GFZ
Helmholtz Centre
POTSDAM

INSTITUTE OF EARTH AND ENVIRONMENTAL SCIENCE

Universität
Potsdam

Constructing global stationary seismicity models from the long-term balance of interseismic strain measurements and earthquake-catalog data

José Antonio Bayona Viveros

Statement of originality

I hereby certify that this thesis is entirely my own original work. Any use of the works of any other author is properly acknowledged.

José Antonio Bayona Viveros

Potsdam, August 2020.

To Adrianita, the child.

Abstract

One third of the world's population lives in areas where earthquakes causing at least slight damage are frequently expected. Thus, the development and testing of global seismicity models is essential to improving seismic hazard estimates and earthquake-preparedness protocols for effective disaster-risk mitigation. Currently, the availability and quality of geodetic data along plate-boundary regions provides the opportunity to construct global models of plate motion and strain rate, which can be translated into global maps of forecasted seismicity. Moreover, the broad coverage of existing earthquake catalogs facilitates in present-day the calibration and testing of global seismicity models. As a result, modern global seismicity models can integrate two independent factors necessary for physics-based, long-term earthquake forecasting, namely interseismic crustal strain accumulation and sudden lithospheric stress release.

In this dissertation, I present the construction of and testing results for two global ensemble seismicity models, aimed at providing mean rates of shallow (0-70 km) earthquake activity for seismic hazard assessment. These models depend on the Subduction Megathrust Earthquake Rate Forecast (SMERF2), a stationary seismicity approach for subduction zones, based on the conservation of moment principle and the use of regional "geodesy-to-seismicity" parameters, such as corner magnitudes, seismogenic thicknesses and subduction dip angles. Specifically, this interface-earthquake model combines geodetic strain rates with instrumentally-recorded seismicity to compute long-term rates of seismic and geodetic moment. Based on this, I derive analytical solutions for seismic coupling and earthquake activity, which provide this earthquake model with the initial abilities to properly forecast interface seismicity. Then, I integrate SMERF2 interface-seismicity estimates with earthquake computations in non-subduction zones provided by the Seismic Hazard Inferred From Tectonics based on the second iteration of the Global Strain Rate Map seismicity approach to construct the global Tectonic Earthquake Activity Model (TEAM). Thus, TEAM is designed to reduce number, and potentially spatial, earthquake inconsistencies of its predecessor tectonic earthquake model during the 2015–2017 period. Also, I combine this new geodetic-based earthquake approach with a global smoothed-seismicity model to create the World Hybrid Earthquake Estimates based on Likelihood scores (WHEEL) model. This updated hybrid model serves as an alternative earthquake-rate approach to the Global Earthquake Activity Rate model for forecasting long-term rates of shallow seismicity everywhere on Earth.

Global seismicity models provide scientific hypotheses about when and where earthquakes may occur, and how big they might be. Nonetheless, the veracity of these hypotheses can only be either confirmed or rejected after prospective forecast evaluation. Therefore, I finally test the consistency and relative performance of these global seismicity models with independent observations recorded during the 2014–2019 pseudo-prospective evaluation period. As a result, hybrid earthquake models based on both geodesy and seismicity are the most informative seismicity models during the testing time frame, as they obtain higher information scores than their constituent model components. These results support the combination of interseismic strain measurements with earthquake-catalog data for improved seismicity modeling. However, further prospective evaluations are required to more accurately describe the capacities of these global ensemble seismicity models to forecast longer-term earthquake activity.

Zusammenfassung

Ein Drittel der Weltbevölkerung lebt in Gebieten, in denen häufig Erdbeben mit zumindest geringen Schäden zu erwarten sind. Daher ist die Entwicklung und das Testen globaler Seismizitätsmodelle für verbesserte Schätzungen der Erdbebengefährdung und Planungen zur Vorbereitung auf Erdbeben für eine wirksame Minderung des Katastrophenrisikos von entscheidender Bedeutung. Derzeit bietet die Verfügbarkeit und Qualität geodätischer Daten entlang der Plattengrenzregionen die Gelegenheit, um globale Modelle der Plattenbewegung und der Dehnungsrate zu erstellen, die in globale Karten der prognostizierten Seismizität übersetzt werden können. Darüber hinaus erleichtert die breite Abdeckung bestehender Erdbebenkataloge in der heutigen Zeit die Kalibrierung und das Testen globaler Seismizitätsmodelle. Infolgedessen können moderne globale Seismizitätsmodelle zwei unabhängige Faktoren integrieren, die für eine physikbasierte Langzeit-Erdbebenvorhersage erforderlich sind, die Ansammlung interseismischer Krustenverformungen und die plötzliche Freisetzung von lithosphärischem Stress.

In dieser Dissertation stelle ich die Konstruktion und die Testergebnisse für zwei globale Ensemble-Seismizitätsmodelle vor, die darauf abzielen, mittlere Raten der Flachbebenaktivität (0-70 km) für die Bewertung der Erdbebengefährdung bereitzustellen. Diese Modelle hängen von dem Subduction Megathrust Earthquake Rate Forecast (SMERF2) ab, einem stationären Seismizitätsmodell für Subduktionszonen, das auf dem Prinzip der Erhaltung des Moments und der Verwendung regionaler "Geodäsie-zu-SeismizitätParameter wie Corner Magnitudes, seismogene Dicken und Subduktionsneigungswinkel basiert. Insbesondere kombiniert dieses Erdbebenmodell geodätische Dehnungsraten mit instrumentell aufgezeichneter Seismizität, um Langzeitraten sowohl des seismischen als auch des geodätischen Moments zu berechnen. Auf dieser Grundlage leite ich analytische Lösungen für die seismische Kopplung und Erdbebenaktivität ab, um mit diesem Erdbebenmodell, die Subduktionsseismizität richtig vorherzusagen. Dann integriere ich SMERF2-Schätzungen an Subduktionsrändern mit Erdbebenberechnungen in Nicht-Subduktionszonen, die von dem Modell SSeismic Hazard Inferred From Tectonics based on the second iteration of the Global Strain Rate Mapßur Erstellung des globalen Tectonic Earthquake Activity Model (TEAM) bereitgestellt werden. Daher ist TEAM darauf ausgelegt, die Anzahl und möglicherweise räumliche Vohersageinkonsistenzen seines tektonischen Erdbebenvorgängermodells im Zeitraum 2015-2017 zu reduzieren. Außerdem kombiniere ich dieses neue geodätische Erdbebenmodell mit einem globalen, geglätteten Seismizitätsmodell, um das World Hybrid Earthquake Estimates based on Likelihood Scores (WHEEL)-Modell zu erstellen. Dieses aktualisierte Hybridmodell dient als alternativer Ansatz zum Global Earthquake Activity Rate (GEAR1)-Modell zur Vorhersage langfristiger Raten flacher Seismizität überall auf der Erde.

Globale Seismizitätsmodelle liefern wissenschaftliche Hypothesen darüber, wann und wo Erdbeben auftreten können und wie groß sie sein können. Die Richtigkeit dieser Hypothesen kann jedoch erst nach prospektiven Tests bestätigt oder abgelehnt werden. Daher teste ich abschließend die Konsistenz und relative Leistung dieser globalen Seismizitätsmodellen gegen unabhängige Beobachtungen, die während des pseudo-prospektiven Evaluierungszeitraums 2014-2019 aufgezeichnet wurden. Hybride Erdbebenmodelle, die sowohl auf Geodäsie als auch auf Seismizität basieren, sind die informativsten Seismizitätsmodelle während des Testzeitraums, da beide höhere Informationswerte als ihre konstituierenden Modellkomponenten erhalten. Diese

Ergebnisse unterstützen die Kombination von interseismischen Dehnungsmessungen mit Erdbebenkatalogdaten für eine verbesserte Seismizitätsmodellierung. Es sind jedoch weitere prospektive Tests erforderlich, um die Kapazitäten dieser globalen Ensemble-Seismizitätsmodelle zur Vorhersage längerfristiger Erdbebenaktivitäten genauer zu bewerten.

Inhaltsverzeichnis

Kapitel 1

Introduction

1.1 Motivation

Earthquakes are the most devastating natural disasters (see Fig. 1.1), as 2.7 billion people, i.e. more than a third of the world's population (Roser et al., 2020), live in areas where seismicity causing at least slight damage is frequently expected (Marti et al., 2019). During the twentieth century, an annual rate of approximately 9000 fatalities was associated with the occurrence of earthquakes worldwide. More recently, 46000 casualties, along with an economic loss of 8 billion US dollars, have been annually derived from these low-frequency, high-impact events (Ritchie and Roser, 2014). By taking into consideration the global population growth observed since the industrial revolution, these numbers indicate that the world has increased its earthquake resilience over time, primarily due to the influence of science and technology on the development of modern societies. In particular, the premise that earthquakes do not cause fatalities, but rather unsafe infrastructures, has encouraged both the earthquake and engineering community to further minimize casualties and losses through, among others, improved seismicity modeling.
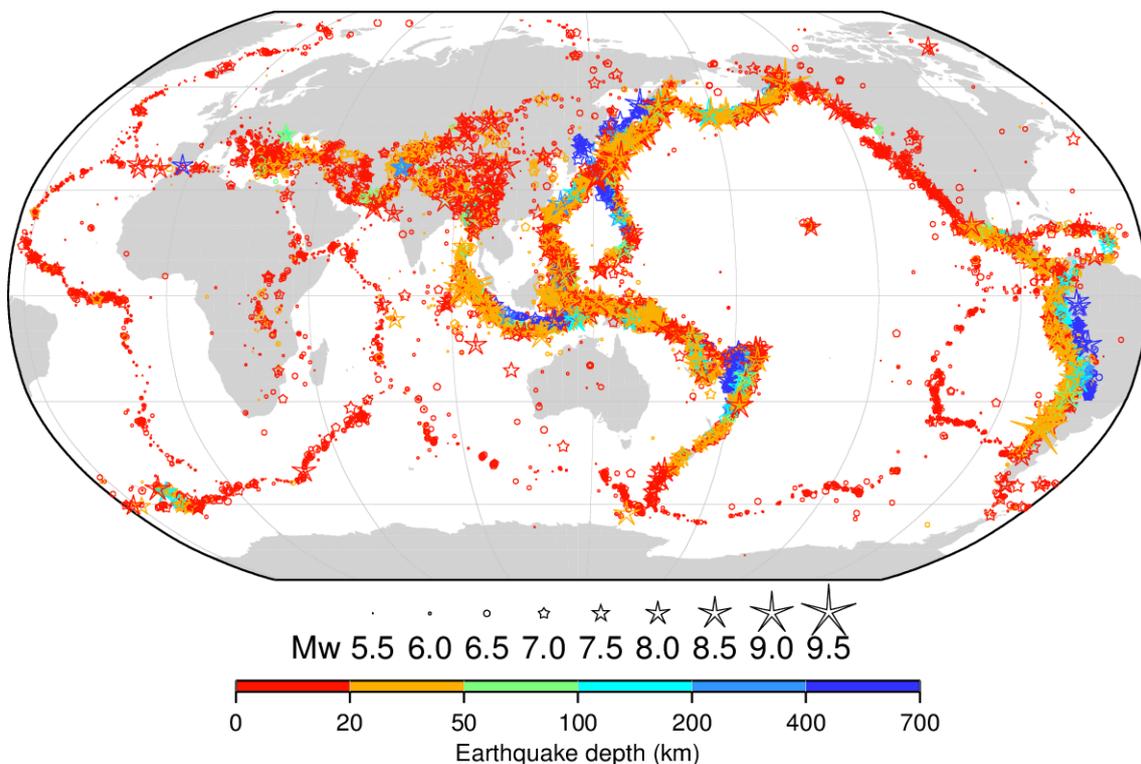
Abbildung 1.1: Map showing more than 28,000 earthquakes listed in the 1904-2014 International Seismological Centre-Global Earthquake Model (ISC-GEM) historical and instrumentally-recorded seismicity catalog.

Seismicity models are descriptions of earthquake activity, aimed at forecasting the rate, location, magnitude, confidence level and probability of occurrence of future earthquakes (Allen, 1976; Geller, 1997). Given a reliable earthquake forecast, societies have the opportunity to better design resilient buildings, and share their risk through insurance and other types of disaster preparedness (Bird and Kreemer, 2015). Thus, earthquake-rate models currently form one of the key ingredients for Probabilistic Seismic Hazard Analysis (PSHA; Cornell, 1968; Esteva, 1968), which extends the field of earthquake forecasting into the domain of ground motion modeling. The Global Seismic Hazard Map (GSHAP; Giardini et al., 1999), for instance, is based on a global seismicity model constructed from historical earthquake descriptions (pre-1900), instrumentally-recorded seismicity (1900-1999) and earthquake-source information. Traditionally, seismicity models are mainly based on the extrapolation of past earthquakes to the future and the characterization of specific seismic sources in terms of their geometries and slip rates. Broadly speaking, these models compute rates of forecasted seismicity by assuming the frequency-magnitude distribution of observed earthquakes to follow for example a characteristic earthquake behavior or a Gutenberg and Richter (1944) model. In addition, seismicity rates are constrained by estimates of corner magnitude, or upper magnitude limits, usually derived from earthquake records and paleoseismological studies. Three major limitations of this approach, however, are a) that seismicity catalogs are too short to include the largest possible earthquakes occurring along faults and plate boundaries (Bird et al., 2015), b) the scarce availability and quality of tectonic-fault data, and c) the lack of a better understanding of the physics of long-term earthquake forecasting.

a) Evidence from uplifted corals and tsunami deposits collected along the Pacific coast of Tohoku indicate that the average return period of megathrust events as large as the 2011 m_{w} = 9.1 Tohoku-oki earthquake is approximately 1,000 years (Minoura et al., 2001). Similarly, paleoseismological investigations of active faults distributed throughout the Trans-Mexican Volcanic Belt, Mexico (e.g. Langridge et al., 2013; Sunye-Puchol et al., 2015; Ortuño et al., 2019) suggest that mean recurrence times of $m > 7$ earthquakes in the region fluctuate between two and fifteen thousand years. Hence, the observation time-window offered by historical and instrumental seismicity catalogs, such as the Global Centroid Moment Tensor Project (CMT, Dziewonski et al., 1981; Ekström et al., 2012) or the International Seismological Centre-Global Earthquake Model (ISC-GEM, Storchak et al., 2013), turns out to be significantly shorter, in comparison with the recurrence rates of big earthquakes. As a result, estimates of seismic moment and forecasted seismicity may be either over- or underestimated, depending on the inclusion or absence of large events within the calibration datasets. In other words, seismicity models based on previous earthquake information may locally fail to forecast a representative level of activity.

b) The study and compilation of active and potentially active faults is useful to overcome the issue of seismicity-data sampling for long-term earthquake forecasting. The United States national seismic hazard model (i.e. Petersen et al., 2015) exemplifies well the great impact of fault characterization on the construction of seismicity models. Nevertheless, fault mapping is yet restricted to a few active continental regions, and it is almost inexistent on several oceanic floors (Bird and Kreemer, 2015).

c) The earthquake size distribution follows a power law, generally described by the Gutenberg-Richter relationship (Ishimoto, 1936; Kagan, 1999). This observational frequency-magnitude distribution, parameterized by the *a*- and *b*-values, has been commonly used as a statistical tool to estimate the recurrence of small and large earthquakes. Over the last decades, temporal and spatial variations in *a* and *b* have been extensively investigated to improve the predictability of future earthquake hazards. Statistically-significant changes in *b*-values, for instances, have been identified as precursors to big events (Smith, 1981), evidence of the fractal dimension of

earthquakes (Hirata, 1989), and indicators of stress variations across different tectonic regimes (Schorlemmer et al., 2005). Moreover, the seismicity level a has been found to empirically correlate with the relative plate velocity (Bird et al., 2009; Ide, 2013). Thus, these observations require physical descriptions that can better explain the causal mechanisms of global earthquake production. Therefore, this doctoral thesis aims to provide analytical solutions for earthquake activity, especially subduction megathrust seismicity, derived from the principle of conservation of moment. Specifically, computations of forecasted seismicity are based on the long-term balance of geodetic and seismic moment rates, obtained from interseismic strain measurements and earthquake-catalog information.

1.2 Global earthquake modeling

1.2.1 Smoothed-seismicity modeling

Since the mid 1960s, the creation and operation of the World-Wide Standardized Seismograph Network (WWSSN; Oliver and Murphy, 1971) allowed the acquisition of high-quality global earthquake data to increase. Given the usual level of global seismicity, these data have been used for various applications of observational seismology, such as the parameterization of seismic sources from waveform inversions (Dziewonski et al., 1981) and the calibration and testing of global earthquake-rate forecasts (Giardini et al., 1999). Long-term and short-term global seismicity forecasts rely on global homogenized catalogs of earthquakes, whose hypocentral locations and moment magnitude estimates are selected from local and global seismicity bulletins. Thus, the Kagan and Jackson (2011) Smoothed-Seismicity (KJSS) model uses the 1977-2008 global CMT catalog to average out the exceedance rates of 30,000 $m \geq 5.8$ earthquakes, and smooths their locations with an anisotropic kernel function:

$$f(r) = \frac{1}{\pi} \times \frac{1}{r^2 + r_s{}^2}, \tag{1.1}$$

from which r is the epicentroid distance from the center of a grid cell to an earthquake and $r_s = 2.5$ km is a spatial scale parameter.
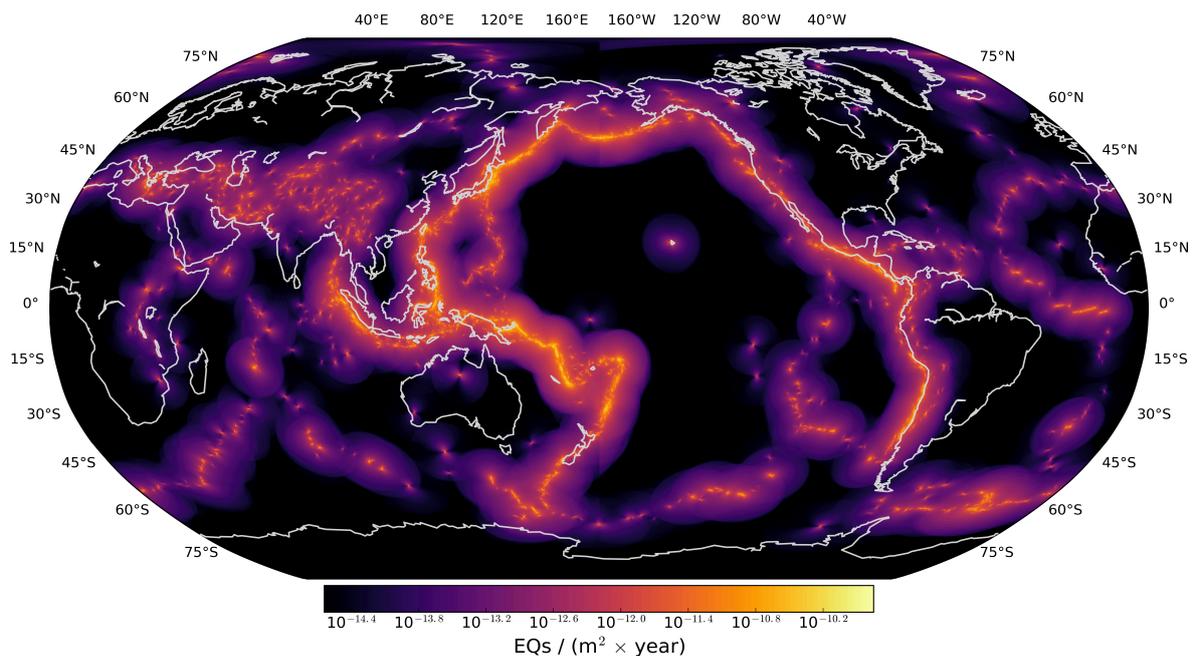
Abbildung 1.2: Forecast map of $m \geq 5.95$, $h \leq 70$ km global earthquake rates, derived from the Kagan and Jackson (2011) Smoothed-Seismicity (KJSS) model.

Empirical normalization factors are then applied as weightings to each earthquake, so that the total expected activity rate matches the total seismicity rate observed from the catalog. As a result, KJSS provides earthquake rates per unit area, time and magnitude to describe seismicity patterns worldwide (see Fig. 1.2).

1.2.2 Strain-based earthquake modeling

The theory of plate tectonics explains the kinematics of a set of rigid volumes of lithosphere, referred to as tectonic plates. Plate boundaries, i.e. subduction zones, seafloor-spreading ridges, continental rift areas and transform margins, cover approximately 15% of the Earth's surface (Gordon and Stein, 1992). Nevertheless, the most prominent seismic and volcanic activity, as well as the greatest portion of crustal deformation, are globally confined to these tectonic margins (Kreemer et al., 2002; Bird, 2003; DeMets et al., 2010). The increasing availability of geodetic measurements in plate boundary regions over the last decades has led to the creation of high-resolution models of plate motion and strain rate, which can be translated in present-day into global maps of forecasted seismicity. The Global Strain Rate Map (GSRM; Kreemer et al., 2003) is a velocity-gradient tensor field and interseismic-strain model, with a spatial resolution of 0.6° in longitude and 0.5° in latitude. This resolution stems from the interpolation of approximately 5,200 horizontal interseismic Global Positioning System (GPS) velocities, Quaternary fault slip rates and earthquake focal mechanisms. GSRM was designed to support the creation of geodetic-based seismicity models, as interseismic strain rates imply seismic-moment accumulation, and earthquakes indicate seismic-moment release (Scholz and Campos, 2012). In addition, interseismic strain measurements can be used to quantify earthquake potential without the need to determine the geometry and slip rates of every seismic source, as required by traditional earthquake-based models.

Authors like Kostrov (1974), Ward (1994) and Savage and Simpson (1997) were the pioneers to propose methods to relate geodetic strain rates with estimates of seismic moment. Essentially, these methods assume that the long-term seismic moment of any deforming volume of lithosphere linearly scales with interseismic strain rates. The proportionality relationship between these physical quantities depends on geomechanical parameters, such as the seismic coupling coefficient, the seismogenic thickness and the fault-dipping orientation. According to the degree of coupling, or the fraction of geodetic moment that will be seismically released, the seismic moment rate can then be related to rates of earthquake activity by assuming a particular magnitude-frequency distribution. Bird et al. (2010) translated GSRM into a global forecast of shallow ($h \leq 70$ km) seismicity at or above a moment or magnitude threshold $m \geq 5.66$. For this purpose, the modelers followed the hypotheses, equations and assumptions of the Seismic Hazard Inferred from Tectonics (SHIFT; Bird and Liu, 2007) model. Specifically, they used global "geodesy-to-seismicityconversion factors reported by Bird and Kagan (2004) and the tectonic regionalization scheme of Kreemer et al. (2002) for this translation. Thus, the authors produced the Seismic Hazard Inferred from Tectonics based on the Global Strain Rate Map (SHIFT_GSRM) seismicity model. In accordance with 1977–2009 retrospective test results, SHIFT_GSRM is consistent with earthquakes observed in active continental regions like the San Andreas fault system in southern California or the active faults in China. However, the raw forecast underestimates seismicity rates in subduction zones by a mean factor of approximately three during the same evaluation period. In this regard, Bird et al. (2010) explained that earthquake model underestimations are mainly due to geometric effects, poorly captured by the global mean subduction dip angle employed in the formulation. As a result, the SHIFT_GSRM2f authors compensate for the lack of specific (more regional) subduction dip angles in the model moment rate by applying empirical adjustment factors.

Four years later, Kreemer et al. (2014) published a major update of GSRM: the GSRM2.1 model. The new GSRM is a large improvement over its predecessor, because it incorporates 22,415 interseismic GPS velocities (almost five times more than GSRM), measured between 1 January 1996 and 31 December 2013. In this manner, GSRM2.1 offers a finer spatial strain-rate distribution than GSMR to be used in high-resolution seismicity forecasts, such as the global map of recurrence times for $m_\mathrm{w} = 7.5$ characteristic earthquakes, presented in the same study. Based on GSRM2.1, Bird and Kreemer (2015) generated the SHIFT_GSRM2f seismicity model. Among other improvements, SHIFT_GSRM2f uses a new tectonic regionalization framework (i.e. Kagan et al., 2010) and velocity-dependent seismogenic thicknesses for subduction margins and continental convergent zones (i.e. Bird et al., 2009). In this occasion, a SHIFT_GSRM2f forecast was retrospectively and pseudo-prospectively contrasted with actual 1977–2004 $m \geq$ 5.767, $h \leq 70$ km observed seismicity. As a result, the uncorrected forecast underpredicts global earthquake activity, mainly due to underestimations of subduction-zone seismicity. Hence, the SHIFT_GSRM authors empirically corrected the forecast again as the last step in constructing this seismicity model (see Fig. 1.3).
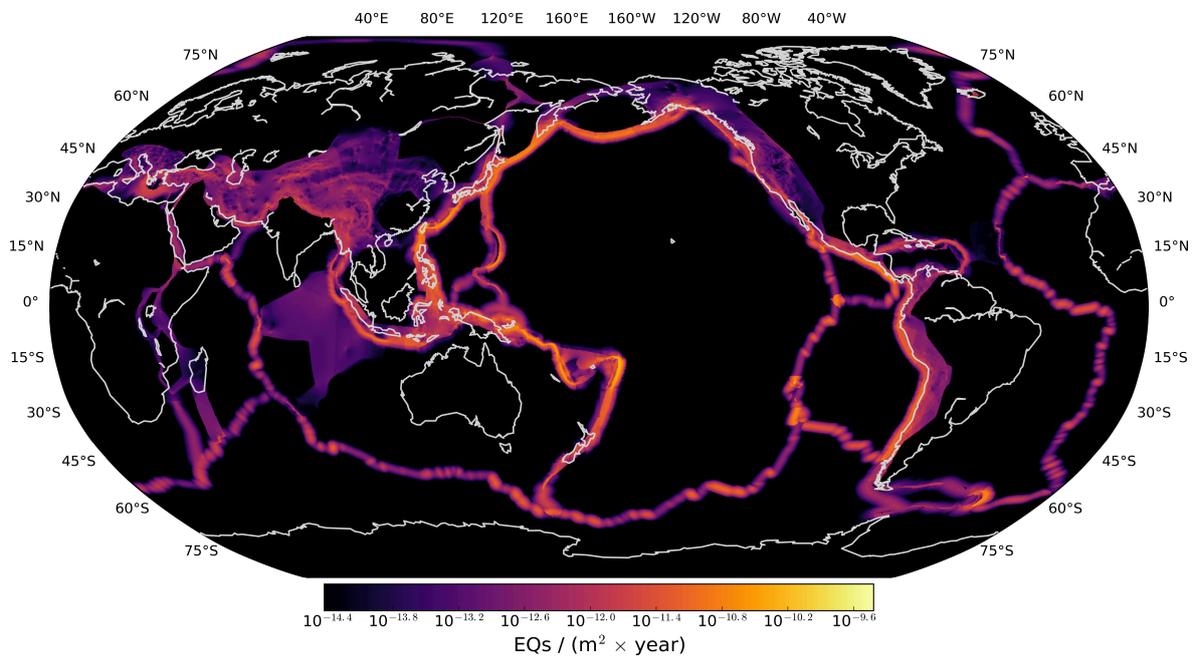
Abbildung 1.3: Forecast map of $m \geq 5.95$, $h \leq 70$ km global earthquake rates, derived from the Seismic Hazard Inferred from Tectonics based on the second iteration of the Global Strain Rate Map (SHIFT_GSRM2f; Bird and Kreemer, 2015) model.

1.2.3 Ensemble earthquake-rate modeling

The Global Earthquake Activity Rate (GEAR1; Bird et al., 2015) model is a global hybrid earthquake approach, resulting from a multiplicative log-linear blend of SHIFT_GSRM2f and KJSS. The optimized combination between these parent forecasts was determined by maximizing the I_1 (success) information score of Kagan (2009) from actual 2005–2012 seismicity. Thus, Bird et al. (2015) found that the most informative GEAR1 forecast stems from a multiplicative blend of earthquake forecast components, with exponent $d = 0.6$ on KJSS (Seismicity) and $c = 1 - d = 0.4$ on SHIFT_GSRM2f (Tectonics):

$$H = S^{0.6} \times T^{0.4} \tag{1.2}$$

Based on this combination, the GEAR1 modelers enhanced the preferred hybrid seismicity model by recomputing the Seismicity and Tectonics parent forecasts using the complete 1977–2013 Global CMT catalog. As a result, the update of such a preferred hybrid earthquake-rate model is what they refer to as the GEAR1 seismicity forecast (see Fig. 1.4). In this manner, GEAR1 captures the major advantages of its constituent forecast components (see Table 1.1), as SHIFT_GSRM2f computes high-resolution rates of seismicity along tectonic plate boundaries and KJSS constrains earthquake activity in intraplate regions, where strain rates are assumed to be zero (i.e. tectonic plates are modeled to be rigid) according to GSRM2.1, but earthquakes with magnitudes up to 7.4 have been recorded (Ekström et al., 2012).
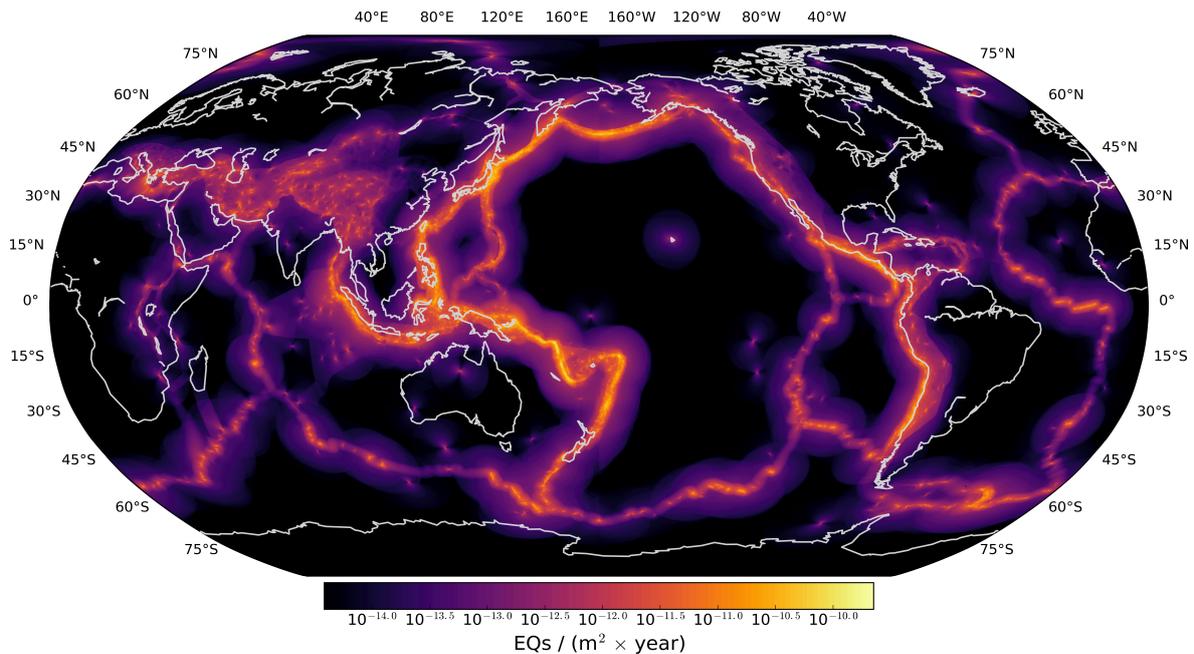
Abbildung 1.4: Forecast map of $m \geq 5.95$, $h \leq 70$ km global earthquake rates, derived from the Global Earthquake Activity Rate (GEAR1; Bird et al., 2015) model.

Tabelle 1.1: Evolution of global seismicity models described in the introduction of this doctoral thesis.

Model	Nature	Based on	Reference
KJSS	Seismicity	CMT catalog	Kagan and Jackson (2011)
SHIFT_GSRM	Tectonic	GSRM and CMT catalog	Bird et al. (2010)
SHIFT_GSRM2f	Tectonic	GSRM2.1 and CMT catalog	Bird and Kreemer (2015)
GEAR1	Hybrid	KJSS and SHIFT_GSRM2f	Bird et al. (2015)

1.3 Earthquake model evaluation

GEAR1, KJSS, SHIFT_GSRM and SHIFT_GSRM2f provide valuable scientific hypotheses about when and where earthquakes may occur and how big they might be. Nevertheless, given the strongly stochastic nature of earthquakes and the limited sampling of large events, the veracity of these hypotheses can only be either confirmed or rejected after prospective forecast experiments (Werner et al., 2011; Taroni et al., 2014). In this regard, a major step towards prospective forecast evaluation was taken by the Working Group on Regional Earthquake Likelihood Models (RELM; Field, 2007; Schorlemmer and Gerstenberger, 2007; Schorlemmer et al., 2007, 2010). This initiative consisted of inviting five-year seismicity forecasts for California to be prospectively tested within a unified, transparent and reproducible environment. Later, the

Collaboratory for the Study of Earthquake Predictability (CSEP; Zechar et al., 2010; Schorlemmer et al., 2018) expanded the RELM's scientific task by receiving regional and global seismicity forecasts for retrospective, pseudo-prospective and prospective evaluation.

1.3.1 Consistency tests

Over time, earthquake consistency tests have been implemented within CSEP to assess the performance of participating seismicity forecasts. These tests are based on the likelihood of observed seismicity patterns, given forecasted earthquake numbers. A forecast's log-likelihood score is a metric, based on the Poisson distribution, used to evaluate the consistency of forecasted seismicity patterns with observed earthquakes (Schorlemmer et al., 2010). Greater log-likelihood scores indicate greater consistency, corresponding to a higher probability of the forecast generating a seismicity distribution similar to observations. This in turn implies a greater ability of the seismicity model to forecast earthquakes.

Consistency with observed seismicity, expressed as the log-likelihood score, can be decomposed into three dimensions, i.e. number of earthquakes, magnitude and spatial distributions, which are directly derived from the likelihood, or L-test (Zechar and Rhoades, 2010). Thus, a forecast's log-likelihood score is most impacted by the number of earthquakes. Additionally, one can apply the conditional likelihood (CL) test, which provides information about a forecast's spatial and magnitude distribution while removing information regarding the total number of earthquakes (Werner et al., 2011).

1.3.2 Comparative tests

The relative performance of forecast pairs can be tested by measuring the rate-corrected information gain per earthquake of one forecast over another (Rhoades et al., 2011). For each global forecast pair, one applies the Student's paired T-test. This comparative test is based on the null hypothesis that two forecasts perform similarly, and the alternate hypothesis that one forecast significantly outperforms the other. As a result, one forecast is considered more informative than the other if the mean information gain significantly differs from the scaled difference in forecasted earthquake numbers between two forecasts. The Student's paired T-test requires the assumption that the information gain scores at observed earthquake locations are normally distributed, which is not always the case. For this reason, one can also apply the non-parametric W-test, evaluating the median information gain per earthquake rather than the mean. This test only requires that the information gain distribution is symmetric, and increases in power with increasing numbers of observed earthquakes (Rhoades et al., 2011). Based on these tests, CSEP has been able to identify the most informative seismicity models among more than 400 participant earthquake approaches (Michael and Werner, 2018; Schorlemmer et al., 2018). In California, for instances, the strain-rate driven forecast of Shen et al. (2007) and the Uniform California Earthquake Rupture Forecast v.2 (UCERF2; Field et al., 2009) have greater forecasting skills than most other RELM models (Strader et al., 2017). In Italy, the most consistent forecast until today is the one derived from the MPS04_AFTER (MPS-Working-Group, 2004) earthquake-rate model (Taroni et al., 2018). In New Zealand, the smoothed seismicity model based on the Proximity to Past Earthquakes method (PPE; Jackson and Kagan, 1999; Rhoades and Evison, 2006) is the best-performing forecast, according to 10 years of prospective testing (Rhoades et al., 2018).

1.4 Scientific questions and outline

The prospective assessment of seismicity models is important, because earthquake forecasts like UCERF2 and MPS04_AFTER provide authorities with actual guidance on hazard estimates to be considered for enhanced building design and updated insurance rating. Nonetheless, these particular test results are still preliminary, as the test areas are too small to observe a statistically representative number of large events in just a decade. In fact, only global $m \geq 5.8$ forecasts can offer prospective test results to be obtained in 1–8 years (Bird et al., 2015). Hence, the first scientific question of this doctoral thesis rises:

1 Which is the most consistent global seismicity forecast with independent earthquake observations?

To respond to this question, my research colleagues and I prospectively evaluated the performance of global earhquake-rate models over a 2-year observation period (the broadest independent dataset available by that time). In this study (see Chapter 2), we contrasted long-term estimates of seismicity provided by SHIFT_GSRM, SHIFT_GSRM2f, KJSS and GEAR1 with actual earthquake activity. Specifically, we evaluated the number, magnitude, and spatial distributions of forecasted events. In addition, we compared the forecasting skills of each earthquake-rate model to that of GEAR1, our reference seismicity model.

During the twentieth century, subduction megathrust seismicity accounted for approximately 90% of the global seismic moment (Pacheco and Sykes, 1992), and comprised 60% of the earthquake activity observed worldwide (Bird et al., 2010). Thus, several authors (e.g. Ruff and Kanamori, 1980; Peterson and Seno, 1984; Pacheco et al., 1993) have largely investigated the great diversity of subduction margins, in terms of mechanical behavior, age and geometry to better explain these observations. McCaffrey (2008) reported that subduction zones unevenly contribute to the total seismic moment globally released. Also, Ide (2013) determined a linear relationship between relative plate velocity and background (declustered) interface seismicity. Moreover, the three-dimensional Slab1 and Slab2 models of Hayes et al. (2012, 2018) show indisputable variations among subduction-area geometries. In this manner, seismicity data suggest that subduction zones are tectonically so distinctive that regional earthquake parameters are required to more accurately describe their specific seismicity patterns. As a result, the core scientific questions of this doctoral thesis rise:

2a Could the regionalization of geodesy-to-seismicity parameters significantly reduce raw-SHIFT_GSRM and SHIFT_GSRM2f earthquake underpredictions in subduction zones?

2b If true, which parameter(s) more extensively contribute(s) to the improvement of these global geodetic seismicity models? Is it indeed the subduction dip angle?

To give answers to these research questions, I developed and tested a complementary model to the methods developed by Bird et al. (2010) and Bird and Kreemer (2015) to forecast global interface seismicity. This model, referred to as Subduction Megathrust Earthquake Rate Forecast (SMERF; see Chapter 3), is a physics-derived, stationary and testable earthquake approach for 37 subduction zones, based on the conservation of moment principle and the use of regional geodesy-to-seismicity parameters. The physics-based method of SMERF consists of balancing long-term rates of geodetic and seismic moment, computed from interseismic-strain data and earthquake-catalog information. Among others, this approach allows to derive analytical solutions for seismic coupling and earthquake activity. Moreover, the iterative method of SMERF is suitable to determine quantitatively, which

regional parameters more effectively reduce raw-SHIFT earthquake-rate underpredictions. So, I additionally tested the earthquake-number consistency of SMERF with observations recorded during the 1977–2014 and 2015–2018 retrospective and pseudo-prospective evaluation periods.

The main difference between SMERF and its predecessors SHIFT_GSRM and SHIFT_GSRM2f models is the redistribution of seismicity rates within subduction margins through the application of regional, rather than global, seismicity parameters. Due to the prominent seismic activity observed within subduction margins, I hypothesize that SMERF could have a measurable impact on the construction of global earthquake-rate models. Thus, the final scientific question of this doctoral thesis gives rise:

3 What is the impact of the new interface-seismicity model on the development, update and improvement of global earthquake-rate models like those of Bird and Kreemer (2015) and Bird et al. (2015)?

   To answer this question, I produced and evaluated two global ensemble seismicity models based on SMERF2, a revised version of SMERF that depends on a reduced number of earthquake parameters (see Chapter 4). Specifically, I combined SMERF2 estimates in subduction zones with SHIFT_GSRM2f computations outside of these tectonic margins to create the global Tectonic Earthquake Activity Model (TEAM). In this manner, TEAM is an updated geodetic-based earthquake model, aimed at providing high-resolution rates of seismicity along plate boundaries, with special focus on subduction regions. Using the same multiplicative blend of parent seismicity forecasts as GEAR1, I moreover combined TEAM with the KJSS forecast to generate the World Hybrid Earthquake Estimates based on Likelihood scores (WHEEL) seismicity model. I designed WHEEL to serve as an alternative method to GEAR1 to forecast earthquakes globally. Therefore, I also compared its relative consistency and the relative consistency of all global seismicity forecasts to that of GEAR1.

Having said this, I finally invite the reader to dive into Chapters 2–4 for further information regarding the creation and testing of global seismicity models (see Table 1.2). Moreover, I encourage her/him to also follow Chapters 5–6, in which I describe the major findings, advantages, limitations and prospects of this doctoral dissertation.

Tabelle 1.2: Summary of long-term seismicity models, constructed and tested for this doctoral thesis.

Model	Nature	Coverage	Based on	Reference
SMERF	Tectonic	37 Subduction zones	GSRM2.1 and CMT catalog	Chapter 3
SMERF2	Tectonic	14 Subduction zones	GSRM2.1 and CMT catalog	Chapter 4
TEAM	Tectonic	Global	SMERF2 and SHIFT_GSRM2f	Chapter 4
WHEEL	Hybrid	Global	TEAM and KJSS	Chapter 4

1.5 Author's publications and contributions

The Chapters 2–4 of this doctoral thesis are peer-reviewed/submitted articles in scientific journals. These articles are listed below:

- Chapter 2:

  Strader, A., Werner, M., **Bayona, J.**, Maechling, P., Silva, F., Liukis, M. and Schorlemmer, D., 2018. Prospective evaluation of global earthquake forecast models: 2 yrs of observations provide preliminary support for merging smoothed seismicity with geodetic strain rates. Seismological Research Letters, 89(4), pp.1262-1271.
  https://doi.org/10.1785/0220180051

  I contributed with the description and plotting of global seismicity models. Moreover, I provided comments to improve the manuscript.

- Chapter 3:

  Bayona Viveros, J.A., Specht, S., Strader, A., Hainzl, S., Cotton, F. and Schorlemmer, D., 2019. A Regionalized Seismicity Model for Subduction Zones Based on Geodetic Strain Rates, Geomechanical Parameters, and Earthquake-Catalog Data. Bulletin of the Seismological Society of America, 109(5), pp.2036-2049.
  https://doi.org/10.1785/0120190034

  I created the theoretical method described in this article. Also, I developed Python codes to construct and test the seismicity model for subduction zones. Furthermore, I wrote the manuscript, including comments from my Ph.D advisors and co-authors.

- Chapter 4:

  Bayona, J.A., Savran, W., Strader, A., Hainzl, S., Cotton, F. and Schorlemmer, D., 2020. Two global ensemble seismicity models obtained from the combination of interseismic strain measurements and earthquake-catalogue information. Geophysical Journal International,224(3), pp.1945-1955.
  https://doi.org/10.1093/gji/ggaa554.

  I designed the theoretical framework presented in this study. In addition, I developed Python codes to produce global ensemble earthquake-rate models. Moreover, I wrote the manuscript, incorporating comments from my Ph.D advisors and co-authors.

The following work has been published in the course of my doctoral studies. However, I will not include it as a part of this thesis:

- **Bayona Viveros, J.A.**, Suárez Reynoso, G. and Ordaz Schroeder, M.G., 2017. A probabilistic seismic hazard assessment of the Trans-Mexican Volcanic Belt, Mexico based on historical and instrumentally recorded seismicity. Geofísica Internacional, 56(1), pp.87-101. http://dx.doi.org/10.19155/geofint.2017.056.1.7

  I developed the theoretical approach introduced in this research. Also, I wrote the manuscript, including comments from my Ph.D adviser and co-author.

Kapitel 2

Prospective Evaluation of Global Earthquake Forecast Models: Two Years of Observations Provide Preliminary Support for Merging Smoothed Seismicity with Geodetic Strain Rates

Abstract

The Global Earthquake Activity Rate (GEAR1) seismicity model uses an optimized combination of geodetic strain rates, hypotheses about converting strain rates to seismicity rates from plate tectonics, and earthquake-catalog data to estimate global $m_w \geq 5.767$ shallow (≤ 70 km) seismicity rates. It comprises two parent models: a strain rate-based model, and a smoothed-seismicity based model. The GEAR1 model was retrospectively evaluated and calibrated using earthquake data from 2005–2012, resulting in a preferred log-linear, multiplicative combination of the parent forecasts. Since October 1, 2015, the GEAR1 model has undergone prospective evaluation within the Collaboratory for the Study of Earthquake Predictability (CSEP) testing center, forecasting $m_w \geq 5.95$ seismicity. We present initial prospective forecast test results for the GEAR1 model, its tectonic and seismicity components, and for the first iteration of the strain rate-based model, during the period October 1, 2015 to September 7, 2017. During the evaluation period, observed earthquakes are consistent with the GEAR1 forecast and comparative test results likewise support that GEAR1 is more informative than either of its components alone. Based on a combination of retrospective and prospective testing, the tectonic forecasts do not effectively anticipate observed spatial earthquake distribution, largely due to over-localization of the model with respect to observed earthquake distributions.[1]

2.1 Introduction

Earthquake-source models that reliably forecast long-term seismicity rates are imperative to seismic risk mitigation. As they form the foundation of seismic hazard models (Anderson and Biasi, 2016; Cornell, 1968), the development and evaluation of seismicity models is essential to improving ground shaking estimates. Such estimates are directly used worldwide as an input for building codes, including nuclear facilities, and for urban planning. Accurate and reliable forecasting of long-term seismicity rates allows societies to identify regions at risk of catastrophic earthquake damage and effectively invest in building and infrastructure safety. Global seismicity models are advantageous in this respect because sufficient earthquakes occur to rank global forecasts with a 5.8–7.0 magnitude threshold after only 1-8 years of testing (Bird et al., 2015).

Until recently, global seismicity models used as input for Probabilistic Seismic Hazard Analysis (PSHA) were primarily based on a combination of earthquake-catalog and fault data (Giardini, 2014; Giardini et al., 1999). These seismicity models extrapolated previous seismicity to the future, and assumed that all faults that could potentially yield large, catastrophic earthquakes were known. In recent years, the quality and quantity of available geodetic data have increased considerably and are beginning to be incorporated into earthquake source models such as the Uniform California Earthquake Rupture forecast (UCERF; Field et al., 2009) and (UCERF3; Field et al., 2015). These data can provide information about tectonic deformation that cannot be supplied by instrumental or historical earthquake catalogs, particularly in regions where only few seismicity data are available (Kreemer et al., 2014; Bird and Kreemer, 2015; Bird et al., 2015).

2.1.1 Seismicity models

To use geodetic data in forecasting time-invariant global seismicity rates, Bird and Kreemer (2015) developed the SHIFT_GSRM2f model. SHIFT_GSRM2f is based on the second iteration of the Global Strain Rate Map (GSRM2.1), a global model of continuous strain rates constrained by 22,415 interseismic GPS velocities (Kreemer et al., 2014). Using the Seismic Hazard Inferred from Tectonics (SHIFT) hypotheses introduced by Bird and Liu (2007), Bird and Kreemer converted geodetic strain rates to long-term seismicity rates by multiplying each strain rate tensor by the elastic shear modulus, spatial bin area, geometric factor, and coupled seismogenic thickness. The authors produced six tectonics-based seismicity models of varying complexity (SHIFT_GSRM2a-f). SHIFT_GSRM2f, the most complex of the resulting seismicity models, recognizes five tectonic zones, applies asymmetric Gaussian smoothing to subduction zone strain rates and symmetric smoothing to all other offshore plate boundaries, and accounts for velocity-dependent coupled seismogenic thickness.

The Global Earthquake Activity Rate Model (GEAR1) is a hybrid seismicity model that uses a combination of geodetic strain rates and earthquake-catalog data to forecast long-term seismicity (Bird et al., 2015). It is a multiplicative log-linear blend of the SHIFT_GSRM2f and Kagan-Jackson Smoothed Seismicity (KJSS) models (Kagan and Jackson, 1994, 2000, 2011). The KJSS model was developed by smoothing shallow (≤ 70 km) $m_w \geq 5.767$ earthquakes during the period 1977–2004 from the Global Centroid Moment Tensor (CMT) earthquake catalog (Ekström et al., 2012). Although strain rates capture seismicity in regions where earthquake-catalog data are scarce, earthquake catalogs provide information about intraplate seismicity, where strain rates are assumed to be zero according to the GSRM2.1.

Both components of the GEAR1 model were developed from undeclustered seismicity, due to the lack of a consensus on an optimal declustering scheme, and potential misclassifications of mainshocks and aftershocks resulting from earthquake catalog boundaries in space, time and magnitude (Bird et al., 2015). For the tectonic models, seismicity rates are quite sensitive to the corner magnitude (increasing the corner magnitude by one increases seismicity rates by a factor of three); corner magnitudes were assigned for both GEAR1 component models based on tectonic region (Bird and Kagan, 2004; Kagan et al., 2010). For consistency, all tests are conducted using undeclustered earthquake catalogs.

Using global CMT seismicity from 2005–2012, Bird et al. (2015) retrospectively evaluated various combinations of the SHIFT_GSRM2f and KJSS models. They tested three main combinations of tectonics and seismicity: a weighted linear combination, a multiplicative log-linear combination, and selection in each spatial bin of the model with the greater forecasted seismicity rate. They optimized the GEAR1 model by maximizing the I1 (ßuccess") score, or mean information gain per earthquake of GEAR1 over a uniform Poisson model (Kagan, 2009). Testing multiple weightings of tectonics and smoothed seismicity, they determined that the optimal GEAR1 model was a multiplicative log-linear combination of SHIFT_GSRM2f and KJSS, with a weight of 0.6 (d=0.6 in the following equation) assigned to smoothed seismicity:

$$H_{ij} = N\Big\{\max[(S_{ij}^d \cdot T_{ij}^{1-d}), f]\Big\} \tag{2.1}$$

where H_{ij} is the total forecasted earthquake rate in the spatial bin with latitude i and longitude j. S_{ij} is the rate forecasted by KJSS and T_{ij} is the rate forecasted by SHIFT_GSRM2f. f is a baseline seismicity rate defined as $\min[\min(S_{ij}), \min(T_{ij})]$, in order to account for earthquakes in unexpected locations. $N\{\}$ is an operator that normalizes the total earthquake rate to equal the total rate observed during 1977–2003 according to the CMT catalog, so that the models project the mean global seismicity rate into the future. This normalization is applied after constraining the corner magnitudes for each tectonic region, such that an increase in the corner magnitude does not result in a decreased total activity rate. Seismicity data used to normalize the total seismicity rates for the global models are completely independent from data used to calibrate or retrospectively/prospectively test the models. Figure 2.1 displays maps of earthquake rates calculated from all three global models, as well as SHIFT_GSRM (i.e. Bird et al., 2010), the predecessor of SHIFT_GSRM2f.

2.1.2 Seismicity forecast evaluation

Although successful retrospective test results are necessary to detect initial seismicity model limitations, such tests are prone to bias because data included in calibrating the model are also used to test it. Only truly prospective evaluation is considered rigorous, because it has the ability to estimate the forecasting power of a seismicity model (Schorlemmer et al., 2007; Zechar et al., 2010; Strader et al., 2017). Since October 1, 2015, seismicity forecasts produced by GEAR1, SHIFT_GSRM2f, and KJSS have undergone prospective testing within the Collaboratory for the Study of Earthquake Predictability (CSEP). We also evaluate the SHIFT_GSRM model, based on the first version of the GSRM (Kreemer et al., 2003). GSRM2.1 improves upon the original GSRM through its incorporation of additional geodetic information, systematic data processing, additional modeled plates and plate boundaries, and a finer spatial grid (Bird and Kreemer, 2015; Kreemer et al., 2014).

2.1.3 Recent results

In this study, we present prospective testing results from October 1, 2015 to September 7, 2017. We evaluate the consistency of each forecast's total number of earthquakes, spatial earthquake distribution and magnitude distribution against earthquakes observed during the evaluation period. To determine each GEAR1 component's contribution to constraining seismicity rates and prospectively assess the stability of GEAR1's optimization, we compare GEAR1's forecast performance with those of SHIFT_GSRM2f and KJSS. We also compare the performance of SHIFT_GSRM and SHIFT_GSRM2f to determine if the additional geodetic data used in developing SHIFT_GSRM2f significantly improves forecast performance during the evaluation period. Because the approximately two-year prospective testing period is considerably shorter than the eight-year GEAR1 model calibration period and global earthquake rates fluctuate considerably over two-year intervals, we cannot identify major model strengths and weaknesses using only current prospective test results. To better understand the temporal stability of two-year global model performance, we include a retrospective analysis of the total forecasted number and spatial distribution of earthquakes for five two-year intervals during 2004-2013. We also compare I1 scores for the entire prospective testing period and one-year intervals with the retrospective one-year tests conducted by (Bird et al., 2015), to determine if the average information gains per earthquake over a uniform Poisson global seismicity model are consistent.

2.1.4 Data

209 $m_w \geq 5.95$ earthquakes (0–70 km depth) occurred during the evaluation period, considerably fewer than in the past decade. From 2004–2013 (during the time the two-year retrospective tests were conducted), an average of 266 earthquakes occurred during consecutive two-year time intervals, with a standard deviation of 26.8 earthquakes. By contrast, the normalization rate for the duration of the prospective testing period determined by Bird et al. (2015) using 1977–2003 seismicity data results in 225 forecasted earthquakes from each global model. We used the undeclustered CMT catalog, consistent with earthquakes used to generate and calibrate GEAR1 (Bird et al., 2015). Notable earthquakes during the prospective testing period include the April 16, 2016 m_w 7.8 Ecuador earthquake, November 14, 2016 m_w 7.8 Kaikoura earthquake in New Zealand, and December 17, 2016 m_w 7.9 Papua New Guinea earthquake.
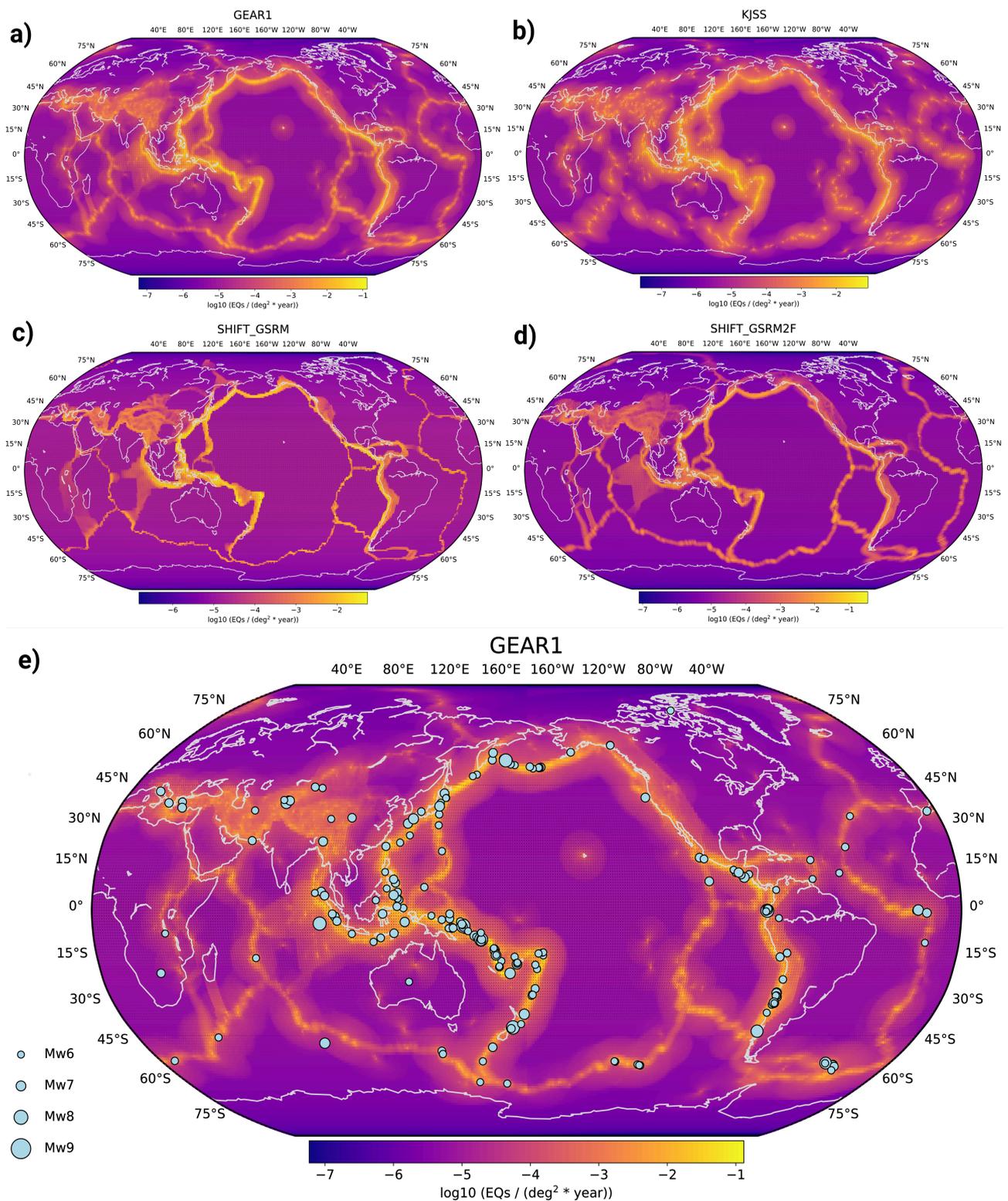
Abbildung 2.1: Forecast maps showing $m_w \geq 5.95$ earthquake epicentroid rates (log10(eqs/year) in each $0.1° \times 0.1°$ spatial bin): a) the preferred GEAR1 model (Bird et al., 2015), b) KJSS (Kagan and Jackson, 2011), c) SHIFT_GSRM (Bird et al., 2010), d) SHIFT_GSRM2f (Bird and Kreemer, 2015). We determine epicentroid rate densities from a loglinear, multiplicative combination of the SHIFT_GSRM2f and KJSS parent forecasts, with exponent d = 0.6. Bright yellow and orange areas indicate regions with elevated seismicity rates. Map e) displays epicenters of observed earthquakes during the evaluation period (overlying the GEAR1 forecast map).

2.2 Methods

2.2.1 CSEP testing center framework

Each CSEP testing center (Schorlemmer and Gerstenberger, 2007) allows for earthquake-rate forecasts generated from seismicity models to be evaluated in a specified region during a specified evaluation period. In a forecasting experiment, the testing region is divided into spatiomagnitude bins based on longitude, latitude and magnitude increments. Seismicity modelers submit forecasts to CSEP in the form of expected numbers of earthquakes (assuming a Poisson seismicity distribution in each spatial bin) during the evaluation period in each spatiomagnitude bin (Werner et al., 2010; Eberhard et al., 2012; Tsuruoka et al., 2012). Forecast experiments are conducted in a transparent, reproducible environment, with earthquake data used to test forecasts provided by an authoritative and independent source (Schorlemmer and Gerstenberger, 2007; Zechar et al., 2010). A brief overview of the consistency and comparative tests (with corresponding literature) used to evaluate earthquake forecasts is provided below.

Comparative Tests

One evaluates the relative performance between forecasts by measuring the rate-corrected information gain per earthquake of one forecast over another (Rhoades et al., 2011). We apply the Student's paired T-test, defining the null hypothesis that two forecasts do not perform significantly differently, and the alternate hypothesis that one forecast significantly outperforms the other. If the mean information gain significantly differs from the scaled difference in forecasted earthquake numbers between two forecasts, one forecast is significantly more informative than the other.

In the case that the information gain scores at observed earthquake locations are not normally distributed, the non-parametric W-test evaluates the median information gain per earthquake rather than the mean. This test requires that the information gain distribution is symmetric, and increases in power with increasing numbers of observed earthquakes (Rhoades et al., 2011).

2.2.2 Consistency tests

Using a suite of likelihood consistency tests, CSEP evaluates the consistency of forecasted and observed seismicity during the experiment's evaluation period. Tests are based on the likelihood of observed seismicity patterns, given forecasted seismicity rates. A forecast's log-likelihood score is a metric, based on the Poisson distribution, used to evaluate the consistency of forecasted seismicity patterns with observed earthquakes (Schorlemmer et al., 2007). Greater log-likelihood scores indicate greater consistency, corresponding to a higher probability of the forecast generating a seismicity distribution similar to observations (indicating greater ability to forecast earthquakes).

Consistency with observed seismicity, or the log-likelihood score, can be decomposed into three dimensions: number of earthquakes, magnitude and spatial distributions. The tests of consistency for these dimensions (N-test, M-test and S-test, respectively) are derived directly from the likelihood, or L-test (Zechar and Rhoades, 2010). Because the number of earthquakes has the greatest impact on a forecast's log-likelihood score, the conditional likelihood (CL) test provides information about a forecast's spatial and magnitude distribution with information regarding the total number of earthquakes removed (Werner et al., 2011).

2.3 Results

2.3.1 Model comparison and ranking

GEAR1 significantly outperforms both of its individual components during the evaluation period, according to the T- and W-test results (Figure 2.2). This is based on the information gain at a selected significance level of 0.05. The forecast's superior performance to its individual components during the prospective evaluation supports the near-even weighting of KJSS and SHIFT_GSRM2f, as well as the selected log-linear combination of the models. Given that the multiplicative model captures two independent factors necessary for earthquake triggering, i.e. continuous lithospheric deformation as well as sudden static or dynamic stress changes (Bird et al., 2015), the stability of the optimized GEAR1 model over two years is promising. However, because Bird et al. (2015) selected the optimized GEAR1 model based on I1 score variations during an eight-year retrospective period, a similar prospective time period will likely be necessary to conclusively validate the model selection.

SHIFT_GSRM displays a small mean information gain over SHIFT_GSRM2f; nonetheless, this difference is not statistically significant during the prospective evaluation period.
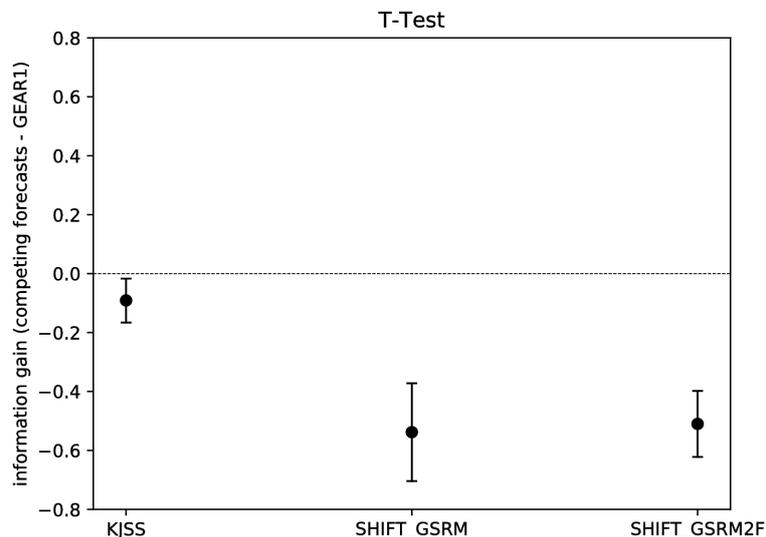
Abbildung 2.2: T-test results comparing GEAR1's performance with those of KJSS, SHIFT_GSRM and SHIFT_GSRM2f during the prospective evaluation period. GEAR1 significantly outperforms both tectonics forecasts as well as KJSS, supporting a near-even combination of geodetic and earthquake-catalog data to best constrain short- to mid-term seismicity rates. Circles correspond to observed mean information gain per earthquake, while vertical lines display the range of 0.05 significance level mean information gain values; the ranges for all tectonic and seismicity component forecasts are all below zero. The mean information gain per earthquake is displayed on the y-axis; the dashed horizontal line corresponds to the scaled mean information gain (that is, no difference in performance between forecast pairs). Although not displayed, the W-test results corroborate all T-test results.

2.3.2 Consistency tests

The observed total earthquake number is consistent with all forecasts during the prospective evaluation period, according to the N-test results (Figure 2.3 and Table 2.1). Because all models had similar spatial seismicity rate integrals, it is unsurprising that the models passed the N-test simultaneously. Bird et al. (2015) normalized all models to forecast total earthquake rates equal to the 1977-2003 global CMT rate at the $m_{\mathrm{w}} \geq 5.767$ threshold, so that annual fluctuations in seismicity would not result in one forecast being favored over another during short

(1-2 year) evaluation periods. For the forecasts in this work, the normalization rate determined from 1977–2003 for $m_w \geq 5.95$ is 225 earthquakes over the length of the prospective testing period. Retrospective test results reflect large observed fluctuations in seismicity rates over two-year time intervals (Table 2.2); the GEAR1 and SHIFT_GSRM2f forecasts are consistent with observed seismicity during two out of the five retrospective testing periods. Because the observation catalogs were undeclustered, aftershocks from large earthquakes (for example, the M9.0 Tohoku event) increased seismicity rates considerably from 2006-2011, compared to the average rate calculated from the 1977–2003 global CMT catalog. Both GEAR1 and SHIFT_GSRM2f underpredicted seismicity during these testing periods.
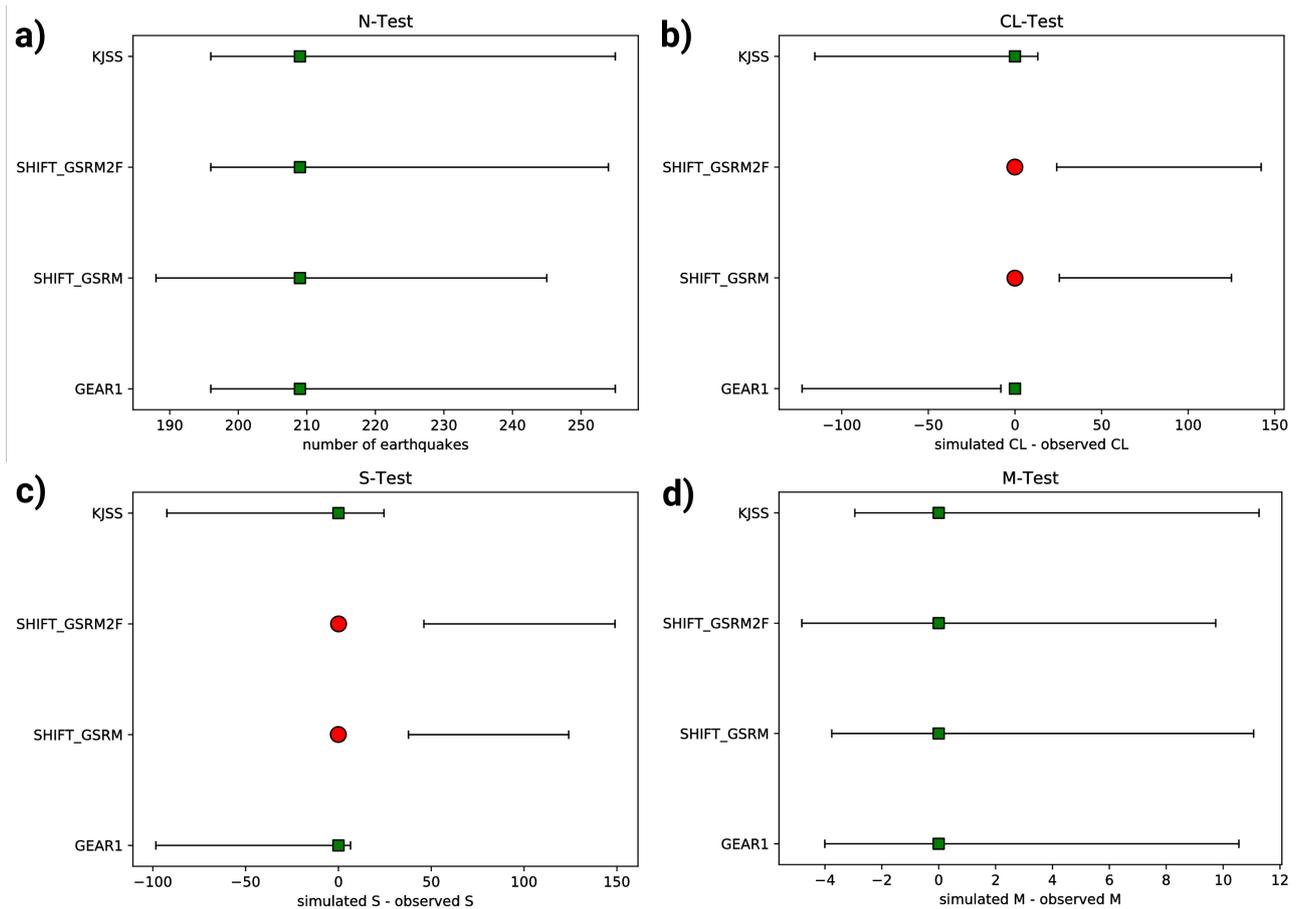
Abbildung 2.3: Consistency test results for the GEAR1, SHIFT_GSRM, SHIFT_GSRM2f, and KJSS forecasts during the period October 1, 2015–September 7, 2017: a) N-test, b) CL-test, c) S-test, d) M-test. Green squares indicate that observations are consistent with the forecasts, while red circles indicate significant inconsistencies between forecasted and observed seismicity. All forecasts pass the N- and M-tests; however, SHIFT_GSRM and SHIFT_GSRM2f forecast spatial earthquake distributions inconsistent with observed seismicity, thus failing the CL- and S-tests. For the N-test, the ordinate displays the number of earthquakes; for all other tests, the log-likelihood score is displayed. The squares and circles indicate number of observed earthquakes (in the case of the N-test) or the log-likelihood score calculated using observed earthquakes (all other tests). The horizontal lines indicate the range of observed earthquake numbers or log-likelihood scores within the 0.05 significance level.

The observed earthquake magnitude distribution is consistent with those of all forecasts during the prospective testing period, as indicated by the M-test results. Prior to combining the KJSS and SHIFT_GSRM2f models, Bird et al. (2015) scaled the KJSS model to have an equal seismicity rate to the SHIFT_GSRM2f model. Both models' magnitude-frequency distributions comprised unions of various tapered Gutenberg-Richter distributions based on the SHIFT philosophy (Bird and Kagan, 2004). Therefore, the four models have similar log-likelihood (M) scores.

Tabelle 2.1: Prospective N-, CL-, S- and M-test statistics for all global seismicity models. The test statistics δ_1 and δ_2 give the probabilities of observing at least and at most the number of observed earthquakes, respectively. If $\delta_1 < 0.025$, the forecast underpredicts; if $\delta_2 < 0.025$, the forecast overpredicts. ξ, ζ and κ give the percentage of earthquake catalogs simulated from the seismicity model with a lower log-likelihood score than the score for the observation catalog. At the 0.05 significance level, if these statistics fall below 0.05, the forecasted earthquake activity (magnitude and spatial patterns, spatial pattern, and magnitude pattern, respectively) is considered inconsistent with that of observed seismicity. The last column, IG, shows the information gain of each model over GEAR1 during the prospective testing period. Inconsistencies between forecasted and observed seismicity are highlighted in bold.

Model	δ_1	δ_2	ξ	ζ	κ	IG (over GEAR1)
GEAR1	0.87	0.15	0.98	0.96	0.12	0.00
KJSS	0.87	0.15	0.93	0.88	0.10	-0.09
SHIFT_GSRM	0.69	0.33	**0.00**	**0.00**	0.12	-0.53
SHIFT_GSRM2F	0.86	0.16	**0.00**	**0.00**	0.17	-0.51

Both GEAR1 and KJSS forecast spatial earthquake distributions consistent with observations during the prospective testing period, as indicated by the S-test; however, both tectonic models do not. One reason for the difference in performance is that forecasted seismicity in SHIFT_GSRM2f is highly concentrated along plate boundaries compared to KJSS. This effect is indicated by large differences in log-likelihood scores for earthquakes near plate boundaries but located just beyond regions of elevated forecasted seismicity. Concentration plots for both forecasts (Figure 2.4) display cumulative observed and forecasted seismicity distributions. The forecasted seismicity distribution for SHIFT_GSRM2f is shifted slightly to the right of the observed seismicity distribution, suggesting that in SHIFT_GSRM2f seismicity is too localizedälong plate boundaries compared to KJSS. The use of undeclustered seismicity also impacted the difference in forecasted and observed spatial likelihood scores, in that simulated earthquake catalogs rarely contained more than one earthquake per spatial bin, whereas four spatial bins contained two observed earthquakes each. Bins located just outside of modeled active seismic margins, containing multiple earthquakes due to clustering, lowered the observed spatial log-likelihood score substantially relative to simulated spatial log-likelihood scores.

Tabelle 2.2: Retrospective N- and S-test results for GEAR1 and SHIFT_GSRM2f over two-year intervals during 2004-2013. The test statistics δ_1 and δ_2 give the probabilities of observing at least and at most the number of observed earthquakes, N, respectively. If $\delta_1 < 0.025$, the forecast underpredicts; if $\delta_2 < 0.025$, the forecast overpredicts. The test statistic ζ give the percentage of earthquake catalogs simulated from the seismicity model with a lower spatial log-likelihood score than the score for the observation catalog. If $\zeta < 0.05$, the forecasted spatial seismicity pattern is considered inconsistent with that of observed seismicity. The N-test results are quite unstable over time, largely due to inclusion of aftershocks from large earthquakes in the observation catalog. Both models' S-test results are stable over the retrospective and prospective time intervals at the 0.05 significance level. Inconsistencies between forecasted and observed seismicity are denoted by light gray shading.

Test Period	N	δ_1 (GEAR1)	δ_2 (GEAR1)	ζ (GEAR1)	δ_1 (SHIFT_GSRM2f)	δ_2 (SHIFT_GSRM2f)	ζ (SHIFT_GSRM2f)
2004-2005	239	0.34	0.68	1.00	0.33	0.69	**0.00**
2006-2007	284	**0.00**	1.00	1.00	**0.00**	1.00	**0.02**
2008-2009	281	**0.00**	1.00	1.00	**0.00**	1.00	**0.00**
2010-2011	291	**0.00**	1.00	1.00	**0.00**	1.00	0.20
2012-2013	235	0.44	0.58	1.00	0.43	0.60	**0.00**

Retrospective analysis shows that the spatial forecast performances for SHIFT_GSRM2f and GEAR1 are mostly stable over two-year testing intervals between 2004–2013 for the 0.05 significance level (Table 2.2). For all retrospective time intervals, GEAR1's spatial forecast is

consistent with observed seismicity; for SHIFT_GSRM2f, the spatial forecast is inconsistent with observed earthquake activity for all time intervals except for 2010-2011. This suggests that SHIFT_GSRM2f's prospective spatial forecast performance may not entirely be due to temporal fluctuations in seismicity patterns, and warrants further investigation into model improvement.

2.3.3 Information scores (I1)

Because the earthquake catalogs used to develop and calibrate GEAR1 were undeclustered (Bird et al., 2015), the earthquakes used to evaluate the global models cannot be assumed independent. Therefore, the CSEP test results should be corroborated by test results that do not require such an assumption. Bird et al. (2015) retrospectively evaluated the relative performances of GEAR1 and KJSS over one-year intervals at the $m_\mathrm{w} \geq 5.767$ and $m_\mathrm{w} \geq 7.1$ thresholds using the I1 information score (Kagan, 2009), because it compares forecasted and observed seismicity patterns independent of earthquake counts.

I1 scores for $m_\mathrm{w} \geq 5.95$ seismicity during the entire prospective testing period, as well as two one-year intervals within the testing period (October 1, 2015 to September 30, 2016 and September 8, 2016 to September 7, 2017), are displayed in Table 2.3. For all prospective time intervals, GEAR1 has the maximum I1 score among the global models. Differences between the I1 scores for GEAR1 and KJSS are consistent with those observed by Bird et al. (2015) for one-year retrospective tests. These results support that, over time windows similar to the eight-year GEAR1 calibration period, short-term (two-year) spatial forecast performance is stable for GEAR1 and SHIFT_GSRM2f.

Tabelle 2.3: I1 scores for all global models during the entire prospective evaluation period and one-year intervals October 1, 2015 to September 30, 2016 (Period I) and September 8, 2016 to September 7, 2017 (Period II). GEAR1 and KJSS have greater information gains per earthquake over uniform seismicity than the tectonic forecasts, in agreement with comparative test results (Figure 2.2). Differences between I1 scores for GEAR1 and KJSS are consistent with one-year retrospective tests conducted by Bird et al. (2015).

Model	I1 (Period I + Period II)	I1 (Period I)	I1 (Period II)
GEAR1	4.56	4.95	4.18
KJSS	4.43	4.78	4.06
SHIFT_GSRM	3.84	4.22	3.47
SHIFT_GSRM2f	3.84	4.26	3.44

2.4 Discussion and Conclusions

We prospectively tested forecasts generated from the GEAR1 model, its components KJSS and SHIFT_GSRM2f, and the original SHIFT_GSRM model from October 1, 2015 to September 7, 2017 within the CSEP testing center. The total earthquake number, spatial and magnitude distributions forecasted by GEAR1 were all consistent with observed seismicity during the prospective testing period. Additionally, GEAR1 outperformed its component models KJSS and SHIFT_GSRM2f. These results with additional retrospective testing support the inclusion of both geodetic strain rates and earthquake-catalog data to better constrain spatial seismicity patterns. However, instability in N-test results during retrospective two-year intervals indicates that a longer testing period (at least equal to the GEAR1 calibration period) is necessary to

understand how effectively the optimal GEAR1 model constrains mid- to long-term seismicity rates.

GEAR1's superior spatial performance to the SHIFT_GSRM and SHIFT_GSRM2f models supports that geodetic data supplements, but does not replace earthquake-catalog data, and also supports using ensemble modeling to improve upon individual forecast performance (Marzocchi et al., 2012; Taroni et al., 2014). This is corroborated by the rejection of KJSS in favor of GEAR1 during the prospective testing period, as well as GEAR1's consistently higher I1 scores during both one-year prospective periods and the retrospective tests by Bird et al. (2015). However, these retrospective tests were conducted during the GEAR1 calibration interval, leaving only two years of prospective test results to validate the selection of the optimal GEAR1 model without bias. All global models continue to be evaluated by CSEP, and future test results for prospective evaluation periods comparable to that of the calibration period will more clearly elucidate how geodetic data contribute to constraining earthquake rates over longer time intervals.

Prospective and retrospective I1 scores suggest that geodetic data can improve forecasted short-term global spatial seismicity patterns; the differences in information gain between GEAR1 and its components are stable over one-year time intervals. However, retrospective S-test results indicate that observed limitations of the SHIFT_GSRM2f model during the prospective testing period are not completely due to temporal variations in seismicity patterns. Despite the incorporation of numerous geodetic data and additional tectonic plates and plate boundaries, SHIFT_GSRM2f also did not significantly outperform SHIFT_GSRM during the prospective testing period. Although SHIFT_GSRM2f incorporates more geodetic velocities and provides a higher-resolution spatial strain-rate distribution than SHIFT_GSRM, the spatial forecast is too localized compared to observed seismicity near plate boundaries (see the concentration plot from Figure 2.4). Future analysis of the GEAR1 and tectonic models in low-seismicity regions is necessary to define a "buffer zoneßurrounding highly seismic margins that accounts for the majority of earthquakes occurring near these margins.

Inconsistencies in spatial seismicity patterns between the SHIFT_GSRM or SHIFT_GSRM2f model and observed seismicity may also be caused by inaccurate or low-resolution physical input parameters used to convert strain rates to seismic moment. Bird et al. (2015) indicate that the SHIFT_GSRM and SHIFT_GSRM2f models tend to underpredict seismicity within subduction zones prior to applying empirical calibration factors to account for the effect of the dip angle on the geometric factor. We are currently investigating how varying input parameters in the SHIFT_GSRM model (for example: subduction dip angle, coupled seismogenic thickness, and corner magnitude) affects ratios of forecasted to observed earthquakes in subduction zones (see Chapter 3). Further testing of individual regions and the subsequent development and prospective evaluation of updated SHIFT_GSRM and GEAR models will indicate the extent to which forecast performance is sensitive to parameters used to convert strain rates to seismicity.

Two-year prospective evaluation of the GEAR1 model and its individual components supports (with the corroboration of retrospective testing) the incorporation of geodetic strain rates in global earthquake source models to better constrain spatial seismicity patterns. However, temporal fluctuations in seismicity rates over two-year intervals indicate that we do not yet have a sufficient amount of earthquake-catalog data to prospectively evaluate the global models' potential for forecasting long-term seismicity rates that are used as input in PSHA. These findings are similar to those from regional forecasting experiments such as the RELM experiment (Strader et al., 2017). Seismicity models such as UCERF2 that incorporate geodetic strain rates not only forecast seismicity patterns consistent with observations, but their forecast performance over consecutive five-year time intervals is more consistent than for models solely

Abbildung 2.4: Concentration plots displaying the cumulative forecasted versus observed seismicity distribution for SHIFT_GSRM2f (red curves) and KJSS (blue curves). The observed seismicity distribution for SHIFT_GSRM2f is shifted slightly to the left of the forecasted distribution, indicating that the forecasted seismicity is too localized. For KJSS, the two distributions are in closer agreement, though the forecasted seismicity distribution is slightly over-smoothed.

based on smoothed seismicity (Helmstetter et al., 2007). In the future, the predictive skill of SHIFT_GSRM2f could be improved further with more region-specific calibration of parameters used to convert strain rates to seismicity rates, as well as further investigation of strain rate variations in regions just outside of defined active seismic margin boundaries.

Data and Resources

The observation earthquake catalog was obtained from the global CMT earthquake catalog by the CSEP testing center. All forecast evaluation tests except for the concentration plots were conducted using miniCSEP, available as open-source software from CSEP.

Acknowledgments

Kapitel 3

A Regionalized Seismicity Model for Subduction Zones based on Geodetic Strain Rates, Geomechanical Parameters and Earthquake-catalog Data

Abstract

The Seismic Hazard Inferred from Tectonics (SHIFT_GSRM) earthquake forecast was designed to provide high-resolution estimates of global shallow seismicity to be used in seismic hazard assessment. This model combines geodetic strain rates with global earthquake parameters to characterize long-term rates of seismic moment and earthquake activity. SHIFT_GSRM properly computes seismicity rates in active continental regions. However, it underestimates earthquake rates in subduction zones by an average factor of approximately three. We present a complementary method to SHIFT_GSRM to more accurately forecast earthquake rates in 37 subduction segments, based on the conservation of moment principle and the use of regional interface seismicity parameters, such as subduction dip angles, corner magnitudes and coupled seismogenic thicknesses. In seven progressive steps, we find that SHIFT_GSRM earthquake-rate underpredictions are mainly due to the utilization of a global probability function of seismic moment release that poorly captures the great variability among subduction-zone megathrust interfaces. Retrospective test results show that the forecast is consistent with the observations during the 1 January 1977–31 December 2014 period. Moreover, successful pseudo-prospective evaluations for the 1 January 2015–31 December 2018 period demonstrate the power of the regionalized earthquake model to properly estimate subduction-zone seismicity.[1]

3.1 Introduction

Megathrust earthquakes accounted for approximately 90% of the seismic moment released worldwide over the last century (Pacheco and Sykes, 1992). Thus, the construction of well-constrained seismicity models for subduction zones is necessary to more effectively estimate earthquake and tsunami hazards on a global scale. Until recently, global seismicity models were mainly based on the characterization of seismic sources and extrapolation to the future of historical and instrumentally-recorded seismicity (e.g. Giardini et al., 1999). Nonetheless, this method depends on the availability and quality of fault- and earthquake-catalog data. Furthermore, existing seismicity catalogs are too short to sample the largest possible earthquakes along faults and plate boundaries (Bird and Kreemer, 2015).

Interseismic strain measurements, however, permit to quantify earthquake potential in areas where seismic sources may still remain unmapped (Bird and Liu, 2007), and currently provide high-resolution coverage for almost all deforming tectonic plate limits (Kreemer et al., 2014). Combined with earthquake information, geodetic strain rates are convenient to generate better-constrained seismicity models to be used for seismic hazard assessment. Whereas strain rates indicate fault stress accumulation, seismicity rates reveal stress release, so that an optimal combination between these data sets is desired to more accurately determine moment balance for long-term earthquake forecasting (Hashimoto et al., 2009; Bird et al., 2015).

The Global Strain Rate Map (GSRM; Kreemer et al., 2003) is a global model of horizontal velocities, rotation rates and horizontal strain rates on the Earth's surface primarily obtained from Global Positioning System (GPS) data, and originally designed to support the creation of global geodetic-based earthquake models. Bird et al. (2010) converted GSRM to an indefinite-term earthquake rate model by applying the hypotheses, assumptions and equations of the Seismic Hazard Inferred From Tectonics (SHIFT; Bird and Liu, 2007) model. The modelers refer to this product as the SHIFT_GSRM seismicity forecast. SHIFT_GSRM assumes that the long-term seismic moment rate of any permanently deforming volume of lithosphere $\dot{M}_{\text{GEO}}$ linearly scales with the product of principal geodetic strain rates and global geomechanical earthquake parameters (similar to Kostrov (1974), Ward (1994) and Savage and Simpson (1997). Accordingly, rates of predicted seismicity are computed by equating the SHIFT_GSRM geodetic moment rate $\dot{M}_{\text{GEO}}$ with the model seismic moment rate $\dot{M}^{\text{CMT}}$ (integral of the best-fitting tapered Gutenberg-Richter distribution of observed earthquakes over all moments), and using actual seismicity rates.

During the 1977–2009.03 period, SHIFT_GSRM properly estimates earthquake rates in seismically-active continental regions (the average ratio between actual and forecasted seismicity is ≈ 1). Nonetheless, the model underpredicts seismicity rates within subduction zones by a mean factor of approximately 3 for the same evaluation period (Bird et al., 2010). The model's authors suggest that such discrepancies may stem from the use of inappropriate geometric factors in the moment-rate equation, a time-dependence of global seismic activity and a velocity-dependence of interface subduction seismicity. Based on these assumptions, Bird et al. (2010) justify the use of empirical correction factors to improve SHIFT_GSRM earthquake rates in subduction zones. The corrected SHIFT_GSRM forecast was submitted to the Collaboratory for the Study of Earthquake Predictability (CSEP; Schorlemmer and Gerstenberger, 2007) testing center for independent retrospective, pseudo-prospective and prospective evaluation.

Years later, Kreemer et al. (2014) presented a major update of GSRM: the GSRM2.1 model. GSRM2.1 incorporates almost five times more interseismic GPS velocities, a finer spatial grid and a more detailed tectonic plate boundaries model. Based on GSRM2.1, Bird and Kreemer (2015) generated the SHIFT_GSRM2f global seismicity forecast, which additionally employs a

new tectonic regionalization scheme (i.e. Kagan et al., 2010), a spatial smoothing of model strain rates around offshore plate limits and velocity-dependent coupled thicknesses for continental convergent boundaries and subduction zones. Despite these improvements, SHIFT_GSRM2f also underpredicts subduction-zone seismicity rates when retrospectively evaluated during the 2005–2012 period. The SHIFT_GSRM2f authors sustain that earthquake underestimations are mainly due to the use of inadequate geometric factors, since the moment-rate equation implicity assumes that fault planes slip at a mean dip angle of 45°. Although such an approximation seems to be reasonable on a global scale, it poorly describes mechanical processes occuring within subduction interfaces at much shallower down-dip angles (23° on average; Heuret et al., 2011).

Fault-plane orientation plays a significant role in the accomodation of plate motion in interface subduction environments (Bilek and Lay, 2018). Particularly, low-angle subduction dipping facilitates the transmission of compressive stress to the overriding plate through increased contact area between the plates (Lallemand et al., 2005). Moreover, the strain regime positively correlates with shallow dip angles (Ruff and Kanamori, 1980), so that the use of specific fault-dipping angles is required to better constrain geodetic-based estimates of seismic moment and rates of shallow subduction-zone seismicity.

We thus propose a complementary method to SHIFT_GSRM to compute long-term rates of shallow seismicity within 37 subduction segments, based on the conservation of moment principle and the use of regional interface earthquake parameters, such as subduction dip angles, corner magnitudes and coupled seismogenic thicknesses. We generate seismicity models in seven progressive steps (seismicity models A-G) to quantify the contribution of each new regional subduction parameter on the computations. Finally, we employ global subduction-zone seismicity recorded between 1977 and 2014, and from 2015 to 2018 to retrospectively and pseudoprospectively test the consistency of each earthquake-rate model.

3.1.1 Data

We use a global tectonic regionalization scheme (KRS2010; Kagan et al., 2010) to estimate earthquake rates exclusively in subduction zones. This model divides the surface of the Earth into five zones of different tectonic styles, based on an algorithm that detects variations of seismicity parameters, such as seismic coupling and corner magnitude.

We also utilize the Global Strain Rate Model 2.1 presented in a regular longitude/latitude grid, with a spatial resolution of 0.1°. GSRM2.1 is a global velocity gradient tensor field and continuous strain rates model for plate boundary zones, based on 22,415 interseismic GPS velocities (Kreemer et al., 2014). According to this model, about 14% of the Earth's surface is allowed to deform and the rest is assumed to be rigid crust, i.e. geodetic strain rates are considered to be zero in intraplate regions.

We additionally employ the Global CMT catalog to calibrate and test our regionalized tectonic forecasts. We consider all earthquakes with hypocentral depths ≤ 70 km and moment magnitude threshold $m_T \geq 5.66$ to assure magnitude completeness (Bird and Kagan, 2004). Our periods of observation comprise from January 1, 1977 to December 31, 2014 for calibration and retrospective evaluation, and from January 1, 2015 to December 31, 2018 for pseudoprospective testing.

3.2 Methods

3.2.1 Model A: The raw SHIFT_GSRM earthquake model

SHIFT_GSRM is a global seismicity forecast based on the use of geodetic strain rates. This model computes long-term seismic moment rates $\dot{M}_{\mathrm{GEO}}$ per deforming cell as:

$$\dot{M}_{\mathrm{GEO}} = A_{\mathrm{S}} \ c \ z \ \mu \left\{ \begin{array}{ll} 2 \ \dot{\varepsilon}_3; & \text{if } \dot{\varepsilon}_2 < 0 \text{ or} \\ -2 \ \dot{\varepsilon}_1; & \text{if } \dot{\varepsilon}_2 \geq 0, \end{array} \right. \tag{3.1}$$

in which A_{S} is the surface area of each GSRM grid cell, c is the mean seismic coupling coefficient, z is the seismogenic thickness, μ is the average elastic shear modulus, ± 2 is a geometric factor accounting for the effect of fault dipping, and $\dot{\varepsilon}_1$, $\dot{\varepsilon}_2$, and $\dot{\varepsilon}_3$ are the minimum, intermediate and maximum principal strain rates, respectively.

Principal strain rates are obtained from GSRM by assuming incompressibility (the sum of the trace of each strain-rate tensor is equal to zero). So, the computation of $\dot{M}_{\mathrm{GEO}}$ is maximum, because the maximum absolute value of the principal strain rates is always chosen, according to Eq. 3.1. Also, seismicity parameters are assigned to each cell, based on the tectonic regionalization scheme of Kreemer et al. (2002). This regionalization model separates the Earth's crust into five types of plate boundaries by inferring significant variations of tectonic moment rates from geodetic velocities and Quaternary fault slip rates.

Activity rates $\dot{N}$ are then obtained by normalizing the estimates of $\dot{M}_{\mathrm{GEO}}$ by the model seismic moment rate $\dot{M}^{\mathrm{CMT}}$ (integral of the best-fitting tapered Gutenberg-Richter distribution of the appropriate Global CMT subcatalog for each type of plate boundary), and multiplying by the number of earthquakes $\dot{N}^{\mathrm{CMT}}$ in each subcatalog at or above a magnitude threshold m_{T}:

$$\dot{N}(m \geq m_{\mathrm{T}}) = (\dot{M}_{\mathrm{GEO}}/\dot{M}^{\mathrm{CMT}}) \ \dot{N}^{\mathrm{CMT}} \tag{3.2}$$

In this model, long-term rates of seismic moment and subduction-zone seismicity are computed by using mean global earthquake parameters (see Table 3.1), obtained from ≤ 70 km interface earthquakes recorded during the 01.01.1977–30.09.2002 period (Bird and Kagan, 2004).

Tabelle 3.1: Global earthquake parameters used to compute SHIFT_GSRM estimates for subduction zones. SHIFT_GSRM computes long-term rates of seismic moment and shallow seismicity, based on the results of Bird and Kagan (2004).

Seismicity Parameter	Global Mean Value
Down-Dip Angle θ	$45°$
Seismic Coupling Coefficient c	0.69
Seismogenic Thickness z	26 km
Elastic Shear Modulus μ	49 GPa
$m_{\mathrm{T}} \geq 5.66$ Seismicity Rate $\dot{N}^{\mathrm{CMT}}$	79.7 EQa^{-1}
Long-Term Seismic Moment Rate $\dot{M}^{\mathrm{CMT}}$	9×10^{21} Nma^{-1}
Beta Value β	0.64
Corner Magnitude m_{cg}	9.58

3.2.2 Model B: A new Global Strain Rate Map

We start with complementing SHIFT_GSRM by substituting the Global Strain Rate Map, used in the seismicity model A, for the updated version GSRM2.1. GSRM2.1 incorporates more geodetic velocities and thus it offers a higher resolution spatial strain-rate distribution to more accurately estimate rates of shallow seismicity.

3.2.3 Model C: A new global regionalization scheme

Our earthquake models aim to contribute to the improvement of global seismicity models, such as the Global Earthquake Activity Rate (GEAR1; Bird et al., 2015) model. GEAR1 is a hybrid seismicity forecast resulting from the optimized log-linear combination of geodetic strain rates and smoothed earthquake data. Specifically, GEAR1 uses both the smoothed seismicity model of Kagan and Jackson (2011) and the SHIFT_GSRM2f tectonic forecast of Bird and Kreemer (2015). SHIFT_GSRM2f, the tectonic parent component of GEAR1, is based on the regionalization framework of Kagan et al. (2010). Accordingly, we replace the global regionalization scheme of Kreemer et al. (2002), employed in the seismicity model B, by the tectonic regionalization model of Kagan et al. (2010) to fairly compare our global subduction-zone earthquake model with these seismicity forecasts in future work.

3.2.4 Model D: Regional subduction dip angles

We then generalize the moment-rate equation for any fault-dipping angle by applying the principle of conservation of moment. We express long-term geodetic moment rates per deforming cell in terms of the down-dip angle θ as:

$$\dot{M}_{\text{GEO}} = A_{\text{S}} \ c \ z \ \mu \ f(\theta) \begin{cases} \dot{\varepsilon}_3; & \text{if } \dot{\varepsilon}_2 < 0 \text{ or} \\ -\dot{\varepsilon}_1; & \text{if } \dot{\varepsilon}_2 \geq 0. \end{cases} \tag{3.3}$$

The long-term rate of seismic moment release is traditionally modeled as the product of the seismic coupling coefficient c, the mean elastic shear modulus μ, the average slip rate $\dot{D}$, the fault length L and the fault width W (Haskell, 1963; Burridge and Knopoff, 1964; Brune, 1968):

$$\dot{M}_{\text{TECT}} = c \ \mu \ \dot{D} \ L \ W \tag{3.4}$$

By assuming that all geodetic deformation is elastic, on the long-term and after several seismic cycles, the summation of Eq. 3.3 over all cell areas ($A_{\text{X}} = \sum A_{\text{S}}$) should be equal to Eq. 3.4, so that:

$$f(\theta) = \frac{\dot{D} \ A_{\text{F}}}{\dot{\varepsilon} \ z \ A_{\text{X}}} \tag{3.5}$$

where $A_{\text{F}} = L \ W$ is the source area, $A_{\text{X}} = L \ W_{\text{X}}$ is the projection to the surface of the source area and $\dot{\varepsilon}$ the spatial-averaged maximum principal strain rate.

Given that $A_{\text{F}} = A_{\text{X}}/\cos(\theta)$ and $W \cos(\theta) = W_{\text{X}}$ (see Fig. 3.1), we write $f(\theta)$ as:

$$f(\theta) = \frac{\dot{D}}{\dot{\varepsilon} \ z \ \cos(\theta)} = \frac{\dot{D} \ W}{\dot{\varepsilon} \ z \ W_{\text{X}}} \tag{3.6}$$

Additionally, $z = W \sin(\theta)$ and $V_{\text{X}} = \dot{D} \cos(\theta)$. Hence,

$$f(\theta) = \frac{\dot{D}}{\dot{\varepsilon} \ W_{\text{X}} \ \sin(\theta)} = \frac{V_{\text{X}}}{\dot{\varepsilon} \ \cos(\theta) \ \sin(\theta) \ W_{\text{X}}} \tag{3.7}$$

In agreement with Bird et al. (2010), we relate the strain rate to the plate velocity by assuming that the relative horizontal plate velocity V_{X} is accommodated by differential deformation at the surface within a distance X. The main deformation is expected to occur on top of the coupled zone within distance W_{X}, but minor deformation also occurs beyond this zone leading to $X \geq W_{\text{X}}$. Consequently, the strain rate can be estimated by $\dot{\varepsilon} = V_{\text{X}}/X \leq V_{\text{X}}/W_{\text{X}}$. Thus, Eq.
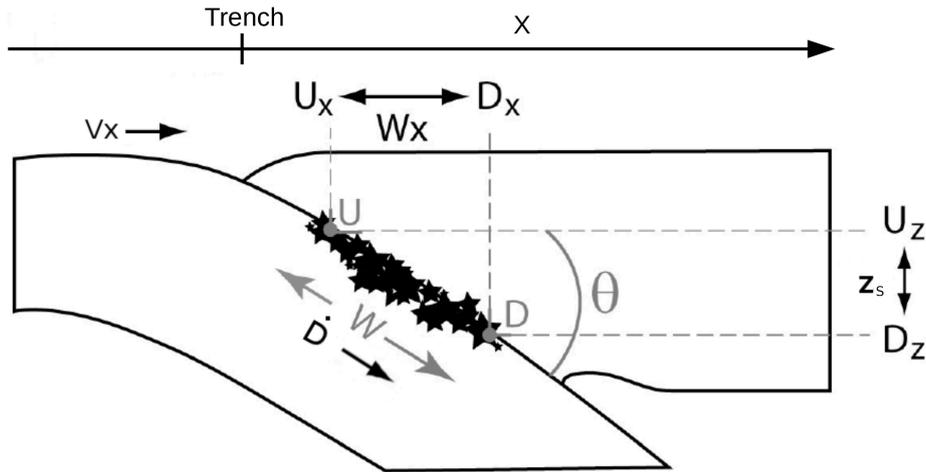
Abbildung 3.1: Geo-mechanical parameters defining the subduction plate interface seismogenic zone geometry. Modified from Heuret et al. (2011).

3.7 becomes:

$$f(\theta) \geq \frac{1}{\cos(\theta) \, \sin(\theta)} \qquad (3.8)$$

By substituting the lower bound $f(\theta) = 1 \, / \cos(\theta) \, \sin(\theta)$ in Eq. 3.3, we present a fault dip-dependent equation to compute geodetic-based rates of seismic moment release per subduction zone:

$$\dot{M}_{\text{GEO}} = \frac{A_{\text{X}} \; c \; z \; \mu}{\cos(\theta) \, \sin(\theta)} \begin{cases} \dot{\varepsilon}_3; & \text{if } \dot{\varepsilon}_2 < 0 \text{ or} \\ -\dot{\varepsilon}_1; & \text{if } \dot{\varepsilon}_2 \geq 0. \end{cases} \qquad (3.9)$$

Note that we verified the derived dip-dependence by numerical simulations, in which we modeled the coupling by backslip on a subduction interface with different dip angles in an elastic half-space.

For our computations, we use regional subduction dip angles obtained from the trench segmentation model proposed by Heuret et al. (2011) (see Fig. 3.2). We exclude subduction zones for which not all needed regional earthquake parameters have been reported yet (e.g. Cascades, Sandwich or Aegean).

Heuret et al. (2011) determined the geometry of the seismogenic zone, particularly the dip angle θ and the seismogenic depth z_{s}, by assuming that it coincides with the distribution of shallow ($\leq$ 70 km) and intermediate-magnitude ($5.5 \leq m_{\text{w}} < 7.0$) subduction thrust earthquakes (i.e. Pacheco et al., 1993). These authors used both the 1976–2007 CMT Harvard (i.e. Dziewonski et al., 1981; Ekström et al., 2012) and the EHB (i.e. Engdahl et al., 1998) earthquake catalogs to identify $m_{\text{w}} \geq 5.5$ subduction earthquakes and improve the location of each identified interface event, respectively.

Based on the estimations of Heuret et al. (2011), we use mean regional down-dip angles θ (see Table 3.2) to quantify the kinematic contribution of the rupture-plane orientation to global estimates of seismic moment rate $\dot{M}_{\text{GEO}}$. Whereas the apparent seismic coupling coefficient c, the seismogenic thickness z and the elastic shear modulus μ remain global (Table 3.1), we vary down-dip angles within each subduction zone.
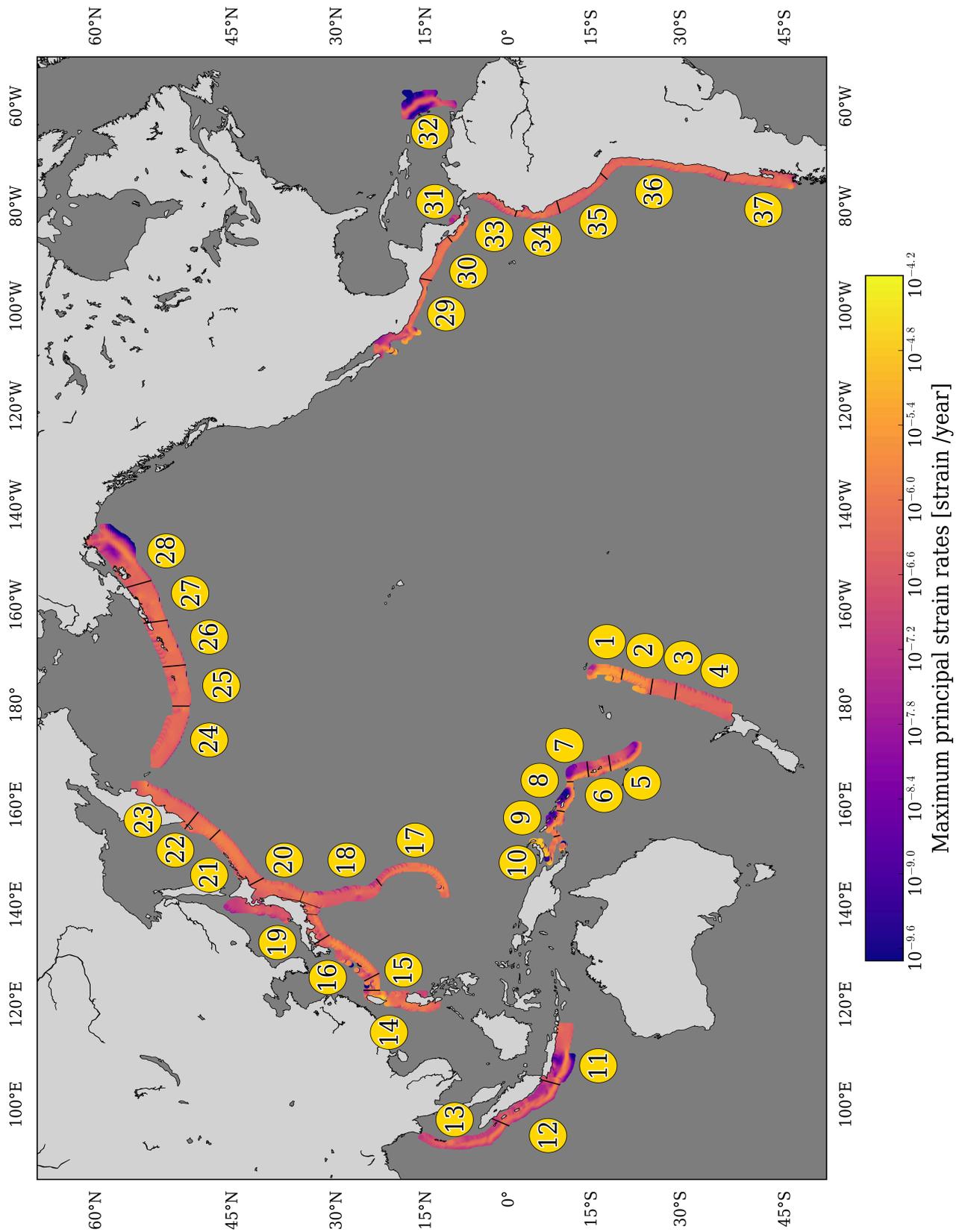
Abbildung 3.2: Subduction zones considered in this study to compute rates of $m_T \geq 5.66$ shallow seismicity, based on the trench segmentation proposed by Heuret et al. (2011). In our model, we rename Manila-Taiwant-the subduction zone referred as Manila"by the model's authors, Tohokuthe subduction zone previously called "Japan", Colombia-Ecuadorthe trench formerly labeled as Colombia", and we exchange Cocosfor Costa Rica"(see Table 3.2). We also present maximum principal strain rates used to compute long-term estimates of seismic moment release per subduction interface, according to Eq. 3.9

3.2.5 Model E: Regional estimates of elastic shear modulus

Next, we propose a method to compute regional estimates of elastic shear modulus μ_s, based on the ak135-f velocity reference Earth model (i.e. Kennett et al., 1995). Such a method consists of first calculating the elastic shear modulus μ_i at any depth z_i by multiplying the square of the corresponding shear-wave velocity v_{si} by the density ρ_i, at such depth as:

$$\mu_i = v_{si}^2 \; \rho_i \tag{3.10}$$

We then estimate mean weighted values of elastic shear modulus per subduction interface as:

$$\mu_s = \frac{\sum_{i=1}^{n} \mu_i \Delta z_i}{\sum_{i=1}^{n} \Delta z_i}, \tag{3.11}$$

in which n is the number of ak135 layers between the upper U_z and lower D_z boundaries of each seismogenic zone (see Table 3.2). Finally, we substitute the mean global elastic shear modulus $\mu = 49$ GPa (Bird and Kagan, 2004), used in the seismicity model D, for these mean regional estimates of elastic rigidity (see Fig. 3.5 and Table 3.5) to analyze the effect of this parameter on the estimation of subduction-zone earthquake rates.

3.2.6 Model F: Regional long-term rates of seismic moment release

Starting from the seismicity model E, we replace the global model seismic moment rate $\dot{M}^{\mathrm{CMT}}$ with regional estimates of seismic moment release $\dot{M}_{\mathrm{SEIS}}$ for each subduction-zone interface, based on the results of Kagan and Jackson (2016). Kagan and Jackson (2016) provide regional β values and corner magnitudes m_{cg} (see Table 3.2) for 18 subduction segments. Moreover, they introduce a theoretical moment-rate equation described by the Gamma distribution with parameters β and M_{cg} to model frequency-moment distributions of instrumentally-recorded subduction-zone seismicity observed from 1977 to 2014:

$$\dot{M}_{\mathrm{SEIS}} = \frac{\dot{N}(m \geq m_{\mathrm{T}}) \, M_{\mathrm{T}}^{\beta} \; \beta}{1 - \beta} \; M_{cg}^{1-\beta} \, \Gamma(2 - \beta) \, \xi_g. \tag{3.12}$$

In Eq. 3.12, M_{T} is the threshold seismic moment corresponding to the lower magnitude threshold $m_{\mathrm{T}} = 5.66$ (i.e. Hanks and Kanamori, 1979), β is the asymptotic spectral slope of the earthquake distribution, M_{cg} is the corner moment, Γ is the gamma distribution and ξ_g is a correction coefficient ≈ 1, if $M_{cg} \gg M_{\mathrm{T}}$.

For this earthquake model, we group together some of the Heuret et al. (2011) subduction segments (e.g. Southwestern Aleutians, Central Aleutians, Eastern Aleutians, Western Alaska and Eastern Alaska) to relate them to the Kagan and Jackson (2016) subduction interfaces (e.g. Aleutian arc). Additionally, we assume that Eqs. 3.9 and 3.12 are equal on the long-term, based on the conservation of moment principle. So, we introduce a new equation to estimate predicted rates of shallow seismicity per subduction zone:

$$\dot{N}(m \geq m_{\mathrm{T}}) = \frac{(1 - \beta) \, A_{\mathrm{X}} \, c \, z \, \mu_s \, \dot{\varepsilon}}{M_{\mathrm{T}}^{\beta} \, \beta \, M_{cg}^{1-\beta} \, \Gamma(2 - \beta) \, \xi_g \, \cos(\theta) \, \sin(\theta)} \tag{3.13}$$

Eq. 3.13 is an innovative approach to forecast subduction-zone earthquake activity $\dot{N}$, because long-term estimates of shallow interface seismicity are obtained not only from earthquake data (β, M_{T}, M_{cg}), but also from geomechanical (A_{X}, $c\,z$, μ_s, θ) and geodetic information ($\dot{\varepsilon}$).

Tabelle 3.2: Mean regional interface-earthquake parameters used in this study. We estimate regional seismic moment rates $\dot{M}_{\text{GEO}}$, theoretical rates of seismic moment release $\dot{M}_{\text{SEIS}}$ and $m_{\text{T}} \geq 5.66$ shallow subduction-zone seismicity by employing interface-earthquake parameters obtained from (Heuret et al., 2011) and Kagan and Jackson (2016).

#	Subduction Zone	$\theta[°]$	z_s[km]	U_z[km]	D_z[km]	β	m_{cg}
1	Northern Tonga	21	34	8	42	0.801 ± 0.039	9.128 ± 0.277
2	Southern Tonga	31	41	4	45	0.801 ± 0.039	9.128 ± 0.277
3	Northern Kermadec	23	43	10	53	0.801 ± 0.039	9.128 ± 0.277
4	Southern Kermadec	29	58	6	64	0.801 ± 0.039	9.128 ± 0.277
5	Southern New-Hebrides	31	43	10	53	0.595 ± 0.032	8.939 ± 0.277
6	D'Entrecasteux	22	28	14	42	0.595 ± 0.032	8.939 ± 0.277
7	Northern New-Hebrides	33	54	10	64	0.595 ± 0.032	8.939 ± 0.277
8	Solomon	24	35	11	46	0.593 ± 0.031	8.947 ± 0.277
9	Bougainville	33	53	10	63	0.593 ± 0.031	8.947 ± 0.277
10	New Britain	22	46	17	63	0.593 ± 0.031	8.947 ± 0.277
11	Java	13	42	15	57	0.661 ± 0.044	9.582 ± 0.280
12	Sumatra	11	33	20	53	0.731 ± 0.063	9.345 ± 0.286
13	Andaman	9	39	11	50	0.731 ± 0.063	9.345 ± 0.286
14	Taiwan-Manila	24	40	12	52	0.645 ± 0.069	9.192 ± 0.292
15	Southern Ryukyu	32	38	12	50	0.667 ± 0.096	9.814 ± 0.307
16	Northern Ryukyu	16	33	20	53	0.667 ± 0.096	9.814 ± 0.307
17	Marianas	28	43	8	51	0.854 ± 0.095	9.888 ± 0.297
18	Izu-Bonin	20	35	10	45	0.854 ± 0.095	9.888 ± 0.297
19	Nankai	10	23	12	35	0.667 ± 0.096	9.814 ± 0.307
20	Tohoku	18	50	10	60	0.646 ± 0.031	9.296 ± 0.277
21	Southern Kuril	22	38	16	54	0.646 ± 0.031	9.296 ± 0.277
22	Northern Kuril	26	41	10	51	0.646 ± 0.031	9.296 ± 0.277
23	Kamchatka	27	50	11	61	0.646 ± 0.031	9.296 ± 0.277
24	Southwestern Aleutians	31	37	11	48	0.662 ± 0.043	9.367 ± 0.280
25	Central Aleutians	35	42	14	56	0.662 ± 0.043	9.367 ± 0.280
26	Eastern Aleutians	33	40	10	50	0.662 ± 0.043	9.367 ± 0.280
27	Western Alaska	24	37	18	55	0.662 ± 0.043	9.367 ± 0.280
28	Eastern Alaska	15	47	7	54	0.662 ± 0.043	9.367 ± 0.280
29	Mexico	24	30	11	41	0.591 ± 0.051	9.132 ± 0.284
30	Cocos	14	34	14	48	0.638 ± 0.054	9.219 ± 0.285
31	Costa-Rica	28	48	15	63	0.591 ± 0.051	9.132 ± 0.284
32	Antilles	12	34	14	48	0.646 ± 0.093	9.369 ± 0.307
33	Colombia-Ecuador	20	39	11	50	0.638 ± 0.054	9.219 ± 0.285
34	Northern Peru	17	34	10	44	0.571 ± 0.038	9.681 ± 0.279
35	Southern Peru	24	33	10	43	0.571 ± 0.038	9.681 ± 0.279
36	Northern Chile	22	39	12	51	0.571 ± 0.038	9.681 ± 0.279
37	Southern Chile	14	45	5	50	0.571 ± 0.038	9.681 ± 0.279

3.2.7 Model G: Regional hybrid coupled thicknesses

Coupled thicknesses *cz* employed for SHIFT_GSRM computations are based on seismicity data (Bird and Kagan, 2004). Nonetheless, the calibration of coupling coefficients based on instrumentally-recorded seismicity can be either under- or overestimated, since the rate of seismic moment is dominated by the occurence of the largest events, with recurrence times usually much larger than observation periods. Accordingly, we calibrate hybrid (geodetic- and seismicity-based) coupled thicknesses c_{H} to better capture the long-term effect of earthquake forecasting within subduction margins. For this aim, we first use regional seismogenic depths z_s, defined from the upper boundary U_z to the lower limit D_z of the seismogenic zone (see

3.1), obtained from Heuret et al. (2011). Whereas Bird and Kagan (2004) use a mean global seismogenic thickness for subduction zones $z \approx 26$ km, we employ seismogenic depths with a considerably larger mean value of 40 km (see Table 3.2), in agreement with the 49 km $\pm$ 11 km reported by Pacheco et al. (1993).

We then use actual $m_T \geq 5.66$ earthquake rates $\dot{N}_s$ recorded from 1977 to 2014 in each subduction zone to solve for hybrid coupling coefficients c_H as:

$$c_H = \frac{\dot{N}_s \, M_T^{\beta} \, \beta \, M_{cg}^{1-\beta} \, \Gamma(2 - \beta) \, \xi_g \, \cos(\theta) \, \sin(\theta)}{(1 - \beta) \, A_X \, z_s \, \mu_s \, \dot{\varepsilon}} \tag{3.14}$$

We consider that corner magnitudes M_{cg} and their corresponding standard deviations (i.e. Kagan and Jackson, 2016) predominantly control the variability of each c_H. Therefore, we report mean hybrid coupling coefficients and 90% confidence ranges (see Fig. 3.5 and Table 3.5). Also, estimates of hybrid coupling per fault area c_H may be larger than the unity, because earthquakes employed for calibration could actually have ruptured the same area more than once. For these scenarios, we finally set mean hybrid coupling coefficients to 1.0.

3.2.8 Model evaluation: N-test

The N-test determines if the forecasted number of earthquakes in a test region is consistent with the observed number of earthquakes. This test assumes that earthquakes within each 0.1° × 0.1° grid cell, subdivided by 0.1 magnitude bins, are independent and thus seismicity rates are considered to be Poisson distributed (Schorlemmer et al., 2007; Zechar and Rhoades, 2010). Each study region (in this case each subduction zone) contains a set of n spatial cells. Each cell i contains a forecasted number of earthquakes λ_i. Accordingly, the total earthquake forecast in a subduction zone is the sum of the forecasted number of events in all cells:

$$N_{fore} = \sum_{i=1}^{n} \lambda_i \tag{3.15}$$

Similarly, each cell also contains a number of earthquakes ω_i. The total number of earthquakes in a subduction zone is subsequently the sum of the observed number of earthquakes in all cells:

$$N_s = \sum_{i=1}^{n} \omega_i \tag{3.16}$$

For each seismicity model, we compute the test metric δ_2 to describe the probability of observing at most ω earthquakes, according to the forecasted rate λ. δ_2 is based on the right-continuous Poisson cumulative distribution with expectation λ and evaluated at k as:

$$\delta_2 = F(N_s | N_{fore}) = e^{-\lambda} \sum_{k=0}^{\omega} \frac{\lambda^k}{k!} \tag{3.17}$$

Moreover, we calculate the probability of observing at least ω earthquakes as:

$$\delta_1 = 1 - F((N_s - 1) | N_{fore}) \tag{3.18}$$

The N-test is two-tailed because a forecast can be inconsistent with observed seismicity by either over- or underpredicting the total number of events. CSEP forecasting experiments typically define the significance level as 0.05, so that the criteria to evaluate earthquake-model consistencies are:

- If $\delta_1 < 0.025$: The earthquake model underpredicts observed seismicity.

- If $\delta_2 < 0.025$: The earthquake model overestimates actual seismicity.

- If $\delta_1 > 0.025$ and $\delta_2 > 0.025$: The earthquake forecast is consistent with recorded seismicity.

The calibration of regional earthquake parameters is carried out by using interface seismicity observed until 31 December, 2014. Thus, we finally employ shallow subduction-zone earthquakes recorded during the 1 January, 1977–31 December, 2014 and 1 January, 2015–31 December, 2018 periods to retrospectively and pseudoprospectively evaluate the consistency of each earthquake forecast for subduction zones.

3.3 Results and Discussion

Retrospective and pseudoprospective test results show that the seismicity model G better describes subduction-zone seismicity, in comparison with the reference earthquake model A (see Figs. 3.3 and 3.4, and Tables 3.3 and 3.4). For the retrospective test period, the consistency between the forecast and observations is unsurprising, because we calibrated hybrid coupling coefficients c_H from actual interface seismicity rates $\dot{N}_s$ (Eq. 3.14). Nonetheless, successful pseudoprospective test results reveal the forecasting capability of the regionalized seismicity model. Earthquake-model improvements primarily stem from the use of regional estimates of seismic moment $\dot{M}_{SEIS}$ computed by Kagan and Jackson (2016). According to retrospective and pseudo-prospective evaluations, the seismicity model F yields ratios of observed and predicted seismicity with mean factors of 0.60 and 0.57, respectively. These improvements consequently indicate that the employment of a global model seismic moment rate $\dot{M}^{CMT}$ for all subduction zones (see Table 3.1) is the major reason for raw SHIFT_GSRM earthquake underestimations.

Bird and Kagan (2004) computed such a global probability function of seismic moment release for subduction zones $\dot{M}^{CMT}$ by fitting a magnitude-frequency distribution of global interface events and integrating over all moments. This global approximation, however, does not properly scale with geodetic moment rates $\dot{M}_{GEO}$ in Eq. 3.1 and thus it regionally underestimates subduction-zone seismicity. These results reveal the great variability among subduction zones reported by other authors (e.g. Ruff and Kanamori, 1980; Pacheco et al., 1993; Scholz and Campos, 2012), and exhibit the apparent non-proportionality observed between rates of seismic moment and seismicity discussed in other studies (e.g. Gutenberg and Richter, 1944; Heuret et al., 2011; Ide, 2013).
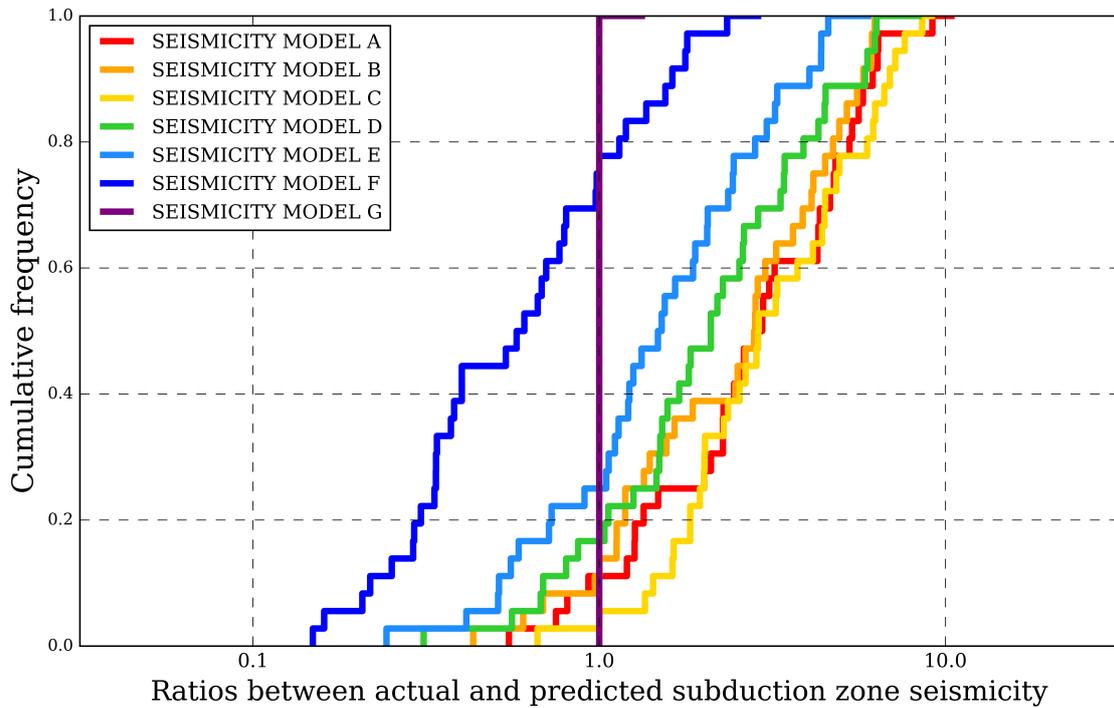
Abbildung 3.3: Cumulative distributions of the ratios between observed and forecasted seismicity, according to seven earthquake models presented in this study. We use actual subduction-zone seismicity observed from 1 January, 1977 to 31 December, 2014 to retrospectively test the consistency of each earthquake model within 37 subduction zone interfaces.
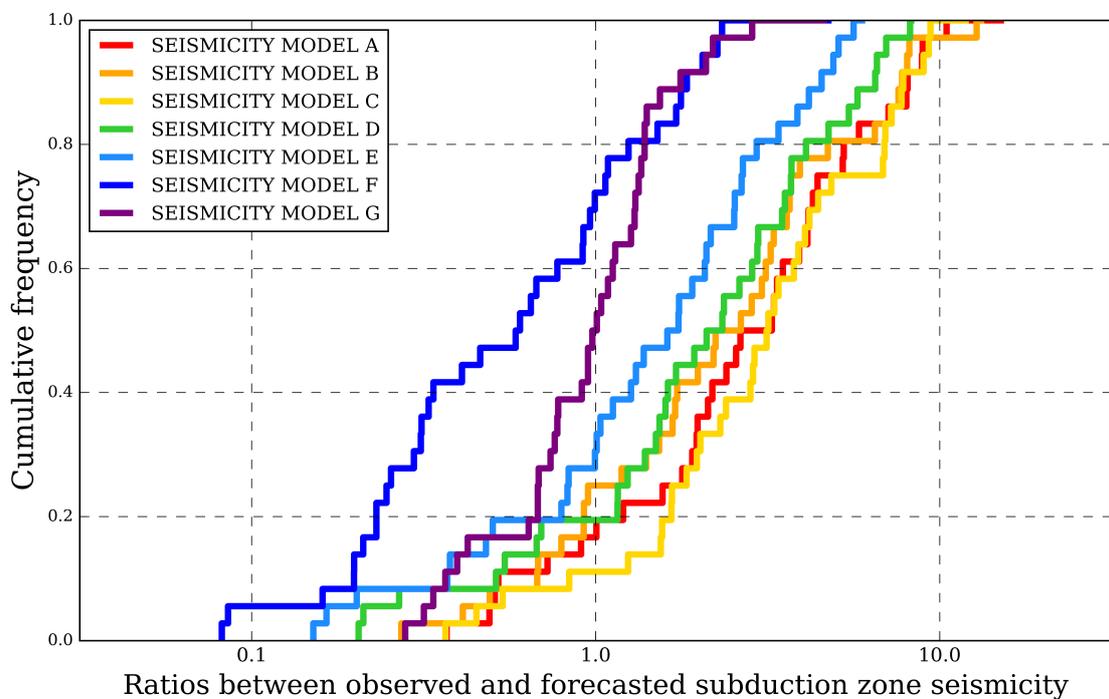
Abbildung 3.4: Cumulative distributions of the ratios between actual and forecasted seismicity, according to seven earthquake models A–G. We employ subduction-zone events recorded during the 1 January 1, 2015–31 December, 2018 period to pseudoprospectively evaluate the consistency of each seismicity forecast within subduction zones.

Tabelle 3.3: $m_T \geq 5.66$ observed N_s and predicted (seismicity models A–G) earthquakes in 37 subduction zones during the retrospectively test period. The retrospective evaluation period comprises from 1 January 1977 to 31 December 2014.

#	Subduction Zone	N_s	A	B	C	D	E	F	G
1	Northern Tonga	171	53.3	103.6	72.6	108.5	141.1	1149.2	171.0
2	Southern Tonga	119	38.4	63.9	41.4	46.9	58.1	472.9	119.0
3	Northern Kermadec	86	16.0	18.2	13.7	19.1	26.3	214.4	86.0
4	Southern Kermadec	178	220.1	295.2	25.8	30.4	40.7	331.2	178.0
5	Southern New-Hebrides	205	33.3	36.9	28.6	32.4	44.7	114.2	205.0
6	D'Entrecasteux	135	21.2	22.0	21.8	31.4	44.3	113.1	135.0
7	Northern New-Hebrides	179	17.2	23.1	19.6	21.4	30.0	76.4	179.0
8	Solomon	109	25.5	30.1	24.3	32.6	44.9	111.6	109.0
9	Bougainville	70	14.5	25.0	18.7	20.5	28.6	71.2	70.0
10	New Britain	212	33.2	36.6	24.8	35.6	52.4	130.4	212.0
11	Java	70	47.3	50.0	38.3	87.3	126.0	239.6	70.0
12	Sumatra	167	35.1	43.2	34.5	92.0	136.7	494.7	167.0
13	Andaman	170	18.6	27.9	28.6	92.6	128.2	556.3	170.0
14	Taiwan-Manila	171	69.7	143.9	85.2	114.7	160.6	426.5	171.0
15	Southern Ryukyu	40	31.6	41.2	24.6	27.3	38.2	58.7	40.0
16	Northern Ryukyu	70	58.2	62.3	42.7	80.6	119.6	183.8	70.0
17	Marianas	68	29.9	60.6	37.1	44.8	59.8	423.7	68.0
18	Izu-Bonin	126	43.8	47.6	38.6	60.0	81.5	577.6	126.0
19	Nankai	68	33.9	43.5	33.8	98.8	132.4	203.4	68.0
20	Tohoku	271	46.9	52.2	35.5	60.5	84.1	198.3	271.0
21	Southern Kuril	169	36.4	40.7	34.2	49.2	71.7	168.9	169.0
22	Northern Kuril	51	19.8	20.6	17.7	22.4	30.8	72.5	51.0
23	Kamchatka	90	30.5	29.8	27.8	34.4	48.2	113.6	90.0
24	Southwestern Aleutians	70	127.7	161.8	15.9	18.0	24.8	61.2	70.0
25	Central Aleutians	119	21.2	19.0	17.9	19.0	27.2	67.1	119.0
26	Eastern Aleutians	66	25.2	23.4	23.2	25.4	34.9	86.0	66.0
27	Western Alaska	27	21.3	22.7	18.9	25.4	37.7	93.1	27.0
28	Eastern Alaska	24	25.8	34.9	17.8	35.4	47.0	116.0	24.0
29	Mexico	137	46.0	48.4	59.8	80.5	109.2	206.3	137.0
30	Cocos	150	34.6	36.5	33.5	71.5	101.5	246.9	150.0
31	Costa Rica	60	26.3	23.9	22.7	27.4	39.6	74.8	60.0
32	Antilles	20	3.8	4.5	7.9	19.3	27.5	59.2	20.0
33	Colombia-Ecuador	31	13.6	10.8	15.9	24.7	34.2	83.1	31.0
34	Northern Peru	26	19.4	19.3	26.0	46.4	62.9	45.0	26.0
35	Southern Peru	50	23.8	15.4	24.8	33.3	45.0	32.2	50.0
36	Northern Chile	253	58.9	51.2	61.1	87.9	123.0	88.0	190.7
37	Southern Chile	14	18.7	14.0	21.1	45.0	57.6	41.2	14.0

Tabelle 3.4: $m_\mathrm{T} \geq 5.66$ recorded N_s and forecasted (earthquake models A–G) seismicity in 37 subduction zones during the pseudo-prospective test period. The pseudo-prospective evaluation period comprises from 1 January 2015 to 31 December 2018.

#	Subduction Zone	N_s	A	B	C	D	E	F	G
1	Northern Tonga	24	5.6	10.9	7.6	11.4	14.9	121.0	18.0
2	Southern Tonga	8	4.1	6.7	4.4	4.9	6.1	49.8	12.5
3	Northern Kermadec	7	1.7	1.9	1.5	2.0	2.8	22.6	9.1
4	Southern Kermadec	21	23.2	31.1	2.7	3.2	4.3	34.9	18.7
5	Southern New-Hebrides	28	3.5	3.9	3.0	3.4	4.7	12.0	21.6
6	D'Entrecasteux	18	2.2	2.3	2.3	3.3	4.7	11.9	14.2
7	Northern New-Hebrides	8	1.8	2.4	2.1	2.3	3.2	8.1	18.9
8	Solomon	24	2.7	3.2	2.6	3.4	4.7	11.8	11.5
9	Bougainville	8	1.5	2.6	2.0	2.2	3.0	7.5	7.4
10	New Britain	31	3.5	3.9	2.6	3.8	5.5	13.7	22.3
11	Java	5	5.0	5.3	4.0	9.2	13.3	25.2	7.4
12	Sumatra	12	3.7	4.5	3.6	9.7	14.4	52.1	17.6
13	Andaman	5	2.0	2.9	3.0	9.8	13.5	58.6	17.9
14	Taiwan-Manila	14	7.3	15.1	9.0	12.1	16.9	44.9	18.0
15	Southern Ryukyu	4	3.3	4.3	2.6	2.9	4.0	6.2	4.2
16	Northern Ryukyu	13	6.1	6.6	4.5	8.5	12.6	19.3	7.4
17	Marianas	11	3.1	6.4	3.9	4.7	6.3	44.6	7.2
18	Izu-Bonin	18	4.6	5.0	4.1	6.3	8.6	60.8	13.3
19	Nankai	7	3.6	4.6	3.6	10.4	13.9	21.4	7.2
20	Tohoku	26	4.9	5.5	3.7	6.4	8.9	20.9	28.5
21	Southern Kuril	6	3.8	4.3	3.6	5.2	7.6	17.8	17.8
22	Northern Kuril	7	2.1	2.2	1.9	2.4	3.2	7.6	5.4
23	Kamchatka	7	3.2	3.1	2.1	3.6	5.1	12.0	9.5
24	Southwestern Aleutians	7	13.4	17.0	1.7	1.9	2.6	6.5	7.4
25	Central Aleutians	13	2.2	2.0	1.9	2.0	2.9	7.1	12.5
26	Eastern Aleutians	7	2.7	2.5	2.4	2.7	3.7	9.1	7.0
27	Western Alaska	4	2.2	2.4	2.0	2.7	4.0	9.8	2.8
28	Eastern Alaska	1	2.7	3.7	1.9	3.7	5.0	12.2	2.5
29	Mexico	20	4.9	5.1	6.3	8.5	11.5	21.7	14.4
30	Cocos	12	3.6	3.9	3.5	7.5	10.7	26.0	15.8
31	Costa Rica	2	2.8	2.5	2.4	2.9	4.2	7.9	6.3
32	Antilles	6	0.4	0.5	0.8	2.0	2.9	6.2	2.1
33	Colombia-Ecuador	15	1.4	1.1	1.7	2.6	3.6	8.8	3.3
34	Northern Peru	1	2.0	2.0	2.7	4.9	6.6	4.7	2.7
35	Southern Peru	6	2.5	1.6	2.6	3.5	4.7	3.4	5.3
36	Northern Chile	44	6.2	5.4	6.4	9.3	13.0	9.3	20.1
37	Southern Chile	1	2.0	1.5	2.2	4.7	6.1	4.3	1.5

We then present mean estimates of hybrid coupling coefficient c_H that average out the spatial variability of coupling degree along each subduction segment (see Fig. 3.5 and Table 3.5). Additionally, we summarize differences between mean seismic coupling coefficients obtained for 37 subduction zones by Heuret et al. (2011) and in this study. Heuret et al. (2011) computed estimates of c by taking the ratio between the average seismic slip rate $\dot{D}$ estimated from $m_w \geq 7.0$ earthquakes recorded during the 1900–2007 period and the mean subduction velocity V_x (Fig. 3.1) obtained from global plate models. In contrast, we constrain hybrid coupling coefficients by efficiently combining geodetic strain rates, earthquake-catalog data and geomechanical information (Eq. 3.14). As a result, we detect the largest discrepancies between our estimates of coupling and Heuret's in the Eastern-Alaska, Southern-Chile and Andaman subduction zones, in which giant earthquakes have taken place during the last century (i.e. Christensen and Beck, 1994; Lomnitz, 2004; Stein and Okal, 2005). Furthermore, we identify significant variations for the Northern-Chile and Tohoku interfaces, where megathrust events (i.e. Vigny et al., 2011; Simons et al., 2011) occurred after the sampling period used by Heuret et al. (2011) to determine c. Thus, we interpret these differences to be due to effects of data sampling and the occurrence of megathrust earthquakes.
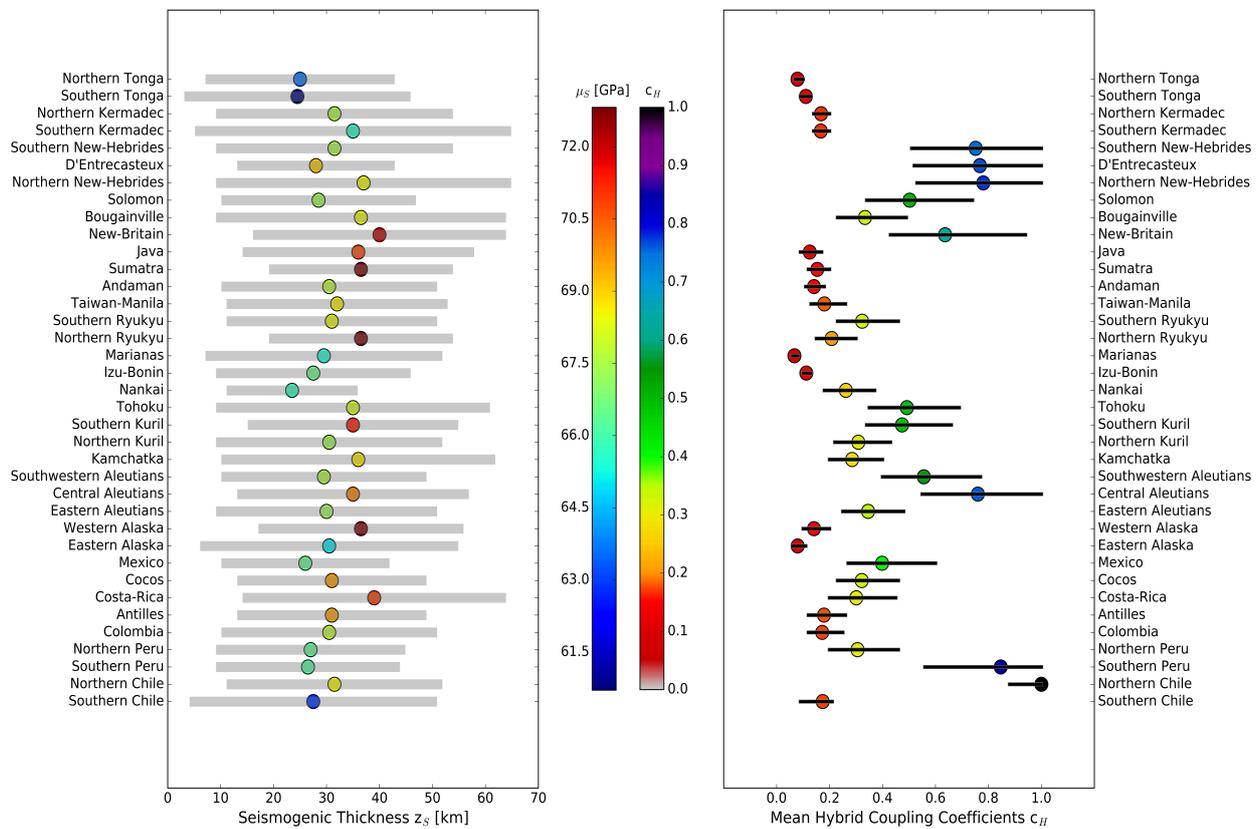
Abbildung 3.5: Mean regional earthquake parameters computed in this study. On the left hand side, we present as circles mean weighted elastic shear moduli μ_s, obtained from Eqs. 3.10 and 3.11. Additionally, we show regional seismogenic thicknesses z_s, defined from the upper U_z to the lower D_z bondaries of the seismogenic zone, as grey thick bars. On the right hand side, we report mean long-term hybrid coupling coefficients c_H (circles) with their corresponding standard deviations (thin black bars).

Tabelle 3.5: Mean regional earthquake parameters computed in this study. We also include seismic coupling coefficients c estimated by Heuret et al. (2011) to compare both methods to calibrate seismic locking for subduction interfaces.

#	Subduction Zone	c	c_H	μ_s[GPa]
1	Northern Tonga	0.04	0.08 [0.07 - 0.10]	64
2	Southern Tonga	0.12	0.11 [0.09 - 0.13]	61
3	Northern Kermadec	0.20	0.17 [0.14 - 0.20]	68
4	Southern Kermadec	0.01	0.17 [0.14 - 0.20]	66
5	Southern New-Hebrides	0.11	0.75 [0.51 - 1.00]	68
6	D'Entrecasteux	0.14	0.77 [0.52 - 1.00]	69
7	Northern New-Hebrides	0.08	0.78 [0.53 - 1.00]	68
8	Solomon	0.11	0.50 [0.34 - 0.74]	67
9	Bougainville	0.19	0.33 [0.23 - 0.49]	68
10	New Britain	0.20	0.64 [0.43 - 0.94]	72
11	Java	0.32	0.13 [0.09 - 0.17]	71
12	Sumatra	0.22	0.15 [0.12 - 0.20]	73
13	Andaman	0.98	0.14 [0.11 - 0.18]	68
14	Taiwan-Manila	0.02	0.18 [0.13 - 0.26]	69
15	Southern Ryukyu	0.14	0.32 [0.23 - 0.46]	68
16	Northern Ryukyu	0.04	0.21 [0.15 - 0.30]	73
17	Marianas	0.02	0.07 [0.06 - 0.08]	65
18	Izu-Bonin	0.05	0.11 [0.10 - 0.13]	67
19	Nankai	0.22	0.26 [0.18 - 0.37]	66
20	Tohoku	0.24	0.49 [0.35 - 0.69]	68
21	Southern Kuril	0.68	0.47 [0.34 - 0.66]	71
22	Northern Kuril	0.26	0.31 [0.22 - 0.43]	67
23	Kamchatka	1.00	0.29 [0.20 - 0.40]	69
24	Southwestern Aleutians	0.70	0.56 [0.40 - 0.77]	68
25	Central Aleutians	0.79	0.76 [0.55 - 1.00]	70
26	Eastern Aleutians	0.12	0.35 [0.25 - 0.48]	67
27	Western Alaska	0.10	0.14 [0.10 - 0.20]	73
28	Eastern Alaska	1.00	0.08 [0.06 - 0.11]	65
29	Mexico	0.43	0.40 [0.27 - 0.60]	66
30	Cocos	0.05	0.32 [0.23 - 0.46]	70
31	Costa Rica	0.03	0.30 [0.20 - 0.45]	71
32	Antilles	0.08	0.18 [0.12 - 0.26]	70
33	Colombia-Ecuador	0.58	0.17 [0.12 - 0.25]	68
34	Northern Peru	0.02	0.31 [0.20 - 0.46]	66
35	Southern Peru	0.28	0.85 [0.56 - 1.00]	66
36	Northern Chile	0.59	1.00 [0.88 - 1.00]	69
37	Southern Chile	1.00	0.14 [0.09 - 0.21]	63

Our computations of seismic coupling suggest that large earthquakes in subduction zones like New-Hebrides, Peru and New-Britain are expected in order to compensate for the elastic rebound. Moreover, these estimates serve as first insights of long-term mean coupling within some subduction interfaces like Solomon, Antilles and Marianas. Future improvements in the calibration of c_H will be achieved either with the acquisition of more earthquake-catalog data or with more precise kinematic inversions of interseismic geodetic data used to further regionalize the earthquake model.

We also discover that the regionalization of elastic shear moduli μ_s, implemented in the seismicity model E, noticeably reduces earthquake-rate underestimations within subduction zones (see Figs. 3.3 and 3.4). The cycle of strain accumulation and stress release within these tectonic margins involves both crustal and manle rocks, so that the optimal method to average their elastic shear moduli remains rather unclear. Therefore, we propose an alternative scheme to compute mean weighted regional rigidity values μ_s, based on the ak135-f global velocity reference model of Kennett et al. (1995) (Eqs. 3.10 and 3.11). Our estimates of regional μ_s are significantly larger than the uncertain mean global $\mu = 49 \pm 21$ reported by Bird and Kagan (2004) (see Table 3.5 and Fig. 3.5). Interestingly, these values well fit the upper uncertainty limit of such global estimate, are in agreement with other studies exploring rigidity variations with depth in subducting slabs (Bilek and Lay, 1999; Ji et al., 2010), and also minimize the epistemic uncertainies associated with the calibration of hybrid coupling coefficients c_H, according to sensitivity analyses carried out in this study.

We moreover find that the employment of regional subduction dip angles, a new regionalization scheme and an updated version of GSRM (seismicity models B–D) are insufficient to significantly reduce SHIFT_GSRM earthquake underestimations within subduction zones (see Figs. 3.3 and 3.4). The inclusion of specific subduction dip angles into the seismicity model D, for instance, slightly improves earthquake underpredictions to mean retrospective and pseudo-prospective values of 2.08 and 1.99. These improvements are mainly due to the contribution of the fault-plane orientation on long-term estimates of seismic moment rate for low-angle subduction zones. Within shallow megathrust interfaces such as Andaman, Nankai or Java, computations of $\dot{M}_{GEO}$ increase 69%, 66% and 56% by only including $f(\theta)$ into the moment-rate Eq. 3.9. In contrast, rates of seismic moment release calculated for steep subduction zones (e.g. Central Aleutians, Northern New-Hebrides or Southern Tonga) remain almost invariant after to applying geometric adjustment factors. Although $f(\theta)$ (Eq. 3.8) is the minimum correction factor needed to account for geometric effects in earthquake forecasting, we analytically demonstrate that the suggested use of specific fault-dipping angles (i.e. Bird et al., 2010; Bird and Kreemer, 2015) does not fully solve the problem of SHIFT_GSRM underestimation factors in subduction zones.

We additionally observe that the incorporation of the global regionalization scheme of Kagan et al. (2010) into the seismicity model C detectably increases the ratios of actual and predicted interface earthquakes. This regionalization model uses global seismicity to define tectonic zones based on variations of earthquake parameters such as corner magnitude and seismic coupling. Distinctively, the global regionalization framework of Kreemer et al. (2002), utilized to construct the seismicity model B, computes tectonic moment rates from geodetic velocities and rates of Quaternary fault slip. Thus, we interpret that minor differences between seismicity models B and C are due to the use of earthquake-catalog data and geodetic strain rates to distinguish subduction-zone margins. Geodetic measurements of the interseismic period provide a more robust comprehension of the plate interface than seismicity does (Scholz and Campos, 2012). However, we utilize the global regionalization scheme of Kagan et al. (2010) for seismicity models C–G to fairly compare estimates of subduction-zone seismicity with other global earthquake models like SHIFT_GSRM2f and GEAR1 in future work.

We furthermore report that the use of GSRM2.1 in the seismicity model B shifts the ratios of actual and forecasted interface seismicity to mean values of 2.44 and 2.33, according to retrospective and pseudo-prospective evaluations. Similar results have previously been described for SHIFT_GSRM2f (i.e. Bird and Kreemer, 2015), which based on GSRM2.1 also underpredicts rates of subduction-zone earthquake activity, despite the incorporation of a higher resolution strain-rate model.

Tabelle 3.6: N-test results for earthquake models A−G during the 1 January 1977−31 December 2014 retrospective period. Here, ▼ stands for underestimation, ▲ for overprediction, and ★ for consistency.

#	Subduction Zone	A	B	C	D	E	F	G
1	Northern Tonga	▼	▼	▼	▼	▼	▲	★
2	Southern Tonga	▼	▼	▼	▼	▼	▲	★
3	Northern Kermadec	▼	▼	▼	▼	▼	▲	★
4	Southern Kermadec	▲	▲	▼	▼	▼	▲	★
5	Southern New-Hebrides	▼	▼	▼	▼	▼	▼	★
6	D'Entrecasteux	▼	▼	▼	▼	▼	▼	★
7	Northern New-Hebrides	▼	▼	▼	▼	▼	▼	★
8	Solomon	▼	▼	▼	▼	▼	★	★
9	Bougainville	▼	▼	▼	▼	▼	★	★
10	New Britain	▼	▼	▼	▼	▼	▼	★
11	Java	▼	▼	▼	★	▲	▲	★
12	Sumatra	▼	▼	▼	▼	▼	▲	★
13	Andaman	▼	▼	▼	▼	▼	▲	★
14	Taiwan-Manila	▼	▼	▼	▼	★	▲	★
15	Southern Ryukyu	★	★	▼	▼	★	▲	★
16	Northern Ryukyu	★	★	▼	★	▲	▲	★
17	Marianas	▼	★	▼	▼	★	▲	★
18	Izu-Bonin	▼	▼	▼	▼	▼	▲	★
19	Nankai	▼	▼	▼	▲	▲	▲	★
20	Tohoku	▼	▼	▼	▼	▼	▼	★
21	Southern Kuril	▼	▼	▼	▼	▼	▲	★
22	Northern Kuril	▼	▼	▼	▼	▼	▲	★
23	Kamchatka	▼	▼	▼	▼	▼	▲	★
24	Southwestern Aleutians	▲	▲	▼	▼	▼	★	★
25	Central Aleutians	▼	▼	▼	▼	▼	▼	★
26	Eastern Aleutians	▼	▼	▼	▼	▼	▲	★
27	Western Alaska	★	★	★	★	★	▲	★
28	Eastern Alaska	★	★	★	★	▲	▲	★
29	Mexico	▼	▼	▼	▼	▼	▲	★
30	Cocos	▼	▼	▼	▼	▼	▲	★
31	Costa Rica	▼	▼	▼	▼	▼	★	★
32	Antilles	▼	▼	▼	★	▲	▲	★
33	Colombia-Ecuador	▼	▼	▼	★	▲	▲	★
34	Northern Peru	★	★	★	▲	▲	★	★
35	Southern Peru	▼	▼	▼	▼	★	▼	★
36	Northern Chile	▼	▼	▼	▼	▼	▼	▼
37	Southern Chile	★	★	★	▲	▲	▲	★

Tabelle 3.7: N-test results for seismicity models A−G during the 1 January 2015−31 December 2018 pseudo-prospective period. Here, ▼ stands for underestimation, ▲ for overprediction, and ★ for consistency.

#	Subduction Zone	A	B	C	D	E	F	G
1	Northern Tonga	▼	▼	▼	▼	▼	▲	★
2	Southern Tonga	★	★	★	★	★	▲	★
3	Northern Kermadec	▼	▼	▼	▼	▼	▲	★
4	Southern Kermadec	★	★	▼	▼	▼	▲	★
5	Southern New-Hebrides	▼	▼	▼	▼	▼	▼	★
6	D'Entrecasteux	▼	▼	▼	▼	▼	★	★
7	Northern New-Hebrides	▼	▼	▼	▼	▼	★	▲
8	Solomon	▼	▼	▼	▼	▼	▼	▼
9	Bougainville	▼	▼	▼	▼	▼	★	★
10	New Britain	▼	▼	▼	▼	▼	▼	★
11	Java	★	★	★	★	▲	▲	★
12	Sumatra	▼	▼	▼	★	★	▲	★
13	Andaman	★	★	★	★	▲	▲	▲
14	Taiwan-Manila	▼	★	★	★	★	▲	★
15	Southern Ryukyu	★	★	★	★	★	★	★
16	Northern Ryukyu	▼	▼	▼	★	★	★	★
17	Marianas	▼	★	▼	▼	★	▲	★
18	Izu-Bonin	▼	▼	▼	▼	▼	▲	★
19	Nankai	★	★	★	★	★	▲	★
20	Tohoku	▼	▼	▼	▼	▼	★	★
21	Southern Kuril	★	★	★	★	★	▲	▲
22	Northern Kuril	▼	▼	▼	▼	★	★	★
23	Kamchatka	★	★	★	★	★	★	★
24	Southwestern Aleutians	★	▲	▼	▼	▼	★	★
25	Central Aleutians	▼	▼	▼	▼	▼	★	★
26	Eastern Aleutians	▼	▼	▼	▼	★	★	★
27	Western Alaska	★	★	★	★	★	★	★
28	Eastern Alaska	★	★	★	★	★	▲	★
29	Mexico	▼	▼	▼	▼	▼	★	★
30	Cocos	▼	▼	▼	★	★	▲	★
31	Costa-Rica	★	★	★	★	★	▲	★
32	Antilles	▼	▼	▼	▼	★	★	▼
33	Colombia-Ecuador	▼	▼	▼	▼	▼	★	▼
34	Northern Peru	★	★	★	★	▲	★	★
35	Southern Peru	★	▼	★	★	★	★	★
36	Northern Chile	▼	▼	▼	▼	▼	▼	▼
37	Southern Chile	★	★	★	★	▲	★	★

Mean ratios of actual and predicted seismicity may actually be misleading, because both over- and underpredictions are observed. Thus, we finally complement our earthquake-model evaluations with N-tests to more reliably describe the consistency of each seismicity forecast (see Tables 3.6 and 3.7). N-test results show that seismicity models G and F significantly outperform their predecessor earthquake models during the evaluation periods. In general, the seismicity model F overpredicts earthquake activity in subduction zones with low degree of predicted mean coupling (e.g. Northern Tonga, Java and Izu Bonin). Additionally, seismicity rates obtained from the seismicity model G during the pseudoprospective period are consistent with the observed number of earthquakes in almost all subduction interfaces. These results reveal the potential of this approach to accurately estimate long-term shallow interface seismicity. Nevertheless, independent prospective evaluations through the CSEP testing center are needed to more objectively evaluate the forecasting power of this regionalized earthquake model.

The lack of regional earthquake parameters in a few subduction zones, the use of empirical calibration factors to constrain long-term seismic coupling coefficients and the interpolation of inland geodetic strain rates to oceanic floors are the major limitations of the seismicity model G. However, this earthquake model provides higher-resolution estimates of shallow seismicity within subduction zones (see Fig. 3.6) in comparison with global earthquake-based tectonic forecasts (e.g. Giardini et al., 1999), because it optimally combines geodetic strain rates, regional geomechanical parameters and earthquake-catalog data.

3.4 Conclusions and Prospects

The SHIFT_GSRM seismicity model is a tectonic forecast constructed from geodetic strain rates and mean global geomechanical parameters. The use of well-constrained earthquake parameters is thus essential to reliably compute long-term estimates of seismic moment release and shallow seismicity. In this study, we show in seven progressive steps that SHIFT_GSRM earthquake-rate underestimations in 37 subduction zones are mainly due to the use of a global probability function for seismic moment release $\dot{M}^{\mathrm{CMT}}$ that poorly captures the variability among subduction zones. By replacing this global function with regional estimates of seismic moment $\dot{M}_{\mathrm{SEIS}}$, we shift the ratios between observed and forecasted seismicity rates to an average factor of 0.57, according to pseudo-prospective evaluations (see Table 3.8). Independent prospective tests are required to better describe the forecasting skills of this earthquake model. Nonetheless, sucessful retrospective and pseudo-prospective test results show the initial capacity of this regionalized earthquake model to properly forecast subduction-zone seismicity.

Tabelle 3.8: Summary of subduction-zone earthquake models introduced in this study. Here, MRR and MPR stand for average retrospective and pseudo-prospective ratios, between observed and predicted seismicity, respectively.

Model	New Input	MRR	MPR
A	Raw SHIFT_GSRM Earthquake Model	2.82	2.70
B	Global Strain Rate Map 2.1	2.44	2.33
C	Regionalization Scheme KRS2010	3.14	3.00
D	Regional Subduction Dip Angles	2.08	1.99
E	Regional Elastic Shear Moduli	1.50	1.43
F	Regional Model Seismic Moment Rates	0.60	0.57
G	Regional Seismogenic Thicknesses	1.01	0.96
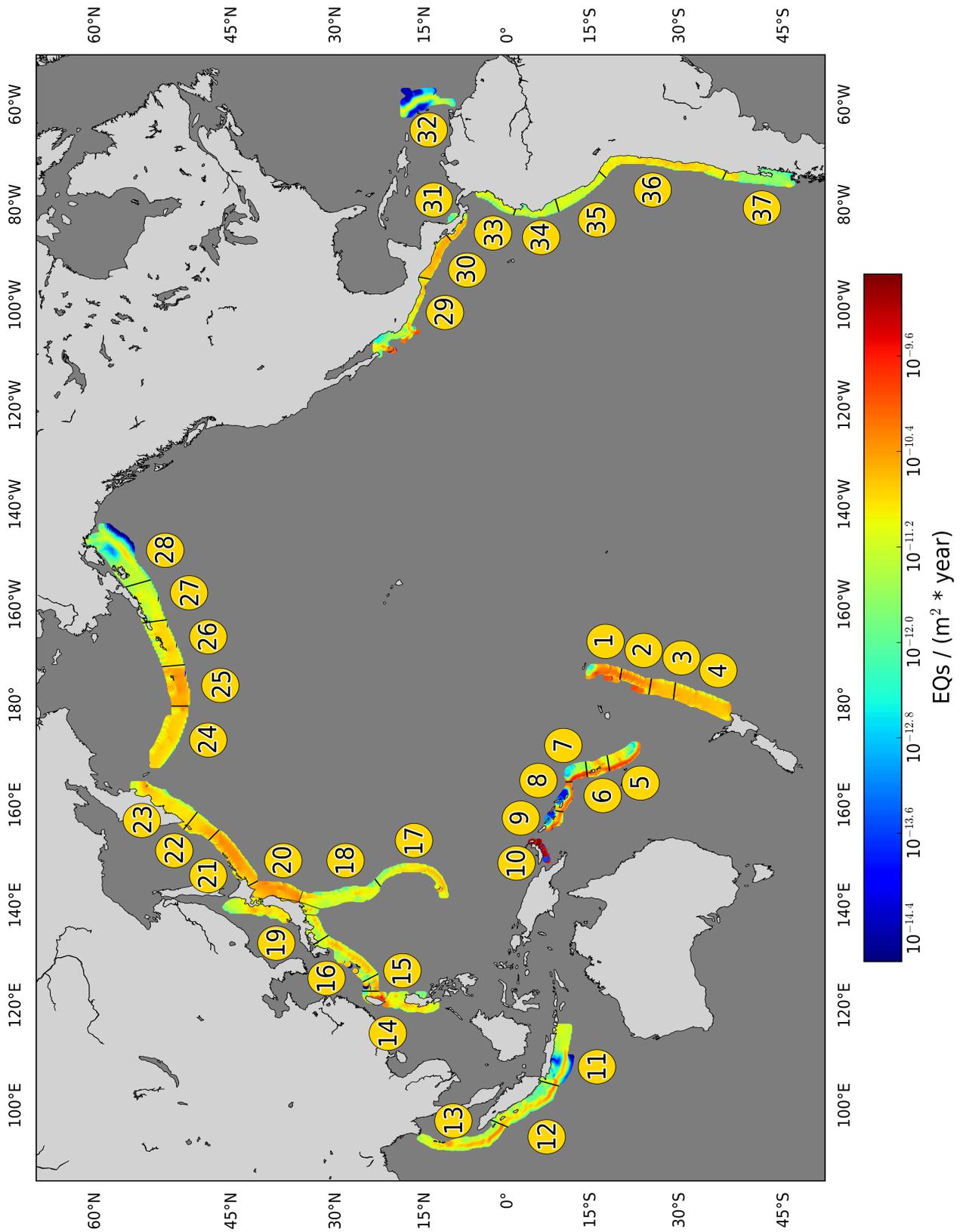
Abbildung 3.6: Long-term rates of shallow seismicity at or above a magnitude threshold $m_{\mathrm{T}} \geq 5.66$ within 37 subduction zones, according to the seismicity model G. We express earthquake-rate densities as epicentroids per square meter per year.

In future work, we will refer to the seismicity model G as the Subduction Megathrust Earthquake Rate Forecast (SMERF; see Fig. 3.6) model, and we will integrate it into a global seismicity model to generate a global earthquake forecast based on the principle of conservation of moment (see Chapter 4). Recently, Strader et al. (2018) prospectively evaluated the GEAR1 model and its parent components within CSEP, and found significant discrepancies in spatial seismicity patterns between SHIFT_GSRM & SHIFT_GSRM2f predictions and observations, mainly due to underpredictions of subduction-zone earthquake activity. Thus, we conclude that SMERF may actually contribute to the development, update and improvement of global hybrid seismicity models like GEAR1.

Data and Resources

The observation earthquake catalog was obtained from the global CMT catalog available at www.globalcmt.org/CMTsearch.html (last accessed on January 1, 2019). Additionally, the Global Strain Rate Map 2.1 was provided by Cornë Kreemer, and the tectonic regionalization grid is available as electronic supplement to Kagan et al. (2010).

Acknowledgments

Kapitel 4

Two Global Ensemble Seismicity Models Obtained from the Combination of Interseismic Strain Measurements and Earthquake-catalog Information

Abstract

Global seismicity models provide scientific hypotheses about the rate, location and magnitude of future earthquakes to occur worldwide. Given the aleatory variability of seismic activity and epistemic uncertainties in seismicity forecasting, the veracity of these hypotheses can only be either confirmed or rejected after prospective forecast evaluation. In this study, we present the construction of and test results for two updated global earthquake models, aimed at providing mean estimates of shallow ($d \leq 70$ km) seismicity for seismic hazard assessment. These approaches, referred to as the Tectonic Earthquake Activity Model (TEAM) and the World Hybrid Earthquake Estimates based on Likelihood scores (WHEEL) model, use the Subduction Megathrust Earthquake Rate Forecast (SMERF2), an earthquake-rate model for subduction zones constrained by geodetic strain measurements and earthquake-catalog information. Thus, TEAM and WHEEL are global ensemble seismicity models, which capture two independent components necessary for long-term earthquake modeling: interseismic crustal strain accumulation and sudden lithospheric stress release. The calibration period for TEAM and WHEEL extends from 1 January 1977 to 31 December 2013. Hence, we use earthquakes recorded during the 2014–2019 period to pseudo-prospectively evaluate the forecasting skills of these earthquake-rate models, and statistically compare their performances to that of the Global Earthquake Activity Rate (GEAR1) model. As a result, GEAR1 and WHEEL are the most informative seismicity models during the pseudo-prospective testing period, as both rank with the highest information scores among all participant earthquake-rate forecasts. Nonetheless, further prospective evaluations are required to more accurately describe the abilities of these global ensemble seismicity models to forecast longer-term earthquake activity.[1]

4.1 Introduction

Until recently, there has been some skepticism within the earthquake and engineering community about the capacity of earthquake-rate models to reliably characterize seismicity patterns. This lack of credulity primarily stems from the strongly stochastic nature of earthquakes and the relatively limited disposal of target data to independently calibrate and test seismicity models (Werner et al., 2011). Accordingly, only prospective forecast experiments are considered rigorous enough to objectively describe the forecasting skills of any seismicity model (Taroni et al., 2014). Nonetheless, one inconvenience of prospective forecast tests is the time that one might have to wait to sample a statistically representative number of large events. Hence, there have been multiple efforts to reduce prospective forecast evaluation periods by expanding earthquake forecasting areas, leading to the generation of global seismicity models. As a result, global $m \geq 5.8$ earthquake forecasts currently offer prospective test results that can be obtained in only 1–8 years (Bird et al., 2015).

The Seismic Hazard Inferred from Tectonics based on the first iteration of the Global Strain Rate Map (SHIFT_GSRM; Kreemer et al., 2003; Bird and Liu, 2007; Bird et al., 2010) model is a global seismicity model designed to improve estimates of seismic hazard worldwide. SHIFT_GSRM converts interseismic strain rates into long-term rates of seismic moment using global "geodesy-to-seismicityparameters, such as fault-dipping angles, coupled seismogenic thicknesses and elastic shear moduli (i.e. Bird and Kagan, 2004). Depending on the degree of coupling, seismicity rates are thereupon obtained by assuming a tapered Gutenberg-Richter frequency-magnitude distribution. During the 1977–2009 retrospective test period, the raw SHIFT_GSRM forecast is consistent with the observations in continental convergent boundaries. However, it significantly underpredicts earthquake rates in subduction zones, presumably due to geometrical effects, inappropriately captured by the global average subduction dip angle employed in the formulation. Thus, the SHIFT_GSRM authors empirically corrected the forecast as the last step in constructing this seismicity model.

A few years later, Bird and Kreemer (2015) presented a revised version of SHIFT_GSRM: the SHIFT_GSRM2f earthquake model. SHIFT_GSRM2f is based on the updated Global Strain Rate Map (GSRM2.1; Kreemer et al., 2014), a global velocity gradient tensor field and continuous strain-rate model obtained from 22,415 GPS interseismic velocities. Among other improvements, SHIFT_GSRM2f incorporates a spatial smoothing of model strain rates around offshore plate boundaries to provide high-resolution estimates of earthquake activity. Nevertheless, the uncorrected SHIFT_GSRM2f forecast underestimates global earthquake rates during the 2005–2012 retrospective evaluation period, primarily due to underpredictions of subduction-zone seismicity. As a result, the SHIFT_GSRM2f modelers applied empirical calibration factors to improve the forecast in subduction zones by assuming that discrepancies between this new earthquake-rate forecast and the observations are mainly derived from the absence of specific fault-dipping angles in the moment-rate balance equation.

SHIFT_GSRM2f was then combined with the Kagan-Jackson Smoothed-Seismicity (KJSS; Kagan and Jackson, 2011) model to generate the Global Earthquake Activity Rate (GEAR1; Bird et al., 2015) model. KJSS, the seismicity parent component of GEAR1, averages out the exceedance rates of 30,000 earthquakes reported in the 1977–2008 Global Centroid Moment Tensor Project (CMT; Dziewonski et al., 1981; Ekström et al., 2012) catalog, and smoothes their locations with an anisotropic kernel function to estimate global earthquake activity, with special focus on intraplate regions. In comparison, SHIFT_GSRM2f, the tectonic earthquake forecast of GEAR1, computes high-resolution rates of seismicity along faults and tectonic plate boundaries based on GSRM2.1. As a result, GEAR1 is a hybrid seismicity model aimed at forecasting $m6-9$ earthquakes everywhere on Earth.

GEAR1 and its individual components were submitted to the Collaboratory for the Study of Earthquake Predictability (CSEP; Jordan, 2006; Schorlemmer and Gerstenberger, 2007; Zechar et al., 2010) testing center for independent evaluation. According to 2-year prospective test results, GEAR1 significantly outperforms both of its parent forecasts, providing preliminary support to combine geodetic strain rates with smoothed-seismicity data for long-term earthquake forecasting (Strader et al., 2018). Specifically, the total earthquake number N-, spatial S-, and magnitude M-distributions forecasted by GEAR1 were all consistent with observed seismicity. Nonetheless, both SHIFT_GSRM and SHIFT_GSRM2f failed the spatial S-test evaluation, in spite of the high spatial strain-rate resolution offered by GSRM and GSRM2.1.

During the twentieth century, subduction interface seismicity released almost 90% of the global seismic moment rate, and comprised approximately 60% of the earthquake activity observed worldwide (Pacheco and Sykes, 1992; Bird et al., 2010). Based on this, Bayona Viveros et al. (2019) constructed the Subduction Megathrust Earthquake Rate Forecast (SMERF) as a complementary approach to the raw SHIFT_GSRM and SHIFT_GSRM2f forecasts to improve their forecasted number, and potentially spatial, distributions in subduction zones. Similar to SHIFT_GSRM2f, SMERF computes long-term budgets of seismic moment from interseismic strain rates and instrumentally-recorded seismicity. Nevertheless, SMERF uses mean regional —not global— geodesy-to-seismicity conversion parameters to account for the great diversity of earthquake patterns among subduction margins (i.e. Heuret et al., 2011; Kagan and Jackson, 2016). As a result, SMERF holds the initial capacity to properly forecast subduction-zone seismicity, according to retrospective and pseudo-prospective N-test results for the 1977–2014 and 2015–2018 periods.

One limitation of SMERF, however, is the relatively large number of seismicity parameters it needs to be constructed and the current availability of earthquake-catalog information to individually calibrate them. Hence, we present a revised version of SMERF that depends on a reduced amount of earthquake parameters, referred to as the SMERF2 model. In addition, we integrate SMERF2 estimates in subduction zones with SHIFT_GSRM2f computations outside of these tectonic margins to produce the global Tectonic Earthquake Activity Model (TEAM). Furthermore, we combine KJSS with this updated tectonic earthquake model to create the World Hybrid Earthquake Estimates based on Likelihood scores (WHEEL) model; an alternative hybrid seismicity approach to GEAR1. Finally, we utilize $h \leq 70$ km, $m \geq 5.95$ actual earthquakes recorded during the 2014–2019 period to pseudo-prospectively evaluate the consistency of these global ensemble seismicity models with the observations, and quantitatively compare their forecasting skills with GEAR1, our brenchmark global earthquake-rate model.

4.2 Seismicity Models

4.2.1 The Global Earthquake Activity Rate (GEAR1) model

GEAR1 is a global hybrid earthquake model resulting from a multiplicative log-linear blend of smoothed-seismicity data and geodetic strain rates:

$$H_{ij} = N\left\{\max[(S_{ij}^d \cdot T_{ij}^{1-d}), f]\right\}. \tag{4.1}$$

In this equation, H_{ij} represents the hybrid seismicity forecast in each grid cell centered on longitude i and latitude j, N is a normalization factor adjusting the global predicted earthquake rate to the global rate of observed seismicity according to the 1977–2004 Global CMT catalog, max stands for the maximum value, S_{ij} denotes the smoothed seismicity KJSS parent forecast, d is an optimization parameter to be determined, T_{ij} refers to the tectonic SHIFT_GSRM2f

earthquake component, and f symbolizes a baseline seismicity rate defined as the $\min[\min(S_{ij}),$ $\min(T_{ij})]$. The optimized combination between parent forecast components for GEAR1 was determined by maximizing the I_1 (success) information score of Kagan (2009) from actual 2005–2012 seismicity. In this manner, Bird et al. (2015) found that the most informative GEAR1 forecast is derived from a multiplicative blend of parent components, with exponent $d = 0.6$ on KJSS (Seismicity) and $c = 1 - d = 0.4$ on SHIFT_GSRM2f (Tectonics). Based on this, the GEAR1 modelers enhanced the preferred hybrid seismicity model by recomputing the Seismicity and Tectonics parent forecasts using the complete 1977–2013 Global CMT catalog. As a result, the update of such a preferred hybrid earthquake-rate model is what they refer to as the GEAR1 seismicity forecast.

4.2.2 The updated Subduction Megathrust Earthquake Rate Forecast (SMERF2)

One of the most basic ways that SMERF differs from SHIFT_GSRM2f is that SMERF identifies 37 subduction margins, according to variations of seismicity parameters (i.e. Heuret et al., 2011; Kagan and Jackson, 2016). In contrast, SHIFT_GSRM2f distinguishes one subduction zone from other tectonic plate boundaries based on discontinuities of relative plate velocity (i.e. Bird et al., 2009). Thus, SMERF posesses a relatively large number of degrees of freedom, specifically regional seismic coupling coefficients, to be individually constrained by available earthquake-catalog data. Therefore, we present a revised version of SMERF that depends on a reduced number of seismicity parameters, referred to as the SMERF2 model. Same as SMERF, SMERF2 uses the trench segmentation model of Heuret et al. (2011) to assign geometric parameters, such as seismogenic thicknesses, trench lengths and fault-dipping angles to 37 subduction areas. Nonetheless, this updated model version employs only one average elastic shear modulus $\bar{\mu}_s$ = 68 GPa for subduction interfaces, derived from the results of Bayona Viveros et al. (2019). Moreover, SMERF2 utilizes the subduction segmentation model of Kagan and Jackson (2016) to re-calibrate mean regional seismic coupling coefficients c_H from estimates of corner magnitude computed in such a study. For this aim, we group together some of the Heuret et al. (2011) subduction zones (e.g. Northern Peru, Southern Peru, Northern Chile, Southern Chile) to relate them to the Kagan and Jackson (2016) subduction segments (e.g. Andean S. America). As a result, SMERF2 reduces the number of seismic coupling coefficients from 37 to 14 (see Table 4.1) to provide high-resolution estimates of subduction-zone seismicity (see Fig. 4.1).

4.2.3 The Tectonic Earthquake Activity Model (TEAM)

Strictly speaking, SHIFT_GSRM2f and SMERF2 are hybrid seismicity models, because both unavoidably employ earthquake information to translate geodetic strain rates into long-term estimates of shallow seismicity. In particular, the predicted number of earthquakes is calibrated from observed seismicity, and the spatial earthquake distribution is provided by interseismic strain measurements. Nonetheless, we assume SHIFT_GSRM2f and SMERF2 to be geodetic-based earthquake models. Thus, we combine SMERF2 estimates of subduction-zone seismicity with SHIFT_GSRM2f computations of earthquake activity outside of SMERF2 subduction margins to generate the global Tectonic Earthquake Activity Model (TEAM; see Figs. 4.2 and 4.3). To achieve it, we first preserve the annual earthquake rate for subduction zones predicted by SMERF2, as seismic coupling coefficients c_H employed to build the model are empirically calibrated from actual 1977–2013 interface seismicity (see Eq. 3.14). Then, we multiply SHIFT_GSRM2f earthquake rate densities in non-SMERF2 subduction areas by a factor of approximately 0.8 to adjust the mean global rate to approximately 176 $m \geq 5.767$ earthquakes per year, provided by the 1977–2013 Global CMT catalog.

4.2.4 The World Hybrid Earthquake Estimates based on Likelihood scores (WHEEL) model

We finally blend the new TEAM seismicity forecast with KJSS to create an updated global hybrid earthquake model. Same as GEAR1, this updated model uses a multiplicative log-linear combination of earthquake parent components, with exponent $d = 0.6$ on the Seismicity forecast (KJSS) and $c = 1 - d = 0.4$ on the Tectonic constituent (now TEAM). As a result, the World Hybrid Earthquake Estimates based on Likelihood scores (WHEEL) model serves as an alternative approach to GEAR1 to compute global earthquake activity, with special focus on subduction-zone seismicity (see Figs. 4.4 and 4.5).

Tabelle 4.1: Mean estimates of seismic coupling coefficient c_{H} for 14 Kagan and Jackson (2016) subduction interfaces. Based on Eq. 3.14, we empirically calibrate these values to construct SMERF2.

#	Heuret's subduction zone	KJ's subduction zone	c_{H}
1	Northern Tonga Southern Tonga Northern Kermadec Southern Kermadec	Kermadec- Tonga- Samoa	**0.12** [0.10 - 0.14]
2	Southern New-Hebrides D'Entrecasteux Northern New-Hebrides	New Hebrides Is.	**0.82** [0.56 - 1.00]
3	Solomon Bougainville New Britain	Bismarck- Solomon Is.	**0.56** [0.38 - 0.83]
4	Java	Sunda Arc	**0.16** [0.11 - 0.22]
5	Sumatra Andaman	Andaman Is.- Sumatra	**0.16** [0.12 - 0.21]
6	Taiwan-Manila	Taiwan	**0.21** [0.15 - 0.30]
7	Southern Ryukyu Northern Ryukyu Nankai	S.E. Japan- Ryukyu Is.	**0.26** [0.18 - 0.37]
8	Tohoku Southern Kuril Northern Kuril Kamchatka	Japan- Kamchatka	**0.47** [0.33 - 0.66]
9	Marianas Izu-Bonin	Guam-Japan	**0.10** [0.09 - 0.12]
10	Southwestern Aleutians Central Aleutians Eastern Aleutians Western Alaska Eastern Alaska	Alaska- Aleutian Arc	**0.36** [0.26 - 0.50]
11	Mexico Cocos	Mexico- Guatemala	**0.47** [0.32 - 0.70]
12	Costa Rica Colombia-Ecuador	Central America	**0.21** [0.15 - 0.30]
13	Antilles	Caribbean Loop	**0.17** [0.11 - 0.24]
14	Northern Peru Southern Peru Northern Chile Southern Chile	Andean S.America	**0.79** [0.52 - 1.00]
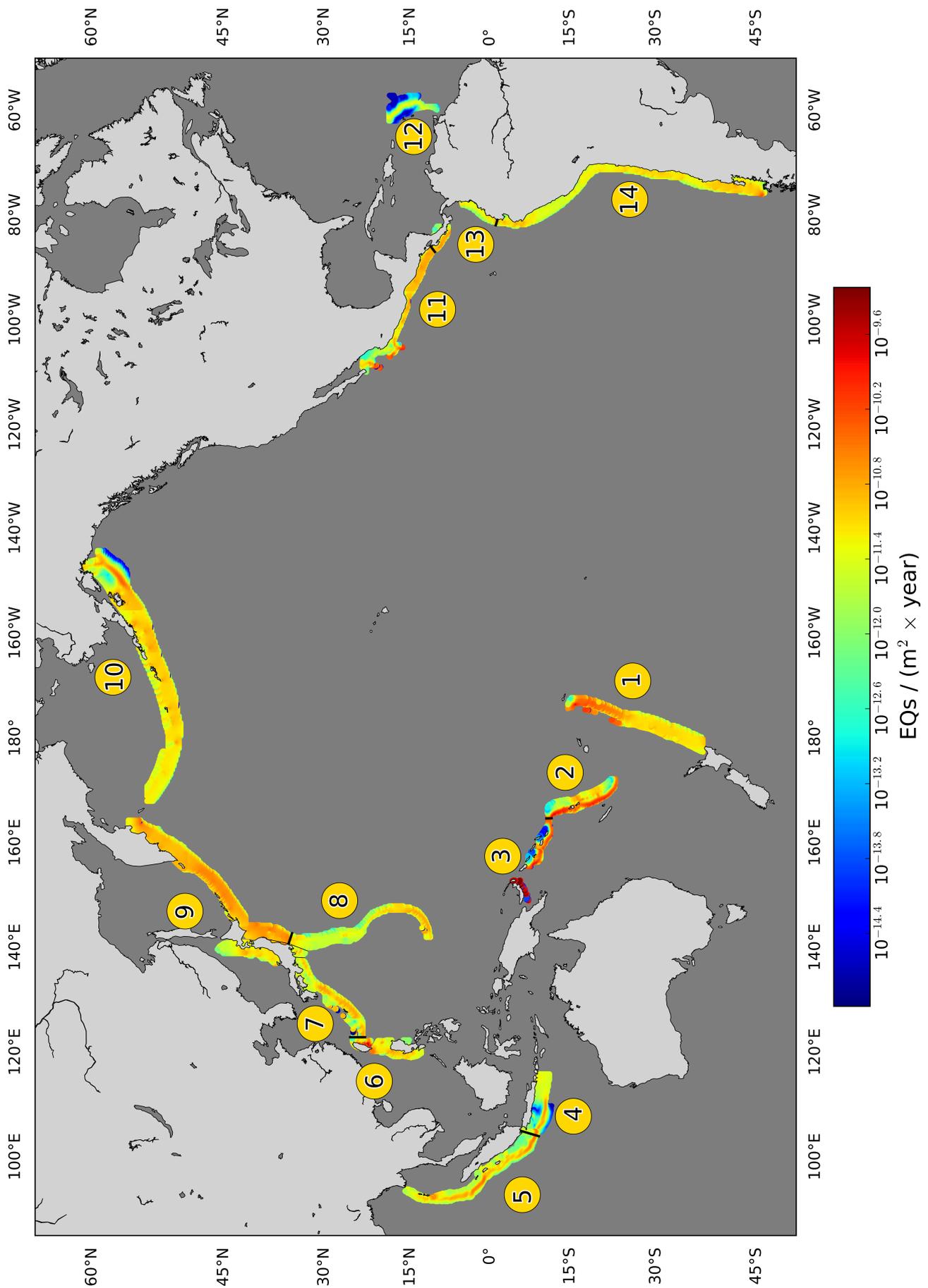
Abbildung 4.1: Annual density rates of shallow $d \leq 70$ km, $m \geq 5.95$ interface seismicity, provided by the updated Subduction Megathrust Earthquake Rate Forecast (SMERF2).
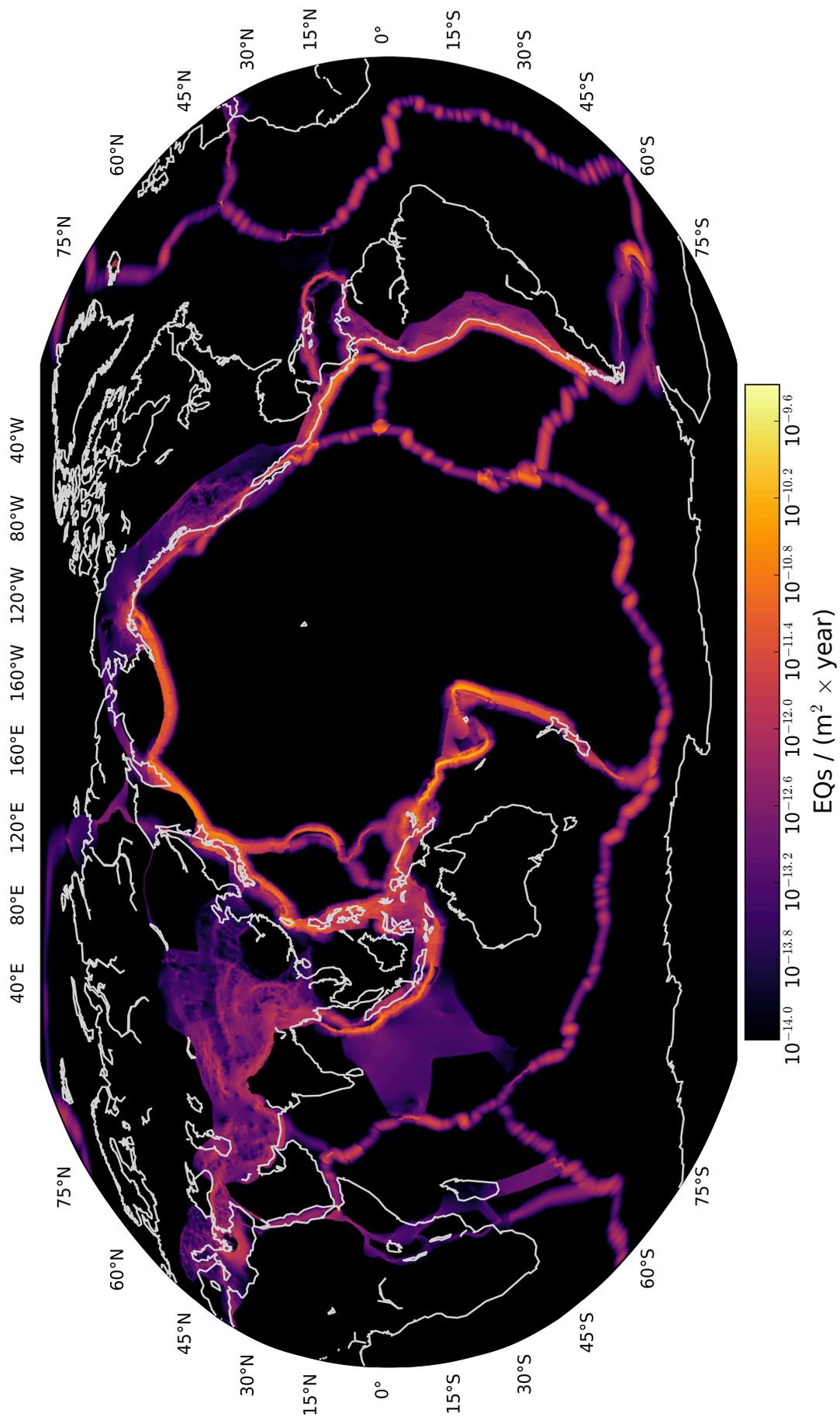
Abbildung 4.2: Forecast map showing annual $d \leq 70$ km, $m \geq 5.95$ earthquake rate densities (per m^2), derived from the global Tectonic Earthquake Activity Model (TEAM).
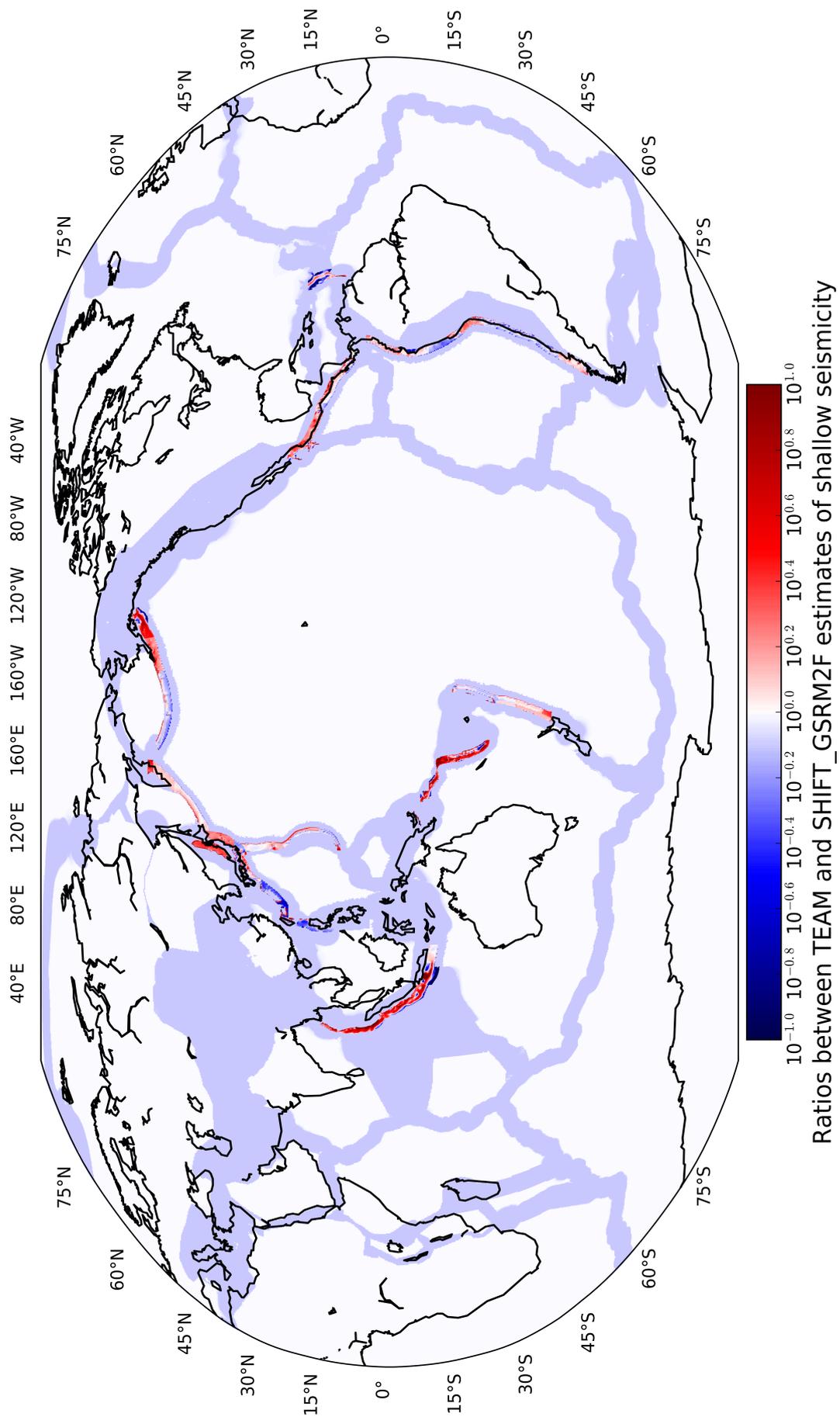
Abbildung 4.3: Variations in $m \geq 5.95$ earthquake rate densities forecasted by TEAM and SHIFT_GSRM2f. In red regions, TEAM estimates a larger number of earthquakes per year than SHIFT_GSRM2f. On the contrary, blue points denote locations where SHIFT_GSRM2f computes larger seismicity rates than TEAM.
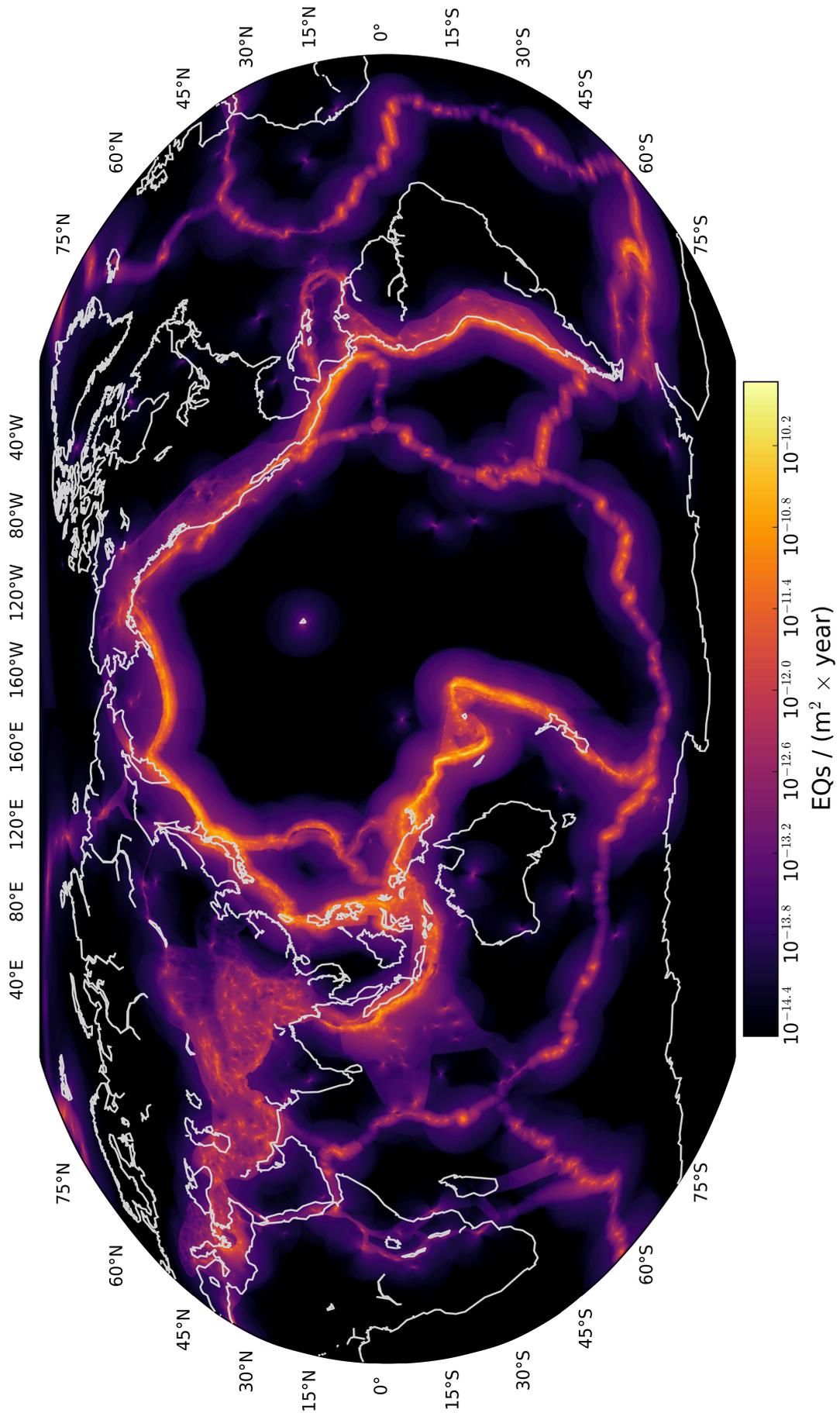
Abbildung 4.4: Mean estimates of annual $d \leq 70$ km, $m \geq 5.95$ earthquake rate densities (per m^2), computed by the World Hybrid Earthquake Estimates based on Likelihood scores (WHEEL) model.
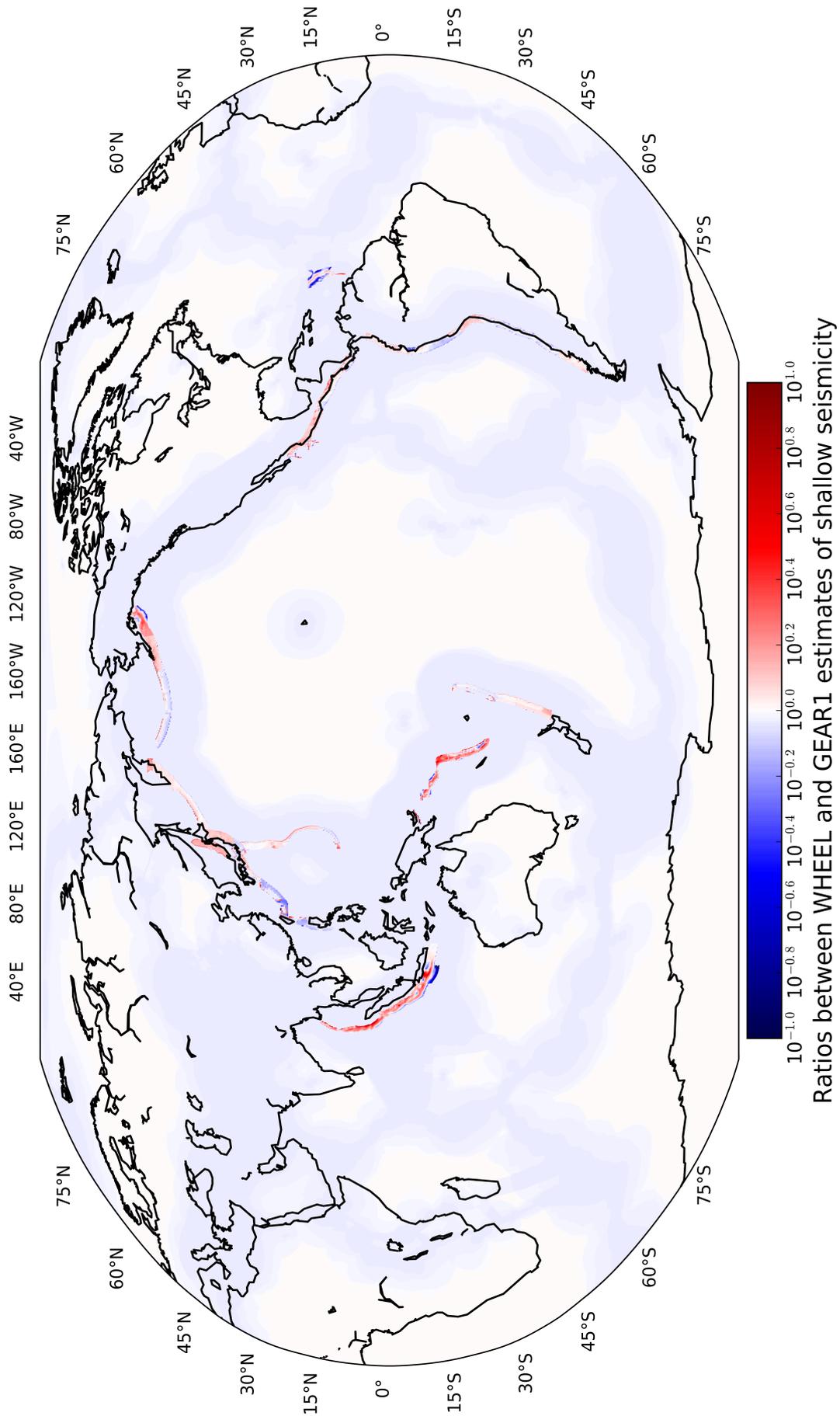
Abbildung 4.5: Variations in $m \geq 5.95$ seismicity rates forecasted by the WHEEL and GEAR1 earthquake models. In red, WHEEL estimates a larger number of earthquakes than GEAR1. In blue, oppositely, GEAR1 computes larger seismicity rates than WHEEL.

4.3 Earthquake Model Evaluations

SHIFT_GSRM2f, KJSS, GEAR1, TEAM and WHEEL provide scientific hypotheses about when and where earthquakes may occur, and how big they might be (see Table 4.2). Nonetheless, the veracity of these hypotheses can only be either confirmed or rejected after prospective forecast experiments. In this study, we evaluate the initial consistency of these global seismicity forecasts with observations recorded during the 1 January 2014-31 December 2019 pseudo-prospective test period. Moreover, we statistically compare the performance of each participant global earthquake approach with GEAR1, our selected brenchmarck seismicity model. To do so, we divide the global study region into spatio-magnitude bins with increments of 0.1 units in longitude, latitude and magnitude. Within each bin, we specify expected numbers of earthquakes during the evaluation period, which are assumed to be Poisson distributed (Schorlemmer et al., 2007; Zechar and Rhoades, 2010). Below, we provide a brief overview of the consistency and comparative tests implemented to describe the forecasting capacities of contestant global seismicity models.

Tabelle 4.2: Summary of seismicity models described in this doctoral dissertation.

Model	Nature	Coverage	Based on	Reference
SHIFT_GSRM	Tectonic	Global	GSRM and CMT catalog	Bird et al. (2010)
SHIFT_GSRM2f	Tectonic	Global	GSRM2.1 and CMT catalog	Bird and Kreemer (2015)
KJSS	Seismicity	Global	CMT catalog	Kagan and Jackson (2011)
GEAR1	Hybrid	Global	KJSS and SHIFT_GSRM2f	Bird et al. (2015)
SMERF	Tectonic	Subduction zones	GSRM2.1 and CMT catalog	Chapter 3
SMERF2	Tectonic	Subduction zones	GSRM2.1 and CMT catalog	Chapter 4
TEAM	Tectonic	Global	SMERF2 and SHIFT_GSRM2f	Chapter 4
WHEEL	Hybrid	Global	TEAM and KJSS	Chapter 4

4.3.1 Comparative tests

We test the relative performance of forecast pairs by measuring the rate-corrected information gain per earthquake of one forecast over another (Rhoades et al., 2011). For each global forecast pair, we apply the Student's paired T-test. This comparative test is based on the null hypothesis that two forecasts perform similarly, and the alternate hypothesis that one forecast significantly outperforms the other. As a result, one forecast is considered more informative than the other if the mean information gain significantly differs from the scaled difference in forecasted earthquake numbers between two forecasts. The Student's paired T-test requires the assumption that the information gain scores at observed earthquake locations are normally distributed, which is not always the case. For this reason, we also apply the non-parametric W-test, evaluating the median information gain per earthquake rather than the mean. This test only requires that the information gain distribution is symmetric, and increases in power with increasing numbers of observed earthquakes (Rhoades et al., 2011).

4.3.2 Consistency tests

Using a suite of likelihood consistency tests, we assess the consistency of forecasted and observed seismicity during the experiment's evaluation period. These tests are based on the likelihood of observed seismicity patterns, given forecasted earthquake numbers. A forecast's log-likelihood score is a metric, based on the Poisson distribution, used to evaluate the consistency of forecasted seismicity patterns with observed earthquakes (Schorlemmer et al., 2010). Greater log-likelihood scores indicate greater consistency, corresponding to a higher probability of the forecast generating a seismicity distribution similar to observations. This in turn implies a greater ability of the seismicity model to forecast earthquakes.

Consistency with observed seismicity, expressed as the log-likelihood score, can be decomposed into three dimensions: number of earthquakes, magnitude and spatial distributions. So, we apply tests of consistency for each of these dimensions (N-test, M-test, and S-test, respectively), which are directly derived from the likelihood, or L-test (Zechar and Rhoades, 2010). A forecast's log-likelihood score is most impacted by the number of earthquakes. Therefore, we also apply the conditional likelihood (CL) test, which provides information about a forecast's spatial and magnitude distribution while removing information regarding the total number of earthquakes (Werner et al., 2011).

4.4 Results and Discussion

T-test results show that GEAR1 and WHEEL are the most informative seismicity forecasts during the 2014–2019 pseudo-prospective testing period, as they obtain the highest information gain scores among all participant earthquake-rate forecasts (see Fig. 4.6 and Table 4.3). In agreement with Strader et al. (2018), the outperformance of these hybrid seismicity models over their individual forecast parent components indicate that the combination of instrumentally-recorded seismicity and geodetic strain data is suitable for long-term earthquake modeling. As formerly explained, the multiplicative log-linear blend of seismicity forecasts is based on the maximization of the Kagan (2009) I_1 sucess information score during the 2005–2012 retrospective period. Thus, comparative T-test results support the selection of the preferred GEAR1 model over a similar test period, suggest its stability over time as described by Bird (2018), and serve as new evidence to promote the creation and development of global ensemble seismicity models, as concluded by Marzocchi et al. (2012) and Marzocchi and Taroni (2014).

Comparative T-test results also exhibit that KJSS exceeds the forecasting skills of TEAM and SHIFT_GSRM2f during the 6-year pseudoprospective assessment period. Moreover, we observe that TEAM and SHIFT_GSRM2f are equally informative during the same testing period, as the difference between their information gain scores is statistically insignificant. Nonetheless, these outcomes might vary with time, because TEAM and SHIFT_GSRM2f are expected to have longer-term effects on earthquake forecasting than KJSS due to the incorporation of interseismic strain data into their formulations. Accordingly, we recommend further prospective evaluations to observe the forecasting potential of these global geodetic-based earthquake models.

Tabelle 4.3: Pseudo-prospective N-, M- S- and CL-test results for participant global seismicity models. The N-test metrics δ_1 and δ_2 describe the probabilities of observating at least and at most the actual number of earthquakes, respectively. If $\delta_1 < 0.025$, the forecast underpredicts observed seismicity; if $\delta_2 < 0.025$, the forecast overestimates actual seismicity; if $\delta_1 > 0.025$ and $\delta_2 > 0.025$ the forecast is consistent with the observations. The magnitude κ, spatial ζ, and magnitude-spatial ξ statistics provide the percentage of simulated seismicity catalogs, derived from each earthquake model, with a lower log-likelihood score than the score computed for the observed earthquake catalog. At a 0.05 confidence level, if these metrics are greater than 0.025, the forecasted seismicity distributions are consistent with actual earthquake activity. Thus, we indicate in bold values inconsistencies of earthquake-rate forecasts with the observations. Finally, we display the information gain score IG and confidence intervals (in brackets) over GEAR1, obtained for each contestant earthquake model during the 2014–2019 pseudo-prospective test period.

Model	δ_1	δ_2	κ	ζ	ξ	IG (Over GEAR1)
GEAR1	0.96	0.04	0.55	1.00	1.00	0.00
WHEEL	0.96	0.04	0.57	0.98	1.00	0.01 [-0.02 – 0.04]
KJSS	0.96	0.04	0.50	0.98	1.00	-0.14 [-0.22 – -0.09]
SHIFT2F_GSRM	0.96	0.05	0.59	**0.00**	**0.00**	-0.45 [-0.55 – -0.32]
TEAM	0.96	0.05	0.62	**0.00**	**0.00**	-0.44 [-0.56 – -0.31]
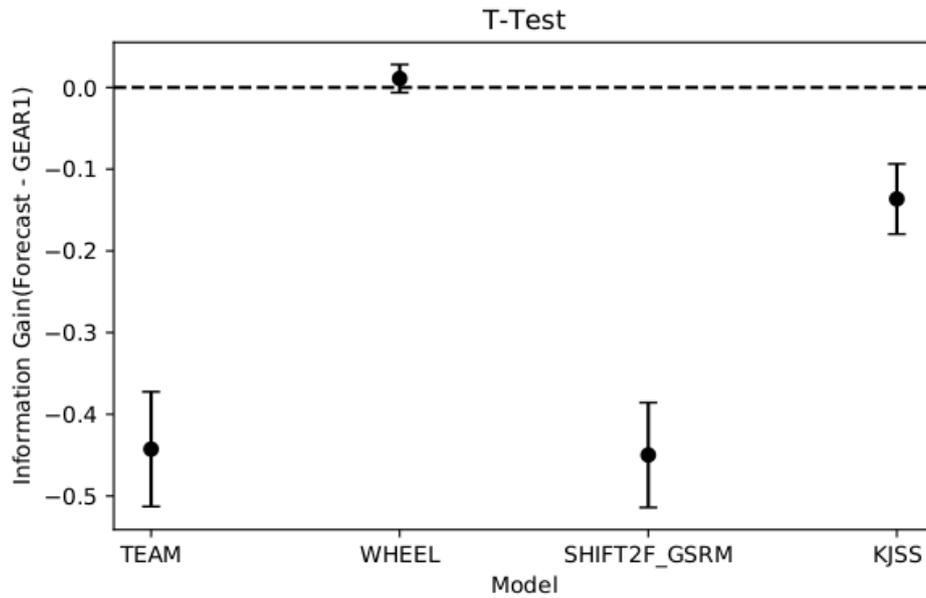
Abbildung 4.6: Comparative T-test results for contestant global earthquake-rate models during the 2014–2019 pseudo-prospective evaluation period. We present the mean information gain per earthquake as circles, and the 95% confidence interval with vertical lines. GEAR1 and WHEEL are the most informative models among all participant earthquake-rate approaches, as both rank with the highest information scores. Although not displayed, W-test results corroborate all T-tests.

N-test results show that all global seismicity forecasts are consistent with earthquakes recorded during the 2014–2019 test period (see Fig. 4.7). This is unsurprising because all contestant global forecasts were normalized to compute the global annual seismicity rate provided by the 1977–2013 Global CMT catalog. However, it is remarkable that such an observed global earthquake rate is systematically near the minimum bound of the number Poisson distribution with 95% confidence range computed for each forecast. Thus, we interpret these results to be due to either a measurable decrease of global earthquake activity during 2014–2019 in comparison with the 1977–2013 period, or the effect of including the sequences of megathrust events, such as the 2004 $m_w = 9.1$ Sumatra-Andaman, the 2010 $m_w = 8.8$ Maule and the 2011 Tohoku-Oki $m_w = 9.1$ earthquakes in the calibration datasets of these earthquake-rate models.

Similar to the N-tests, all contestant seismicity forecasts simultaneously pass the magnitude M-test during the pseudo-prospective evaluation period (see Fig. 4.7). These results are due to the fact that all models use unions of tapered Gutenberg-Richter functions to distribute their computed seismic moment rates among rates of forecasted seismicity. Distinctively, only the spatial earthquake distributions forecasted by WHEEL, GEAR1 and KJSS are consistent with observed seismicity during the testing period at the 0.05 significance level (see Fig. 4.7). Both tectonic models, TEAM and SHIFT_GSRM2f, forecast spatial seismicity patterns inconsistent with observations during the evaluation time frame. As formerly discussed in Chapters 2 and 3, SHIFT_GSRM2f spatial inconsistencies during the 2015–2017 period are assumed to potentially be due to initial underestimations of earthquake numbers in most subduction zones. Accordingly, TEAM attempts to compensate for this issue by concentrating more of the forecasted seismicity density within subduction margins by applying localized geomechanical parameters, such as the dip angle, seismogenic depth and coupling coefficient to these regions. Nevertheless, both TEAM and SHIFT_GSRM2f are limited in their abilities to globally forecast spatial earthquake patterns, suggesting that S-test results are, in fact, due to effects of spatial undersmoothing.
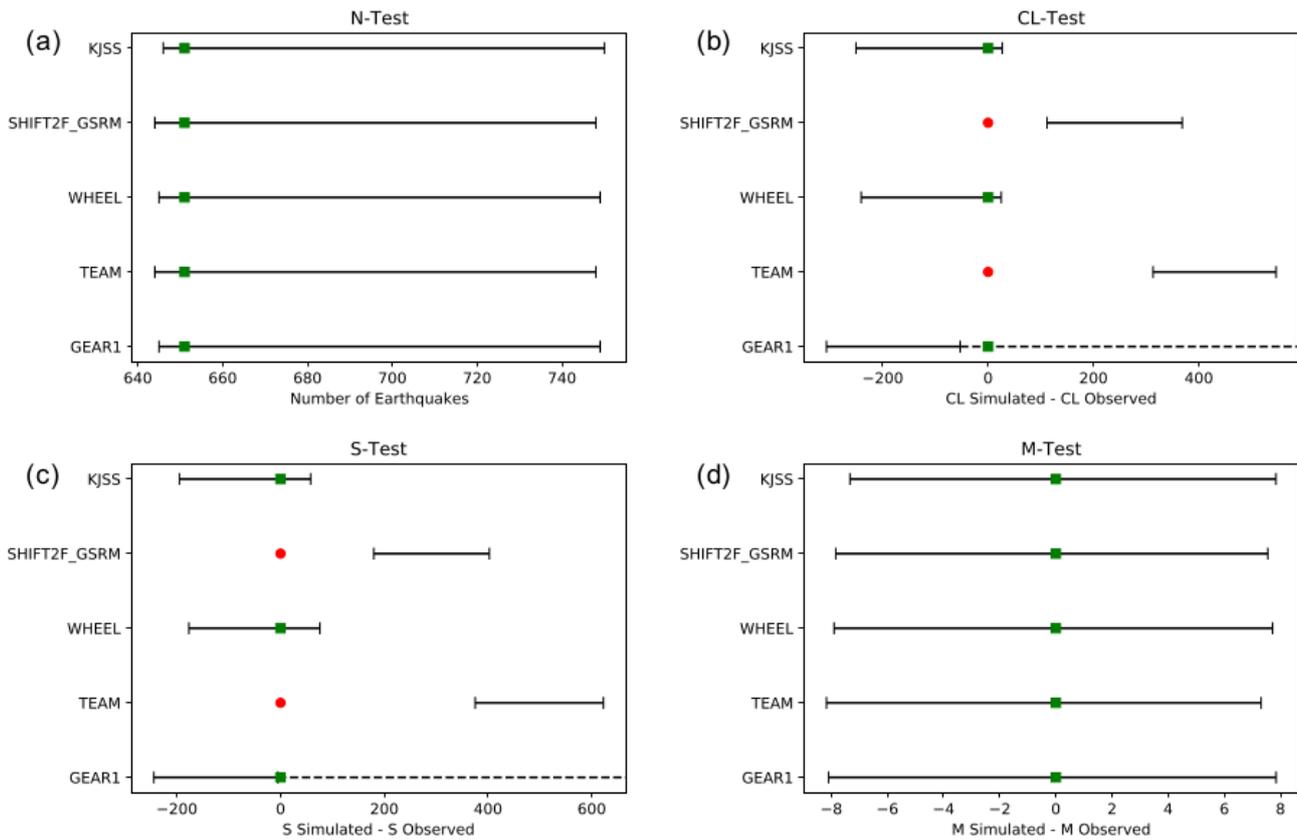
Abbildung 4.7: Consistency evaluation results for participant global earthquake-rate models during the 1 January 2014-31 December 2019 pseudo-prospective testing period. We show (a) number N-, (b) conditional likelihood CL-, (c) spatial S-, and (d) magnitude M-test results. Green squares indicate that earthquake forecasts are consistent with the observations. In contrast, red circles denote that models forecast earthquake patterns inconsistent with actual seismicity. Horizontal lines stand for the 95% confidence intervals. Dashed lines represent the one-sided form of the test distribution. Thus, seismicity models with a quantile score greater than the distribution upper limit are not rejected. For the N-test, the x-axis exhibits the number of earthquakes. For all other evaluations, the x-axis displays log-likelihood scores.

To give an answer to this rising hypothesis, we generate concentration plots for both tectonic forecasts, showing their normalized cumulative observed and forecasted seismicity distributions (see Fig. 4.8). For both models, the forecasted seismicity distribution is shifted to the right of the observed earthquake distribution, with a greater difference observed for TEAM. This indicates that seismicity forecasted by TEAM and SHIFT_GSRM2f is too localized along plate boundaries compared to KJSS, GEAR1 and WHEEL. As a result, if an earthquake occurs just outside of a subduction zone boundary, the forecasted rate will abruptly decrease and cause a large drop in the S-likelihood score.

Further pseudo-prospective analysis shows that the number and spatial test results for GEAR1 and WHEEL tend to remain stable over 2-year intervals, overlapping during the 2014–2019 testing period (see Table 4.4). Interestingly, these forecasts overestimate global earthquake activity during the 1 January 2016-31 December 2017 period, despite the occurrence of intermediate-size tectonic events, such as the 2016 m_{w} = 7.9 Papua New Guinea, the 2016 m_{w} = 7.8 Solomon Islands and the 2017 m_{w} = 8.2 Mexico earthquakes. Nevertheless, we interpret these results to be due to temporal fluctuations in earthquake activity, which should be addressed in future research for time-dependent seismicity modeling.
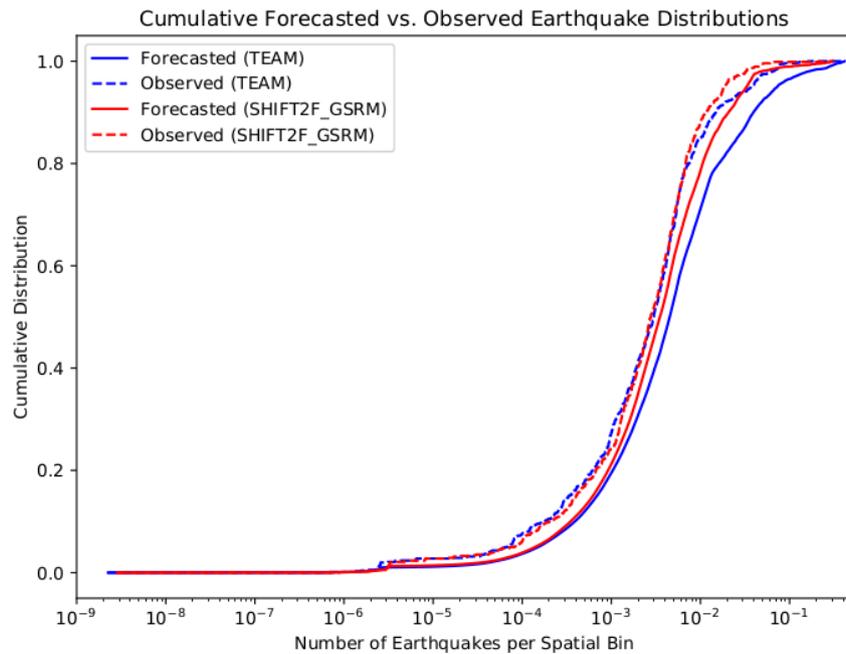
Abbildung 4.8: Concentration plots displaying the normalized cumulative distribution of observed seismicity (dashed lines), in comparison to that of the forecasted earthquake activity (solid lines), according to TEAM (blue) and SHIFT_GSRM2f (red).

Tabelle 4.4: Pseudo-prospective N- and M-test results for GEAR1 and WHEEL over 2-yr intervals, overlapping during the 2014–2019 period. These results show that N-test results are unstable during the 1 January 2016-31 December 2017 time frame, as GEAR1 and WHEEL overestimate rates of global seismicity.

Evaluation start date	δ_1 (GEAR1)	δ_2 (GEAR1)	ζ (GEAR1)	δ_1 (WHEEL)	δ_2 (WHEEL)	ζ (WHEEL)
01-01-2014	0.56	0.47	1.00	0.56	0.47	0.99
01-01-2015	0.89	0.12	1.00	0.89	0.12	0.98
01-01-2016	0.99	**0.01**	0.69	0.99	**0.01**	0.40
01-01-2017	1.00	**0.01**	0.29	1.00	**0.01**	0.08
01-01-2018	0.73	0.29	0.98	0.73	0.29	0.91

4.5 Conclusion and Prospects

In this study, we describe the construction and pseudo-prospective evaluation of two updated global seismicity models aimed at providing mean estimates of earthquake activity for seismic hazard assessment. These approaches are based on SMERF2, an earthquake-rate model for subduction zones constrained by interseismic strain measurements and earthquake-catalog information. TEAM, the first of them, is an ensembled tectonic seismicity model resulting from the combination of SMERF2 estimates in subduction margins with SHIFT_GSRM2f computations in non-SMERF2 interplate regions. Similar to GEAR1, WHEEL is a hybrid seismicity model created from the optimized multiplicative blend of the tectonic parent TEAM and the KJSS seismicity forecast.

GEAR1 and its individual forecast components use instrumentally-recorded seismicity during the 1977–2013 period to adjust and retrospectively evaluate their forecasted rates of earthquake activity. For fair comparison, we utilize the same training dataset to calibrate TEAM and WHEEL. Moreover, we employ actual earthquakes recorded during the 1 January 2014-31 December 2019 period to pseudo-prospectively test the forecasting skills of all global earthquake models, and statistically compare their performances with respect to GEAR1, our benchmark global seismicity model. As a result, GEAR1 and WHEEL are the most informative seismicity

forecasts during the pseudo-prospective test period, as they obtain the highest information gain scores among all contestant earthquake forecasts (see Figure 4.6 and Table 4.3). Interestingly, the difference between their information gain scores is statistically so insignificant, that the two models can be considered equally informative. Hence, we strongly recommend further prospective evaluations of all contestant global seismicity models to more accurately describe their capacities to forecast longer-term earthquake activity.

Pseudo-prospective S-test results indicate that TEAM and SHIFT_GSRM2f inappropriately characterize global spatial seismicity patterns during the 2014–2019 period (see Fig. 4.7). Surprisingly, both tectonic forecasts obtain almost the same spatial likelihood score (-4858.75 for TEAM and -4857.09 for SHIFT_GSRM2f) during the testing period, despite substantial variations among their spatial earthquake distributions (see Fig. 4.9). Thus, these findings reject the working hypothesis that SHIFT_GSRM2f failed the S-test evaluation during the 2015–2017 period due to initial SHIFT_GSRM2f underestimations of subduction-zone seismicity (Bird and Kreemer, 2015). Instead, spatial model inconsistencies seem to be related to effects of undersmoothing, according to concentration plots of normalized cumulative observed and forecasted seismicity distributions displayed for TEAM and SHIFT_GSRM2f (see Fig. 4.8). This possible explanation is in accordance with the positive impact of earthquake-catalog smoothing on spatial seismicity distributions forecasted by hybrid earthquake-rate models, reported by Akinci et al. (2018).

TEAM exceeds the spatial forecasting abilities of SHIFT_GSRM2f in subduction areas where both forecasted and observed earthquake rate densities are high (e.g. the western portion of the Aleutian arc, the coastline of central Chile, and the New-Hebrides and Bismarck-Solomon subduction zones; see Fig. 4.9). In contrast, SHIFT_GSRM2f better characterizes spatial patterns than TEAM in interplate regions with low forecasted seismicity rates, hosting relatively few earthquakes during the testing period (e.g. the Andaman-Sumatra subduction margin, the middle American trench along southern Mexico, and the eastern sector of the Aleutian arc). We expect these results to converge into a more stable solution in the future, as geodetic strain rates are thought to have longer-term effects on earthquake forecasting than six years. As a result, we will investigate the time dependence of these global seismicity models to more accurately describe short- and long-term earthquake patterns. Moreover, we will create and test a new global ensemble earthquake model, capturing the major advantages of TEAM, SHIFT_GSRM2f and KJSS for enhanced seismicity forecasting. Finally, we will submit TEAM and WHEEL to the CSEP testing centre for further prospective evaluation.

CSEP is currently transitioning into a new phase of earthquake forecast experiments designed to evaluate the forecasting abilities of seismicity models (Savran, Schorlemmer and Werner, personal communication, 2020). The pseudo-prospective evaluation of global earthquake-rate models presented in this Chapter is the first testing exercise of this under-construction project. Previously, CSEP has been able to identify the most informative seismicity models among more than 400 participant earthquake approaches (Michael and Werner, 2018; Schorlemmer et al., 2018). In this manner, seismicity models like GEAR1 and WHEEL are proven to indeed possess the forecasting potential to reliably characterize seismicity patterns worldwide, despite some skepticism.
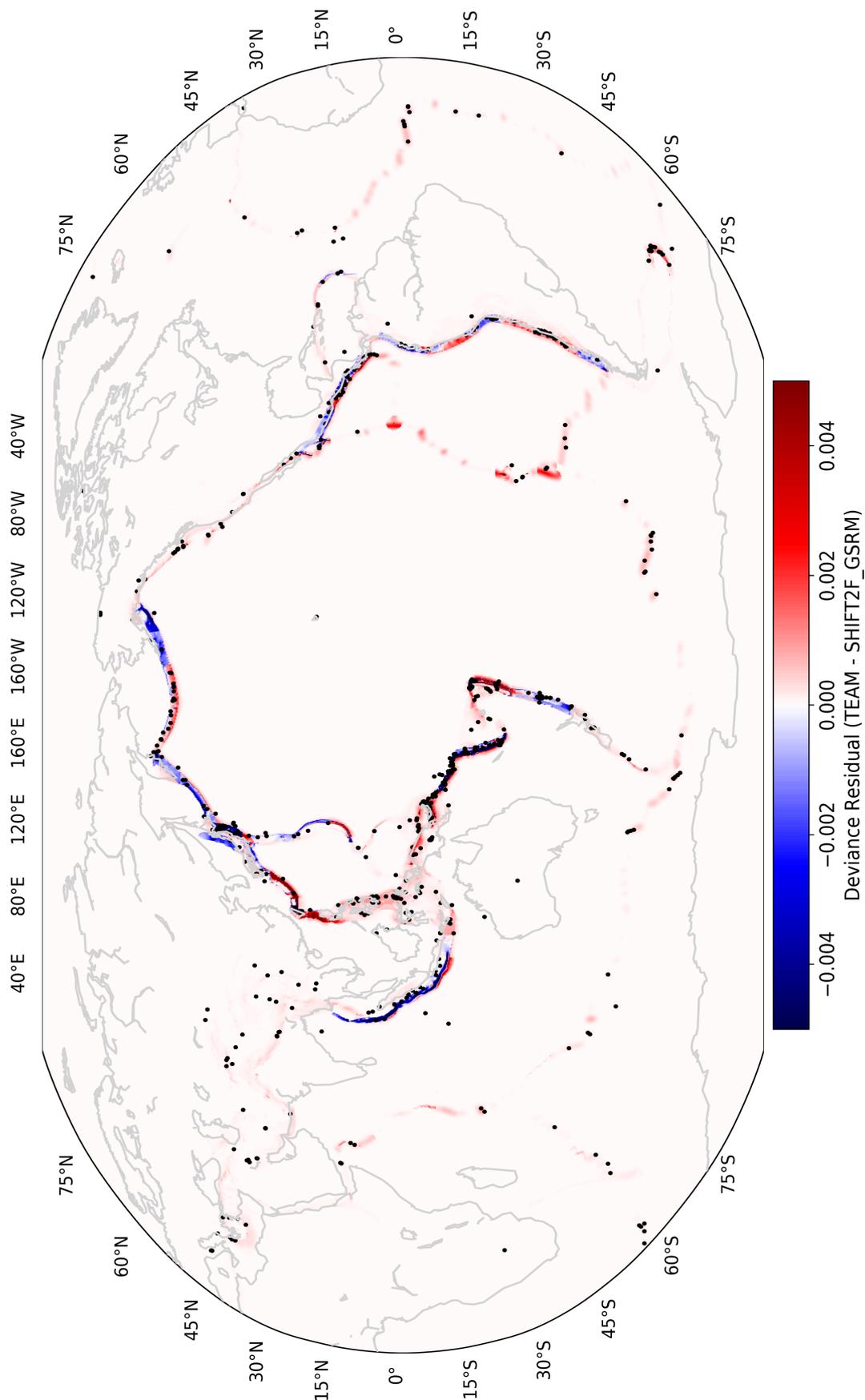
Abbildung 4.9: Residuals between spatial log-likelihood scores forecasted by TEAM and SHIFT_GSRM2f during the pseudo-prospective test period. In red areas, ratios between spatial likelihood scores obtained by these models are positive, indicating that TEAM is spatially more informative than SHIFT_GSRM2f. In contrast, blue zones denote locations where the spatial distribution forecasted by SHIFT_GSRM2f is more similar to the observations than the spatial distribution computed by TEAM. We include as black circles $d \leq 70$ km, $m_{\mathrm{w}} \geq$ 5.95 earthquakes reported from 1 January 2014 to 31 December 2019 in the global CMT catalog.

Acknowledgments

Kapitel 5

Results and Discussion

1. Prospective test results show that GEAR1 is the most informative global earthquake-rate forecast during the evaluation period 2015–2017 evaluation period.

According to T- and W-test results, GEAR1 significantly outperformed both SHIFT_GSRM2f and KJSS during the 1 October 2015-7 September 2017 evaluation period (see Fig. 2.2 and Table 2.3). The forecast's superior performance to its individual parent components shows that the combination of interseismic strain rates with earthquake-catalog information is suitable for stationary seismicity modeling. These results, with the corroboration of retrospective testing, also support the selection of the preferred GEAR1 model. Nonetheless, instability in number test results during retrospective two-year intervals (see Table 2.2) indicates that a longer testing period, at least equal to the GEAR1 calibration period, is necessary to better understand how effectively the optimal GEAR1 model constrains mid- to long-term seismicity patterns.

The total earthquake number computed by all participant seismicity forecasts is consistent with the observations during the prospective evaluation period. This is unsurprising, because all models were normalized to forecast total earthquake rates equal to the 1977-2003 global CMT rate at the $m_\mathrm{w} \geq 5.767$ threshold. Moreover, all contestant forecasts simultaneously passed the magnitude test during the prospective testing period. These results are due to the fact that all models use unions of Gutenberg-Richter functions to distribute their estimates of seismic moment among rates of predicted seismicity. Interestingly, only the spatial earthquake distributions forecasted by GEAR1 and KJSS are consistent with observed earthquakes. Both the SHIFT_GSRM and SHIFT_GSRM2f geodetic models failed the spatial test, despite the high spatial resolution of interseismic-strain data offered by GSRM and GSRM2.1. One reason for the difference in performance is that forecasted seismicity in SHIFT_GSRM2f is highly concentrated along plate boundaries compared to KJSS. This effect is indicated by large differences in log-likelihood scores for earthquakes near plate boundaries but located just beyond regions of elevated forecasted seismicity. Concentration plots for these forecasts (Fig. 2.4) display cumulative observed and forecasted seismicity distributions. The forecasted earthquake distribution for SHIFT_GSRM2f is shifted slightly to the right of the observed seismicity distribution, suggesting that seismicity in SHIFT_GSRM2f is too localizedälong tectonic plate boundaries compared to KJSS.

Inconsistencies in spatial seismicity patterns between the SHIFT_GSRM or SHIFT_GSRM2f model and observed seismicity may also be caused by inaccurate or low-resolution physical input parameters used to convert strain rates to seismic moment. Bird et al. (2015) indicate that the SHIFT_GSRM and SHIFT_GSRM2f models tend to underpredict seismicity within subduction zones prior to applying empirical calibration factors to account for the effect of the dip angle on the geometric factor. Thus, further testing of individual regions and the subsequent development and prospective evaluation of updated SHIFT_GSRM and GEAR models will indicate the extent to which forecast performance is sensitive to parameters used to convert strain rates to estimates of earthquake activity.

2 Retrospective and pseudo-prospective test results for the 1977–2013 and 2014–2018 testing periods show that the regionalized SMERF earthquake model considerably improves initial SHIFT_GSRM and SHIFT_GSRM2f estimates of seismicity in subduction margins.

Based on the iterative method implemented to construct SMERF and SMERF2, I found that estimates of corner magnitude and interseismic coupling, rather than subduction dip angles, play the major roles in reducing raw-SHIFT_GSRM and SHIFT_GSRM2f interface-earthquake underestimations (see Figs. 3.3 and 3.4, and Tables 3.3 and 3.4). Particularly, forecasted rates of subduction-zone seismicity increase by a factor of three when increasing the corner magnitude by a factor of one (Gutenberg and Richter, 1944; Strader et al., 2018). Therefore, magnitude distributions of great earthquakes calculated by Kagan and Jackson (2016) constitute the main scientific basis, on which both SMERF and SMERF2 are based. Kagan and Jackson (2016) computed estimates of corner magnitude by using the moment conservation principle that equates the seismic moment rate with the tectonic moment rate inferred from geology and geodesy. This approach is consistent with the physics-based nature of SMERF and SMERF2. As a result, estimates of corner magnitude control the uncertainties associated with the calibration of seismic coupling coefficients c_H. In this regard, I observe that SMERF2 coupling coefficients are in overall agreement with estimates of seismic locking reported in other studies. This is particularly interesting, because I calculated mean values of interseismic locking by balancing geodetic and seismic moment rates (see Eq. 3.14), in comparison with other authors who estimated coupling by calculating the ratio between the observed seismic moment rate and the rate obtained from plate tectonic velocities (earthquake-based c_E), or inverting interseismic strain accumulation rates of the upper plate (geodetic-based c_G), exclusively (see Table 5.1). Along the Tohoku subduction zone, for instances, Suwa et al. (2006), Hashimoto et al. (2009) and Loveless and Meade (2010) computed an average coupling c_G = 0.59–0.65, based on kinematic inversions of GPS interseismic velocities. Similarly, Bürgmann et al. (2005) and Scholz and Campos (2012) estimated a mean coupling coefficient c_G = 0.48–0.67 at a depth range of 10 to 50 km within the Kamchatka subduction margin. Thus, these values together are consistent with the SMERF2 coupling coefficient c_H = 0.33–0.66, calculated for the entire Japan-Kamchatka trench.

Discrepancies between estimates of seismic coupling along subduction zone interfaces have formerly been discussed in the literature (e.g. Pacheco et al., 1993; Heuret et al., 2011; Scholz and Campos, 2012). Thus, I also detect inconsistencies in a few subduction areas like Sumatra-Andaman, where the high degree of locking described by Chlieh et al. (2008) significantly differs from the hybrid coupling coefficient c_H = 0.12–0.20 obtained in this research. In general, these differences primarily stem from the scarce sampling of large earthquakes providing actual guidance on corner magnitudes for coupling calibration, the sensitivity of coupling calculations to input parameters, and the interpolation of inland deformation rates to oceanic floors. Hence, I argue that more stable solutions for seismic coupling will be obtained in the future with the acquisition of more earthquake-catalog information, and the analysis of oceanic-based interseismic strain data. Computations of interseismic coupling suggest that large earthquakes are expected to occur in subduction zones, such as New-Hebrides, Andean S. America (Peru) and Bismarck-Solomon Is. (New-Britain), as the total available seismic moment budget has not yet been used. Moreover, these results serve as first insights into geodetic-based coupling coefficients within some oceanic-oceanic subduction interfaces like Kermadec-Tonga-Samoa, Bismarck-Solomon Islands and Guam-Japan. Furthermore, these outcomes support the hypothesis that subduction margins are tectonically so distintive that specific regional seismicity parameters are required to better understand their unique earthquake patterns. According to sucessful retrospective and pseudo-prospective test results, SMERF holds the initial capacity to properly forecast

subduction-zone seismicity. Nonetheless, further prospective evaluations are needed to more objectively describe the forecasting abilites of this regionalized interface-seismicity model.

Tabelle 5.1: Estimates of interseismic coupling for 14 Kagan and Jackson (2016) subduction interfaces, obtained from earthquake-catalog data (c_E), GPS measurements (c_G) and the combination of both datasets (c_H).

Subduction zone	c_E	c_G	c_H	References
Kermadec-Tonga-Samoa	**0.16**[*] [0.01 - 0.66]		**0.12** [0.10 - 0.14]	*Peterson and Seno (1984) *Pacheco et al. (1993) *Scholz and Campos (1995) *Heuret et al. (2011)
New Hebrides Is.	**0.11**[*] [0.08 - 0.14]		**0.82** [0.56 - 1.00]	*Heuret et al. (2011)
Bismarck-Solomon Is.	**0.16**[*] [0.11 - 0.20]		**0.56** [0.38 - 0.83]	*Pacheco et al. (1993) *Heuret et al. (2011)
Sunda Arc	**0.01**[*] [0.00 - 0.02]	**0.45**[**] [0.20 - 0.70]	**0.16** [0.11 - 0.22]	*Peterson and Seno (1984) *Pacheco et al. (1993) *Heuret et al. (2011) **Hanifa et al. (2014) **Koulali et al. (2017)
Andaman Is.-Sumatra	**0.44**[*] [0.00 - 0.98]	**1.00**[**]	**0.16** [0.12 - 0.21]	*Peterson and Seno (1984) *Pacheco et al. (1993) **Chlieh et al. (2008) *Heuret et al. (2011) *Scholz and Campos (2012)
Taiwan	**0.01**[*] [0.00 - 0.20]	**0.40**[**]	**0.21** [0.15 - 0.30]	*McCann et al. (1979) *Heuret et al. (2011) **Hsu et al. (2012)
S.E. Japan-Ryukyu Is.	**0.11**[*] [0.04 - 0.22]	**0.56**[**] [0.00 - 1.00]	**0.26** [0.18 - 0.37]	*Peterson and Seno (1984) **Ito et al. (1999) *Pacheco et al. (1993) **Sagiya and Thatcher (1999) **Mazzotti et al. (2000) *Heuret et al. (2011)
Japan-Kamchatka	**0.67**[*] [0.50 - 1.00]	**0.60**[**] [0.48 - 0.67]	**0.47** [0.33 - 0.66]	*Peterson and Seno (1984) *Pacheco et al. (1993) **Suwa et al. (2006) **Hashimoto et al. (2009) **Loveless and Meade (2010) *Heuret et al. (2011)
Guam-Japan	**0.02**[*] [0.01 - 0.05]		**0.10** [0.09 - 0.12]	*Peterson and Seno (1984) *Pacheco et al. (1993) *Heuret et al. (2011)
Alaska-Aleutian Arc	**0.61**[*] [0.10 - 1.00]	**0.52**[**] [0.00 - 0.96]	**0.36** [0.26 - 0.50]	**Savage et al. (1998) **Fournier and Freymueller (2007) **Cross and Freymueller (2007) **Cross and Freymueller (2008) *Heuret et al. (2011) *Scholz and Campos (2012)
Mexico-Guatemala	**0.49**[*] [0.05 - 1.00]	**0.39**[**] [0.20 - 1.00]	**0.47** [0.32 - 0.70]	**Azúa et al. (2002) **Kostoglodov et al. (2003) **Franco et al. (2005) *Heuret et al. (2011) **Franco et al. (2012) *Scholz and Campos (2012)
Central America	**0.30**[*] [0.03 - 0.58]	**0.29**[**] [0.20 - 0.40]	**0.21** [0.15 - 0.30]	**White et al. (2003) **Norabuena et al. (2004) **Correa-Mora et al. (2009) *Heuret et al. (2011) **Scholz and Campos (2012)
Caribbean Loop	**0.08**[*]	**0.26**[**] [0.00 - 0.55]	**0.17** [0.11 - 0.24]	**Manaker et al. (2008) *Heuret et al. (2011)
Andean S. America	**0.47**[*] [0.02 - 1.00]	**0.81**[**] [0.41 - 1.00]	**0.79** [0.52 - 1.00]	**Chlieh et al. (2004) **Moreno et al. (2010) *Heuret et al. (2011) **Chlieh et al. (2011) **Métois et al. (2012) **Métois et al. (2013)

3 The impact of the new subduction-zone seismicity model remains unmeasurable on a global scale after six years of pseudo-prospective evaluation.

WHEEL and GEAR1 are global ensemble seismicity models, aimed at combining the advantages of their earthquake forecast constituents. Whereas their tectonic forecast components compute seismicity rates along plate boundaries, their seismicity forecast parent estimates earthquake activity in the surrounding intraplate areas, where geodetic strain rates are assumed to be zero according to GSRM2.1. As a result, these models are the most informative global seismicity approaches during the 2014–2019 pseudo-prospective testing period, as both rank with the highest information scores among all tested earthquake-rate models (see Fig. 4.6 and Table 4.3). Interestingly, the difference between their information gain scores is statistically so insignificant, that the two models can be considered equally informative. In agreement with Strader et al. (2018), the outperformance of these hybrid seismicity models over their individual forecast parent components indicate that the combination of instrumentally-recorded seismicity with geodetic strain data is suitable for stationary earthquake forecasting. As explained in previous chapters, the multiplicative log-linear blend of seismicity forecasts is based on the maximization of the Kagan (2009) I_1 sucess information score during the 2005–2012 retrospective period. Thus, comparative T-test results support the selection of the preferred GEAR1 model during a similar test period, suggest its stability over time as described by Bird (2018), and serve as new evidence to promote the creation and development of global ensemble seismicity models, as concluded by Marzocchi et al. (2012) and Marzocchi and Taroni (2014).

Comparative T-test results also exhibit that the seismicity forecast component KJSS exceeds the forecasting skills of the geodetic forecast parents TEAM and SHIFT_GSRM2f during the 6-year pseudo-prospective testing period. This is mainly due to the fact that both geodetic earthquake models are limited in their abilities to forecast spatial seismicity patterns globally (see Fig. 4.7). I created SMERF2 to reduce number, and potentially spatial, inconsistencies of its predecessor SHIFT_GSRM2f model with the observations. Specifically, I redistributed earthquake rates within subduction margins through the application of regional seismicity parameters. However, spatial earthquake inconsistencies seem to be primarily due to either a spatial undersmoothing of too localizedforecasted seismicity along tectonic plate boundaries (see Fig. 4.8) or limitations in testing procedures ranking the spatial distributions of forecasted seismicity. In this regard, it is well known that the Poisson assumption is questionable in presence of clusters of seismicity. Lombardi and Marzocchi (2010), for instances, showed that the use of the Poisson hypothesis can lead to a greater rejection of correctmodels, as their derived earthquake forecasts could possess a significantly higher variability than that of a Poisson process, invalidating one of the main assumptions that stands behind the CSEP/RELM experiments. Therefore, future seismicity models might need to incorporate more physics-based testing procedures and smoothing kernels to better understand the clustering nature of earthquakes in and outside plate boundary regions. Currently, global geodetic-based earthquake models like SHIFT_GSRM2f and TEAM are restricted to provide high-resolution estimates of seismicity along tectonic plate boundaries, given the global distribution of interseismic strain data (i.e. Kreemer et al., 2014). Nevertheless, these areas, comprising only 15% of the Earth's surface (Gordon and Stein, 1992), host the most prominent seismic activity observed worldwide (Bird and Kagan, 2004; Ekström et al., 2012; Bilek and Lay, 2018). Thus, the study of physics-derived forecast metrics (e.g. Parsons, 2020) might be useful to more accurately describe the impact of SMERF2 on the development and testing of global ensemble seismicity models.

Kapitel 6

Conclusions and Perspectives

In this doctoral dissertation, I developed and tested a stationary seismicity model for subduction zones, referred to as the Subduction Megathrust Earthquake Rate Forecast (SMERF2). This model is based on the premise that subduction margins are tectonically so distinctive, that regional seismicity parameters are required to adequately describe their particular earthquake patterns. Thus, SMERF2 uses geodetic strain rates, earthquake-catalog data and regional geomechanical parameters, such as subduction dip angles, seismogenic thicknesses and seismic coupling coefficients, to estimate mean rates of seismic moment and interface seismicity. As a result, SMERF2 captures two independent factors necessary for long-term earthquake modeling, namely interseismic crustal strain accumulation and sudden lithospheric stress release. Among others, the long-term balance of geodetic and seismic moment rates allows to derive analytical solutions for interseismic coupling and earthquake activity (Eqs. 3.13 and 3.14) that depend on physical quantities rather than empirical correlations. Therefore, I consider these equations to be scientific contributions to the development of physics-based seismicity models.

I integrated SMERF2 computations in subduction zones with estimates of its predecessor SHIFT_GSRM2f in non-subduction margins to create the TEAM seismicity model. In addition, I combined this new tectonic earthquake approach with the Smoothed-Seismicity model of Kagan and Jackson (2011) to produce the global WHEEL hybrid seismicity model. After six years of pseudo-prospective evaluation, comparative T-test results show that both WHEEL and the Global Earthquake Activity Rate (GEAR1; Bird et al., 2015) are the most informative global earthquake-rate models until today. Interestingly, variations in their information gain scores are statistically insignificant, despite the use of different earthquake parameters with which these models characterize subduction-zone seismicity. Therefore, I strongly recommend further prospective evaluations of all contestant global seismicity models to more accurately describe their capacities to forecast longer-term earthquake activity. Furthermore, I speculate that at least one decade of prospective testing is required to assess the performance of global ensemble earthquake models presented in this dissertation.

SMERF2, TEAM and WHEEL are certainly subject to be improved. A potential improvement is the creation and testing of an updated hybrid seismicity model, capturing the forecasting benefits of these earthquake-rate approaches. In this regard, I think that the development of ensemble earthquake-rate models will experience a noticable boost in the nearby future after evidence of consistent test results with the observations (e.g. Marzocchi et al., 2012; Taroni et al., 2014; Akinci et al., 2018). An alternative source of improvement is the combination of Eq. 3.13 describing earthquake rates in subduction zones with the analytical solution of Petruccelli et al. (2019) for the b value, to translatethe observational Gutenberg-Richter frequency-magnitude distribution into a more physics-based earthquake-rate model. On one hand, Eq. 3.13 predicts that earthquake rates in subduction zones linearly correlate with the plate convergence velocity. On the other hand, Petruccelli et al. (2019) demonstrated that estimates of b systematically vary with the stress regime or the style of tectonic faulting. Thus, I believe that the Gutenberg-

Richter distribution may be expressed, among others, in terms of the relative plate velocity controlling the tectonic deformation, and the faulting geometry regulating the transmission of stress through contact area between plates (Lallemand et al., 2005). A third approach worth exploring is the time dependence of earthquake activity, connecting the long-term background seismicity with short-term aftershock sequences. Tentatively, I would like to confront Eq. 3.13 with the total earthquake rate $\lambda(t)$ of the Epistemic Type Aftershock Sequence (ETAS; Ogata, 1988; Zhuang et al., 2002) model to explore possible correlations between interface earthquake parameters and ETAS parametric values within subduction margins (e.g. Zhang et al., 2020).

Long-term seismicity forecasting is a fundamental component of earthquake science that form the basis for earthquake hazard quantification. Modern Probabilistic Seismic Hazard Analyses (PSHA) require that the seismogenic source be characterized in terms of a finite rupture in order to translate a long-term stationary earthquake forecast into a PSHA calculation (e.g. Field et al., 2015; Bayona Viveros et al., 2017). In practice, this requires that for each earthquake source, a set of potential rupture mechanisms and a set of hypocentral depths must be defined (Weatherill and Pagani, 2011). On a global scale, this information can be obtained from available catalogs of seismicity, which currently provide hypocentral locations and focal mechanisms of thousands of earthquakes (e.g. Ekström et al., 2012; Storchak et al., 2013). In this manner, the rupture area can be computed for each magnitude in the magnitude-frequency distribution using a magnitude-area scaling relation (e.g. Wells and Coppersmith, 1994). Moreover, an upper and lower seismogenic thickness can be set to constrain the seismogenic thickness through which the finite rupture propagates. With this in mind, I conclude that global seismicity models described in this dissertation might be useful in improving global hazard modeling, as they possess key ingredients for hazard estimation, such as point-earthquake sources that could be converted into finite rupture areas (i.e. geodetic-strain data and smoothed-seismicity information) and geomechanical parameters for subduction zones. Thus, I finally hope that the development and testing of these global ensemble seismicity models could serve as a solid scientific basis to more accurately assess seismic hazard, and hence more efficiently mitigate earthquake risk worldwide.

Literaturverzeichnis

Akinci, A., Moschetti, M., and Taroni, M. (2018). Ensemble smoothed seismicity models for the new Italian probabilistic seismic hazard map. *Seismological Research Letters*, 4(89):1277–1287.

Allen, C. R. (1976). Geological criteria for evaluating seismicity. *Developments in Geotechnical Engineering*, 15:31–69.

Anderson, J. G. and Biasi, G. P. (2016). What is the Basic Assumption for Probabilistic Seismic Hazard Assessment? *Seismological Research Letters*, 87(2A):323–326.

Azúa, B., DeMets, C., and Masterlark, T. (2002). Strong interseismic coupling, fault afterslip, and viscoelastic flow before and after the Oct. 9, 1995 Colima-Jalisco earthquake: Continuous GPS measurements from Colima, Mexico. *Geophysical Research Letters*, 8(29):122–1.

Bayona Viveros, J., Suárez Reynoso, G., and Ordaz Schroeder, M. (2017). A probabilistic seismic hazard assessment of the Trans-Mexican Volcanic Belt, Mexico based on historical and instrumentally recorded seismicity. *Geofísica internacional*, 1(56):87–101.

Bayona Viveros, J., von Specht, S., Strader, A., Hainzl, S., Cotton, F., and Schorlemmer, D. (2019). A Regionalized Seismicity Model for Subduction Zones Based on Geodetic Strain Rates, Geomechanical Parameters, and Earthquake-Catalog Data. *Bulletin of the Seismological Society of America*, 5(109):2036–2049.

Bilek, S. and Lay, T. (1999). Rigidity variations with depth along interplate megathrust faults in subduction zones. *Nature*, 400(6743):443–446.

Bilek, S. and Lay, T. (2018). Subduction zone megathrust earthquakes. *Geosphere*, 14(4):1468–1500.

Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, 4(3).

Bird, P. (2018). Ranking some global forecasts with the Kagan information score. *Seismological Research Letters*, 89(4):1272–1276.

Bird, P., Jackson, D. D., Kagan, Y. Y., Kreemer, C., and Stein, R. S. (2015). GEAR1: A global earthquake activity rate model constructed from geodetic strain rates and smoothed seismicity. *Bulletin of the Seismological Society of America*, 105(5):2538–2554.

Bird, P., Kagan, Y., Jackson, D., Schoenberg, F., and Werner, M. (2009). Linear and Nonlinear Relations between Relative Plate Velocity and SeismicityLinear and Nonlinear Relations between Relative Plate Velocity and Seismicity. *Bulletin of the Seismological Society of America*, 99(6):3097–3113.

Bird, P. and Kagan, Y. Y. (2004). Plate-tectonic analysis of shallow seismicity: Apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings. *Bulletin of the Seismological Society of America*, 94(6):2380–2399.

Bird, P. and Kreemer, C. (2015). Revised tectonic forecast of global shallow seismicity based on version 2.1 of the Global Strain Rate Map. *Bulletin of the Seismological Society of America*, 105(1):152–166.

Bird, P., Kreemer, C., and Holt, W. E. (2010). A long-term forecast of shallow seismicity based on the Global Strain Rate Map. *Seismological Research Letters*, 81(2):184–194.

Bird, P. and Liu, Z. (2007). Seismic hazard inferred from tectonics: California. *Seismological Research Letters*, 78(1):37–48.

Brune, J. N. (1968). Seismic moment, seismicity, and rate of slip along major fault zones. *Journal of Geophysical Research*, 73(2):777–784.

Bürgmann, R., Kogan, M., Steblov, G., Hilley, G., Levin, V., and Apel, E. (2005). Interseismic coupling and asperity distribution along the Kamchatka subduction zone. *Journal of Geophysical Research: Solid Earth*, 110(B7).

Burridge, R. and Knopoff, L. (1964). Body force equivalents for seismic dislocations. *Bulletin of the Seismological Society of America*, 54(6A):1875–1888.

Chlieh, M., Avouac, J., Sieh, K., Natawidjaja, D., and Galetzka, J. (2008). Heterogeneous coupling of the Sumatran megathrust constrained by geodetic and paleogeodetic measurements. *Journal of Geophysical Research: Solid Earth*, 113(B5).

Chlieh, M., De Chabalier, J., Ruegg, J., Armijo, R., Dmowska, R., Campos, J., and Feigl, K. (2004). Crustal deformation and fault slip during the seismic cycle in the North Chile subduction zone, from GPS and InSAR observations. *Geophysical Journal International*, 158(2):695–711.

Chlieh, M., Perfettini, H., Tavera, H., Avouac, J., Remy, D., Nocquet, J., Rolandone, F., Bondoux, F., Gabalda, G., and Bonvalot, S. (2011). Interseismic coupling and seismic potential along the Central Andes subduction zone. *Journal of Geophysical Research: Solid Earth*, 116(B12).

Christensen, D. H. and Beck, S. L. (1994). The rupture process and tectonic implications of the great 1964 Prince William Sound earthquake. *Pure and Applied Geophysics*, 142(1):29–53.

Cornell, C. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5):1583–1606.

Correa-Mora, F., DeMets, C., Alvarado, D., Turner, H., Mattioli, G., Hernandez, D., Pullinger, C., Rodriguez, M., and Tenorio, C. (2009). GPS-derived coupling estimates for the Central America subduction zone and volcanic arc faults: El Salvador, Honduras and Nicaragua. *Geophysical Journal International*, 179(3):1279–1291.

Cross, R. and Freymueller, J. (2007). Plate coupling variation and block translation in the Andreanof segment of the Aleutian arc determined by subduction zone modeling using GPS data. *Geophysical Research Letters*, 34(6).

Cross, R. and Freymueller, J. (2008). Evidence for and implications of a Bering plate based on geodetic measurements from the Aleutians and western Alaska. *Journal of Geophysical Research: Solid Earth*, 113(B7).

DeMets, C., Gordon, R. G., and Argus, D. F. (2010). Geologically current plate motions. *Geophysical Journal International*, 181(1):1–80.

Dziewonski, A. M., Chou, T. A., and Woodhouse, J. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research: Solid Earth*, 86(B4):2825–2852.

Eberhard, D. A. J., Zechar, J. D., and Wiemer, S. (2012). A Prospective Earthquake Forecast Experiment in the Western Pacific. *Geophysical Journal International*, 190(3):1579–1592.

Ekström, G., Nettles, M., and Dziewoński, A. M. (2012). The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200:1–9.

Engdahl, E. R., van der Hilst, R., and Buland, R. (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bulletin of the Seismological Society of America*, 88(3):722–743.

Esteva, L. (1968). Basis for the formulation of decisions on seismic design. *PhD Thesis (in Spanish)*. Universidad Nacional Autónoma de México.

Field, E. H. (2007). Overview of the working group for the development of regional earthquake likelihood models (RELM). *Seismological Research Letters*, 78(1):7–16.

Field, E. H., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., Johnson, K. M., Jordan, T. H., Madden, C., Michael, A. J., Milner, K. R., Page, M. T., Parsons, T., Powers, P. M., Shaw, B. E., Thatcher, W. R., Weldon II, R. J., and Zeng, Y. (2015). Long-Term Time-Dependent Probabilities for the Third Uniform California Earthquake Rupture Forecast (UCERF3). *Bulletin of the Seismological Society of America*, 105(2A):511–543.

Field, E. H., Dawson, T. E., Felzer, K. R., Frankel, A. D., Gupta, V.and Jordan, T. H., Parsons, T., Petersen, M. D., Stein, R. S., Weldon, R. J., and Wills, C. J. (2009). Uniform California earthquake rupture forecast, version 2 (UCERF 2). *Bulletin of the Seismological Society of America*, 99(4):2053–2107.

Fournier, T. and Freymueller, J. (2007). Transition from locked to creeping subduction in the Shumagin region, Alaska. *Geophysical Research Letters*, 6(34).

Franco, A., Lasserre, C., Lyon-Caen, H., Kostoglodov, V., Molina, E., Guzman-Speziale, M., Monterosso, D., Robles, V., Figueroa, C., Amaya, W., and Barrier, E. (2012). Fault kinematics in northern Central America and coupling along the subduction interface of the Cocos Plate, from GPS data in Chiapas (Mexico), Guatemala and El Salvador. *Geophysical Journal International*, 3(189):1223–1236.

Franco, S., Kostoglodov, V., VC, M., Manea, M., and JA, S. (2005). Propagation of the 2001-2002 silent earthquake and interplate coupling in the Oaxaca subduction zone, Mexico. *Earth, planets and space*, 10(57):973–985.

Geller, R. J. (1997). Earthquake prediction: a critical review. *Geophysical Journal International*, 131(3):425–450.

Giardini, D. (2014). Mapping Europe's Seismic Hazard. *EOS*, 95(29):261–268.

Giardini, D., Grünthal, G., Shedlock, K. M., and Zhang, P. (1999). The GSHAP global seismic hazard map. *Annals of Geophysics*, 42(6).

Gordon, R. G. and Stein, S. (1992). Global tectonics and space geodesy. *Science*, 256(5055):333–342.

Gutenberg, B. and Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4):185–188.

Hanifa, N., Sagiya, T., Kimata, F., Efendi, J., Abidin, H., and Meilano, I. (2014). Interplate coupling model off the southwestern coast of Java, Indonesia, based on continuous GPS data in 2008-2010. *Earth and Planetary Science Letters*, 401:159–171.

Hanks, T. C. and Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research: Solid Earth*, 84(B5):2348–2350.

Hashimoto, C., Noda, A., Sagiya, T., and Matsu'ura, M. (2009). Interplate seismogenic zones along the Kuril-Japan trench inferred from GPS data inversion. *Nature Geoscience*, 2(2):141–144.

Haskell, N. (1963). Radiation pattern of Rayleigh waves from a fault of arbitrary dip and direction of motion in a homogeneous medium. *Bulletin of the Seismological Society of America*, 53(3):619–642.

Hayes, G., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., and Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. *Science*, 362(6410):58–61.

Hayes, G. P., Wald, D. J., and Johnson, R. L. (2012). Slab1. 0: A three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research: Solid Earth*, 117(B1).

Helmstetter, A., Kagan, Y. Y., and Jackson, D. D. (2007). High-Resolution Time-Independent Grid-Based Forecast for $M \geq 5$ Earthquakes in California. *Seismological Research Letters*, 78(1):78–86.

Heuret, A., Lallemand, S., Funiciello, F., Piromallo, C., and Faccenna, C. (2011). Physical characteristics of subduction interface type seismogenic zones revisited. *Geochemistry, Geophysics, Geosystems*, 12(1).

Hirata, T. (1989). A correlation between the b value and the fractal dimension of earthquakes. *Journal of Geophysical Research: Solid Earth*, 94(B6):7507–7514.

Hsu, Y., Yu, S., Song, T., and Bacolcol, T. (2012). Plate coupling along the Manila subduction zone between Taiwan and northern Luzon. *Journal of Asian Earth Sciences*, 51:98–108.

Ide, S. (2013). The proportionality between relative plate velocity and seismicity in subduction zones. *Nature Geoscience*, 6(9):780–784.

Ishimoto, M. (1936). Observations of earthquakes registered with the microseismograph constructed recently. *Bull. Earthquake Res. Inst. Univ. Tokyo*, 17:443–478.

Ito, T., Yoshioka, S., and Miyazaki, S. (1999). Interplate coupling in southwest Japan deduced from inversion analysis of GPS data. *Physics of the Earth and Planetary Interiors*, 115(1):17–34.

Jackson, D. D. and Kagan, Y. Y. (1999). Testable earthquake forecasts for 1999. *Seismological Research Letters*, 70(4):393–403.

Ji, S., Sun, S., Wang, Q., and Marcotte, D. (2010). Lamé parameters of common rocks in the Earth's crust and upper mantle. *Journal of Geophysical Research: Solid Earth*, 115(B6).

Jordan, T. (2006). Earthquake predictability, brick by brick. *Seismological Research Letters*, 77(1):3–6.

Kagan, Y. Y. (1999). Universality of the seismic moment-frequency relation. *In Seismicity patterns, their statistical significance and physical meaning*, pages 537–573.

Kagan, Y. Y. (2009). Testing long-term earthquake forecasts: likelihood methods and error diagrams. *Geophysical Journal International*, 177(2):532–542.

Kagan, Y. Y., Bird, P., and Jackson, D. D. (2010). Earthquake patterns in diverse tectonic zones of the globe. *Pure and applied geophysics*, 167(6-7):721–741.

Kagan, Y. Y. and Jackson, D. D. (1994). Long-Term Probabilistic Forecasting of Earthquakes. *Journal of Geophysical Research: Solid Earth*, 99(B7):13685–13700.

Kagan, Y. Y. and Jackson, D. D. (2000). Probabilistic Forecasting of Earthquakes. *Geophysical Journal International*, 143(2):438–453.

Kagan, Y. Y. and Jackson, D. D. (2011). Global earthquake forecasts. *Geophysical Journal International*, 184(2):759–776.

Kagan, Y. Y. and Jackson, D. D. (2016). Earthquake rate and magnitude distributions of great earthquakes for use in global forecasts. *Geophysical Journal International*, 206(1):630–643.

Kennett, B. L., Engdahl, E. R., and Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1):108–124.

Kostoglodov, V., Singh, S., Santiago, J., Franco, S., Larson, K., Lowry, A., and Bilham, R. (2003). A large silent earthquake in the Guerrero seismic gap, Mexico. *Geophysical Research Letters*, 30(15).

Kostrov, V. V. (1974). Seismic moment and energy of earthquakes, and seismic flow of rock. *Izv. Acad. Sci. USSR Phys. Solid Earth, Engl. Transl*, 1:23–44.

Koulali, A., McClusky, S., Susilo, S., Leonard, Y., Cummins, P., Tregoning, P., Meilano, I., Efendi, J., and Wijanarto, A. (2017). The kinematics of crustal deformation in Java from GPS observations: Implications for fault slip partitioning. *Earth and Planetary Science Letters*, 458:69–79.

Kreemer, C., Blewitt, G., and Klein, E. C. (2014). A geodetic plate motion and Global Strain Rate Model. *Geochemistry, Geophysics, Geosystems*, 15(10):3849–3889.

Kreemer, C., Holt, W. E., and Haines, A. J. (2002). The global moment rate distribution within plate boundary zones. *Plate boundary zones*, 30:173–190.

Kreemer, C., Holt, W. E., and Haines, A. J. (2003). An integrated global model of present-day plate motions and plate boundary deformation. *Geophysical Journal International*, 154(1):8–34.

Lallemand, S., Heuret, A., and Boutelier, D. (2005). On the relationships between slab dip, back-arc stress, upper plate absolute motion, and crustal nature in subduction zones. *Geochemistry, Geophysics, Geosystems*, 6(9).

Langridge, R. M., Persaud, M., Zúñiga, F. R., de Jesús Aguirre-Díaz, G., Villamor, P., and Lacan, P. (2013). Preliminary paleoseismic results from the Pastores fault and its role in the seismic hazard of the Acambay graben, Trans-Mexican Volcanic Belt, Mexico. *Revista Mexicana de Ciencias Geológicas*, 30(3):463–481.

Lombardi, A. and Marzocchi, W. (2010). The assumption of Poisson seismic-rate variability in CSEP/RELM experiments. *Bulletin of the Seismological Society of America*, 100(5A):463–481.

Lomnitz, C. (2004). Major earthquakes of Chile: a historical survey, 1535-1960. *Seismological Research Letters*, 75(3):368–378.

Loveless, J. and Meade, B. (2010). Geodetic imaging of plate motions, slip rates, and partitioning of deformation in Japan. *Journal of Geophysical Research: Solid Earth*, 115(B2).

Manaker, D., Calais, E., Freed, A., Ali, S., Przybylski, P., Mattioli, G., Jansma, P., Prépetit, C., and De Chabalier, J. (2008). Interseismic plate coupling and strain partitioning in the northeastern Caribbean. *Geophysical Journal International*, 174(3):889–903.

Marti, M., Stauffacher, M., and Wiemer, S. (2019). Difficulties in explaining complex issues with maps: evaluating seismic hazard communication-the Swiss case. *Natural Hazards and Earth System Sciences*, 19(12):2677–2700.

Marzocchi, W. and Taroni, M. (2014). Some thoughts on declustering in probabilistic seismic-hazard analysis. *Bulletin of the Seismological Society of America*, 4(104):1838–1845.

Marzocchi, W., Zechar, J. D., and Jordan, T. H. (2012). Bayesian Forecast Evaluation and Ensemble Earthquake Forecasting. *Bulletin of the Seismological Society of America*, 102(6):2574–2584.

Mazzotti, S., Le Pichon, X., Henry, P., and Miyazaki, S. (2000). Full interseismic locking of the Nankai and Japan-west Kurile subduction zones: An analysis of uniform elastic strain accumulation in Japan constrained by permanent GPS. *Bulletin of the Seismological Society of America*, B6(105):13159–13177.

McCaffrey, R. (2008). Global frequency of magnitude 9 earthquakes. *Geology*, 36(3):263–266.

McCann, W., Nishenko, S., Sykes, L., and Krause, J. (1979). Seismic gaps and plate tectonics: seismic potential for major boundaries. *In Earthquake Prediction and Seismicity Patterns*, pages 1082–1147.

Métois, M., Socquet, A., and Vigny, C. (2012). Interseismic coupling, segmentation and mechanical behavior of the central Chile subduction zone. *Journal of Geophysical Research: Solid Earth*, 117(B3).

Métois, M., Socquet, A., Vigny, C., Carrizo, D., Peyrat, S., Delorme, A., Maureira, E., Valderas-Bermejo, M., and Ortega, I. (2013). Revisiting the North Chile seismic gap segmentation using GPS-derived interseismic coupling. *Geophysical Journal International*, 194(3):1283–1294.

Michael, A. and Werner, M. (2018). Preface to the Focus Section on the Collaboratory for the Study of Earthquake Predictability (CSEP): New results and future directions. *Seismological Research Letters*, 89(4):1226–1228.

Minoura, K., Imamura, F., Sugawara, D., Kono, Y., and Iwashita, T. (2001). The 869 Jogan tsunami deposit and recurrence interval of large-scale tsunami on the Pacific coast of northeast Japan. *Journal of Natural Disaster Science*, 23(2):83–88.

Moreno, M., Rosenau, M., and Oncken, O., . (2010). Maule earthquake slip correlates with pre-seismic locking of Andean subduction zone. *Nature*, 467(7312):198–202.

MPS-Working-Group (2004). Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003. Rapporto conclusivo per il Dipartimento della Protezione Civile, INGV, Milano-Roma, April 2004 (MPS04), (in Italian). 23(2):65–70. http://zonesismiche.mi.ingv.it (last accessed February 2020).

Norabuena, E., Dixon, T., Schwartz, S., DeShon, H., Newman, A., Protti, M., Gonzalez, V., Dorman, L., Flueh, E., Lundgren, P., and Pollitz, F. (2004). Geodetic and seismic constraints on some seismogenic zone processes in Costa Rica. *Journal of Geophysical Research: Solid Earth*, 109(B11).

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, 83(401):9–27.

Oliver, J. and Murphy, L. (1971). WWNSS: Seismology's global network of observing stations. *Science*, 174(4006):254–261.

Ortuño, M., Corominas, O., Villamor, P., Zúñiga, R. F., Lacan, P., Aguirre-Díaz, G., Perea, H., Štěpančíková, P., and Ramírez-Herrera, M. T. (2019). Evidence of recent ruptures in the central faults of the Acambay Graben (central Mexico). *Geomorphology*, 326:17–37.

Pacheco, J. F. and Sykes, L. R. (1992). Seismic moment catalog of large shallow earthquakes, 1900 to 1989. *Bulletin of the Seismological Society of America*, 82(3):1306–1349.

Pacheco, J. F., Sykes, L. R., and Scholz, C. H. (1993). Nature of seismic coupling along simple plate boundaries of the subduction type. *Journal of Geophysical Research: Solid Earth*, 93(B8):14133–14159.

Parsons, T. (2020). On the use of receiver operating characteristic tests for evaluating spatial earthquake forecasts. *Geophysical Research Letters*, p.e2020GL088570.

Petersen, M. D., Moschetti, M. P., Powers, P. M., Mueller, C. S., Haller, K. M., Frankel, A. D., Zeng, Y., Rezaeian, S., Harmsen, S. C., Boyd, O. S., and Field, N. (2015). The 2014 United States national seismic hazard model. *Earthquake Spectra*, 31(S1):S1–S31.

Peterson, E. and Seno, T. (1984). Factors affecting seismic moment release rates in subduction zones. *Journal of Geophysical Research: Solid Earth*, 89(B12):10233–10248.

Petruccelli, A., Schorlemmer, D., Tormann, T., Rinaldi, A., Wiemer, S., Gasperini, P., and Vannucci, G. (2019). The influence of faulting style on the size-distribution of global earthquakes. *Earth and Planetary Science Letters*, 527:115791.

Rhoades, D. A., Christophersen, A., Gerstenberger, M. C., Liukis, M., Silva, F., Marzocchi, W., Werner, M. J., and Jordan, T. H. (2018). Highlights from the first ten years of the New Zealand earthquake forecast testing center. *Seismological Research Letters*, 89(4):1229–1237.

Rhoades, D. A. and Evison, F. F. (2006). The EEPAS forecasting model and the probability of moderate-to-large earthquakes in central Japan. *Tectonophysics*, 417(1-2):119–130Bird.

Rhoades, D. A., Schorlemmer, D., Gerstenberger, M. C., Christophersen, A., Zechar, J. D., and Imoto, M. (2011). Efficient Testing of Earthquake Forecasting Models. *Acta Geophysica*, 59(4):728–747.

Ritchie, H. and Roser, M. (2014). Natural Disasters. *Our World in Data.* https://ourworldindata.org/natural-disasters (last accessed March 2020).

Roser, M., Ritchie, H., and Ortiz-Ospina, E. (2020). World Population Growth. *Our World in Data.* https://ourworldindata.org/world-population-growth (last accessed March 2020).

Ruff, L. and Kanamori, H. (1980). Seismicity and the subduction process. *Physics of the Earth and Planetary interiors*, 23(3):240–252.

Sagiya, T. and Thatcher, W. (1999). Coseismic slip resolution along a plate boundary megathrust: The Nankai Trough, southwest Japan. *Journal of Geophysical Research: Solid Earth*, 104(B1):1111–1129.

Savage, J., Svarc, J., Prescott, W., and Gross, W. (1998). Deformation across the rupture zone of the 1964 Alaska earthquake, 1993âĂŞ1997. *Journal of Geophysical Research: Solid Earth*, 103(B9):21275–21283.

Savage, J. C. and Simpson, R. W. (1997). Surface strain accumulation and the seismic moment tensor. *Bulletin of the Seismological Society of America*, 87(5):1345–1353.

Scholz, C. H. and Campos, J. (1995). On the mechanism of seismic decoupling and back arc spreading at subduction zones. *Journal of Geophysical Research: Solid Earth*, 100(B11):22103–22115.

Scholz, C. H. and Campos, J. (2012). The seismic coupling of subduction zones revisited. *Journal of Geophysical Research: Solid Earth*, 117(B5).

Schorlemmer, D. and Gerstenberger, M. C. (2007). RELM testing center. *Seismological Research Letters*, 78(1):30–36.

Schorlemmer, D., Gerstenberger, M. C., Wiemer, S., Jackson, D. D., and Rhoades, D. A. (2007). Earthquake likelihood model testing. *Seismological Research Letters*, 78(1):17–29.

Schorlemmer, D., Werner, M. J., Marzocchi, W., Jordan, T. H., Ogata, Y., Jackson, D. D., Mak, S., Rhoades, D. A., Gerstenberger, M. C., Hirata, N., and Liukis, M. (2018). The collaboratory for the study of earthquake predictability: achievements and priorities. *Seismological Research Letters*, 89(4):1305–1313.

Schorlemmer, D., Wiemer, S., and Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature*, 437(7058):539–542.

Schorlemmer, D., Zechar, J. D., Werner, M. J., Field, E. H., Jackson, D. D., Jordan, T. H., and RELM-Working-Group (2010). First results of the regional earthquake likelihood models experiment. *Seismogenesis and Earthquake Forecasting: The Frank Evison Volume II*, pages 5–22. Springer, Basel.

Shen, Z. K., Jackson, D. D., and Kagan, Y. Y. (2007). Implications of geodetic strain rate for future earthquakes, with a five-year forecast of M5 earthquakes in southern California. *Seismological Research Letters*, 78(1):116–120.

Simons, M., Minson, S. E., Sladen, A., Ortega, F., Jiang, J., Owen, S. E., Meng, L., Ampuero, J. P., Wei, S., Chu, R., and Helmberger, D. V. (2011). The 2011 magnitude 9.0 Tohoku-Oki earthquake: Mosaicking the megathrust from seconds to centuries. *science*, 332(6036):1421–1425.

Smith, W. (1981). The b-value as an earthquake precursor. *Nature*, 289(5794):136–139.

Stein, S. and Okal, E. A. (2005). Speed and size of the Sumatra earthquake. *Nature*, 434(7033):581–582.

Storchak, D. A., Di Giacomo, D., Bondár, I., Engdahl, E. R., Harris, J., Lee, W. H., Villaseñor, A., and Bormann, P. (2013). Public release of the ISC-GEM global instrumental earthquake catalogue (1900-2009). *Seismological Research Letters*, 84(5):810–815.

Strader, A., Schneider, M., and Schorlemmer, D. (2017). Prospective and retrospective evaluation of five-year earthquake forecast models for California. *Geophysical Journal International*, 211(1):239–251.

Strader, A., Werner, M., Bayona, J., Maechling, P., Silva, F., Liukis, M., and Schorlemmer, D. (2018). Prospective evaluation of global earthquake forecast models: 2 yrs of observations provide preliminary support for merging smoothed seismicity with geodetic strain rates. *Seismological Research Letters*, 89(4):1262–1271.

Sunye-Puchol, I., Lacan, P., Ortuño, M., Villamor, P., Audin, L., Zúñiga, F., Langridge, R. M., Aguirre-Díaz, G. D. J., and Lawton, T. F. (2015). La falla San Mateo: nuevas evidencias paleosismológicas de fallamiento activo en el graben de Acambay, México. *Revista mexicana de ciencias geológicas (in Spanish)*, 32(3):361–376.

Suwa, Y., Miura, S., Hasegawa, A., Sato, T., and Tachibana, K. (2006). Interplate coupling beneath NE Japan inferred from three-dimensional displacement field. *Journal of Geophysical Research: Solid Earth*, 111(B4).

Taroni, M., Marzocchi, W., Schorlemmer, D., Werner, M. J., Wiemer, S., Zechar, J. D., Heiniger, L., and Euchner, F. (2018). Prospective CSEP Evaluation of 1-Day, 3-Month, and 5-Yr Earthquake Forecasts for Italy. *Seismological Research Letters*, 89(4):1251–1261.

Taroni, M., Zechar, J. D., and Marzocchi, W. (2014). Assessing annual global M 6+ seismicity forecasts. *Geophysical Journal International*, 196(1):422–431.

Tsuruoka, H., Hirata, N., Schorlemmer, D., Euchner, F., Nanjo, K. Z., and Jordan, T. H. (2012). CSEP Testing Center and the First Results of the Earthquake Forecast Testing Experiment in Japan. *Earth Planets Space*, 64(8):661–.

Vigny, C., Socquet, A., Peyrat, S., Ruegg, J. C., MÃľtois, M., Madariaga, R., Morvan, S., Lancieri, M., Lacassin, R., Campos, J., and Carrizo, D. (2011). The 2010 Mw 8.8 Maule megathrust earthquake of central Chile, monitored by GPS. *Science*, 332(6036):1417–1421.

Ward, S. N. (1994). A multidisciplinary approach to seismic hazard in southern California. *Bulletin of the Seismological Society of America*, 84(5):1293–1309.

Weatherill, G. and Pagani, M. (2011). From smoothed seismicity forecasts to probabilistic seismic hazard: insight and challenges from a global perspective. *Second European Conference on Earthquake Engineering and Seismology*, Istanbul.

Wells, D. and Coppersmith, K. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the seismological Society of America*, 84(4):974–1002.

Werner, M. J., Helmstetter, A., Jackson, D. D., and Kagan, Y. Y. (2011). High-resolution long-term and short-term earthquake forecasts for California. *Bulletin of the Seismological Society of America*, 101(4):1630–1648.

Werner, M. J., Zechar, J. D., Marzocchi, W., Wiemer, S., and the CSEP-Italy Working Group (2010). Retrospective Evaluation of the Five-Year and Ten-Year CSEP-Italy Earthquake Forecasts. *Annals of Geophysics*, 53(3):11–30.

White, S., Trenkamp, R., and Kellogg, J. (2003). Recent crustal deformation and the earthquake cycle along the Ecuador-Colombia subduction zone. *Earth and Planetary Science Letters*, 216(3):231–242.

Zechar, J. D., Schorlemmer, D., Liukis, M., Yu, J., Euchner, F.and Maechling, P. J., and Jordan, T. H. (2010). The Collaboratory for the Study of Earthquake Predictability perspective on computational earthquake science. *Concurrency and Computation: Practice and Experience*, 22(12):1836–1847.

Zechar, J. D.and Gerstenberger, M. C. and Rhoades, D. A. (2010). Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts. *Bulletin of the Seismological Society of America*, 100(3):1184–1195.

Zhang, L., Werner, M., and Goda, K. (2020). Variability of ETAS Parameters in Global Subduction Zones and Applications to Mainshock-Aftershock Hazard Assessment. *Bulletin of the Seismological Society of America*.

Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, 97(458):369–380.

Abbildungsverzeichnis

Tabellenverzeichnis