OPINION

Predicting or Forecasting Earthquakes and the Resulting Ground-Motion Hazards: A Dilemma for Earth Scientists

Recent earthquakes, particularly the 2008 Wenchuan (China), 2009 L'Aquila (Italy), 2010 Haiti, 2011 Christchurch (New Zealand), and 2011 Tohoku (Japan), events, have renewed debate among scientists (e.g., Jordan *et al.*, 2011; Stein *et al.*, 2011; Peresan *et al.*, 2012; Stirling, 2012; Jordan, 2013; Jordan *et al.*, 2014; Wang and Rogers, 2014) on predicting or forecasting earthquakes and their resulting ground-motion hazards. The main reason for this debate is societal demand for predicting or forecasting. In other words, scientists, seismologists in particular, feel compelled to provide predictions or forecasts to society. The initial sentencing of six years in prison for six Italian scientists following the 2009 L'Aquila earthquake demonstrates this. However, it is well understood that earthquakes and their resulting ground-motion hazards cannot be predicted or forecasted reliably. Thus, the statement by Jordan *et al.* (2014, p. 959), "though communicating OEF (operational earthquake forecast) and its uncertainties is a difficult issue, not communicating is hardly an option," illustrates the dilemma.

Scientists, seismologists in particular, feel compelled to provide earthquake predictions or forecasts to society. The differences between predicting and forecasting have to do with how they quantify and communicate uncertainty.

EARTHQUAKE PREDICTION VERSUS FORECASTING

Location, magnitude, and recurrence interval of earthquakes have large uncertainty. For example, the estimated magnitudes for earthquakes in the New Madrid Seismic Zone range from **M** 6.6 to **M** 8.0, and estimates of the recurrence interval range from 500 to 50,000 years (Petersen *et al.*, 2014). Uncertainties of this scale indicate that earthquakes cannot be predicted or forecasted reliably. One such example is Iben Browning's forecast: a 50% chance of a major earthquake with a magnitude of about 7 in the New Madrid Seismic Zone within a few days of 3 December 1990 (Stein, 2010). Even though the 1975 Haicheng, China, earthquake has been claimed as a successful prediction, it was not predicted scientifically (Wang *et al.*, 2006).

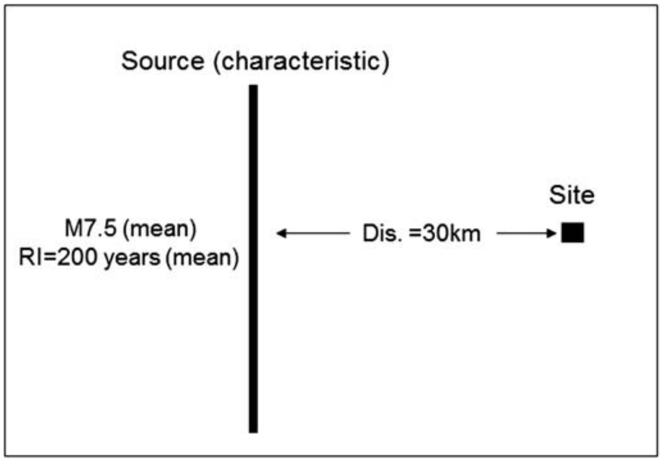
According to Jordan *et al.* (2011), a "prediction" is a prospective deterministic statement (an assertion that one or more earthquakes will occur in a specified subdomain of space and future time), whereas a "forecast" is a prospective probabilistic statement (a probability that earthquakes will occur in the space–time subdomains). Jordan *et al.* (2011, p. 328) also said, "statements about future earthquakes are inherently uncertain, and no forecast or prediction can be complete without a description of this uncertainty." Thus, there is no qualitative difference between deterministic predicting and probabilistic forecasting; because "uncertainty is expressed in terms of probabilities, both deterministic predictions and probabilistic forecasts need to be stated and evaluated using probabilistic concepts" (Jordan *et al.*, 2011, p. 328). Therefore, the differences between predicting and forecasting have to do with how they quantify and communicate uncertainty, as illustrated in Figure 1.

Generally, the location, magnitude, and recurrence interval of earthquakes and their respective uncertainties are quantified by a probability model (distribution), such as a Gaussian (normal) model with a mean and standard deviation or by a logic tree (e.g., Petersen *et al.*, 2014). A mean magnitude of **M** 7.5 and mean recurrence interval of 200 years were assumed for the characteristic fault in Figure 1. The prediction for this case may be that an **M** 7.5 earthquake "could occur" along the fault in the next month, next year, or in 50 years or that an **M** 7.5 earthquake "will probably occur" along the fault in the next month, next year, or in 50 years. To make a forecast, a probability model has to be introduced to describe earthquake occurrence in time (e.g., the Poisson, empirical, Brownian passage time, or time predictable). Although the Poisson model (i.e., time independent) contradicts the generally accepted physical model (i.e., Reid's elastic rebound theory), it is the most commonly used model for estimating earthquake probability (e.g., Jordan *et al.*, 2011). The Poisson model assumes the exceedance probability (PE) of the **M** 7.5 earthquake occurrence along the fault over a specified time period (t) can be estimated by

$$\mathrm{PE} = 1 - e^{-\frac{t}{\tau}} \approx \frac{t}{\tau}, \tag{1}$$

in which τ is the mean recurrence interval of the earthquake. For time periods of 1 month, 1 year, and 50 years, equation (1) yields PEs of about 0.042%, 0.5%, and 22.1%, respectively. Thus, the forecast that an **M** 7.5 earthquake will occur along the fault within the next month is 0.042%; the probability that an **M** 7.5 earthquake will occur within the next 50 years is 22.1%.

doi: 10.1785/0220140211



▲ **Figure 1.** A characteristic fault with an earthquake having mean magnitude of **M** 7.5 and mean recurrence interval of 200 years.

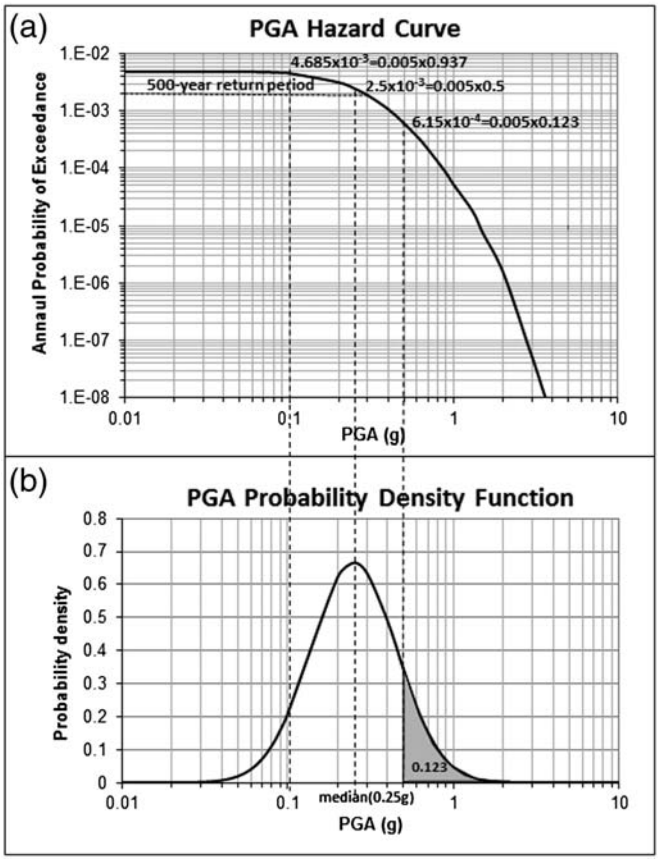
GROUND-MOTION PREDICTING VERSUS FORECASTING

To predict or forecast a ground motion at a site, a ground-motion attenuation model—also called a ground-motion prediction equation (GMPE)—is needed. A GMPE is assumed to follow a lognormal distribution and has the mathematical form

$$\ln(Y) = f(\mathbf{M}, R) + \delta, \qquad (2)$$

in which Y is a median ground motion, **M** is earthquake magnitude, R is the source-to-site distance, and δ is uncertainty or residual and follows a normal distribution with a standard deviation of σ (Fig. 2b). Currently, deterministic seismic-hazard analysis (DSHA) is most commonly used for predicting ground motion, and probabilistic seismic-hazard analysis (PSHA) is used for forecasting. DSHA and PSHA use the same seismological and statistical data but define and calculate ground-motion hazard differently. The differences between DSHA and PSHA can be illustrated by predicting and forecasting the ground motion at a site 30 km from the characteristic fault shown in Figure 1. A median peak ground acceleration (PGA) of 0.25g and standard deviation of 0.6 (in natural logarithms) from the **M** 7.5 earthquake were assumed for the site.

DSHA determines seismic hazard as the ground motion of a certain percentile from single or a set of earthquakes that have maximum impact (Krinitzsky, 1995). For a single characteristic source, ground-motion prediction using DSHA is simple and straightforward. Using DSHA, the prediction for the site might be that a median peak ground acceleration (PGA) of 0.25g or a PGA of 0.46g (84th percentile) could be expected if the **M** 7.5 earthquake occurs along the fault or that a median PGA of 0.25g or a PGA of 0.46g (84th percentile) could be expected if the **M** 7.5 earthquake occurs along the fault in the next month, next year, or in 50 years.

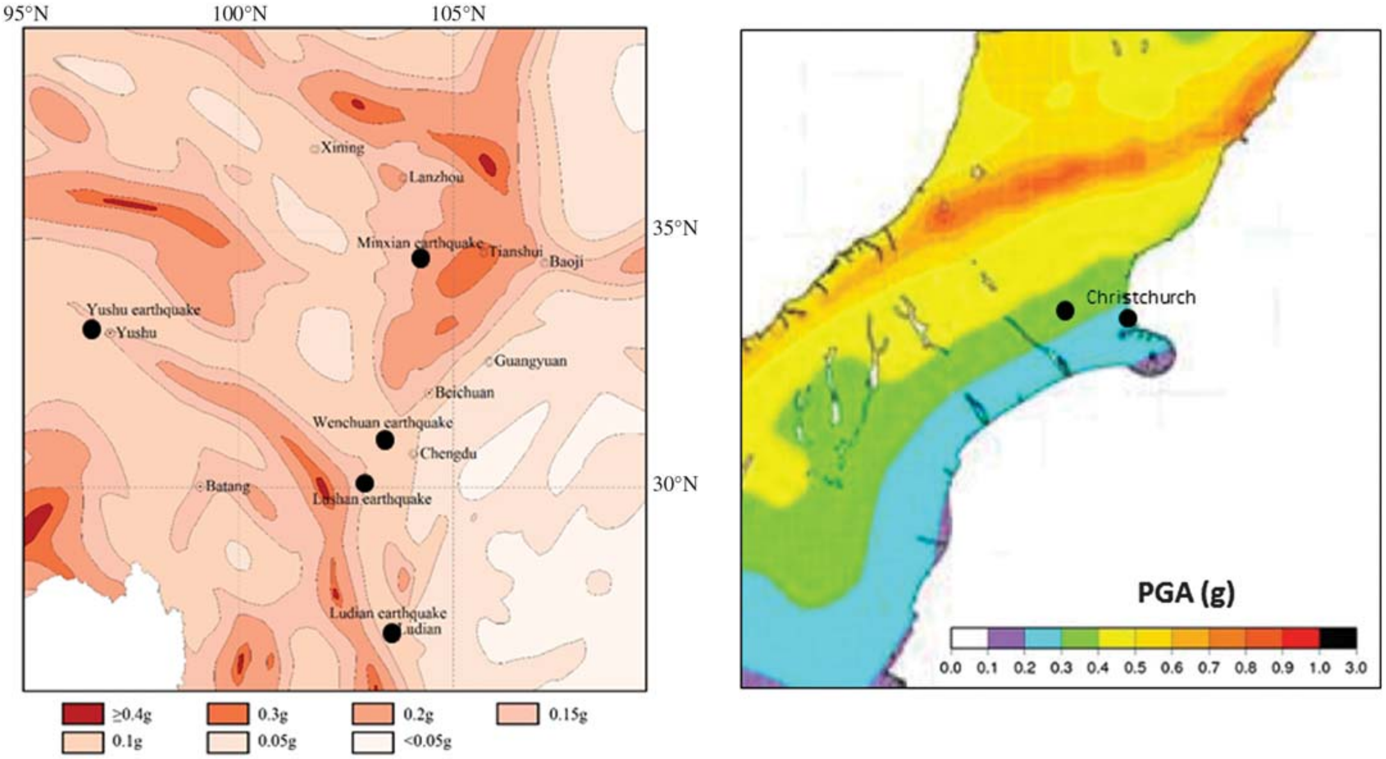


▲ **Figure 2.** (a) Peak ground acceleration (PGA) hazard curve for a site 30 km from the characteristic fault (Fig. 1) and (b) PGA uncertainty distribution (i.e., a lognormal distribution). A median PGA of 0.25g and standard deviation of 0.6 (in natural logarithms) were assumed.

PSHA determines seismic hazard as the annual PE (i.e., the PE in 1 year) for a given ground motion and calculates it from a probability analysis (Cornell, 1968). To determine the probability, several probability models have to be assumed for describing the occurrences of earthquake and the resulting ground motion in time and space. According to Cornell (1968), under the assumptions of (1) equal likelihood of earthquake occurrence (single point) along a line or over an areal source, (2) a constant-in-time average occurrence rate of earthquakes, and (3) a Poisson distribution for earthquake occurrence in time, the annual PE for a given ground motion y from a single characteristic source is

$$P_a[Y \geq y] \approx \frac{t(1\ \text{year})}{\tau} \times \left[1 - \Phi\left(\frac{\ln y - \ln y_{mr}}{\sigma}\right)\right], \qquad (3)$$

in which $\ln y_{mr} = f(m, r)$, 1–$\Phi(x)$ is the PE for ground motion. As shown in equation (3), the annual PE is equal to the annual PE for the earthquake, determined from equation (1) for $t = 1$ yr (annual), times the PE for ground motion. For example, the annual PE (0.0025) for PGA of 0.25g is equal to the annual PE (1/200 or 0.005) for the earthquake times the PE (0.5) for the ground motion (Fig. 2a). Thus, using PSHA, the forecast could be that the annual PE for PGA of 0.25g at the site is 0.25% or that the probability the ground motion exceeds 0.25g PGA in one year is 0.25%.



▲ **Figure 3.** PGA maps with 500-year return periods (10% PE in 50 years) for (a) southwestern China (PRCNS, 2001) and (b) Christchurch, New Zealand (Stirling *et al.*, 2012). The solid circles indicate the locations of recent damaging earthquakes.

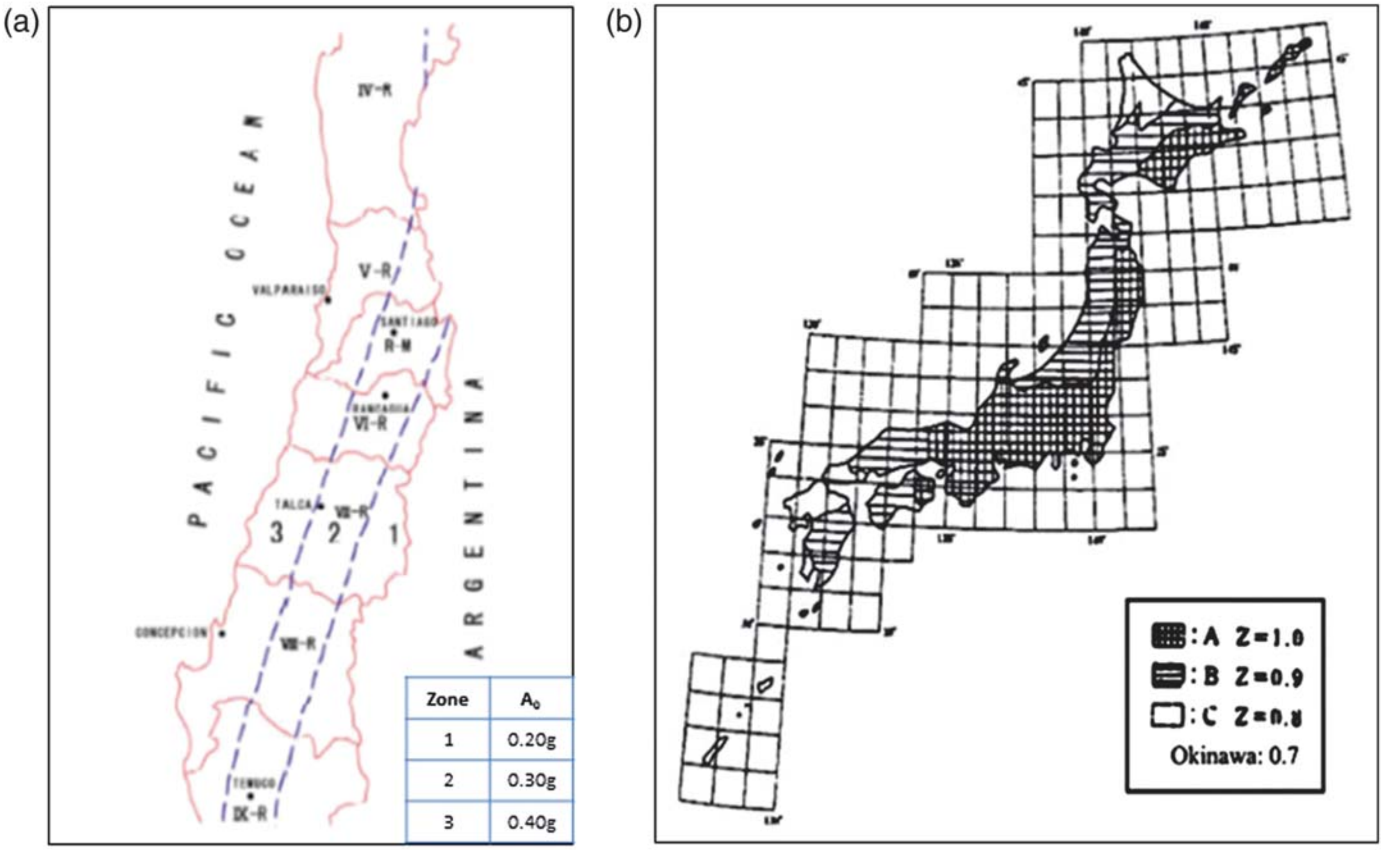
PROBABILISTIC FORECAST OR DETERMINISTIC PREDICTION?

"Which is better for quantifying and communicating earthquakes and their resulting ground-motion hazards: probabilistic forecasting or deterministic predicting?" A critical issue is transparency (e.g., Jordan *et al.*, 2014; McNutt, 2014). As demonstrated here, deterministic prediction quantifies and communicates the key scientific information (i.e., mean, median, or a certain percentile) in a way that is transparent and easy to understand by users. Probabilistic forecasting, however, quantifies and communicates the probability in a way that is compounded from many assumed probability models, some of which are bad (Stein *et al.*, 2011) or even all wrong (Field, 2013). Particularly in PSHA, the probability (i.e., the annual PE) has been erroneously interpreted as the frequency or rate of ground-motion occurrence and used to numerically generate another PE, such as 10%, 5%, and 2% PEs in 50 years, under another Poisson assumption (Wang, 2011, 2012). Thus, deterministic prediction is more transparent and physically comprehensible than probabilistic forecasting.

Recent worldwide observations illustrate the problems with probabilistic forecasting, and with PSHA in particular. Figure 3a shows the 2001 Chinese national PGA hazard map with 10% PE in 50 years for the southwest region of China (People's Republic of China National Standard [PRCNS], 2001). The locations of recent earthquakes, including the 2008 Wenchuan, 2010 Yushu, 2011 Lushan and Minxian, and 2014 Ludian earthquakes are also shown. All the earthquakes occurred in areas with low forecasted PGA hazards. Similarly, in New Zealand, the 2010 Darfield and 2011 Christchurch earthquakes occurred in an area with relative low hazards forecasted by the New Zealand national hazard map, with 10% PE in 50 years (Stirling *et al.*, 2012) (Fig. 3b). In contrast, Figure 4 shows seismic design zone maps for Chile and Japan. There was no significant building and infrastructure damage in Chile during the 2010 Maule earthquake (**M** 8.8) or in Japan during the 2011 Tohoku earthquake (**M** 9.0). It should be noted that seismic design ground motion in coastal California is deterministic ground motion from the maximum considered earthquake, not the probabilistic ground motion from PSHA (American Society of Civil Engineers [ASCE], 2010). These worldwide observations show that deterministic prediction performs better than probabilistic forecasting in helping society to mitigate earthquake impacts and avoid disaster.

CONCLUSION

Deterministic prediction and probabilistic forecast use the same statistical data but quantify and communicate uncertainty differently. A prediction implicitly communicates uncertainty without a quantified probability, whereas a forecast explicitly communicates uncertainty with a quantified probability. Several probability models have to be introduced to make a forecast. For some users, such as the insurance industry, a numerical probability might be preferred. In other words, some of the models might be useful. However, the probabilities are not reliable, because the models are bad or even all wrong. Introducing more probability



Chilean seismic zone (NCh433 1996)

Seismic zone factor in Japan (Kuramoto, 2006).

▲ **Figure 4.** (a) Chilean national seismic design zones (Instituto Nacional de Normalización [INN], 1996) and (b) Japanese national seismic design zones (Kuramoto, 2006).

(mathematical) models adds more layers of uncertainty and leads to more difficulty in communicating and understanding the results. This is particularly true for PSHA, because the annual PE (a dimensionless quantity) has been erroneously communicated and used as the annual frequency or rate of exceedance (a dimensional quantity with the unit of 1/yr).

Although debate about probabilistic forecasting versus deterministic predicting will continue among scientists, hazard mitigation policies must be improved to reduce the risk of another disaster like Wenchuan, L'Aquila, Haiti, or Christchurch. All scientists, seismologists in particular, must quantify and communicate about earthquakes and their resulting hazards in a transparent, scientifically defensible, and understandable way. The critical issue is how to quantify and communicate the inherent uncertainty. As demonstrated here, deterministic prediction is more transparent and easier to understand than probabilistic forecasting, because more layers of uncertainty are introduced to make probabilistic forecasts. Worldwide observations show that deterministic prediction performs better than probabilistic forecasting.

Although debate about probabilistic forecasting versus deterministic predicting will continue among scientists, hazard mitigation policies must be improved to reduce the risk of another disaster like Wenchuan, L'Aquila, Haiti, or Christchurch. All scientists, seismologists in particular, must quantify and communicate about earthquakes and their resulting hazards in a transparent, scientifically defensible, and understandable way.

ACKNOWLEDGMENTS

I appreciate comments and suggestions from Giuliano Panza and Seth Carpenter and review comments from Seth Stein and an anonymous reviewer, which helped to improve the manuscript greatly. I thank Meg Smath of the Kentucky Geological Survey for editorial help.

REFERENCES

American Society of Civil Engineers (ASCE) (2010). *Minimum Design Loads for Buildings and Other Structures: ASCE/SEI 7-10*, ASCE, Reston, Virginia, 608 pp.

Cornell, C. A. (1968). Engineering seismic risk analysis, *Bull. Seismol. Soc. Am.* **58,** 1583–1606.

Field, E. H. (2013). The Uniform California Earthquake Rupture Forecast version 3—Key assumptions and testability, *the 2013 Seismological Society of America Annual Meeting Presentation (abstract)*, 17–19 April 2013, Salt Lake City, Utah, http://www.seismosoc.org/members/annual-meetings/2013/presentation.php?id=13-174 (last accessed September 2014).

Instituto Nacional de Normalización (INN) (1996). *Earthquake Resistant Design of Buildings*, Official Chilean Code NCh433. Of96.

Jordan, T. (2013). Lessons of L'Aquila for operational earthquake forecasting, *Seismol. Res. Lett.* **84,** 4–7.

Jordan, T., Y. Chen, P. Gasparini, R. Madariaga, I. Main, W. Marzocchi, G. Papadopoulos, G. Sobolev, K. Yamaoka, and J. Zschau (2011). Operational earthquake forecasting: State of knowledge and guidelines for utilization, *Ann. Geophys.* **54,** no. 4, 316–391, doi: 10.4401/ag-5350

Jordan, T., W. Marzocchi, A. J. Michael, and M. C. Gerstenberger (2014). Operational earthquake forecasting can enhance earthquake preparedness, *Seismol. Res. Lett.* **85,** 955–959.

Krinitzsky, E. L. (1995). Deterministic versus probabilistic seismic hazard analysis for critical structures, *Eng. Geol.* **40,** 1–7.

Kuramoto, H. (2006). Seismic design codes for buildings in Japan, *J. Disaster Res.* **1,** 341–356.

McNutt, M. (2014). Journals unite for reproducibility, *Science* **346,** no. 6210, 679.

People's Republic of China National Standard (PRCNS) (2001). *Seismic Ground Motion Parameter Zonation Map of China*, scale 1:8,000,000, GB 18306–2001.

Peresan, A., V. G. Kossobokov, and G. F. Panza (2012). Operational earthquake forecast/prediction, *Rendiconti Lincei* **23,** 131–138, doi: 10.1007/s12210-012-0171-7.

Petersen, M. D., M. P. Moschetti, P. M. Powers, C. S. Mueller, K. M. Haller, A. D. Frankel, Y. Zeng, S. Rezaeian, S. C. Harmsen, O. S. Boyd, N. Field, R. Chen, K. S. Rukstales, N. Luco, R. L. Wheeler, R. A. Williams, and A. H. Olsen (2014). Documentation for the 2014 update of the United States national seismic hazard maps, *U.S. Geol. Surv. Open-File Rept. 2014–1091*, 243 pp.

Stein, S. (2010). *Disaster Deferred*, Columbia University Press, New York, New York, 282 pp.

Stein, S., R. Geller, and M. Liu (2011). Bad assumptions or bad luck: Why earthquake hazard maps need objective testing, *Seismol. Res. Lett.* **82,** 623–626.

Stirling, M. (2012). Earthquake hazard maps and objective testing: The hazard mapper's point of view, *Seismol. Res. Lett.* **83,** 231–232.

Stirling, M., G. McVerry, M. Gerstenberger, N. Litchfield, R. VanDissen, K. Berryman, P. Barnes, L. Wallace, P. Villamor, R. Langridge, G. Lamarche, S. Nodder, M. Reyners, B. Bradley, D. Rhoades, W. Smith, A. Nicol, J. Pettinga, K. Clark, and K. Jacobs (2012). National seismic hazard model for New Zealand: 2010 update, *Bull. Seismol. Soc. Am.* **102,** 1514–1542.

Wang, K., and G. C. Rogers (2014). Earthquake preparedness should not fluctuate on a daily or weekly basis, *Seismol. Res. Lett.* **85,** 569–571.

Wang, K., Q. Chen, S. Sun, and A. Wang (2006). Predicting the 1975 Haicheng earthquake, *Bull. Seismol. Soc. Am.* **96,** 757–795.

Wang, Z. (2011). Seismic hazard assessment: Issues and alternatives, *Pure Appl. Geophys.* **168,** 11–25.

Wang, Z. (2012). Comment on "PSHA Validated by Quasi Observational Means" by R. M. W. Musson, *Seismol. Res. Lett.* **83,** 714–716.

Zhenming Wang
Kentucky Geological Survey
504 Rose Street
Lexington, Kentucky 40506 U.S.A.
zmwang@uky.edu