Ⓔ

What Googling Trends Tell Us about Public Interest in Earthquakes

by Yen Joe Tan and Rijan Maharjan

ABSTRACT

Previous studies have shown that immediately after large earthquakes, there is a period of increased public interest. This represents a window of opportunity for science communication and disaster-relief fundraising efforts to reach more people. However, how public interest varies for different earthquakes has not been quantified systematically on a global scale. We analyze how global search interest for the term "earthquake" on Google varies following destructive earthquakes from 2004 to 2016. We find that there is a spike in search interest after destructive earthquakes followed by an exponential temporal decay. The duration and time constant of increased search interest correlate with death toll and damages but did not correlate with earthquake magnitude, estimated population exposed to very strong shaking, and number of U.S. Geological Survey (USGS) "Did You Feel It?" (DYFI) responses. Furthermore, we obtain similar time constants of increased search interest when analyzing just the U.S. search interest following destructive earthquakes outside of U.S.A., Canada, and Mexico. This suggests that a significant portion of the increased search interest comes from people who did not feel the shaking. Our observations are consistent with more destructive earthquakes receiving more media coverage which leads to a longer duration of elevated public interest in earthquakes. Of the 73 earthquakes that resulted in an increase in global search interest that fit our selection criteria, only 11 (15%) resulted in an elevated search interest of more than a week. Therefore, to take advantage of these short durations of increased public interest, science communication and disaster-relief fundraising efforts have to act promptly following devastating earthquakes.

Electronic Supplement: Tables of characteristics for the 73 earthquakes included in the study as well as exponential and power-law parameters for the 10 earthquakes with the longest duration of increased search interest, and figures comparing N and T with earthquake magnitude, estimated population exposed to very strong shaking (modified Mercalli intensity [MMI] ≥ VII), and the number of U.S. Geological Survey (USGS) "Did You Feel It?" (DYFI) responses.

INTRODUCTION

Science website traffic (Bossu *et al.*, 2011, 2014; Quigley and Forte, 2017), earthquake-related Twitter messages (Earle *et al.*, 2012), and earthquake-notification mailing-list subscriptions (Schwarz, 2004) spike and decay after large earthquakes. This illustrates that immediately after an earthquake, there is a period of increased public interest during which outreach and disaster-relief fundraising efforts might reach more people.

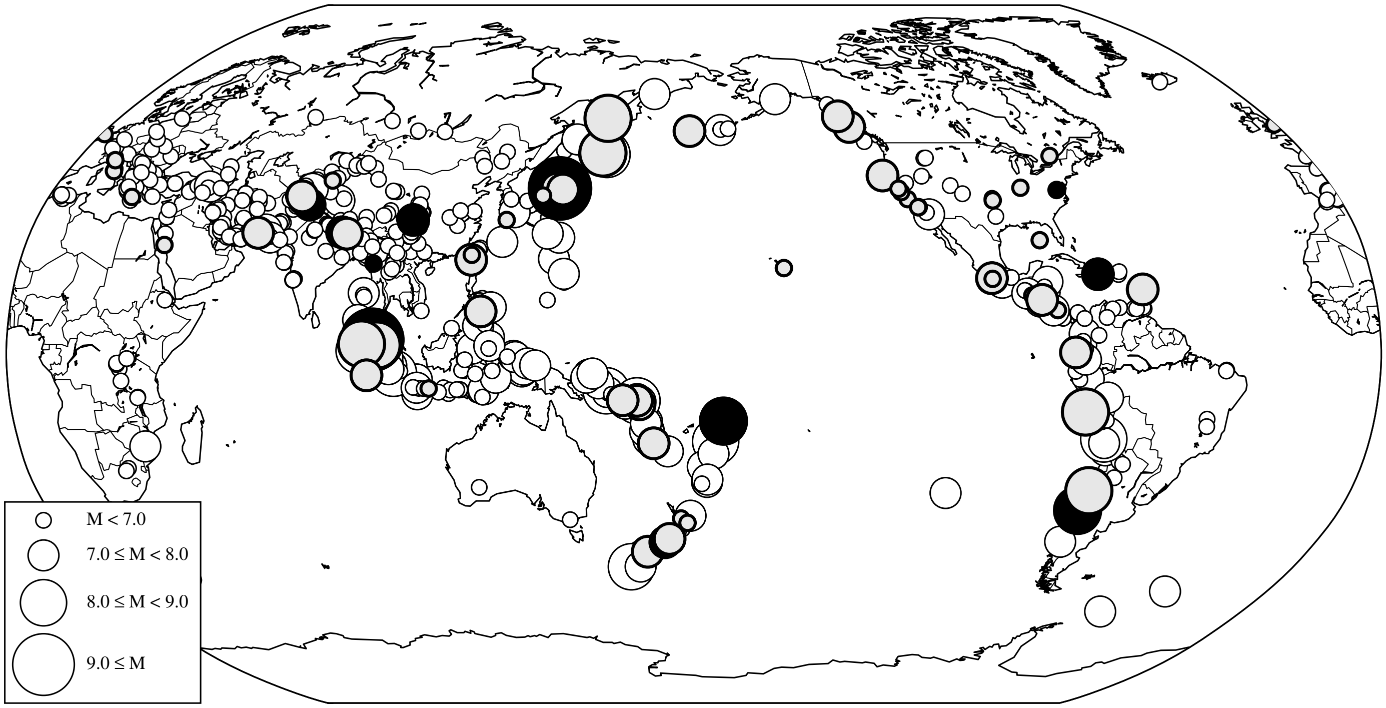
Looking at the U.S. Geological Survey (USGS) Southern California Earthquake Hazards Program's electronic mailing lists between 2001 and 2003, Schwarz (2004) found that the subscription rate increased above background rates for approximately two weeks following major earthquakes in California. Looking at the number of visitors at four science websites during the 2010–2012 Canterbury earthquake sequence in New Zealand, Quigley and Forte (2017) found that after large earthquakes, there is an immediate increase in website traffic followed by a power-law-characterized decay. The traffic perturbation scales with earthquake magnitude and peak ground acceleration. However, we are unaware of a previous attempt to systematically quantify how public interest varies for different earthquakes on a global scale.

In this study, we analyze how global search interest for the term "earthquake" on Google varies following destructive earthquakes from 2004 to 2016. We then explore if earthquake magnitude, death toll, damages, and the population size exposed to very strong shaking affect how search interest varies following different earthquakes.

GLOBAL GOOGLE SEARCH INTEREST

Google processes more than 3.5 billion searches per day (Sullivan, 2013). Since 2004, Google Trends has been providing the daily proportion of searches on Google for any term, scaled on a 0–100 range over the time interval of interest. The search interest provided is rounded off and hence consists only of integer values. In this study, we use the global search interest for the term "earthquake" provided by Google Trends as a measure of public interest in earthquakes. The search interest provided is not case sensitive and includes searches such as "Nepal

doi: 10.1785/0220170116



▲ **Figure 1.** Map showing all the earthquakes included in our analysis. The marker size is scaled by earthquake magnitude. The marker color varies by N, the total number of days of increased search interest: Black for $N > 7$ and dark gray for $N \leq 7$. Open circles are events not followed by a significant increase in search interest that fit our selection criteria.

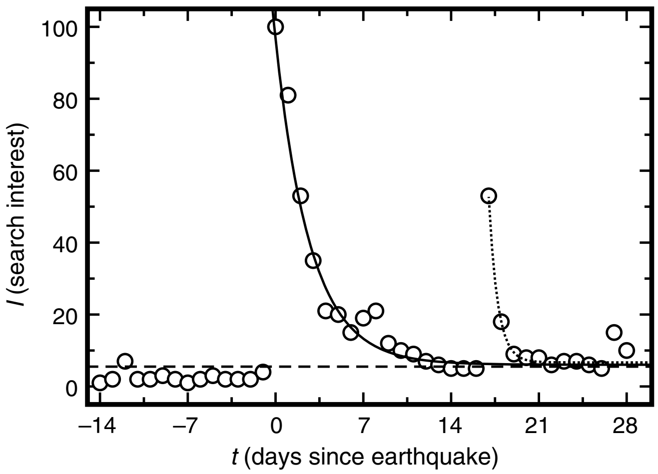
Earthquake" because the term "earthquake" is present. However, the search interest provided does not include searches for the term "earthquake" in other languages.

We look at how the global Google search interest changes in response to destructive earthquakes from 2004 to 2016 that are listed in the Significant Earthquake Database (see Data and Resources), available through the National Centers for Environmental Information of the National Oceanic and Atmospheric Administration. The catalog includes earthquakes that meet at least of one of the following criteria: caused deaths, caused damages ≥ 1 million U.S., generated a tsunami, of $M_w \geq 7.5$, or modified Mercalli intensity (MMI) ≥ X. The database also provides the death toll and damages, when available. When there is more than one earthquake on the same day, we only include the largest-magnitude event. Therefore, a total of 685 earthquakes is included in our analysis (Fig. 1).

For each earthquake, we download the global Google search-interest time series starting from two weeks before to four weeks after the earthquake (see Data and Resources). Because the search interest provided is scaled on a 0–100 range over the time interval of interest, the day with the largest search frequency in the six-week time range will have a normalized search interest of 100. We then obtain a background noise level, which we define as two standard deviation from the mean search interest in the two weeks before the earthquake (Fig. 2). This threshold signifies a 95% confidence level that a point above this background noise level represents an increase in search interest. We consider an earthquake to have resulted in a significant increase in search interest if the search interest on the day of the earthquake is the largest within the six-week time window and above the background noise level. For these events, we further check if there is an earthquake on the day before that is of a larger magnitude. If so, we reassign this increase in search interest to that earthquake. This additional step aims to circumvent the limitation of having only a daily Google search-interest time series, for which a peak in search interest could have resulted from an earthquake that occurred late in the previous day and would otherwise be wrongly assigned to a different earthquake. Finally, we use two parameters to quantify how long the search interest remains elevated after the earthquake.

The first parameter N is the total number of days of increased search interest. Following a spike in search interest on the day of the earthquake, we consider search interest to have dissipated when it falls below the background noise level for two consecutive days. For the four events with initially calculated $N > 21$ (three weeks), we extend the Google search-interest time series to eight weeks after the earthquake for reanalysis, so as not to have an artificial upper limit for the duration of increased search interest (see Ⓔ Fig. S1, available in the electronic supplement to this article). Although parameter N is an intuitive measure of the duration of increased search interest, it can be affected by aftershocks that create subsequent spikes in search interest (Fig. 2).

The second parameter T is the time constant of increased search interest. We fit the search-interest time series starting



▲ **Figure 2.** Search interest as a function of days since the 25 April 2015 M_w 7.9 earthquake in Nepal. Search interest is scaled on a 0–100 range, for which the day with the largest search frequency in the six-week time range will have a normalized search interest of 100. Dashed line marks the background noise level. Solid line marks the exponential fit to the search-interest time series. This earthquake was followed by an M_w 7.2 aftershock that resulted in a subsequent spike in search interest at $t = 17$. Dotted line marks the exponential fit to the aftershock search-interest time series.

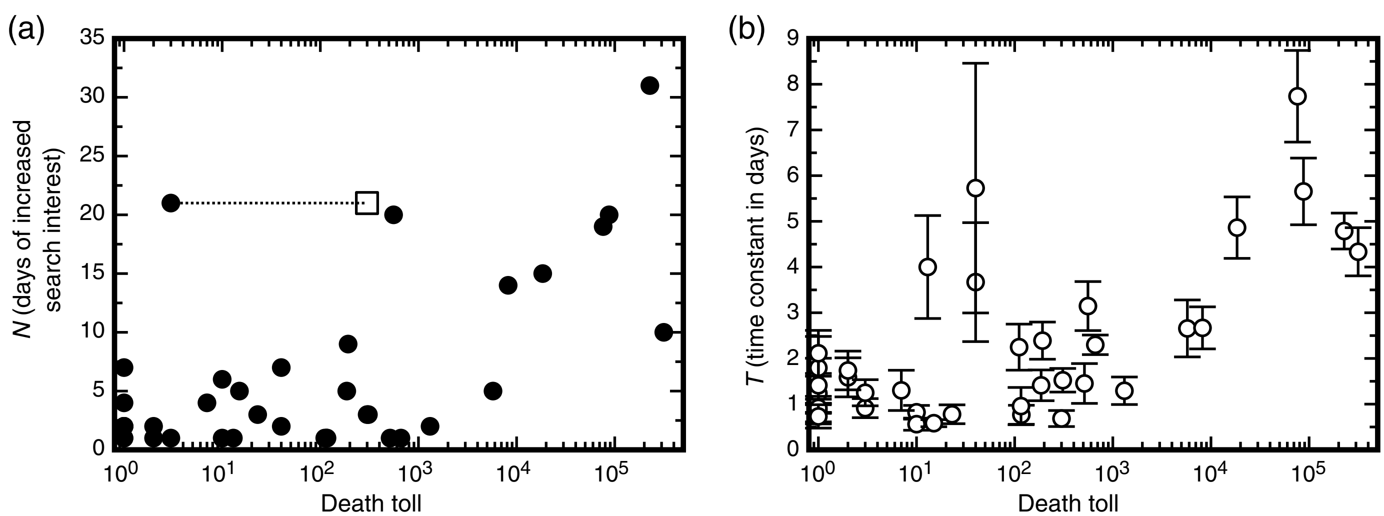
from the day of the earthquake to four weeks after with an exponential function

$$I = m \times \exp(-t/T) + c, \tag{1}$$

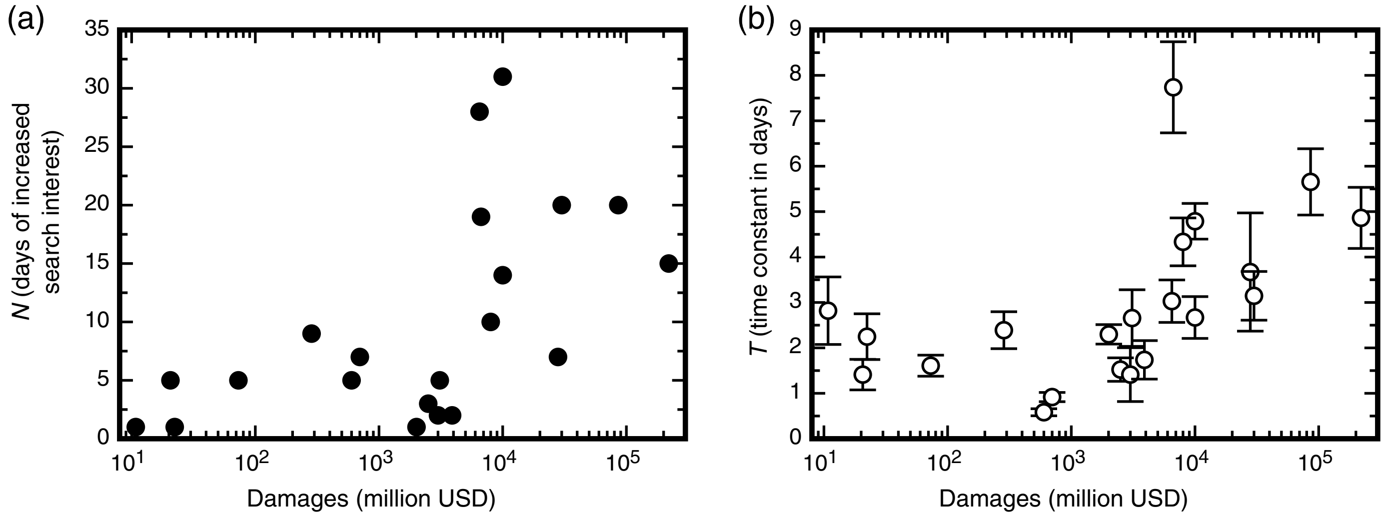
in which I is the search interest, m is the proportionality constant, t is the number of days since the earthquake, T is the time constant of increased search interest in days, and c is the constant. We obtain the fit parameters with meaningful uncertainties by adjusting the input errors until the reduced $\chi^2 \sim 1$. We only keep events for which the uncertainty is less than 50% of the measured T. We chose this parameter because it provides an intuitive quantification of how search interest decays with time; that is, T represents the time for search interest to fall below ~63% of the peak. In addition, T is less affected by subsequent spikes in search interest from aftershocks (Fig. 2) and is thus more robust than N for comparing different earthquakes. Finally, we find that, visually, the fall-off in search interest after an initial spike is well described by an exponential function (Fig. 2). We further verified that exponential fit is better than general power-law fit, based on their residual sum of squares (RSS) for 8 out of the 10 earthquakes with the longest duration of increased search interest, that is, largest Ns (see the Ⓔ electronic supplement). While further testing which functions, including the Omori law (Quigley and Forte, 2017), best describe how search interest decays with time would be beneficial, it is beyond the scope of this study.

RESULTS

Of the 685 earthquakes analyzed, 73 resulted in a significant increase in search interest that fit our selection criteria (Fig. 1). Of these, only 11 earthquakes (15%) resulted in increased search interest for more than a week (Figs. 3a and 4a). We investigate the factors affecting the duration of increased search interest by evaluating how N and T vary with earthquake magnitude, death toll, damages, estimated population exposed to very strong (MMI ≥ VII) shaking (Wald *et al.*, 2010), and number of USGS "Did You Feel It?" (DYFI) responses (Wald *et al.*,



▲ **Figure 3.** (a) Number of days of increased global search interest as a function of death toll. The one outlier event is the Myanmar M_w 6.8 earthquake on 24 August 2016. On the same day, the smaller M_w 6.2 Italy earthquake occurred and had a larger death toll. Associating the increase in search interest to the Italy earthquake instead would shift the point to that marked by the open square. (b) Time constant of increased global search interest as a function of death toll. Error bars represent 68% confidence interval. Only the 38 earthquakes with death toll information as provided in the Significant Earthquake Database are included. Both N and T increase systematically with earthquake death toll, especially when death toll is greater than 1000.

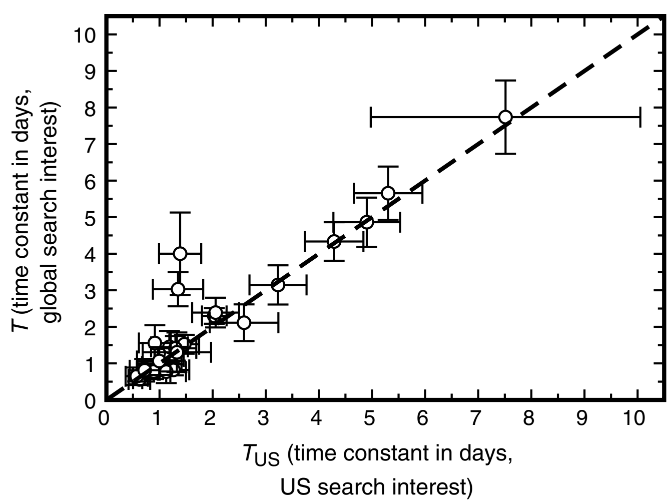


▲ **Figure 4.** (a) Number of days of increased global search interest as a function of damages. (b) Time constant of increased global search interest as a function of damages. Error bars represent 68% confidence interval. Only the 21 earthquakes with damages information as provided in the Significant Earthquake Database are included. Both N and T increase systematically with earthquake damages, especially when damages are greater than \$1 billion U.S.

2011). We find that neither N nor T shows a clear relationship with earthquake magnitude, estimated population exposed to very strong shaking, and number of USGS DYFI responses (see Ⓔ Figs. S3–S5). However, both N and T increase systematically with earthquake death toll and damages (Figs. 3 and 4). The correlation looks stronger for T, presumably because the estimated T is less affected by aftershocks. Therefore, we infer that earthquake death toll and damages are the main underlying factors that influence the duration of increased search interest.

We have two hypotheses explaining why Google search interest is elevated following devastating earthquakes and why the duration of increased search interest correlates with earthquake death toll and damages. First, more destructive earthquakes tend to have a larger population that felt the shaking, who then proceeded to obtain more information online. However, this explanation is inconsistent with N and T not correlating with estimated population exposed to very strong shaking and number of USGS DYFI responses. Nevertheless, the lack of correlation could be due to poor estimates of population exposed to very strong shaking, as well as uneven adoption of the USGS DYFI program globally.

Another potential explanation is that more destructive earthquakes receive more media coverage and public attention and hence have a larger population that learned of it, who then proceeded to obtain more information online. To test this hypothesis, which implies that a significant portion of the increased global Google search interest following devastating earthquakes comes from people who did not feel the shaking due to the earthquakes, we analyzed the Google search interest for just the U.S. population following devastating earthquakes in other countries. Of the 73 earthquakes that resulted in a significant increase in global search interest, 56 occurred outside of the U.S.A., Canada, and Mexico. We proceeded to analyze these events using the U.S. Google search-interest time series, following the same procedures as our global Google search-interest analysis. We find a strong correlation between the time constant T measured for the U.S. search interest and that for the global search interest (Fig. 5). Because the elevated search interest in the U.S.A. comes from people who did not feel the shaking due to the earthquakes, this is consistent with the hypothesis that more destructive earthquakes receive more media coverage, which leads to a longer duration of increased search interest. In addition, the strong correlation between the time constant T measured for the U.S. search interest and that



▲ **Figure 5.** Relationship between time constant of increased global search interest and time constant of increased U.S. search interest for earthquakes in countries other than Canada, U.S.A., and Mexico. Error bars represent 68% confidence interval. Dashed line marks a slope of 1. Only the 27 earthquakes with significant increase in both global and U.S. search interest are included.

for the global search interest might reflect U.S. searches being a significant portion of the global searches.

DISCUSSION AND CONCLUSIONS

Although Google search queries do not come from a truly random and unbiased sample of the world's population, it provides a valuable tool for us to quantitatively estimate public interest in various topics, including natural disasters such as earthquakes. This could provide useful information that can aid decision-making in science communication and disaster-relief fundraising efforts.

Our analyses suggest that following devastating earthquakes, global public interest spikes and then decays exponentially. The duration of increased global public interest increases systematically with death toll and damages. We obtain similar durations when analyzing the search interest for just the U.S. population following earthquakes outside of Canada, U.S.A., and Mexico. This suggests that the duration of increased public interest in earthquakes correlating with death toll and damages might reflect differences in media coverage. Of the 73 earthquakes that resulted in an increase in global search interest, only 11 (15%) resulted in an elevated search interest of more than a week. Therefore, to take advantage of these short durations of increased public interest, science communication and disaster-relief fundraising efforts have to act promptly following destructive earthquakes.

DATA AND RESOURCES

Google Trends time series were obtained from http://www.google.com/trends/ (last accessed November 2017). Estimated population exposed to strong shaking data were obtained from http://earthquake.usgs.gov/pager/ (last accessed November 2017), whereas "Did You Feel It?" (DYFI) data were obtained from http://earthquake.usgs.gov/dyfi/ (last accessed November 2017). Earthquake magnitudes, death toll, and damages were obtained from National Oceanic and Atmospheric Administration's (NOAA's) Significant Earthquake Database, doi: 10.7289/V5TD9V7K.

ACKNOWLEDGMENTS

The authors thank Eric Brown, Maya Tolstoy, and Felix Waldhauser for their careful reading of the article and valuable feedback. The authors thank David Wald, Lisa Wald, and Michael Hearne for their help and advice on extracting data from the U.S. Geological Survey (USGS) database. The authors thank Editor-in-Chief Zhigang Peng, Rémy Bossu, and Mark Quigley for their constructive comments. This study was partially funded by the Division of Marine Geology and Geophysics of the Lamont-Doherty Earth Observatory.

REFERENCES

Bossu, R., S. Gilles, G. Mazet-Roux, F. Roussel, L. Frobert, and L. Kamb (2011). Flash sourcing, or rapid detection and characterization of earthquake effects through website traffic analysis, *Ann. Geophys.* **54,** no. 6, 716–727.

Bossu, R., S. Lefebvre, Y. Cansi, and G. Mazet-Roux (2014). Characterization of the 2011 Mineral, Virginia, earthquake effects and epicenter from website traffic analysis, *Seismol. Res. Lett.* **85,** no. 1, 91–97.

Earle, P. S., D. C. Bowden, and M. Guy (2012). Twitter earthquake detection: Earthquake monitoring in a social world, *Ann. Geophys.* **54,** no. 6, 708–715.

Quigley, M. C., and A. M. Forte (2017). Science website traffic in earthquakes, *Seismol. Res. Lett.* **88,** no. 3, 867–874.

Schwarz, S. (2004). Cyberseismology and teachable moments, *Seismol. Res. Lett.* **75,** no. 6, 749–750.

Sullivan, D. (2013). Google still world's most popular search engine by far, but share of unique searchers dips slightly, available at http://searchengineland.com/google-worlds-most-popular-search-engine-148089 (last accessed May 2017).

Wald, D. J., K. S. Jaiswal, K. D. Marano, M. Hearne, and D. Bausch (2010). PAGER—Rapid assessment of an earthquake's impact, *USGS Fact Sheet 2010-3036*, 4 pp.

Wald, D. J., V. Quitoriano, C. B. Worden, M. Hopper, and J. W. Dewey (2011). USGS "Did You Feel It?" Internet-based macroseismic intensity maps, *Ann. Geophys.* **54,** no. 6, doi: 10.4401/ag-5354.

Yen Joe Tan
Lamont-Doherty Earth Observatory
Columbia University
61 Route 9 West, P.O. Box 1000
Palisades, New York 10964 U.S.A.
yjt@ldeo.columbia.edu

Rijan Maharjan
Department of Mechanical Engineering and Materials Science
Yale University
9 Hillhouse Avenue, M11
New Haven, Connecticut 06520 U.S.A.
rijan.maharjan@yale.edu

Published Online 24 January 2018