Bulletin of the Seismological Society of America, Vol. 105, No. 1, pp. 152–166, February 2015, doi: 10.1785/0120140129

Ⓔ

Revised Tectonic Forecast of Global Shallow Seismicity Based on Version 2.1 of the Global Strain Rate Map

by Peter Bird and Corné Kreemer*

Abstract A major upgrade of the Global Strain Rate Map, version 2.1, uses far more geodetic data, systematic data processing, more modeled plates and plate boundaries, an improved algorithm, and a finer spatial grid than version 1 (Kreemer *et al.*, 2014). We convert this model to an indefinite-term tectonic forecast of shallow seismicity on a fine global grid, using the assumptions of Bird *et al.* (2010) and similar algorithms. One new feature is smoothing of model strain rates, and thus forecast seismicity, around offshore plate boundaries where strain rates are not controlled by local geodetic data. Another is incorporation of velocity-dependent seismic coupling in subduction zones and continental convergent boundaries (Bird *et al.*, 2009). The seismicity model is constructed in six progressively more complex versions. Only catalog years 1977–2004 are used for calibration, leaving years 2005–2012 available for retrospective tests. Tests of forecast success that are independent of total earthquake rate and that use no declustering show success comparable to that of one mature, optimized smoothed-seismicity algorithm.

Online Material: FORTRAN 90 codes to generate seismicity forecasts.

Introduction

Long-term forecasting of seismicity is an essential service that seismology provides to the global community. It permits buildings and other infrastructure to be designed for an appropriate level of shaking and permits individuals and communities to share their risk through insurance and other kinds of disaster preparedness. The present situation is that several national governments have invested in national long-term forecasts of great sophistication, but the most recent global hazard map to be widely circulated and accepted is that of Giardini *et al.* (1999).

An essential (and usually dominant) component of past global seismicity models has been the extrapolation of past seismicity to the future, with appropriate spatial smoothing. A weakness of this method is that global earthquake catalogs are too short to record the largest possible earthquakes in most areas and may locally fail to record a representative level of activity. Therefore, there have also been large efforts to catalog and study active and potentially active faults that are potential earthquake sources, and such fault-based models play an important role in some national models, such as the recent U.S. National Seismic Hazard Map (Petersen *et al.*, 2014). Yet fault mapping is sadly incomplete in many countries, and fault mapping on the ocean floor presents another set of formidable challenges.

Two recent scientific developments give promise of a third approach, which might capture the ongoing deformation of the Earth's surface without the need to map every fault. One is the theory of plate tectonics, which clarifies how the greatest part of the deformation is organized. The other is geodesy, which now permits secular relative motions of benchmarks to be measured with precisions better than 1 mm/year. The third approach, which we are exploring, is to use the secular velocities to build global models of the strain rates of the Earth's surface and then convert these to maps of estimated long-term seismicity by applying appropriate conversion factors (such as the thickness of the seismogenic lithosphere) and the fraction of seismogenic coupling within this friction-dominated part of the system. Although geodesy is still primarily a land-based technology, the theory of plate tectonics provides guidance on how to extrapolate relative velocities observed on land to adjacent seafloor. This makes it possible to construct global velocity, strain rate, and seismicity models. Creation of global seismicity models implies the exciting possibility of prospective forecast tests that can give preliminary results in a year and definitive results in a decade.

Previously, we reported (Bird *et al.*, 2010) the conversion of the original Global Strain Rate Map (GSRM) of Kreemer *et al.* (2003) to a stationary forecast of global shallow

*Also at Nevada Seismological Laboratory, University of Nevada, Reno, Nevada.

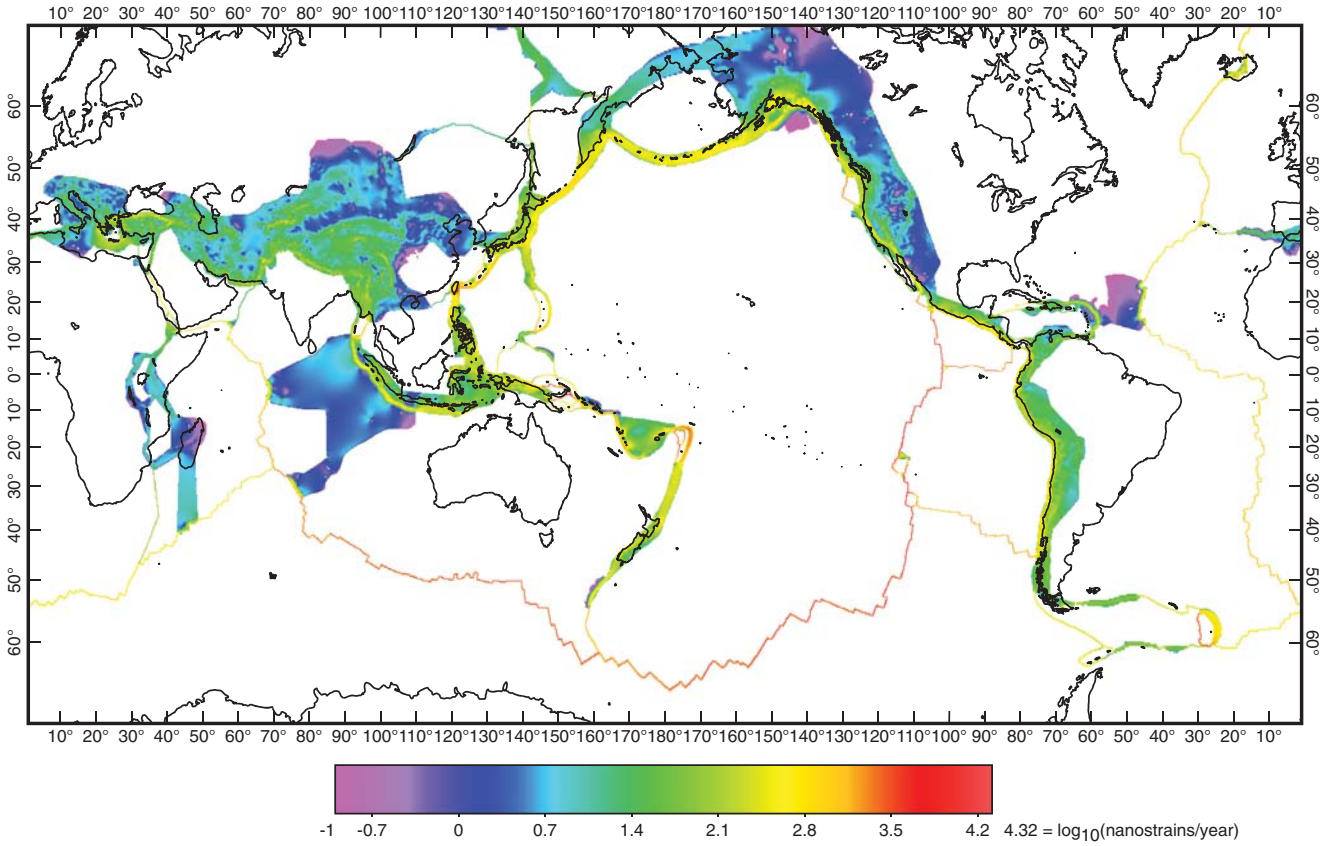

Figure 1. A Mercator projection showing the second invariant of 2D surface strain rates from Global Strain Rate Map, version 2.1. Colored areas are composed of cells 0.20° high in latitude and 0.25° wide in longitude, in each of which the average strain rate is shown. A logarithmic color scale is used to display variations across five orders of magnitude. All velocity gradients integrate across plate-boundary zones to equal relative plate rotation rates that were estimated simultaneously. White areas are modeled as plate-interior regions of zero secular strain rate. Additional details are available in Kreemer *et al.* (2014).

seismicity (in which "shallow" refers to all hypocentroids located no more than 70 km below sea level). That forecast was submitted to the Collaboratory for the Study of Earthquake Predictability (CSEP) for independent prospective testing; however, at the time this was written, only the simple *N*-test had been implemented. Meanwhile, Kreemer *et al.* (2014) reported a major upgrade of the GSRM: version 2.1 (GSRM2.1) uses far more geodetic data, systematic data processing, more modeled plates and plate boundaries, an improved algorithm, and a finer spatial grid. Here we report (1) a revised long-term tectonic forecast of global shallow seismicity based on GSRM2.1, including some improvements to the algorithm for this translation, and (2) results of retrospective testing of this forecast using information scores I_0 and I_1 of Kagan (2009) and the area skill score (ASS) of Zechar and Jordan (2008).

Version 2.1 of the Global Strain Rate Map

Whereas the previous GSRM was constrained by ~5200 geodetic velocities and additional constraints from Quaternary faulting and earthquake focal mechanisms, GSRM2.1 is constrained solely by 22,415 geodetic velocities, of which 6705 were determined from continuous Global Positioning System (GPS) stations exclusively for the project, and the rest came from 233 studies. Geodetic data showing aseismic transients were not used. The geometry of the deforming zones was redefined (partly after Bird, 2003), and the size of individual grid cells was reduced from $0.5° \times 0.6°$ to $0.20° \times 0.25°$ such that the total number of grid cells increased from ~24,000 to 144,086. Although deforming zones on land were mostly expanded compared with the previous model, the number of assumed rigid plates and blocks was also increased from 25 to 50. These new plates were mostly introduced in places where rigidity can be assumed but where there is not sufficient GPS coverage to model the separation between rigid blocks and deforming zones properly. In an attempt to properly fit the data everywhere, an empirical Bayesian approach was used in which the strain-rate model from a spatially damped solution was used to give prior variance inputs to the final model. The second invariant of the model strain-rate field is mapped in Figure 1. Further details are presented in Kreemer *et al.* (2014).

Scaling of Strain Rates to Shallow Seismicity

To convert GSRM2.1 to a forecast of long-term shallow seismicity, we apply the hypotheses, assumptions, and

equations of Bird and Liu (2007), who called them the seismic hazard inferred from tectonics (SHIFT) hypotheses: (1) the long-term seismic moment rate of any tectonic fault, or any large volume of permanently deforming lithosphere, is approximately that computed using the coupled seismogenic thickness (i.e., dimensionless seismic coupling coefficient × seismogenic thickness) of the most comparable class of plate boundary; and (2) the long-term rate of earthquakes generated along any tectonic fault, or within any large volume of permanently deforming lithosphere, is approximately that computed from its SHIFT moment rate (of method 1 above) using the frequency–magnitude distribution of the most comparable class of plate boundary. The necessary seismicity factors were estimated by Bird and Kagan (2004) for each of these plate-boundary classes: oceanic spreading ridge (OSR), oceanic transform fault (OTF), oceanic convergent boundary (OCB), subduction zone (SUB), continental convergent boundary (CCB), continental transform fault (CTF), and continental rift boundary (CRB). The division of all global plate-boundary steps into these seven classes was a feature of the PB2002 model of Bird (2003), who used bathymetry, seafloor age, the angle between relative plate velocity and the boundary, and some subsidiary criteria for subduction (volcanic arcs and Benioff zones) in this classification. The same paper identified 13 regions as orogens in which the model was known to be incomplete and/or inaccurate, so that calibration of seismicity factors could avoid using data from those regions.

Because this new forecast applies the SHIFT hypotheses to the GSRM2.1 strain rates, we refer to it as the SHIFT-GSRM2 model (allowing for the possibility that others will wish to convert GSRM2.1 to seismicity in other ways).

In this conversion, we faced four conceptual and/or practical difficulties. First, the strain rates available are not always the kinds of strain rates that would be preferred. The strain rates in GSRM2.1 are determined by geodetic velocities as well as assumed plate geometries. Except where faults creep (or have many small earthquakes during the measurement period), the strain rates inferred from differentiation of GPS velocities are dominantly elastic strain rates associated with rising deviatoric stresses. However, the strain rates input to SHIFT calculations should ideally be long-term permanent strain rates with no elastic components. Geodetic strain rates typically have smoother map patterns than long-term permanent strain rates which include singularities along fault traces. However, in a 2D Earth-surface model based on plate-tectonic concepts, cross-boundary line integrals of these two kinds of velocity gradient across any boundary are the same, because both are equal to the relative plate-velocity vector. The use of available GSRM2.1 strain rates, which include elastic components, should not greatly affect regional averages of long-term seismicity computed in a SHIFT model, but only smooth the forecast seismicity. On the scale of global maps and forecasts, this smoothing is relatively insignificant. Thus we disregard this distinction in practice.

Second, GSRM2.1 treats plate interiors as perfectly rigid and predicts zero strain rates in these regions. Yet a global seismicity forecast with zero rates in plate interiors would be both unrealistic and irresponsible. Our solution is to forecast a uniform low seismicity rate in all plate interiors, which is based on the collective frequency–magnitude distribution of these regions in the years covered by a reliable catalog.

Third, the basic SHIFT hypotheses do not specify how to decide which is the most comparable class of plate boundary for each spatial cell. These must be determined by subsidiary rules or hypotheses appropriate to the data and/or models available. We cannot use all of the decision rules suggested in table 2 of Bird and Liu (2007) because they assumed that all subduction zones and spreading ridges would be represented by discrete fault traces, which is not the case in GSRM2.1. We tested six increasingly sophisticated versions of these assignment rules, producing six variants of the SHIFT-GSRM2 forecast, which will be distinguished by suffix letters "a" through "f" in the discussion below.

Finally, we found (both in Bird *et al.*, 2010, and in the current work) that our raw SHIFT forecasts were underpredicting global shallow seismicity, primarily due to underpredictions of subduction zone seismicity. The greatest problem was the use of inappropriate geometric factors in the moment-rate formula for many subduction zones for which the faults dip much less than 45°. As discussed fully in Bird *et al.* (2010), we compensate for the absence of defined fault dips in the GSRM2.1 tectonic model by using a few empirical calibration factors in the last step of forecast computation, the values of which are based on retrospective counting of shallow earthquakes in each tectonic zone. The simplest models presented here divide the Earth's surface into only two tectonic zones, whereas the more complex models distinguish five tectonic zones. Thus, either 2 or 5 degrees of freedom in the map pattern of our forecast are determined by prior seismicity (measured at the lowest practical threshold magnitude), whereas all the other 144,671 degrees of freedom in each map are determined by the GSRM2.1 data and the SHIFT method.

Data Sets

The seismic catalog we use is the Global Centroid Moment Tensor (Global CMT; see Data and Resources) catalog (e.g., Ekström *et al.*, 2012). We consider all shallow earthquakes, defined as those with nominal centroid depth ≤70 km below sea level. We truncate the catalog at a threshold seismic moment ($M_T = 5.01 \times 10^{17}$ N·m, equivalent to moment magnitude threshold of $m_T = 5.767$) to assure that the subcatalog we use is nearly complete (Kagan, 2003; Bird and Kagan, 2004). We use the moment-to-magnitude conversion of Hanks and Kanamori (1979):

$$m = (2/3)[\log_{10}(M) - 9.05]. \quad (1)$$

GSRM2.1 is presented by Kreemer *et al.* (2014) and represented by file GSRM_average_strain_v2.1.txt. Specifically,

we use horizontal-plane (three-component) strain rates that are average values, stated at the centers of $0.20° \times 0.25°$ (latitude/longitude) cells forming a regular latitude/longitude grid. Cells not described in this file are intraplate regions with assumed zero strain rate.

Constants needed for the SHIFT calculations are obtained from table 5 of Bird and Kagan (2004). This table has columns for each of seven classes of plate boundary, each previously defined in the PB2002 model of Bird (2003). It lists empirical estimates of mean coupled thickness of seismogenic lithosphere $\langle cz \rangle$ (i.e., mean value of dimensionless seismic coupling coefficient × seismogenic lithosphere thickness), elastic shear modulus μ, corner magnitude m_c in the tapered Gutenberg–Richter frequency/seismic moment relation (Jackson and Kagan, 1999; Kagan and Jackson, 2000), asymptotic spectral slope β (at low magnitudes) in the same law, the number of shallow CMT earthquakes N^{CMT} (outside orogens and above threshold) associated with each class of plate boundary, and the threshold seismic moment $M_{\mathrm{T}}^{\mathrm{CMT}}$ used in counting those events. Bird and Kagan (2004) determined coupled thicknesses $\langle cz \rangle$ by comparing their subcatalogs based on the Global CMT catalog to the PB2002 kinematic model of relative plate motions. Because they restricted such comparisons to nonorogen regions in which motions are relatively well known, and because relative plate rotations in GSRM2.1 are quite similar to those in PB2002, we do not consider it necessary to repeat those calculations.

Plate boundaries from the PB2002 model (Bird, 2003) are represented by published file PB2002_steps.dat, which contains classes, elevations, locations, and relative velocities (see Data and Resources). However, plate-boundary locations and strain rates from GSRM2.1 take precedence over these older estimates; this file is used primarily to provide boundary classes, elevations, and azimuths of relative plate motion in subduction zones.

In models SHIFT-GSRM2d–f, we refer to a classification of the Earth's surface into five mutually exclusive tectonic zones (0, intraplate; 1, active continent; 2, slow ridge; 3, fast ridge; 4 trench) that was introduced by Kagan *et al.* (2010), who computed the zones with an objective algorithm. Their assignments are in the published file PB2002_tectonic_zones.grd (see Data and Resources).

SHIFT Calculations

To forecast the seismicity of one spatial cell, we start with the horizontal-plane components of the strain-rate tensor. We determine the vertical strain rate $\dot{\varepsilon}_{rr}$ by invoking incompressibility:

$$\dot{\varepsilon}_{\phi\phi} + \dot{\varepsilon}_{\theta\theta} + \dot{\varepsilon}_{rr} = 0, \tag{2}$$

in which ϕ (longitude), θ (colatitude), and r (radius) define a spherical coordinate system for the Earth. (This is one of the steps in which the strain rate is assumed to be permanent.) Because $\dot{\varepsilon}_{r\phi}$ and $\dot{\varepsilon}_{r\theta}$ should vanish at the shear-traction-free surface, $\dot{\varepsilon}_{rr}$ is a principal strain rate. The two principal strain rates in the horizontal plane ($\dot{\varepsilon}_{1\mathrm{h}} \le \dot{\varepsilon}_{2\mathrm{h}}$ by convention) are

$$\dot{\varepsilon}_{1\mathrm{h}} = \frac{\dot{\varepsilon}_{\phi\phi} + \dot{\varepsilon}_{\theta\theta}}{2} - \sqrt{\dot{\varepsilon}_{\phi\theta}^2 + (\dot{\varepsilon}_{\phi\phi} - \dot{\varepsilon}_{\theta\theta})^2/4} \quad \text{and}$$
$$\dot{\varepsilon}_{2\mathrm{h}} = \frac{\dot{\varepsilon}_{\phi\phi} + \dot{\varepsilon}_{\theta\theta}}{2} + \sqrt{\dot{\varepsilon}_{\phi\theta}^2 + (\dot{\varepsilon}_{\phi\phi} - \dot{\varepsilon}_{\theta\theta})^2/4}. \tag{3}$$

These three principal strain rates are also arranged as an ordered triplet of eigenvalues:

$$\dot{\varepsilon}_1 \le \dot{\varepsilon}_2 \le \dot{\varepsilon}_3. \tag{4}$$

Next we assign the most comparable class of plate boundaries. (Variant ways of doing this are discussed below.) From this assignment, we infer the applicable values of coupled seismogenic thickness $\langle cz \rangle$ and shear modulus μ for use in the conversion of strain rate to seismic moment rate; we also infer the values of corner magnitude m_c and asymptotic spectral slope β of the frequency/moment relation for use in conversion of seismic moment rate to earthquake rate. A global map of $\langle cz \rangle$ is shown by Kreemer *et al.* (2014).

We then compute the expected SHIFT seismic moment rate using equation (7b) of Bird and Liu (2007),

$$\dot{M} = A\langle cz \rangle \mu \begin{cases} 2\dot{\varepsilon}_3; & \text{if } \dot{\varepsilon}_2 < 0, \text{ or} \\ -2\dot{\varepsilon}_1; & \text{if } \dot{\varepsilon}_2 \ge 0 \end{cases}, \tag{5}$$

in which A is the area of one GSRM2.1 spatial cell. This equation is based on a kinematic model: that volume-preserving deformation is equivalent to slip on faults falling into (one or two) conjugate sets. The more active conjugate fault set bisects the angles between the principal strain-rate axes $\pm\hat{\varepsilon}_1$ and $\pm\hat{\varepsilon}_3$; the less active conjugate fault set bisects the angles between principal strain-rate axes $\pm\hat{\varepsilon}_2$ and $\pm\hat{\varepsilon}_3$ (if $\dot{\varepsilon}_2 < 0$) or between $\pm\hat{\varepsilon}_1$ and $\pm\hat{\varepsilon}_2$ (if $\dot{\varepsilon}_2 \ge 0$). The factor of ± 2 appearing in equation (5) is the smallest coefficient possible and comes from the assumption that these fault planes make angles of 45° with the principal strain-rate axes. Because this assumption is often wrong, we found it necessary to apply regional recalibration factors for each tectonic zone in a final step.

Once the SHIFT seismic moment rate ($\dot{M}$, in $\mathrm{N \cdot m \cdot s^{-1}}$) of a grid cell is determined, its expected SHIFT shallow earthquake rate is obtained in two steps. First, we divide the SHIFT moment rate of the cell by the model moment rate (integral of best-fitting tapered Gutenberg–Richter distribution) of the appropriate Bird and Kagan (2004) orogen-free subcatalog of the Global CMT catalog. We then multiply by the number of events in that subcatalog to determine the rate of earthquakes in the cell that will exceed the threshold magnitude of that subcatalog:

$$\dot{N}(m > m_{\mathrm{T}}^{\mathrm{CMT}}) = (\dot{M}/\dot{M}^{\mathrm{CMT}})\dot{N}^{\mathrm{CMT}}. \tag{6}$$

We adjust the forecast rate to any desired threshold magnitude m_{T} using the tapered Gutenberg–Richter model (see also equation 9 of Bird and Kagan 2004):

$$\dot{N}(m > m_{\mathrm{T}}) = \dot{N}(m > m_{\mathrm{T}}^{\mathrm{CMT}})\left[\frac{M(m_{\mathrm{T}})}{M(m_{\mathrm{T}}^{\mathrm{CMT}})}\right]^{-\beta} \times \exp\left[\frac{M(m_{\mathrm{T}}^{\mathrm{CMT}}) - M(m_{\mathrm{T}})}{M(m_{\mathrm{c}})}\right], \quad (7)$$

in which corner magnitude m_{c} describes seismicity of the analog plate-boundary class.

Finally, we sum the raw forecast earthquake rate of all the cells in each of the tectonic zones and compare this zone total with the actual Global CMT shallow earthquake rate in the same tectonic zone, during a preforecast calibration period, to determine a correction factor for that zone. This calibration is performed at the catalog-completeness threshold of $m_{\mathrm{T}} = 5.767$, regardless of the intended threshold for the forecast that is to be computed. These factors are dimensionless; in model SHIFT-GSRM2f, they range from 0.86 (zone 1: active continent) to 2.90 (zone 4: trench) for the calibration period 1977–2004. This empirical correction corrects for a number of potential biases, especially including incorrect dips of dip-slip faults assumed in equation (5) and any differences between the seismic coupling and/or frequency–magnitude relations of Bird and Kagan (2004) orogen-free subcatalogs and the wider tectonic zones of this article.

Variant Versions of the Forecast

SHIFT-GSRM2a (or "a") is the simplest version, as it divides the Earth into only two tectonic zones, exactly as mapped in GSRM2.1: a plate boundary zone (all deforming cells of GSRM2.1) and an intraplate zone (all other cells). The plate-boundary seismicity forecast by SHIFT-GSRM2a is not displayed because it looks just like Figure 1, except for a change in color scale. The global level of intraplate seismicity at the catalog-completeness threshold of 5.767 is determined from the Global CMT catalog, and for other desired thresholds it is projected using $m_{\mathrm{c}} = 9$ and $\beta = 0.63$ as in the previous SHIFT_GSRM forecast of Bird *et al.* (2010). Because more than half of Earth's shallow plate-boundary seismicity is generated by subduction zones (including associated shallow plate-bending earthquakes), we assign the seismicity properties of the SUB category of Bird and Kagan (2004) to all plate-boundary cells in model a. These include a high corner magnitude $m_{\mathrm{c}} = 9.58$ for the global plate-boundary zone, which causes this simple forecast to overpredict the seismicity of nonsubduction plate boundaries at high magnitudes. A more glaring defect became apparent when we plotted the calibration-period earthquakes that were automatically classified as intraplate. Instead of being uniformly distributed, they were strongly concentrated along offshore plate boundaries and amounted to 20.8% of total shallow seismicity.

This problem occurs because most offshore plate boundaries in GSRM2.1 (other than most subduction zones) are only one cell wide (i.e., 0.20° × 0.25°, or ~25 km). Whereas the original GSRM assigned wider offshore plate boundaries (guided by catalog seismicity), GSRM2.1 adopts a strict plate-tectonic model that plate boundaries are single discrete faults, except where geodetic data prove otherwise. The shortage of seafloor geodetic benchmarks thus leaves most offshore plate boundaries modeled as simple linear traces. As discussed in Bird and Kagan (2004), there are at least four reasons why mapped plate-boundary traces and offshore shallow seismicity might fail to overlap: (1) errors in the mapping of the plate boundaries, (2) errors in the location of earthquake hypocenters, (3) nonvertical dips of plate-boundary faults, and/or (4) finite width of the plate boundary due to multiple active faults in a network. We lack the high-resolution fault maps and earthquake hypocentroids needed to determine which of these factors is most responsible, especially under the ocean. Fortunately, Bird and Kagan (2004) presented empirical half-widths of offshore plate-boundary seismicity in their figure 1 and table 1 that give the net effect of all four sources of misfit.

Forecast version SHIFT-GSRM2b ("b") keeps the two-part tectonic zonation of version a but modifies the GSRM2.1 strain-rate map by isotropic Gaussian smoothing of the strain rates of offshore plate boundaries (other than those subduction zones that already have finite width) according to the aforementioned empirical half-widths determined by Bird and Kagan (2004). The strain rates in a GSRM2.1 cell are only redistributed if five criteria are met: (1) the nearest PB2002 plate-boundary step is not in class SUB or is a SUB mapped as only one cell wide; (2) the nearest PB2002 plate-boundary step is underwater; (3) the nearest PB2002 step is less than 186 km away; (4) this GSRM2.1 cell was adjacent (before smoothing) to a nondeforming intraplate cell; and (5) both the donor and the receiver cell are underwater. The Gaussian smoothing parameter σ_{m} is chosen according to which class of PB2002 plate-boundary step is closest to the offshore cell: OSR 66 km; OTF 64 km; OCB 93 km; SUB 88.75 km; CRB 77 km; CTF 128.5 km; and CCB 94.5 km. (The last three assignments occur only locally where plate boundaries cross coastlines.) Examining maps of calibration-period earthquakes that were automatically classified as intraplate under this method, we saw that the problem of mislabeled oceanic earthquakes was corrected along spreading ridges and oceanic transform faults but persisted along subduction zones, particularly in the outer-rise regions on the seaward sides of trenches.

Forecast version SHIFT-GSRM2c (or "c") adds unilateral smoothing of the activity (i.e., strain rates and, consequently, forecast seismicity) of the submarine parts of subduction zones toward their outer rises. We chose unilateral smoothing for two reasons. First, many subduction zones have land, and geodetic benchmarks, on the down-dip sides of trenches; we did not wish to artificially smooth the strain rates from GSRM2.1 in land areas where they were (often) constrained by GPS data. Second, subduction zones exhibit a great number of shallow earthquakes on the seaward sides of their trenches, which are generated by nonelastic bending of the downgoing lithosphere below the outer rises. Fortunately, we have three empirical constraints on this outer-rise seismicity from the

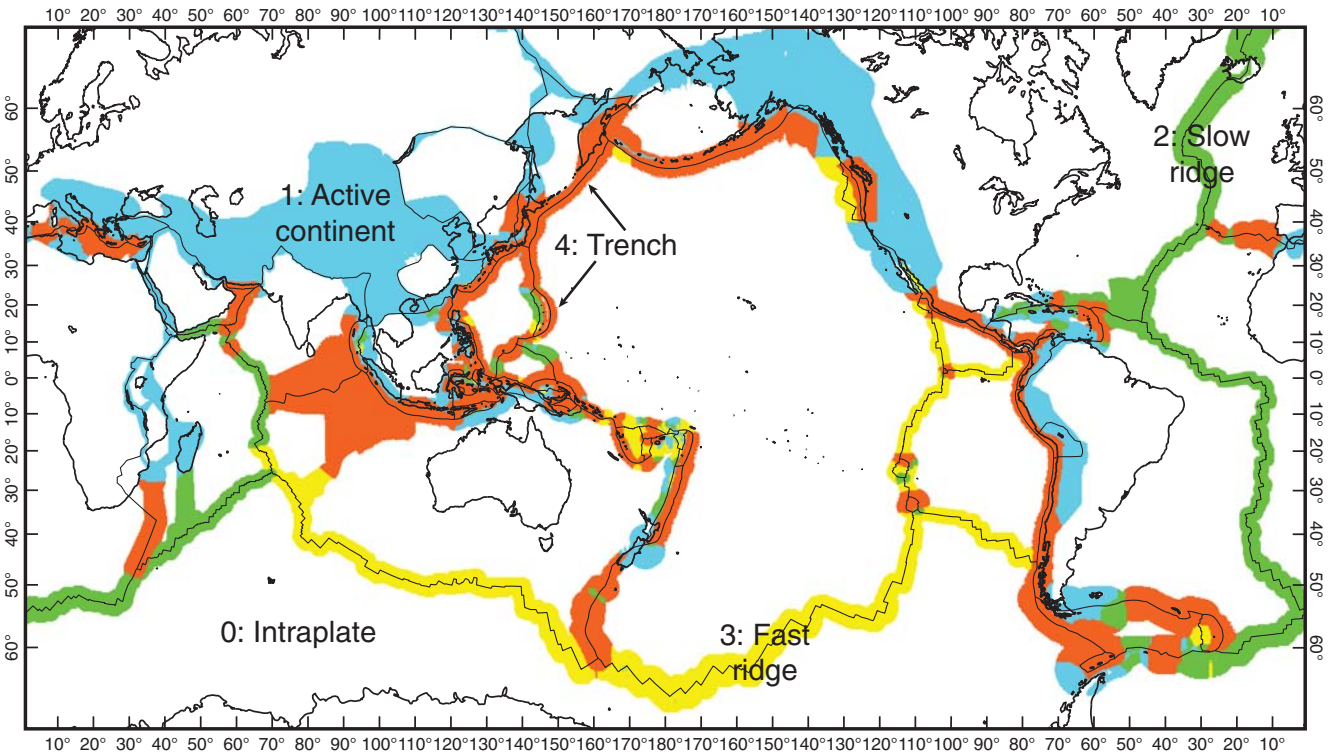

Figure 2. Tectonic zones used in models SHIFT-GSRM2d–f. The algorithm and datasets defining these zones were presented by Kagan *et al.* (2010). Because some deforming plate-boundary belts in SHIFT-GSRM2d–f are locally wider than the zone belts as originally defined, some zones shown here have been laterally expanded by a nearest-neighbor voting algorithm. Note that zone 4 (trench) includes all oceanic zones of compressive deformation, not just mature subduction zones. Zones 2 (slow ridge) and 3 (fast ridge) are not subdivided into spreading ridge segments versus transform fault segments, because this would require spatial resolution exceeding that of global seismic catalogs. Zone 1 (active continent) includes some continental shelves, slopes, and rises.

global-mean profile of subduction seismicity presented in figure 13 of Bird *et al.* (2009): 15% of shallow subduction zone seismicity is on the wrong (i.e., up-dip or seaward) side of the trench; the root mean square distance of these wrong-side earthquakes from the trench is 67 km; and wrong-side seismicity decays exponentially with distance from the trench. We devised a simple algorithm to transfer strain rate (and thus forecast seismicity) along the direction of relative plate motion (according to the PB2002 model) to meet these objectives. Small quanta of tensor strain rate are removed from convex regions near the maxima of cross-trench profiles of scalar strain rate and deposited in the outer-rise region after displacement by distances having a normal distribution with mean of 66 km. (Metaphorically, this can be thought of as blowing grains of sand off the crest of a dune and then redepositing them slightly downwind.) In addition to this displacement of quanta of activity along the azimuth of relative plate convergence, 20% of the displaced quanta are deposited one cell to the left or to the right of their nominal target to prevent the representation of outer-rise plate-bending activity from having unrealistically sharp terminations along the trends of trenches. This algorithm was successful in expanding the offshore parts of subduction plate boundaries to encompass virtually all of their associated plate-bending earthquakes. Thus, the map of calibration-period earthquakes classified as intraplate in this model is relatively free of the former concentration near known plate boundaries.

All forecast versions discussed above (a, b, and c) share common defects, which are that a single common coupled thickness is used for all plate-boundary cells and a single common corner magnitude is used when projecting their seismicity to high magnitudes. Both of these oversimplifications cause the seismicity of oceanic ridges and transforms to be systematically overpredicted. There are also overpredictions in regions of continental rifting and transform faulting. These deficiencies were addressed in model versions d, e, and f.

Forecast version SHIFT-GSRM2d ("d") adds recognition of the five tectonic zones (Fig. 2) of Kagan *et al.* (2010), both when assigning seismicity parameters to a grid cell and when grouping grid cells into a zone for the final seismicity-calibration step (which now uses 5 degrees of freedom instead of only 2). Cells in tectonic zone 1 (active continent) are assigned the seismicity parameters of subcatalog CTF of Bird and Kagan (2004). Cells in tectonic zone 2 (slow ridge) are assigned seismicity parameters of subcatalog slow-OTF. Cells in tectonic zone 3 (fast ridge) are assigned seismicity parameters of subcatalog fast-OTF. Cells in tectonic zone 4 (trench) get their parameters from subcatalog SUB. One practical difficulty we encountered is that some plate-boundary belts in SHIFT-GSRM2d–f are broader than the corresponding

tectonic zone belts defined by Kagan *et al.* (2010). We solved this with an automated algorithm that expands the nonzero tectonic zones iteratively, by nearest-neighbor voting methods, until all SHIFT-GSRM2d–f cells with a nonzero strain rate have a tectonic zone assignment different from 0 (intraplate). This algorithm appeared to work well over most of the globe but required some *ad hoc* adjustments in western North America, where plate-boundary topology has been rapidly changing. (Ⓔ Details are in the source code provided in programs S1 and S2 of the electronic supplement to this article.)

Forecast version SHIFT-GSRM2e ("e") adds a more fine-grained assignment of plate-boundary analog classes for cells in tectonic zones 1–3 (active continent, slow ridge, and fast ridge). The idea is that any strain-rate tensor for which two horizontal principal values have opposite signs implies a mixture of strike-slip and dip-slip faulting but that these will typically occur on distinct faults of different dip and perhaps different coupling. As in the original SHIFT-GSRM forecast of Bird *et al.* (2010), we partition most strain-rate tensors in zones 1–3 into a strike-slip part (with vertical principal strain rate of zero) and a dip-slip part (with one horizontal principal strain rate of zero). Then, we compute the forecast seismicity of the strike-slip component in active continents using analog CTF of Bird and Kagan (2004); the dip-slip component is converted to seismicity using either analog CRB (where there is horizontal extension) or CCB (where there is horizontal compression), depending on its sense. In a similar way, the strike-slip components in zones 2–3 are modeled using OTF parameters, whereas the dip-slip components are modeled either with analog OSR (where there is horizontal extension) or OCB (where there is horizontal compression). In model e (and also in model f, below), the only sites at which we forecast large earthquakes along the East Pacific Rise and Pacific–Antarctic ridge (Fig. 3) are some exceptional environments where rotating microplates are overthrusting their neighbors (Bird, 2003); these oceanic compressional environments have been assigned to zone 4. However, slow ridges can still generate very occasional $m > 8$ earthquakes through their strike-slip slow-OTF components. In this forecast version ("e"), strain rates in tectonic zone 4 (trench) are not partitioned; however, one additional distinction is made. If the nearest plate boundary of model PB2002 is a SUB boundary, then appropriate subduction parameters are used; otherwise, a nonsubduction convergent boundary in the oceans is modeled using analog OCB of Bird and Kagan (2004). The collision of rotating microplates along the East Pacific Rise, mentioned above, is one example in which seismicity parameters are derived from class OCB instead of class SUB.

Forecast version SHIFT-GSRM2f ("f"; Fig. 4) adds one more complexity, based on the discovery of Bird *et al.* (2009) that subduction zones (SUB of Bird and Kagan, 2004) and continental convergent boundaries (CCB of Bird and Kagan, 2004) of well-known relative velocity (i.e., outside orogens defined by Bird, 2003) both exhibit velocity-dependent coupled thickness, with coupled thickness about twice as great at higher relative plate velocities. The reasons may include both (1) velocity-dependent pore pressures in megathrusts and (2) velocity-dependent viscous strength limits in megathrusts competing with velocity-independent frictional strength limits. Using these empirical results of Bird *et al.* (2009; their figs. 6 and 7 and related discussion), we apply an additional seismicity factor of 1.206 to subduction zone cells with relative plate velocity greater than 66 mm/year but apply factor 0.589 to slower subduction zones. Similarly, we apply seismicity factor 1.383 to continental convergent boundary cells with relative plate velocity over 24 mm/year but factor 0.606 to slower continental convergent cells. These velocity factors are applied prior to the retrospective calibration of overall seismicity in zones 4 (trench) and 1 (active continent), respectively; thus, they only redistribute forecast seismicity within the zone but do not change the overall seismicity of the zone in the final forecast.

Retrospective Testing

We computed each of these six forecast variants at two threshold magnitudes. Magnitude 5.767 is the completeness threshold of Global CMT, and the forecasts at this threshold can be tested most reliably because test earthquakes occur frequently. The second set of forecasts had threshold magnitude 7.0, so they are harder to test but have greater relevance to seismic hazard and risk. We calibrated each forecast using only the years 1977–2004 of the Global CMT catalog, leaving eight years (2005–2012) for testing. During these test years, there were 1694 and 90 shallow earthquakes at or above threshold 5.767 (Fig. 5) and 7.0, respectively.

We also tested one mature smoothed-seismicity forecast (Fig. 6) provided by Yan Y. Kagan and David D. Jackson. Their method (Kagan and Jackson, 1994, 2010a,b, 2011, 2014) does not use any plate-tectonic velocities or strain-rate information, but instead applies an optimized smoothing footprint to each previous shallow earthquake in the learning portion of the Global CMT catalog. The smoothing footprint for each source event with known mechanism is anisotropic and projects greater future seismicity along the strike of the source fault. The particular forecast tested was based on $m \geq 5.767$ shallow earthquakes in Global CMT years 1977–2004 and used smoothing parameter $r = 6$ km and a global grid of $0.1° \times 0.1°$ cells. Such forecasts are already being tested at CSEP. All forecasts project the same total earthquake rate, because all were calibrated to the same years (1977–2004) of the Global CMT catalog.

Our preferred measures of forecast specificity and success are two of the information scores (I) defined by Kagan (2009). I_0 is the sum over all forecast cells of the normalized forecast rate times the base 2 logarithm of the ratio of normalized forecast rate to normalized cell area. This measure I_0 (herein called "specificity") is the mean number of binary bits of information gain (per virtual, expected earthquake) over an ignorant model that has only a single global earthquake rate. Measure I_1 (herein called "success") is the

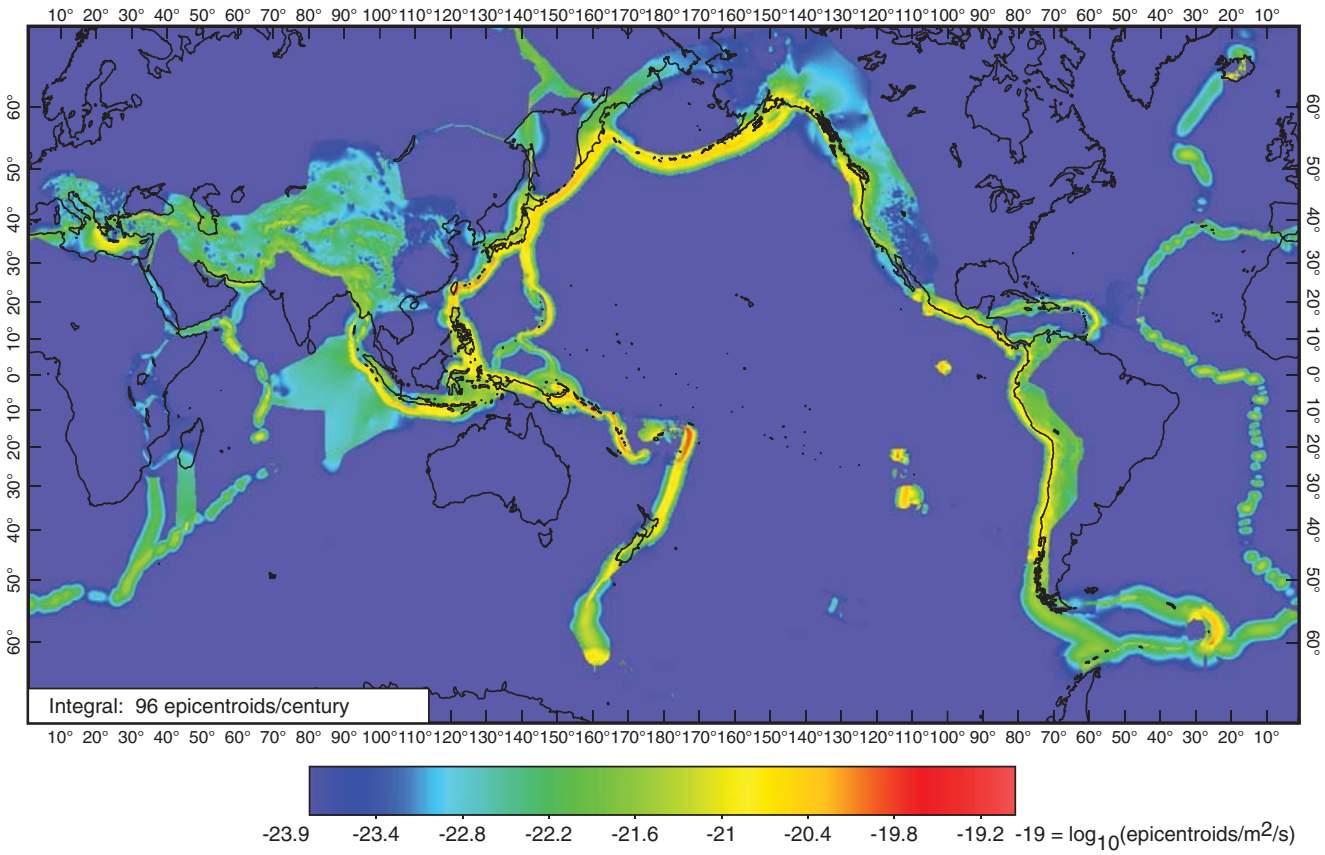

Figure 3. A Mercator projection of the SHIFT-GSRM2e tectonic forecast of the rate densities of shallow epicentroids at or above threshold $m_T = 8.00$ for years 2005 and after. Rate densities are expressed as epicentroids per square meter per second, including aftershocks. The colors use a logarithmic scale to express variations across ~5 orders of magnitude from peak subduction zone rates to very low intraplate rates. In this model, fast-spreading ridges and their transform faults do not create any seismicity at $m > 8$. Slow-spreading ridges produce occasional m 8 epicentroids but only along their transform faults. The contributions of continental rifts are visible but very small. These features persist in the preferred model version f.

mean, over all test earthquakes, of the base 2 logarithm of the ratio of conditional probability density of the cell in which the test earthquake epicentroid occurred to the mean conditional probability density in the whole forecast region. Thus, I_1 is the mean number of binary bits of information gain per actual test earthquake over an ignorant model that has only a single global earthquake rate. These information scores have several advantages: scores are independent of any difference between the total counts of forecast and test earthquakes, no simulated virtual catalogs (based on the forecast) are needed, no random perturbations of the test catalog are needed, declustering is not needed, scores are independent of forecast cell sizes (so long as the global map of conditional probabilities is unchanged), the scores have normal distributions (if multiple nonoverlapping tests are performed) when the number of test earthquakes is large (Kagan, 2009), and the results are on a simple absolute scale (i.e., binary bits of information gain). Results of these tests are displayed in Table 1.

We also scored each forecast using the ASS of Zechar (2008) and Zechar and Jordan (2008). The ASS test uses the earthquake rates forecast for cells only to rank the cells, after which a Molchan diagram of fraction-of-failures-to-predict versus fraction-of-ranked-forecast-domain-considered is constructed. The ASS is the normalized area above the descending curve, and high values are best.

To make a meaningful comparison between the ASS scores of these two kinds of forecast, they should be tested on the same spatial grid. For this purpose, we resampled all the SHIFT-GSRM2x forecast variants onto the $0.1° \times 0.1°$ global grid of the Kagan and Jackson (2014) smoothed-seismicity forecast (described above). Because this finer spatial grid is well aligned with the GSRM2.1 grid of cells, 80% of the seismicity forecast values (when expressed as epicentral rate densities, in epicentroids $m^{-2} \cdot s^{-1}$) can simply be transferred from the GSRM2.1 grid of SHIFT-GSRM2x to the finer grid. The remaining 20% of resampled values are equally weighted averages of the epicentral rate densities forecast for the western cell and the eastern cell in an adjacent SHIFT-GSRM2x pair of cells that are straddled by a single $0.1° \times 0.1°$ cell.

If the fraction-of-ranked-forecast domain is measured by summed map areas, then all reasonable forecasts have high scores and discrimination may be difficult. To make the test more challenging, and perhaps more informative, we follow the recommendation of Zechar and Jordan (2008) to define

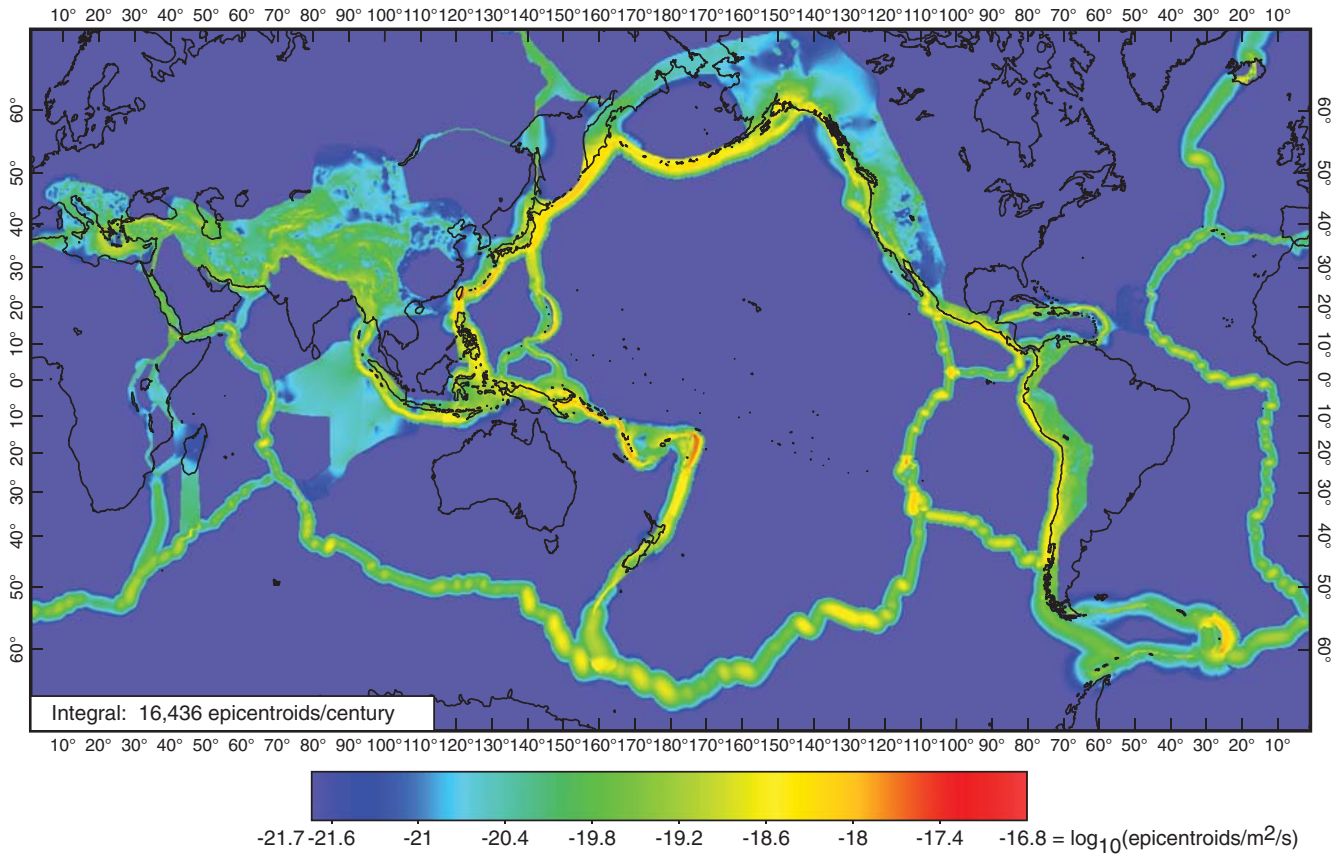

Figure 4. SHIFT-GSRM2f tectonic forecast of the rate densities of shallow epicentroids at or above threshold $m_T = 5.767$ for years 2005 and after. Map conventions as in Figure 3. This is our preferred tectonic model, shown at the completeness threshold for the shallow part of the Global CMT catalog.

the fraction-of-forecast domain by prior relative intensity (RI), which is just the seismicity rate of each cell in the pre-test part of the catalog. Thus, this version of the ASS test is a measure of quality of forecast map pattern relative to binned-but-unsmoothed extrapolation of past seismicity. Clearly, results depend on spatial cell sizes. One may wonder how this test behaves when the spatial cells are very small (i.e., $0.1° \times 0.1°$ in our test) and the magnitude threshold is high (i.e., 7.0, as in the right column of Table 1) so that prior RI is 0 in most cells and only 1 in most of the remaining cells. To minimize this problem, we use RI measured above completeness threshold m 5.767 in 1977–2004 (4602 earthquakes) as the reference RI model for both the m 5.767+ and the m7+ forecast tests.

Discussion

Test results in Table 1 show that specificity and success both increase with model sophistication in the series SHIFT-GSRM2a through SHIFT-GSRM2f. Models a and b, which lack spatial smoothing of the activity of all offshore plate boundaries or subduction plate boundaries, respectively, are systematically deficient because they interpret too many calibration-period earthquakes as intraplate (20.8% and 3.8%, respectively) and then project this artificially high intraplate seismicity over the whole Earth. In contrast, models c–f attribute only 1.35% of calibration-period earthquakes to intraplate seismicity, and our tests show that this is more realistic. Models d–f add very valuable distinctions between tectonic zones. Some of the improvement in test scores is due to the change from 2 calibrated degrees of freedom to 5 (reflecting the increased number of tectonic zones employed); however, a more important part is due to the recognition that different tectonic environments have different coupled seismogenic thicknesses and corner magnitudes. The importance of this latter improvement is demonstrated by the greater increase in success for the m 7+ forecasts as compared to the m 5.767+ forecasts (Table 1). The incremental improvement in the last variant (f, relative to e) is mainly in I_0 specificity; there is only a tiny improvement in I_1 success. The added feature in variant f (velocity dependence of coupled thickness in megathrusts) is based on a study (Bird *et al.*, 2009) that used all available years of the Global CMT catalog, not just the eight years of these tests, and demonstrated high statistical significance. This feature has potential to add valuable information to tectonic forecasts at the m 8+ level, although seismic catalogs are currently too short to conduct tests of that potential. Therefore, we designate the most complete model variant, SHIFT-GSRM2f, as our preferred tectonic model.

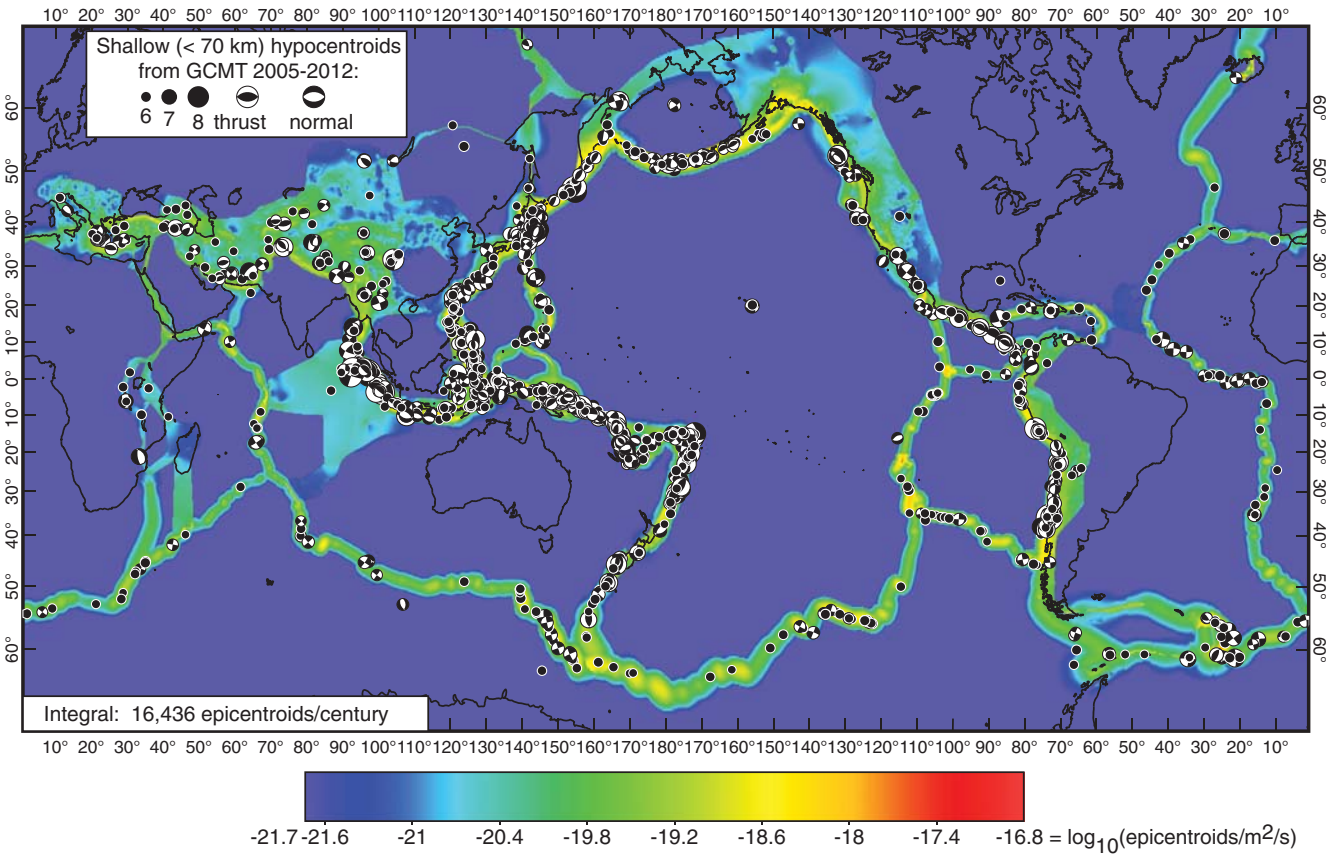

Figure 5. Comparison of tectonic model SHIFT-GSRM2f (exactly as in Fig. 4) to 1694 epicentroids corresponding to shallow hypocentroids at or above threshold $m_T = 5.767$ from years 2005–2012 in the Global CMT catalog. The focal mechanism is shown on the lower focal hemisphere for events with magnitude $m > 6.0$. If intraplate seismicity was correctly estimated as 1.35%, then about 23 of these test events would be expected to occur in deep-blue intraplate areas; the actual intraplate count was 18. See Table 1 for scores.

Some perspective on the meaning of the I_1 success scores can be obtained, in bootstrap fashion, by considering the range of contributions from individual test earthquakes before they are averaged together. For example, in testing our preferred model against shallow Global CMT $m\,5.767+$ earthquakes in 2005–2012, individual earthquake contributions to I_1 went from -5.80 (for each intraplate earthquake in an unexpected place) to $+10.48$ (for yet another subduction earthquake off Taiwan, where we assigned very high conditional probabilities). The average success of $+3.68$ is thus much better than the worst possible result, which would be -5.80 (among models which share the same level of intraplate seismicity). The high-side range of achievable successes is less clear. One benchmark is the model specificity of $I_0 = 4.042$; if the future seismicity of the Earth were to exactly follow the model (except for aleatory finite-sample effects), then future I_1 successes should average 4.042. Yet this is not an upper limit; one occasionally gets results of $I_1 > I_0$ when test earthquakes occur in the most notorious subduction zones and (temporarily) avoid all plate boundaries of lower forecast seismicity. Also, one can conceive of time-dependent models that would have both higher specificity and higher success (i.e., if they were successful earthquake predictions); however, whether a stationary forecast of very high specificity and success could even exist is unclear, as it depends on the relative frequency of characteristic repeating earthquakes, which is currently controversial. A necessary caveat is that all of this discussion relates to the context of global modeling; if the model domain were restricted to a high-seismicity region like Japan or California, then the range of potential successes would be different.

Another natural question is how the success of this preferred model relates to that of our previous model, SHIFT-GSRM of Bird *et al.* (2010). This previous model had specificity of $I_0 = 3.951$ and (when tested against shallow Global CMT earthquakes of $m\,5.767+$ in 2005–2012) scores a success of $I_1 = 3.68$. Thus, it seems there has been no improvement in scores. However, it must be remembered that the original GSRM model of Kreemer *et al.* (2003) used plate boundaries of variable width, and in the oceans these widths were primarily based on the map of previous Global CMT epicentroids. That map pattern was inherited by the previous seismicity model SHIFT-GSRM of Bird *et al.* (2010), which was, in effect, a hybrid tectonic/smoothed-seismicity model. In creating the recent GSRM version 2.1, Kreemer *et al.* (2014) elected to represent all offshore plate boundaries (other than certain subduction zones) as thin bands of single cell width (i.e., 0.20° or 0.25°). In doing this, they ceased to rely on

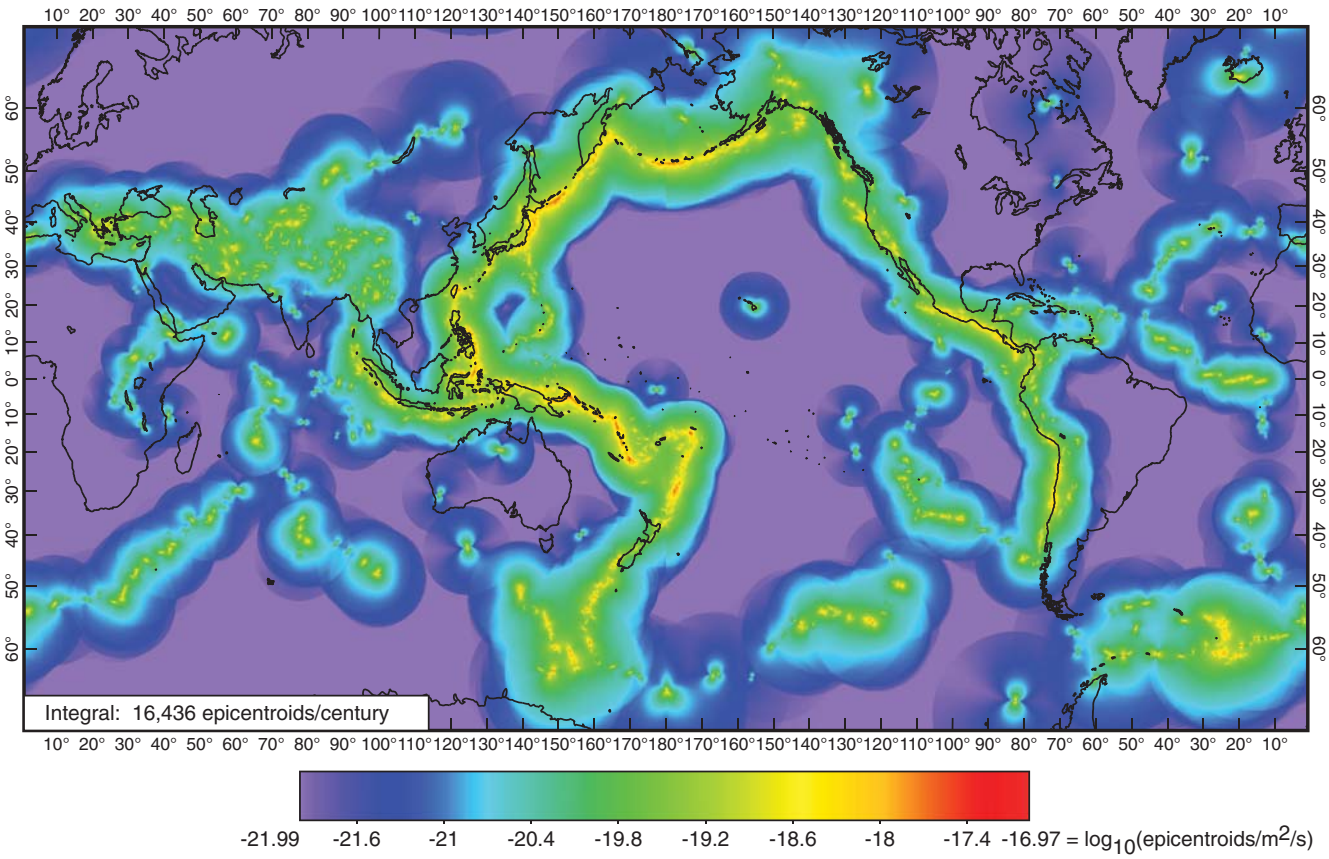

Figure 6. Smoothed-seismicity forecast of the rate densities of shallow epicentroids at or above threshold $m_T = 5.767$ for years 2005 and after, by Yan Y. Kagan and David D. Jackson. See Kagan and Jackson (1994, 2010a,b, 2011, 2014) for construction methods. The map conventions and color scale are as in Figures 3–5. Scores for this model are also reported in Table 1.

the map of instrumental seismicity, creating a conceptually simpler and more purely tectonic model. What these I scores show is that the improvement in forecasting seismicity on land (due to more and more precise GPS data and better algorithms) is approximately offset by its lesser success in forecasting offshore seismicity (when the map of prior epicentroids is replaced by generic plate-boundary bands of constant width). This simplification of the tectonic model is actually a step forward, because now we can more easily and thoroughly investigate whether adding a smoothed-seismicity component to the present tectonic model will yield further improvement. We will present such work in a future article.

The ASS statistics in Table 1 generally parallel the I_1 successes, even though these two measures compare our model performance with different null models (uniform seismicity for I_1; prior RI for ASS). The ASS scores of our preferred model SHIFT-GSRM2f and the smoothed-seismicity model are virtually identical at both magnitude thresholds, so from this ASS-test perspective our new model is comparable in performance to the alternative smoothed-seismicity method. In the future, it might be interesting to use a modified version of ASS for pairwise comparison of all models in a class, rather than harboring doubts about whether the ASS statistics might contain potential biases related to the deficiencies of the instrumental record of RI.

Still, the smoothed-seismicity forecast of Kagan and Jackson (2014) has the best success (I_1 in Table 1) for threshold $m \geq 5.767$. (Our preferred model f performs equally well at threshold $m \geq 7.00$, but this latter comparison is less definitive due to the limited number of $m \geq 7.00$ test earthquakes in eight years.) Consequently, we do not advocate our tectonic forecast as a replacement for smoothed-seismicity methods of forecasting seismicity, but rather as a supplement. There are at least three reasons why tectonic forecasts may add value, even if they are not the highest scoring in a particular eight-year test:

1. Design of buildings and other critical infrastructure requires forecasts that look ahead for the expected life of the structure, perhaps 50–200 years. Part of the success of smoothed-seismicity forecasts in short-term tests is due to the continuation into the test period of aftershock sequences that started in the learning period. However, this contribution to their success can be expected to fade as the forecast time window gets longer. Perhaps, somewhere within the range of 8–200 years of forecast window length, tectonic forecasts may become more accurate?
2. The data sources, assumptions, and algorithmic pathways of tectonic forecasts and smoothed-seismicity forecasts are fundamentally different. Therefore, each can serve

Table 1
Retrospective Global Tests against Shallow Global Centroid Moment Tensor 2005–2012

Model	$m \geq 5.767$ ($N = 1694$)			$m \geq 7.000$ ($N = 90$)		
	Specificity (I_0)	Success (I_1)	Area Skill Score	Specificity (I_0)	Success (I_1)	Area Skill Score
SHIFT-GSRM2a	3.355	2.50	0.889	3.304	2.3	0.89
SHIFT-GSRM2b	3.436	3.10	0.945	3.422	2.5	0.90
SHIFT-GSRM2c	3.548	3.27	0.958	3.544	3.2	0.96
SHIFT-GSRM2d	3.833	3.47	0.961	4.101	3.8	0.97
SHIFT-GSRM2e	3.914	3.67	0.965	4.228	3.9	0.97
SHIFT-GSRM2f	4.042	3.68	0.965	4.357	3.8	0.97
Kagan and Jackson smoothed seismicity	3.829	3.93	0.964	3.829	3.8	0.97

as a check against systematic and random errors in the other, and critical societal decisions should be informed by both estimates.

3. The specific geographic areas of greatest discrepancy between tectonic and smoothed seismicity may be fruitful sites for focused research. For example, where the tectonic forecast exceeds the smoothed-seismicity forecast, there may be faults of special petrology or geometry that are creeping aseismically. This could be checked with local high-resolution geodetic and seismic studies. Future research into the physics of aseismic creep could benefit from knowing the geography and geology of creeping areas. Also, tectonic forecasts of seismicity could then be further improved by special treatment of creeping faults using locally reduced coupled-thickness coefficients.

Global Earthquake Rates at High Magnitudes

The previous discussion has focused entirely on evaluating the map patterns of forecasts, independent of their global earthquake rates. A good reference model should also have a global frequency–magnitude distribution that is not falsified by the best and longest catalogs. However, our preferred model SHIFT-GSRM2f forecasts global earthquake rates in a certain range of high magnitudes ($7.65 < m < 9.0$), which are slightly above those in the Global CMT catalog (Fig. 7). This probably occurs because the SHIFT forecast is based on frequency–magnitude curves for different classes of plate boundary estimated by Bird and Kagan (2004). Although their primary source for calibrations was Global CMT 1977–2002, they did not always accept the low apparent corner magnitudes (in nonorogen regions, in those years) from the Global CMT catalog; instead, they broadened the search for the largest earthquakes to include orogens and to include the years 1900–1976 from the $m\,7+$ catalog of Pacheco and Sykes (1992). It is possible that this caused them to overestimate some corner magnitudes through reliance on inaccurate magnitudes of a few critical early events. Supporting this view, the new International Seismological Centre-Global Earthquake Model catalog (Storchak *et al.*, 2012), which provides calibrated estimates of moment magnitudes for all listed events in 1918–1976, gives a frequency–magnitude curve that is compatible with that of the Global CMT for 1977–2012 (Fig. 7). Another consideration is that our preferred SHIFT model gives a global shallow seismic moment rate (4.34×10^{14} N·m·s^{-1}) that is 36% higher than the real seismic moment rate (3.21×10^{14} N·m·s^{-1}) from 1 January 1918 to 31 December 2013 (Fig. 8). From this point of view, our SHIFT model may forecast too much activity at high magnitudes.

Another view might be that the years 1977–2003, which comprise ~72% of the Global CMT record, were unusually quiet and not representative. For example, the largest earthquake during that time had $m = 8.41$ (23 June 2001, off Peru). Global earthquake rates have been distinctly higher in years 2004–2012 (Fig. 7), especially at high magnitudes. The rates before and after 2004 cannot be reconciled in a simple Poisson model of independent earthquakes; it is necessary to decluster the catalog (Kagan, 2010) and discount earthquakes according to their probability of being aftershocks, after which it can be argued the model mainshock rates of the two periods are not distinguishable with high confidence. However, as the distribution of total earthquake counts (over consecutive time windows) does not follow the binomial distribution, but instead follows the broader negative-binomial distribution (Kagan, 2010) with at least one additional unknown parameter, we simply do not know how much the Global CMT global rate in 1977–2012 might differ from the true long-term rate. From this point of view, it might be reasonable to accept the SHIFT frequency–magnitude curve as a plausible merger of the best available information from different catalogs and different decades.

Conclusions and Prospects

In this project, we used the latest update (version 2.1) of the GSRM of Kreemer *et al.* (2014) and converted it to a global long-term model of shallow seismicity as a function of latitude, longitude, and magnitude. This was done under the assumptions that long-term seismic moment release rates will be proportional to tectonic deformation rates and that well-studied simple plate boundaries can serve as analogs for more complex zones, if we partition their deformation and apply seismicity coefficients from the most appropriate plate-boundary class to each part. We developed the forecast in six versions of increasing complexity and realism to show that each added factor either enhances forecasting skill or,

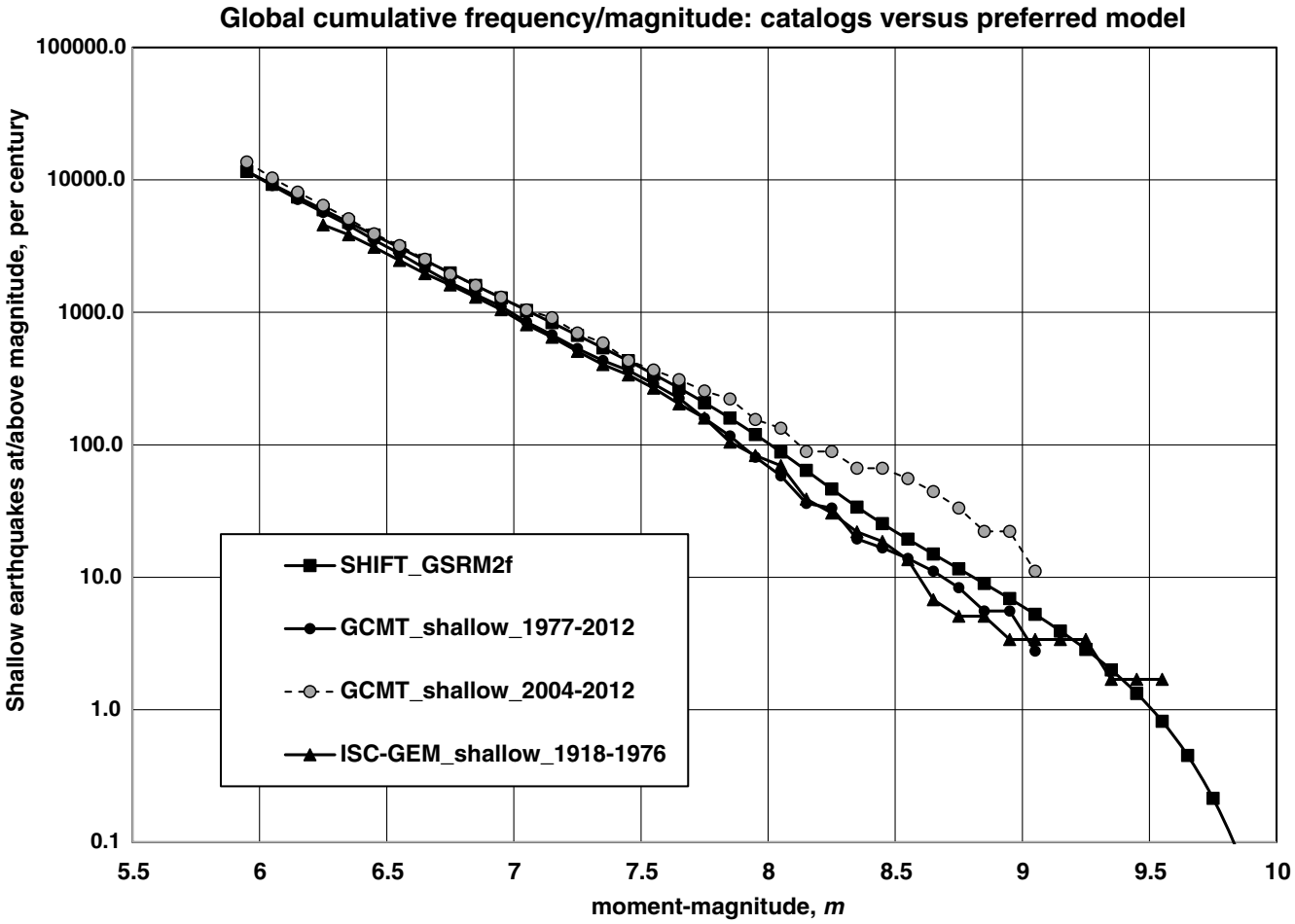

Figure 7. Global cumulative frequency–magnitude distribution of preferred model SHIFT-GSRM2f compared with the distributions of two global seismic catalogs: the International Seismological Centre-Global Earthquake Model (ISC-GEM) catalog of Storchak *et al.* (2012) for years 1918–1976 (including its supplement) and the Global Centroid Moment Tensor (CMT) catalog of Ekström *et al.* (2012) for years 1977–2012. The increased Global CMT rate in 2004–2012 is also shown with a dashed curve. All curves are normalized to 100 years of observation and are restricted to shallow earthquakes, with hypocentroids no deeper than 70 km.

at least, does not detract. Our preferred tectonic forecast is available (and will be submitted) for independent prospective testing, which will help to judge the validity of our assumptions. However, our own retrospective testing has already shown that this model performs about as well as one mature, optimized, smoothed-seismicity model. Thus, available GPS data (supplemented by a plate-tectonic model, especially offshore) now have about the same value for forecasting conditional probabilities (map patterns) of future earthquakes as modern seismic catalogs.

An important change in GSMR2.1 (in addition to the greatly improved GPS dataset on land) was the deliberate simplification of most offshore plate boundaries to narrow lines. This eliminated the somewhat subjective use of instrumental seismicity maps to locally vary the widths of plate boundaries that was a feature of the original GSRM of Kreemer *et al.* (2003) and was inherited by the previous global seismicity model of Bird *et al.* (2010). Although we found in this study that we had to smooth the activity of these offshore plate boundaries to obtain good forecast scores, we have now done this in a more systematic way that uses less than a dozen parametric measures from the instrumental seismic catalog, but not its complete map. Thus, the present update more closely approaches the abstract ideal of being purely tectonic, whereas the previous SHIFT-GSRM seismicity model of Bird *et al.* (2010) had a greater component of future projection of past instrumental seismicity in its map patterns.

This sets the stage for systematically and formally considering the potential added value of hybrid models with both tectonic and smoothed-seismicity components, which we intend to present in a future article. Practical earthquake forecasting may require all relevant information to be combined; yet, scientific hypothesis testing is more straightforward when distinct kinds of information are also developed separately into alternative models.

Data and Resources

The files necessary to fully represent our seismicity forecasts are much larger than the files needed to create them;

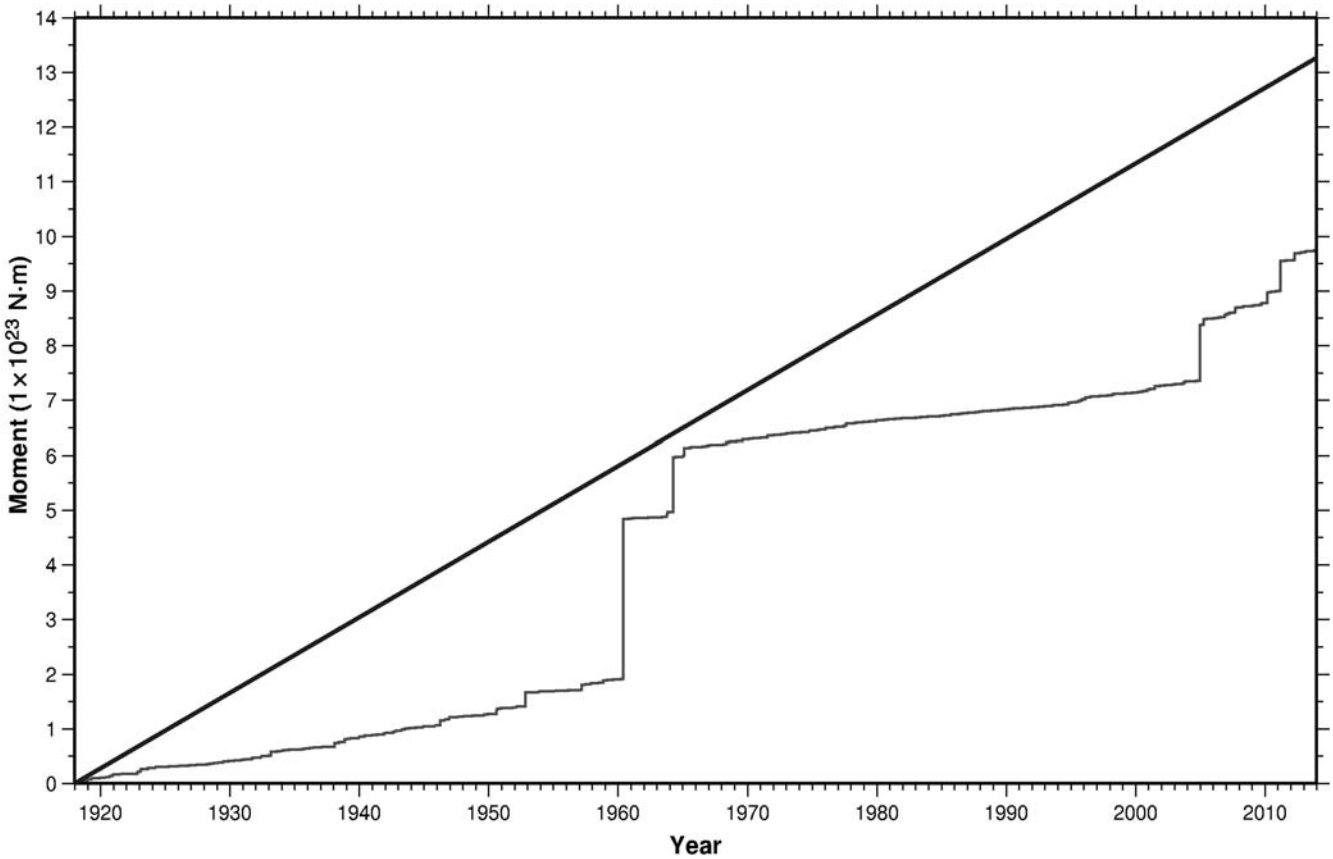

Figure 8. The stair-stepped graph line shows the global cumulative seismic moment release from shallow (≤70 km below sea level) hypocentroids in the ISC-GEM catalog of Storchak *et al.* (2012) from 1 January 1918 until 31 December 1976 and the Global CMT catalog of Ekström *et al.* (2012) from 1 January 1977 until 31 December 2013. The largest step, in 1960, represents the seismic moment of the *m* 9.64 Valdivia subduction earthquake offshore Chile. We increased the moment of the 2004 Sumatra earthquake to 1.0×10^{23} N·m (*m* 9.3) to be consistent with the latest findings (Stein and Okal, 2005, 2007; Tsai *et al.*, 2005). The straight diagonal line is the constant shallow-earthquake seismic moment rate implied by our preferred model SHIFT-GSRM2f.

therefore we provide source code and necessary input datasets. (Also, source code must be provided when a forecast is registered at Collaboratory for the Study of Earthquake Predictability [CSEP], so this is an emerging standard practice.) Ⓔ FORTRAN 90 source code is available in two versions (SHIFT_GSRM2x.f90 and SHIFT_GSRM2f_for_CSEP.f90), included as programs S1 and S2 of the electronic supplement to this paper, respectively. The former program computes any of the six model variants and outputs gridded epicentral rate densities (epicentroids per square meter per second) above a user-selected magnitude threshold, in gridded data (GRD) format, on a 0.20° × 0.25° grid. Effective coupled thickness of cells is also output in GRD format. The GRD file format is documented at http://peterbird.name/guide/grd_format.htm (last accessed November 2014); in this project, each GRD file is of managable size (13 MB), but a large number of files would be needed to completely represent the variation of forecast seismicity with threshold magnitude. The latter program computes only the preferred model (f) and provides epicentral rates in a set of 31 magnitude bins, in CSEP's own XML format, on a 0.1° × 0.1° grid; this output file will be about 3.7 GB in size. GSRM2.1 average strain rates of cells may be obtained from the website of the first author (P. B.).

The Global CMT catalog is available at http://www.globalcmt.org; this database was downloaded from www.globalcmt.org/CMTfiles.html as monthly summaries. (Our last access for downloading calibration and test catalogs occurred April 2013; our last access for computing the 96-year moment rate shown in Fig. 8 occurred May 2014.) For use with our programs, this catalog (or any other) must be expressed in EarthQuake Catalog (EQC) format, which was documented by Bird and Kagan (2004) and described at http://peterbird.name/oldFTP/2003107-esupp/eqc_format.pdf (last accessed November 2014). Input files describing plate boundaries (PB2002_steps.dat) and tectonic zones (PB2002_tectonic_zones.grd) are available as electronic supplements to Bird (2003) and Kagan *et al.* (2010), respectively. A logical GRD file (underwater.GRD) based on the digital elevation model named ETOPO5 is available from the website of the first author (http://peterbird.name/oldFTP/SHIFT_GSRM2_forecast/underwater.grd.zip, last accessed May 2014).

For prospective testing, all available catalog years should be used for calibration. This will change the global seismicity rate of the model, as well as the balance between zones. For example, computing a model based on Global CMT years 1977–2013 gives a global shallow seismicity rate that is 6.5% greater than when only years 1977–2004 are used for calibration; this is because of the 26.7% increase in global shallow seismicity that occurred at the end of 2004.

The applicable time window of this forecast starts immediately after the calibration period and is conceptually open (infinitely extended) into the future, although for any currently existing CSEP test class an end date must be chosen. In practical terms, we might expect that continued collection of GPS data and improved knowledge of plate boundaries will eventually result in a replacement model.

Acknowledgments

C. K. was supported by a grant from the Global Earthquake Model (GEM) Foundation and by National Science Foundation (NSF) Grant EAR-0911754. Yan Kagan and David Jackson provided a customized version of their smoothed-seismicity forecast for comparison purposes. All opinions, expressed or implied, are those of the authors and do not reflect official positions of NSF or the GEM Foundation.

References

Bird, P. (2003). An updated digital model of plate boundaries, *Geochem. Geophys. Geosyst.* **4,** no. 3, 1027, doi: 10.1029/2001GC000252.

Bird, P., and Y. Y. Kagan (2004). Plate-tectonic analysis of shallow seismicity: Apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings, *Bull. Seismol. Soc. Am.* **94,** no. 6, 2380–2399.

Bird, P., and Z. Liu (2007). Seismic hazard inferred from tectonics: California, in *Regional Earthquake Likelihood Models*, S. E. Hough and K. B. Olsen (Guest Editors), special issue, *Seismol. Res. Lett.* **78,** no. 1, 37–48.

Bird, P., Y. Y. Kagan, D. D. Jackson, F. P. Schoenberg, and M. J. Werner (2009). Linear and nonlinear relations between relative plate velocity and seismicity, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3097–3113, doi: 10.1785/0120090082.

Bird, P., C. Kreemer, and W. E. Holt (2010). A long-term forecast of shallow seismicity based on the Global Strain Rate Map, *Seismol. Res. Lett.* **81,** no. 2, 184–194.

Ekström, G., M. Nettles, and A. M. Dziewonski (2012). The Global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. Int.* **200/201,** 1–9.

Giardini, D., G. Grunthal, K. Shedlock, and P. Zhang (1999). Global Seismic Hazard Map, 1:35,000,000. *Global Seismic Hazard Assessment Project*, United Nations International Decade of Natural Disaster Reduction, International Lithosphere Program.

Hanks, T. C., and H. Kanamori (1979). A moment magnitude scale, *J. Geophys. Res.* **84,** no. B5, 2348–2350.

Jackson, D. D., and Y. Y. Kagan (1999). Testable earthquake forecasts for 1999, *Seismol. Res. Lett.* **70,** no. 4, 393–403.

Kagan, Y. Y. (2003). Accuracy of modern global earthquake catalogs, *Phys. Earth Planet. In.* **135,** nos. 2/3, 173–209.

Kagan, Y. Y. (2009). Testing long-term earthquake forecasts: Likelihood methods and error diagrams, *Geophys. J. Int.* **177,** no. 2, 532–542.

Kagan, Y. Y. (2010). Statistical distributions of earthquake numbers: Consequence of branching process, *Geophys. J. Int.* **180,** no. 3, 1313–1328.

Kagan, Y. Y., and D. D. Jackson (1994). Long-term probabilistic forecasting of earthquakes, *J. Geophys. Res.* **99,** 13,685–13,700.

Kagan, Y. Y., and D. D. Jackson (2000). Probabilistic forecasting of earthquakes, *Geophys. J. Int.* **143,** no. 2 (Leon Knopoff volume), 438–453.

Kagan, Y. Y., and D. D. Jackson (2010a). Short- and long-term earthquake forecasts for California and Nevada, *Pure Appl. Geophys.* **167,** nos. 6/7 (Frank Evison volume), 685–692, doi: 10.1007/s00024-010-0073-5.

Kagan, Y. Y., and D. D. Jackson (2010b). Earthquake forecasting in diverse tectonic zones of the globe, *Pure Appl. Geophys.* **167,** nos. 6/7 (Frank Evison volume), 709–719, doi: 10.1007/s00024-010-0074-4.

Kagan, Y. Y., and D. D. Jackson (2011). Global earthquake forecasts, *Geophys. J. Int.* **184,** no. 2, 759–776.

Kagan, Y. Y., and D. D. Jackson (2014). Statistical earthquake focal mechanism forecasts, *Geophys. J. Int.* **197,** no. 1, 620–629, doi: 10.1093/gji/ggu015.

Kagan, Y. Y., P. Bird, and D. D. Jackson (2010). Earthquake patterns in diverse tectonic zones of the globe, *Pure Appl. Geophys.* **167,** nos. 6/7 (Frank Evison volume), 721–741, doi: 10.1007/s00024-010-0075-3.

Kreemer, C., G. Blewitt, and E. C. Klein (2014). A geodetic plate motion and global strain rate model, *Geochem. Geophys. Geosyst.* doi: 10.1002/2014GC005407.

Kreemer, C., W. E. Holt, and A. J. Haines (2003). An integrated global model of present-day plate motions and plate boundary deformation, *Geophys. J. Int.* **154,** 8–34.

Pacheco, J. F., and L. R. Sykes (1992). Seismic moment catalog of large, shallow earthquakes, 1900–1989, *Bull. Seismol. Soc. Am.* **82,** 1306–1349.

Petersen, M. D., Y. Zeng, K. M. Haller, R. McCaffrey, W. C. Hammond, P. Bird, M. Moschetti, Z. Shen, J. Bormann, and W. Thatcher (2014). Geodesy- and geology-based slip-rate models for the western United States (excluding California) national seismic hazard maps, *U.S. Geol. Surv. Open-File Rept. 2013-1293*, 80 pp., doi: 10.3133/ofr20131293.

Stein, S., and E. A. Okal (2005). Speed and size of the Sumatra earthquake, *Nature* **434,** 581–582, doi: 10.1038/434581a.

Stein, S., and E. A. Okal (2007). Ultralong period seismic study of the December 2004 Indian Ocean earthquake and implications for regional tectonics and the subduction process, *Bull. Seismol. Soc. Am.* **97,** S279–S295.

Storchak, D. A., D. Di Giacomo, I. Bondar, J. Harris, E. R. Engdahl, W. H. K. Lee, A. Villasenor, P. Bormann, and G. Ferrari (2012). ISC-GEM Global Instrumental Earthquake Catalog (1900–2009), *GEM Tech. Rept. 2012-01*, Global Earthquake Model Foundation, Pavia, Italy, 128 pp., available at http://www.globalquakemodel.org/resources/publications/?page=2 (last accessed November 2014).

Tsai, V. C., M. Nettles, G. Ekström, and A. Dziewonsk (2005). Multiple CMT source analysis of the 2004 Sumatra earthquake, *Geophys. Res. Lett.* **32,** L17304, doi: 10.1029/2005GL023813.

Zechar, J. D. (2008). Methods for evaluating earthquake predictions, *Ph.D. dissertation*, University of Southern California, Los Angeles.

Zechar, J. D., and T. H. Jordan (2008). Testing alarm-based earthquake predictions, *Geophys. J. Int.* **172,** 715–724, doi: 10.1111/j.1365-246X.2007.03676.x.

Department of Earth, Planetary, and Space Sciences
University of California
Los Angeles, California 90095-1567
pbird@epss.ucla.edu
(P.B.)

Nevada Bureau of Mines and Geology
University of Nevada
Reno, Nevada 89557-0178
kreemer@unr.edu
(C.K.)

Manuscript received 8 May 2014;
Published Online 23 December 2014