualization of error. The spatial uncertainty of data gaps and locations with high variance are captured in the resulting surface as the highest levels of radioactivity are in the areas with the least quantity of measurements. A few areas are left blank in which no prediction is given.

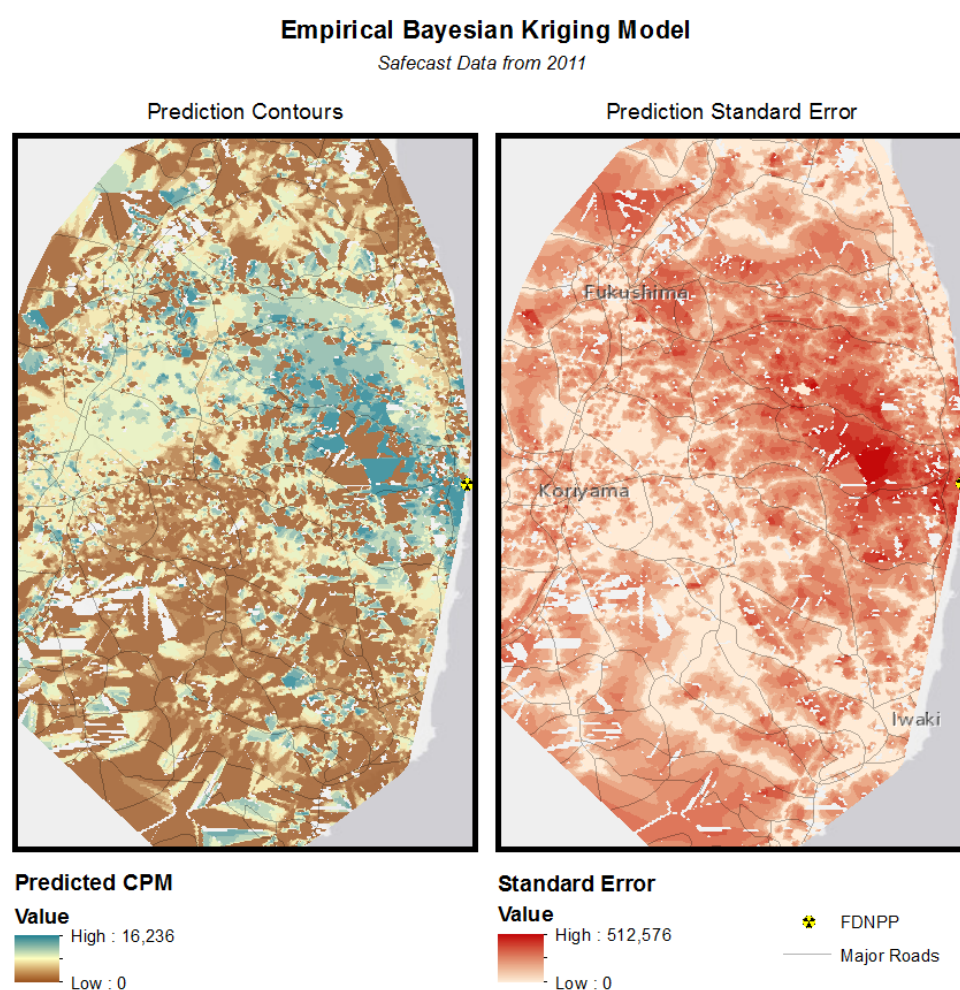


Figure 11. Bayesian model results of predicted values in CPM (left) and standard error of prediction (right).

In the current analysis, Bayesian methods indicated more certainty when measurements are consistent in each location so we also use this technique to consider the probability of collecting a data point that is above certain radiation level thresholds. Measurements at spatial distances are considered independently, but also probabilistically in relation to the other nearest measurements. This approach provides a more detailed understanding of uncertainty in spatio-temporal data to make better predictions. Bayesian modeling takes into account geostatistical relationships among the data for probabilistic prediction.

Figure 12, produced in GIS, illustrates this via a probability map showing the likelihood of collecting differing intervals of high values measurements over specified thresholds. The intervals are standardized throughout all of the maps. The model produces statistical outputs of the probability of the radiation level exceeding a specified limit. While there are many areas in which low radiation levels could be observed, setting higher thresholds allows for the isolation of the plume region of the highest concentrations of radiation. Due to the contours of the domains, only the areas that have observations of the highest levels are highlighted.

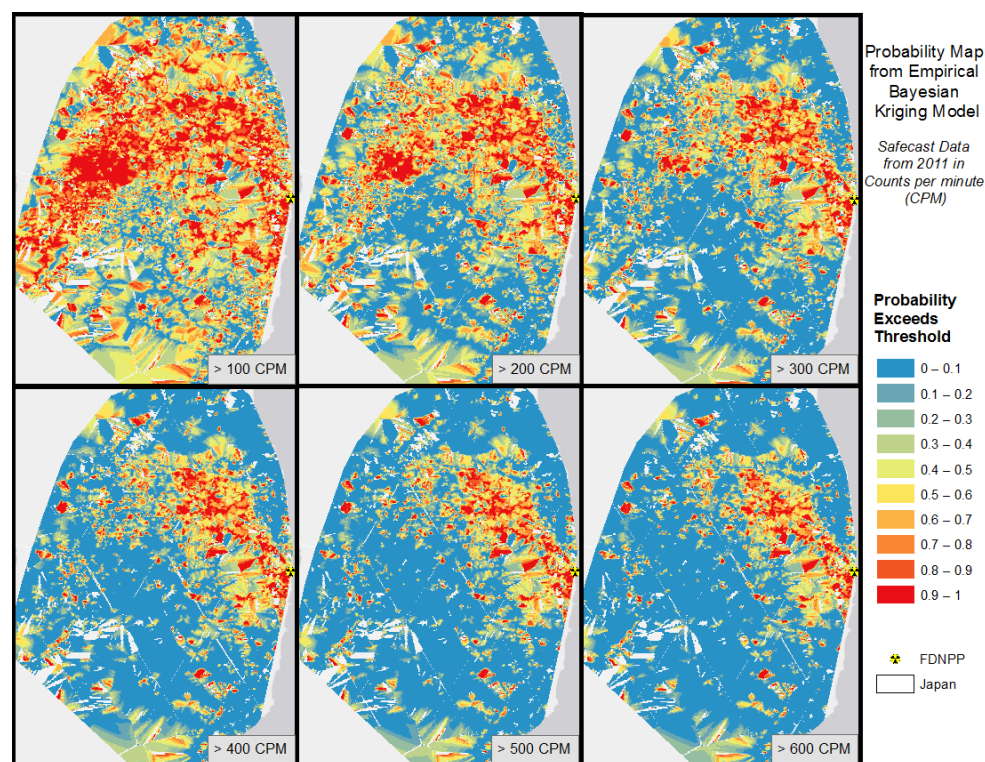


Figure 12. Probability that a Safecast measurement will exceed different threshold values in CPM.

In summary, the Bayesian kriging approach identifies some localized areas in which there are extremely high values recorded even outside of the plume. When aggregated using a rasterized approach, these ‘pockets’ of extreme values are obscured. When using classical kriging the high levels are erroneously expanded beyond areas in which these extreme values were actually recorded. As the surface is not consistently sampled, it can be insightful to evaluate extreme values in crowdsourced data with representation of uncertainty and probabilistic outputs.

5. Discussion

Bayesian methods take into account spatial relationships among the data for probabilistic prediction, which provides a way to represent uncertainty [36,37]. A Bayesian approach provides a statistical analysis of how reliable crowdsourced is and what errors may be present in the interpolated surface. Crowdsourced data could be evaluated for other aspects of the data such as the likelihood of collection in a certain time and place, trends of extreme values, measurement error, geographic error, and other forms of quality metrics. Future work could evaluate, as done by Tomassini et al. [66] for surface temperature, how much effect natural variability of radiation has on the uncertainty analysis.

Quantitative methods and comparisons can be helpful to express uncertainty inherent in the data [27], but this assessment needs to be grounded in an understanding of characteristics of the datasets and how the data are collected. Some aspects of novel data sources can also be evaluated from a qualitative perspective, such as the origins of a crisis project,

motivations, types of quality, resources, and if the training or expertise of the volunteers could influence the quality of collection [25]. Quality assurance protocols can be developed to handle crowdsourced [26] and enable contributors to flag potential issues in their own inputs helps to provide context even when using sensor devices. Even prior to crisis, establishing connections between citizen projects and organizations willing to perform evaluation could encourage the assessments of how projects can have their data used.

The interpretation of data is linked to characteristics of the data and the methods being used. These techniques should be critically considered when making use of data sources, particularly when the data relate to environmental hazards and are used for decision-making at any level from the public to authorities. Critical consideration should be given to the representation of spatial data, the selection of spatial analysis methods to present data-derived surfaces [35], and the techniques used for visualization. All of these choices can have a significant impact on the interpretation of the phenomenon. In general, geostatistical methods are needed that are scalable to large spatio-temporally varying datasets that do not have consistent sampling. Bayesian methods work well when sparseness is a factor [40]. Bayesian methods and metrics can be used to clearly visualize uncertainty [36,37] that is inherent in opportunistic spatial data. Building off of the work of Wainwright et al. [39], we assess that it is important to indicate observational gaps in data and the probability of elevated radiation levels when people could make decisions based on these maps.

Statistically derived surfaces are interpreted based on assumptions of the kriging model, the underlying data, and its appropriate visual representation. Further work could evaluate the influence that of these components have in crowdsourced mapping and use during hazards. This involves considerations for sampling and evaluating the issues of statistical assessment of opportunistic data using classical geostatistical techniques. In addition to the methods demonstrated here, other methods from systematic bias could be adopted from fields that are accustomed to adjusting for spatially and temporally non-random observations [67]. Additionally, the influence of spatio-temporal density should be evaluated in order to understand the spatio-temporal variance of the phenomenon, as well as the uncertainty present in the data. Further work could evaluate how the use of various statistical classification intervals to represent crowdsourced data values reveal different spatial patterns. Data could be visualized with various numerical classification methods to compare how spatial patterns of extreme values in crowdsourced data are influenced by selected methods.

Methods to visualize point data, such as aggregations or interpolations, have inherent assumptions behind these surfaces which should be evaluated and clearly described in light of the intended purpose of the map. It is important to note the temporal and spatial dimensions of the data collection included as opportunistic data are not sampled as a coherent survey. Spatial data points can be aggregated into grid cells of various sizes using a selection of statistical measures. In addition, clipping data to subsets based on location can lead to varied spatial distributions and resampling artifacts. Likewise, as a whole, datasets can be characterized by very different statistical distributions, but when a standardized subset is taken, they can exhibit similar properties. Bayesian spatial analysis and modeling can handle very large datasets and be used to capture statistical variations [68]. This research demonstrates approaches to evaluate the spatial and temporal dimensions of social big data when applying geostatistical techniques to contributed data. Bayesian spatial analysis could be useful for future studies on the spatial distribution of phenomena for other forms of crowdsourced data. Further work is needed in this field to capture and represent uncertainty, particularly for environmental monitoring and hazards.

6. Conclusions

Citizen-led initiatives producing spatio-temporal big data are potential sources of useful information during disasters [69]. Yet there is a scientific need to understand the uncertainty of big spatial data from a multitude of sources. Crowdsourced data is often opportunistic, so it may not capture a complete representation of a phenomenon,

and the statistical variations in the datasets are not typically assessed. Big data challenges add complexity to the use of geostatistical methods that were created for traditionally sampled datasets. As opportunistic environmental monitoring data are not intentionally sampled to be generalized, scalable techniques are needed to evaluate diverse forms of large volumes of data collected over spatio-temporal dimensions. Creating large spatio-temporally standardized datasets to overcome the infrequency of crowdsourced data in certain areas can broaden the scope and improve the reliability of insights gained.

Bayesian methods provide a way to represent uncertainty across the spatial distribution of large spatio-temporally varying opportunistic data. We demonstrate that the geostatistical variability of spatial data through Bayesian methods [35,40] can quantify uncertainty in the spatial distribution of crowdsourced data. The Bayesian approach that handles patchy data and also represents the spatial uncertainty [40] in derived surfaces is a promising direction for further work on assessing contributed data. In addition to comparison with sources of intentionally sampled dataset, spatially sampled subsets of the crowdsourced data can be compared to determine what data are necessary to have a better representation of the phenomenon. Beyond visualization techniques, inherent assumptions internal to geostatistical methods play a critical role in how the results will be interpreted. It is important to recognize factors that are inherent to the formation of the datasets and models that contribute to the differences between sources. These approaches need to be considered especially when these results could be used to make decisions about environmental hazards. While collecting more data in specific locations of interest could provide more coverage, ethical considerations should be evaluated prior to tasking the public collection of measurements in hazardous areas.

Estimation of geostatistical variations in the contributed dataset improves our scientific understanding of unevenly sampled data. Bayesian methods can quantify uncertainty and produce probabilistic outputs on the spatial distribution which in turn provides more representative visualizations. The reliability of using a visualization of opportunistic data is not just about the individual accuracy of measurements, but a product of how the collection was spatially sampled, the aggregation or interpolation mapping techniques applied, cartographic decisions, and the spatio-temporal properties of the phenomenon. We provide this approach with the recommendation that it be used as a framework for assessing other types of contributed data. A Bayesian approach provides a flexible methodology with unique solutions for big data problems which arise in the evaluation of crowdsourced data.

Author Contributions: Conceptualization, Carolynne Hultquist and Zita Oravecz; Data curation, Carolynne Hultquist and Guido Cervone; Formal analysis, Carolynne Hultquist; Funding acquisition, Carolynne Hultquist; Investigation, Carolynne Hultquist; Methodology, Carolynne Hultquist and Zita Oravecz; Supervision, Zita Oravecz and Guido Cervone; Visualization, Carolynne Hultquist; Writing—original draft, Carolynne Hultquist; Writing—review & editing, Zita Oravecz and Guido Cervone. All authors have read and agreed to the published version of the manuscript.

Funding: This material is based on work supported by the National Science Foundation under IGERT Grant DGE-1144860, Big Data Social Science.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: <https://safecast.org/data/download/> accessed on 7 December 2018. Safecast data is available under a CC0 designation—Public Domain Dedication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Flanagan, A.J.; Metzger, M.J. The credibility of volunteered geographic information. *GeoJournal* **2008**, *72*, 137–148. [CrossRef]
2. Connors, J.P.; Lei, S.; Kelly, M. Citizen Science in the Age of Neogeography: Utilizing Volunteered Geographic Information for Environmental Monitoring. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 37–41. [CrossRef]
3. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [CrossRef]

4. Tzavella, K.; Fekete, A.; Fiedrich, F. Opportunities provided by geographic information systems and volunteered geographic information for a timely emergency response during flood events in Cologne, Germany. *Nat. Hazards* **2018**, *91*, 29–57. [\[CrossRef\]](#)
5. Van Oudheusden, M.; Abe, Y. Beyond the grassroots: Two trajectories of “citizen sciencization” in environmental governance. *Citiz. Sci. Theory Pract.* **2021**, *6*, 1–15. [\[CrossRef\]](#)
6. Bonney, R.; Shirk, J.L.; Phillips, T.B.; Wiggins, A.; Ballard, H.L.; Miller-Rushing, A.J.; Parrish, J.K. Citizen science: Next steps for citizen science. *Science* **2014**, *343*, 1436–1437. [\[CrossRef\]](#)
7. Cooper, C.B.; Hawn, C.L.; Larson, L.R.; Parrish, J.K.; Bowser, G.; Cavalier, D.; Dunn, R.R.; Haklay, M.; Gupta, K.K.; Jelks, N.O.; et al. Inclusion in citizen science: The conundrum of rebranding. *Science* **2021**, *372*, 1386–1388. [\[CrossRef\]](#)
8. Bonney, R.; Ballard, H.; Jordan, R.; McCallie, E.; Phillips, T.; Shirk, J.; Wilderman, C.C. *Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education*; Technical Report July; Center for Advancement of Informal Science Education (CAISE): Washington, DC, USA, 2009.
9. Shirk, J.L.; Ballard, H.L.; Wilderman, C.C.; Phillips, T.; Wiggins, A.; Jordan, R. Public Participation in Scientific Research: A Framework for Deliberate Design. *Ecol. Soc.* **2012**, *17*, 1–20. [\[CrossRef\]](#)
10. Flagg, B.N. Contribution of Multimedia to Girls’ Experience of Citizen Science. *Citiz. Sci. Theory Pract.* **2016**, *1*, 1–13. [\[CrossRef\]](#)
11. Gray, S.; Jordan, R.; Crall, A.; Newman, G.; Hmelo-Silver, C.; Huang, J.; Novak, W.; Mellor, D.; Frensley, T.; Prysby, M.; et al. Combining participatory modelling and citizen science to support volunteer conservation action. *Biol. Conserv.* **2017**, *208*, 76–86. [\[CrossRef\]](#)
12. Haklay, M.; Fraisl, D.; Greshake Tzovaras, B.; Hecker, S.; Gold, M.; Hager, G.; Ceccaroni, L.; Kieslinger, B.; Wehn, U.; Woods, S.; et al. Contours of citizen science: A vignette study. *R. Soc. Open Sci.* **2021**, *8*, 1–24. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Schaefer, T.; Kieslinger, B. Supporting emerging forms of citizen science: A plea for diversity, creativity and social innovation. *J. Sci. Commun.* **2016**, *15*. [\[CrossRef\]](#)
14. Jordan, R.; Crall, A.; Gray, S.; Phillips, T.; Mellor, D. Citizen Science as a Distinct Field of Inquiry. *BioScience* **2015**, *65*, 208–211. [\[CrossRef\]](#)
15. Rickles, P.; Haklay, M.; Ellul, C.; Skarlatidou, A. Citizen Science with GIS & T. *Geogr. Inf. Sci. Technol. Body Knowl.* **2017**. [\[CrossRef\]](#)
16. Ishigaki, Y.; Matsumoto, Y.; Ichimiya, R.; Tanaka, K. Development of Mobile Radiation Monitoring System Utilizing Smartphone and Its Field Tests in Fukushima. *IEEE Sens. J.* **2013**, *13*, 3520–3526. [\[CrossRef\]](#)
17. Hemmi, A.; Graham, I. Hacker science versus closed science: Building environmental monitoring infrastructure. *Inf. Commun. Soc.* **2014**, *17*, 830. [\[CrossRef\]](#)
18. Brown, A.; Franken, P.; Bonner, S.; Dolezal, N.; Moross, J. Safecast: Successful citizen-science for radiation measurement and communication after Fukushima. *J. Radiol. Prot.* **2016**, *36*, S82–S101. [\[CrossRef\]](#)
19. Bonner, S.; Brown, A.; Cheung, A. *The Safecast Report*; Technical Report; The Safecast Team. 2015. Available online: <https://safecast.org/about/research/> (accessed on 26 November 2021).
20. Brown, A.; Franken, P.; Bonner, S. *The Safecast Report*; Technical Report; The Safecast Team. 2016. Available online: <https://safecast.org/about/research/> (accessed on 26 November 2021).
21. Brown, A.; Franken, P.; Moross, J.; Dolezal, N.; Bonner, S. *The Safecast Report*; Technical Report; The Safecast Team. 2017. Available online: <https://safecast.org/about/research/> (accessed on 26 November 2021).
22. Safecast. *About Safecast*. 2021. Available online: <https://safecast.org/about/> (accessed on 26 November 2021).
23. Monroe, B.L. The Five Vs of Big Data Political Science: Introduction to the Virtual Issue on Big Data in Political Science Political Analysis. *Political Anal.* **2012**, *5*, 1–9. [\[CrossRef\]](#)
24. Hultquist, C.; Sava, E.; Cervone, G.; Waters, N. Damage Assessment of the Urban Environment during Disasters using Volunteered Geographic Information. In *Big Data for Regional Science*; Shintler, L., Chen, Z., Eds.; CRC Press: Boca Raton, FL, USA, 2017; Chapter 18.
25. See, L.; Comber, A.; Salk, C.; Fritz, S.; Velde, M.V.D.; Perger, C.; Schill, C.; McCallum, I.; Kraxner, F.; Obersteiner, M. Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts. *PLoS ONE* **2013**, *8*, e069958. [\[CrossRef\]](#)
26. Samulowska, M.; Chmielewski, S.; Raczko, E.; Lupa, M.; Myszkowska, D.; Zagajewski, B. Crowdsourcing without Data Bias: Building a Quality Assurance System for Air Pollution Symptom Mapping. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 46. [\[CrossRef\]](#)
27. Freitag, A.; Meyer, R.; Whiteman, L. Strategies Employed by Citizen Science Programs to Increase the Credibility of Their Data. *Citiz. Sci. Theory Pract.* **2016**, *1*, 2. [\[CrossRef\]](#)
28. Xin, Y. Anomaly detection for volunteered geographic information: A case study of Safecast data. *Int. J. Geogr. Inf. Sci.* **2021**, 1–20. [\[CrossRef\]](#)
29. Wiggins, A.; Newman, G.; Stevenson, R.D.; Crowston, K. Mechanisms for Data Quality and Validation in Citizen Science. In Proceedings of the IEEE Seventh International Conference on e-Science, Stockholm, Sweden, 5–8 December 2011; pp. 1–6.
30. Panek, J.; Marek, L.; Paszto, V.; Valuch, J. The Crisis Map of the Czech Republic: The nationwide deployment of an Ushahidi application for disasters. *Disasters* **2016**, *41*, 649–671. [\[CrossRef\]](#)
31. Jackson, S.P.; Mullen, W.; Agouris, P.; Crooks, A.; Croitoru, A.; Stefanidis, A. Assessing Completeness and Spatial Error of Features in Volunteered Geographic Information. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 507–530. [\[CrossRef\]](#)
32. Catlin-Groves, C.L. The citizen science landscape: From volunteers to citizen sensors and beyond. *Int. J. Zool.* **2012**, *2012*, 1–14. [\[CrossRef\]](#)

33. Welvaert, M.; Caley, P. Citizen surveillance for environmental monitoring: Combining the efforts of citizen science and crowdsourcing in a quantitative data framework. *SpringerPlus* **2016**, *5*, 1–14. [CrossRef]
34. Hultquist, C.; Simpson, M.; Huang, Q.; Cervone, G. Using Nightlight Remote Sensing Imagery and Twitter Data to Study Power Outages. In Proceedings of the ACM SIGSPATIAL, Bellevue, WA, USA, 3–6 November 2015; pp. 1–6. [CrossRef]
35. Diggle, P.J.; Tawn, J.A.; Moyeed, R.A. Model-based Geostatistics. *Appl. Stat.* **1998**, *47*, 299–350. [CrossRef]
36. Leasure, D.R.; Jochem, W.C.; Weber, E.M.; Seaman, V.; Tatem, A.J. National population mapping from sparse survey data: A hierarchical Bayesian modeling framework to account for uncertainty. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 24173–24179. [CrossRef]
37. Wu, X.; Marshall, L.; Sharma, A. Quantifying input error in hydrologic modeling using the Bayesian error analysis with reordering (BEAR) approach. *J. Hydrol.* **2021**, *598*, 126202. [CrossRef]
38. Mirzaei, R.; Sakizadeh, M. Comparison of interpolation methods for the estimation of groundwater contamination in Andimeshk-Shush Plain, Southwest of Iran. *Environ. Sci. Pollut. Res.* **2016**, *23*, 2758–2769. [CrossRef]
39. Wainwright, H.M.; Seki, A.; Chen, J.; Saito, K. A multiscale Bayesian data integration approach for mapping air dose rates around the Fukushima Daiichi Nuclear Power Plant. *J. Environ. Radioact.* **2017**, *167*, 62–69. [CrossRef]
40. De Risi, R.; De Luca, F.; Gilder, C.E.; Pokhrel, R.M.; Vardanega, P.J. The SAFER geodatabase for the Kathmandu valley: Bayesian kriging for data-scarce regions. *Earthq. Spectra* **2021**, *37*, 1108–1126. [CrossRef]
41. Li, Y.; Hernandez, J.H.; Aviles, M.; Knappett, P.S.; Giardino, J.R.; Miranda, R.; Puy, M.J.; Padilla, F.; Morales, J. Empirical Bayesian Kriging method to evaluate inter-annual water-table evolution in the Cuenca Alta del Río Laja aquifer, Guanajuato, México. *J. Hydrol.* **2020**, *582*, 124517. [CrossRef]
42. Fairbairn, D.; Al-Bakri, M. Using Geometric Properties to Evaluate Possible Integration of Authoritative and Volunteered Geographic Information. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 349–370. [CrossRef]
43. Safecast. *bGeigie Nano—Safecast*. 2021. Available online: <https://safecast.org/devices/bgeigie-nano/> (accessed on 26 November 2021).
44. Safecast. *Licenses We Use—Safecast*. 2021. Available online: <https://safecast.org/about/licenses/> (accessed on 26 November 2021).
45. Cervone, G.; Hultquist, C. Calibration of Safecast dose rate measurements. *J. Environ. Radioact.* **2018**, *190–191*, 51–65. [CrossRef]
46. Department of Energy. *US DOE/NNSA Response to 2011 Fukushima Incident—Raw Aerial Data and Extracted Ground Exposure Rates and Cesium Deposition*; Department of Energy. 2011. Available online: <https://catalog.data.gov/dataset/us-doe-nnsa-response-to-2011-fukushima-incident-raw-aerial-data-and-extracted-ground-expos> (accessed on 26 November 2021).
47. Lyons, C.; Colton, D. Aerial Measuring System in Japan. *Health Phys.* **2012**, *102*, 509–515. [CrossRef] [PubMed]
48. Masoudi, P.; Le Coz, M.; Cazala, C.; Saito, K. Spatial properties of soil analyses and airborne measurements for reconnaissance of soil contamination by ¹³⁷Cs after Fukushima nuclear accident in 2011. *J. Environ. Radioact.* **2019**, *202*, 74–84. [CrossRef]
49. Coletti, M.; Hultquist, C.; Kennedy, W.G.; Cervone, G. Validating Safecast data by comparisons to a U. S. Department of Energy Fukushima Prefecture aerial survey. *J. Environ. Radioact.* **2017**, *171*, 9–20. [CrossRef] [PubMed]
50. Hultquist, C.; Cervone, G. Citizen monitoring during hazards: Validation of Fukushima radiation measurements. *GeoJournal* **2017**, *83*, 189–206. [CrossRef]
51. Japan Atomic Energy Agency. *Airborne Monitoring in the Distribution Survey of Radioactive Substances*; Japan Atomic Energy Agency. 2014. Available online: <http://emdb.jaea.go.jp/emdb/en/portals/b224/> (accessed on 26 November 2021).
52. Hultquist, C.; Cervone, G. Comparison of simulated radioactive atmospheric releases to citizen science observations for the Fukushima nuclear accident. *Atmos. Environ.* **2019**, *198*, 478–488. [CrossRef]
53. Sarkar, D.; Andrews, F.; Wright, K.; Klepeis, N.; Larsson, J.; Murrell, P. *Lattice: Trellis Graphics for R*, R Package Version 0.20-45. 2021. Available online: <https://cran.r-project.org/package=lattice> (accessed on 26 November 2021).
54. Newirth, E. *RColorBrewer: ColorBrewer Palettes*; R Package Version 1.1-2. 2014. Available online: <https://cran.r-project.org/package=RColorBrewer> (accessed on 26 November 2021).
55. Hijmans, R.J. *Raster: Geographic Data Analysis and Modeling*; R Package Version 3.4-13. 2021. Available online: <https://cran.r-project.org/package=raster> (accessed on 26 November 2021).
56. Evans, J.S. *spatialEco*, R Package Version 1.3-6. 2021. Available online: <https://cran.r-project.org/package=spatialEco> (accessed on 26 November 2021).
57. Carr, D.; Lewin-Koh, N.; Maechler, M. *Hexbin: Hexagonal Binning Routines*; R Package Version 1.28.2. 2021. Available online: <https://cran.r-project.org/package=hexbin> (accessed on 26 November 2021).
58. Finley, A.; Banerjee, S.; Hjelle, Ø. *MBA: Multilevel B-Spline Approximation*, R Package Version 0.0-9. 2017. Available online: <https://CRAN.R-project.org/package=MBA> (accessed on 26 November 2021).
59. Nychka, D.; Furrer, R.; Paige, J.; Sain, S. *Fields: Tools for Spatial Data*; R Package Version 12.5. 2017. Available online: <https://CRAN.R-project.org/package=fields> (accessed on 26 November 2021).
60. Gribov, A.; Krivoruchko, K. Empirical Bayesian kriging implementation and usage. *Sci. Total Environ.* **2020**, *722*, 137290. [CrossRef] [PubMed]
61. Pilz, J.; Spöck, G. Why do we need and how should we implement Bayesian kriging methods. *Stoch. Environ. Res. Risk Assess.* **2008**, *22*, 621–632. [CrossRef]
62. Krivoruchko, K.; Gribov, A. Evaluation of empirical Bayesian kriging. *Spat. Stat.* **2019**, *32*, 100368. [CrossRef]

-
63. Aelion, C.M.; Davis, H.T.; Liu, Y.; Lawson, A.B.; McDermott, S. Validation of Bayesian Kriging of Arsenic, Chromium, Lead, and Mercury Surface Soil Concentrations Based on Internode Sampling. *Environ. Sci. Technol.* **2009**, *43*, 4432–4438. [[CrossRef](#)] [[PubMed](#)]
 64. Recta, V. A Model-Based Analysis of Semicontinuous Spatial Data. Ph.D. Thesis, The Pennsylvania State University, State College, PA, USA, 2009. Available online: <https://etda.libraries.psu.edu/catalog> (accessed on 26 November 2021).
 65. Brewer, C. Spectral Schemes: Controversial Color Use on Maps. *Cartogr. Geogr. Inf. Sci.* **1997**, *24*, 203–220. [[CrossRef](#)]
 66. Tomassini, L.; Reichert, P.; Knutti, R.; Stocker, T.F.; Borsuk, M.E. Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods. *J. Clim.* **2007**, *20*, 1239–1254. [[CrossRef](#)]
 67. Kosmala, M.; Wiggins, A.; Swanson, A.; Simmons, B. Assessing data quality in citizen science. *Front. Ecol. Environ.* **2016**, *14*, 551–560. [[CrossRef](#)]
 68. Katzfuss, M.; Cressie, N. Bayesian hierarchical spatio-temporal smoothing for very large datasets. *Environmetrics* **2012**, *23*, 94–107. [[CrossRef](#)]
 69. Degrossi, L.C.; De Albuquerque, J.P.; Fava, M.C.; Mendiando, E.M. Flood citizen observatory: A crowdsourcing-based approach for flood risk management in Brazil. In Proceedings of the International Conference on Software Engineering and Knowledge Engineering (SEKE), Vancouver, BC, Canada, 1–3 July 2014; pp. 570–575.