Tenth U.S. National Conference on Earthquake Engineering
Frontiers of Earthquake Engineering
July 21-25, 2014
Anchorage, Alaska

10NCEE

RAPID ESTIMATE OF GROUND SHAKING INTENSITY BY COMBINING SIMPLE EARTHQUAKE CHARACTERISTICS WITH TWEETS

L. Burks[1]*, M. Miller[1]*, and R. Zadeh[2]

ABSTRACT

Here we demonstrate a model that combines Tweets following significant earthquakes with basic site and earthquake characteristics to produce a rapid estimate of the ground shaking intensity. We used all geo-tagged Tweets around the world containing the keyword "earthquake" or "tsunami" in several languages that occurred in the first 10 minutes following Japanese earthquakes of magnitude 6 or greater from 2011 to 2012. Using different regression models, we found relevant features and the best model by minimizing the mean squared error between model predictions and historical estimates of shaking intensity. We found that the model with lowest error was based on a combination of earthquake- and Tweet-based features, such as local site conditions, source-to-site distance, and number of Tweets within a certain radius. Ground shaking intensity estimates from our model are comparable with historical recordings and conventional estimates provided, for example, by the United States Geological Survey ShakeMaps. Therefore Tweets may be a useful additional data source following significant earthquakes, especially in regions without an extensive network of recording stations.

[1] PhD Candidate, Dept. of Civil Eng., Stanford University, Stanford, CA USA

[2] Consulting Professor, Institute of Computational and Mathematical Eng., Stanford University, Stanford, CA USA
* The first two authors contributed equally to this work.

Rapid estimate of ground shaking intensity by combining simple earthquake characteristics with tweets

L. Burks[3]*, M. Miller[1]*, and R. Zadeh[4]

ABSTRACT

Here we demonstrate a model that combines Tweets following significant earthquakes with basic site and earthquake characteristics to produce a rapid estimate of the ground shaking intensity. We used all geo-tagged Tweets around the world containing the keyword "earthquake" or "tsunami" in several languages that occurred in the first 10 minutes following Japanese earthquakes of magnitude 6 or greater from 2011 to 2012. Using different regression models, we found relevant features and the best model by minimizing the mean squared error between model predictions and historical estimates of shaking intensity. We found that the model with lowest error was based on a combination of earthquake- and Tweet-based features, such as local site conditions, source-to-site distance, and number of Tweets within a certain radius. Ground shaking intensity estimates from our model are comparable with historical recordings and conventional estimates provided, for example, by the United States Geological Survey ShakeMaps. Therefore Tweets may be a useful additional data source following significant earthquakes, especially in regions without an extensive network of recording stations.

Introduction

The ability to rapidly estimate the distribution of shaking intensity across a large area following a significant earthquake is essential for efficient emergency response, loss estimation, and communication to the public [1,2]. Decision makers need accurate information as quickly as possible about the spatial distribution of damage so they can quickly send aid to appropriate locations. Previous study shows the importance of having accurate shaking intensity estimates for post-earthquake response, such as estimating the amount of international aid needed or prioritizing building inspections. However, getting information quickly is often challenging, especially if the area in which the earthquake occurred has few strong motion recording instruments.

Currently, ShakeMaps produced by the United States Geological Survey (USGS) represent the state-of-the-art in rapid shaking intensity estimation [1]. When an earthquake happens, a ShakeMap is typically produced in a matter of minutes using a combination of recordings, a simple ground motion prediction equation (GMPE), and geological site correction factors. As time progresses, the ShakeMap is continually updated as new information becomes available, including "did you feel it?" data, which are qualitative first-hand accounts of the earthquake collected via online surveys [3,4]. While this qualitative information is extremely

[3] PhD Candidate, Dept. of Civil Eng., Stanford University, Stanford, CA USA

[4] Consulting Professor, Institute of Computational and Mathematical Eng., Stanford University, Stanford, CA USA
* The first two authors contributed equally to this work.

informative, it can take anywhere from hours to days to weeks to collect and is therefore not typically included in very early shaking intensity estimates.

Other sources for this type of information are rapidly growing, particularly Twitter, which is a service that enables everyone to create and share ideas and information instantly, without barriers online in real-time. Twitter has reached more than 255 million active users worldwide, and including regions of high seismicity like the United States, Japan, Indonesia, Mexico, and the Philippines [5]. Researchers have recognized the potential of Tweets for improving earthquake response [6], and the USGS is investigating using Tweets for rapid earthquake detection [7]. Other event detection algorithms can estimate the location of the earthquake epicenter from Tweets [8] and suggest that Tweets within 10 minutes following an earthquake can provide a good approximation of the impacted area [9]. While these studies show great potential for using Tweets in earthquake detection and quickly identifying impacted areas, they do not provide a quantitative description of shaking intensity. Quantitative shaking intensity (measured on the Modified Mercalli Intensity scale, i.e. MMI) is instead provided via ShakeMaps, which do not yet include data from Tweets.

Here, we use Tweets that occurred in the 10 minutes following an earthquake along with simple earthquake characteristics to produce an early estimate of the shaking intensity across a region. We use three different learning algorithms, including linear, elastic net, and k-nearest neighbors regression, to find a model that minimizes the error between historical observations and our predicted value. The prediction is based on Tweet features like number of Tweets and length of Tweet in characters within certain radii of a location, simple earthquake characteristics like source-to-site distance and local site conditions. Using only these simple features, our model is able to predict shaking intensity with reasonable accuracy relative to the historical recordings, indicating that Tweets may aid in early intensity estimates and could potentially reduce uncertainty in existing models. The intent of this model is to show that Tweets contain relevant information that could be of use to refine current estimates following a significant earthquake, and is in no way intended as a replacement for current ShakeMaps.

Dataset

The data used for this project primarily came from Twitter and the USGS. We collected data over multiple earthquakes to build a set of observations, where each observation corresponds to an estimate of shaking intensity from a recording station during an earthquake in Japan and associated features from Twitter and the USGS. Therefore the data may contain multiple observations from the same earthquake, but at different locations.

Twitter Dataset

We used a dataset of Tweets following significant earthquakes in Japan. First, we compiled a list of 86 earthquakes with magnitude greater than 6 that occurred in Japan between 2011 and 2012, including an earthquake id, magnitude, epicenter location, and time at start of rupture [10]. Then we scraped all geo-tagged Tweets that occurred anywhere in the world in the 10 minutes following each of these earthquakes that included the keyword "earthquake" or "tsunami" in English, Japanese, Spanish, Turkish, or Tagalog, specifically *quake, terremoto, maremoto, gempa, 地震, 大震災, 震災, tsunami, じしん, つなみ, 津波, deprem, zelzele, and lindol. We included both original Tweets and Retweets, which are copies of Tweets posted by another user.

In just the 10 minutes after each of these 86 earthquakes that occurred in Japan, people created 6.6 million geo-tagged Tweets that used at least one of the earthquake-related keywords.

For each earthquake, we observed trends in the Tweets such as an increase in the rate of Tweets and constant locations of Tweets. These trends are demonstrated using the 2011 M9.0 Tohoku earthquake (i.e. earthquake id c0001xgp). The number of geo-tagged Tweets containing the earthquake keywords rapidly increased after the start of the event (Figure 1a), with people tweeting from across Japan, particularly in urban areas such as around Sendai, Tokyo, and Osaka (Figure 1c). By studying the distance between each Tweet and the epicenter as a function of time after the earthquake (Figure 1b), we clearly see urban centers: Sendai around 130km from the epicenter, the Tokyo region around 300-400km away, and Osaka around 730km away. Relatively little Tweet chatter about earthquakes occurred in the first 30 seconds because seismic waves from the offshore epicenter had not yet reached mainland Japan. Then, however, people started to feel strong shaking and the number of Tweets dramatically increases.
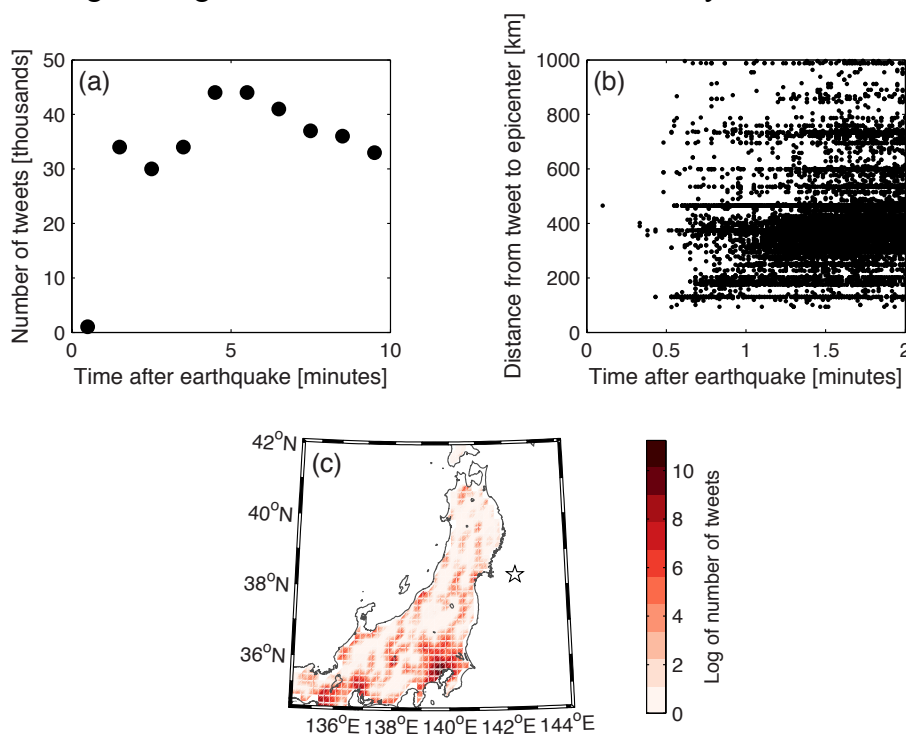

Figure 1.   For the Tohoku (c0001xgp) earthquake: (a) number of geo-tagged Tweets containing an earthquake keyword per minute after the event, (b) distance between each Tweet and the epicenter as a function of time, and (c) map showing the number of Tweets, where the star represents the epicenter location.

Shaking Intensity Dataset

We obtained the true shaking intensity for each Japanese earthquake at all recording stations, as reported by the USGS [1]. Shaking intensity is measured according to the Modified Mercalli Intensity scale (MMI), which has 12 levels between imperceptible shaking and catastrophic destruction. From the 86 Japanese earthquakes described in the Twitter dataset, there was a cumulative total of 7,627 observations of intensity from recording stations, where each of these observations represents an example in the dataset we used for analysis.

Features

We considered two classes of features: earthquake-based and Tweet-based. We defined 6 earthquake-based features including moment magnitude (M), source-to-site distance (R), shear-wave velocity in the top 30 meters (V_{S30}), community decimal intensity (CDI), latitude, and longitude. V_{S30} captures local site effects that may amplify shaking intensity. CDI is estimated using the simple GMPE defined in the baseline model section below and depends on M and R. Latitude and longitude are included to capture the site location because shaking intensity is often spatially correlated [11].

We supplemented the earthquake-based features with Tweet-based features, which is the main contribution of this work. Intuitively, a Twitter user is a noisy sensor because he or she may report an earthquake and the Tweet content may provide clues to what he or she experienced. Furthermore, all Tweets are time-stamped and all Tweets we considered are marked with location (i.e. latitude and longitude).

We defined 8 Tweet-based features, each considered within varying radii of a recording station, including count of Tweets, average number of tokens, average number of characters, average index of first mention of an earthquake keyword, average number of appearances of earthquake keywords, average number of exclamation points, average number of dashes, and average number of appearances of a selected Japanese keyword. The first 3 Tweet-based features capture general information about Tweets, where tokens were defined as a word, punctuation mark, or emoticon for alphabet-based languages, but simply each character for East Asian languages. The next 5 Tweet-based features are content-specific language markers. Exclamation points are often used to show emphasis or strong emotion in alphabet-based languages, and similarly dashes are used in East Asian languages. The Japanese keywords were selected by a native Japanese speaker as words used to indicate fear and other related earthquake terms, including 揺れ, 揺れてる, 避難, 余震, 怖い, 恐い, こわい, 無事, 震度. Each Tweet-based feature was considered at 9 different radii from the recording station (1km, 2km, 5km, 10km, 20km, 50km, 100km, and whole world), resulting in a total of 72 Tweet-based features. We also experimented with normalizing the Tweet counts by the average number of Tweets on days without any significant earthquakes, but we saw no significant impact on accuracy.

Each observation in our dataset corresponds to a single intensity estimate from a recording station during an earthquake in Japan. Therefore multiple observations may come from the same earthquake, but at different locations. For each observation, we built a feature vector with the 78 features described above. Some features are constant for different observations from the same earthquake, such as any Tweet-based feature calculated for the whole world, but this feature varies for different earthquakes.

While we have created a variety of features, not all features are necessarily "relevant" to the model and we expect some overfitting if too many features are included compared to the number of training observations. For each regression model described in the next section, we varied the number of features from one to all. Features with the best scores were selected based on cross-correlation between the other features and the desired value [12].

Models for Shaking Intensity

The goal was to develop one model that can reasonably estimate the shaking intensity for an

arbitrary location in Japan when given different features. Because shaking intensity is a numerical value, we focused on regression models. Our training set consists of actual MMI observations at recording stations during historical Japanese earthquakes and the features described in the previous section.

Baseline

The baseline approach to calculating the shaking intensity at a location uses a GMPE that is empirically fit from a set of historical earthquakes. In this paper, we use a GMPE for community decimal intensity (CDI) provided in the documentation for USGS "did you feel it?" post-earthquake community survey [13]. CDI was designed to quantify simple descriptions of ground shaking and is comparable to MMI [14]. The form of this GMPE is

$$CDI = 1.15 + 1.01M - 0.00054R - 1.72\log_{10} R \tag{1}$$

where M is the moment magnitude and R is the distance between the epicenter and the query location. We use this GMPE as the baseline model for simplicity, but it is not a direct comparison with early versions of USGS ShakeMaps, which typically include more sophisticated GMPEs such as [15] and site correction factors.

Regression Models

Instead of using the provided baseline equation, we demonstrate how to use the features described previously to train a new model. We considered three different models including linear regression, elastic net regression, and k-Nearest-Neighbors regression (kNN). Linear and elastic net regression approximate the desired value (y) as a linear combination of the given observations (x). For this application, the desired y-value is the natural logarithm of MMI (log MMI), which is consistent with GMPEs that typically model ground motion intensity as a lognormal distribution [16]. Elastic net also applies $l1$- and $l2$-regularization penalty terms to avoid overfitting, where ρ is the mixing parameter of $l1$ and $l2$ penalty [17]. Using the $l1$ penalty term, elastic net can perform additional feature selection by zeroing the weights of some features. Finally, kNN approximates the desired y-value by computing a distance metric between a given observation and all others, finding the k observations with smallest distance to the given observation, and then computing a weighted average of the y-values at the k closest neighbors [17]. The value of ρ for elastic net and k for kNN was chosen using modified 10-fold cross validation discussed in the next section.

Error metric

We used mean-squared error (MSE) between the true and predicted MMI (not log MMI) for all observations as a metric to select the best model and compare against the baseline. To compute MSE, we used a modified 10-fold cross validation approach. Normally, 10-fold cross validation randomly splits the observations into 10 subsets, trains the model on all but one subset, tests on the held-out subset, and is repeated 10 times. However, our dataset contains multiple observations from the same earthquake but at different locations, which are highly correlated to one another. Thus normal cross validation would be inappropriate because observations from the

same earthquake might be used to both train and test the model. So we implement a modified approach that divides the set of earthquakes into 10 subsets, trains on all observations except those from earthquakes in the held-out subset, tests on observations from earthquakes in the held-out subset, and is repeated 10 times. The final error of a model is the average of all MSE values over all 10 folds.

Results

We used the generated features for all observations along with the three regression models for two experiments. First, we selected and validated our models by comparing the testing error of predicted MMI with actual observations from stations in Japan. The goal was to find the choice of features and model that best estimated historical recordings. Second, we extrapolated the model to other locations to produce an estimated MMI map 10 minutes after three case study earthquakes. This MMI map can provide an early estimate of shaking intensity as well as supplement existing USGS ShakeMap estimates of MMI at sites far away from recording stations. The goal was to provide a realistic estimate of MMI at all locations impacted by an earthquake, even where there are no recording stations.

Model Selection and Validation

To choose a model that best estimates historical MMI observations, we varied the model, model parameters, and features. We considered the baseline model, linear regression, elastic net regression, and kNN-regression. As for model parameters, for linear we varied no parameters, for elastic net we varied ρ (the mixing parameter of *l1* and *l2*-regularization) between 0 and 1, and for kNN we varied k (the number of neighbors) between 2 and 80. For features, we used three sets for comparison: tweet-based features only, earthquake-based features only, and combined tweet- and earthquake-based features. We also considered selecting between 1 and all features for each set. For each model, choice of parameter, and target number of features, we selected features and fit the models using the python package sklearn [18].

The baseline model (simple GMPE only) had the highest error, models using only Tweet-based features had relatively large error compared to models using only earthquake-based features, and models using a combination of earthquake-based and Tweet-based features, i.e., selecting from all features, generally performed best (Figure 2a). We found that elastic net regression using all features performed best over the whole dataset with a mean-squared error of 0.99 using modified 10-fold cross validation. This result indicates a reasonable estimate of the MMI because MMI values range from 1 to 12. Furthermore, our model showed a significant reduction over the baseline model (MSE of 0.99 vs. 5.57, respectively).

From modified cross validation, we found the best choice of elastic net model included 24 features and $\rho=0.5$. As the number of features increased, we observed a reduction in testing error, leading eventually to overfitting (Figure 2b). The step-like behavior indicates that additional features are set with zero weights until enough are added to make a step in error. For the chosen model, the exact weights on the selected features vary depending on the cross validation fold. However, we find that all models are magnitude-independent. Specifically, most folds have the following features with non-zero weights: number of Tweets at radii of 5km, 10km, 20km, 50km, 100km, and 1000km, number of characters at radius 100km, source-to-site distance (R), and V_{S30}.
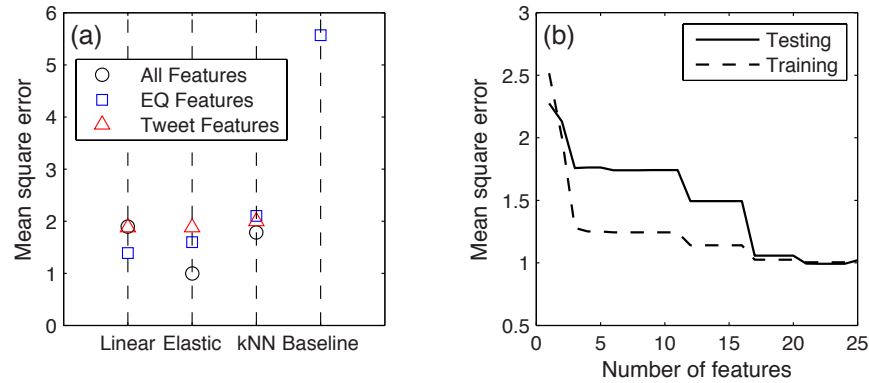
Figure 2. (a) Mean-squared error in the MMI value (on a scale of 1 to 12) for the baseline, linear regression, elastic net regression, and kNN-regression models using modified 10-fold cross validation, and trained using earthquake- and Tweet-based features (i.e. All Features), earthquake-based features only, or Tweet-based features only. (b) Variation in training and test error with increasing features for elastic net regression with all features where the mixing parameter, ρ, equals 0.5.

Since the best model overall, which selects from both Tweet- and earthquake-based features, has a lower error than the best model using only tweet-based features and the best model using only earthquake-based features (MSE=0.99 versus 1.39 and 1.88, respectively), we conclude that combining Tweets and earthquake-based data leads to the best estimate. Furthermore, by examining the selected features, we see that readily available Tweet count-based features provided useful information. In other words, knowing the sheer number of people tweeting at a location about earthquakes aids in predicting how intensely an earthquake impacted a region. Furthermore, the number of relevant Tweet around a location has more predictive power than magnitude according to our model.

Predicted MMI Maps

We used the elastic net regression model and corresponding features to estimate MMI maps for three historical events:

- M9.0, March 2011 Tohoku earthquake (c0001xgp), which had observations from 1194 recording stations. This earthquake represents an extreme case to test the limits of our model as well as a well-studied event for validation.
- M7.1, April 2011 earthquake (c0002ksa), which had observations from 922 recording stations. This earthquake represents a more likely event where the actual shaking intensity was well recorded.
- M6.4, July 2011 earthquake (c000531r), which had observations from only 9 recording stations. This earthquake represents a case where additional data such as from Tweets is particularly relevant.

First, we considered the Tohoku earthquake (Figure 3). By visual inspection, the elastic net regression model using both Tweet- and earthquake-based features produced the MMI map most representative of the USGS ShakeMap. This Shakemap is particularly robust because it was updated over 5 months as more recordings, human reports of intensity ("did you feel it?" data),

and fault models were incorporated. By comparing the MMI map produced using only earthquake-based features to both Tweet- and earthquake-based features, we saw that the Tweet-based features effectively scaled the entire MMI map obtained using earthquake-based features only (Figure 3a,c). For example, a similar spatial pattern in Western Japan exists for both, but with different intensities. This scaling is particularly important for high intensity events where existing models break down, but people, as captured by Tweets, add valuable information. In comparison, Tweet-based features alone, particularly as shown here without being normalized based on population or normal Twitter activity, were poor predictors of shaking intensity. Thus, compared to the baseline model and models using only earthquake- or Tweet-based features, the proposed method captured spatial variation and reduced error across a region.
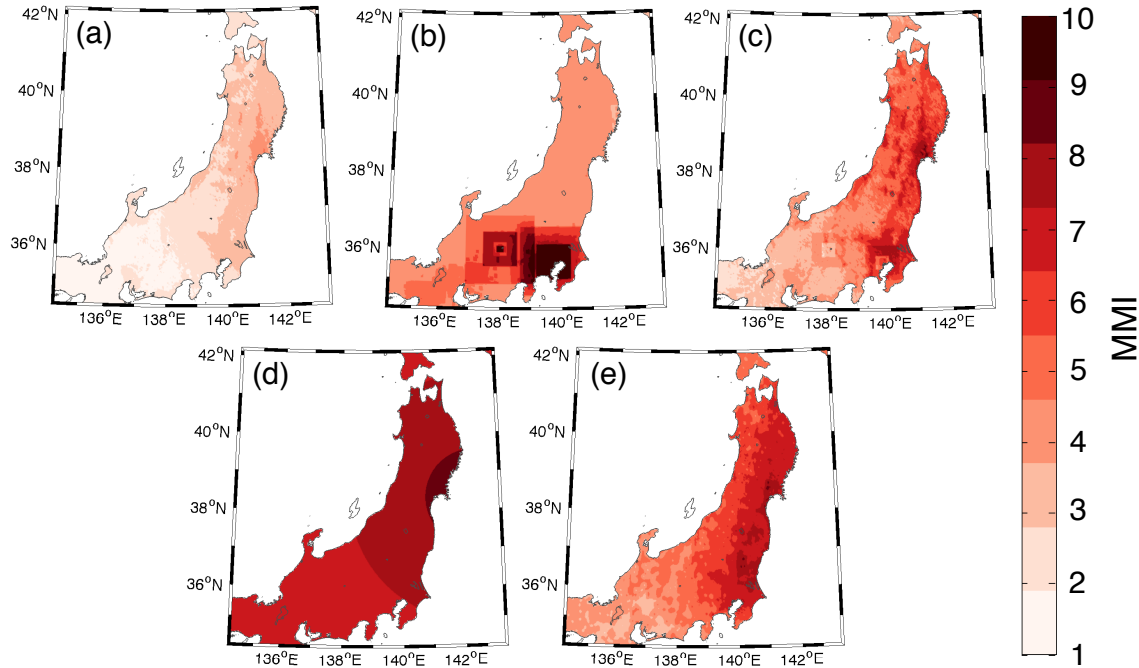

Figure 3. Estimated MMI maps for the Tohoku earthquake (c0001xgp) by (a) elastic net regression using earthquake-based features only, (b) elastic net regression using Tweet-based features only, (c) elastic net regression using both earthquake- and Tweet-based features, (d) baseline model, and (e) USGS ShakeMap.

Second, we compared MMI estimates from the baseline model, best elastic net regression model, and USGS ShakeMap for all three case study earthquakes (Figures 3 and 4). As expected from the MSE results, the baseline model typically missed the local variations in intensity present in ShakeMaps. We saw that in all three events, our best elastic net regression model accurately captured some local effects, such as the North to South streak of higher MMI seen in Figure 3a,c,e and Figure 4b,c,e that coincides with the Kitikami River basin, marked by both lower V_{S30} values (which amplifies shaking intensity) and more people (i.e. more Tweets). The MMI maps using our best elastic net regression were reasonably representative of the ShakeMaps for Tohoku and c0002ska, which include many observations of MMI from recordings stations and thus have relatively low uncertainty. The c000531r earthquake was an event with few observations from recording stations, so the Shakemap more closely resembles the baseline model. In contrast, our elastic net model showed more local variations in intensity because we included V_{S30} in addition to Tweet-based features.
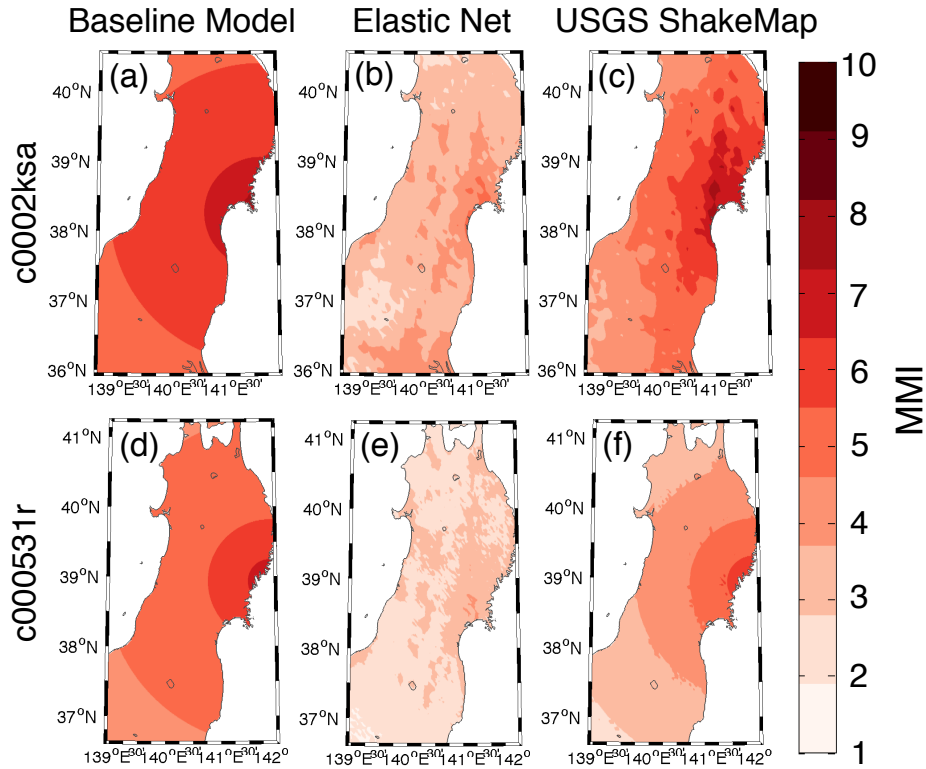
Figure 4.   Estimated MMI maps using different models for different earthquake events. Each row represents an earthquake event and each column represent a different model. Elastic Net model is calculated with Tweet-based and earthquake-based features.

Conclusions

We have presented a regression model using Tweets that occurred in the 10 minutes following an earthquake coupled with simple earthquake characteristics to estimate the shaking intensity (MMI) at particular sites in Japan. The key contribution of our work was to explore Tweets as an additional data source for estimating shaking intensity that is available almost immediately after an earthquake and can supplement existing methods. We have shown that counts of geo-tagged Tweets at different radii around a particular site are good predictors of shaking intensity when used in combination with earthquake-based features such as V_{S30} and source-to-site distance. In practice, this is important because additional data sources such as on-the-ground inspections by engineers, measurements from recording stations not connected to the internet, and online surveys such as "did you feel it?" may not be available before some decisions about relief aid need to be made. While we have estimated MMI values across a region using only simple earthquake characteristics and Tweet-based features for illustration, this data could potentially be combined with existing data sources, such as initial measurements from recording stations, for a more refined shaking intensity estimate with lower uncertainty. Finally, while we have shown results for Japan as a case study, our results depend neither on an earthquake fault model nor on results from pre-existing earthquake recording stations, so with some tuning of the model based on baseline Twitter activity, this approach could be applied to other regions. The approach has significant potential for regions that may not have an extensive network of existing earthquake recording stations but have many active Twitter users.

References

[1] Wald DJ, Worden BC, Quitoriano V, Pankow KL. ShakeMap Manual: Technical Manual, Users Guide, and Software Guide. 2006; US Geological Survey; Boulder, CO.

[2] Wald D, Lin K-W, Porter K, Turner L. ShakeCast: Automating and Improving the Use of ShakeMap for Post-Earthquake Decision-Making and Response. *Earthquake Spectra* 2008; **24** (2): 533–553.

[3] Wald DJ, Quitoriano V, Dengler LA, Dewey JW. Utilization of the Internet for Rapid Community Intensity Maps. *Seismological Research Letters* 1999; **70** (6): 680–697.

[4] Atkinson GM, Wald DJ. "Did You Feel It?" Intensity Data: A Surprisingly Good Measure of Earthquake Ground Motion. *Seismological Research Letters* 2007; **78** (3): 362–368.

[5] Campagne J-C, Dux J, Guyot P, Julien D. Twitter reaches half a billion accounts - More than 140 million in the U.S. *Semiocast* 2012; Paris, France. Available from: http://semiocast.com/publications

[6] Earle P, Guy M, Buckmaster R, Ostrum C, Horvath S, Vaughan A. OMG Earthquake! Can Twitter Improve Earthquake Response? *Seismological Research Letters* 2010; **81** (2): 246–251.

[7] Earle PS, Bowden DC, Guy M. Twitter earthquake detection: earthquake monitoring in a social world. *Annals of Geophysics* 2011; **54** (6): 708–715.

[8] Sakaki T, Okazaki M, Matsuo Y. Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. *Proceedings of the World Wide Web Conference (WWW)* 2010. Raleigh, NC. p. 851–860.

[9] Crooks A, Croitoru A, Stefanidis A, Radzikowski J. #Earthquake: Twitter as a Distributed Sensor System. *Transactions in GIS* 2012; **17** (1): 124–147.

[10] Taskan G, Knappe N. Seismi API. http://www.seismi.org/api/

[11] Jayaram N, Baker JW. Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering & Structural Dynamics* 2009; **38** (15): 1687–1708.

[12] Hall MA. Correlation-based feature selection for machine learning. PhD Dissertation 1999. University of Waikato.

[13] United States Geological Survey. Did you feel it? The science behind the maps. *Earthquake Hazards Program* 2013; http://earthquake.usgs.gov/research/dyfi/

[14] Dengler LA, Dewey JW. An intensity survey of households affected by the Northridge, California, earthquake of 17 January 1994. *Bulletin of the Seismological Society of America* 1998; **88** (2): 441–462.

[15] Youngs RR, Chiou SJ, Silva WJ, Humphrey JR. Strong ground motion attenuation relationships for subduction zone earthquakes. *Seismological Research Letters* 1997; **68**: 58–73.

[16] Abrahamson N, Atkinson G, Boore D, Bozorgnia Y, Campbell K, Chiou B, Idriss IM, Silva W, Youngs R. Comparisons of the NGA Ground-Motion Relations. *Earthquake Spectra* 2008; **24** (1): 45–66.

[17] Cover TM. Estimation by the nearest neighbor rule. *IEEE Transactions on Information Theory* 1968; **14** (1): 50–55.

[18] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 2011; **12**: 2825–2830.