Mak, S., Cotton, F., Schorlemmer, D. (2017): Measuring the Performance of Ground-Motion Models: The Importance of Being Independent. - *Seismological Research Letters*, *88*, pp. 1212—1217.

DOI: http://doi.org/10.1785/0220170097

SSA

THE SEISMOLOGICAL SOCIETY OF AMERICA
400 Evelyn Ave., Suite 201
Albany, CA 94706-1375
(510) 525-5474; FAX (510) 525-7204
www.seismosoc.org

OPINION

Measuring the Performance of Ground-Motion Models: The Importance of Being Independent

ABSTRACT

The results of data-driven evaluation of ground-motion models could be ambiguous if the test data are not independent of all the evaluated models. In such a case, the results describe both the explanatory and the predictive powers of the models. As an example, we demonstrate how a superseded ground-motion model appears to perform better than its successor, an anti-intuitive result. We hope to raise the awareness of the seismic-hazard community on the importance of data independence when conducting and interpreting a data-driven evaluation of ground-motion models. The evaluation can still be useful even if test data cannot be made entirely independent. but the result should be interpreted with care.

DATA-DRIVEN EVALUATION: EXPLANATORY OR PREDICTIVE?

Selecting one or multiple ground-motion models (GMMs) suitable for a seismic-hazard model is a consequential task for a seismic-hazard modeler. GMMs are often selected by a group of experts after carefully evaluating the merits and shortcomings of each shortlisted model (not too different from the process of filling an important tenure-track position, albeit reference letters may not be available). Criteria have been proposed to guide this selection process (Cotton *et al.*, 2006; Bommer *et al.*, 2010). Recently, data-driven (i.e., empirical) GMM evaluations have become popular (see Table 1 for an incomplete survey; see also table 1 of Mak *et al.*, 2017). Data-driven GMM evaluations could inform the selection process by providing objective evidence about the performance of a GMM. Some studies of this kind aim at highlighting the difference between models to help the hazard modeler to capture epistemic uncertainty; others aim at identifying reasons for good/poor model performance that will lead to model improvements.

It is the predictive power of a GMM, not its explanatory power, that is relevant to seismic hazard (Bindi, 2017). A model that explains the physics well does not necessarily predict the future well (Shmueli, 2010). Therefore, a strict GMM evaluation should ideally be based only on prospective data (i.e., data collected after the GMM is created; see Schorlemmer and Gerstenberger, 2007, for examples of prospective test). Ground-motion observations are always scarce. Meeting the strict definition of prospectiveness (as used in, say, the Collaboratory for the Study of Earthquake Predictability; Jordan, 2006) is difficult. Data that are independent of the GMM (i.e., data not used in creating the GMM, although the earthquakes may occur before the GMM is created) are a pragmatic choice to demonstrate the predictive power of a GMM.

The applicability of foreign models to a local region is often the focus of a data-driven GMM evaluation (Table 1) because models specifically developed for a region are often not available. For studies of this purpose, data independence is less controversial because the models and the test data are naturally independent. Similarly, some studies compare the ground-motion observations with the predictions from simulation-based stochastic GMMs. If the parameters of the GMMs are prescribed instead of empirically derived, then the issue of data independence is also irrelevant. It is not uncommon, however, that the test data have been used in developing some of the candidate GMMs (Table 1). The interpretation of the results of this kind of evaluation could be problematic because the resulting performances of some of the candidate models represent their predictive powers (i.e., for independent data), whereas those for other candidate models represent their explanatory powers (i.e., for data used to develop the model). It is inappropriate to directly compare these two types of performance.

In this article, we demonstrate, by an example based on real data and models, what could happen if the issue of data independence is ignored. We hope to raise the awareness of the seismic-hazard community on this issue when conducting evaluation studies and interpreting the results of such studies.

EXAMPLE: PREDECESSOR VERSUS SUCCESSOR

Among the five GMMs developed under the Next Generation Attenuation (NGA) project (2008), four of them are updated versions of GMMs used in the 1996 version of the U.S. Geological Survey National Seismic Hazard Model. For example, Boore and Atkinson (2008; hereafter, 2008BA) is an update of Boore *et al.* (1997; hereafter, 1997BJF). The NGA-West2

doi: 10.1785/0220170097

Table 1
Ground-Motion Model Evaluation Studies

Study	Region	Independent Data?*
Lee *et al.* (2000)	Southern California	No
Bindi *et al.* (2006)	Umbria–Marche, Italy	Yes; foreign models
Douglas *et al.* (2006)	French Antilles	Yes; foreign models
Drouet *et al.* (2007)	Pyrenees, Spain	Yes; foreign models
Hintersberger *et al.* (2007)	Central Europe	Yes; foreign models
Stafford *et al.* (2008)	Euro-Mediterranean	No
Delavaud *et al.* (2009)	California	No
Douglas and Mohais (2009)	French Antilles	Yes; foreign models
Scasserra *et al.* (2009)	Italy	No
Nishimura (2010)	Japan	Yes
Shoja-Taheri *et al.* (2010)	Iran	Yes; foreign models
Çağnan *et al.* (2011)	Turkey	No
Kaklamanos and Baise (2011)	California	Yes
Ornthammarath *et al.* (2011)	South Iceland	No
Uchiyama and Midorikawa (2011)	Japan	Yes; foreign models
Arango *et al.* (2012)	South and Central America	Yes
Beauval *et al.* (2012)	Southern and Eastern France	Yes; foreign models
Delavaud *et al.* (2012)	Global	No
Joshi *et al.* (2012)	Himalaya	Yes; foreign models
Massa *et al.* (2012)	Italy	No
Mousavi *et al.* (2012)	Iran	No
Vilanova *et al.* (2012)	Southwestern Iberia	Yes; foreign models
Edwards and Douglas (2013)	Cooper Basin, Australia	Yes; stochastic models
Vacareanu *et al.* (2013)	Eastern Romania	Yes; foreign models
Mousavi *et al.* (2014)	Iran	Yes; foreign models
Ogweno and Cramer (2014)	Central and Eastern United States	No
Zafarani and Mousavi (2014)	Ahar–Varzaghan, Iran	No
Allen and Brillon (2015)	Haida Gwaii, Canada	Yes; foreign models
Drouet and Cotton (2015)	France Alps	Yes; stochastic models
Haendel *et al.* (2015)	Northern Chile	Yes
Mak *et al.* (2015)	Italy	Yes[†]
Danciu *et al.* (2016)	Middle East	No
Roselli *et al.* (2016)	Italy	No
Salic *et al.* (2017)	Western Balkan	Yes; foreign models
Van Houtte (2017)	New Zealand	No
Zafarani and Farhadi (2017)	Iran	No

*“No” indicates studies in which a significant portion of the test data has been used to develop some of the evaluated models. “foreign models” indicates that foreign models are evaluated using local tests, so the data are naturally independent of the models; similarly for “stochastic models.”

[†] Regarding Mak *et al.* (2015), the dataset for small and moderate (respectively, large) earthquake data was independent (respectively, not independent) of the evaluated models.

flatfile (see Data and Resources) includes the data used to develop 1997BJF (hereafter, the subset pre-97; these data were also used, together with many other new data, to develop 2008BA), as well as a subset of data that have not been used in both 1997BJF and 2008BA (hereafter, the subset post-NGA). We used this pair of GMMs and data subsets to demonstrate the importance of data independence in data-driven GMM evaluations. 1997BJF is the only predecessor of the NGA GMMs that provides the complete dispersion information regarding the within- and between-event sigmas, so that a

Table 2
Data Used in This Study

Earthquake ID*	Subset	Earthquake Name*	Date (yyyy/mm/dd)	Magnitude	$N^\dagger$
6	pre-97	Imperial Valley-02	1940/05/19	6.95	1
12	pre-97	Kern County	1952/07/21	7.36	1
25	pre-97	Parkfield	1966/06/28	6.19	4
30	pre-97	San Fernando	1971/02/09	6.61	15
33	pre-97	Point Mugu	1973/02/21	5.65	1
45	pre-97	Santa Barbara	1978/08/13	5.92	2
48	pre-97	Coyote Lake	1979/08/06	5.74	10
50	pre-97	Imperial Valley-06	1979/10/15	6.53	31
53	pre-97	Livermore-01	1980/01/24	5.80	5
118	pre-97	Loma Prieta	1989/10/18	6.93	39
125	pre-97	Landers	1992/06/28	7.28	14
176	post-NGA	Tottori, Japan	2000/10/06	6.61	18
177	post-NGA	San Simeon, California	2003/12/22	6.50	6
179	post-NGA	Parkfield-02, California	2004/09/28	6.00	74
180	post-NGA	Niigata, Japan	2004/10/23	6.63	20
278	post-NGA	Chuetsu-oki	2007/07/16	6.80	68
279	post-NGA	Iwate	2008/06/13	6.90	56
280	post-NGA	El Mayor–Cucapah	2010/04/04	7.20	22

*As given in the Next Generation Attenuation (NGA)-West2 flatfile.
[†]Number of records within 40 km from the rupture plane.

fair comparison between the predecessor–successor pair (see a detailed discussion in Mak *et al.*, 2017) is possible. We, therefore, use only the 1997BJF–2008BA pair in our analysis.

We compared the relative performance of 1997BJF and 2008BA (for peak ground acceleration prediction) using these two datasets. Because 1997BJF is designed for magnitudes above 5.5, we removed events smaller than this threshold from the two subsets. We also used only the records within 40 km from the rupture plane for two reasons. First, near-field records are more relevant than far-field records for most engineering interests. Second, although the modelers of 1997BJF specified that their model is applicable for distance within 80 km and about half of the data they used to develop the model are of distance greater than 40 km, the majority (~90%) of the far-field records comes from only two earthquakes (Loma Prieta and Landers; see Boore *et al.*, 1997, their table 5 and fig. 1). A model developed from such a highly unbalanced far-field dataset is not expected to perform well in the far field. To avoid the effect of distance distribution on the model performance and focus our discussion on the effect of data independence, we used only near-field data. The filtered subsets pre-97 and post-NGA contained 123 records from 11 events and 264 records from 7 events, respectively (Table 2).

The performance evaluation based on the subset post-NGA is a true test of the predictive power of the two GMMs, whereas the evaluation based on the subset pre-97 is expected to favor 1997BJF because the test data overlap much with the data that were used to develop the model. 1997BJF was originally developed from 23 earthquakes containing 271 records (Boore *et al.*, 1997, their table 5); the subset pre-97 is about half the size of the original dataset. The discrepancy between the sizes of the test data and the original data 1997BJF used is due to our exclusion of far-field data. On the other hand, 2008BA was originally developed from 58 earthquakes containing 1674 records (Boore and Atkinson, 2008, their table 1); the size of the subset pre-97 is less than one-tenth of the size of the original dataset.

The evaluation method we used follows Mak *et al.* (2017). This approach is based on the multivariate logarithmic score, an extension of the widely used log-likelihood score (Scherbaum *et al.*, 2009) to include the correlation structure of a GMM, so that the information carried by the within- and between-event sigmas can be fully incorporated. It also uses the cluster bootstrap to assess the variability of the evaluation result. It shows the relative performance of two models by the distinctness index (DI), a value ranging from −1 to 1. A positive (respectively, negative) DI means one GMM is more (respectively, less) often better than the other, given the variability of the available data. The extreme case of DI = 1 (respectively, −1) means a model is always better (respectively, worse) than the other. The use of the multivariate logarithmic score, similar to the widely used mixed-effect model for GMM development, has the advantage that the result will not be dominated by well-recorded earthquakes. This is important for our analysis because the data we used are somewhat unbalanced (Table 2).

Based on this evaluation method, the performance of 1997BJF was found to be better than 2008BA when the subset pre-97 was used (DI = 0.72 for 1997BJF with respect to 2008BA). When using the subset post-NGA, however, 2008BA was found to perform better (DI = −0.83 for 1997BJF with respect to 2008BA). We reiterate that the subset pre-97 is not independent of 1997BJF, and the subset post-NGA is independent of both GMMs.

THE IMPORTANCE OF BEING INDEPENDENT

The presented example is, of course, not of practical importance in its specific sense, because superseded GMMs are often not considered better than their successors and so should not be used (Bommer *et al.*, 2010). The focus here is the effect of data independence in a data-driven GMM evaluation. Although 2008BA has better predictive power, as demonstrated by the test using the post-NGA subset (which is both independent and prospective), an opposite (and, presumably, incorrect) result could occur if the test data are not independent of 1997BJF.

GMM evaluations (Table 1) often involve both foreign and local models, whereas the test data are often local. Although the test data are seldom identical to that used for developing the evaluated models, it is not uncommon for the two datasets to significantly overlap. Because large earthquakes are rare, it is understandably difficult to compile a set of completely independent data of engineering significance (e.g., large magnitude and short distance) when the model comparison involves local models. Even with nonindependent data, data-driven GMM evaluation could still be useful: if a local model, when favored by the local data, is not found to be unambiguously better than a foreign one, it will be a strong argument for the good performance of the foreign model (e.g., Stafford *et al.*, 2008). The problem is in the opposite case that the local model is shown to be better (e.g., Mousavi *et al.*, 2012; Zafarani and Farhadi, 2017). In such a case, the data-driven evaluation may be inconclusive.

A data-driven GMM evaluation informs model selection. It is, however, seldom sufficient to select models and assign weights (assume using the logic-tree approach) purely based on empirical evidence because the available data seldom cover all situations of practical interest. Nevertheless, a data-driven evaluation helps to clearly separate the portion of the decision made based on empirical evidence from that based on expert judgment, enhancing the transparency of the decision-making process. When the test data are not independent of the evaluated models, the evaluation result should be interpreted with care and should not be taken entirely as empirical evidence.

This article focuses on the issue of data independence of data-driven evaluation of GMMs. The principle of the need to use independence data also applies to the data-driven evaluation of seismic-hazard models (e.g., Mak and Schorlemmer, 2016). The issue of data independence, however, may be less critical for evaluating seismic-hazard models because hazard models, unlike common GMMs, are often not direct products of statistical model fittings. The seismicity model used in a seismic-hazard model may rely heavily on geological and geodetic information that is largely independent of the test data, which is ground-motion observation.

DATA AND RESOURCES

The Next Generation Attenuation (NGA)-West2 flatfile was downloaded from http://peer.berkeley.edu/ngawest2/databases/ (last accessed January 2017). The ground-motion model (GMM) computation was conducted using OpenQuake Hazard Library (https://github.com/gem/oq-hazardlib, last accessed January 2017). An accelerograph records three-component time series. A peak ground acceleration (PGA) value is derived from combining the peak values of the two horizontal time series. The details of how this combination is done in the NGA-West2 flatfile, the data used in Boore and Atkinson (2008; hereafter, 2008BA), and the data used in Boore *et al.* (1997; hereafter, 1997BJF) are different. Moreover, some records in the flatfile may have been reprocessed so that they are different from what has been used in 2008BA and 1997BJF. We assume these differences did not affect our analysis.

ACKNOWLEDGMENTS

The authors thank the careful reviews by James Kaklamanos and Associate Editor Eric Thompson. This study is supported by the Global Earthquake Model Foundation (www.globalquakemodel.org, last accessed January 2017).

REFERENCES

Allen, T. I., and C. Brillon (2015). Assessment of ground-motion models for use in the British Columbia north coast region, Canada, *Bull. Seismol. Soc. Am.* **105,** no. 28, 1193–1205, doi: 10.1785/0120140266.

Arango, M. C., F. O. Strasser, J. J. Bommer, J. M. Cepeda, R. Boroschek, D. A. Hernandez, and H. Tavera (2012). An evaluation of the applicability of current ground-motion models to the south and central American subduction zones, *Bull. Seismol. Soc. Am.* **102,** no. 1, 143–168, doi: 10.1785/0120110078.

Beauval, C., H. Tasan, A. Laurendeau, E. Delavaud, F. Cotton, P. Guéguen, and N. Kuehn (2012). On the testing of ground-motion prediction equations against small-magnitude data, *Bull. Seismol. Soc. Am.* **102,** no. 5, 1994–2007, doi: 10.1785/0120110271.

Bindi, D. (2017). The predictive power of ground-motion prediction equations, *Bull. Seismol. Soc. Am.* **107,** no. 2, 1005–1011, doi: 10.1785/0120160224.

Bindi, D., L. Luzi, F. Pacor, G. Franceschina, and R. R. Castro (2006). Ground-motion predictions from empirical attenuation relationships versus recorded data: The case of the 1997–1998 Umbria-Marche, central Italy, strong-motion data set, *Bull. Seismol. Soc. Am.* **96,** no. 3, 984–1002, doi: 10.1785/0120050102.

Bommer, J. J., J. Douglas, F. Scherbaum, F. Cotton, H. Bungum, and D. Fäh (2010). On the selection of ground-motion prediction equations for seismic hazard analysis, *Seismol. Res. Lett.* **81,** no. 5, 783–793, doi: 10.1785/gssrl.81.5.783.

Boore, D. M., and G. M. Atkinson (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV, and

5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthq. Spectra* **24,** no. 1, 99–138, doi: 10.1193/1.2830434.

Boore, D. M., W. B. Joyner, and T. E. Fumal (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work, *Seismol. Res. Lett.* **68,** no. 1, 128–153, doi: 10.1785/gssrl.68.1.128.

Çağnan, Z., S. Akkar, and P. Gülkan (2011). A predictive ground-motion model for Turkey and its comparison with recent local and global GMPEs, in *Earthquake Data in Engineering Seismology*, in Geotechnical, Geological, and Earthquake Engineering, chapter 4, Springer, Amsterdam, The Netherlands, 29–52, ISBN: 978-94-007-0151-9, doi: 10.1007/978-94-007-0152-6_4.

Cotton, F., F. Scherbaum, J. J. Bommer, and H. Bungum (2006). Criteria for selecting and adjusting ground-motion models for specific target regions: Application to Central Europe and rock sites, *J. Seismol.* **10,** no. 2, 137–156, doi: 10.1007/s10950-005-9006-7.

Danciu, L., Ö. Kale, and S. Akkar (2016). The 2014 earthquake model of the Middle East: Ground motion model and uncertainties, *Bull. Earthq. Eng.* doi: 10.1007/s10518-016-9989-1.

Delavaud, E., F. Scherbaum, N. Kuehn, and T. Allen (2012). Testing the global applicability of ground-motion prediction equations for active shallow crustal regions, *Bull. Seismol. Soc. Am.* **102,** no. 2, 707–721, doi: 10.1785/0120110113.

Delavaud, E., F. Scherbaum, N. Kuehn, and C. Riggelsen (2009). Information-theoretic selection of ground-motion prediction equations for seismic hazard analysis: An applicability study using Californian data, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3248–3263, doi: 10.1785/0120090055.

Douglas, J., and R. Mohais (2009). Comparing predicted and observed ground motions from subduction earthquakes in the Lesser Antilles, *J. Seismol.* **13,** 577–587, doi: 10.1007/s10950-008-9150-y.

Douglas, J., D. Bertil, A. Roullé, P. Dominique, and P. Jousset (2006). A preliminary investigation of strong-motion data from the French Antilles, *J. Seismol.* **10,** no. 3, 271–299, doi: 10.1007/s10950-006-9016-0.

Drouet, S., and F. Cotton (2015). Regional stochastic GMPEs in low-seismicity areas: Scaling and aleatory variability analysis—Application to the French Alps, *Bull. Seismol. Soc. Am.* **105,** no. 4, 1883–1902, doi: 10.1785/0120140240.

Drouet, S., F. Scherbaum, F. Cotton, and A. Souriau (2007). Selection and ranking of ground motion models for seismic hazard analysis in the Pyrenees, *J. Seismol.* **11,** 87–100, doi: 10.1007/s10950-006-9039-6.

Edwards, B., and J. Douglas (2013). Selecting ground-motion models developed for induced seismicity in geothermal areas, *Geophys. J. Int.* **195,** no. 2, 1314–1322, doi: 10.1093/gji/ggt310.

Haendel, A., S. Specht, N. M. Kuehn, and F. Scherbaum (2015). Mixtures of ground-motion prediction equations as backbone models for a logic tree: An application to the subduction zone in northern Chile, *Bull. Earthq. Eng.* **13,** 483–501, doi: 10.1007/s10518-014-9636-7.

Hintersberger, E., F. Scherbaum, and S. Hainzl (2007). Update of likelihood-based ground-motion model selection for seismic hazard analysis in western central Europe, *Bull. Earthq. Eng.* **5,** 1–16, doi: 10.1007/s10518-006-9018-x.

Jordan, T. (2006). Earthquake predictability, brick by brick, *Seismol. Res. Lett.* **77,** no. 1, 3–6, doi: 10.1785/gssrl.77.1.3.

Joshi, A., A. Kumar, C. Lomnitz, H. Castaños, and S. Akhtar (2012). Applicability of attenuation relations for regional studies, *Geofis. Int.* **51,** no. 4, 349–363.

Kaklamanos, J., and L. G. Baise (2011). Model validations and comparisons of the Next Generation Attenuation of ground motions (NGA-West) project, *Bull. Seismol. Soc. Am.* **101,** no. 1, 160–175, doi: 10.1785/0120100038.

Lee, Y., J. G. Anderson, and Y. Zeng (2000). Evaluation of empirical ground-motion relations in southern California, *Bull. Seismol. Soc. Am.* **90,** no. 6B, S136–S148.

Mak, S., and D. Schorlemmer (2016). A comparison between the forecast by the United States national seismic hazard maps with recent ground motion records, *Bull. Seismol. Soc. Am.* **106,** no. 4, 1817–1831, doi: 10.1785/0120150323.

Mak, S., R. A. Clements, and D. Schorlemmer (2015). Validating intensity prediction equations for Italy by observations, *Bull. Seismol. Soc. Am.* **105,** no. 6, 2942–2954, doi: 10.1785/0120150070.

Mak, S., R. A. Clements, and D. Schorlemmer (2017). Empirical evaluation of hierarchical ground motion models: Score uncertainty and model weighting, *Bull. Seismol. Soc. Am.* **107,** no. 2, 949–965, doi: 10.1785/0120160232.

Massa, M., L. Luzi, F. Pacor, D. Bindi, and G. Ameri (2012). Comparison between empirical predictive equations calibrated at global and national scale and Italian strong-motion data, *Boll. Geof. Teor. Appl.* **53,** no. 1, 37–53, doi: 10.4430/bgta0018.

Mousavi, M., A. Ansari, H. Zafarani, and A. Azarbakht (2012). Selection of ground motion prediction models for seismic hazard analysis in the Zagros region, Iran, *J. Earthq. Eng.* **16,** no. 8, 1184–1207, doi: 10.1080/13632469.2012.685568.

Mousavi, M., H. Zafarani, S. Rahpeyma, and A. Azarbakht (2014). Test of goodness of the NGA ground-motion equations to predict the strong motions of the 2012 Ahar-Varzaghan dual earthquakes in northwestern Iran, *Bull. Seismol. Soc. Am.* **104,** no. 5, 2512–2528, doi: 10.1785/0120130302.

Nishimura, T. (2010). Conformity of the attenuation relationships in Japan with those by the NGA-project, *Summaries of Technical Papers of Annual Meeting of Architectural Institute of Japan*, Toyama, Japan, 9–11 September 2010 (in Japanese).

Ogweno, L. P., and C. H. Cramer (2014). Comparing the CENA GMPEs using NGA-East ground-motion database, *Seismol. Res. Lett.* **85,** no. 6, 1377–1393, doi: 10.1785/0220140045.

Ornthammarath, T., J. Douglas, R. Sigbjörnsson, and C. G. Lai (2011). Assessment of ground motion variability and its effects on seismic hazard analysis: A case study for Iceland, *Bull. Earthq. Eng.* **9,** 931–953, doi: 10.1007/s10518-011-9251-9.

Roselli, P., W. Marzocchi, and L. Faenza (2016). Toward a new probabilistic framework to score and merge ground-motion prediction equations: The case of the Italian region, *Bull. Seismol. Soc. Am.* **106,** no. 2, 720–733, doi: 10.1785/0120150057.

Salic, R., M. A. Sandikkaya, Z. Milutinovic, Z. Gulerce, L. Duni, V. Kovacevic, S. Markusic, J. Mihaljevic, N. Kuka, N. Kaludjerovic, *et al.* (2017). BSHAP project strong ground motion database and selection of suitable ground motion models for the Western Balkan region, *Bull. Earthq. Eng.* **15,** 1319–1343, doi: 10.1007/s10518-016-9950-3.

Scasserra, G., J. P. Stewart, P. Bazzurro, G. Lanzo, and F. Mollaioli (2009). A comparison of NGA ground-motion prediction equations to Italian data, *Bull. Seismol. Soc. Am.* **99,** no. 5, 2961–2978, doi: 10.1785/0120080133.

Scherbaum, F., E. Delavaud, and C. Riggelsen (2009). Model selection in seismic hazard analysis: An information-theoretic perspective, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3234–3247, doi: 10.1785/0120080347.

Schorlemmer, D., and M. Gerstenberger (2007). RELM testing center, *Seismol. Res. Lett.* **78,** no. 1, 30–36, doi: 10.1785/gssrl.78.1.30.

Shmueli, G. (2010). To explain or to predict? *Stat. Sci.* **25,** no. 3, doi: 10.1214/10-STS330.

Shoja-Taheri, J., S. Naserieh, and G. Hadi (2010). A test of the applicability of NGA models to the strong ground-motion data in the Iranian plateau, *J. Earthq. Eng.* **14,** 278–292, doi: 10.1080/13632460903086051.

Stafford, P. J., F. O. Strasser, and J. J. Bommer (2008). An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region, *Bull. Earthq. Eng.* **6,** 149–177, doi: 10.1007/s10518-007-9053-2.

Uchiyama, Y., and S. Midorikawa (2011). A study of the applicability of NGA models to strike-slip earthquakes in Japan, *Summaries of Tech-*

nical Papers of Annual Meeting of Architectural Institute of Japan, Tokyo, Japan, 23–25 August 2011 (in Japanese).

Vacareanu, R., F. Pavel, and A. Aldea (2013). On the selection of GMPEs for Vrancea subcrustal seismic source, *Bull. Earthq. Eng.* **11,** no. 6, 1867–1884, doi: 10.1007/s10518-013-9515-7.

Van Houtte, C. (2017). Performance of response spectral ground-motion models against New Zealand data, *Bull. New Zeal. Natl. Soc. Earthq. Eng.* **50,** no. 1, 21–38.

Vilanova, S. P., J. F. B. D. Fonseca, and C. S. Oliveira (2012). Ground-motion models for seismic-hazard assessment in western Iberia: Constraints from instrumental data and intensity observations, *Bull. Seismol. Soc. Am.* **102,** no. 1, 169–184, doi: 10.1785/0120110097.

Zafarani, H., and A. Farhadi (2017). Testing ground-motion prediction equations against small-to-moderate magnitude data in Iran, *Bull. Seismol. Soc. Am.* **107,** no. 2, 912–933, doi: 10.1785/0120160046.

Zafarani, H., and M. Mousavi (2014). Applicability of different ground-motion prediction models for northern Iran, *Nat. Hazards* **73,** no. 3, 1199–1228, doi: 10.1007/s11069-014-1151-2.

Sum Mak
Fabrice Cotton
Danijel Schorlemmer
Helmholtz-Zentrum Potsdam—Deutsches GeoForschungs-Zentrum (GFZ)
Section 2.6, Telegrafenberg
14473 Potsdam, Germany
smak@gfz-potsdam.de