UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE GEOFÍSICA

RELACIÓN ENTRE COMPORTAMIENTO SISMOGÉNICO DEL MEGATHRUST Y PARÁMETROS DE SUBDUCCIÓN: ANÁLISIS ESTADÍSTICO GLOBAL MEDIANTE INTELIGENCIA ARTIFICIAL

Antonio Lucas Crisosto Urrutia

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Concepción para optar al grado académico de Magíster en Geofísica

Marzo 2024
Concepción, Chile

Profesor Guía: Dr. Andrés Tassara Oddo
Comisión evaluadora:
Dr. Matthew Miller - Dr. Joaquín Cortés - Dr. Sergio Ruiz

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE GEOFÍSICA

RELACIÓN ENTRE COMPORTAMIENTO SISMOGÉNICO DEL MEGATHRUST Y PARÁMETROS DE SUBDUCCIÓN: ANÁLISIS ESTADÍSTICO GLOBAL MEDIANTE INTELIGENCIA ARTIFICIAL

Antonio Lucas Crisosto Urrutia

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Concepción para optar al grado académico de Magíster en Geofísica

Profesor Guía: Dr. Andrés Tassara
Comisión evaluadora:
Dr. Sergio Ruiz - Dr. Joaquín Cortés - Dr. Matthew Miller

Marzo 2024
Concepción, Chile

A mi familia.

AGRADECIMIENTOS

En primer lugar, quiero agradecer a mi familia, Javi y amigos; de quienes siempre recibí el apoyo necesario para en este proceso. A José e Isa quienes me instaron a dar el paso en esta etapa de mi formación.

A Andrés Tassara, un pilar fundamental en esta tesis y a lo largo del trayecto del magíster; por su orientación, inspiración y amistad, mis agradecimientos.

A la gente de ISTerre, quienes contribuyeron importantemente en este trabajo durante la estadía en Grenoble: Diego Molina, Giuseppe Constantino y Anne Socquet. Gracias por el gran recibimiento.

A la gente del DGEO, docentes, administrativos y compañeros, por recibirme, educarme y apoyarme siempre con la mejor voluntad.

Finalmente, al Núcleo Milenio Cyclo por su continuo apoyo y financiamiento a lo largo de este trabajo .

Resumen

En este estudio investigamos la relación entre el comportamiento sismogénico de megathrusts globales y diferentes parámetros de subducción. Realizamos un enfoque paramétrico implementando tres algoritmos de Machine Learning (ML) basados en árboles de decisión (XGBoost, GradientBoosting y CatBoost) para predecir el b-value de la relación frecuencia-magnitud de la sismicidad como una combinación no lineal de variables de subducción (edad y rugosidad de la placa subductante, ángulo de buzamiento de la subducción, velocidad y azimut de la convergencia, distancia a la dorsal y límite de placa más cercanos). Utilizando SHAP values para interpretar los resultados de ML, observamos que la edad de la placa y el ángulo de subducción poco profundo son las variables más influyentes. Los resultados sugieren que las placas más antiguas y con buzamientos menos profundos contribuyen a bajos b-values, lo que indica un mayor stress en el megathrust. Este patrón se atribuye a la mayor rigidez de las placas más antiguas, que aumentan la resistencia a la flexión y generan un ángulo de penetración poco profundo, aumentando el área de fricción entre placas e intensificando el estado de stress del megathrust. Estos resultados ofrecen nuevas perspectivas sobre la complejidad no lineal del comportamiento sísmico a escala global.

Keywords – Parámetros de subducción, Megathrust, b-value, Machine Learning

Abstract

We investigate the relationship between the seismogenic behavior of global megathrusts and various subduction parameters. We performed a parametric approach by implementing three decision tree-based Machine Learning (ML) algorithms (XGBoost, GradientBoosting and CatBoost) to predict the b-value of the frequency-magnitude relationship of seismicity as a non-linear combination of subduction variables (subducting plate age and roughness, subduction dip angle, convergence speed and azimuth, distance to closest ridge and plate boundary). Using the SHAP value to interpret the ML results, we observe that plate age and shallow subduction dip are the most influential variables. The results suggest that older, shallow-dipping plates contribute to low b-values, indicating higher megathrust stress. This pattern is attributed to the higher rigidity of older plates, increasing flexural strength and generating a shallow penetration angle, increasing the frictional interplate area and intensifying the megathrust stress. These findings offer new insights into the non-linear complexity of seismic behaviour on a global scale.

Keywords – Subduction parameters, Megathrust, b-value, Machine Learning

Índice general

Índice de cuadros

Índice de figuras

Capítulo 1

Introducción

1.1. Introducción

Los grandes terremotos se producen en los límites de las placas convergentes a lo largo de la zona sismogénica del megathrust de subducción. Las propiedades físicas de la interfaz de subducción varían según la región y afectan al estado de stress que, a su vez, influye en el comportamiento sismogénico (Nishikawa and Ide, 2014). Este comportamiento sismogénico hace referencia a la distribución espacio-temporal de la sismicidad para un tiempo o lugar en particular. Para caracterizar el comportamiento de la sismicidad en una cierta región diferentes proxies han sido usados en la literatura, tales como la máxima magnitud registrada o la tasa de sismicidad (e.g., Ruff and Kanamori, 1980; Brizzi et al., 2018; Heuret et al., 2012). Un parámetro también relevante para la caracterización del comportamiento sismgenico es el *b-value* de la ley de Gutenberg-Richter (Gutenberg and Richter, 1944). Experimentos de laboratorio sugieren que el estado de *stress* y el *b-value* tienen una correlación negativa (Scholz, 1968; Schorlemmer et al., 2005). Según la definición del *b-value*, zonas que tienen valores bajos tienen más probabilidades de estar asociadas a grandes terremotos que las que tienen valores altos (Brodsky, 2019). Distintos autores han reportado diferencias globales en este parámetro, lo que refleja cambios en el estado de *stress* en el *megathrust* a lo largo de distintas zonas de subducción (e.g., Carter and Berg, 1981; Nanjo et al., 2012; Kagan and Jackson, 2013; Nishikawa and Ide, 2014). Las causas exactas de estas variaciones del *stress* siguen siendo inciertas (Dielforder et al., 2023). No obstante, varios estudios, a través de distintos enfoques paramétricos, han intentado dilucidar los

factores que influyen en el estado de *stress* y, por tanto, en el comportamiento sismogénico (e.g., Heuret et al., 2012; Schellart and Rawlinson, 2013). Por una parte, estudios pioneros (e.g., Ruff and Kanamori, 1980; Kanamori, 1983) han sugerido que la magnitud máxima esperada de un terremoto está dada por rangos favorables de la tasa de convergencia y la edad de la placa subductante. Sin embargo, esta suposición no conversa con la sismicidad documentada en el siglo XXI (Stein and Okal, 2007). No obstante, Scholz (2015) y Nishikawa and Ide (2014) han hallado notables correlaciones entre el estado de *stress* y la edad de la placa. Estos resultados sugieren que una menor edad de la placa en subducción estaría asociada a una mayor flotabilidad, lo que genera un mayor *stress* normal en la placa superior. Por otro lado, Brizzi et al. (2018), a partir de un análisis estadístico multivariante determinó que la extensión paralela a la fosa y el espesor sedimentos en la fosa proporcionarían la mayor capacidad para discriminar entre áreas que han experimentado terremotos de alta y baja magnitud. Trabajos previos han sido principalmente basados en el reconocimiento y cuantificación de posibles correlaciones mediante regresiones lineales entre el *b-value* y parámetros de subducción. Sin embargo, la relación real entre estos parámetros y el comportamiento sismogénico es probablemente no lineal, lo que justifica la aplicación de algortimos de *Machine Learning* (ML). Los métodos basados en ML han sido recomendados para entender la interdependencia no lineal entre varios factores posibles que actúan para controlar un determinado proceso, tal como el comportamiento sismogénico de diferentes zonas (e.g., Saeed et al., 2011; Jones et al., 2020; Xiong et al., 2021; Corbi et al., 2019). El Aprendizaje Automático o *Machine Learning* es el estudio de algoritmos capaces de mejorar o aprender automáticamente a través de la experiencia y el uso de datos (Mitchell, 1997). Los enfoques más comunes del aprendizaje automático son el supervisado y el no supervisado. En el aprendizaje supervisado, los algoritmos son capaces de generalizar o aprender en base a datos con etiqueta, ya sea para problemas de clasificación o regresión (Hastie et al., 2001). Por otro lado, el aprendizaje no supervisado intenta extraer características y patrones desde datos sin etiquetar (Hastie et al., 2001). El trabajo de Schäfer and Wenzel (2019) destaca entre estos métodos para el entendimiento del comportamiento de la sismicidad a través de la aplicación de ML no supervisado. En su estudio, se intenta clusterizar zonas que podrían experimentar una misma máxima magnitud en base a la entrada de parámetros de subducción y similitud entre segmentos según diferentes propiedades.

Sin embargo, a pesar de las ventajas de los algoritmos de ML en general, en la comunidad científica estos métodos son criticados por su naturaleza de "caja-negra" (Salas et al., 2022), constituyendo la principal barrera para adoptar estos métodos en otras áreas (Molnar, 2020). Por esta razón, se hace necesario incorporar un análisis de interpretabilidad de las predicciones para generar confianza en los resultados del modelo (Molnar, 2022; Salas et al., 2022). Dentro de los métodos comunes de interpretabilidad en la literatura destaca *Shapley Additive Explanations* (SHAP). Esta metodología, desarrollada por Lundberg and Lee (2017), está basada en conceptos de teoría de juegos cooperativos lo cual permite que los valores SHAP (Shapley et al., 1953) puedan dar una distribución justa de importancia a las variables de entrada dependiendo de sus contribuciones individuales. En (Lundberg and Lee, 2017) se definen los valores SHAP como una medida unificada de la importancia de variables, que son la solución a la ecuación del valor de *Shapley*. Dada una instancia del modelo generado por algún algoritmo de ML, el objetivo es descomponer la predicción y asignar los valores SHAP a las variables individuales de la instancia. De esta forma, los valores SHAP pueden medir las contribuciones de las variables de entrada a la predicción del modelo y asignar distintos niveles de importancia en las variables de entrada (Rozemberczki et al., 2022). En esta investigación, se llevaron a cabo mediciones de diversos parámetros de subducción a lo largo de 157 transectas que cubren diferentes la mayoría de las zonas de subducción a escala global. El objetivo principal de estas mediciones fue caracterizar ciertas propiedades geodinámicas de las zonas de subducción para posteriormente evaluar el impacto combinado de las variables en el *stress* del *megathrust*, representado por el *b-value*. Para ello, se utilizaron diversos algoritmos de ML supervisado de regresión (*XGboost*, *GradientBoosting* y *CatBoost*) con el fin de analizar patrones entre las variables de entrada y el *b-value* estimado en cada una de las transectas. Posteriormente, se llevó a cabo una interpretación de los modelos de ML generados utilizando los SHAP *values*. Este enfoque permitió comprender las tendencias de las variables y su contribución a la predicción del *b-value*, contribuyendo así a una comprensión integral de los procesos que regulan el comportamiento sismogénico en los *megathrust* de subducción.

1.2. Hipótesis y objetivos

El comportamiento sismogénico del *megathrust* es controlado, de manera combinada, por factores cinemáticos y geométricos asociados al proceso de subducción.

1.2.1. Objetivo general

Determinar el grado de influencia de parámetros cinemáticos y geométricos de subducción en el comportamiento sismogénico para segmentos a escala global.

1.2.2. Objetivos específicos

• Caracterizar el comportamiento sismogénico de cada segmento de subducción a través del parámetro *b-value* de la Ley de Gutenberg-Richter.

• Analizar el grado de correlación entre variables medidas y el *b-value* para cada segmento con la aplicación de algoritmos de *Machine Learning*.

• Construir un modelo conceptual sobre el fenómeno físico asociado a la correlación entre los parámetros medidos.

Capítulo 2

Relating megathrust seismogenic behavior and subduction parameters via Machine Learning at global scale.

2.1. Abstract

We investigate the relationship between the seismogenic behavior of global megathrusts and various subduction parameters. We performed a parametric approach by implementing three decision tree-based Machine Learning (ML) algorithms (XGBoost, GradientBoosting and CatBoost) to predict the b-value of the frequency-magnitude relationship of seismicity as a non-linear combination of subduction variables (subducting plate age and roughness, subduction dip angle, convergence speed and azimuth, distance to closest ridge and plate boundary). Using the SHAP value to interpret the ML results, we observe that plate age and shallow subduction dip are the most influential variables. The results suggest that older, shallow-dipping plates contribute to low b-values, indicating higher megathrust stress. This pattern is attributed to the higher rigidity of older plates, increasing flexural strength and generating a shallow penetration angle, increasing the frictional interplate area and intensifying the megathrust stress. These findings offer new insights into the non-linear complexity of seismic behaviour on a global

scale.

2.2. Introduction

The largest earthquakes on Earth occur at convergent plate boundaries along the seismogenic zone of subduction megathrust. The physical properties of the subduction zones vary according to the region and affect the stress state which, in turn, influences the seismogenic behavior of these zones (Nishikawa and Ide, 2014). To characterize the stress state, different proxies have been used in the literature, such as the maximum recorded magnitude, the seismicity rate or the b-value of the Gutenberg-Richter law (Gutenberg and Richter, 1944). Regarding this latter, laboratory experiments and natural examples suggest that the stress state and the b-value have a negative correlation (Scholz, 1968; Schorlemmer et al., 2005). Several authors have reported global differences in this parameter, reflecting variations in the stress state along the megathrust (e.g., Carter and Berg, 1981; Nanjo et al., 2012; Kagan and Jackson, 2013; Nishikawa and Ide, 2014). Furthermore, several studies have attempted to clarify the factors that influence the stress state and thus the seismogenic behaviour of the megathrust (e.g., Heuret et al., 2012; Schäfer and Wenzel, 2019; Brizzi et al., 2018). Pioneering studies (Ruff and Kanamori, 1980; Kanamori, 1983) have suggested that the maximum expected magnitude of an earthquake is determined by favourable ranges of convergence rate and age of the subducting plate. However, this assumption would be inconsistent with the seismicity documented during the XXI century (Stein and Okal, 2007). On the other hand, Nishikawa and Ide (2014) and Scholz (2015) have found remarkable correlations between stress levels measured by the b-value and both plate age and slab pull force, as a function of slab age, respectively. These results suggest that a younger subducting plate would be associated with a higher buoyancy, which generates a higher normal stress on the upper plate and therefore a lower b-value. Previous works have been mainly based on the recognition and quantification of possible correlations via linear regression between b-value and subduction parameters. However, the actual relationship between these parameters and the seismogenic behavior is likely non-linear which justifies the implementation of Machine Learning (ML) algorithms. ML-based methods has been recommended to understand the nonlinear interdependence between several possible factors acting to control a given processes, such as the seismogenic behaviour of different

areas (e.g., Saeed et al., 2011; Jones et al., 2020; Xiong et al., 2021; Corbi et al., 2019). Among these methods, the work of Schäfer and Wenzel (2019) stands out, where an attempt is made to cluster zones of maximum magnitude based on input of subduction parameters and similarity between segments according to different properties. In this research, detailed measurements of various subduction parameters and *b*-value were carried out for each of our 157 transects (Fig. 2.3.1a) that cover most of the subduction zones worldwide. The main purpose of these measurements was to evaluate the combined impact of these variables on the stress of the megathrust, represented by the *b*-value. For this purpose, three supervised regression ML algorithms were used to analyse patterns among the input variables and, consequently, to predict the *b*-value. Subsequently, an interpretation of the generated ML models was carried out using SHAP values (Lundberg and Lee, 2017). This approach allowed to understand the trends of the variables and their contribution to the prediction of the *b*-value, thus contributing to a comprehensive understanding of the processes that regulate the stress state in the megathrust.

2.3. Data and Methods

2.3.1. Quantification of geometric and kinematic parameters of subduction zones

We measured different parameters for each of the 157 transects perpendicular to the trench (Fig. 2.3.1a). Convergence velocity (vc_10 in Fig.2.4.4 and azimuth angle (and_conv in Fig. 2.4.4) were derived from the plate kinematics model of Müller et al. (2016), which reconstructs the kinematics of plate tectonics from the Jurassic to the present using GPlates. Subducting plate age at the trench (age in Fig. 2.4.4) is also derived from this model. Seafloor roughness was derived from the GEBCO bathymetric model Lemenkova (2020). To quantify the roughness, the standard deviation of the bathymetry along a transect perpendicular to the trench was calculated oceanward (roughness in Fig. 2.4.4). To measure the distance along the trench between transects and both the plate edge and the nearest ridge (Dse and Dcr in Fig. 2.4.4), ArcGIS Pro software was implemented directly with its base map as a reference. Finally, the subduction angle between 0 and 60 km depth (ang_60 in Fig. 2.4.4) was obtained from the Slab2.0 model of Hayes et al. (2018).

2.3.2. Estimation of *b*-value

We use the seismicity catalogue provided by the International Seismological Center (ISC) between years 1900 and 2022. To estimate one *b*-value for each studied transect we consider earthquakes with epicentres located at a maximum distance of 200 km perpendicular to the transect´s azimuth. We consider a 25 % overlap between each transect in order to capture the spatial variability of seismic activity (Fig. 2.3.1b). We also tested applying no overlap between transects. Four sub-catalogues were then created for each transect considering either all the recorded events or earthquakes located around the slab upper surface at depths between ±5, ±10 and ±15 km of the Slab2.0 model (Hayes et al., 2018, see Figure 2.3.1c). From these sub-catalogues, magnitude differences between events were calculated and the *b*-value was estimated using the b-positive method proposed by van der Elst (2021). This method, which follows the same form as the maximum likelihood estimator (Aki, 1965), only considers positive magnitude differences to avoid incompleteness problems and the sequences of aftershocks associated with the seismic catalogue. After exploring different criteria for estimating the *b*-value, we decided to show results considering 25 % overlap between transects and considering events within ±10 km of the slab.

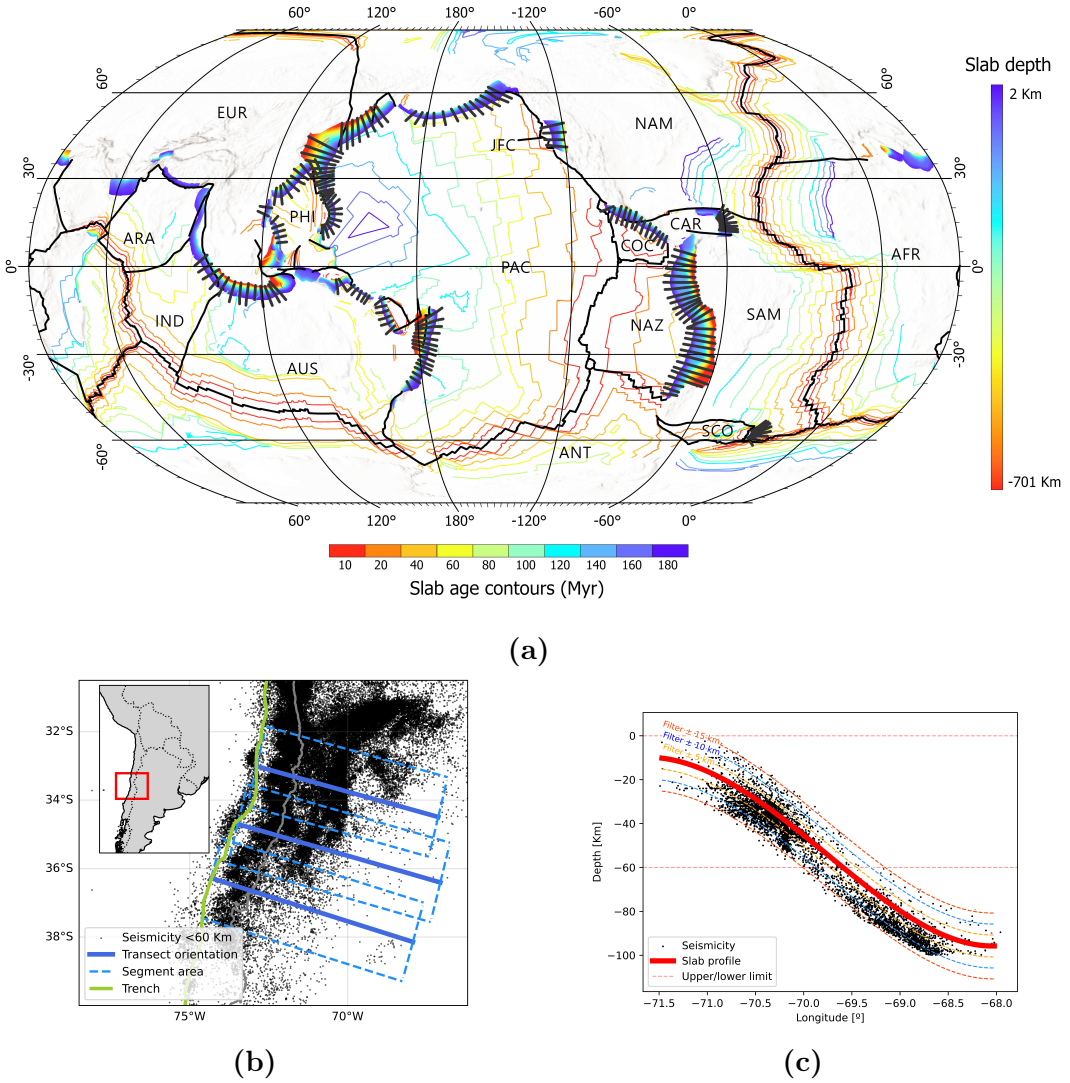

Figura 2.3.1: Distribution of transects perpendicular to the trench for the quantification of subduction parameters and *b*-value. In Figure 2.3.1a, the overall distribution of profiles in major subduction zones is depicted (dark lines), based on the Slab2.0 model (Hayes et al., 2018), in addition with seafloor age contours provided by the grid of Müller et al. (2016). Figure 2.3.1b provides a close-up view of the segments from each transect along central Chile, emphasising the 25 % overlap with neighbouring segments. The estimation of the *b*-value for each transect relies on the consideration of seismicity in relation to these segments. Figure 2.3.1c illustrates an exemplary depth profile of seismicity for one of the transects. Different filters at distances of ±5, ±10, and ±15 km relative to the slab are applied in order to evaluate the variability in the *b*-value estimation.

2.3.3. Machine Learning

In this study we applied three widely used machine learning algorithms that are based on decision trees. These algorithms are implemented in Python packages as XGBoost (Chen and Guestrin, 2016), GradientBoosting (Natekin and Knoll, 2013) and CatBoost (Prokhorenkova et al., 2018). A brief description of them is provided in the Supplementary Material (Text in Apendice). Focused on regression problems, these algorithms aim to predict a target variable (b-value in our case) from a set of input features (subduction parameters). The use of three different supervised machine learning algorithms is driven by our quest for convergence in conclusions, ensuring consistency in results and strengthening the reliability of interpretations. These implemented ensemble methods focus on the sequential construction of decision trees, with the primary goal of improving the predictive accuracy of the model and capturing complex relationships in the data.

For the model's construction, the data are initially randomly split into training and test sets with a 90 % and 10 % proportion, respectively. Subsequently, a cross-validation process is carried out on the training set, with the aim of building and validating models using training and validation subsets (Fig.2.3.2, more details can be found in the supplementary material). During this phase, an optimal set of hyperparameters is determined for each algorithm, thus defining the models. Once a model with optimal performance has been built, we proceed to evaluate its performance using test data not previously seen by the model. During this evaluation, we compare the model predictions to the benchmark data using various performance metrics such as R2, RMSE and MAE (Text in Apendice). These metrics provide insight into both the percentage of variability explained by the independent variables and the absolute and root mean square differences of the mean squared errors between the model predictions and the actual benchmark values (Text in Apendice). To interpret the inner functioning of the model, the SHAP values method is then implemented. These values provide a way to explain the results of any machine learning model using a game-theoretic approach by measuring the contribution of each feature to the final outcome (Apendice). This approach examines the effect of each feature on the predicted outcomes by controlling for the presence of features, which allows us to better understand the decision-making process of the model. In other words, SHAP value allow us to quantify the influence of each feature on outcome prediction. Finally, in order

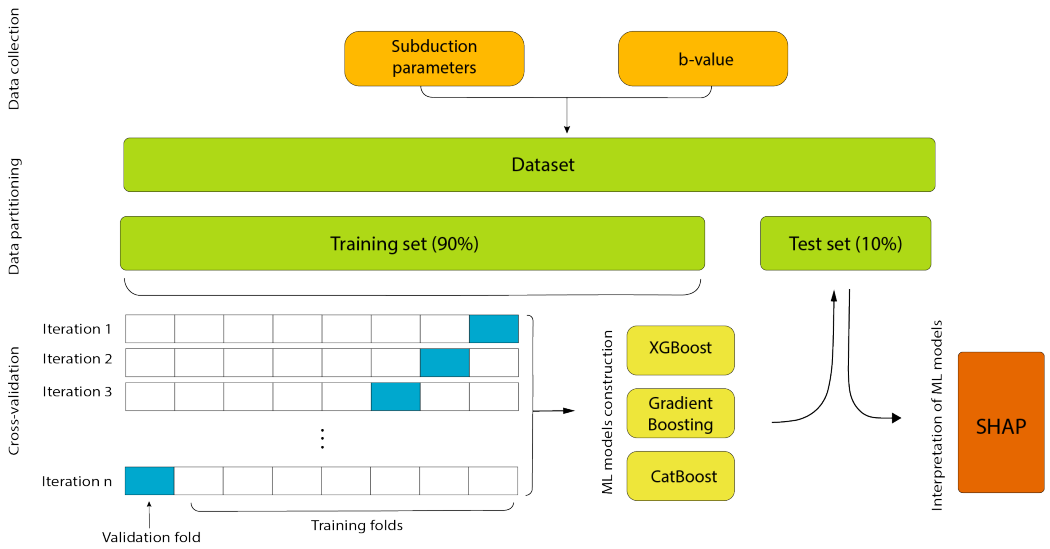

Figura 2.3.2: Methodology diagram showing the steps from data collection to the implementation of ML algorithms. The first step is to collect the essential data for the study, forming a set that is then divided into a training set (90 %) and a test set (10 %). Cross-validation is then performed, partitioning the training set into k-folds for robust evaluation. During each iteration of cross-validation, the model is trained on k-1 subsets and evaluated on the remaining one, calculating the average error of all iterations. Then, for each algorithm (XGBoost, GradientBoosting and CatBoosting), a model built with different hyperparameters is selected, the choice being based on the performance evaluated by metrics such as R2, RMSE and MAE. Next, each model is tested on the test set with previously unseen data, also evaluated by performance metrics, to truly verify the generalisability of the model created for each algorithm. Finally, the inner workings of the model are interpreted using the SHAP value.

to analyse the stability of the feature importance in the interpretation of the models, additional tests are performed using random partitions of the data in 80/20 and 70/30 training and test ratios. This approach, applied to a limited dataset of 157 observations, allows us to evaluate the robustness of the previously built models and to determine if there is a general pattern in the data or if the model is sensitive to specific partitions. Figure 2.3.2 depicts the methodology carried out throughout this study.

2.4. Results

The map in Figure 2.4.1 shows the global distribution of the estimated b-values for different subduction zones. A significant variation in the b-values is observed in different regions of the world. For the South American zone, a high variability is observed, with values close to 0.8 predominating and areas of increased b-value coincidinge with the subduction of the Juan Fernandez and Carnegie ridge. Likewise, in Cascadia, Sumatra and Aleutians, low b-values ($<$0.75) predominate, indicating high stress of the megathrust. b-values close to 1 representing moderate stress are found in the Marianas, Philippines and Tonga-Kermadec. For the Sandwich, Caribbean, Philippines and Central America zones, trends towards b-values higher than 1 are observed. A significant increase in the b value, indicating lower stress, is observed particularly for the Central American zone.

The performance of each of the three machine learning algorithms used by us, reflecting its predictive ability, was evaluated using standard metrics, including R^2, RMSE and MAE. Specific details on the metrics and performance of each algorithm are available in the supplementary material (Table 1).

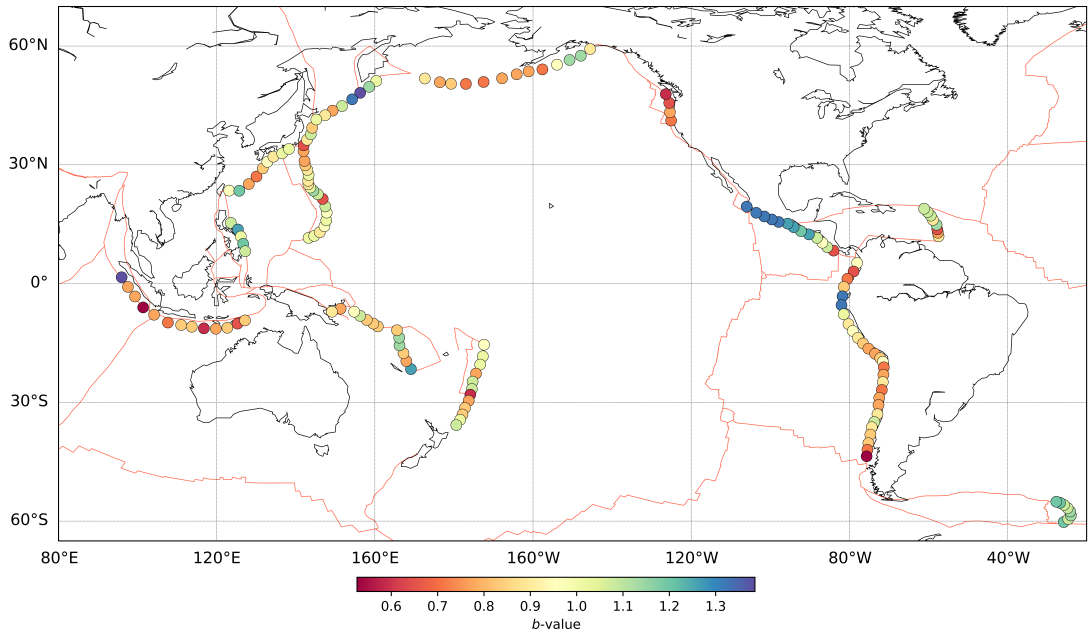

Figura 2.4.1: Computed b-values for each transect considering 25 % overlap between them and seismicity recorded within ± 10km of the slab upper surface.

The performance of each of the three machine learning algorithms used by us,

reflecting its predictive ability, was evaluated using standard metrics, including R^2, RMSE and MAE. To ensure the reliability of our models, we conducted several tests with different random partitions of the dataset, using ratios of 90/10, 80/20 and 70/30 for training and testing. Specific details on the metrics and performance of each algorithm are available in the Apendice (Table A1.1). Below, the performance of the three ML algorithms is detailed focusing on R^2 as a measure of the percentage of variability explained by the independent variables in the target variable.

Overall, at a ratio of 90/10, all three algorithms were found to have considerable predictive ability, with R^2 values of 0.83, 0.88 and 0.82 for XGBoost, GradientBoosting and CatBoost, respectively (Figure 2.4.2 and Table A1.1 in Apendice).With a training and test data ratio of 80/20, there was an overall decrease in performance, with XGBoost and GradientBoosting achieving an R^2 of 0.52, while CatBoost achieved 0.64 (Table A1.1 in Apendice).With a partition of 70/30 in the dataset, an even lower performance of the algorithms is observed with XGBoost having an R^2 of 0.51, followed by GradientBoosting ($R^2 = 0.48$) and CatBoost ($R^2 = 0.41$) (Table A1.1 in Apendice).

When interpreting the ML models using SHAP values, regardless of the algorithm and the proportion of training and test data used, a consistency in the data patterns can be seen, despite variations in performance between partitions. In Figure 2.4.4, we present the detailed interpretation of the models with SHAP values for a 90/10 partition of the data, showing in a revealing way how the input variables contribute to the prediction of the output variable. SHAP values for 80/20 and 70/30 partitions can be found in Figures 2.4.5 and 2.4.6. From the bar plots in Figures 2.4.4a, c, and e, we observe that the subduction variables having the largest impact in predicting the b-value for the three machine learning algorithms are consistently the plate age, the subduction angle (ang_60), and the distance to the closest slab edge (Dse). In both XGBoost and GradientBoosting (2.4.4a and c), the plate age and subduction angle are indicated as the most important in first and second place, respectively, while in CatBoost (Figure 2.4.4e), this order is inverted. Notably, when examining the summary plot for the three models, we can discern a clear trend in the impact of plate age and subduction angle. For instance, we can see that older subducting plates (red dots) are associated with negative SHAP values that mean a predicted low b-values, and vice versa. Conversely, the

impact of the subduction angle is observed in the opposite way, where smaller dip angles (blue dots) have negative contributions in the SHAP values and therfore in low b-values, and vice versa. The trend for the impact of the distance to the closest slab edge (Dse) is less clear than the other two variables, showing some variability and outliers in its impact on predictions (no clear trend from red to blue or viceversa along the x-axis).

The remaining variables (ang_conv, vc_10, Dcr, and roughness) reveals distinct patterns and less relevant contributions to the predictive models. Convergence azimuth angle (ang_conv), while displaying a generally low impact, exhibits a noteworthy trend where smaller to medium angles (i.e. orthogonal to semi-oblique convergence) consistently contribute to low b-values. In the case of convergence velocity (vc_10), all three algorithms present an unclear trend. High values contribute both positively and negatively, rendering its impact ambiguous. The absence of a discernible pattern with convergence velocity highlights the complexity of its influence on predicting b-values.

For Dcr, a consistent observation emerges, particularly pronounced in CatBoost and Gradient Boosting: predominantly low Dcr (i.e. when the transect is closer to a subducting ridge) contribute positively to predictions and therefore are associated with high b-values, while large Dcr have a negative impact predicting low b-values. Finally, the subducting plate roughness is consistently indicated as the variable with the least impact across all three algorithms. In addition, its relationship with b-value via SHAP value remains unclear, adding an element of complexity to its role in shaping the predictive accuracy of the models.

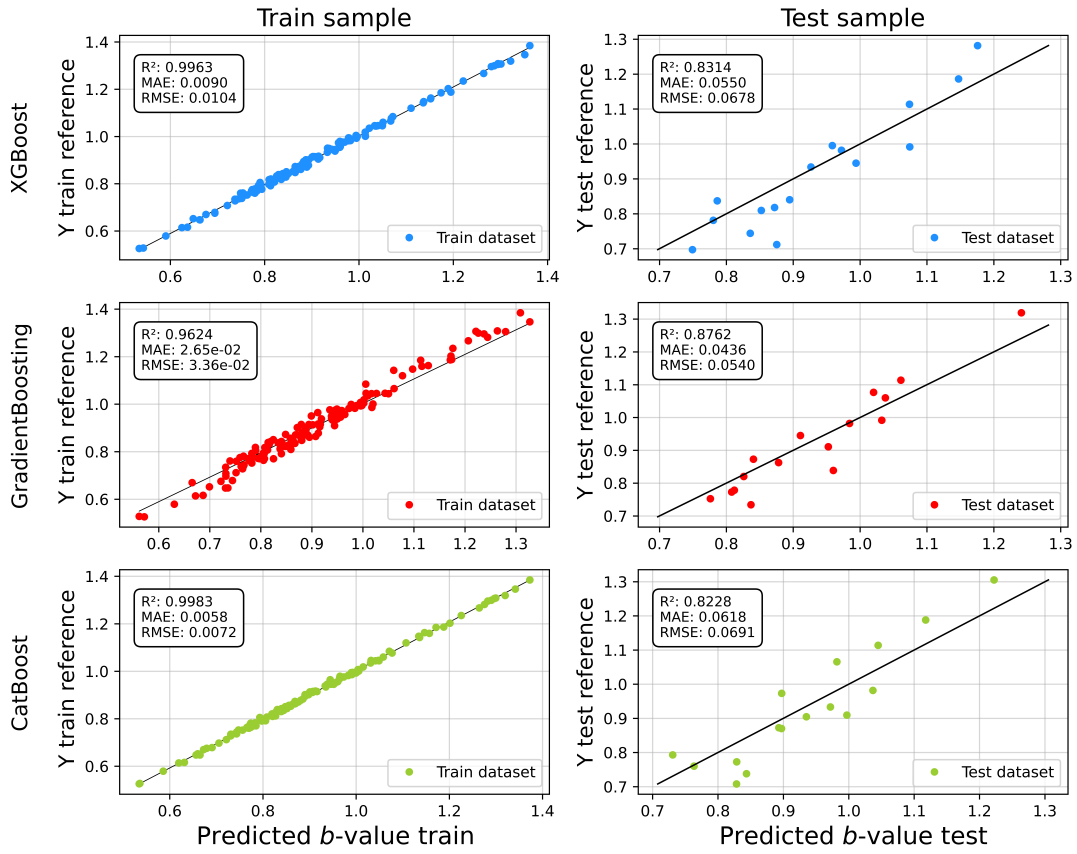

Figura 2.4.2: Visualization of the performance of machine learning models with a 90/10 training/test data split. The figure shows the comparison between predictions and actual benchmark values, highlighting model performance, for training and test data. The performance metrics R2, MAE and RMSE indicate the assessment of the predictive ability of the model.

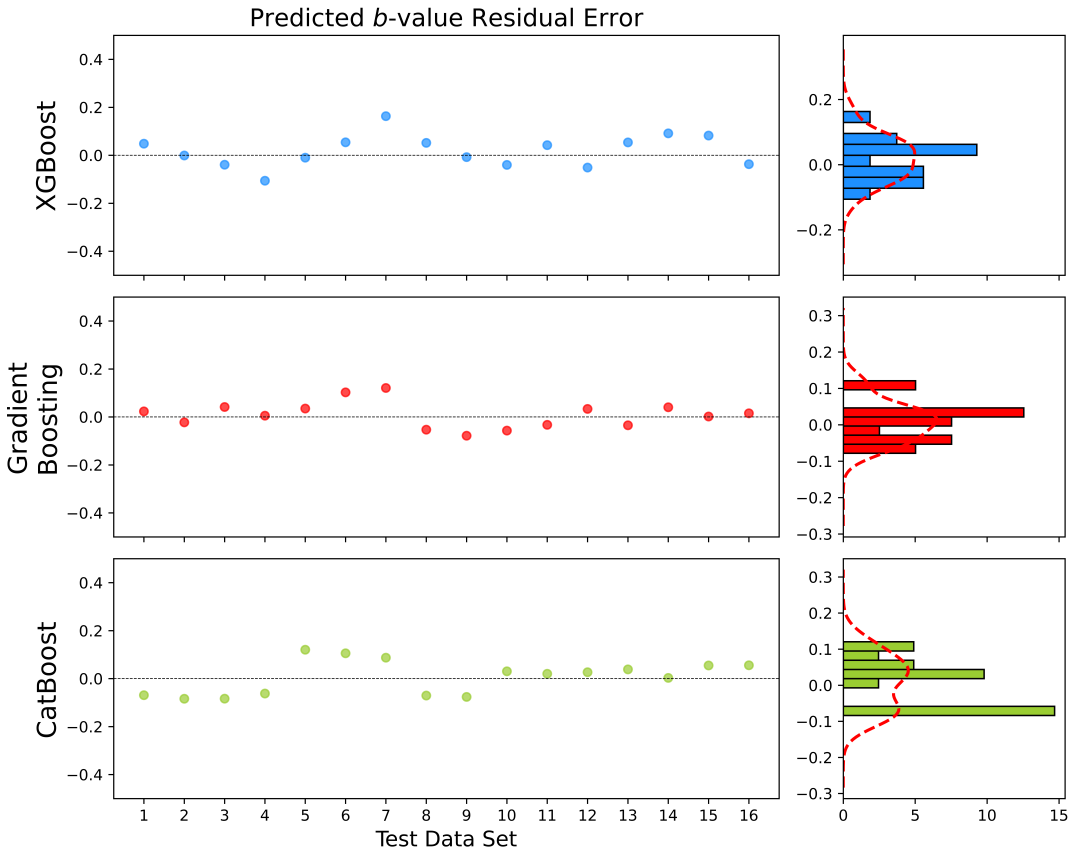

Figura 2.4.3: Residual prediction errors generated by the three Machine Learning models, using a 90/10 training/test data split

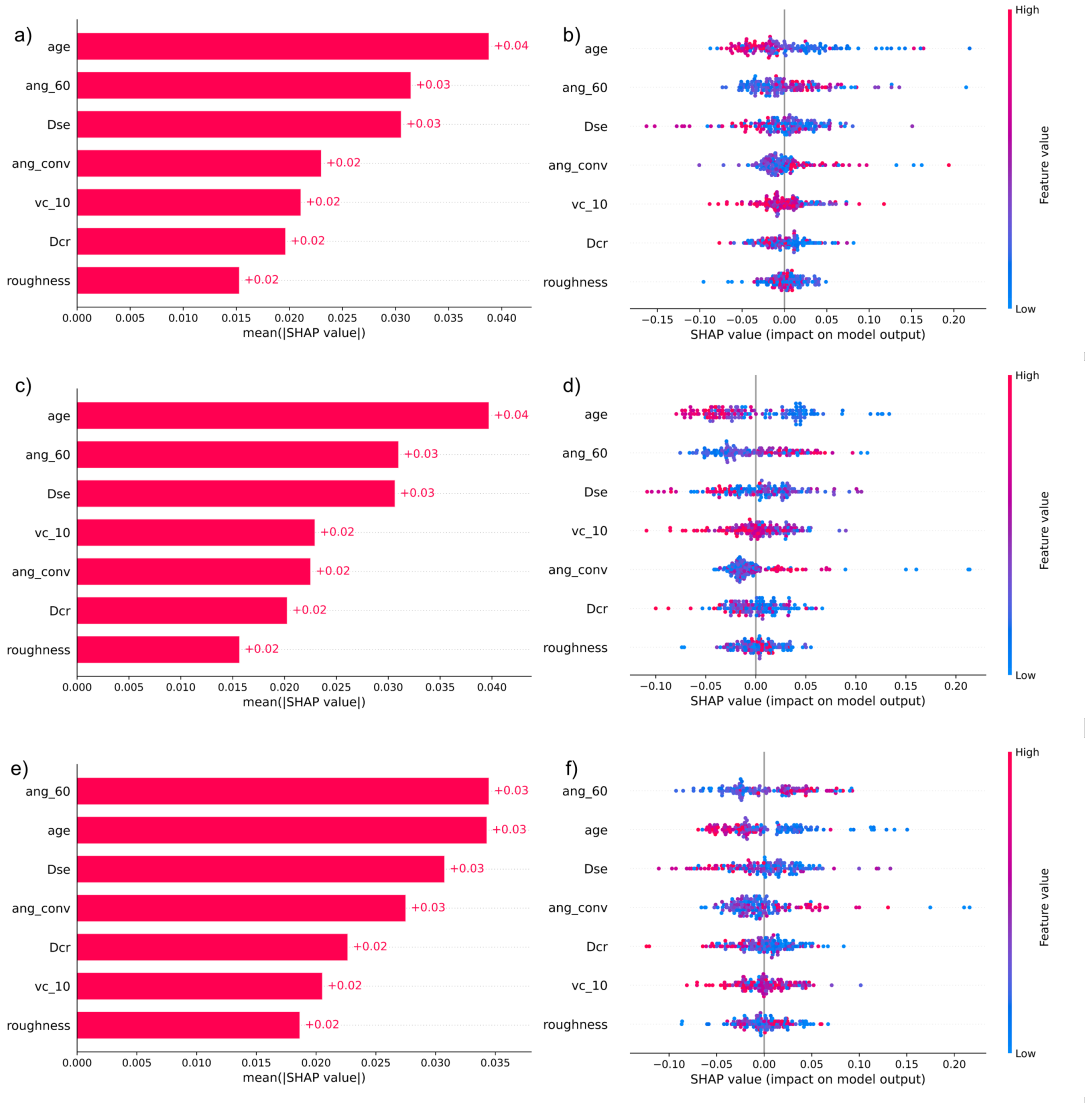

Figura 2.4.4: Comparison of features importance in predicting the *b*-value for three different ML models, each trained with a 90/10 train-test partition and using each of the three supervised ML algorithms: XGBoost (a, b), GradientBoosting (c, d), and CatBoost (e, f). Figures 2.4.4a, c, and e show the mean absolute SHAP values for each variable for each model, indicating the impact of variables ordered by highest to lowest relevance. Figures 2.4.4b, d, and f show the relative contribution of each feature to the predictions of the machine learning model. The points on the horizontal axis represent the magnitude of the impact of each feature, where positive SHAP values contribute to higher predictions and negative SHAP values contribute to a lower prediction in the model. The color of each point indicates the value of the feature for that sample, with blue for low values and red for high values. The vertical line in the center reflects the mean value of the model's predictions. ang_60 = subduction angle between 0 – 60 km depth; ang_conv = convergence azimuth; vc_10 = convergence velocity: Dse = distance between each transect and the closest slab edge along the trench; Dcr = distance between each transect and the closest subducting ridge along the trench, roughness = seafloor roughness 250 km seaward from the trench.

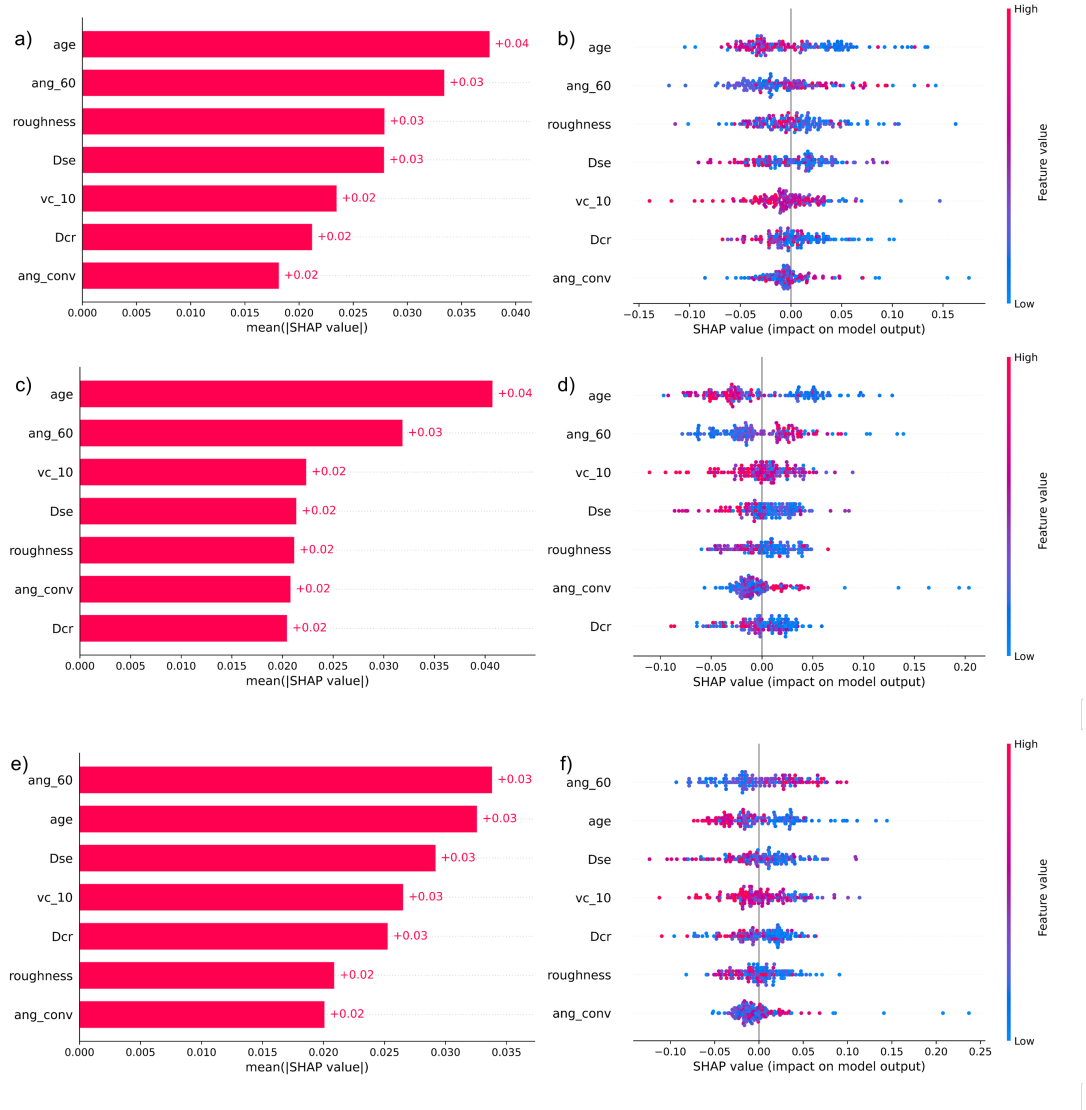

Figura 2.4.5: SHAP value barplots and summary plots employed for the interpretation of Machine Learning models. SHAP values are depicted for XGBoost (a, b), GradientBoosting (c, d), and CatBoost (e, f) models. All models were trained using an 80/20 split of training/testing data.

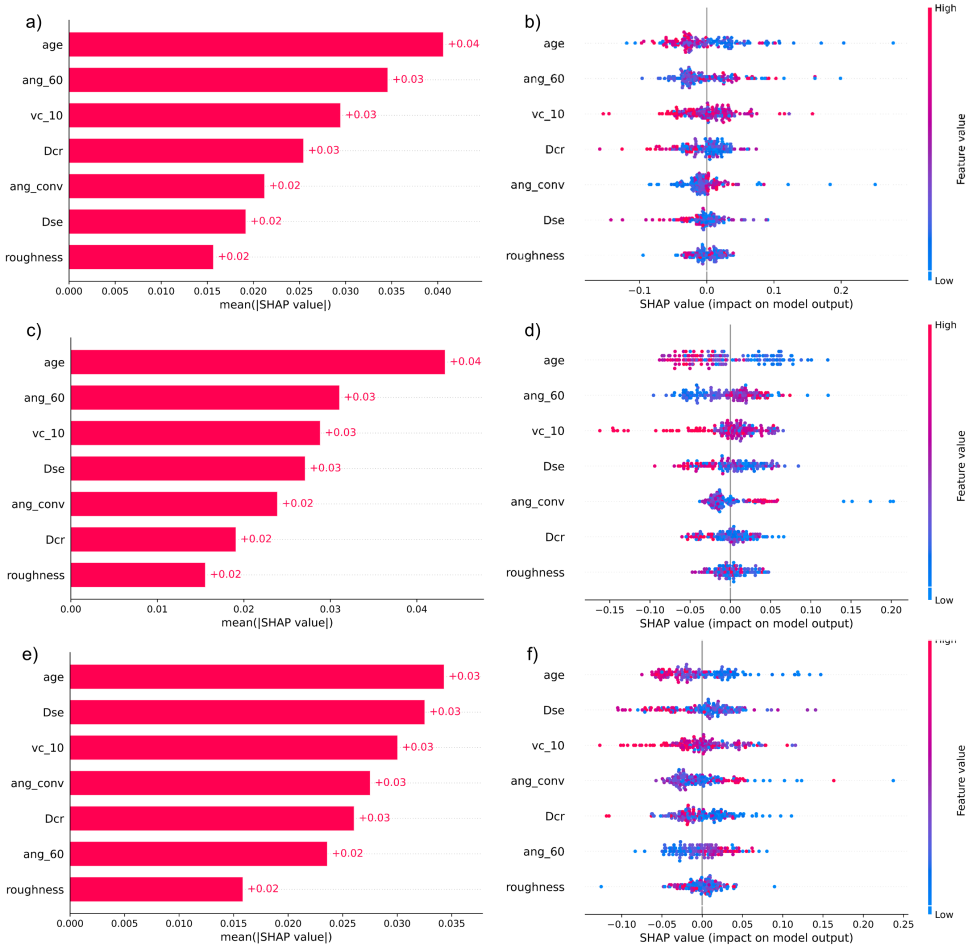

Figura 2.4.6: SHAP value barplots and summary plots employed for the interpretation of Machine Learning models. SHAP values are depicted for XGBoost (a, b), GradientBoosting (c, d), and CatBoost (e, f) models. All models were trained using a 70/30 split of training/testing data.

2.5. Discussions and Conclusions

Results obtained in this study reveal that oceanic plate age at the trench is the subduction parameter with a greater influence on the *b*-value and therefore on stress state of the megathrust. In a first glance, this conclusion agrees with Nishikawa & Ide (2014, herein N&I14), who found that plate age has the highest correlation coefficient (0.60) in a linear regression against *b*-value, with convergence velocity and upper plate velocity away from the trench having a rather week or null correlation. However, the positive correlation between slab age and *b*-value observed by N&I14, which for them implies a dominance of the age-dependent slab buoyancy on megathrust stress state, is at odd with our results since younger subducting plates (blue dots in Figs. 2.4.4b, d and f) are associated to positive SHAP values translating into grater *b*-values, and vice versa.

Although we believe that using a linear univariate correlation approach to analyze the likely complex non-linear interaction of different variables is less efficient than using machine learning, we still computed a linear correlation between our estimates of *b*-value and subducting plate age at the trench (Figure 2.5.2) just to repeat the analysis of N&I14 and to have a better base for comparison. We found a very week and negative correlation, with a coefficient of -0.12. This notable disagreement, which challenges the main conclusions of N&I14, can be due to several factors. First, our linear correlation (Fig. 2.5.2) was computed considering almost two times more data points than N&I14 (157 versus 75), covering subduction segments that were excluded from their analysis (Alaska-Aleutians, Cascadia, Southern Chile, Lesser Antilles, Sandwich). Some of these segments (e.g. Alaska-Aleutians, Sandwich) are characterized in our work (2.4.1) by an inverse relation between plate age and *b*-value, contributing to the discrepancy. We also note that for some segments included in both analysis (e.g. Sumatra, Central America) we obtain very different estimates of *b*-value compared with N&I14. These differences likely own to differences in: the seismicity catalogue used by both studies (ANSS by N&I14 v/s ISC by us), the time interval considered (1978-2009 by N&I14 v/s 1900-2022 by us), the hypocentral depths of considered events (all events by N&I14 v/s only those around the slab upper surface by us), and the method to compute the *b*-value (maximum likelihood without declustering of aftershock sequences by N&I14 v/s b-positive by us). Particularly this latter point can be significant, since

considering only the positive magnitude differences between events to perform the b-positive method van der Elst (2021), instead of all absolute magnitudes as the classical maximum likelihood method Aki (1965), means that aftershock sequences are naturally avoided ensuring that they cannot contaminate the overall estimate of the *b*-value, something particularly relevant in segments that experienced great earthquakes during the considered time interval (like in Sumatra for instance).

Accepting that our *b*-value estimates are well-computed and they can be considered a good representation of the stress state at subduction megathrusts, then we must discuss an alternative conceptual model to the one proposed by Nishikawa and Ide (2014). For this we also consider the large impact that our machine learning models unravel for the subduction angle as a predictor of the *b*-value (high average SHAP values in Figs. 2.4.4a, c and e). Moreover, our results indicate a positive correlation between both parameters, with shallower/smaller subduction angles (blue dots in Figs. 2.4.4b, d and f) associated to negative SHAP values meaning lower *b*-values. The combined trend of *b*-value being negatively correlated to plate age and positively correlated with the subduction angle indirectly implies a reverse correlation between these two subduction parameters, something that is partially supported by recent linear regression analysis at global scale (i.e. Hu and Gurnis, 2020), although a role of plate motion in controlling slab dip seems to be dominant (Cruciani et al., 2005; Lallemand et al., 2005). Into this framework, we propose a novel conceptual model (Fig. 2.5.1) where the oceanic plate age exerts its dominance via a control on flexural rigidity of the slab, more specifically on the elastic thickness of the plate. In our model, the elastic core of older plates is thicker than for younger plates, and therefore they tend to subduct with larger radius of curvature generating shallower subduction angles (Wu et al., 2008; Capitanio and Morra, 2012). This further implies a larger contact area between both converging plates across the megathrust, augmenting the width of the seismogenic zone, its overall shear stress level and increasing the potential for larger earthquakes to occur.

Our results also suggest that other parameters might play a secondary role modulating the stress state of the megathrust in addition to the dominance of slab age at the trench and shallow subduction angle. The distance to the lateral boundaries of subducting plates (Dse in Fig. 2.4.4) seems to be only marginally less significant than the subduction angle, with segments faraway from

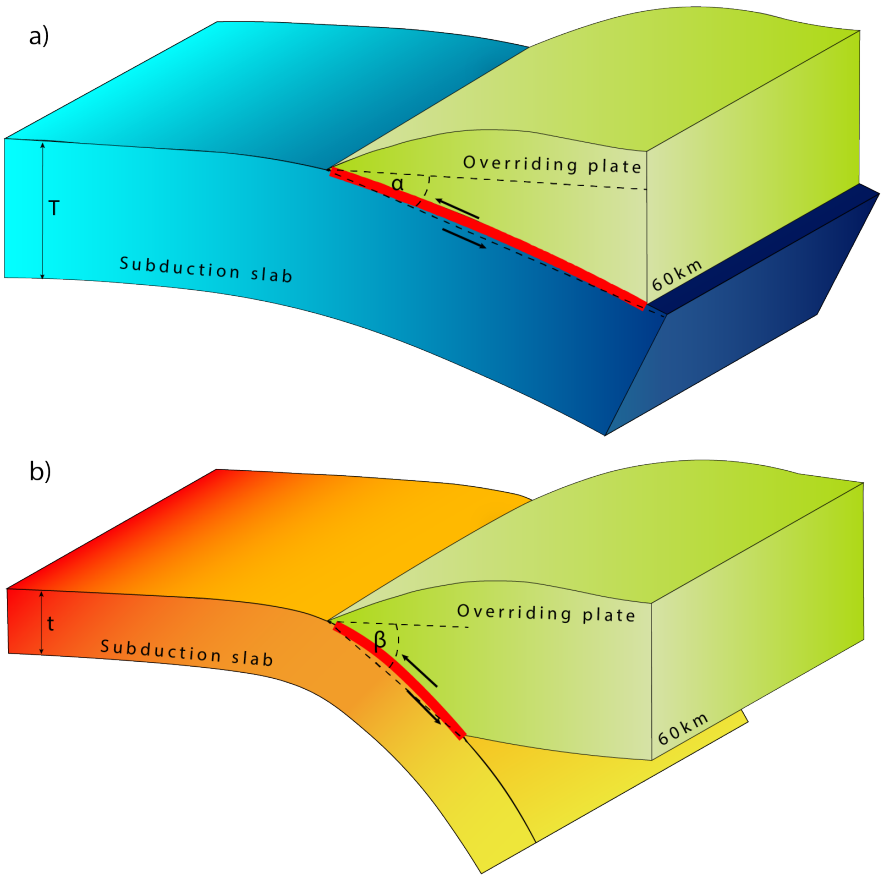

Figura 2.5.1: Comparison of the subduction of an older and younger plate. a) An older, thicker (T), more rigid plate subducts at a shallower angle (α), which increases the contact surface (red) and the stress on the megathrust. b) A younger, thinner (t), more flexible plate subducts at a steeper angle (β), which reduces the interplate contact surface (red) and the stress on the megathrust.

boundaries having the lowest *b*-values and therefore highest stresses. This is in agreement with previous researchers (i.e. Schellart and Rowlison, 2013) that found a relative large linear univariate correlation of Dse with the maximum magnitude of megathrust earthquakes. Plate convergence appears to have a secondary impact compared to previously discussed parameters, somehow in line with global linear regressions (Nishikawa and Ide, 2014; Hu and Gurnis, 2020). However, it stands in Figs.2.4.4b, d and f that most rapid and orthogonal convergence favors low *b*-values and large megathrust stresses, as can intuitively be supposed. Although the calculated *b*-values seems to be much less sensitive to the proximity to a subducting aseismic ridge and the roughness of the oceanic crust, our results suggest that megathrust strength tend to be lower (i.e. higher *b*-values) in subduction segments dominated by ridge subduction. This can be also appreciated in Fig. 2.4.1 for

South America for example, where subduction of the Carnegie Ridge near 5^{o}S and Juan Fernandez Ridge at 33^{o}S are clearly related to locally augmented b-values compared to adjacent segments, something that has been observed by previous studies (Legrand et al., 2012) and support the notion that subducting rough bathymetry associated to seamount chains decrease the strength of the megathrust and favor convergence absorption via creep and aseismic slip (i.e., Wang and Bilek, 2014; Basset and Watts, 2015).

The complexity of the likely non-linear interactions between subduction variables in terms of their integrated effect over the megathrust stress state means that using machine learning approaches, as done here, to analyze the possible influence of each variable in the context of all other existing variables is superior compared to previous uni- or multi-variate linear regressions. This underscores the need for a more holistic approach when interpreting seismic phenomena, highlighting the importance of the interrelation of multiple factors in predicting the seismic behavior of the megathrust. Future works in this line should include other parameters that have been also indicated as significantly affecting the seismogenic behavior, as the thickness of subducting sediment (e.g. Brizzi et al., 2021), gravity anomalies (i.e., Basset and Watts, 2015; Molina et al., 2021) or temperature (Hyndman, 2023). These considerations highlight the importance of future research that expands the range of factors considered, thus contributing to a more comprehensive understanding of the complexity of the interaction between subduction variables and megathrust stress.

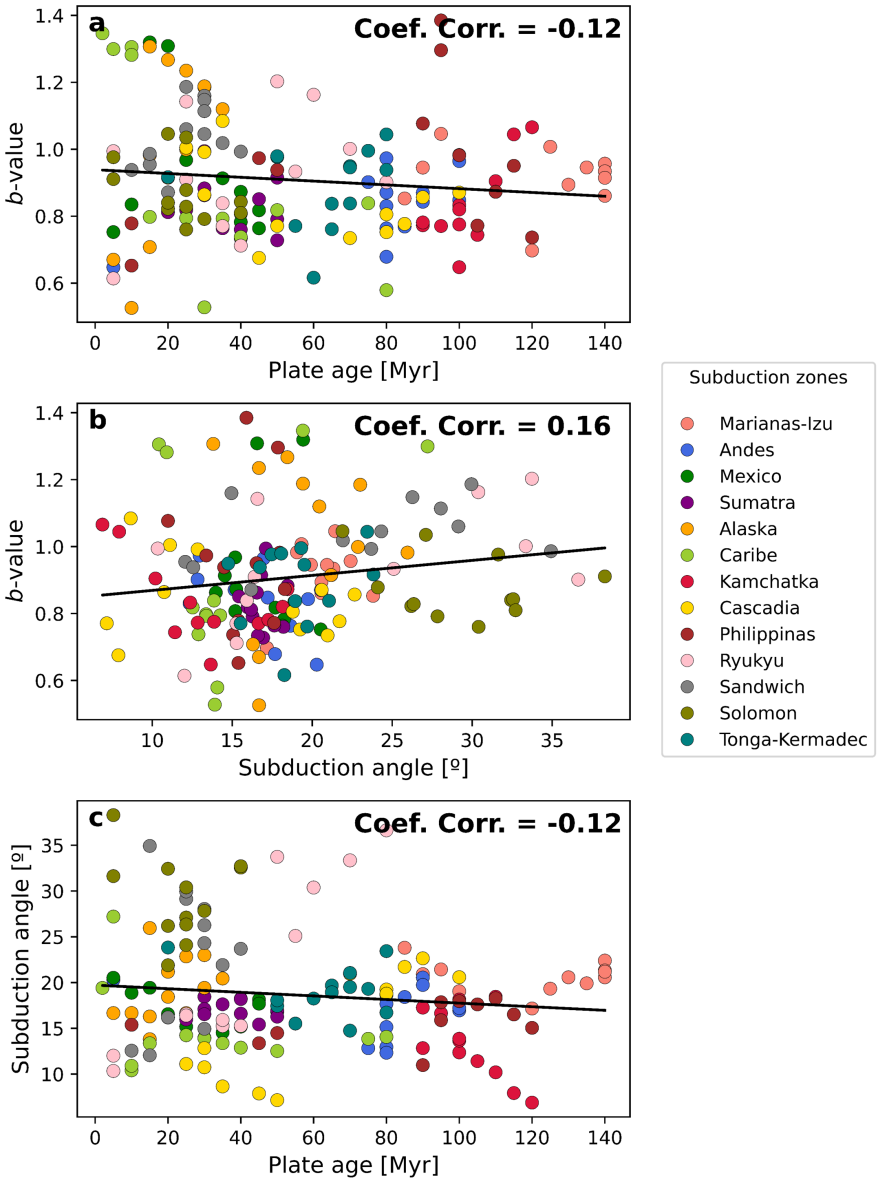

Figura 2.5.2: Comparison between most important parameters and b-value showing the lack of correlation. a) Relationship between plate age and b-value. b) Relationship between subduction angle age and b-value. c) Relationship between plate age and subduction angle.

Capítulo 3

Discusión y Conclusión

3.1. Discusión y Conclusión

El principal objetivo de este estudio es determinar el impacto de variables de subducción en el estado de *stress* del *megathrust* a escala global. Para ello llevamos a cabo la implementación la implementación de tres algoritmos de *Machine Learning* supervisado basados en árboles de decisión: *XGBoost*, *GradientBoosting* y *CatBoost*. Hemos creado y evaluado un modelo para cada algoritmo, utilizando diferentes proporciones de datos de entrenamiento y de prueba (90/10, 80/20 y 70/30). Posteriormente, para interpretar el funcionamiento interno de los modelos y entender el imapcto de las variables de entrada en la salida de los modelos, hemos utilizado el algoritmo SHAP *value* , que permite analizar cómo contribuyen las variables de subducción a la predicción del *b-value* según cada modelo.

Los resultados indican que, con una partición de 90/10 de datos de entrenamiento y prueba, los tres modelos se ajustan bien a los datos y son capaces de predecir el *b-value* con un alto grado de precisión y bajo error (R^2 XGBoost = 0.83, R^2 GradientBoosting = 0.88 y R^2 CatBoost = 0.82), lo que significa que el porcentaje de variabilidad del *b-value* explicado por las variables independientes es sobre el 82 %. Con una proporción de datos de entrenamiento-prueba menores (80/20 y 70/30) el rendimiento general de los algoritmos en la predicción disminuye (R^2 entre 0.64 y 0.44). Esto indicaría que una reducción en el porcentaje de datos de entrenamiento no permite que los modelos cuenten con una cantidad de datos

representativa para aprender de patrones complejos o que se indique una mayor sensibilidad a valores atípicos, dada la aleatoriedad de la partición. A pesar de las diferencias en el rendimiento de los algoritmos cuando de acuerdo con el porcentaje en la partición de datos de entrenamiento y prueba, al implementar el SHAP *value* para la interpretación de los modelos, las relaciones encontradas son bastante consistentes independiente del algoritmo y de la partición (Figuras 2.4.4, 2.4.5 y 2.4.6), lo que da luces de un patrón dentro de los datos y expone la complejidad del efecto de variables de subducción en el *stress* del *megathrust* a escala global. Para todas las particiones los gráficos SHAP muestran que para los modelos *XGBoost* y *GradientBoosting*, la edad de la placa y el ángulo de subducción son las variables con mayor influencia en la predicción, mientras que para el modelo *CatBoost*, el ángulo de subducción es la más importante, seguida de la edad de la placa. Las tendencias observadas sugieren que las zonas de subducción de placas más antiguas y de menor ángulo contribuyen a la predicción de *b-values* más bajos, lo que implica un mayor *stress* en el *megathrust*. Se propone que la mayor rigidez de las placas más antiguas puede aumentar la resistencia a la flexión, expresada como un menor ángulo de penetración y, por tanto, una mayor superficie de fricción en la interfaz de las placas, lo que se traduce en un aumento del *stress* en el *megathrust*. También se pone de relieve la relación entre la subducción ortogonal a subortogonal y un mayor *stress*, teniendo en cuenta la tendencia del ángulo de convergencia.

En conclusión, los hallazgos indicados por SHAP *value*, permiten sostener, por una parte, que el objetivo principal de la investigación fue alcanzado y, por otro lado, confirmar la hipótesis propuesta inicialmente. Aunque nuestros resultados apoyan algunos aspectos de la literatura existente y enriquecen la comprensión no lineal del comportamiento sismogénico global, es importante reconocer las limitaciones de este estudio. Entre ellas se incluye la necesidad de incluir variables adicionales como el espesor de los sedimentos, la temperatura y las heterogeneidades del ante-arco para proporcionar una perspectiva más completa y detallada (p. ej. (Brizzi et al., 2018; Molina et al., 2021; Hyndman, 2023). Estas consideraciones ponen de relieve la importancia de futuras investigaciones que amplíen el abanico de factores considerados, contribuyendo así a una comprensión más completa de la complejidad de la interacción entre las variables de subducción y el estado de *stress* del *megathrust*.

Bibliografía

Aki, K. (1965). Maximum likelihood estimate of b in the formula log n= a-bm and its confidence limits. *Bull. Earthquake Res. Inst., Tokyo Univ.*, 43:237–239.

Basset, D. and Watts, A. (2015). Gravity anomalies, crustal structure, and seismicity at the subduction zones: 2. interrrelationships between fore-arc structure and seismogenic behavior. geochemestry. *Geophysics*, 16(5).

Brizzi, S., Sandri, L., Funiciello, F., Corbi, F., Piromallo, C., and Heuret, A. (2018). Multivariate statistical analysis to investigate the subduction zone parameters favoring the occurrence of giant megathrust earthquakes. *Tectonophysics*, 728:92–103.

Brodsky, E. E. (2019). Determining whether the worst earthquake has passed.

Capitanio, F. and Morra, G. (2012). The bending mechanics in a dynamic subduction system: Constraints from numerical modelling and global compilation analysis. *Tectonophysics*, 522:224–234.

Carter, J. A. and Berg, E. (1981). Relative stress variations as determined by b-values from earthquakes in circum-pacific subduction zones. *Tectonophysics*, 76(3-4):257–271.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

Corbi, F., Sandri, L., Bedford, J., Funiciello, F., Brizzi, S., Rosenau, M., and Lallemand, S. (2019). Machine learning can predict the timing and size of analog earthquakes. *Geophysical Research Letters*, 46(3):1303–1311.

Cruciani, C., Carminati, E., and Doglioni, C. (2005). Slab dip vs. lithosphere age: no direct function. *Earth and Planetary Science Letters*, 238(3-4):298–310.

Dielforder, A., Bocchini, G., Kemna, K., Hampel, A., Harrington, R., and Oncken, O. (2023). Megathrust stress drop as trigger of aftershock seismicity: Insights from the 2011 tohoku earthquake, japan. *Geophysical Research Letters*, 50(3):e2022GL101320.

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

Hastie, T., Tibshirani, R., and Friedman, J. (2001). The elements of statistical learning. springer series in statistics. *New York, NY, USA*.

Heuret, A., Conrad, C., Funiciello, F., Lallemand, S., and Sandri, L. (2012). Relation between subduction megathrust earthquakes, trench sediment thickness and upper plate strain. *Geophysical Research Letters*, 39(5).

Hyndman, R. (2023). The thermal regime of nw canada and alaska, and tectonic and seismicity consequences. *Geochemistry, Geophysics, Geosystems*, 24(7):e2022GC010570.

Jones, G. P., Hickey, J. M., Di Stefano, P. G., Dhanjal, C., Stoddart, L. C., and Vasileiou, V. (2020). Metrics and methods for a systematic comparison of fairness-aware machine learning algorithms. *arXiv preprint arXiv:2010.03986*.

Kagan, Y. Y. and Jackson, D. D. (2013). Tohoku earthquake: A surprise? *Bulletin of the Seismological Society of America*, 103(2B):1181–1194.

Kanamori, H. (1983). Magnitude scale and quantification of earthquakes. *Tectonophysics*, 93(3-4):185–199.

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).

Legrand, D., Tassara, A., and Morales, D. (2012). Megathrust asperities and clusters of slab dehydration identified by spatiotemporal characterization of seismicity below the andean margin. *Geophysical Journal International*, 191(3):923–931.

Lemenkova, P. (2020). Gebco gridded bathymetric datasets for mapping japan trench geomorphology by means of gmt scripting toolset. *Geodesy and Cartography*, 46(3):98–112.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.

Mitchell, T. M. (1997). Machine learning.

Molina, D., Tassara, A., Abarca, R., Melnick, D., and Madella, A. (2021). Frictional segmentation of the chilean megathrust from a multivariate analysis of geophysical, geological, and geodetic data. *Journal of Geophysical Research: Solid Earth*, 126(6):e2020JB020647.

Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.

Müller, R. D., Seton, M., Zahirovic, S., Williams, S. E., Matthews, K. J., Wright, N. M., Shephard, G. E., Maloney, K. T., Barnett-Moore, N., Hosseinpour, M., et al. (2016). Ocean basin evolution and global-scale plate reorganization events since pangea breakup. *Annual Review of Earth and Planetary Sciences*, 44:107–138.

Nanjo, K., Hirata, N., Obara, K., and Kasahara, K. (2012). Decade-scale decrease inb value prior to the m9-class 2011 tohoku and 2004 sumatra quakes. *Geophysical Research Letters*, 39(20).

Natekin, A. and Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7:21.

Nishikawa, T. and Ide, S. (2014). Earthquake size distribution in subduction zones linked to slab buoyancy. *Nature Geoscience*, 7(12):904–908.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018). Catboost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31.

Rozemberczki, B., Watson, L., Bayer, P., Yang, H.-T., Kiss, O., Nilsson, S., and Sarkar, R. (2022). The shapley value in machine learning. *arXiv preprint arXiv:2202.05594*.

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

Saeed, M., Villarroel, M., Reisner, A. T., Clifford, G., Lehman, L.-W., Moody, G., Heldt, T., Kyaw, T. H., Moody, B., and Mark, R. G. (2011). Multiparameter intelligent monitoring in intensive care ii (mimic-ii): a public-access intensive care unit database. *Critical care medicine*, 39(5):952.

Salas, P., De la Fuente, R., Astroza, S., and Carrasco, J. A. (2022). A systematic comparative evaluation of machine learning classifiers and discrete choice models for travel mode choice in the presence of response heterogeneity. *Expert Systems with Applications*, 193:116253.

Schäfer, A. M. and Wenzel, F. (2019). Global megathrust earthquake hazard—maximum magnitude assessment using multi-variate machine learning. *Frontiers in Earth Science*, 7:443496.

Schellart, W. P. and Rawlinson, N. (2013). Global correlations between maximum magnitudes of subduction zone interface thrust earthquakes and physical parameters of subduction zones. *Physics of the Earth and Planetary Interiors*, 225:41–67.

Scholz, C. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bulletin of the seismological society of America*, 58(1):399–415.

Scholz, C. H. (2015). On the stress dependence of the earthquake b value. *Geophysical Research Letters*, 42(5):1399–1402.

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

Shapley, L. S. et al. (1953). A value for n-person games.

Stein, S. and 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(1A):S279–S295.

van der Elst, N. J. (2021). B-positive: A robust estimator of aftershock magnitude distribution in transiently incomplete catalogs. *Journal of Geophysical Research: Solid Earth*, 126(2):e2020JB021027.

Wang, K. and Bilek, S. L. (2014). Invited review paper: Fault creep caused by subduction of rough seafloor relief. *Tectonophysics*, 610:1–24.

Wu, B., Conrad, C. P., Heuret, A., Lithgow-Bertelloni, C., and Lallemand, S. (2008). Reconciling strong slab pull and weak plate bending: The plate motion constraint on the strength of mantle slabs. *Earth and Planetary Science Letters*, 272(1-2):412–421.

Xiong, Q., Xu, M., Li, J., Liu, Y., Zhang, J., Xu, Y., and Dong, W. (2021). Clinical sequelae of covid-19 survivors in wuhan, china: a single-centre longitudinal study. *Clinical microbiology and infection*, 27(1):89–95.

Apéndice A

A1. Notas sobre métodos relacionados al *Machine Learning*

XGboost

XGBoost (Extreme Gradient Boosting) propuesto por Chen and Guestrin (2016) es un algoritmo de aprendizaje automático potente y versátil. Se basa en la construcción secuencial de árboles de decisión débiles donde en cada iteración, un nuevo árbol se adapta para corregir los errores residuales del modelo existente. La contribución de cada árbol se pondera mediante un factor de aprendizaje y se incorpora una regularización para controlar la complejidad del modelo. XGBoost destaca por su capacidad para manejar conjuntos de datos grandes y complejos, su eficacia a la hora de seleccionar puntos de división óptimos y su flexibilidad al permitir la definición de funciones objetivo personalizadas. Además, su capacidad para construir árboles más profundos facilita la captura de patrones complejos en los datos.

GradientBoosting

El algoritmo GradientBoosting (Natekin and Knoll, 2013) también se basa en la construcción iterativa de árboles de decisión débiles, pero su enfoque es más sencillo en términos de regularización y ajuste de hiperparámetros. De forma secuencial, cada nuevo árbol se ajusta para minimizar la derivada de la función de pérdida con respecto a la predicción del modelo existente. Este enfoque permite mejorar el modelo centrándose en los errores residuales de iteraciones

anteriores.Gradient Boosting tiende a construir árboles menos profundos que XGBoost, lo que limita la complejidad del modelo. La introducción de un factor de aprendizaje, conocido como contracción, controla la contribución de cada árbol al modelo global, mejorando la generalización.

CatBoost

CatBoost, desarrollado por Prokhorenkova et al. (2018), es también un algoritmo basado en GDBT (Gradient Boosting Decision Trees). En comparación con otros algoritmos de gradient boosting, la fuerza de CatBoost radica en su capacidad para manejar características categóricas. A diferencia de otros algoritmos, CatBoost no requiere preprocesamiento adicional para estas variables, ya que implementa técnicas avanzadas como la codificación por orden y la estadística de las categorías. Esto simplifica enormemente el flujo de trabajo y hace que el algoritmo sea especialmente eficiente en conjuntos de datos del mundo real con variables categóricas.CatBoost también destaca por su optimización integrada de hiperparámetros, que facilita el ajuste del modelo sin necesidad de realizar ajustes manuales. Además, el proceso de optimización en paralelo mejora la eficiencia computacional, lo que convierte a CatBoost en una opción atractiva para una gran variedad de problemas de aprendizaje automático de regresión.

SHAP value

SHAP (SHapley Additive ExPlanations, (Lundberg and Lee, 2017) es una técnica de explicabilidad basada en la teoría de juegos y los valores de Shapley (Shapley et al., 1953) para explicar el resultado de los modelos de aprendizaje automático. Permite asignar una contribución de importancia a cada característica de entrada en una predicción específica, teniendo en cuenta todas las interacciones posibles. Además, permite visualizar las contribuciones de cada característica mediante diagramas de importancia. El algoritmo SHAP explora exhaustivamente todas las combinaciones posibles de características para calcular la contribución de cada una de ellas. Este enfoque garantiza que cada característica reciba una contribución justa teniendo en cuenta todas las interacciones posibles.Para cada combinación de características, el algoritmo evalúa el modelo y calcula la diferencia entre la predicción del modelo con la característica y la predicción media del modelo

mediante la siguiente ecuación:

$$\phi_i(f,x) = \sum_{z' \in x'} \frac{|z'|!(M-|z'|-1)!}{M!}[f_x(z') - f_x(z' \backslash i)]$$

donde ϕ_i corresponde a la contribución de la característica de entrada i, en función de f que corresponde al modelo generado previamente por cada uno de los algoritmos (XGboost, GradientBoosting y CatBoost) y x que sería una fila concreta dentro del conjunto de datos. $f_x(z')$ y $f_x(z' \backslash i)$ corresponden a los modelos con y sin la variable de interés respectivamente, donde la diferencia entre ambos dirá cuánto contribuye esa variable a la predicción en el subconjunto de datos (Lundberg and Lee, 2017).

Validación cruzada

Para evitar el sobreajuste en la predicción de datos y maximizar la eficacia de los algoritmos, es necesario ajustar los hiperparámetros que determinan el proceso de aprendizaje del algoritmo. Para ello, se realiza una validación cruzada del algoritmo. La validación cruzada es una técnica que permite evaluar los modelos de ML y seleccionarlos con mayor precisión y fiabilidad. En este proceso, el conjunto de datos de entrenamiento se divide en k conjuntos más pequeños denominados "foldsz, a continuación, se utiliza cada uno de ellos como conjunto de validación para evaluar el rendimiento del modelo entrenado con los otros pliegues como conjunto de entrenamiento. Por ejemplo, en una técnica de validación cruzada de 5 $folds$ (como en nuestro caso), el conjunto de datos de entrenamiento se divide en 5 $folds$ iguales. Para cada uno de los segmentos se evaluará una vez con todos los rangos de valores determinados en un diccionario para cada uno de los hiperparámetros. En cada una de las 5 iteraciones (para cada uno de los segmentos del conjunto de datos de entrenamiento), este segmento aislado ($fold$) se utilizará para validar y estimar el rendimiento del modelo (R^2) a partir del entrenamiento del modelo generado con los 4 pliegues restantes. Finalmente, con el R^2 más alto se determinarán los hiperparámetros óptimos en función de la variable de salida a predecir. Este proceso garantiza que el modelo se evalúe y pruebe exhaustivamente con diferentes conjuntos de datos. Además, permite utilizar el conjunto completo de datos de entrenamiento disponibles para construir y validar el modelo, evitando así una posible sobreestimación de los resultados.

Métricas de evaluación

Las métricas de evaluación son parámetros que miden y ayudan a comprender el rendimiento de los modelos de ML. Estas métricas proporcionan una evaluación cuantitativa de lo bien que un modelo se ajusta a los datos y realiza predicciones midiendo la precisión, robustez y capacidad de generalizar nuevos datos de un algoritmo. Utilizando medidas como el coeficiente de determinación (R^2), el error cuadrático medio (RMSE) y el error medio absoluto (MAE), es posible analizar el poder explicativo y la precisión del modelo, lo que permite tomar decisiones informadas sobre la eficacia y aplicabilidad del algoritmo en cuestión.

El **coeficiente de determinación**(R^2) es una medida del porcentaje de variabilidad de la variable dependiente que explican las variables independientes del modelo. R^2 oscila entre 0 y 1, donde 1 indica que el modelo explica perfectamente la variabilidad de los datos. En términos sencillos, R^2 cuantifica hasta qué punto las predicciones del modelo se ajustan a los datos reales. El coeficiente de determinación se calcula del siguiente modo:

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

donde SS_{res} es la suma de los cuadrados de las diferencias entre las predicciones del modelo y los valores reales (residuos) y SS_{tot} es la suma de los cuadrados de las diferencias entre los valores reales y la media de los valores reales.

El **Error cuadrático medio** (RMSE) es una medida de la precisión del modelo que representa la raíz cuadrada de los cuadrados medios de las diferencias entre las predicciones del modelo y los valores reales. Esta métrica penaliza los errores grandes de forma más significativa, ya que el cálculo implica elevar al cuadrado las diferencias individuales y, a continuación, llevar la raíz cuadrada de nuevo a la unidad original. El RMSE se calcula mediante la siguiente fórmula:

$$RSME = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

donde, n es el número de observaciones del conjunto de datos, y_i son los valores reales de la variable dependiente e $\hat{y}_i$ son las predicciones del modelo de la variable dependiente.

El **Error Medio Absoluto** (MAE) es otra medida de precisión y mide la media de las diferencias absolutas entre las predicciones del modelo y los valores reales. A diferencia del RMSE, que da más peso a los errores grandes y, por tanto, es sensible a los valores atípicos, el MAE trata todos los errores por igual, independientemente de su tamaño, y se calcula mediante la siguiente fórmula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|,$$

donde, n es el número de observaciones del conjunto de datos, y_i son los valores reales de la variable dependiente e $\hat{y}_i$ son las predicciones del modelo de la variable dependiente.

ML model	**Hiperparámetros**	**Data split**	R^2	**RMSE**	**MAE**
XGBoost	Depth_max = 5, n_estimators = 500, learning_rate = 0.3, reg_alpha = 0, reg_lambda = 0.001, gamma = 0.001, sub_sample = 0.8, random_state = 4719, colsample_bytree=0.8, min_child_weight=1	90/10	0.831	0.067	0.055
		80/20	0.528	0.112	0.095
		70/30	0.510	0.118	0.087
GradientBoosting	Max_depth = 4, n_estimators = 100, learning_rate = 0.1, min_samples_leaf = 4, min_samples_split=2, random_state = 1975	90/10	0.876	0.054	0.044
		80/20	0.528	0.101	0.086
		70/30	0.481	0.110	0.087
CatBoost	Depth = 4, iterations = 500, learning_rate = 0.2, reg_alpha = 0, gamma = 0.001, l2_leaf_reg= 4, loss_function='RMSE', random_state = 421	90/10	0.823	0.069	0.062
		80/20	0.643	0.094	0.076
		70/30	0.407	0.122	0.100

Cuadro A1.1: Métricas de rendimiento para los modelos de ML evaluados en el conjunto de test para diferentes particiones de entrenamiento/prueba.