LAWRENCE LIVERMORE NATIONAL LABORATORY

LLNL-TR-841783

Analysis of Data Fusion between Waveform Events and Radionuclide Detections Reported in 2021 by the International Data Centre

M. G. Fernandez, D. D. Lucas, S. C. Myers

October 27, 2022

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Lawrence Livermore National Laboratory

Analysis of Data Fusion between Waveform Events and Radionuclide Detections Reported in 2021 by the International Data Centre

M. Giselle Fernández-Godino, Donald D. Lucas, and Stephen C. Myers

Lawrence Livermore National Laboratory,
P.O. Box 808, Livermore, CA 94551-0808

October 27, 2022

Disclaimer

Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344.

Table of Contents

Motivation

During this project, we assessed the percentage of events reported by the International Data Centre (IDC) that are candidates for evaluation using data fusion methods. We calculated the percentage of waveform (seismic, hydroacoustic, infrasound, SHI) events that may be associated with radionuclide (RN) detections and vice versa. For this task, data fusion means identifying RN detections that may be associated with SHI events.

The task of producing a fused SHI and RN bulletin is included in the IDC operational manual. As of 2022, the production of a fused event bulletin remains incomplete, largely because methods for producing the bulletin are under development. Nonetheless, simple methods for data fusion have been developed and preliminary software tools are available at the IDC to associate SHI events and RN detections. This report is a first look at the percentage of SHI events that may be candidates for fusion with RN detections or bulletins. One year of IDC bulletin data (2021) is considered in this report. It is important to note that there were no known nuclear explosions during this time period. Hence, we expect that events during this time period would generate SHI and RN signals. If many SHI events are candidates for fusion with RN data, greater priority should be placed on establishing screening procedures before and/or during the fusion process. By contrast, if the number of candidate events is diminishingly small, then data fusion is not a driver for further development of event or detection screening at the IDC.

LLNL reviewed the events reported in the Standard Event Bulletin (SEB) and determined which events are candidates for fusion with the Reviewed Radionuclide Reports (RRR). Much of the work involves establishing a metric or metrics for determining whether a Reviewed Event Bulletin (REB) event is a candidate for data fusion.

IDC Data Analysis

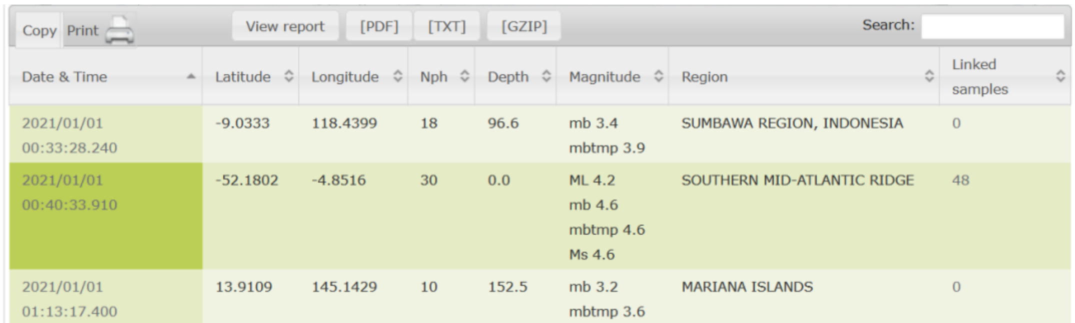
We start with SHI and RN bulletin data that IDC repots on their data fusion website as shown in Fig. 1 and Fig. 2. The SHI, radionuclide, and geographic datasets are described below, along with the postprocessing steps we have taken. The data for the analysis was obtained from the CTBTO's IDC Products > Fusion webpage[1], which implements a concept of data fusion between SEB SHI events and radionuclide detections. An SHI event is connected to a radionuclide sample if it overlaps with the relevant Field of Regard, which is calculated from backward atmospheric modeling. Two separate datasets are available from the IDC Products > Fusion webpage, one from the waveform side and another from the RN side, both of which are described below.

SHI Event Data Analysis

SHI event data for calendar year 2021was downloaded from the IDC fusion website in five-day intervals and concatenated together to form a single data set. A screenshot of the first few entries in January is shown below in Fig. 1. Each entry contains information about the date, time, and

[1] https://swp.ctbto.org/web/swp/fusionr (RN webpage), https://swp.ctbto.org/web/swp/fusionw (SHI webpage).

other characteristics of the event, as derived from International Monitoring System (IMS) SHI data. In contrast to RN detections (which are detections), SHI event locations pertain to a physical source. The rightmost column in Fig. 1 shows the number of RN detections linked (a.k.a. associated).



Copy Print View report [PDF] [TXT] [GZIP] Search:

Date & Time	Latitude	Longitude	Nph	Depth	Magnitude	Region	Linked samples
2021/01/01 00:33:28.240	-9.0333	118.4399	18	96.6	mb 3.4 mbtmp 3.9	SUMBAWA REGION, INDONESIA	0
2021/01/01 00:40:33.910	-52.1802	-4.8516	30	0.0	ML 4.2 mb 4.6 mbtmp 4.6 Ms 4.6	SOUTHERN MID-ATLANTIC RIDGE	48
2021/01/01 01:13:17.400	13.9109	145.1429	10	152.5	mb 3.2 mbtmp 3.6	MARIANA ISLANDS	0

Figure 1. Screenshot of the summary of SHI events from the IDC's fusion waveform webpage.

As illustrated by the first few entries in Fig. 1, there are many SHI events with zero linked RN detections. We, therefore, removed all SHI events without any linked samples. According to the reported data on the IDC webpage, out of a total of 38,037 SHI events recorded at 574 unique regions during the year 2021, 20,924 may be linked to an RN detection, which represents 55% of the data. If there is at least one link between an SHI event and RN detection, the number of SHI events that are linked to RN detections has a mean of 23.3, a median of 19.0, and a standard deviation of 19.3 (see Fig. 2 right). The number of resulting linked SHI events normalized by the number of seismic phases (SPh) have a mean of 2.5, a median of 1.6, and a standard deviation of 2.7 (see Fig. 2, left).

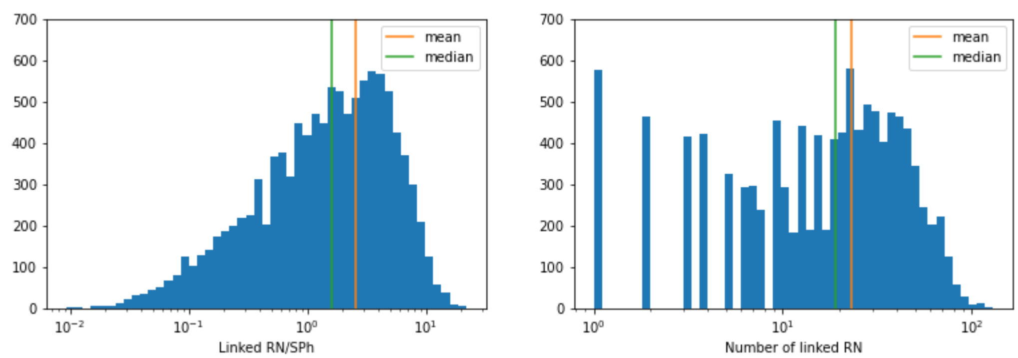


Figure 2. The number of RN/SPh (left subfigure) and of RN (right subfigure) detections linked with an SHI event. The orange line represents the mean, while the green represents the median.

Radionuclide Data Analysis

Similarly, RN detection data for the calendar year 2021 was downloaded from the IDC fusion webpage in ten-day intervals and concatenated to form a single data set. The first eight entries are shown in Fig. 3. Each entry in the dataset shows information about a single RN detection at an IMS station, including the station where the sample was collected, the sample ID, the collection period (12 or 24 hours), the level of detection, links to radiological reports for the sample (ARR and RRR), whether the sample was sent to a radiological laboratory for analysis

(RLR), and the number of SEB events that could be associated with the RN detection. The final column is the number of SHI events that fall within each sample's Field of Regard (FOR), up to 14 days before the sample collection. Even in the limited set of detections shown in Fig. 3, the JPX38 station has the most associated SEB events. JPX38 is located in Takasaki, Japan, which is in a seismically active region and downwind of many potential sources of background radionuclides.

Search: Copy Print

Station	SID	Cstart(GMT)	Cstop(GMT)	Category	Reports	RLR	SEB events
AUX04	5993847	2020/12/31 20:39	2021/01/01 08:39	Level C	ARR RRR		5
DEX33	6004160	2021/01/07 06:00	2021/01/08 06:00	Level C	ARR RRR		2
DEX33	5996940	2021/01/02 06:00	2021/01/03 06:00	Level C	ARR RRR		1
DEX33	5995495	2021/01/01 06:00	2021/01/02 06:00	Level C	ARR RRR		
JPX38	6003846	2021/01/07 18:42	2021/01/08 06:42	Level C	ARR RRR		122
JPX38	6003118	2021/01/07 06:42	2021/01/07 18:42	Level C	ARR RRR		143
JPX38	5994450	2021/01/01 06:42	2021/01/01 18:42	Level C	ARR RRR		57
JPX38	5993731	2020/12/31 18:42	2021/01/01 06:42	Level C	ARR RRR		60

Figure 3. Screenshot of the summary of radionuclide detections from the IDC's fusion RN webpage.

The first step for analyzing the RN detection data was to remove the entries without associated SEB events (blanks on the last column in Fig. 3). A total of 37,728 RN detections are included in the dataset from 32 certified and operational noble gas IMS stations and 72 certified particulate IMS stations around the globe throughout the year 2021. A total of 15,448 detections are linked with more than one SEB event (41% of the detections). Note that the atmospheric data was pulled from the IDC's fusion RN webpage using the 14-day option (not the 60-day option). That is, SHI events up to 14 days before the RN detection were considered when reporting the linked SEB events. The number of SHI events linked to an RN detection (excluding detections with no linked SHI events) has a mean of 27, a median of 7, and a standard deviation of 53 (see Fig. 4).

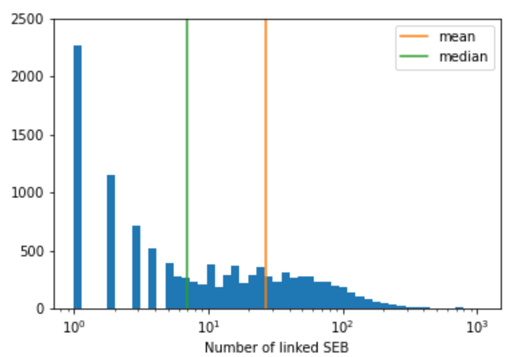


Figure 4. The number of SHI events linked with a RN detections. The orange line represents the mean while the green represents the median.

Naïve Data Fusion Analysis

This section describes a naive algorithm to determine which of the SHI events and RN detections reported by the IDC may be candidates for data fusion. For this analysis, we only considered events and detections that the IDC reported with non-zero/non-black linked events. That is 20,924 SHI events and 15,448 Rn detections. We first identified the linked SHI events and RN detections that had the largest number of radionuclide links (last column in Fig. 1) and the number of SEB events (last column in Fig. 3). The cutoff for the seismic data was to consider linked events such that RN/SPh >10 (RN/SPh is the ratio between "Linked samples" and "Nph" columns in Fig. 1). This represents 0.8% of the SHI events dataset. The cutoff for the atmospheric data is SEB events > 168 ("SEB events" is the last column in Fig. 3). This also was chosen to represent 0.8% of the RN detections dataset. According to our naïve algorithm, RN detections and SHI events are considered linked if they are within 10 degrees latitude, 10 degrees longitude and the SHI events occurred up to 14 days before the RN detections. The longitude and latitude are included in the SHI event data (columns "Longitude" and "Latitude" in Fig. 1). The longitude and latitude for the RN detections data represent the location of the station (column "Station" in Fig. 3), which was retrieved from the IDC webpage as well. For the SHI event time we used the "Date & Time" column in Fig. 1. For RN detection time, we considered the one reported in the "Cstart" column in Fig. 3.

SHI Events Associated with RN Detections

Using the naïve algorithm described above, for each SHI event, we looked at the RN dataset to see if we could find a station within 10 degrees in latitude and longitude that happened at the time up to 14 days after the seismic detection. For all the RN detections found using this criterium, we computed the total number of "SEB events" (last column Fig. 3) associated with the found detections. Using this approach, we found that from the total 38,037 SHI events that the IDC reported, 11,037, 29% could be associated with an RN detection.

Radionuclide Detections Associated with SHI Events

For the first analysis, we use the latitude, longitude and time of each of the considered SHI events For each SHI event, we looked at the RN dataset to see if we can find a station within 10 degrees in latitude and longitude that happened at the time up to 14 days after the seismic detection. For all the RN detections found using these criteria, we computed the total number of "SEB events" (last column Fig. 3) associated with the found detections. Using this approach, we found that from the total 37,728 RN detections that the IDC reported, 11,144, 30% could be associated with an SHI event.

Geographic Trends in SHI Event and RN Detection Associations

There are certain geographic regions where associations between SHI events and RN detections are particularly strong during the year 2021. Table 1 summarizes the top five stations in the number of SEB event and RN detection links found by our algorithm based on the data reported by the IDC. As mentioned before, JPX38 (East Pacific Rim) leads the board with an order of magnitude more RN detection links and six times more SEB events. Recognizing geographic locations with a high number of SHI events and human-made RN activity (see Fig. 5) is crucial because they represent a challenge for the data fusion process. In these regions, it would be difficult to separate the background from a real nuclear event if it were to happen unless the SHI event or RN detection are unambiguously characteristic of a nuclear explosion. For example, in a nuclear explosion, isotope ratios may preclude a release from a nuclear reactor or isotope facility.

Region	**Station**	**Linked events**	**SEB events**
East Pacific Rim	JPX38	16506	5789
Central North America	USX74	1262	1223
Southern Mid-Atlantic Ocean	GBP68/GBX68	73/1037	408/4306
East Europe	SEX63	132	204

Table 1. The top five strongest links between SHI events and RN detections according to our naïve algorithm and based on the IDC-reported associations.

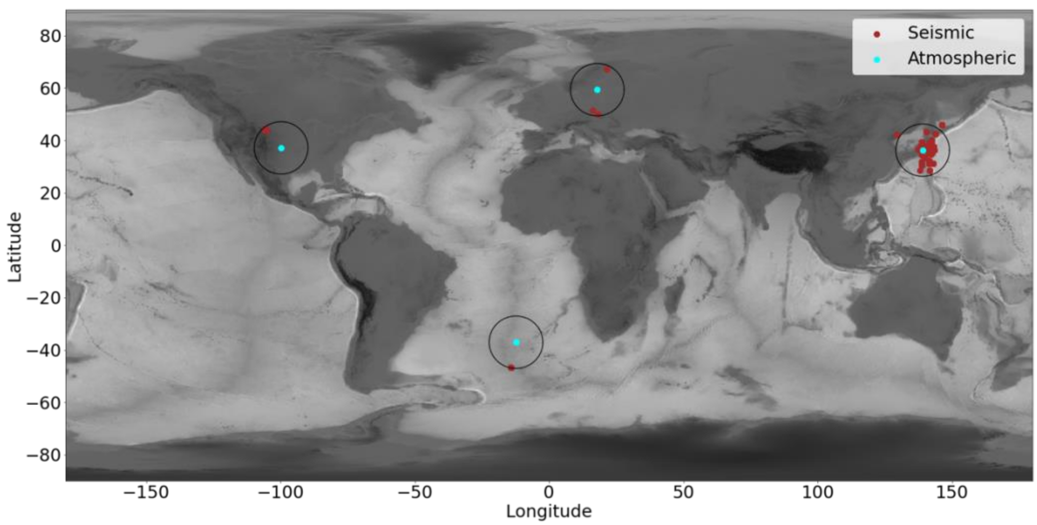


Figure 5. The geographical location of the top five links between RN and SHI events was calculated using our naïve algorithm and based on the IDC-reported associations. In cyan are shown the RN detections, and in red, the SHI events.

Summary and Recommendations

Based on results for 2021listed on the IDC data fusion website, 55% of SHI events reported in the SEB may be linked to an RN detection. The number of RN detections linked to a SEB event has a mean of 23.3, a median of 19.0, and a standard deviation of 19.3 (see Fig. 2 right). Similarly, for the year 2021, the IDC data fusion tool reported that 41% of RN detections might be linked to a SEB event. The number of SEB events linked to an RN detection has a mean of 27, a median of 7, and a standard deviation of 53 (see Fig.4).

In 2021 there are no records of a nuclear test, yet the IDC data fusion webpage reports 40% of RN detections may be linked to an SEB event through the field of regard. Since all of these potential links are almost certainly false associations, it is clear that more efforts should be put into event and detection screening, either prior to or during the data fusion process. In the meantime, the exceedingly high probability of false association between SHI events and RN detections should be understood by scientists and policymakers who examine the output of the data fusion webpage.

A naïve algorithm based on event/detection proximity in time and space was used to compare the two datasets (SHI event dataset and RN detection dataset). The algorithm was successful in reporting that 30% of the SHI events may be linked with RN detections and that approximately 30% of the RN detections may be linked with SHI events. But the algorithm consistency for the overall calculation deteriorates if we look closer solely into regions with high seismic and nuclear activity (see Table 1).

Accurate association of SHI events and RN detections can be particularly challenging in certain geographic regions due to a combination of high seismic activity and the presence of numerous nuclear reactors and isotope facilities. East Pacific Rim, East Europe, Central North America, and the Southern Mid-Atlantic Ocean may be particularly problematic regions (Fig. 5). The large number of high false positive associations in these areas will hamper efforts to properly associate SHI events and RN detections, complicating the correct identification of nuclear tests. Further analysis has been performed for this report and can be found summarized in the figures in the Appendix.

Appendix

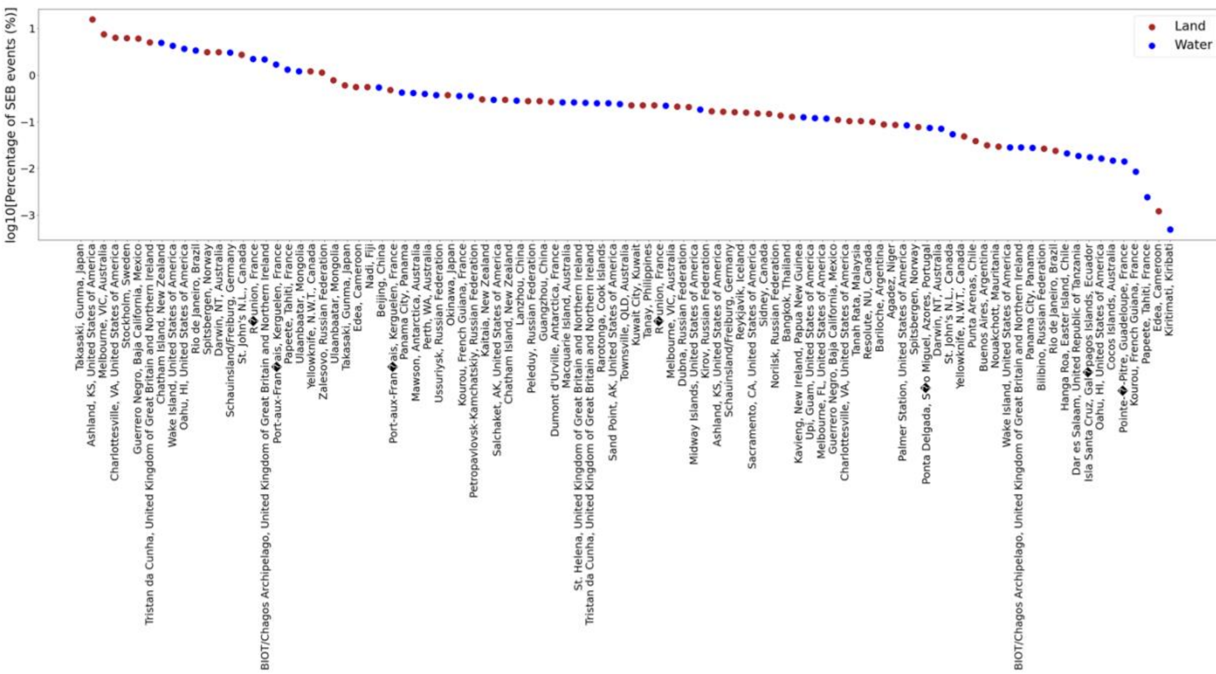


Figure A1. Percentage of SHI events linked with RN data around the world for the year 2021, according to the IDC RN webpage. Base 10 logarithm was applied to the vertical axis for visualization purposes.

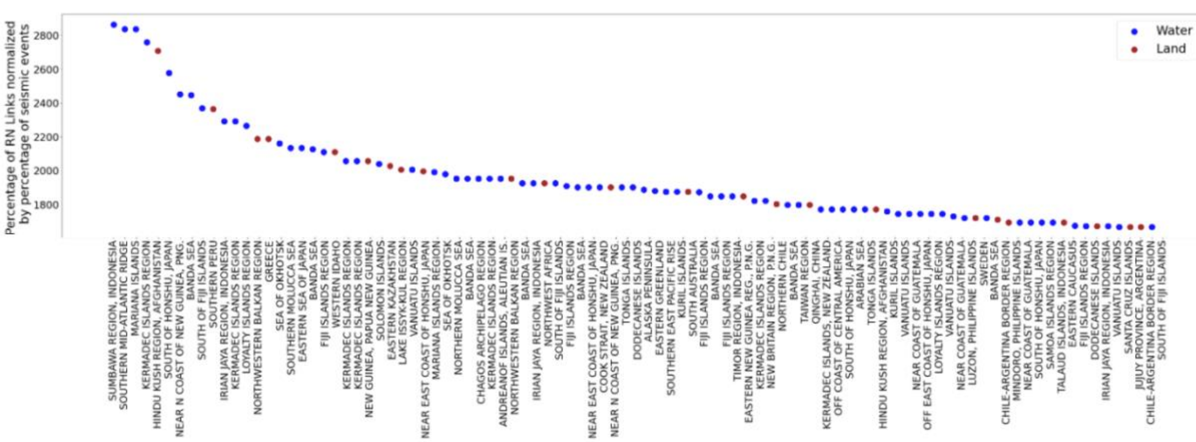


Figure A2. Percentage of RN links normalized by the percentage of SHI events around the world for the year 202,1 according to the IDC seismic webpage.

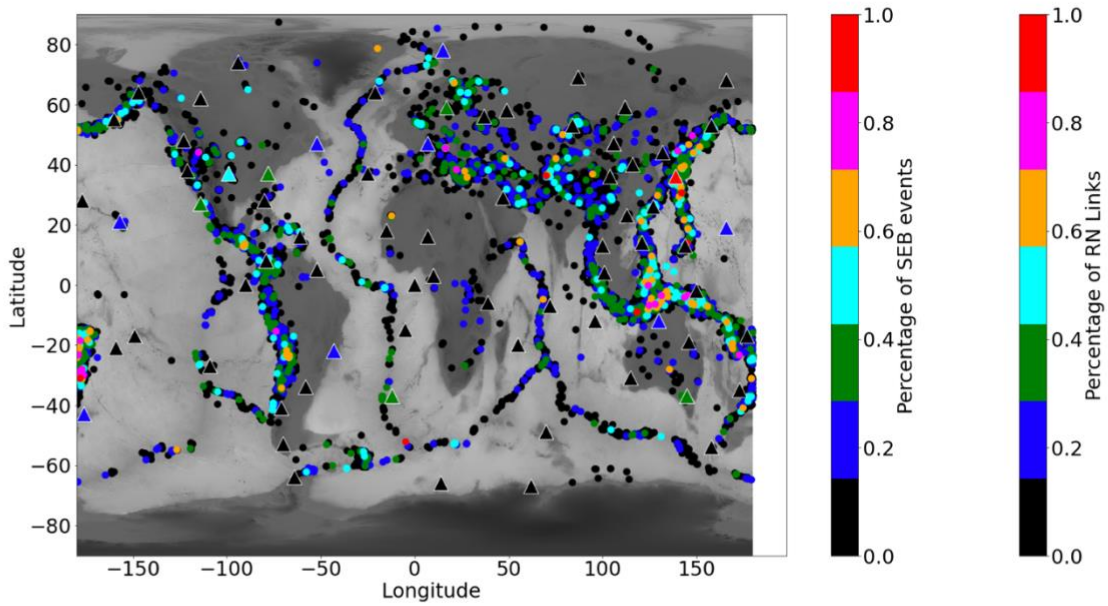


Figure A3. Percentage of RN links (circles) and seismic links (triangles) around the world reported by the IDC webpages for the year 2021.

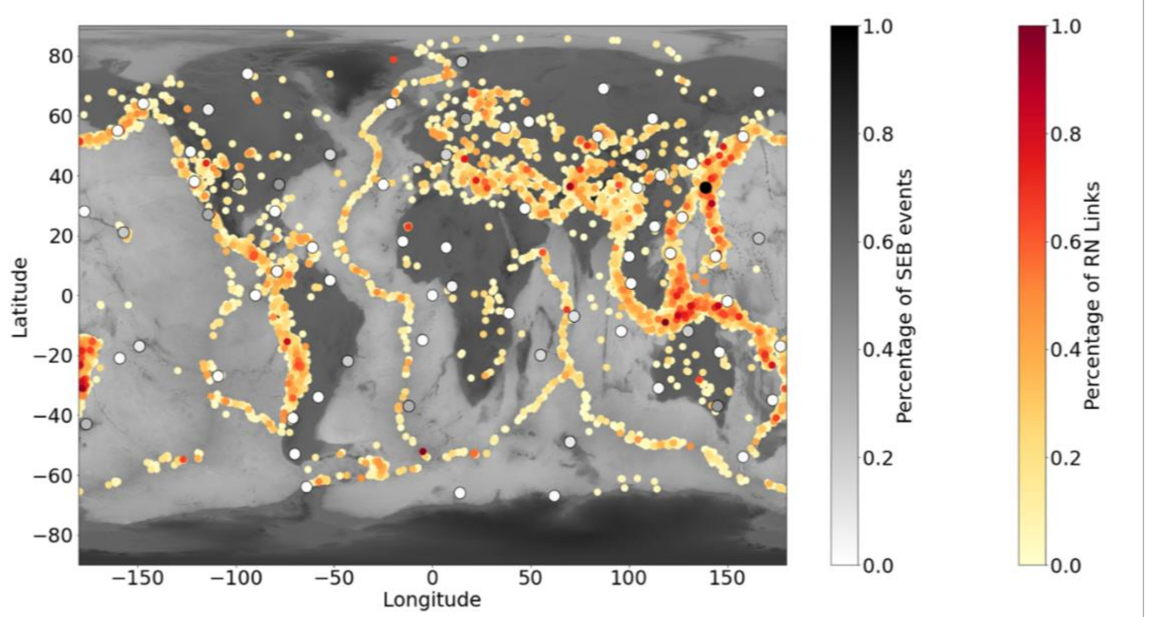


Figure A4. Percentage of RN links (yellow-orange-red scale) and seismic links (grayscale) around the world reported by the IDC webpages for the year 2021.