# Machine Learning can predict the timing and size of analog earthquakes

**F. Corbi[1*], L. Sandri[2], J. Bedford[3], F. Funiciello[1], S. Brizzi[1,4], M. Rosenau[3], S. Lallemand[5]**

[1] Università "Roma TRE", Dip Scienze, Laboratory of Experimental Tectonics, Rome, Italy.
[2] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Bologna, Italy.
[3] Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, Potsdam, Germany.
[4] University of Parma, Natural and Experimental Tectonics research group, Dip. SCVSA, Parma, Italy.
[5] Géosciences Montpellier, CNRS, University of Montpellier, Montpellier, France.

Corresponding author: Fabio Corbi (fabio.corbi3@gmail.com)

**Key Points:**

- We simulate multiple seismic cycles in a laboratory-scale subduction zone with two asperities.
- Slip-deficit appears to be diagnostic of the location of highest slip but is poorly informative of the size of next event.
- Deciphering the spatially and temporally complex surface deformation history, ML predicts the timing and size of analog earthquakes.



This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2018GL081251

**Abstract**

Despite the growing spatio-temporal density of geophysical observations at subduction zones, predicting the timing and size of future earthquakes remains a challenge. Here, we simulate multiple seismic cycles in a laboratory-scale subduction zone. The model creates both partial and full margin ruptures, simulating magnitude Mw 6.2-8.3 earthquakes with a coefficient of variation in recurrence intervals of 0.5, similar to real subduction zones. We show that the common procedure of estimating the next earthquake size from slip-deficit is unreliable. On the contrary, Machine Learning predicts well the timing and size of laboratory earthquakes by reconstructing and properly interpreting the spatio-temporally complex loading history of the system. These results promise substantial progress in real earthquake forecasting, as they suggest that the complex motion recorded by geodesists at subduction zones might be diagnostic of earthquake imminence.

**Plain Language Summary**

Large and devastating subduction earthquakes, such as the 2011 magnitude 9.0 Tohoku-oki earthquake (Japan), are currently considered unpredictable. Scientists lack a long enough seismic catalogue that is necessary for drawing statistical insights and developing predictions. For this reason, we simulate tens of earthquakes using a small-scale experimental replica of a subduction zone. We show that Machine Learning (ML - a group of algorithms that make predictions based on the "information" acquired in past "experience") can predict when, where and how big the next experimental earthquake will be. The "information" in our study is provided by the slow deformation accumulating in the analog tectonic plates during the periods in-between earthquakes. Since such slow deformation is also measured by means of space geodesy along real subduction zones, there is the possibility that, in the future, variations of this ML approach can predict the timing and size of natural subduction earthquakes.

## 1 Introduction

The subduction megathrust - the interface between subducting and overriding tectonic plates – has hosted the largest earthquakes of the last century, e.g., the 2011 $M_w$ 9.0 Tohoku-Oki earthquake (Japan). Such events originate from sudden slip episodes that propagate for several hundreds of kilometers along the megathrust, displacing the Earth's crust by tens of meters and relaxing the elastic stress that has accumulated due to the combination of tectonic plate convergence and friction acting along the megathrust.

In the ideal case of *(i)* perfectly elastic plates, *(ii)* constant coupling between overriding and subducting plate, *(iii)* constant plate convergence velocity, and *(iv)* homogeneous frictional properties and pre-stress along the megathrust, earthquakes should occur at regular time intervals, with only small variations in magnitude (i.e., the characteristic earthquake model; Reid 1910). Other earthquake recurrence models, including the time- and slip-predictable models (Shimazaki & Nakata, 1980) and the non-periodic non-characteristic model (Satake and Atwater, 2007), take into account temporal variations of fault strength, stress drop and loading rate. Such variations move the system towards a more complex recurrence pattern (e.g., Brodsky and Kanamori, 2004). Subduction megathrusts follow irregular rupture patterns with recurrence times for large earthquakes spanning from a few hours to centuries or more (Ando 1970; Sieh et al 2008; Goldfinger 2017). Such variability can be explained by the interplay between spatio-temporal variations of interplate coupling (e.g., Moreno et al., 2011; Tsang 2015), spatial variations of frictional properties (e.g., Cubas et al., 2013), differences of interplate roughness (e.g., Lallemand et al., 2018) and non-steadiness of plate motions (Heki and Mitsui, 2013; Loveless and Meade 2016).

Despite the international efforts to instrument and monitor subduction zones with unprecedented resolution (McGuire & Plank, 2017) it is still unclear how to constrain the timing and size of future earthquakes. There is increasing seismological evidence of phenomena related to the failure of highly locked zones (aka asperities) on the megathrust such as the increase of foreshocks (Schurr et al., 2014) and repeating events that are indicative of a slow slip on the interface surrounding the asperity, pushing the megathrust to failure (Kato et al., 2012). Geodesy also yields useful information for understanding megathrust dynamics. By measuring the ratio between the overriding and incoming plate velocities, geodesists can identify coupled (locked) portions of the megathrust, where stress builds up interseismically and where this stress will likely be released through future earthquakes. Moreover, the recent identification of transients (i.e., interseismic accelerations and decelerations) in the geodetic time series prior to large earthquakes (e.g. Mavrommatis et al., 2014) suggests that the continental surface velocity is likely a good indicator of when a given portion of the megathrust is ready to fail. The recent progress in seismic and geodetic identification of earthquake precursory phenomena is accompanied by a rapidly increasing number of studies where Machine Learning (ML) algorithms have been used for earthquake related problems, such as predicting laboratory earthquakes (Rouet-Leduc et al., 2017), estimating lab-scale fault friction (Rouet-Leduc et al., 2018), predicting GPS displacement rates associated to slow slip events (Rouet-Leduc et al., 2018b), and forecast of aftershock locations (DeVries et al., 2018).

Currently there are too few documented instances of precursory phenomena to facilitate a robust assessment of earthquake imminence. Moreover, seismic and geodetic records span only a small fraction of the seismic cycle duration, and therefore it is a challenge to determine which features of the data are indicative of the late-interseismic phase.

To overcome the lack of long time series in real subduction zones, we profit from recent developments in analog seismotectonic modeling that allows for the simulation of multiple seismic cycles in a convenient experimental time and scale (Rosenau et al., 2017). We show that, while slip-deficit alone has a low informative power with respect to the size and timing

of future earthquakes, ML can decipher the geodetic-like signal preceding slip events, allowing for the prediction of both the timing and size of future events.

## 2 The state of the art: Inferring the pattern of future earthquakes using interseismic coupling

Megathrust earthquakes grow into $M_w$>9 events by unzipping multiple asperities along strike of the subduction zone (Lay and Kanamori 1981; Moreno et al., 2009; Subarya et al., 2006). Inferring the number of asperities that may fail is thus a first-order predictor for the size of future events. Estimates on potential along-strike rupture extent and, in turn, earthquake magnitude are aided by maps of interseismic coupling (ISC) generated from geodetic measurements (e.g., Avouac 2015). Higher landward velocities of coastal sites indicate more strongly coupled zones of the megathrust, where most of elastic energy and slip-deficit accumulate during the interseismic observation period.

According to the slip-predictable model, coseismic slip equals slip-deficit, while the timing remains unknown (Shimazaki & Nakata, 1980). Therefore, the pattern of ISC may serve as a proxy for constraining the size of future earthquakes. However, this simple model is supported by only a few observations from natural subduction zones (Moreno et al., 2010, Schurr et al., 2014) and numerical simulations (Kaneko et al., 2010). In contrast, recent investigations reveal that high ISC zones may span the whole range of coseismic slip magnitudes (Metois et al., 2016; Barnhart et al., 2016) and that earthquakes may leave large portions of highly coupled megathrust unruptured (e.g., Konca et al., 2008). Moreover, some earthquakes [e.g., $M_w$ 8.4 Sumatra 2007] propagated or even nucleated into areas of low/no coupling (Konca et al., 2008) highlighting the potential for temporal variations in the ISC pattern as well as uncertainties associated to the ISC inversion method. Hence, the concept that slip-distribution mirrors ISC cannot be generalized. The 2010 $M_w$ 8.8 Maule earthquake (Chile) demonstrates the challenges of setting up and interpreting ISC models. For Maule, slip-distributions have been suggested to correlate with ISC either very well (Moreno et al., 2010; Melnick et al., 2012) or poorly (Lorito et al., 2011). This situation is partly due to uncertainties associated with inversion methods, inconsistent data availability before and after the event, and missing spatial coverage of the offshore seismogenic region, meaning that improving the robustness of spatial correlation between slip and locking models is an ongoing area of research (e.g., Loveless and Meade, 2011; Barnhart et al., 2016). Besides modeling issues, the major problem in determining the spatio-temporal relationship between slip-deficit and earthquake slip is the limited number of natural cases where large subduction zone earthquakes have been recorded by dense geodetic networks: this is where analog seismotectonic models can fill the gap.

## 3 New analog models of megathrust earthquakes

Analog modeling is an experimental technique allowing for the simulation of a given geologic process in a convenient spatial and temporal scale. The recent improvements in imaging and monitoring techniques associated to specific scaling (Rosenau et al., 2017) allows lab modelers to simulate the main features of the megathrust seismic cycle in a simplified but realistic way (e.g., Corbi et al., 2013). To conduct our model, we use a 52 x 34 x 11 $cm^3$ gelatin wedge (analog of the overriding plate) underthrusted at constant rate of 0.01 cm/s by a 10° dipping, flat rigid plate (analog of the subducting plate, Fig. 1a). The analog megathrust embeds two velocity weakening patches (asperities) of equal size and friction separated by a velocity strengthening patch (barrier). Velocity weakening and strengthening behaviors are achieved with the gelatin on sandpaper and gelatin on plastic contacts, respectively (Corbi et al., 2013). After an initial stress build-up phase, the model experiences

stick-slip behavior consisting of periods of stress build-up interrupted by spontaneous nucleation of frictional instabilities propagating at the contact between the gelatin and the plate (Corbi et al., 2013). These instabilities are the analog earthquakes. The model produces ruptures across single (Fig. 1b) or twin asperities (Fig. 1c) which relative number depends on the barrier to asperities length ratio (Corbi et al., 2017). The analysis of the modeled surface deformation is performed via image cross correlation (MatPiv; Sveen, 2004). This analysis provides us the velocity field between consecutive images, discretized in 1350 interrogation windows that are analogs of homogeneously distributed "synthetic GPS stations" (1 station every 1.2 cm, equivalent to 1 station every 7.5 km in nature) above the whole model surface (Fig. 1d).

## 4 Spatial correlation between slip and slip-deficit

To evaluate the correlation between slip- and slip-deficit maps we select a 7 min long interval (Fig. S1; data are published open access in Corbi et al. 2019) during which the experiment produces 40 seismic cycles with average duration of about 10.5 s (standard deviation 4.9 s, Fig. 1e). The corresponding coefficient of variation (CV=standard deviation/mean) of ca. 0.47 indicates a recurrence behavior intermediate between quasi-periodic (CV<0.1) and random (CV=1), and which reproduces the general behavior of natural subduction zones [e.g. Cascadia, CV=0.5 (Kulkarni et al., 2013); South Peru- North Chile, CV=0.3 (Comte & Pardo 1991), South Chile, CV=0.4 (Bookhagen et al. 2006)]. The key advantages of experimental over natural data are straightforward: (*i*) a denser spatial coverage of synthetic GPS stations, (*ii*) availability of the offshore (hardly accessible) region up to the trench, and (*iii*) long time-series spanning tens of seismic cycles (hundreds of thousands of years). In order to avoid inversion bias, we directly compare interseismic and coseismic surface motion rather than slip at depth. Accordingly, we will hereafter refer to the seaward surface motion as slip (i.e., seaward motion) and the landward surface motion as slip-deficit. Slip and slip-deficit maps of each cycle are constructed by summing incremental displacement fields during coseismic and interseismic phases, respectively.

Our synthetic catalog of slip maps consists of 30 single asperity ruptures with scaled-to-nature Mw [Mw is computed first scaling to nature the rupture area A and then using the following magnitude-rupture area proportionality (Strasser et al., 2010): Mw = 4.441 + 0.846*(log10(A))] spanning from 6.2 to 8.0 (Fig. 1b) and 10 twin-asperity ruptures with scaled-to-nature $M_w$ 8.3 (i.e., ruptures that saturate the entire model length; 2c). Slip-deficit maps are generally characterized by two patches of relatively higher deficit located above the two velocity weakening regions. The average ISC across all cycles is approximately 0.5 (Fig. S2).

We quantified the similarity between the slip-deficit pattern and the subsequent slip by means of spatial correlation computed both within the slip area (Rsa) and over the whole seismogenic zone (Rsz). We found that in 40% of seismic cycles, Rsa and Rsz show good correlations (i.e., both Rsz and Rsa >0.5; Fig. 2a), and in 50% of cycles, Rsa is high and Rsz is low (i.e., Rsa > 0.5 and Rsz < 0.5; Fig. 2b). The former case corresponds to a scenario of slip mirroring slip-deficit, the latter case corresponds to a scenario where slip matches well the slip-deficit pattern within the ruptured asperity but the second region of high slip-deficit remains unruptured. Over the analyzed 40 seismic cycles (Fig. S3), Rsa generally shows higher values (> 0.4), while Rsz spans homogeneously from 0 to 1 (Fig. 2c). This indicates that: a) velocity weakening regions are areas where slip-deficit accumulates interseismically and where the bulk of slip occurs (e.g., Kaneko et al., 2010); and b) slip-deficit represents only a weak proxy for constraining the lateral extent of future earthquakes because locked patches do not necessarily rupture in a given seismic cycle. Based on our modelling results,

we conclude that slip-deficit alone has low predictive potential with respect to the size of future events in a twin-asperity system.

This variability of Rsz can be explained by the quasi-periodic behavior of each asperity (CV<0.1; Fig. S4). The asperities have mean recurrence times of 19.2±0.9 s and 16.1±1.3 s (where error is one standard deviation). This recurrence behavior of the two asperities results in occasional large (dual asperity) ruptures that correlate highly with the slip deficit of the whole subduction margin, whereas most of the smaller, single asperity ruptures correlate with the slip deficit local to the rupture.

## 5 Machine Learning predicts the timing and size of laboratory earthquakes

Our next modeling approach is inspired by the recent laboratory study of Rouet-Leduc et al. (2017) in which ML was used to determine which features of the continuous acoustic data of a block slider experiment were most predictive of the time until failure. Following this supervised learning framework, we take the continuous "synthetic" geodetic signals from our experiment and construct a ML problem to determine which features of these data (if any) can predict time until failure.

We use the Gradient Boosted Regression Trees (GBRT; Friedman 2001; Supplementary Information S1), a ML regression algorithm, to predict the time to the subsequent earthquake (hereafter time to failure, TTF) based on 94 features describing the surface deformation measured in the analog model (Table S1, none of the features refers to time). The GBRT is first trained on a portion of data: in this phase, the GBRT learns the relationships between data features and the target TTF (Supplementary Information S2). The trained algorithm is then fed with 'test' data (data not used in the training) and predicts the subsequent TTFs. To quantify the quality of the predictions, we report the correlation coefficient R.

We use a shifting training window of N (N=2,...35) seismic cycles and a single (the subsequent) cycle to test. This data split ensures no "leakage" into the testing data (having a total of 40 seismic cycles available, we can run 40-N shifting tests). For each iteration, we select the most relevant features to be used for the prediction (Supplementary Information S2). With respect to the standard approach of training the algorithm on a large fraction of the dataset and testing on the remaining part (e.g., training on 60% and testing on 40%; Supplementary Information S3; Fig. S5), our approach improves the prediction performances and also allows us to evaluate the influence of the training window length on prediction performances. We found that the best prediction performances are achieved for N=10 (see Supplementary Information S4 for details about the metrics used for discriminating between the prediction performance of models with different N; Fig. S6). In this case, R=0.3 over the whole time series (Fig. 3a; Fig. S7). Considering that in this case the GBRT has to predict TTF for a variety of event locations and magnitudes, this result shows that the model does a decent job of generalizing.

Aiming to improve the fit to observations, we include an additional constraint: we train individual GBRT models for TTF at 9 target points equally distributed parallel to the trench (Fig. 1d). The TTF of individual points appears more regular/periodic (with CV<0.1 for target point 1, 2, 3, and 8; Fig. S4) and easier to interpret for the GBRT, especially those located above the asperities (Fig. S8). In fact, the GBRT is capable of predicting the TTF of individual target points with great accuracy (i.e., half of the target points located within the two asperities show 0.7<R<0.8), even when using a short training window length of N=5. By checking which and how many adjacent points will rupture simultaneously (i.e., analyzing the space-time distribution of TTF; Fig. 3b), this method allows predicting the location, size and timing of upcoming events. For example, in figure 3c we see that at time 3.05 min, the GBRT correctly predicts an upcoming event occurring in the northern part of the trench, involving 4 to 5 target points [≈86-107 km when scaled to nature (Corbi et al., 2013),

corresponding to $M_w$ 7.9/8.0 according to the magnitude-rupture length scaling (Strasser et al., 2010)]. Similarly, the GBRT successfully predicts an upcoming event at time 3.85 min spanning the whole model width (i.e., 7 target points corresponding to a rupture length of 150 km in nature and therefore a $M_w$ 8.3). There are still some misfits (e.g., the false alarm at 5.05 min in the northern asperity) that make our analysis still only semi-quantitative. However, the success of this ML approach raises the challenging possibility that features of real geodetic signals may be used as a proxy for constraining the onset time of future earthquakes.

## 6 Identifying diagnostic features in geodetic time-series

Understanding which features are the most important for a successful prediction is essential for exporting this method to real subduction zones. When ranking our features in terms of predictor importance (Supplementary Information S5) we observe that the top 15 features are representations of the cumulative displacement (or elastic loading) of the plate margin (Fig. S9a). In particular, the algorithm identified the trench parallel component of cumulative displacement measured at target point 2 and 7 (i.e., Dv2 and Dv7) as the most influent features of the whole set. Dv2 and Dv7 are indeed roughly linearly related to TTF so that they became a valid proxy for TTF (Fig. S9b). Therefore, the GBRT is mostly informed by the relative loading history of various positions on the surface of the model (Fig. S10).

The dominant role of cumulative displacement is not a surprise: the elastic wedge cannot be loaded by convergence indefinitely but only as much as its frictional base is able to sustain in agreement with the elastic rebound theory. Rather, what is remarkable, is that the GRBT keeps track of the site specific loading history and reconstructs when the strength limit is approaching at individual surface locations.

## 7 Implications and outlook: Towards earthquake prediction?

We have used an analog model that reproduces multiple subduction megathrust seismic cycles to overcome the limits caused by *(i)* the short temporal span of geodetic data with respect to the duration of natural seismic cycles, *(ii)* the fragmentary coverage of the geodetic network, and *(iii)* uncertainties related to inversion of inland data to the predominantly offshore seismogenic zone.

We observed that slip-deficit accumulated during the last interseismic phase is diagnostic of the location of highest slip. However, it is a poor indicator for constraining the lateral extent of a future rupture, and thus the magnitude of future earthquakes. Extending the analysis of spatial correlation back in time (beyond the latest interseismic period) could result in an improvement of this correlation, although the optimal number of previous interseismic cycles for correlation with the latest coseismic slip is very likely linked to the particular configuration of the asperities such as size, inter-asperity distance, and friction (Kaneko et al., 2010). In the long-term perspective, asperities tend to synchronize if conditions for interaction are met (Kaneko et al., 2010; Corbi et al., 2017) leading the system to produce giant earthquakes. Unfortunately, the stage of the synchronization of a given megathrust is unknown and, most importantly, if and how the pattern of asperities evolves thought subsequent seismic cycles is still debated (Park & Mori, 2007). Another limiting factor of the slip-deficit analysis is that it does not hold any information about the timing of future earthquakes.

Our analysis showed that analog-earthquake prediction requires the reconstruction of the spatio-temporally complex loading history. This reconstruction is implicitly done by ML using the training data. We thus showed that 10 seismic cycles of dense-network observations are required for training the algorithm to maximize its predictive accuracy, and that the training length may be decreased when predicting the TTF of a specific region. It is possible

that these results may be analog-model dependent and that a larger number of cycles might be necessary for training more complicated models that incorporate frictional, rheological, and geometrical variations. However, even a low complexity model such as ours, creates variations in magnitude and recurrence time making the system "realistically unpredictable". Using ML, we were able to counterbalance this unpredictability by training the algorithm to recognize patterns that are difficult to detect by human inspection.

Given that we have shown predictability in an already fairly realistic model, there is hope that similar approaches could be developed for real subduction zones. Currently, the available geodetic record is far shorter than the number of cycles needed by our approach on the analog model, although in nature there are shorter cycles of smaller-magnitude events. This leads to two straightforward paths of future development: *(i)* to develop ML approaches on simulated data (analog or numerical) that can be trained on a variety of models to predict the time to failure on a range of spatio-temporal scales, and *(ii)* to tailor ML approaches at natural subduction zone observatories with the aim of predicting also smaller-magnitude events or the whole spectrum of subduction seismicity including inter-, intra- and upper-plate earthquakes. Additionally, there is considerable scope for investigating further features beyond the ones used in this study as they might not be complete enough for natural cases where e.g. strength variations might occur due to pore fluid pressure variation or where foreshocks, slow earthquakes, or mantle relaxation additionally influence the loading history.

Future work can focus on engineering features and approaches that are more likely to succeed both in lab and nature. Comparable or better predictions can be likely obtained using a fraction of the data currently needed for training with a wider fan of features. Accordingly, we tested the two most important features in a linear implementation (linear regression) of the prediction of TTF (Supplementary Information S4; Fig. S11). The linear implementation, while producing good fits in some cycles, also produced a drift in the prediction of later cycles and unphysical negative TTF values. From this limited investigation, it remains unclear if a linear model trained on a small subset of features could improve upon the success of the GBRT.

Alternatively, changes in the framing of the problem may lead to other types of predictions: This study frames the regression to the time to failure since we choose to first demonstrate that ML recognizes a specific deformation pattern as indicative of a given stage of the interseismic period. Other framings, however, such as the classification of earthquake imminence (e.g., binary classification distinguishing if an event will happen in the next seconds), or regression to the future surface velocity, are all yet to be explored. Another possibility is treating surface deformation data as an image analysis problem (e.g., using Convolutional Neural Networks) and directly training the algorithm to predict the upcoming slip area.

Securing the society against natural disasters is a social priority. In this era of "big data" in earthquake science, advanced data analysis methods, such as ML, could open the door to a situation where earthquakes are increasingly predictable. We have shown that the application of ML to surface velocities at plate boundaries is a promising avenue for further research towards this goal.

**Acknowledgments**

We thank D. Bonacorsi and F. Grigoli for initial suggestions about Machine Learning; In figure 3 we used the perceptually uniform colormap Davos by F. Crameri. We also thank two anonymous reviewers and the editor G. Hayes for their constructive comments. **Funding:** M.R. has benefitted from the inspiring environment of the CRC 1114 “Scaling Cascades in Complex Systems” (funded by Deutsche Forschungsgemeinschaft (DFG)). J.B has is grateful for his funding from DFG Pescados Project grant number MO-2310/3; **Author**

**contributions:** F.C. J.B. and F.F. coordinated the work and wrote the manuscript. F.C. and S.B. carried out the experiments. F.C. L.S. and J.B. carried out the Machine Learning investigation and the statistical analysis. M.R. and S.L. contributed to the statistical analysis and interpretation of the results. All authors have read and revised the manuscript, contributed ideas and input to the research; **Competing interests:** Authors declare no competing interests; and **Data and materials availability:** All data and materials used in the analysis is available through the EPOS data repository.

## References


Ando, M. (1970). Source mechanisms and tectonic significance of historical earthquakes along the Nankai trough, Japan, *Tectonophysics* **27**, 119–140.

Avouac, J. P. (2015). From Geodetic Imaging of Seismic and Aseismic Fault Slip to Dynamic Modeling of the Seismic Cycle. *Annual Review Earth Planetary Sciences* **43**, 1 - 38.

Barhrnard W. *et al.* (2016). Coseismic slip and early afterslip of the 2015 Illapel, Chile, earthquake: Implications for frictional heterogeneity and coastal uplift. *Journ. Geoph. Res. Solid Earth.* , 6172-6191.

Bookhagen, B., Echtler, H., Melnick, D., Strecker, M. & Spencer, J. (2006). Using uplifted Holocene beach berms for paleoseismic analysis on the Santa María Island, south-central Chile. *Geophysical Research Letters* **33,** L15302.

Comte, D. & Pardo, M. (1991). Reappraisal of great historical earthquakes in the northern Chile and southern Peru seismic gaps. *Natural Hazards* **4**, 23–44.

Corbi F. *et al.* (2013). The seismic cycle at subduction thrusts: 1. Insights from laboratory models. *J. Geophys. Res.* **118**, 1–19.

Corbi, F., Funiciello, F., Brizzi, S., Lallemand, S. & Rosenau, M. (2017). Control of asperities size and spacing on seismic behavior of subduction megathrusts. *Geophysical Research Letters* **44**, 8227–8235.

Corbi, F. et al., (2019) Supplementary material to "Machine Learning can predict the timing and size of analog earthquakes". GFZ Data Services. http://doi.org/10.5880/fidgeo.2018.071.

Cubas, N., Avouac, J. P., Souloumiac, P. & Leroy, Y. (2013). Megathrust friction determined from mechanical analysis of the forearc in the Maule earthquake area. *Earth Planetary Science Lett*ers **381**, 92-103.

DeVries, P., Viegas, F., Wattenberg, M., & Meade B. (2018). Deep learning of aftershock patterns following large earthquakes. **Nature** 560. doi.org/10.1038/s41586-018-0438-y.

Friedman, J. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of statistics*. 1189-1232.

Goldfinger, C., Nelson, C. H. & Johnson, J. E. (2003). Holocene earthquake records from the Cascadia subduction zone and northern San Andreas fault based on precise dating of offshore turbidites. *Annu. Rev. Earth Planet. Sci.* **31**, 555–77.

Heki, K. & Mitsui, Y. (2013). Accelerated pacific plate subduction following interplate thrust earthquakes at the Japan trench, *Earth Planetary Science Letters*, **363**, 44–49.

Kanamori, H. & Brodsky, E. E. (2010). The Physics of Earthquakes. *Physics Today*. **67**, 34.

Sieh, K. *et al.* (2008). Earthquake Supercycles Inferred from Corals of West Sumatra, *Science.* **322**, 1674–1678.

Kaneko, Y., Avouac, J. P. & Lapusta N. (2010). Towards inferring earthquake patterns from geodetic observations of interseismic coupling. *Nature Geoscience* **3**, 363–369.

Kato, A. *et al.* (2012). Propagation of Slow Slip Leading Up to the 2011 Mw 9.0 Tohoku-Oki Earthquake. *Science.* **335**, 705–708.

Konca O. *et al.* (2008). Partial rupture of a locked patch of the Sumatra megathrust during the 2007 earthquake sequence. *Nature*. **456**, 631–5.
Kulkarni, R., Wong, I., Zachariasen, J., Goldfinger, C. & Lawrence, M. (2013). Statistical analyses of great earthquake recurrence along the Cascadia subduction zone. *Bullettin Seismological Society American* **103,** 6.
Lallemand, S., Peyret, M., van Rijsingen, E., Arcay, D. & Heuret, A. (2018). Roughness characteristics of oceanic seafloor prior to subduction in relation to the seismogenic potential of subduction zones. *Geochemistry Geophysics Geosystems* 19, 1–26. https://doi.org/10.1029/2018GC007434.
Lay, T. & Kanamori, H. (1981). An asperity model of large earthquake sequences. *Earthquake Prediction* **4,** 579 - 592.
Lorito S. *et al.*, (2011). Limited overlap between the seismic gap and coseismic slip of the great 2010 Chile earthquake. *Nature Geoscience* **4**, 173–177.
Loveless, J. P. & Meade, B. J. (2011). Spatial correlation of interseismic coupling and coseismic rupture extent of the 2011 Mw = 9.0 Tohoku-oki earthquake. *Geophysical Research Letters* **38**, 2–6.
Loveless, J. P. & Meade, B. (2016). Two decades of spatiotemporal variations in subduction zone coupling offshore Japan. *Earth Planetary Science Letters* 436, 19-30.
Mavrommatis, A., Segall, P., Johnson, K. M. (2014). A decadal-scale deformation transient prior to the 2011 Mw 9.0 Tohoku-oki earthquake, *Geophisical Research Letters*. 4486–4494, doi:10.1002/2014GL060139.
McGuire, J. J. & Plank, T. *et al.* (2017). The SZ4D Initiative: Understanding the Processes that Underlie Subduction Zone Hazards in 4D. *Vision Document Submitted to the National Science Foundation*. (The IRIS Consortium), 63 pp.
Melnick, D., Cisternas, M., Moreno, M. & Norambuena, R. (2012). Estimating coseismic coastal uplift with an intertidal mussel: calibration for the 2010 Maule Chile earthquake (Mw = 8.8). *Quat. Sci. Rev.* **42**, 29–42.
Métois M., Vigny C. & Socquet, A. (2016). Interseismic coupling, megathrust earthquakes and seismic swarms along the chilean subduction zone (38°–18°S). *Pure Applied Geophysics* **173**, 1431–1449.
Moreno, M., Bolte, J., Klotz, J. & Melnick, D. (2009). Impact of megathrust geometry on inversion of coseismic slip from geodetic data: Application to the 1960 Chile earthquake. *Geophysical Research Letters* **36**, 1–5.
Moreno, M., Rosenau, M. & Oncken, O. (2010). Maule earthquake slip correlates with pre-seismic locking of Andean subduction zone. *Nature*. **467**, 198–202.
Moreno, M. *et al.* (2011). Heterogeneous plate locking in the South – Central Chile subduction zone : Building up the next great earthquake. *Earth Planetary Science Letters* **305**, 413–424.
Park, S. C. & Mori, J. (2007) Are asperity patterns persistent? Implication from large earthquakes in Papua New Guinea. *Journal Geophysical Research* **112**, 1–16.
Reid, H. F. (1906). On mass-movements in tectonic earthquakes, in *The California Earthquake of April 18, 1906: Report of the State Earthquake Investigation Commission* (Carnegie Inst., Washington, D. C., 1910), pp. 29–32.
Rosenau, M., Corbi, F. & Dominguez, S. (2017). Analogue earthquakes and seismic cycles: experimental modelling across timescales. *Solid Earth*. **8**, 1–12.
Rouet-Leduc, B. *et al.* (2017). Machine Learning predicts laboratory earthquakes. *Geophysical Research Letters* **44**, 9276–9282.
Rouet-Leduc, B. et al. (2018a). Estimating Fault Friction From Seismic Signals in the Laboratory. *Geophysical Research Letters*. **45-3**, 1321-1329. doi: 10.1002/2017GL076708.

Rouet-Leduc, B., Hulbert, C & Johnson, P. A. (2018b). Breaking Cascadia's Silence: Machine Learning Reveals the Constant Chatter of the Megathrust https://arxiv.org/abs/1805.06689.
Satake, K. & Atwater, B. F. (2007). Long-Term Perspectives on Giant Earthquakes and Tsunamis at Subduction Zones. *Annual Review Earth Planetary Sciences* **35**, 349–374.
Schurr, B. *et al.* (2014). Gradual unlocking of plate boundary controlled initiation of the 2014 Iquique earthquake. *Nature*, **512**, 299 – 302.
Shimazaki, K. & Nakata, K. (1980). Time-predictable recurrence model for large earthquakes, *Geophisical Research Letters* **7**, 4.
Strasser, F. O., Arango, M. C. & Bommer, J. J. (2010) Scaling of the Source Dimensions of Interface and Intraslab Subduction-zone Earthquakes with Moment Magnitude. *Seismological Research Letters* **81**, 951–954.
Subarya, C. *et al.* (2006). Plate-boundary deformation associated with the great Sumatra-Andaman earthquake. *Nature*. **440**, 46–51.
Sveen, J. K. (2004). An introduction to MatPIV v.1.6.1 Eprint no. 2, ISSN 0809-4403, Dep. of Mathematics, University of Oslo.
Tsang, L. L. H., Meltzner, A. J., Hill, E. M., Freymueller, J. T. & Sieh, K. (2015). A paleogeodetic record of variable interseismic rates and megathrust coupling at Simeulue Island, Sumatra. *Geophysical Research Letters* **42**, 10585–10594.

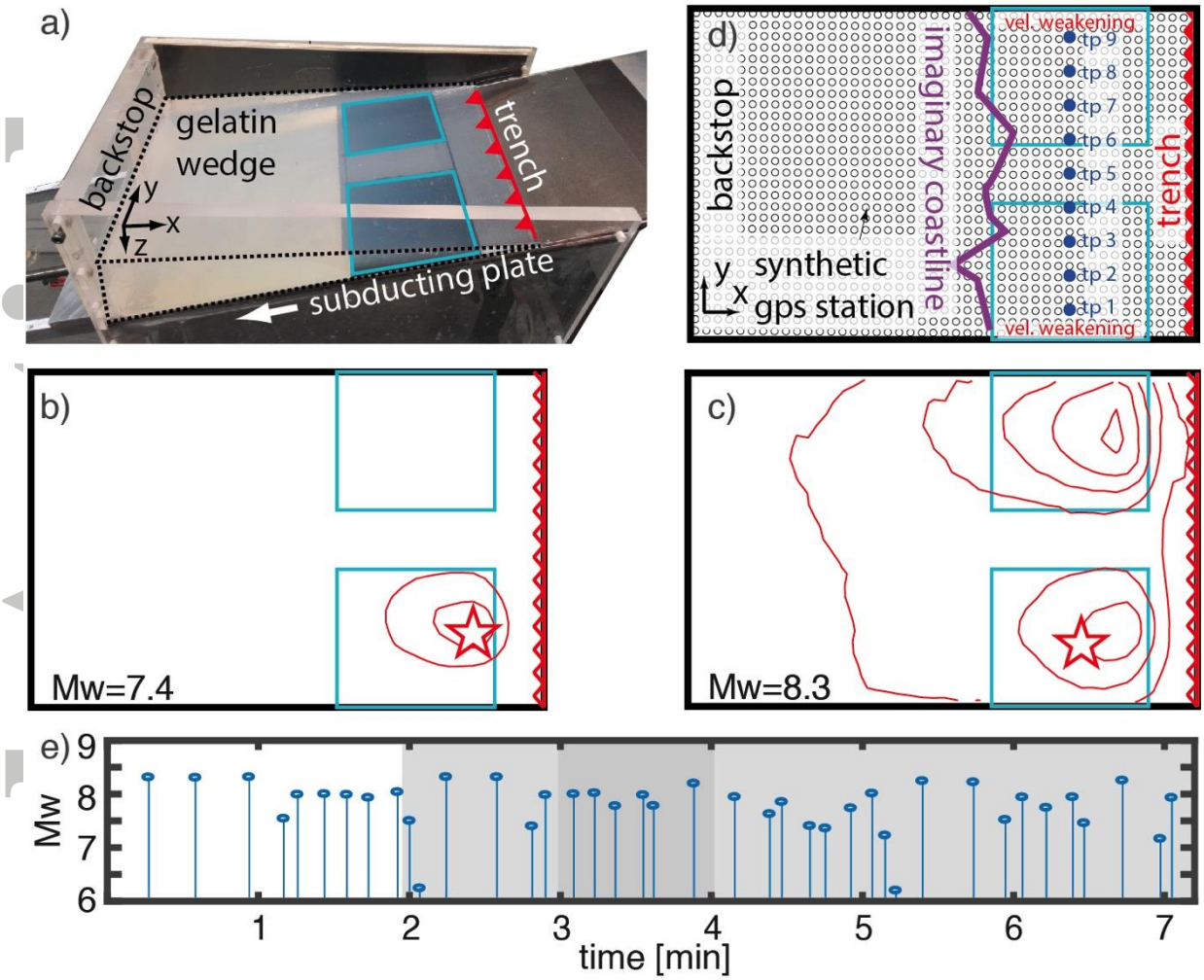


**Figure 1:** Analog model of megathrust earthquakes. Panel a: Oblique view of the experimental setup. Cyan rectangles highlight the two velocity weakening patches. The analog model produces spontaneous ruptures that involve either one (panel b) or both velocity weakening zones (panel c). Ruptures are characterized in terms of apparent slip (i.e., coseismic surface displacement; red contour with interval of 0.15 mm). White/red stars in panels b and c represent the epicenters. Experimental monitoring is performed with a video camera from top view (orthogonal to the x-y plane); image analysis allows dense resolution of the velocity field (white points) also above the offshore seismogenic zone (panel d). The nine blue points highlight target stations used for the Machine Learning analysis. Panel e: Time series of analog earthquakes magnitude. The light and dark gray shading highlight the fraction of the experiment shown in figure 3 panel a - b and the zoom of panel c, respectively.

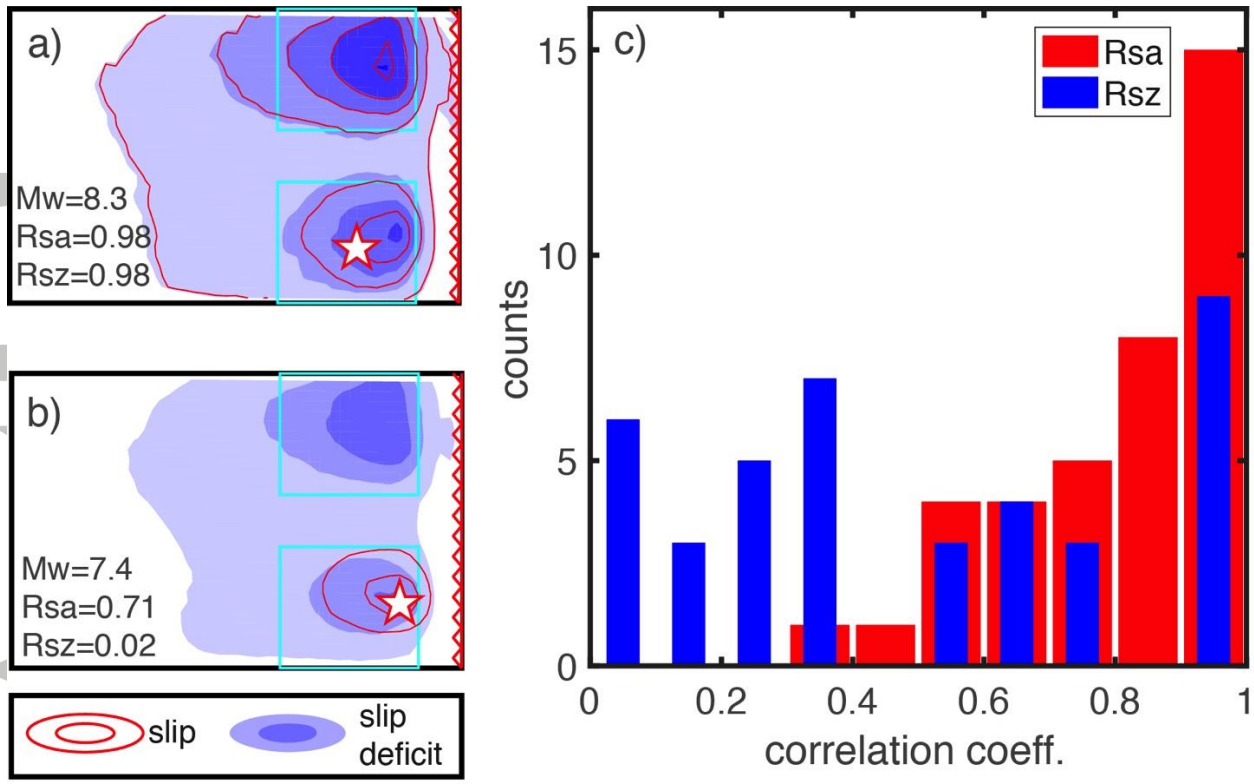


**Figure 2:** Spatial correlation between apparent slip and slip-deficit. Example of seismic cycles where the slip correlates well with slip-deficit both along the whole margin and within the slip area (panel a) and where slip correlates well with slip-deficit only within the slip area (panel b). Slip and slip-deficit maps have the same contour interval of 0.15 mm. The white/red stars represent the epicenters. Histogram quantifying the number of seismic cycles with given correlation coefficient (panel c).

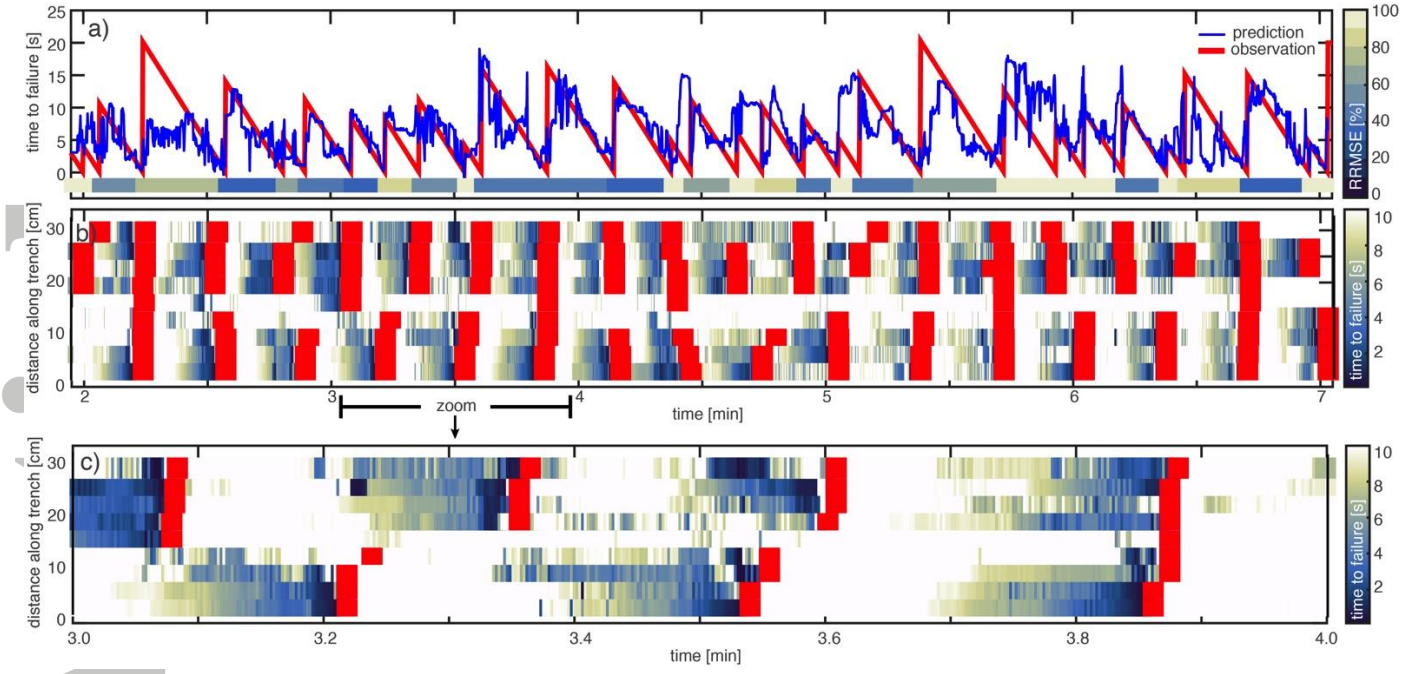


**Figure 3:** Machine learning results. Panel a: TTF versus experimental time. The red and blue time series refer to observation and prediction, respectively. The color-coded horizontal line highlight the performances of the prediction cycle by cycle quantified by means of relative root mean squared error RRMSE. Dark blue RRMSE colors highlight cycles with a good fit. Panel b: Space-time evolution of the predicted time to failure calculated for the 9 target points. The blue shading highlights the prediction of an upcoming event at a given location. The red squares indicate the spatio-temporal distribution of observed analog earthquakes (width of red squares not scaling with time). Panel c: one minute long zoom of panel b. Synchronous existence of low time to failure values for several target points along the margin suggests that a large magnitude earthquake is impending.