An online benchmarking platform for visualizing ionizing radiation doses in different cities

Julián Villegas Computer Arts Lab. University of Aizu Aizu-Wakamatsu, Japan julian at u-aizu.ac.jp

Abstract—A working prototype for alternative visualizations of environmental data (currently, ionizing radiation) measured with bGeigie nano Safecast sensors is presented. Contrary to previous interfaces, in this visualization users have finer control of the displayed data (i.e., can determine date ranges, compare locations, decide the averaging areas, etc.) and more detailed information of the resulting visualization (size of the samples per day and per region, etc.). With this new data visualization, it is easier to compare local environment figures with those of other regions of the planet.

1. Introduction

In the aftermath of the Great East Japan earthquake, a $9.0 M_W$ earthquake and tsunami [1] that caused level 7 meltdowns at three nuclear reactors in Fukushima Daiichi power plant [2], concerns about health and safety were exacerbated by the scarcity of reliable information [3], [4]. Given the magnitude of the problem and its possible consequences, these concerns were shared among local and global communities [5], [6]. Unavoidably, Japan (and especially Fukushima prefecture) were associated with dangerous levels of radiation. Within Japan, this stigma discouraged residents from other areas to accept Fukushima residents willing to relocate [7], deterred Japanese tourist from visiting Fukushima's attractions and investors from participating in business opportunities in the region, etc. [8]. To some extent, the same trends were observed internationally, increasing the difficulties faced by Japan to revitalize the affected region.

Among several efforts to rapidly (and accurately) assess the levels of ionizing radiation (γ -rays, X-rays, high ultraviolet spectrum, etc.) around Japan [9], [10], a group of volunteers (Safecast) started monitoring the levels in different cities and publishing those measurements so anyone could judge how severe the situation was in a given region. A typical data log comprised twelve measurements per minute (one every five seconds), sometimes spanning periods of hours. As more volunteers joined this project, it was increasingly difficult to figure radiation patterns from disparate sensors and locations. To ameliorate this problem, several visualizations of data collected with Safecast sensors were made available: from simple 'pins' (points) in a Google map to sophisticated interpolation maps. Some of these maps are still available from Safecast's website [11].

The above-mentioned visualizations satisfy most of the users' needs such as quick assessment of radiation levels in contiguous regions, visualization of latest captures from web browsers, iOS devices, etc. Because the huge amount of data involved, these visualizations offer rather little opportunities for data exploration. For example, users cannot easily select date ranges, so it is difficult to understand if an observed pattern is the result of a recent or past event. Also, it is difficult to determine if the level being observed corresponds to single or multiple measurements; additionally, for a given zoom level, users cannot control the number of area cells to aggregate the data; finally, the current visualizations present data in radiation units which may not be completely understood by laymen, hindering assessment of risk and comparison between places around the world.

The purpose of this research is two-fold: to fill the breach left by current Safecast visualizations, i.e., empowering the user with a finer control on how data is presented, and to pave the way for future benchmarking of environmental indicators other than ionizing radiation. To illustrate the latter point, when considering moving to a different city, a person could be interested in comparing her current location to the new one in terms of air quality, noise polution, etc. These indicators are usually found from disparate sources and their comparison may be cumbersome. With more sensors integrated into the Safecast kit, a more convenient way to compare them could be implemented using graphic instead of textual information.

The following section further discusses details of Safecast's devices and maps, this is followed by a description of the implemented system with its merits and shortcomings; this section segues into a brief discussion on future work on comparative environmental information and conclusions.

2. Background

Safecast is a successful example of citizen science. This non-governmental/non-profit organization works primarily

⁸th Euro American Conf. on Telematics and Information Systems (EATIS), Cartagena de Indias, Colombia (2016).

in mapping radiation levels and building sensor networks, enabling non-specialists to contribute and use the collected data. In 2013, Safecast was honored with the Good Design Award for its Radiation sensor network.¹ Safecast's data and sensors are open hardware and open software published under a Creative Commons licenses [12]. Originally focused on background radiation in Japan, collecting environmental data on a global level is one of the long-term objectives of Safecast.

As of March 2016, about 42 million measurements are recorded in Safecast's database [11], and the growth rate is accelerating with an increasing number of sensors capturing data. Measurements comprise an individual ID, timestamp, unit, value, and location (latitude, longitude). A great amount of these measurements are captured using bGeigie sensors.

2.1. bGeigie sensors

The bGeigie nano sensor is one of the devices employed in data collection. These sensors are usually created by end users in workshops around the world where Safecast experts guide participants in their construction, so no previous experience with hardware, soldering, etc. is necessary. These sensors fit inside a small bento box and can be used as a dosimeter (Geiger counter), directly pointing the device at a measurement target, or to log radiation levels during a given period. In the former case, measurements are displayed in μ Sv/h (micro-Sieverts per hour), and in the latter, in counts per minute (CPM).

Sievert (Sv) is the unit to measure health impact of exposure to ionizing radiation. The world average dose rate of radiation due to natural causes (cosmic rays, ingestion, Earth radiation, etc.) is about $0.273 \,\mu$ Sv/h. In comparison, the most radioactive measurement found so far in the main building of Fukushima Daiichi plant is about $20 \,\mu$ Sv/h [13]. BGeigie sensors are capable of reporting in CPM units, so estimation of the radiation dose rate is carried out by assuming that most radiation events are produced by γ -rays from m137 Ba, a nuclear isomer produced by the decay of 137 Cs, which in turn, is produced in fission reactions such as those that occurred in the Fukushima Daiichi meltdowns. The accuracy of this approximation is $\pm 15\%$, as confirmed on several laboratories [14].

A typical user would log measurements every 5 s along a path (e.g., a highway, streets in a city, tracks in the countryside) and upload these measurements to Safecast servers at least in two fashions: via email and through the Safecast web API. When submitted by email, a robot replies with a summary of the measurements. This summary comprises descriptive statistics and a color-coded map with average dose rates per tile; these results are not incorporated into the Safecast database. When submitted via the Safecast web API, the log is verified by an authorized person and ultimately, these data are incorporated into the database of measurements.

1. http://www.g-mark.org/award/describe/40575

2.2. Current visualizations

As previously mentioned, there are several visualizations of Safecast data. For sake of brevity, only two of the main visualizations are discussed in this article: Safecast and Drive maps. These visualizations are summarized in Figures 1 and 2. The information provided in this section is based on the detailed explanation found in [15]. The source code of the discussed visualizations is also available there.



Figure 1. Safecast maps. These maps summarize Safecast data and are available via web, iOS, and Mac OS X native applications.

2.2.1. Safecast maps. These maps, illustrated in Figure 1, are officially endorsed by Safecast and can be accessed via web, and native iOS and OS X applications. Whereas the iOS and OS X applications are dynamic (in the sense that the tiles are created each user inquiry), the web map is a collection of static .PNG tiles. To create these maps, data is firstly curated (eliminating suspect measurements, usually through a blacklist manually edited), then CPM values are converted to μ Sv/h using a different conversion constant depending on the device type. Finally, the GPS data are reformatted to Mercator projection at a zoom level of about 19 m (level 13). The most recent data (up to 270 days) in each point are aggregated, so it is possible that for two neighboring tiles, the date ranges differ. The curated dataset is stored on PostgreSQL and SQLite3 databases. The latter are used for the iOS and OS X applications. In these applications, interpolations are created for different zoom levels on the hosting device. The interpolation for the web maps is achieved by resampling the .PNG tiles at zoom level 13.

In these representations, among other features, users can navigate the world by dragging and zooming gestures, select different color schemes for representing radiation levels, several base maps, include radiation readings from some selected sources.



Figure 2. Drive maps. These maps are useful to visually check individual data submissions.

2.2.2. Drive maps. As illustrated in Figure 2, this visualization shows individual maps for captures performed by Safecast volunteers (usually while driving a vehicle). These maps, available from [11], are created by a client-side application written in Javascript. Upon processing a user's query, this application connects to Safecast servers and relies the query. The server response is received via JSON (JavaScript Object Notation) and is parsed, and processed to overlay each point onto a map retrieved using Google Maps API. Given the potential complexity of this task, data processing is multi-threaded using web-workers. As a way to improve the performance of the application, occluded points are not displayed.

In these maps, each datum is represented by a colorcoded circle over a road, terrain, or Satellite map. Fewer CPM measurements are represented as less saturated circles; by hovering over these circles additional information (i.e., captured time, latitude, longitude, CPM, and dose rate in μ S/h) are displayed next to the map. **2.2.3. Problems with current visualizations.** Given the complexity of the visualizations, some compromises were made in order to produce maps in a reasonable amount of time. Although these maps have proven to be very useful, some shortcomings are mentioned here as a way of justification for the newly created visualization:

- Offline interpolation is slow and needs daily updates ($\approx 20 \text{ min}$, for 90 MB, in the case of iOS App).
- Temporal trends are hidden in the current interpolations.
- Users have little control on how the data is displayed.
- It is difficult to understand the overall situation by looking at individual measurements only.
- For naive users, it is difficult to assess the radioactivity risk by only looking at the dose rate figures.

3. Implementation



http://onkyo.u-aizu.ac.jp:3838/safecast

Figure 3. Implementation overview of the proposed visualization system. The core functionalities were developed using R and its libraries.

Following the guidelines provided in [15], the curated data offered there was used for the creation of this visualization. In order to ease real-time computations several simplifications were taken:

- 1) Only CPM data were used.
- 2) Capture times were rounded to the day,
- Azimuth and elevation angles were rounded to four decimal places,
- 4) Sample size N and arithmetic mean radiation V for each datapoint of this newly formed subset were computed and stored in an SQLite database (safecast.db),
- 5) No Mercator projection was carried out (i.e., the stored locations conform the original EPSG:4326 standard).
- 6) Indices were created on azimuth and elevation fields of this database.

Keeping only CPM data is justified by the fact that although the origin of the radiation around Japan is likely to be the decay of 137 Cs, this might not be the case in other parts of the world. Note that by reducing the number of decimal digits of latitude and longitude, the resolution of the map is effectively reduced. In the worst case (at the equator), latitude and longitude would have uncertainties of 11.132 m, and 10.247 m, approximately. Also, note that the Mercator projection depends on the size of the map, which in turn, depends on the zoom level; in this proposal, this projection is carried out at the time of the query. The overview of the implementation is depicted in Figure 3 and can be accessed online via [16]. The source code of this application is freely available (i.e., distributed under a GNU license) via [17].

The core of the application is written in R (v. 3.2.3) [18], an extensible functional language to explore datasets, distributed under a GNU license. The created R program uses the Shiny library (v. 0.13.1) [19], which is an R framework for creating interactive web applications. Displaying plots and maps is achieved by mainly using ggplot2 (v. 2.1.0) [20] and ggmap (v. 2.6.1) [21] libraries, as detailed in the following sections.

3.1. Server side

The Shiny server runs in an Ubuntu platform and listens to port 3838. The server takes care of running all the necessary code to produce the actual visualizations. It reacts to changes in a user's web-form, reflecting the new parameter values in the visualizations. Note that depending on the specifications of some parameters such as time ranges and zoom factor, a visualization can take up to several minutes to be rendered.

When a query is received from a client, the shiny server queries Google Maps, OpenStreetMap, Stamen Maps, or Naver Map servers, depending on the specified map type and rasters if it exists. These maps are saved as temporary .PNG files with a unique identifier per session and query. Being able to choose different kind of maps is an important feature that helps data exploration. For example, a user interested in observing radiation patterns along the path of a river may choose a terrain map, whereas a user preparing an illustration for publication may be interested in a black-and-white map to enhance the color-coded radiation figures. According to the query, the server also retrieves the environmental data corresponding to the data range from the SQLite3 database. With these data, radiation over time, spatial averaging, and city comparison visualizations are created.

Periodically (and automatically), old files are deleted from the server and the local database is updated with new values retrieved from Safecast servers. This process still needs manual verification to ensure that updates on Safecast's blacklists are reflected on the local database.

3.2. Client side

The client side of the application include a web-form from which users can input parameters for modifying visu-



Figure 4. Parameters for data exploration.

alizations. Currently, these parameters include location, indicator, date range, zoom factor, number of averaging cells, and map type for timeline and spatial average visualizations, as shown in Figure 4. For city comparisons, up to three locations can be selected. Throughout all the visualizations, the same color code (a spectral diverging palette with 11 levels) is used to show radiation dose levels.

Locations can be input in several formats such as coordinates (longitude, latitude), addresses, city names, countries, etc.; by default, the visualization starts centered in the city of Aizu-Wakamatsu, Fukushima prefecture (Japan) and displays measurements of the last three months.

Although only radiation data is currently present in the database, several efforts are being conducted within the Safecast community to support other environmental indicators (CO, NOx, Temperature, Humidity, etc.); foreseeing such extensions, users can select the kind of indicator of their interest (currently only radiation is working). Users can also specify, in a very intuitive manner, a range of dates for the visualization (by default, the last three months). In addition, users can modify the zoom factor of a map (small numbers produce large visualization areas), the number of hexagonal cells on which the space averaging should be computed, and the map type (terrain, road, satellite, black-and-white, etc.).

Users can access the web application from several web clients. Currently, Chrome, Opera, and Safari web browsers running on Mac OS X and iOS platforms have been suc-

cessfully tested.

3.3. Radiation over time

This visualization presents the average radiation and number of measurements between specified dates, as shown in Figure 5. An individual vertical bar corresponds to measurements captured in a single day. Its height corresponds to the sample size (number of measurements in that day for the specified region) and its color corresponds to the mean radiation per day. These mean values are actually the weighted arithmetic mean

$$m_w = \sum (VN) / \sum (N), \tag{1}$$

where V is the recorded (arithmetic) mean radiation in CPMs for a given day, longitude, and latitude, and N is the number of measurements from which this average was computed.

3.4. Spatial average

For the spatial averaging, the weighted mean radiation value on the number of specified cells is computed. The results are color-coded with the same color scheme as in the time-line visualization. In the ggplot2 library, data can be averaged on several shapes, including rectangular and hexagonal tiles. The latter were chosen for being the best way to divide the raster into regions of equal area with the smallest total perimeter, i.e., hexagons are more similar to circles than rectangles are. This property allows aggregation of data around the hexagon center more efficiently [22]. These hexagonal cells are overlaid on top of the rasterized map and decrease their alpha value (i.e., transparency) with increasing radiation levels so that the background map is still visible for low radiation levels and "hot" spots readily identified. An example of this visualization is presented in Figure 6.

3.5. Comparison between cities

One of the problems of conveying the risks of ionizing radiation to a general audience is the lack of an anchor to compare with when looking at figures. Health impacts of exposure to ionizing radiation are expressed in Sv, but dose rates can be expressed in different time units (e.g., μ Sv/h, μ Sv/year, etc.); furthermore, different units to measure radiation (e.g., Gy-Gray, REM-Rötgen Equivalent in Man, etc.) are often used, increasing confusion for laymen. Arguably, what most people are interested in is the assessment of the radiation level in a place compared to the level of a known "safe," "clean" location. This is the motivation for the last visualization where users can choose up to three cities to compare their radiation levels. The weighted means are color coded as in the previous visualizations and error bars showing the 95% Confidence Interval are added on top of the bars, as illustrated in Figure 7.

This figure shows that radiation doses in Rome, measured with Safecast sensors, are more than double those recorded over the same period in Aizu-Wakamatsu (a city about 100 km west of the crippled plant), and that the levels in Aizu-Wakamatsu are similar to those found in Cali (Colombia). The origin of the seemingly high levels found in Rome are unknown by the author, but this fact has been reported before by the Italian press (See [23], for example). There is no "safe" amount of radiation exposure, but this finding does not imply that radiation levels in Rome are dangerous; on the contrary, it shows that associating most Fukushima cities with high levels of radiation is unfounded. This association, however, is often made by naive observers.

4. Limitations

The present article reports on the progress of an ongoing visualization project that is far from being mature. Many limitations have been already identified, the main ones being errors in averaging and computation of the radiation doses. Visualization of radiation data uses arithmetic mean values of daily records stored in the database. Note that given the exponential decaying nature of radioactive materials, such average is neither accurate nor appropriate for scientific purposes. Moreover, this error increases as longer periods are considered. However, the purpose of the visualizations proposed in this article is to ease the comparison between different levels at different places, and such error is a necessary trade-off. Also, as previously mentioned, conversion of radiation doses from CPM to μ S/h is done based on the assumption that the ionizing events are produced by γ -rays from m^{137} Ba which may not always be true.

The discussed visualizations use Google Maps API to determine the coordinates of locations of interest. Currently, the free version of this service limits the number of queries to 2,500 a day.

Finally, since the rendering of visualization is made at the time of the request, some compromises have been made to keep timely responses: zoom factor is limited to 8 (i.e., 611 m/pixel) and navigating the globe by dragging gestures (as commonly used in smartphones and tablets) has not been implemented.

5. Future Work

As the number of records increase on the database of Safecast, it is becoming increasingly difficult to deal with such a huge amount of data. To solve that problem, an application based on the pbdR (Programming with Big Data in R) library [24] is currently being evaluated. Alternative visualizations based on the Leaflet library [25], a modern map library recently added to the Shiny library repertoire, are also being investigated to allow dragging gesture navigation.

In anticipation to the incorporation of other environmental indicators into the Safecast sensor kits, different visualization such as radar plots for simultaneously comparing several environment indicators in several cities, are being currently investigated.



Figure 5. Timeline visualization of Safecast measurements centered at Aizu-Wakamatsu (Fukushima prefecture, Japan) with a zoom factor of 8 from 2011 to 2016. Bar heights correspond to the number of measurements in each day, and their color to the mean dose rate over the same period.





Figure 6. Spatial averaging of Safecast radiation measurements centered at Aizu-Wakamatsu (Fukushima Prefecture, Japan) with a zoom factor of 8 (i.e., 611 m/pixel) from data collected between 2011 and 2016. Radiation doses are averaged over 100 hexagonal cells in both vertical and horizontal directions.

Figure 7. Radiation comparison between Aizu-Wakamatsu (Fukushima Prefecture, Japan), Cali (Colombia) and Rome (Italy) from Safecast data captured between January 2015, and March 2016. Error bars correspond to 95% Confidence Intervals.

6. Conclusions

An alternative visualization method for ionizing radiation data captured with Safecast sensors was presented. The benefits and shortcomings of previous and present visualizations were discussed. Relative to the previous visualizations, the one presented here offers a finer control of the data to be visualized both in time and spatial averaging.

Radiation exposure assessments are difficult to understand and it is easy to prejudge regions on basis of speculation. Five years after the catastrophe in Fukushima Daiichi plant, the situation there is still unstable and the general understanding and vigilance of information gathered by independent sources (such as Safecast) is a way to prevent future events of the same kind. It is the hope of the author that tools like the visualizations discussed here could help in this process of social awareness.

7. Acknowledgments

This research has been funded by the Competitive Research Funds of the University of Aizu.

References

- United States Geological Survey (USGS), "Earthquake magnitude 9.03-near the east coast of Honshu, Japan," Retrieved April 5, 2016. Available from http://earthquake.usgs.gov/earthquakes/map/.
- [2] R. Pool, "Fukushima: the facts," *Engineering Technology*, vol. 6, no. 4, pp. 32–36, May 2011.
- [3] L. Pierpoint, "Fukushima, Facebook and Feeds: Informing the Public in a Digital Era," *The Electricity Journal*, vol. 24, no. 6, pp. 53–58, 2011.
- [4] G. J. Rubin, R. Amlôt, S. Wessely, and N. Greenberg, "Anxiety, distress and anger among British nationals in Japan following the Fukushima nuclear accident," *The British Journal of Psychiatry*, vol. 201, no. 5, pp. 400–407, 2012.
- [5] A. Sidorin, "Radiation situation in Kamchatka after the Fukushima nuclear power station accident," *Izvestiya, Atmospheric and Oceanic Physics*, vol. 49, no. 8, pp. 860–869, 2013.
- [6] J. D. Leon, D. Jaffe, J. Kaspar, A. Knecht, M. Miller, R. Robertson, and A. Schubert, "Arrival time and magnitude of airborne fission products from the Fukushima, Japan, reactor incident as measured in Seattle, WA, {USA}," *Journal of Environmental Radioactivity*, vol. 102, no. 11, pp. 1032–1038, 2011.
- [7] J. Yamashita and J. Shigemura, "The great east Japan earthquake, tsunami, and Fukushima Daiichi nuclear power plant accident: A triple disaster affecting the mental health of the country," *Psychiatric Clinics of North America*, vol. 36, no. 3, pp. 351–370, 2013.
- [8] F. Yamane, H. Ohgaki, and K. Asano, "The immediate impact of the Fukushima Daiichi accident on local property values," *Risk Analysis*, vol. 33, no. 11, pp. 2023–2040, 2013.
- [9] Y. Ishigaki, Y. Matsumoto, R. Ichimiya, and K. Tanaka, "Development of mobile radiation monitoring system utilizing smartphone and its field tests in Fukushima," *Sensors Journal, IEEE*, vol. 13, no. 10, pp. 3520–3526, Oct 2013.
- [10] J.-C. Plantin, "The politics of mapping platforms: participatory radiation mapping after the Fukushima Daiichi disaster," *Media, Culture* & *Society*, vol. 37, no. 6, pp. 904–921, 2015.
- [11] "Safecast Website," [Software]. Retrieved April 5, 2016. Available from http://blog.safecast.org/.

- [12] "Safecast Licenses," [Software]. Retrieved April 5, 2016. Available from http://blog.safecast.org/faq/licenses/.
- [13] TEPCO. Tokyo Electric Power Company, "Radiation Dose measured at Monitoring Post of Fukushima Daiichi Nuclear Power Station," [Software]. Retrieved April 5, 2016. Available from http://www.tepco. co.jp/en/nu/fukushima-np/f1/index-e.html.
- [14] "Safecast bGeigie Calibration," [Software]. Retrieved April 5, 2016. Available from http://blog.safecast.org/faq/ about-calibration-and-the-bgeigie-nano.
- [15] "Safecast Methodology," [Software]. Retrieved April 5, 2016. Available from http://safecast.org/tilemap/methodology.html.
- [16] "Safecast data visualization," [Software]. Retrieved April 5, 2016. Available from http://onkyo.u-aizu.ac.jp:3838/safecast.
- [17] "Safecast data visualization source code," [Software]. Retrieved April 5, 2016. Available from https://julovi@bitbucket.org/julovi/ visualization.
- [18] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Retrieved April 5, 2016. Available from http://www.R-project.org/.
- [19] W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson, *shiny: Web Application Framework for R*, 2016, Retrieved April 5, 2016. Available from http://CRAN.R-project.org/package=shiny.
- [20] H. Wickham, ggplot2: Elegant graphics for data analysis. Springer New York, 2009, http://had.co.nz/ggplot2/book.
- [21] D. Kahle and H. Wickham, "ggmap: Spatial visualization with ggplot2," *The R Journal*, vol. 5, no. 1, pp. 144–161, 2013, http: //journal.r-project.org/archive/2013-1/kahle-wickham.pdf.
- [22] D. B. Carr, R. J. Littlefield, W. Nicholson, and J. Littlefield, "Scatterplot matrix techniques for large n," *Journal of the American Statistical Association*, vol. 82, no. 398, pp. 424–436, 1987.
- [23] "Roma più radioattiva di tokyo, non c'è contaminazione," Retrieved April 5, 2016. Available from http://roma.corriere.it/roma/notizie/cronaca/11_marzo_16/ ambasciata-giappone-roma-piu-radioattiva-tokyo-190238152192. shtml?refresh_ce-cp, (Article in Italian).
- [24] G. Ostrouchov, W.-C. Chen, D. Schmidt, and P. Patel, "Programming with Big Data in R," 2012, Retrieved April 5, 2016. Available from http://r-pbd.org/.
- [25] J. Cheng and Y. Xie, Leaflet: Create interactive web maps with the JavaScript 'Leaflet' Library, 2015, R package version 1.0.0. Retrieved April 5, 2016. Available form http://CRAN.R-project.org/package= leaflet.