

Uncovering the relationship between subducting bathymetric ridges and volcanic chains with significant earthquakes using geophysical data mining

T. C. W. LANDGREBE AND R. D. MÜLLER

EarthByte Group, School of Geosciences, Madsen Building, The University of Sydney, NSW 2006, Australia.

E-mail: dietmar.muller@sydney.edu.au,

Tel: +61 2 9351 3625, Fax +61 2 9351 2442

Short running title: Uncovering relationships between subducting features and significant earthquakes

The subduction of bathymetric asperities has been linked with the location and rupture characteristics of significant (moment magnitude > 5.0) thrust earthquakes in several regions, but a global analysis of these relationships is lacking. We carry out a global investigation of well-defined subducting bathymetric features including ridges, fracture zones and seamounts and their spatial association with significant earthquakes. The NGDC global significant earthquake database has been filtered to include only those events occurring in the coupling zone between subducting and over-riding plates, which includes the bulk of megathrust earthquakes. A statistical methodology is used to compare spatial associations between subducting linear asperities and significant earthquakes. Randomly sampled coupling zone locations are used to establish sensitivity/specificity relationships as a function of proximity, ruling out random effects and establishing meaningful spatial interpretations. In a previous study we showed that there is a strong link between the largest earthquakes and proximity to fracture-zone subduction-zone intersections. Here this analysis is extended using a game-formulation methodology, rigorously demonstrating that significant subduction-related earthquakes are more likely to be generated at, or very close to, the intersections between subduction zones and fracture zones, compared to arbitrary locations along subduction zones. A weaker, broader effect is exhibited by subducting volcanic ridges/seamount chains. Fracture zone associations are attributed to the large, localised, linear and mechanically strong bathymetric

anomalies related to fracture zone ridges, leading to strong seismic coupling, whereas associations pertaining to volcanic ridges/seamount chains are explained by inherently broader, less elevated and poorly-defined characteristics, combined with structural weaknesses that may reduce the long-term coupling effect.

KEYWORDS: Bathymetric asperities; megathrust subduction earthquakes; fracture zones; aseismic ridges; seamount chains; data mining

INTRODUCTION

Subduction zones, and more specifically, coupling zones between down going and overriding tectonic plates, are associated with the majority of the largest megathrust earthquakes. Traditional hazard prediction methods have proven to be unsuccessful at identifying localities at risk of generating some of the most significant earthquakes on record (Stein *et al.* 2012), since digital earthquake catalogues do not allow reliable differentiation of regional risk levels if great earthquake cycles are up to an order of magnitude longer than the ~100 year time span covered by these catalogues (Sieh *et al.* 2008).

Lay & Kanamori (1981) developed a conceptual model in which major subduction zone earthquakes are driven by strong coupling between the down going and overriding plates, caused by the subduction of asperities. These are transverse (elongated) structures on the down going plate (at arbitrary angles), such as aseismic ridges, causing strong coupling at the plate interface. The effect of aseismic ridge and seamount subduction on seismic coupling and earthquake rupture behaviour and overriding plate deformation has been demonstrated at several localities (Gorbatov *et al.* 1997; Scholz & Small 1997; Kodaira *et al.* 2000; Bilek *et al.* 2003; von Huene & Ranero 2009), and is underpinned by a physical model for seismic rupture caused by seamounts (Cloos 1992) or other asperities (Ruff 1992). Other investigators used the distribution of basins, basement highs and associated gravity anomalies on the overriding plate as proxies for mapping the location of subduction zone asperities (Wells *et al.* 2003) without explicitly addressing their origin or the implication for seismic risk. A detailed study of the tectonic setting along the Japan Trench by Mochizuki *et al.* (2008) concluded that a large subducting seamount in the region is associated with weak interplate coupling. This observation casts doubt on the idea

that seamounts are the most obvious candidates for barriers, which locally inhibit faulting for long periods of time, leading to great earthquake supercycles, e.g. the large 869 AD Jogan earthquake (Minoura *et al.* 2001) was in the vicinity of the recent 2011 Tohoku-Oki great earthquake, over 1100 years later. Oceanic fracture zones represent another form of subducting asperities, but their effect on earthquake rupture has received much less attention than that of aseismic ridges and seamounts, with the exception of specific fracture zones offshore Chile–Peru (Robinson *et al.* 2006; Contreras-Reyes & Carrizo 2011), Kamchatka (Gorbatov *et al.* 1997), Alaska (Von Huene *et al.* 1999), Sumatra (Lange *et al.* 2010) and the Solomon Islands (Taylor *et al.* 2008). An opportunity for an alternative hazard prediction approach arises from links between bathymetric roughness/anomalies at the subduction trench and the occurrence of earthquakes.

Here we capitalise on the establishment of a number of global digital datasets to develop a quantitative link between some of these asperities and the generation of significant subduction earthquakes (significant earthquakes are defined to be those with moment magnitudes larger than 5.0). In particular, we investigate two different “classes” of asperities, both of which can be mapped reliably via geophysical methods, and have elongated/extensive topographic structure, suggesting a high likelihood of partial subduction for those features in close proximity to subduction zones. The first asperity class pertains to oceanic fracture zones, typically consisting of large linear features, exhibiting prominent bathymetric anomalies that can be involved in altering the coupling between overriding and down going plates. The second asperity class consists of volcanic ridges and seamount chains, which are often broader, more discontinuous and more irregular in shape.

Analysing the associations between the various global datasets in this study calls for statistical and data-mining methodologies that inherently account for variabilities and uncertainties. This allows for quantification of both the “sensitivity” of associations (i.e. how closely related the bathymetric asperities are to significant subduction-based earthquakes), and the “specificity” (i.e. how focused in space the associations are relative to the area occupied by the coupling zone). The study also includes an analysis of primary controlling factors that are typically related to subduction earthquakes, namely the convergence velocity, the age of the sea floor, and the dip angle of the subducting slab.

This study builds upon the methodology developed in Müller & Landgrebe (2012), where a significant relationship between the largest subduction-based earthquakes and fracture-zone subduction-zone intersections was found. The study showed that 13 of the 15 largest (moment magnitude $M_w \geq 8.6$) and half of the 50 largest ($M_w \geq 8.4$) coupling zone earthquakes were associated. In this study we focus on the proximity relationships with significant earthquakes in general to establish (in a rigorous fashion) whether significant earthquakes have a proclivity towards being generated in the vicinity of subducted asperities. Establishing such a bias on this larger collection of known subduction-earthquakes has broad implications on the role of subducting asperities, and improving hazard assessments that can make use of geophysical data in addition to standard techniques. This is appealing because evidence suggests that the slow periodicities of larger earthquakes implies they are poorly sampled in global earthquake catalogues, which are the basis for hazard prediction.

DATASET DESCRIPTIONS

The global analysis of subducting bathymetric ridges and volcanic chains involves the integration of several digital datasets, including earthquakes, subduction zones, lithosphere thickness, and seafloor asperities. A model of the coupling-zone between subducting and over-riding plates was used to compute associations using a large earthquake catalogue focused on thrust-type subduction. Intersections between these two datasets were used to select subduction-based earthquakes, while excluding others. The datasets and associated filtering followed the approach taken in Müller & Landgrebe (2012), with an overview provided in this paper.

Subducting bathymetric ridges and volcanic chains

A recent compilation of global fracture zones has been used for this study, as described by Matthews *et al.* (2011). As discussed in Müller & Landgrebe (2012), intersections with subduction zones have been identified semi-automatically *via* the following process:

- computation of the nearest distances between the digitised geometries to the subduction zones;
- storage of the intersection location; and
- manual verification involving a combination of bathymetry and gravity anomaly data to confirm bathymetric expressions visually. Most

intersections involve fracture zone locations within less than 100 km of a trench, taking into account that sediments on the down-going plate may partially obscure bathymetric expressions of fracture zones.

In total 59 locations were identified, depicted in Figure 1 (purple squares). Note that this study has intentionally chosen these large, well-defined features (as opposed to smaller, less well-defined cracks and fractures occurring along subducting margins), hypothesising that these large, long-lived features pose a predictable and substantial coupling effect at subduction zones. Bathymetric anomalies along subduction plate boundaries have been computed for 11 fracture zones (light dashed lines) associated with large earthquake events occurring in the vicinity (extracted from the ETOPO1 global relief model; Amante *et al.* 2009). The subducting ridges at these 11 locations have elevations above the surround seafloor from ~200–1200 m (see Müller & Landgrebe 2012). This illustrates that substantial bathymetric offsets are involved, providing a physical link to the effect demonstrated in this paper.

Volcanic chains and aseismic volcanic ridges have been compiled following Coffin & Eldholm (1994). As in the fracture-zone case, features on the seafloor in the proximity of subduction zones were classified to be either in the process of being subducted or not. Isolated seamounts do not form part of this dataset. A total of 14 subducting volcanic chains/ridges were identified, as shown in Figure 1 (green squares). In Appendix A, the various latitude–longitude pairs for the intersected fracture zones (Table A.1) and volcanic ridges/chains (Table A.2), respectively, are listed. The sparsity of this dataset is noteworthy when considering the statistical analyses presented later.

Collation of data and models for computing the associations

The study in Müller & Landgrebe (2012) details the various datasets used for this study, as well as how data has been filtered and interpreted. The primary dataset for the associations comprise a significant earthquake catalogue, originating from the NGDC global significant earthquakes database (NGDC/WDC 2011). The various datasets are used to investigate relationships with particular subduction zone target areas, comprising 5539 recorded earthquakes considered to be significant in magnitude. After a filtering step, 3684 samples were extracted, which were filtered further to retain only those earthquakes originating in the subduction coupling zone, reducing the dataset to 1073 samples.

Establishment of a global “coupling zone” is computed by intersecting models of the Lithosphere–Asthenosphere Boundary (LAB) with recent 3-dimensional representations of subduction. These coupling zones allow subduction-related earthquakes to be filtered and tested for spatial association with proximal fracture zones. As described in Müller & Landgrebe (2012), the LAB is defined using a combination of continental and oceanic models. The LAB model is then intersected with three-dimensional models of the down going subduction slabs. Constructing the overall coupling zone involved the integration of 5 global datasets and a statistical study. The derived coupling zones (filled yellow polygons), together with other pertinent datasets integrated for this study, are shown in Figure 1. We acknowledge that using lithospheric thickness to constrain the coupling zone may result in a suboptimally large zone, which may be improved upon using a more complex temperature–depth model as discussed in Oleskevich *et al.* (1999). Another dataset used in this study is a seafloor global age model, consisting of gridded ages of the ocean floor is used investigate age relationships in the vicinity of subduction trenches, based on the model by Müller *et al.* (2008).

Controlling factors due to interplate convergence velocities were analysed by using the GSRMP geodetic strain rate dataset from the global strain rate map project (Kreemer *et al.* 2003). These are used to estimate the convergence rates at subduction zones, required for assessing their role in the context of modulating the effect of subduction asperities on generating significant earthquakes.

METHODOLOGY

Associating subducting bathymetric features with significant earthquakes

The primary analysis presented in this paper investigates spatial relationships between filtered significant earthquakes in the vicinity of fracture zone and volcanic chain/ridge intersections within the coupling zone. The key difference with respect to the analysis in Müller & Landgrebe (2012) is that rather than investigating magnitude-sorted relationships, the coupling-zone filtered significant earthquakes are analysed as a whole. The intention is to establish whether subducting asperities play a significant role in these earthquakes being preferentially generated in their vicinity based on a larger dataset. In this way a robust assessment can be undertaken to distinguish whether there are only arbitrary associations, or whether there is in fact a controlling process. This in turn will be useful for further research in this area, such as

establishing the respective impact it may have on hazard predictions, and understanding the nature of these subduction phenomena.

The primary associations were computed as follows: the orientation of the bathymetric features was used to project their extensions into the nearby coupling zone, maintaining the same azimuth as in their closest geophysical expressions seaward of a given trench. The resultant intersection is bounded by the width of the coupling zone (see Figure 2 for an example region). This results in linear spatial features that serve as a reference for undertaking the analysis of associated earthquakes for a range of proximities. For each of these proximities, spatial buffers are formed around the intersection lines (called *target regions*), in which associated earthquakes can be filtered, and allowing the sensitivity of the association to be quantified by varying the buffer distances. The associations computed in this paper are thus computed adaptively according to 3-dimensional subducting slab models and lithosphere thickness.

The association measure used is called *recall*, which is simply the number of earthquakes (filtered by the coupling zone) within the target regions, normalised by all earthquake samples. The analysis is described formally as follows: the significant earthquakes dataset is denoted \mathbf{E} , consisting of a list of latitude (θ), longitude (λ) pairs such that $\mathbf{E} = [e_1, e_2, \dots, e_{N_e}]$ for a dataset size of N_e , and the i_{th} element of \mathbf{E} is denoted $e_i = (\theta(e_i), \lambda(e_i), M(e_i))$. Target locations are projected into the coupling zones by using the principal direction of the bathymetric feature (at the trench), extended into the coupling zone, resulting in lines traversing the coupling zone. The list of N_t projected target lines pairs is defined as:

$$\mathbf{L}_t = [L_1, L_2, \dots, L_N]$$

The recall measure involves computing the ratio of the number of earthquakes occurring within a specified region of interest (ROI), specified by a buffer distance d_{ROI} in kilometres around \mathbf{L}_t , calculated by summing the number of earthquakes associated within the ROI regions. This procedure involves creating a binary vector \mathbf{A} of length N , defined as $\mathbf{A} = [a_1, a_2, \dots, a_N]$. The i_{th} item in \mathbf{A} is determined as follows:

$$a_i = \begin{cases} 1 & \text{if } \sum_{j=1}^{N_t} F(e_{s(i)}, L_j, d_{ROI}, CZ) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where CZ is the coupling zone polygon geometry, and F considers the association within a thresh-holded buffer around L_j bounded by CZ :

$$F(e_{s(i)}, L_j, d_{ROI}, CZ) = \begin{cases} 1 & \text{if } e_{s(i)} \text{ inpolygon } G(L_j, CZ, d_{ROI}) \\ 0 & \text{otherwise} \end{cases}$$

The function $G(L_j, CZ, d_{ROI})$ creates a buffer polygon of width d_{ROI} km around L_j , clipped by CZ . The target earthquake $e_{s(i)}$ can then be tested for association *via* a standard point-in-polygon test, denoted *inpolygon*. Now that the binary vector \mathbf{A} has been computed, the *recall* score can be determined *via*:

$$recall = \frac{1}{N} \sum_{j=1}^N a_j$$

Assessing the arbitrary case

The locations of the subducting bathymetric feature intersections are widely distributed across subduction zones throughout the globe, and thus it is essential to ascertain the *significance* of the computed associations with earthquakes. A key aspect of the analysis is to ascertain whether the computed associations could occur arbitrarily (i.e. at random), or whether such an occurrence is unlikely.

The approach taken for determining the significance of the associations identified involves computing what would be expected if the associations were to occur arbitrarily, involving repeated samplings of arbitrary partitioning along the coupling zone, thus simulating virtual subduction-zone intersections, as presented in Müller & Landgrebe (2012). This procedure allows the associations identified here to be calibrated against the case in which the same association could occur at random. In order to generate these arbitrary data partitions, a first step involves computing a poly-line geometry along the axis of the estimated coupling zone, providing a reference for partitions to be generated perpendicular to the trench. In a second step the global coupling zone is partitioned into 20 km intervals, providing a set of 2634 adjacent partitions, limited spatially by the extent of the coupling zone (see an indicative example of the segmented coupling zone in Figure 2). In a third step, each partition element is buffered in order to form a spatial region in which to assess the existence of earthquakes during the association-calculation process. This is repeated a number of times for different buffer distances (constrained to fall within the coupling zone extent). During the statistical experimentation at a particular spatial geometry size, the arbitrary case is then computed by repeatedly selecting a total of 59 of these

partitions at random (in the fracture zone case), followed by the same evaluation procedure, and similarly 14 partitions in the subducting volcanic ridge/seamount chain case. The process iterates 200 times, with the resulting statistical distribution used to compare against the target association computed. The number of runs was determined by analysing the stability of the statistical performance measures, with 100–200 runs resulting in stable statistics

An alternative evaluation: a game formulation

The *recall* as a function of buffer distance results in a measure of association strength that is easy to interpret based on summary statistics. In this study we also include an alternative evaluation that provides population comparison statistics computed in a different way in order to improve understanding of the underlying associations, and increase confidence *via* a confirmation of the resultant effect. In these analyses, each of the 200 arbitrary experiments is formulated as a “game”, in which the number of significant earthquakes associated with the targeted subducting features is compared to those associated with the arbitrary case. A majority in favour of the targeted regions is a “win”, whereas the opposite case is a “loss” (and “ties” may also occur). It is expected that strong associations will show consistent wins, whereas weaker/arbitrary associations will tend towards balanced wins/losses. These evaluations are computed simply by summing the **A** vectors (defined above) between a *target* and *arbitrary* run, resulting in an outcome G that is either a win/loss/tie based on majority voting. This is computed as follows, denoting the target vector \mathbf{A}^t , the arbitrary vector \mathbf{A}^a , with the k_{th} arbitrary run denoted a_k^a :

$$G = \begin{cases} \text{win if } \sum_{j=1}^N a_j^t > \sum_{j=1}^N a_j^a \\ \text{tie if } \sum_{j=1}^N a_j^t = \sum_{j=1}^N a_j^a \\ \text{loss otherwise} \end{cases}$$

Thus 200 runs results in 200 “games”, with the final outcomes tallied. The process is repeated across a number of buffer distances to establish the performance as a function of proximity.

Additional subduction parameters

The analysis of the subducting bathymetric features investigates these targeted spatial regions in the context of their role in the generation of significant subduction earthquakes. In order to rule out the possibility that the associations are being

primarily controlled by other physical parameters an additional analysis is undertaken. Previous studies have highlighted relationships between subduction earthquakes and interplate convergence rates (Ruff & Kanamori 1983; McCaffrey 1994; Gutscher & Westbrook 2009) as well as the age of the downgoing plate (Ruff & Kanamori 1983; Jarrard 1986), although these were shown not to be significant by McCaffrey (2007) in the wake of the great 2004 Sumatra–Andaman earthquake, as well as by Stein & Okal (2007). In this study we also use relatively recent models of the 3-dimensional slab structure (Gudmundsson & Sambridge 1998; Hayes *et al.* 2012), allowing the slab dip to be included in the follow-up study.

RESULTS AND INTERPRETATION

The primary associations discussed in this paper are presented in Figures 3 and 4 as a function of the buffer distance around each feature projected into the coupling zone. Computing associations for the range of buffer distances constitutes a sensitivity analysis. This analyses the trade-off between the association strength with how specific it is spatially. Fracture zones are analysed separately from volcanic ridges/chains. Figure 3a shows the absolute *recall* scores as a function of a varying buffer distance, both for the targeted bathymetric features, and for the arbitrary case. As expected, the recall rates increase monotonically with increasing buffer distance as successively more coupling-zone filtered earthquakes become enveloped by the various buffer regions. Figure 4 depicts the same result in a normalised form, in which the difference between the target association is shown with the arbitrary case subtracted, allowing the strength of association to be analysed for different buffer distances. In these plots, associations that are larger than zero imply that earthquakes are biased towards the regions surrounding the targeted subducting bathymetric features. Results in which error bars (20th/80th percentiles shown) fall completely within this category (i.e. above zero) imply associations are significant. Considering the fracture-zone intersection results in Figure 3a, it can be seen that for small buffer distances of up to around 100 km, earthquakes are associated with the fracture-zone intersections. This result implies that earthquakes are more likely to be produced at subduction-zone fracture-zone intersections than at arbitrary locations along subduction zones, quantitatively confirming the relationship between these subducting “asperities” and earthquakes on a global scale. In Figure 4a, the strongest associations (compared to the arbitrary case) occur at buffer distances of 50 km

(median difference of about 4%). Decreasing this buffer distance does not modify our results appreciably, which could be explained by the spatial uncertainty of near-trench fracture zone locations that have been projected into the coupling zone. Considering the fracture zone results for buffer distances exceeding 300 km, an initially surprising result is obtained, i.e. that earthquakes appear to be significantly biased *away* from fracture-zone intersections. However, this is due to the fact that many fracture-zone intersection points occur in close proximity of each other, and thus for increasing buffer distances the coupling zone areas used by the arbitrary case tend to exceed those between fracture zone–subduction zone intersections. The implication of this is that the arbitrary case associations computed are statistically biased *against* the identified associations in this study, implying that in reality the effect we studied is even stronger.

Our interpretation of these observations is that significant subduction-based earthquakes are more likely to be generated at, or very close to, the intersections between subduction zones and fracture zones. This is compared with arbitrary locations along subduction zones. This interpretation can be applied globally. From a hazard analysis point of view, we compute a coupling zone surface area of about $\sim 1.1 \times 10^7 \text{ km}^2$, which is about 2.1% of the Earth's surface area ($\sim 51.1 \times 10^7 \text{ km}^2$). We calculated that for a hazard zone utilising a 50 km buffer around fracture-zone intersections, that these regions would consume only 14.7 % of the coupling zone area, thus reducing this area by almost a factor of 7. This may have an immediate implication for hazard modelling, with this evidence in favour of assigning higher risks to 1/7th of the coupling zone regions along subduction zones.

In Figure 3b and Figure 4b (right), the same analysis is presented in relation to subducting volcanic ridges/seamount chains. Again it can be seen that these subducting bathymetric features are also significantly biased in terms of earthquake generation in the vicinity of the subduction intersection (note the error bars that do not intersect imply higher significance). However, the association is weaker and broader. In the fracture zone case the association decreases radically as the buffer zone increases, but this is not the case for the volcanic ridges/seamount chains, attributed to their broader nature. The reduced association strength may be due to the less well-defined linear nature of these features, resulting in less consistent asperities and consequently less reduced coupling. In Singh *et al.* (2011), this

conclusion was also reached pertaining to a large subducting seamount in Sumatra. These results are, however, less reliable than in the fracture zone case because of the low sample size involved.

An independent statistical evaluation is undertaken (Figure 5) to confirm the relationships established. In these experiments, each of the 200 arbitrary experiments has been formulated as a “game” (as described in the methodology), in which the number of significant earthquakes associated with the targeted subducting features is compared to those associated with the arbitrary case. Of interest is establishing the number of “wins/losses/ties”. The results confirm the relationships established using the primary analysis involving computing distances to target regions, i.e. that there is indeed a strong bias towards the subducting asperities defined in this study. It also confirms the stronger association identified for the fracture zone intersections, and the rapid decline as the buffer distance increases, as compared to the volcanic ridges/seamounts.

We assess relationships of buffer width with earthquake magnitude, undertaken for four buffer distances (Figure 6). It can be seen that for the fracture zone intersections, the associations are significant across all magnitude ranges for small buffer distances, but particularly strong for the largest magnitude group. Interestingly, this association remains strong with increasing buffer distance, a statistic not reflected in the main experiments due to the smaller sample sizes (larger earthquakes become rarer with increasing magnitude). This result provides strong additional support for the suggested link between subducting fracture zones and the generation of very large earthquakes (Müller & Landgrebe 2012), extending it to all significant earthquake magnitudes. For the volcanic ridges/chains, associations appear more significant for smaller magnitude earthquakes.

CONTROLLING PARAMETERS

The statistical analysis of subducting linear bathymetric features shows that the regions surrounding subducting, linear bathymetric features are strongly associated with significant earthquakes. In this section, physical parameters comprising established connections with subduction earthquakes, namely the interplate convergence rates, age of the down going slabs, and slab dip, are computed to assess their additional influence on earthquake populations associated with bathymetric feature subduction-intersection regions. We focus this part of our study

on fracture zones, where the identified associations are very localised in space (the previous section showed how the association rapidly diminishes as the region of interest is increased).

In Figure 7, an analysis involving the parameters of interest are presented, that is relationships with convergence rate, down going slab age, and slab dip. These compare significant earthquakes associated within 50 km of fracture-zone intersections to the remaining data, plotting the computed convergence rates *versus* down-going slab age (Figure 7a), and the estimated slab-dip *versus* the down-going slab age (Figure 7b). No strong relationship with convergence rates and down-going slab age is found, leading to the conclusion that these factors do not affect the main associations demonstrated in this paper. This is in agreement with McCaffrey (2007) and Stein & Okal (2007), which refuted the conclusions in Ruff & Kanamori (1983) who identified convergence rates and down-going slab age as strongly linked to associations as a function of earthquake magnitude.

There is weak relationship between the estimated slab dip (Figure 7b), with the slab dips related to the fracture zone intersections appearing to be generally lower than the alternative case. This effect is investigated further in Figure 8, in which the two slab-dip sets of data (or “populations”) are compared *via* a normalised histogram (the histogram frequencies are divided by the population group sizes respectively). A weak relationship can be seen, that is a tendency for fracture-zone intersections to have a more shallow slab dip. The weak nature of this observed association leads to the conclusion that the relationship is not a primary controlling factor, appearing to play a more secondary role. Other factors identified in the literature include relationships with trench sediment thickness, studied in works such as Ruff (1989). As summarised in Gutscher & Westbrook (2009), the sediment thickness at the trench does appear to play a role in large earthquake and tsunami generation, but is in turn a function of other parameters such as the convergence rate and sediment composition, leading to a more indirect relationship with too few target subduction zone intersections available to study in a statistically significant fashion.

CONCLUSIONS

We consider a growing body of evidence linking the subduction of particular bathymetric asperities with the occurrence of significant earthquakes. Geophysical datasets related to well-defined, linear bathymetric features including fracture zones,

aseismic ridges and volcanic chains have resulted in the digital compilation of these features on a global scale. This study involves identifying localities along subduction zones that intersect with these features (implying they are being subducted), and quantitatively assessing their adjacency associations with significant earthquakes. A global coupling zone model was used to define the interface between down-going slabs and over-riding plates, allowing a large significant earthquake catalogue to be filtered into subduction-related events only. The association analysis reveals that significant earthquakes are biased towards regions involving both subducting fracture zones and volcanic ridges/chains. A novel analysis approach is utilised to validate the significance of the result, in which the assessment of association is posed as a game, involving computing wins, losses and ties. Fracture zone intersections were found to exhibit a stronger association that rapidly diminishes with increasing distance from the targeted regions, whereas the volcanic ridges/chains demonstrated a smaller, broader effect. The latter results were, however, based on a far smaller sample size. A more detailed assessment of associations as a function of earthquake magnitude revealed significant associations for fracture zones with earthquake moment magnitudes above 8.5, whereas the volcanic ridges/chains showed significant associations for moment magnitudes lower than 8.0. An analysis of additional factors related to interplate convergence velocities, subducting sea-floor age and subducting slab dips revealed no strong relationships with significant earthquakes, confirming the primary role being played by linear subducting seafloor features. A data-mining approach proved highly effective in taking account of noise, variability, and large multi-dimensional datasets.

ACKNOWLEDGEMENTS

This research was funded by ARC grant FL0992245. This research was supported by the Australian Research Council (FL099224) and the Science Industry Endowment Fund (RP 04-174) Big Data Knowledge Discovery project.

REFERENCES

AMANTE C., EAKINS B. W. & NGDC. 2009. *ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis*. US Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite,

- Data, and Information Service, National Geophysical Data Center, Marine Geology and Geophysics Division.
- BILEK S. L., SCHWARTZ S. Y. & DESHON H. R. 2003. Control of seafloor roughness on earthquake rupture behavior. *Geology* **31**, 455.
- CLOOS M. 1992. Thrust-type subduction-zone earthquakes and seamount asperities: A physical model for seismic rupture. *Geology* **20**, 601–604.
- COFFIN M. F. & ELDHOLM O. 1994. Large igneous provinces: Crustal structure, dimensions, and external consequences. *Reviews of Geophysics* **32**, 1–36.
- CONTRERAS-REYES E. & CARRIZO D. 2011. Control of high oceanic features and subduction channel on earthquake ruptures along the Chile–Peru subduction zone. *Physics of the Earth and Planetary Interiors* **186**, 49–58.
- GORBATOV A., KOSTOGLODOV V., SUAREZ G. & GORDEEV E. 1997. Seismicity and structure of the Kamchatka subduction zone. *Journal of Geophysical Research* **102**, 17883–17898.
- GUDMUNDSSON Ó. & SAMBRIDGE M. 1998. A regionalized upper mantle (RUM) seismic model. *Journal of Geophysical Research: Solid Earth* **103**, 7121–7136.
- GUTSCHER M. A. & WESTBROOK G. K. 2009. Great earthquakes in slow-subduction, low-taper margins. *In: Lallemand S. & Funiciello F. eds. Subduction Zone Geodynamics*, pp 119–133, Springer-Verlag, Berlin Heidelberg.
- HAYES G. P., WALD D. J. & JOHNSON R. L. 2012. Slab1. 0: A three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research* **117**, B01302.
- JARRARD R. D. 1986. Relations among subduction parameters. *Reviews of Geophysics* **24**, 217–184.
- KODAIRA S., TAKAHASHI N., NAKANISHI A., MIURA S. & KANEDA Y. 2000. Subducted seamount imaged in the rupture zone of the 1946 Nankaido earthquake. *Science* **289**, 104.
- KREEMER C., HOLT W. E. & HAINES A. J. 2003. An integrated global model of present-day plate motions and plate boundary deformation. *Geophysical Journal International* **154**, 8–34.
- LANGE D., TILMANN F., RIETBROCK A., COLLINGS R., NATAWIDJAJA D. H., SUWARGADI B. W., BARTON P., HENSTOCK T. & RYBERG T. 2010. The fine structure of the

- subducted investigator fracture zone in western Sumatra as seen by local seismicity. *Earth Planetary Science Letters* **298 (1-2)**, 47–56.
- LAY T. & KANAMORI H. 1981. An asperity model of large earthquake sequences. *Earthquake Prediction* **4**, 579–592.
- MATTHEWS K., MÜLLER R. D., WESSEL P. & WHITTAKER J. M. 2011. The tectonic fabric of the ocean basins. *Journal of Geophysical Research* **116**, 1–28.
- MCCAFFREY R. 1994. Dependence of earthquake size distributions on convergence rates at subduction zones. *Geophysical Research Letters* **21**, 2327–2330.
- MCCAFFREY R. 2007. The next great earthquake. *Science* **315**, 1675–1676
- MCCALPIN J. P. 2009. Earthquake magnitude scales. *In: James P. M. ed. International Geophysics, Volume 95*, pp 1–3, Academic Press. San Diego.
- MINOURA K. & IMAMURA F. & SUGAWARA D. & KONO Y. & 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**, 83–88.
- MOCHIZUKI K., YAMADA T., SHINOHARA M., YAMANAKA Y. & KANAZAWA T. 2008. Weak interplate coupling by seamounts and repeating M~ 7 earthquakes. *Science* **321**, 1194–1197.
- MÜLLER R. D. & LANDGREBE T. C. W. 2012. The link between great earthquakes and the subduction of oceanic fracture zones. *Solid Earth* **3**, 447–465.
- MÜLLER R. D., SDROLIAS M., GAINA C. & ROEST W. R. 2008. Age, spreading rates, and spreading asymmetry of the world's ocean crust. *Geochemistry, Geophysics, Geosystems* **9**, Q04006.
- NGDC/WDC. 2011. National Geophysical Data Center / World Data Center Significant Earthquake Database. *Boulder, CO, USA*.
- OLESKEVICH D. A., HYNDMAN R. D. & WANG K. 1999. The updip and downdip limits to great subduction earthquakes: Thermal and structural models of Cascadia, south Alaska, SW Japan, and Chile. *Journal of Geophysical Research*. **104**, B7, 14965–14991, doi:10.1029/1999JB900060.
- ROBINSON D., DAS S. & WATTS A. 2006. Earthquake rupture stalled by a subducting fracture zone. *Science* **312**, 1203.
- RUFF L. J. 1989. Do trench sediments affect great earthquake occurrence in subduction zones? *Pure and Applied Geophysics* **129**, 263–282.

- RUFF L. J. 1992. Asperity distributions and large earthquake occurrence in subduction zones. *Tectonophysics* **211**, 61–83.
- RUFF L. J. & KANAMORI H. 1983. Seismic coupling and uncoupling at subduction zones. *Tectonophysics* **99**, 99–117.
- SCHOLZ C. H. & SMALL C. 1997. The effect of seamount subduction on seismic coupling. *Geology* **25**, 487.
- SIEH K., NATAWIDJAJA D. H., ARON J. M., SHEN C.-C., CHENG H., LI K.-S., SUWARGADI B. W., GALETZKA J., PHILIBOSIAN B. & EDWARDS R. L. 2008. Earthquake supercycles inferred from sea-level changes recorded in the corals of West Sumatra. *Science* **322**, 1674–1678.
- SINGH S. C., HANANTO N., MUKTI M., ROBINSON D. P., DAS S., CHAUHAN A., CARTON H., GRATACOS B., MIDNET S., DJAJADIHARDJA Y. & HARJONO H. 2011. Aseismic zone and earthquake segmentation associated with a deep subducted seamount in Sumatra. *Nature Geoscience* **4**, 308–311.
- STEIN S. & OKAL E. A. 2007. Ultralong period seismic study of the December 2004 Indian Ocean earthquake and implications for regional tectonics and the subduction process. *Bulletin of the Seismological Society of America* **97**, S279–S295.
- STEIN S., GELLER R.J. & LIU M. 2012. Why earthquake hazard maps often fail and what to do about it. *Tectonophysics*. **1–25**, 562–563.
- TAYLOR F. W., BRIGGS R. W., FROHLICH C., BROWN A., HORNBAACH M., PAPABATU A. K., MELTZNER A. J. & BILLY D. 2008. Rupture across arc segment and plate boundaries in the 1 April 2007 Solomons earthquake. *Nature Geoscience* **1**, 253–257.
- VON HUENE R. & RANERO C. R. 2009. Neogene collision and deformation of convergent margins along the backbone of the Americas. In: Mahlburg Kay S., Ramos V. A. & Dickinson W. R. eds., *Backbone of the Americas: shallow subduction, plateau uplift, and ridge and terrane collision*, pp 67–84, Geological Society of America. Boulder Co.
- VON HUENE R., KLAESCHEN D., FRUEHN J., SAUBER J. & DMOWSKA R. 1999. Relation between the subducting plate and seismicity associated with the great 1964 Alaska earthquake. *Pure and Applied Geophysics* **154**, 575–592.

WELLS R. E., BLAKELY R. J., SUGIYAMA Y., SCHOLL D. W. & DINTERMAN P. A. 2003.

Basin-centered asperities in great subduction zone earthquakes: A link between slip, subsidence, and subduction erosion. *Journal of Geophysical Research* **108**, 1–30.

Received; accepted December 2014

APPENDIX

Table A.1 Latitude–longitude pairs pertaining to identified fracture-zone intersections with subduction zones.

Table A.2 Latitude–longitude pairs pertaining to identified volcanic ridge/seamount-chain intersections with subduction zones.

FIGURE CAPTIONS

Figure 1 Map with prominent datasets used in this study superimposed on the ETOPO1 global relief model (Amante *et al.* 2009), on a Robinson projected map: subduction coupling zones (light green/yellow bands) (Hayes *et al.* 2012), oceanic fracture zones (light grey) and oceanic volcanic chains and aseismic ridges (maroon outlines) (Matthews *et al.* 2011), intersection points between fracture zones and subduction zones (purple squares), intersection points between volcanic chains and ridges with subduction zones (green squares), and significant earthquakes (beige circles) (NGDC/WDC 2011).

Figure 2 Illustrating the approach taken to compute arbitrary associations in the North-Western Pacific region, with blue lines portraying the subduction depth contours in this region (from Müller & Landgrebe 2012). The coupling zone is segmented (axis-aligned) into 20 km-spacings (yellow lines), allowing the coupling zone to be “queried” in an unbiased fashion, simulating an arbitrary bathymetric feature being subducted. Black lines indicate fracture zones in the region.

Figure 3 Results of the association calculations, comparing the bathymetric features to the arbitrary case for fracture zones (left), and volcanic ridges/chains (right), as a function of the buffer distance around each feature. The measure *recall* is computed as the association between the bathymetric features and the significant

earthquakes dataset. Error bars on the arbitrary case depict 20th/80th percentiles over 200 experimental runs.

Figure 4 Results of the association calculations as in Figure 3, plotting the normalised result (*delta recall*) to compute the association difference between the bathymetric features and the significant earthquakes dataset and the same association computed for the arbitrary case (median). This allows association strength to be visualised in a normalised form for a growing buffer distance. Results above the zero line indicate non-arbitrary associations (called the “biased” case).

Figure 5 Comparing associations with the arbitrary case over 200 runs for fracture zones (left) and volcanic ridges/chains (right) as a function of buffer distance. A “win” is where earthquake associations with the bathymetric features exceed those pertaining to the arbitrary case, and a “loss” is the opposite case.

Figure 6 Assessing associations for four different moment magnitude earthquake data partitions, by comparing results for fracture zones on the left, and volcanic ridges/chains on the right. The same associations are shown for four buffer ranges to demonstrate the insensitivity of the associations. The magnitude partitionings are defined as follows: the label “8.5” consists of data samples with moment magnitude $M_w \geq 8.5$; “8.0” consists of samples with $8.5 < M_w \leq 8.0$; “7.5” consists of samples with $8.0 < M_w \leq 7.5$; “<7.5” consists of samples with $M_w < 7.5$.

Figure 7 Additional controlling parameters typically linked with subduction earthquakes for the fracture zone intersections: statistics related to coupling-zone filtered earthquakes within 50 km of fracture-zone intersections are plotted *versus* other earthquakes (denoted “Not associated”). The left plot depicts global convergence velocities versus the age of the down-going slab closest to the earthquake location. Similarly the right plot shows the estimated slab dip closest to the earthquake hypocenter *versus* the age of the down-going slab.

Figure 8 Normalised histogram showing the population distribution pertaining to the slab-dip associated with earthquakes within 50 km of fracture zone intersections, *versus* others. This plot suggests a weak association at best.

APPENDIX

Table A.1. Latitude-longitude pairs pertaining to identified fracture-zone intersections with subduction zones.

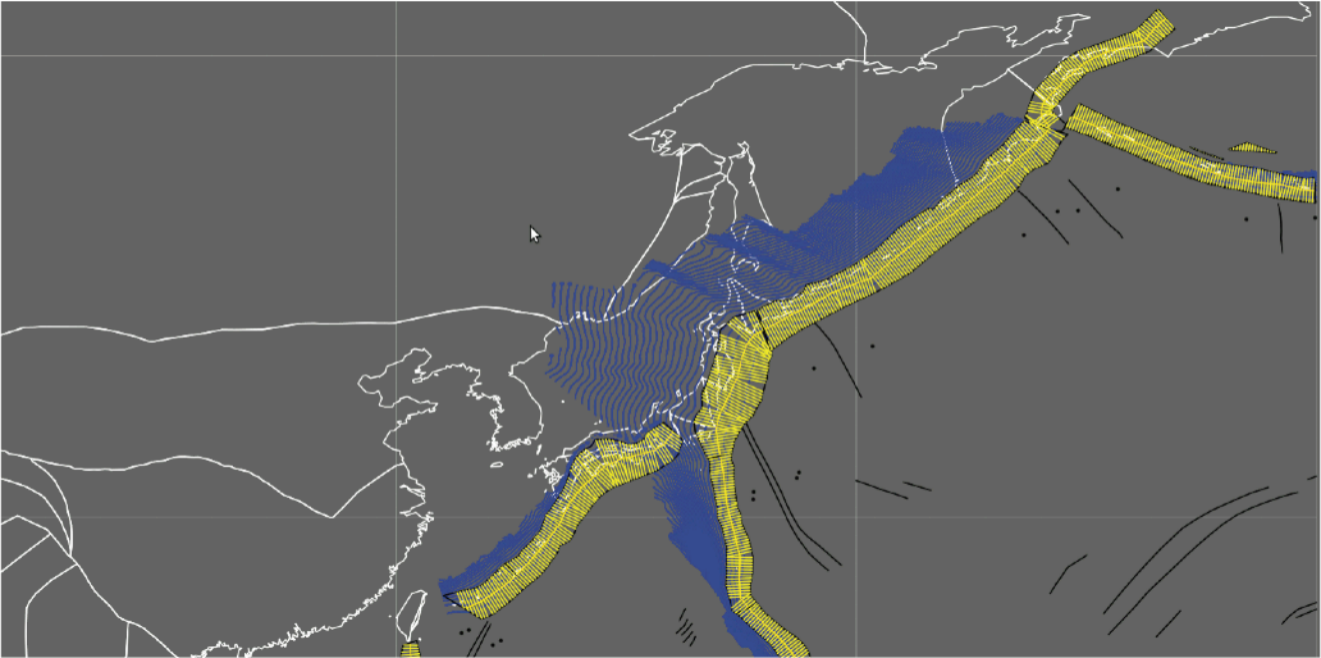
Latitude	Longitude
-2.277	98.529
3.829	93.172
7.75	91.882
2.346	94.585
2.615	94.231
-7.972	104.055
-7.426	102.846
-6.792	102.075
-3.092	99.202
-7.215	102.483
1.757	95.715
17.003	-59.192
13.205	-57.587
12.942	-57.544
13.559	-57.652
13.887	-57.712
55.836	-152.693
56.249	-151.462
14.891	-95.124
-3.047	-81.642
-4.194	-81.977
-4.499	-82.009
-5.103	-82.001

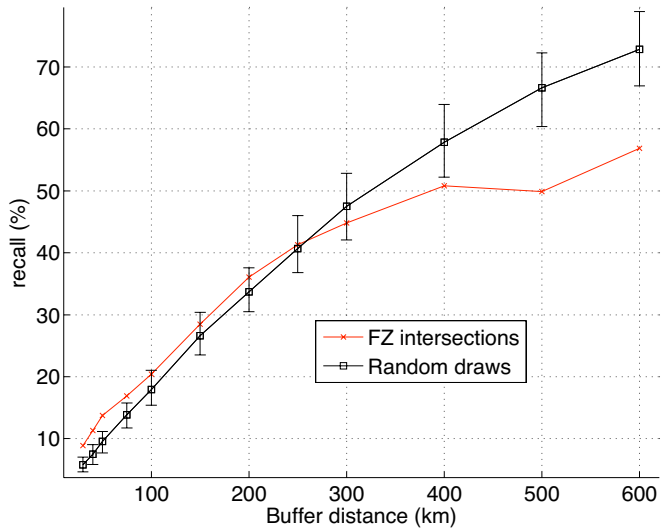
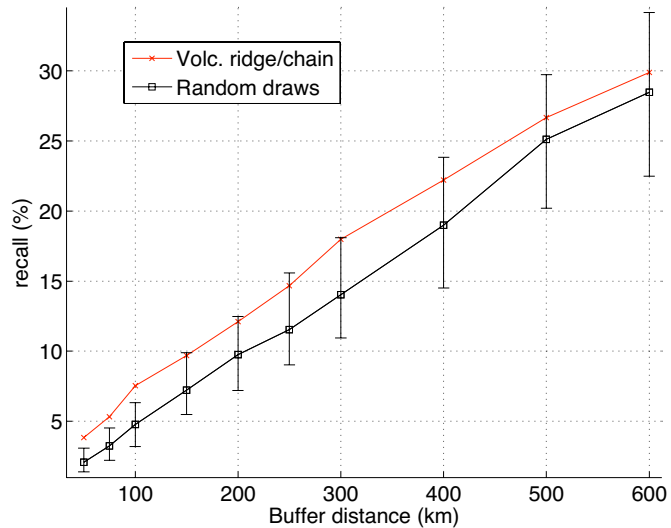
Latitude	Longitude
-5.361	-81.992
-10.087	-80.272
-8.989	-80.898
-10.398	-80.091
15.218	-95.694
13.383	-92.463
-17.632	-73.569
-45.619	-75.997
-45.834	-75.995
-46.604	-75.969
-47.015	-75.993
-48.943	-76.988
-43.892	-75.723
-41.767	-75.396
-40.334	-75.192
-40.024	-75.154
42.641	147.255
36.193	142.769
35.871	142.529
51.336	160.426
53.299	162.616
-37.94	-74.634
-61.481	-61.098
-56.746	-69.786
-61.324	-60.48
23.667	126.257
23.91	126.634
-32.408	-72.789

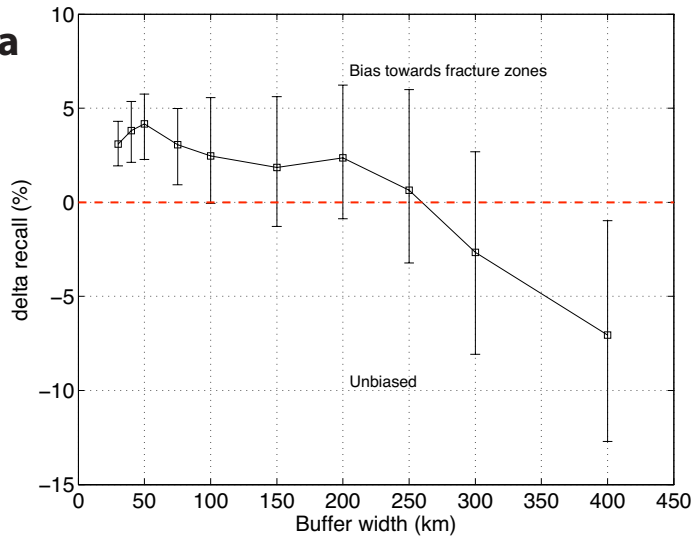
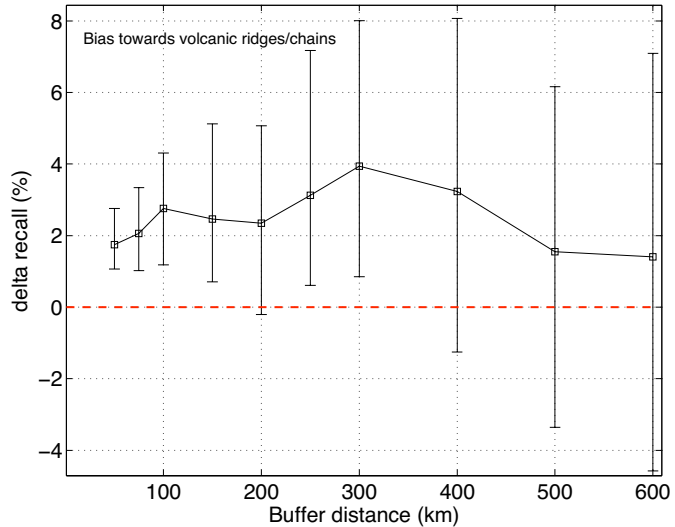
Latitude	Longitude
50.751	177.422
51.023	-172.852
50.525	-176.819
-59.295	-24.365
-58.21	-23.823
-57.217	-24.215
8.58	-82.61
40.3	-124.73

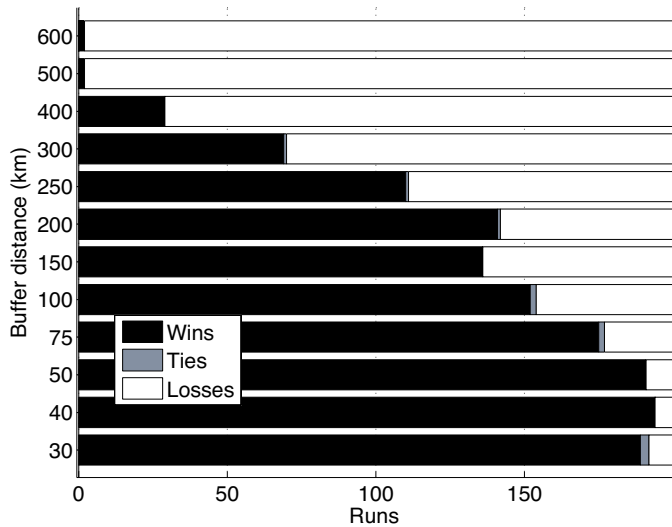
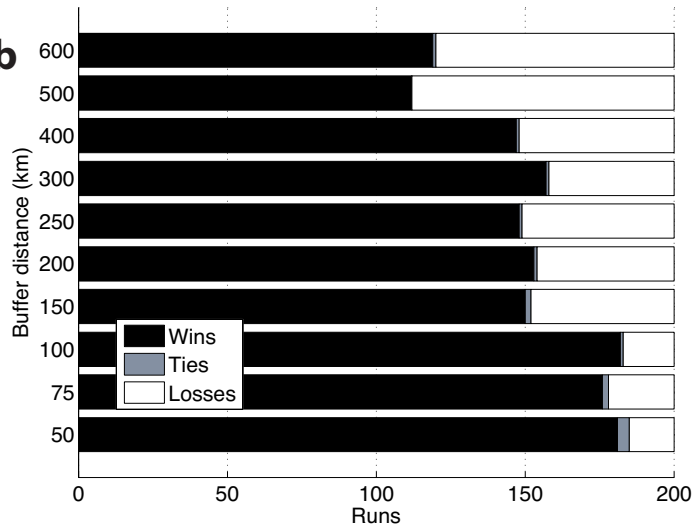
Table A.2. Latitude-longitude pairs pertaining to identified volcanic ridge/seamount-chain intersections with subduction zones.

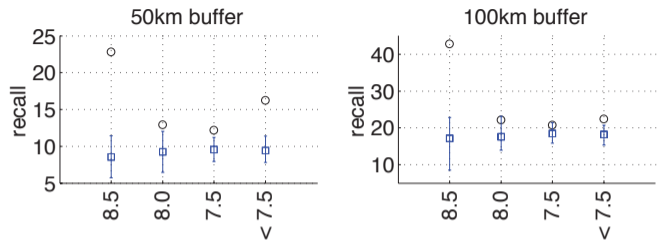
Latitude	Longitude
-0.79	-81.54
7.89	-83.36
15.87	147.76
20.37	147.19
24.39	143.48
54.11	163.1
-15.27	-76.45
-11.28	115.06
-6.53	154.21
-25.19	-175.24
-35.8	-179.13
-40.64	178.5
-9.65	158.94
10.85	139.52



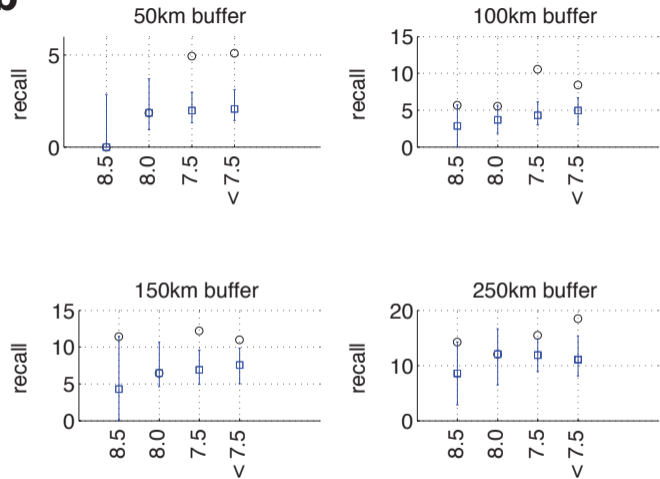
a**b**

a**b**

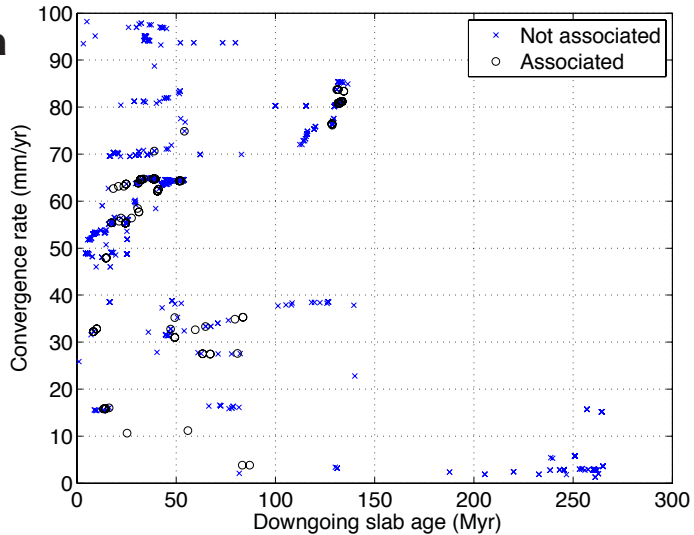
a**b**

a

○ Fracture Zones □ Arbitrary

b

○ Volc. ridge/chain □ Arbitrary

a**b**