Check for updates

USGS "Did You Feel It?"—Science and Lessons From 20 Years of Citizen Science-Based Macroseismology

Vincent Quitoriano and David J. Wald*

National Earthquake Information Center, U.S. Geological Survey, Golden, CO, United States

The U.S. Geological Survey (USGS) "Did You Feel It?" (DYFI) system is an automatic method for rapidly collecting macroseismic intensity (MI) data from internet users' shaking and damage reports and for generating intensity maps immediately following felt earthquakes. DYFI has been in operation for nearly two decades (1999–2019) in the United States, and for nearly 15 years globally. During that period, the amount of data collected is astounding: Over 5 million individual DYFI intensity reports— spanning all magnitude and distance ranges—have been amassed and archived. DYFI allows for macroseismic data collection at rates and quantities never before imagined, and thus high-quality MI maps can be made almost immediately, and with more complete coverage at higher resolution than in the past. DYFI also allows for valuable positive interactions of the citizenry with a Federal science agency. In essence, the widespread adoption of DYFI – along with ShakeMap—has facilitated the general acceptance of the very concept of shaking intensity, fundamentally improving our agency's ability to communicate both hazard and risk to the population. DYFI effectively confirms the importance of reporting and inculcating the public's understanding of intensity – in addition to magnitude – for a proper perspective of earthquake risk-related decision-making. Furthermore, the vast amount of DYFI data allows for data-rich analyses of otherwise intractable seismological, sociological, and earthquake impact studies, such as quantifying the shaking due to induced earthquakes, human response and risk perception, relating recorded shaking metrics to macroseismic effects, and the attenuation of intensity with magnitude and distance. Naturally, web-based data collection also poses challenges. After two decades of experience acquiring data with the DYFI system, we address some of these challenges by documenting refinements to our algorithmic and operational procedures that have evolved over that time. Lastly, we outline new opportune research and development directions for our DYFI approach to citizen seismology.

Keywords: citizen science, seismology, science communication, seismic hazard, open data, macroseismology, earthquake intensity, ShakeMap

INTRODUCTION

An impressive, rapid evolution has taken place in the realm of macroseismic intensity (MI) data collection and assignment since the revision of the European Macroseismic Scale of 1998 (EMS-98; Grünthal, 1998), wherein well-defined building vulnerability classes combined with damage matrices facilitated reliable MI assignments, particularly at high intensities. A more recent revolution in the field began with web-based macroseismic surveys and assignments following Dengler and Dewey (1998) and Wald et al. (1999a) followed by parallel developments in Italy by Sbarra et al. (2010). At the same time, internet-based access to reconnaissance photos and media accounts significantly improved the availability of testimonials and images (as well as their locations) of earthquake effects for analysis. The latest innovation is to reduce data collection to convenient-to-use cartoons that readily allow for a user-selected MI value among a choice of intensity levels visually depicted, as in the LastQuake mobile application by the European-Mediterranean Seismological Centre (EMSC; Bossu et al., 2017).

Over the past two decades, the U.S. Geological Survey (USGS) has relied on the "Did You Feel It?" (DYFI) portal (Wald et al., 2011) to collect shaking and damage reports from internet users immediately following felt events, effectively scaling back and deemphasizing more traditional postal questionnaires and reconnaissance surveys. The USGS has been operating DYFI since 1999 in California, since 2000 for the rest of the United States, and since 2004 globally. DYFI is essential for systematically collecting macroseismic data for all felt seismic events in the United States and has become one of the most popular interactive web sites within the United States Government.

For earthquakes outside the United States, DYFI data rapidly signal or confirm earthquake occurrence for seismic analysts and scientists at the USGS National Earthquake Information Center (NEIC), giving a quick indication of the extent and severity of shaking effects. Intensity data from DYFI are automatically used to provide valuable shaking constraints for the USGS Global ShakeMap system (Wald et al., 1999b), which in turn is the fundamental hazard input for the USGS Prompt Assessment of Global Earthquakes for Response (PAGER; Wald et al., 2008) system that allows the USGS to alert agencies and users around the world of significant earthquakes and their likely impacts.

The data collection and assignment of DYFI-based intensity depart from traditional expert-assigned intensities (e.g., Musson et al., 2010) but they are made more rapidly; provide better coverage and at higher spatial resolutions; and allow citizen input and interaction. A widely felt earthquake near a populated area can provide thousands or tens of thousands of independent observations over a wide geographic extent, far more than can be collected from traditional assessments.

This paper provides an overview of the DYFI system – after 20 years of experience – with emphasis on the citizen science-based macroseismic data that we have collected as well as the resulting research that those data have allowed or facilitated. We first provide background on the current system and processing software, which has been recently reengineered and made open source. We then focus on data collection and how quality assurance is maintained given the nature of internet-based data contributors. Next, we present examples of unique studies that employ DYFI data in both the seismological and social science realms; in particular, the general adoption of MI as a metric for communicating hazard and risk is emphasized since intensity is such a vastly more useful descriptor than earthquake magnitude alone. Lastly, we describe challenges and limitations of DYFI and suggest both potential solutions and new directions that will facilitate even more widespread adoption of DYFI as a citizen-science portal for both societal benefits and scientific advancements.

THE DYFI SYSTEM

The DYFI software package is fully open source, written in Python, and available publicly through GitHub since 2018[1]. Incoming entries from multiple web servers are processed and aggregated over postal ZIP codes (in the United States) and 1-km and 10-km aggregated boxes for every earthquake. These data are used to make interactive maps and plots (e.g., **Figures 1–5**) served via the USGS Earthquake Program web pages[2].

One of the key procedures of the DYFI process is the aggregation of responses within compact spatial areas. Aggregation allows us to combine the observations of many users and fill in the gaps in relevant intensity markers. For example, one contributor might observe objects falling off shelves (the *shelf* index) but have no pictures hanging on their wall. Another nearby might have no objects on shelves but report pictures falling off their wall (the *picture* index). The two relevant questionnaire indices are combined (not averaged) with other users in their community to produce an intensity calculation. For details, see the section below.

Quality Assurance

"Did You Feel It?" maps and products have been updated occasionally, sometimes systematically, over the last two decades of operation. DYFI maps are aggregated models of MI that change over time. Of course, users often contribute data for months or longer after an earthquake. More importantly, while the intensity calculation of Wald et al. (2011) has not changed, operational and postprocessing procedures have improved over the years. DYFI maps and products are expected to change and improve over time. Many DYFI improvements and quality control strategies have been implemented incrementally but have been refined and standardized in the new code base. We now systematically track and document all code changes on GitHub, which preserves all versions of the code and is publicly available online. Previously, changes were done on a more *ad hoc* basis, generally only to improve or fix operational issues.

"Did You Feel It?" originally defined the communities used for aggregation as ZIP codes and cities (outside the United States).

---

[1] https://code.usgs.gov/ghsc/esi/dyfi
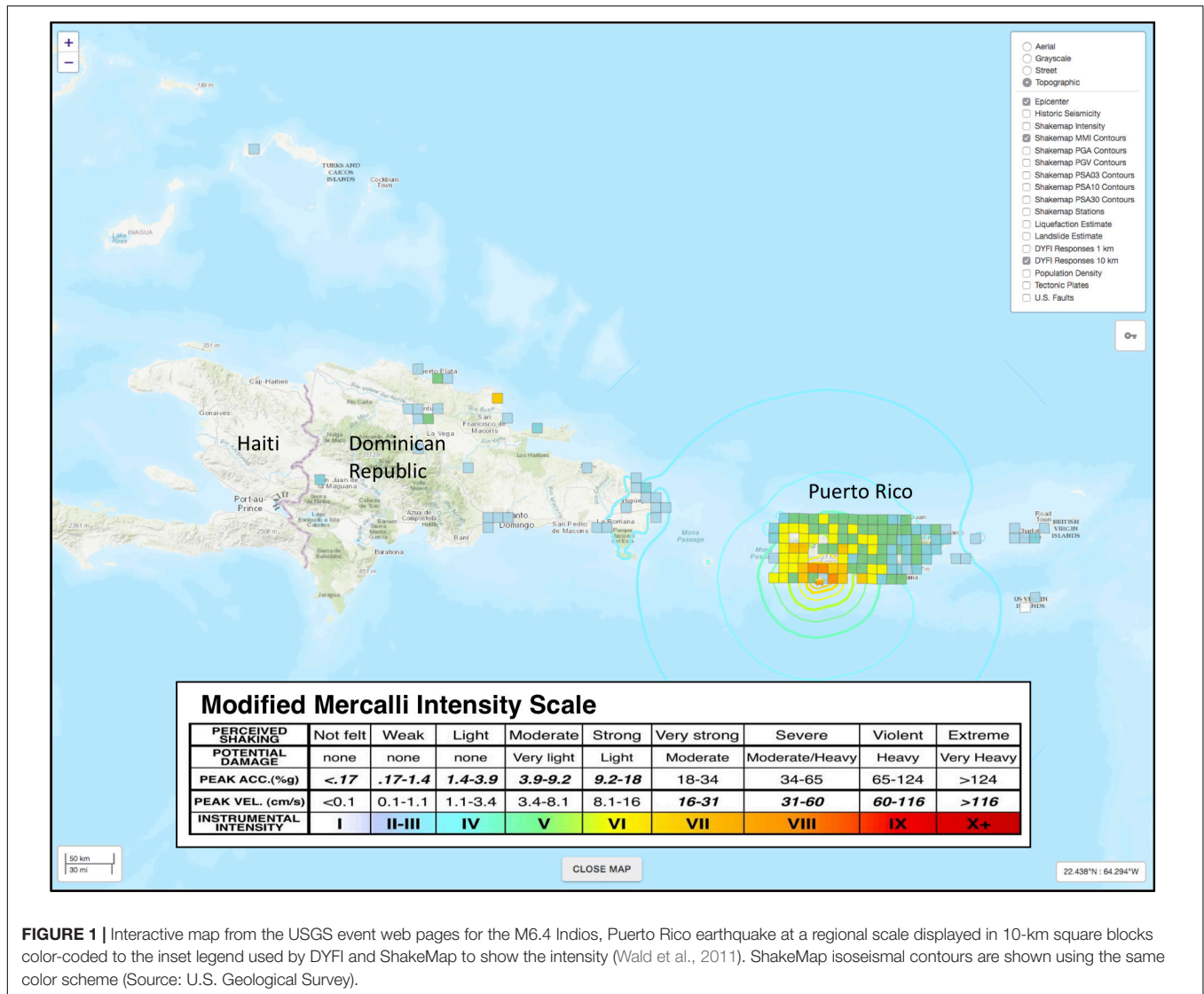
[2] https://earthquake.usgs.gov

FIGURE 1 | Interactive map from the USGS event web pages for the M6.4 Indios, Puerto Rico earthquake at a regional scale displayed in 10-km square blocks color-coded to the inset legend used by DYFI and ShakeMap to show the intensity (Wald et al., 2011). ShakeMap isoseismal contours are shown using the same color scheme (Source: U.S. Geological Survey).

This was problematic as some ZIP codes are much larger in area than others, and cities are not evenly distributed; in addition, both ZIP code boundaries and city names change with time. In order to standardize aggregation sizes, we have now switched to a system based on the Universal Transverse Mercator (UTM) Geographic Grid. DYFI now automatically geolocates each user response (down to the level of street address in most cases), aggregates questionnaire responses using UTM coordinates to define 1-km and 10-km blocks, and computes intensities using the responses within each box. We find that sparsely felt events benefit from using 10-km blocks to combine more responses for each intensity calculation. Alternatively, for events near population centers, 1-km blocks allow us to show fine variations in the felt intensities. Maps and datafiles for both aggregations are produced for every DYFI event. UTM blocks do not change with time, which makes the comparison of earthquake data at different times much less complicated. We encourage researchers to use geocoded UTM- aggregated

datasets instead of older ZIP code aggregated datasets because of these reasons.

A note of caution: We have redone the geocoding of all DYFI entries using modern, online geocoding services. While geocoding tends to yield consistent results, they are not perfect. Some observer entries likely have been moved, added, or removed from their original aggregations in the intervening years as online geocoding services have improved.

We have updated the automatic removal of outlier intensities from the DYFI dataset. Each event is assigned a region-dependent intensity prediction equation (IPE) that is a function of magnitude and distance from the epicenter. Each geocoded block, ZIP code, or city with a computed intensity more than a certain threshold away from the expected intensity is flagged as an outlier and not included in DYFI products. The current filtering threshold is 3 intensity units above or below the value expected from the IPE. For significant events, we sometimes manually flag entries which are obvious outliers,
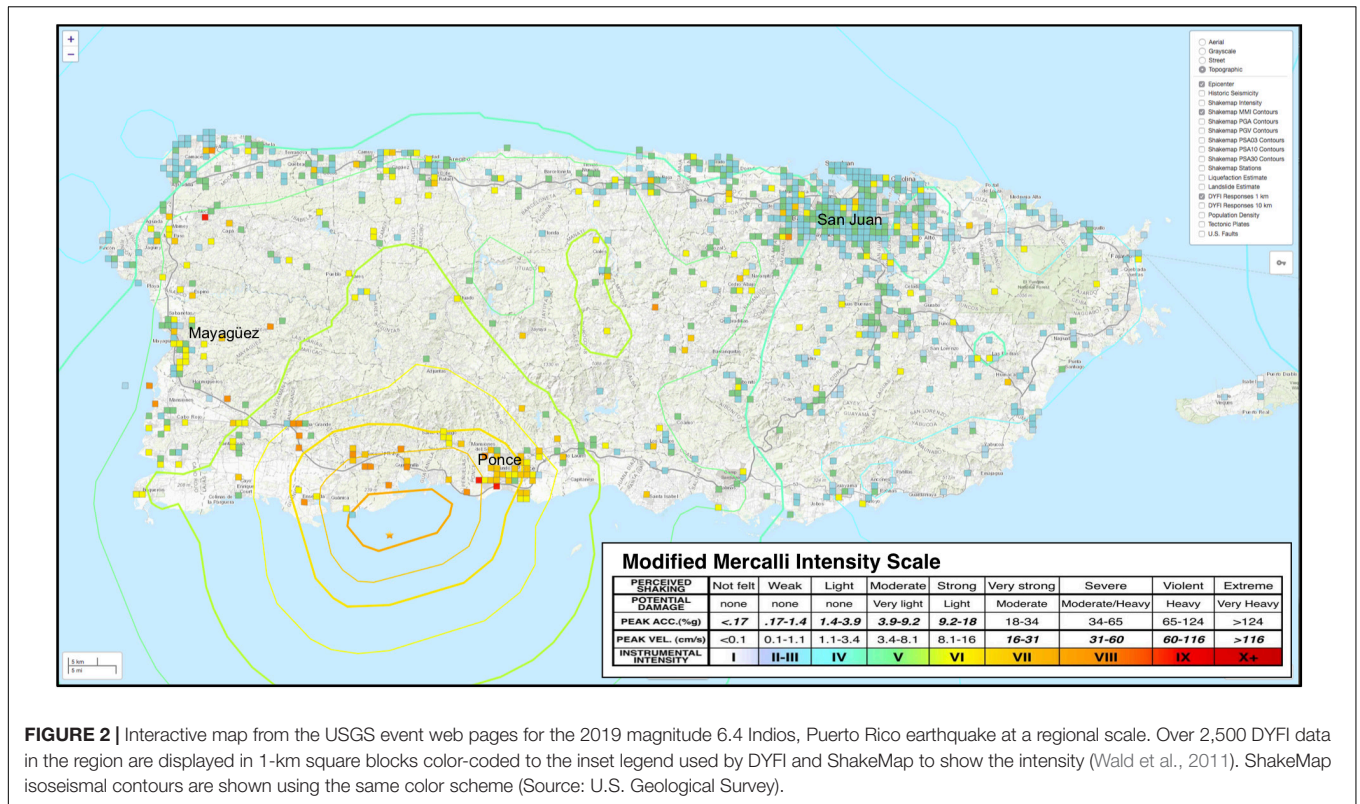
FIGURE 2 | Interactive map from the USGS event web pages for the 2019 magnitude 6.4 Indios, Puerto Rico earthquake at a regional scale. Over 2,500 DYFI data in the region are displayed in 1-km square blocks color-coded to the inset legend used by DYFI and ShakeMap to show the intensity (Wald et al., 2011). ShakeMap isoseismal contours are shown using the same color scheme (Source: U.S. Geological Survey).
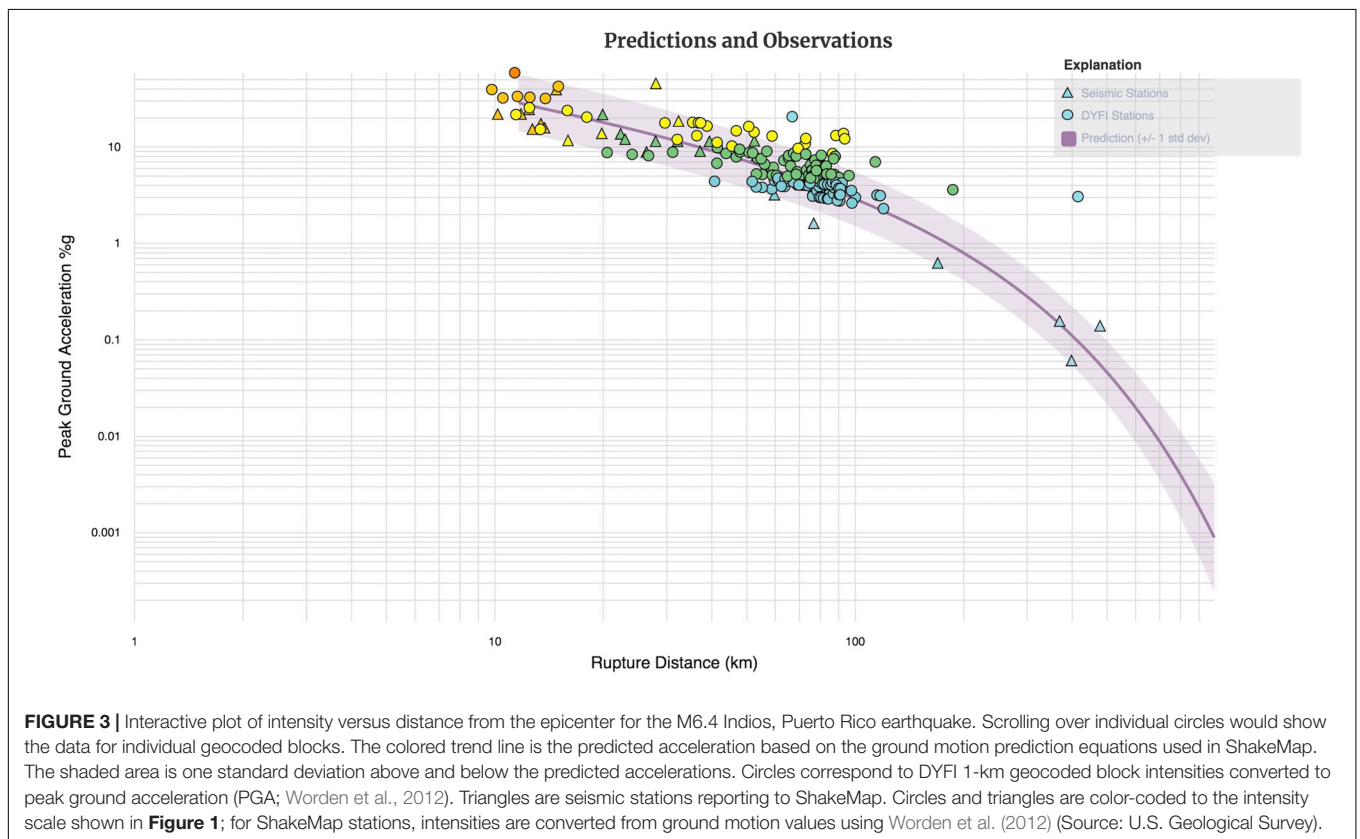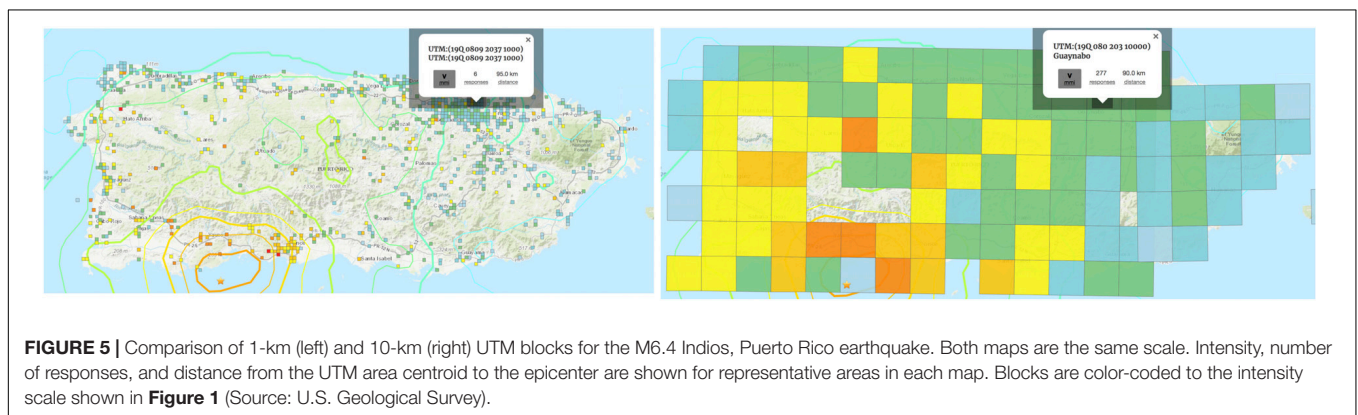

FIGURE 3 | Interactive plot of intensity versus distance from the epicenter for the M6.4 Indios, Puerto Rico earthquake. Scrolling over individual circles would show the data for individual geocoded blocks. The colored trend line is the predicted acceleration based on the ground motion prediction equations used in ShakeMap. The shaded area is one standard deviation above and below the predicted accelerations. Circles correspond to DYFI 1-km geocoded block intensities converted to peak ground acceleration (PGA; Worden et al., 2012). Triangles are seismic stations reporting to ShakeMap. Circles and triangles are color-coded to the intensity scale shown in **Figure 1**; for ShakeMap stations, intensities are converted from ground motion values using Worden et al. (2012) (Source: U.S. Geological Survey).

FIGURE 4 | Response rate of DYFI report entries for the M6.4 Indios, Puerto Rico earthquake. For example, scrolling over the curve would show the number of responses at different times. Shown here is the number of responses submitted up to 1 h after the origin time, when nearly 1,200 of 2,200 responses were received (Source: U.S. Geological Survey).

FIGURE 5 | Comparison of 1-km (left) and 10-km (right) UTM blocks for the M6.4 Indios, Puerto Rico earthquake. Both maps are the same scale. Intensity, number of responses, and distance from the UTM area centroid to the epicenter are shown for representative areas in each map. Blocks are color-coded to the intensity scale shown in **Figure 1** (Source: U.S. Geological Survey).

duplicates, or spurious responses. This usually happens in the days after an event, but occasionally we receive requests to manually check data at a much later date by detail-oriented users.

The problem of associating user responses to the correct event can be complicated during an earthquake sequence with multiple foreshocks and/or aftershocks. DYFI contributors tend to select the most recent earthquake displayed on the USGS website, which might not be the event that corresponds to their observations; or to the "Unknown Event Form," which has no associated event. Whenever DYFI processes an event, it checks for unassociated entries and for other entries that are likely to be associated to that event. We are still examining various ways of improving this process and sometimes resort to manually disentangling entries if possible; but often, little distinguishes between mainshock and aftershock reports for lower intensity observations.

Viewing DYFI products from a particular moment in time is sometimes useful, for example, at a certain period after an earthquake, to compare different versions, or to see the evolution of the DYFI map days after a significant event.

In the past, we replicated these "snapshots" by rerunning the DYFI process on subsets of DYFI entries within the desired timeframe. Recently, the development of the Advanced National Seismic System's Comprehensive Earthquake Catalog (ComCat) has enabled the storage and retrieval of USGS real-time products (Guy et al., 2015). ComCat archives all versions of the products that were sent and displayed online, so previous versions of DYFI products in ComCat can now be accessed easily for comparison.

An additional point about the DYFI intensity algorithm is warranted: Intensity is computed for a consensus (or numerical average) separately for each question in the questionnaire; each question can be answered by more or fewer observers. The consensus values are then weighted and summed to compute the intensity (see Wald et al., 2011, for details). Entries that do not answer a particular question (as opposed to answering "None") do not count for the corresponding index in the computation. For example, let us take a hypothetical community of several observers who did not answer the *picture* index ("Did pictures on walls move or get knocked askew?"), perhaps because they were in rooms with no pictures. In this case, the addition

of one observer who does answer positively will represent an N of 1 for the *picture* index, and the answer for the whole community for that index switches from 0 to 1. Thus, a single added or corrected entry could change the intensity much more substantially than would be expected from simply averaging intensity scores.

Data Sampling Bias

Unlike traditional postal MI questionnaires, DYFI and other internet collection systems are self-selecting. First, coverage is dependent on population and internet access, so areas with dense populations and adequate internet access are overrepresented compared to sparsely populated regions or populations that lack internet access (see, for example, Montalvo-Arrieta et al., 2019). Second, responses are overwhelmingly from users who felt an event. Less than 3% of DYFI responses are from contributors who respond with "not felt." These biases may be mitigated by a statistical approach. Mak and Schorlemmer (2016) model the reliability of DYFI data based on population density and socioeconomic parameters. Tosi et al. (2015) and Boatwright and Phillips (2017) propose methods of combining reporting and non-reporting communities to improve intensity estimates at the lower intensity range.

Another method of reducing non-reporting bias is to reach out to potential contributors. The Istituto Nazionale di Geofisica e Vulcanologia (INGV, Italy) invites users to pre-register on their website then notifies them of earthquakes in their area (Tosi et al., 2015). EMSC benefits from similar notifications using their mobile app. Linking DYFI to third party applications such as the MyShake Early Warning Platform (Allen et al., 2020) would allow us to solicit contributions from registered users near an identified or suspected event.

Effect of Observer Conditions

Observer location, building type, and situation all affect intensity reporting (e.g., Sbarra et al., 2014). Though Sbarra et al. (2014) report a variance of about 0.6 intensity units from such effects, the DYFI questionnaire requests those data but does not use them to systematically correct intensity assignments. Rather, we assume that, with sufficient numbers of observers, such details are averaged out. The tradeoff considered is between (potentially) more precise measurements and additional required questions for each user.

Network Performance

In the minutes after a widely felt earthquake near a populated region, the biggest challenge for DYFI is to accommodate the immediate deluge of web traffic and input data. The sheer number of responses after a significantly felt earthquake puts unprecedented stress on the performance of the USGS internet download and web capacity. As DYFI has grown in popularity, the USGS has radically improved capacity via both hardware and software improvements in order to handle the spike in internet traffic following such events, such as running multiple servers and containerized processing. For the largest events, even these efforts may be insufficient. The collection and storage of incoming user responses are done separately from the backend processes

of aggregating data, computing intensities, and creating DYFI products. Thus, raw data are collected and stored safely even during extreme processing loads. While new responses come in continuously, the backend processes only refresh the online maps every 5 min to avoid processing and network overload.
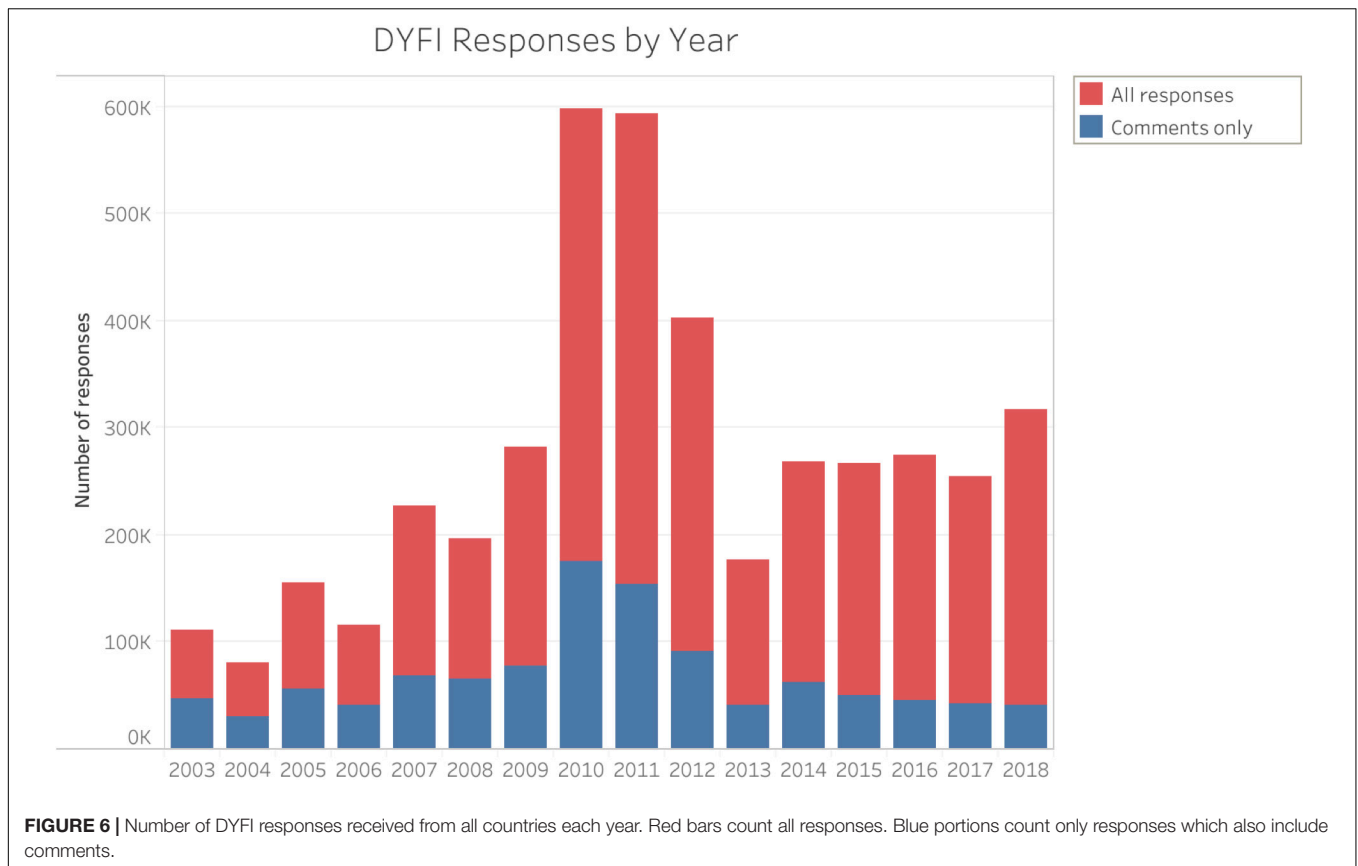
DYFI DATA

A recent example of the nature of DYFI data can be visualized for the 2019 magnitude 6.4 Indios, Puerto Rico earthquake. **Figure 1** shows over 2,500 DYFI responses in the region aggregated in 10-km square blocks. **Figure 2** is zoomed into the area of the island of Puerto Rico, and switches to geocoded data aggregated into 1-km blocks. Also shown are ShakeMap intensity contours for comparison. In fact, it can be seen that the DYFI data, which are much more numerous than the seismic stations used in ShakeMap (see **Figure 3**), play an important role in constraining the ShakeMap intensity contour pattern in **Figures 1**, **2** (see section "Integration into ShakeMap," below). **Figure 3** also provides an indication of the variability of the DYFI data (circles) and their distribution with distance from the epicenter at a regional scale in comparison to the recorded accelerations (triangles). In general, the intensity data fit estimates of ground acceleration quite well, despite being converted from intensity to acceleration (Worden et al., 2012).

Figure 4 provides the rate of DYFI report entries for the Indios, Puerto Rico earthquake: Nearly 1,200 of the total 2,200 responses were submitted in the first hour after the origin time. **Figure 5** provides a comparison of the 1-km and 10-km aggregated DYFI data. As with the other interactive maps, the intensity value, number of responses, and the geocoded box location are easily accessible via mouseover.

More general statistics further attest to the growth of internet-based macroseismic data collection. When it first went online in 2003, DYFI received about 110,000 responses, primarily from California earthquakes. Since then, over five million entries have been amassed over two decades. Currently, 64 events have more than 10,000 responses, and 550 events have over 1,000 responses. In 2018 (which we consider a typical year), more than 300,000 entries were received for 4,500 earthquakes (**Figure 6**). The year with the highest absolute number of responses was 2010, with nearly 600,000; for this time period, the largest impact event was the April 4, 2010, M7.2 Baja California earthquake, with nearly 80,000 responses.

The highest number of responses for a single earthquake is more than 146,000, for the 2011 M5.8 Mineral, Virginia event, which was felt by more Americans than any other in history. Response rates reached 62,000 submissions per hour (more than 1,000 per minute). For the 2014 M6.0 South Napa, California earthquake, 26,000 were received within the first hour and a total of 44,000 were ultimately received. Typically, about 60–90 percent of the entries are received within the first hour of an earthquake. For the largest earthquakes, response times peak at more than 30 responses per second.

The growth of DYFI contributions is accompanied by an evolution of MI reporting through the years. In 2003, 42%

FIGURE 6 | Number of DYFI responses received from all countries each year. Red bars count all responses. Blue portions count only responses which also include comments.

of contributors left a comment in addition to filling out the checkbox portion of the questionnaire. These comments have proven useful to social scientists exploring people's responses to earthquakes (see below). Since then, the number of responses that include comments has fallen to 13% of all responses, even as the absolute number of responses with comments has been relatively steady at roughly 50,000 per year. We attribute this phenomenon to the transformation of internet access from desktop computers toward mobile phones, tablets, and other portable devices and the ubiquity of social media outlets for reporting human experiences.

With two decades of DYFI reports, we can now map out the maximum MI of shaking reported over the entire United States during that time period (**Figure 7**), and nearly every felt earthquake in the United States is or can now be reported. Thus, this map represents the actual distribution of reported shaking intensity over the entire nation for nearly two decades, up through 2018. **Figure 7** depicts several easily recognizable seismological observables: First, most states experience some shaking over this time scale. Second, the pattern of shaking reflects many of the general trends of the USGS Probabilistic Seismic Hazard Assessment maps (PSHA, e.g., Petersen et al., 2020), but many of the intensity reports mapped in the Central United States are dominated by induced earthquakes, which were not explicitly considered in the 2014 or 2018 PSHA assessment. Lastly, felt areas are significantly larger for Central and Eastern events than those in the

West, a well-documented difference in crustal attenuation (e.g., Atkinson and Wald, 2007).

EARTHQUAKE RESPONSE AND SCIENCE USING DYFI DATA

Many of the earlier studies using DYI data were summarized by Wald et al. (2011). Here, we provide a partial summary of subsequent analyses.

Integration Into ShakeMap

Since its inception, ShakeMap has used DYFI intensity observations as proxies for ground motion data in areas with sparse instrumental coverage. The newest version of ShakeMap takes into account the uncertainties of its various inputs using a conditional multivariate normal (MVN) distribution (Worden et al., 2018) in order to combine data from different sources. Macroseismic intensities derived from DYFI observations have an intrinsic variability as a function of the number of responses (Worden et al., 2012). DYFI now automatically computes this uncertainty value as part of its product suite as input to ShakeMap. In addition, we are implementing the technique of Worden et al. (2012) to determine uncertainty functions for other macroseismic collection programs such as EMSC in order to systematize the inclusion of their data into ShakeMap. In other countries,
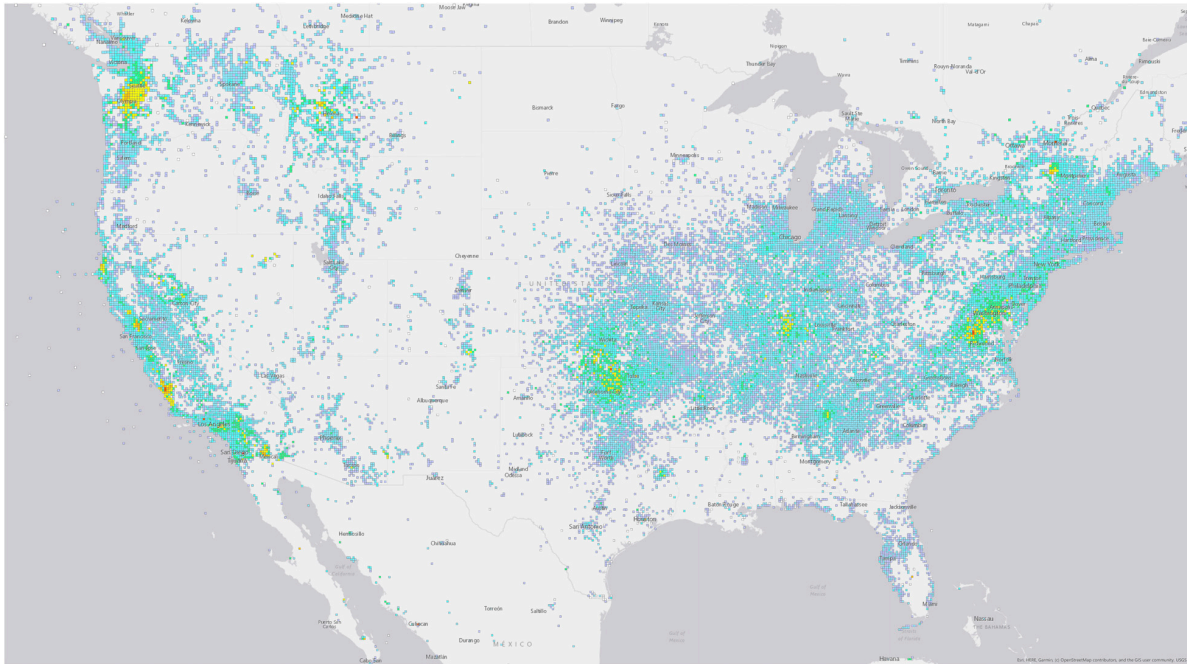
FIGURE 7 | Cumulative maximum DYFI intensities, 10-km geocoded blocks, in the U.S. from 1999 through 2018. Each cell is a 10-km block. Color corresponds to the highest intensity reported in that block among all events in that timespan using the intensity scale shown in **Figure 1**. This map and intensity maps for each year are available online in the Summary Maps section of the DYFI website (Source: U.S. Geological Survey).

DYFI systems or their equivalents are routinely ingested as part of their ShakeMap production; for instance, in Australia (Allen et al., 2019), northwestern Europe (Van Noten et al., 2016), and the French Overseas Territories (A. Schlupp, written communication, 2017).

The ability to incorporate macroseismic data geospatially into ShakeMap serves a very useful purpose. Recall that one attribute of macroseimic data is the connection of the present to the past. Historical macroseismic observations used directly in ShakeMap play a vital role in constraining shaking from significant past earthquakes. In turn, these maps help us elucidate the nature and pattern of shaking behavior, damage, and ground failure that might otherwise remain elusive. With the portfolio of historical macroseismic data and modern DYFI-based MI observations, we have basic constraints on any event that left an impression on the regional population.

New Empirical Relations

Worden et al. (2012) and Caprio et al. (2015) employ DYFI data to derive new relations among a range of peak ground motion parameters and MI data, or ground motion-intensity conversion equations (GMICEs). The development of these relations sets a new standard for ground motion to intensity relations in that the DYFI intensity data used are decimal intensities and inverse relations are provided explicitly. Similarly, DYFI data have been used to derive (IPEs, e.g., Atkinson et al., 2014) to estimate MI directly from magnitude and distance. Both GMICEs and IPEs are important for robust ShakeMap generation (Worden et al., 2018) and hazard evaluations that incorporate historical

macroseismic observations, as well as for improved DYFI real-time filtering. In addition, the spatial coverage of DYFI and the precision provided by geocoding allows researchers to study regional amplification effects from intensities to complement instrumental data (e.g., Van Noten et al., 2016).

Induced Earthquakes

One special subset of the DYFI data is the Induced Events Database, which collects all data received for induced seismicity in the Central United States Over 200,000 observations for these events have been collected in the past decade with 22,000 at epicentral distances less than 20 km. These data have been particularly useful in determining the unique characteristics of induced events and evaluating their potential for damage (e.g., Atkinson et al., 2018). We foresee that this catalog will be useful for improving IPEs and GMICEs specifically for areas at risk of induced seismicity. The catalog also includes tools to create specific subsets of events and allows researchers to download intensity data for their own customized datasets.

Social Science and Behavior Studies

The DYFI portal allows for a participatory experience. Users coming to the USGS for information are empowered to become data providers themselves by contributing valuable observations that benefit the USGS as well as the participants, their local communities, and earthquake responders. DYFI also provides an important human perspective on earthquakes, providing sociological documentation of the way people behave and

respond, and how they perceive risk (e.g., Celsi et al., 2005; Goltz et al., 2020).

What's more, DYFI seems to provide emotional support to citizens who have just had a frightening or even traumatic experience (e.g., Casey et al., 2018). By allowing citizens to share their experiences and enabling them to contribute their observations toward a general public understanding of the phenomenon they have experienced, DYFI provides many with a form of catharsis at an opportune time. Often, users describe the desire to confirm their experience with their larger community. The DYFI system also educates the public on oft- misunderstood seismological concepts like the geographic variations of shaking intensity and the difference between earthquake magnitude and MI (Celsi et al., 2005).

In addition, the DYFI questionnaire includes questions about contributors' situations, experiences, and behaviors that go beyond the calculation of MI. Hence, the DYFI database is a repository of millions of relevant social science observations that is still largely untapped. Some researchers have started to examine DYFI data in this light. Boatwright and Phillips (2017) explore ZIP code population demographics of "Did You Feel It?" responses in California to correct for potential sampling biases as an effort to better estimate the felt area of moderate earthquakes there. Mak and Schorlemmer (2016) ask, "What makes people respond to "Did You Feel It?" They were concerned mostly with the question of data completeness, but conclude that the number of responses depends not only on population and felt intensity but also social factors such as ethnicity, education, and age. Likewise, Goltz et al. (2020) examine response behaviors of DYFI users in various regions during earthquakes of various magnitudes, emphasizing the need for further study of appropriate response during potentially damaging earthquakes while observing that "studies that specifically address the response of persons during earthquake shaking are few in number."

We hope that the creation of specialized, accessible data subsets such as the Induced Events Database, and increasing the accessibility of DYFI generally, will encourage interest from social scientists and researchers from other fields.

Toward a Common Macroseismic Scale

Several other national government agencies employ DYFI software or use the DYFI questionnaire and intensity algorithm for their domestic macroseismic data collection: Geosciences Australia, New Zealand's GeoNet, Natural Resources Canada, the British Geological Survey, and France's Bureau Central Sismologique Français, among others. Having a common code base throughout multiple countries not only facilitates maintenance; it allows disseminating new techniques and best practices among agencies. More importantly, the use of compatible questionnaires makes possible a larger common dataset for scientific research.

At higher MI values (typically, VIII and greater), neither DYFI intensities nor Modified Mercalli Intensity assignments are particularly well defined. This intensity range primarily describes observed structural damage to buildings (e.g., Musson et al., 2010). Building vulnerability and damage grading play a crucial role in assigning high intensities, and assessing these requires a degree of engineering expertise that most DYFI contributors lack. Tosi et al. (2015) group the higher degrees of EMS-98 (>VII) into a single class, maintaining that direct evaluations by experts are needed for correct assessment.

Other than via DYFI, the USGS no longer maintains dedicated staff to assign traditional MI assignments. The USGS is therefore interested in pursuing MI data collection that combines the advantages of DYFI for crowd-sourced, massive MI data collection for lower MI (<VII, which is >95% of all MI data collected) with professional assignments at higher MI based on the more systematic EMS-98 methodology. We aim to support the development of tools for domestic MI collection that utilize engineering expertise via onsite reconnaissance, remote imagery, and other rapid data-collection strategies. Employing EMS-98 domestically will require its adaptation for United States structures, partnering with professionals to calibrate EMS-98 to United States earthquake damage data and developing outreach materials to facilitate its adoption for future domestic earthquakes. Likewise, efforts are ongoing to employ more uniform data collection strategies, with the goal of harmonizing data collection around the globe (e.g., Goded et al., 2018). Some progress to this end has been made on global macroseismic data harmonization through efforts by the European Seismological Commission's Working Group in Macroseismology (Van Noten et al., 2018). Likewise, continuing efforts to develop a Global Macroseismic Scale (Spence and Foulser-Piggott, 2014) continue from time to time.

The use of uncertainty estimates in the newest version of ShakeMap provides a possible solution to combining disparate intensity scales. Any intensity measure can now be turned into DYFI-like intensity "stations" in ShakeMap, and combined with DYFI and other data, as long as that intensity has a computed uncertainty. Intensities assigned by expert observers could be combined this way as well.

NEW OPPORTUNITIES

Increasing DYFI Data Access

The USGS is developing various ways of facilitating access to earthquake data. DYFI has been fully integrated with other USGS earthquake products available from the National Earthquake Information Center (NEIC). Guy et al. (2015) summarizes the various NEIC systems that support earthquake triggering, data processing, and product delivery. For researchers, DYFI data have been made accessible via the ComCat earthquake archives and database. Users can now replicate, filter, and update aggregated datasets via USGS web services, although individual responses are not accessible (except by special request) given that they contain Personally Identifiable Information (PII). A web service and Python library for ComCat are available for researchers to automate queries for DYFI and other products.

We have replaced or enhanced many of the DYFI static maps and products with dynamic versions using modern web tools such as GIS, GeoJSON, and Leaflet (Smoczyk et al., 2017). Web displays of DYFI data are now zoomable and interactive,

allowing users to overlay their choice of layers that include 1- or 10-km geocoded boxes, population density, and ShakeMap contours and stations. Our online products now provide more information by clicking or hovering over different data elements (see **Figures 1**–**5**). We are also providing annual and cumulative aggregation maps showing the highest intensities reported in every UTM block (e.g., **Figure 7**). We have worked to enable usability of all webpages on smartphones and tablets, which are now the primary means of submitting questionnaire responses and viewing DYFI content.

Event Magnitude and Location With DYFI Data

Often small events, typically less than magnitude 3.5, occur near small population centers where observers quickly report via DYFI. We have developed a grid-search algorithm that employs the first set of incoming entries to determine the best-fitting magnitude and location based on region-specific IPEs (Quitoriano and Wald, 2016). NEIC analysts receive notifications of clusters of felt reports and the estimated location and magnitude from DYFI if the solution is sufficiently well constrained. Oftentimes, these approximate solutions provide a heads-up to NEIC seismic analysts of small events that may otherwise take time to locate. The origin time – needed to find the event in the seismic traces – becomes obvious since the first reported time is typically only a minute or two *after* the event's occurrence.

Developing New Tools and Approaches

We are currently developing a voice-activated DYFI questionnaire (currently for an Alexa Skill for Amazon's Alexa Smart Speaker). The difference in listening to and interacting with the questionnaire verbally, as opposed to a screen, necessitates a "Conversational User Interface" (CUI) that is easier to use than simply reading the questionnaire out loud. Considerable effort, including feedback from test users, went into the CUI to allow for more natural conversations. With proper care in skill development, voice-enabled Internet of Things (IoT) devices may allow people to interact and respond more easily with DYFI during real world events and eliminate some of the technological barriers to entry.

Ultimately, such IOT devices will likely all have accelerometers such that colocated human and instrumental measurements could be commonplace. Gathering colocated accelerometric parameters for joint analyses of instrumental and human observations is a holy grail in human-centric ground motion seismology. As mentioned earlier, anticipating the expansion of earthquake reporting on ubiquitous, lower cost – but limited quality – smartphone and speaker sensors, USGS ShakeMap can now accept uncertainty measures for intensity and ground motion parameters and weigh their contributions accordingly.

As part of the development process of the voice interface, we also designed a DYFI Questionnaire Application Program Interface (API) that allows selected third parties to submit questionnaire entries from their own applications, without going through the online questionnaire. While this raises questions of security and data quality, it also opens the potential of increased participation by partnering and integrating with other data users.

Short Versus Long Form Intensity Questionnaires

The evolution from manual, postal macroseismic questionnaires to emailed forms to internet surveys has been accomplished in many regions of the world. Many countries either maintain a manual approach as the primary strategy or reserve the option to augment their web-based approaches with traditional assignments. Several very successful internet-based macroseismic survey systems are now implemented in many countries or regions (see summaries in Wald et al., 2011 and Goded et al., 2018). One recent trend in the collection of felt reports is the rapid collection of large quantities of observer reports using simple picture-based options, which we refer to as the "short form." These are now used for worldwide events by EMSC and in New Zealand (Goded et al., 2018) as an adjunct to the "traditional" long-form questionnaire.

We agree that short form questionnaires have advantages over the long form. Their ease of use on mobile phone apps allows contributors to fill them much more quickly compared to full questionnaires, potentially increasing coverage and user participation. For example, during the 2016 M7.8 Kaikoura Earthquake (New Zealand), "[GNS] got 15,000 felt rapids in first 30 min. (which is an) order of magnitude more than traditional [questionnaires]" (N. Horspool, written communication, 11/17/2016). We believe that, given sufficient calibration, they might be relatively accurate.

However, we have two concerns with the short form. The first is the lack of precision. Since discerning differences of a single intensity unit from pictures is hard, the intensity scales of these short-form questionnaires are necessarily much coarser than a full questionnaire such as DYFI. Some of this concern may be alleviated by our ongoing efforts to quantify the uncertainty of these systems. Bossu et al. (2017) provide a bias correction for EMSC short-form responses to better align with DYFI intensity values.

A more fundamental problem with short-form questionnaires is that they leave no archival record of the actual effects and observations that are essential to establishing higher intensities. Many historical studies (e.g., Ambraseys and Douglas, 2004; Szeliga et al., 2010; Hough, 2013) have relied on the reevaluation of documented accounts of shaking and earthquake effects. Such studies would be impossible from a purely picture-based dataset.

Allowing short form derived intensities to be used in ShakeMap as long as their uncertainties can be quantified could ameliorate the record issue. For significant earthquakes, we could encourage follow-up responses employing DYFI or additional engineering assessments for archival purposes. For example, the MyShake Earthquake Early Warning platform allows its users to make short-form observations. Pointing those users to the DYFI questionnaire to capture more detailed information about their experience is a possibility.

SUMMARY

In this study, we tried to directly address the questions posited by the Editors of this Special Issue on the Power of Citizen Seismology (slightly edited):

- How much does public involvement help awareness and preparation toward seismic impact, and how does it affect public communication?
- How has the Power of Citizen Seismology made a difference in influencing government agency actions?
- What specific scientific advances have been made through data integration and interoperability between projects/across countries?
- What ethical and other challenges have been encountered?

Fundamentally, DYFI relies on input from the general public, rather than trained citizen-scientists, so more properly, DYFI is citizen-based science, rather than citizen science. In conjunction with ShakeMap, DYFI has substantially facilitated the use of MI throughout the United States, educating millions of citizens who experienced earthquakes to think in terms of the varying intensities produced by an earthquake rather than the poorly understood concept of magnitude. ShakeMap and the citizen-based science of DYFI in particular have played an important role in guiding the media and the public toward a more suitable way to describe the variations of earthquake shaking, and thus to better understand the nature of earthquake shaking hazards and risks more generally. And, given the public's uptake of MI domestically following the advent of DYFI and ShakeMap, the USGS' 2019 public release of Earthquake Early Warning (EEW) in the United States considered intensity to be the most suitable metric for warning the populace of imminent shaking with intensity-based depictions of shaking levels (e.g., Given et al., 2018). The ShakeAlert EEW system, now in operation along the United States West Coast, communicates MI using the Modified Mercalli Intensity scale (MMI) and determines which areas to alert using prescribed MMI thresholds. This strategy is consistent with the long-held approach adopted in Japan, where the Japan Meteorological Society (JMA) seismic intensity scale is very well established and understood within the community, media, and decision-makers (e.g., Doi, 2011), and preferred over magnitude as the main earthquake information delivered.

This reintroduction of the concept and use of MI to the general public has, in turn, allowed numerous federal and state government agencies including the USGS and Federal Emergency Management Agency (FEMA), various non-governmental agencies, and earthquake managers and responders to more widely adopt the depiction and communication of shaking intensity, thus promoting a more intuitive understanding of earthquake shaking hazards. This general promotion of the use of intensity can provide benefits to the general public's understanding of earthquake hazards and risk (Celsi et al., 2005).

As a web-based citizen science tool, the data collected with DYFI data are neither immutable nor perfect. And yet, the IPEs developed from DYFI data are remarkably robust and have proven to be invaluable. We have described our strategies for continued quality control. Robust treatment of the DYFI data lead to robust results. One conceivable ethical issue raised over the two-decade experiment with DYFI data collection is in the potential for manipulation of aggregated intensity values for ulterior motives by contributing spurious data. While this issue is addressed in detail by Wald et al. (2011), we add here that the direct use of DYFI data in ShakeMap can accommodate spurious observations by culling outliers and weighting DYFI data according to their uncertainty, and with a second, failsafe strategy: providing additional ShakeMap layers where DYFI data are not utilized in the computation of the ground motion field. To date, we have few examples of spurious DYFI entries during significant events; a few random entries do occasionally show up during quiet times (Wald et al., 2011). Another issue pertaining to DYFI is with respect to user privacy. Since the data are aggregated into 1- km and 10-km cells, individual users are not recoverable from DYFI products. Whereas individual entries are made available for research purposes, they are anonymized by removing PII and truncating the precision of provided locations.

Operationally, the DYFI data, when integrated directly for use in ShakeMap, allow for better-constrained estimates of shaking for significant earthquakes around the globe (to varying degrees, based on the region's uptake of DYFI). In turn, better-constrained ShakeMaps as input for USGS earthquake impact products such as PAGER and ShakeCast (Lin and Wald, 2008) improve our ability to project useful loss estimates immediately following earthquakes worldwide. In this context, interoperability across nations could be further achieved by the types of analyses described herein where both obvious trends and heteroscedastic uncertainties can be accommodated in ShakeMap. Moreover, we have described a wide array of seismological hazard and risk studies that depend primarily on DYFI data, ranging from response-oriented applications (ShakeMap constraints), to better ground motion estimates of intensity and shaking from induced earthquakes, to social and behavioral science.

DATA ACCESS

"Did You Feel It?" can be found online at the website http://earthquake.usgs.gov/dyfi/. Event queries can be made through the ComCat webpage at https://earthquake.usgs.gov/data/comcat/. Command line tools and the Python API for accessing ComCat are available at https://github.com/usgs/libcomcat. The DYFI Induced Events Database is available online in USGS ScienceBase data archives (https://doi.org/10.5066/F7WM1BPC). Specialized DYFI data requests, including (anonymized) user entries and comments, can be made to the authors upon request for educational or research purposes.

AUTHOR CONTRIBUTIONS

DW wrote the first draft of the manuscript. Both authors wrote or rewrote the sections of the manuscript. Both authors contributed to manuscript revision, read, and approved the submitted version.

REFERENCES

Allen, R. M., Kong, Q., and Martin-Short, R. (2020). The MyShake platform: a global vision for earthquake early warning. *Pure Appl. Geophys.* 177, 1699–1712. doi: 10.1007/s00024-019-02337-7

Allen, T., Carapetis, A., Bathgate, J., Ghasemi, H., Pejić, T., and Moseley, A. (2019). "Real-time community internet intensity maps and shakemaps for australian earthquakes," in *Proceedings of the Australian Earthquake Engineering Society 2019 Conference*, Newcastle, NSW, 13.

Ambraseys, N. N., and Douglas, J. (2004). Magnitude calibration of north Indian earthquakes. *Geophys. J. Int.* 159, 165–206.

Atkinson, G. M., and Wald, D. J. (2007). "Did You Feel It?" Intensity data: a surprisingly good measure of earthquake ground motion. *Seis. Res. Lett.* 78, 362–368. doi: 10.1785/gssrl.78.3.362

Atkinson, G. M., Wald, D. J., Worden, C. B., and Quitoriano, V. (2018). The intensity signature of induced seismicity. *Bull. Seism. Soc. Am.* 108, 1080–1086. doi: 10.1785/0120170316

Atkinson, G. M., Worden, C. B., and Wald, D. J. (2014). Intensity prediction equations for north america. *Bull. Seism. Soc. Am.* 104, 3084–3093. doi: 10.1785/0120140178

Boatwright, J., and Phillips, E. (2017). Exploiting the demographics of "Did You Feel It?" Responses to Estimate the felt area of moderate earthquakes in California. *Seism. Res. Lett.* 88, 335–341. doi: 10.1785/0220160041

Bossu, R., Landès, M., Roussel, F., Steed, R., Mazet-Roux, G., Martin, S. S., et al. (2017). Thumbnail-based questionnaires for the rapid and efficient collection of Macroseismic Data from Global Earthquakes. *Seism. Res. Lett.* 88, 72–81. doi: 10.1785/0220160120

Caprio, M., Tarigan, B., Worden, C. B., Wald, D. J., and Wiemer, S. (2015). Ground Motion to Intensity Conversion Equations (GMICEs): a global relationship and evaluation of regional dependency. *Bull. Seism. Soc. Am.* 105, 1476–1490. doi: 10.1785/0120140286

Casey, J. A., Goldman-Mellor, S., and Catalano, R. (2018). Association between Oklahoma earthquakes and anxiety-related Google search episodes. *Environ. Epidemiol.* 2:e016. doi: 10.1097/EE9.0000000000000016

Celsi, R., Wolfinbarger, M., and Wald, D. J. (2005). The effects of magnitude anchoring, earthquake attenuation estimation, measure complexity, hubris, and experience inflation on individuals' Perceptions of felt earthquake experience and perceptions of earthquake risk. *Earthquake Spectra.* 21, 987–1008.

Dengler, L. A., and Dewey, J. W. (1998). An intensity survey of households affected by the Northridge, California, earthquake of 17 January 1994. *Bull. Seism. Soc. Am.* 88, 441–462.

Doi, K. (2011). The operation and performance of earthquake early warnings by the Japan Meteorological Agency, Soil. Dyn. *Earthquake Eng.* 31, 119–126.

Given, D. D., Allen, R. M., Baltay, A. S., Bodin, P., Cochran, E. S., Creager, K., et al. (2018). *Revised Technical Implementation Plan for the ShakeAlert system—An Earthquake Early Warning System for the West Coast of the United States: U.S. Geological Survey Open-File Report 2018–1155*. Reston, VA: U.S. Geological Survey, doi: 10.3133/ofr20181155

Goded, T., Horspool, N., Canessa, S., Lewis, A., Geraghty, K., Jeffrey, A., et al. (2018). New macroseismic intensity assessment method for New Zealand Web Questionnaires. *Seismol. Res. Lett.* 89, 640–652. doi: 10.1785/0220170163

Goltz, J. D., Park, H., Nakano, G., and Yamori, K. (2020). Earthquake ground motion and human behavior: using DYFI data to assess behavioral response to earthquakes. *Earthquake Spectra.* doi: 10.1177/8755293019899958 [Epub ahead of print].

Grünthal, G. (1998). European macroseismic scale 1998 (EMS-98). *Cahiers Centre Europe'en Geodynamique Seismologie Luxembourg* 15, 1–99.

Guy, M., Patton, J., Fee, J. M., Hearne, M., Martinez, E. M., Ketchum, D., et al. (2015). *National Earthquake Information Center Systems Overview And Integration: USGS Open File Report.* Reston, VA: USGS, 2015–1120. doi: 10.3133/ofr20151120

Hough, S. E. (2013). Spatial Variability of "Did You Feel It?" Intensity data: insights into sampling biases in historical earthquake intensity distributions. *Bull. Seism. Soc. Am.* 103, 2767–2781. doi: 10.1785/0120120285

Lin, K. W., and Wald, D. J. (2008). ShakeCast manual: U.S. *Geol. Surv. Open File Rep.* 200:90.

Mak, S., and Schorlemmer, D. (2016). What makes people respond to "Did You Feel It?" *Seismol. Res. Lett.* 87, 119–131.

Montalvo-Arrieta, J. C., Perez-Campos, X., Ramirez-Guzman, L., Sosa-Ramirez, R. L., Ruiz-Esparza, M. C., and Leonardo-Suarez, M. (2019). Macroseismic Intensities from the 19 September 2017 7.1 Puebla-Morelos Earthquake. *Seismol. Res. Lett.* 90, 2142–2153. doi: 10.1785/0220190145

Musson, R. M. W., Grünthal, G., and Stucchi, M. (2010). The comparison of macroseismic intensity scales. *J. Seismol.* 14, 413–428. doi: 10.1007/s10950-009-9172-0

Petersen, M. D., Shumway, A. M., Powers, P. M., Mueller, C. S., Moschetti, M. P., Frankel, A. D., et al. (2020). The 2018 update of the US National Seismic Hazard Model: Overview of model and implications. *Earthq. Spectra* 36, 5–41. doi: 10.1177/8755293019878199

Quitoriano, V., and Wald, D. J. (2016). Use of U. S. Geological Survey's "Did You Feel It?" Responses for Near Real-Time Earthquake Location, Seismo. *Res. Lett.* 87:567.

Sbarra, P., Tosi, P., and De Rubeis, V. (2010). Web-based macroseismic survey in Italy: method validation and results. *Nat. Hazards* 54, 563–581. doi: 10.1007/s11069-009-9488-7

Sbarra, P., Tosi, P., and De Rubeis, V. (2014). How observer conditions impact earthquake perception. *Seismol. Res. Lett.* 85, 306–313. doi: 10.1785/0220130080

Smoczyk, G. M., Wald, D. J., Worden, C. B., Thompson, E. M., Quitoriano, V., and Hearne, M. (2017). "Seismology and GIS: USGS near real-time significant earthquake and earthquake scenario GIS feeds," in *Proceedings of the Seismological Society of America 2017 Annual Meeting*, Denver, CO.

Spence, R., and Foulser-Piggott, R. (2014). "The International Macroseismic Scale – extending EMS-98 for global application," in *Proceedings of the Second European Conference on Earthquake Engineering and Seismology*, Istanbul,

Szeliga, W., Hough, S., Martin, S. S., and Bilham, R. (2010). Intensity, magnitude, location, and attenuation in India for felt earthquakes since 1762. *Bull. Seismol. Soc. Am.* 100, 570–584.

Tosi, P., Sbarra, P., De Rubeis, V., and Ferrari, C. (2015). Macroseismic intensity assessment method for web questionnaires. *Seismol. Res. Lett.* 86, 985–990. doi: 10.1785/0220150127

Van Noten, K., de Rubeis, V., Tosi, P., and Sbarra, P. (2018). "Challenges of the new ESC Working Group Macroseismology—Integrating corrective parameters to merge multiple-sourced online macroseismic data," in *Proceedings of the 6th Colloquium on Historical Earthquakes & Paleoseismology Studies, Han-sur-lesse*, Belgium,

Van Noten, K., Lecocq, T., Sira, C., Hinzen, K.-G., and Camelbeeck, T. (2016). Path and site effects deduced from transfrontier internet macroseismic data of two recent M4 earthquakes in NW Europe. *Solid Earth Discuss* 8, 453–477. doi: 10.5194/se-2016-150

Wald, D. J., Earle, P. S., Porter, K., Jaiswal, K., and Allen, T. I. (2008). "Development of the U.S. Geological Survey's prompt assessment of global earthquakes for Response (PAGER) System," in *Proceedings of the 14th World Conference. Earthquake Engineering*, Beijing.

Wald, D. J., Quitoriano, V., Dengler, L., and Dewey, J. W. (1999a). Utilization of the internet for rapid community intensity maps. *Seismol. Res. Lett.* 70, 680–697.

Wald, D. J., Quitoriano, V., Heaton, T. H., Kanamori, H., Scrivner, C. W., and Worden, C. B. (1999b). TriNet "ShakeMaps": rapid generation of peak ground motion and intensity maps for earthquakes in southern California. *Earthquake Spectra* 15, 537–556.

Wald, D. J., Quitoriano, V., Worden, C. B., Hopper, M., and Dewey, J. W. (2011). USGS "Did You Feel It?" Internet- based macroseismic intensity maps. *Ann. Geophys.* 54, 688–709. doi: 10.4401/ag-5354

Worden, C. B., Gerstenberger, M. C., Rhoades, D. A., and Wald, D. J. (2012). Probabilistic relationships between peak ground motion and Modified Mercalli intensity. *Bull Seism. Soc. Am.* 102, 204–221. doi: 10.1785/01201 10156

Worden, C. B., Thompson, E. M., Baker, J. W., Bradley, B. A., Luco, N., and Wald, D. J. (2018). Spatial and spectral interpolation of ground motion intensity measure observations. *Bull. Seism. Soc. Am.* 108, 866–875. doi: 10. 1785/0120170201