UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

ESTIMACIÓN AUTOMÁTICA DE DESLIZAMIENTOS EN LA INTERFASE DE SUBDUCCIÓN USANDO DATOS GNSS Y MACHINE LEARNING

Por: Valentina Inzunza García

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

Enero 2025

Concepción, Chile

Profesor Guía: Dr. Matt Miller y Dr. Marcos Moreno

Comisión Evaluadora: Dra. Ignacia Calisto, Dr. Francisco Ortega, Dr. Julio Aracena

AGRADECIMIENTOS

Esta tesis fue financiada por el proyecto FONDECYT 1221507.

Agradezco a los profesores que participaron en el desarrollo de este trabajo por sus correcciones y sugerencias, las cuales contribuyeron al progreso y mejora de esta tesis.

También agradezco a mi familia y amistades por estar presentes durante este proceso. Su apoyo y compañía fueron fundamentales a lo largo de este trabajo.

Resumen

La estimación precisa y rápida de deslizamiento cosísmico es esencial para comprender los terremotos y optimizar los sistemas de alerta temprana. En este trabajo, se presenta una metodología basada en redes neuronales para estimar distribuciones de deslizamiento cosísmico, utilizando datos sintéticos de GNSS para un entrenamiento robusto. Este enfoque se distingue por su alta eficiencia computacional, logrando estimaciones de la distribución del deslizamiento en solo 0.07 segundos. Validado con datos reales del terremoto de Illapel de 2015 (M_w 8.3), el modelo demostró una notable precisión. Además, se analizan diversos hiperparámetros y condiciones de los datos, evaluando el desempeño del modelo e identificando oportunidades para optimizar su rendimiento en el futuro.

Índice general

Índice de cuadros

Índice de figuras

Capítulo 1

Introducción

1.1. Introducción

La tectónica de placas, a través de la acumulación y liberación de esfuerzos, es responsable de los terremotos más poderosos de la Tierra, especialmente en zonas de subducción. Estas zonas, donde una placa tectónica se desliza por debajo de otra, tienen el potencial de generar terremotos tsunamigénicos devastadores, como el terremoto de Sumatra-Andamán de 2004 (M_w 9.1) (e.g., Okal and Stein, 2009; Lay et al., 2005; Chlieh et al., 2007) y el terremoto de Tohoku de 2011 (M_w 9.0) (e.g., Fujii et al., 2011; Ozawa et al., 2012; Tajima et al., 2013). En estos márgenes convergentes, la energía elástica se acumula durante décadas o siglos, siendo liberada en gran parte durante terremotos en la porción somera de la interfaz de subducción. Chile es una región altamente activa sísmicamente (e.g., Cisternas et al., 2017), donde han ocurrido algunos de los terremotos de mayor magnitud registrados, a lo largo de la zona de subducción donde la placa de Nazca subduce bajo la placa Sudamericana a una tasa de 66 mm/año (Altamimi et al., 2016). Eventos recientes destacados en este margen incluyen el terremoto de Valdivia de 1960 (M_w 9.5) (e.g., Barrientos and Ward, 2007; Fujii and Satake, 2012; Lorenzo-Martín et al., 2006), el terremoto de Maule de 2010 (M_w 8.8) (e.g., Moreno et al., 2010, 2012; Bedford et al., 2013; Delouis et al., 2010), el terremoto de Iquique de 2014 (M_w 8.2) (e.g., Duputel et al., 2015; Jara et al., 2018; Meng et al., 2015) y el terremoto de Illapel de 2015 (M_w 8.3) (e.g., Melgar et al., 2016; Tilmann et al., 2015; Heidarzadeh et al., 2015).

En las últimas décadas, el uso de estaciones del Sistema Global de Navegación por

Satélite (GNSS) ha mejorado significativamente la observación de desplazamientos superficiales intersísmicos y cosísmicos, avanzando en la comprensión de la dinámica de los terremotos. A medida que las estaciones GNSS se han vuelto más comunes a nivel global, ahora ofrecen mediciones de alta resolución de la deformación cortical, especialmente en zonas de subducción, permitiendo identificar características y comportamientos de grandes terremotos de cabalgadura (e.g., Ruegg et al., 2009; Chen et al., 2015; Luo et al., 2020). La estimación de deslizamiento cosísmico de los terremotos es esencial para comprender las mecánicas subyacentes, incluyendo la liberación de energía, la redistribución de esfuerzos y la dinámica de fallas. Esta comprensión es vital para evaluar el riesgo sísmico, desarrollar estrategias de mitigación e informar esfuerzos de respuesta rápida (Iinuma et al., 2012).

Los enfoques comunes para estimar deslizamiento cosísmico implican la construcción de un modelo directo lineal a través de la generación de funciones de Green basadas en dislocaciones elásticas (e.g., Okada, 1985; Nikkhoo and Walter, 2015). A pesar de la linealidad del modelo directo, la estimación del deslizamiento cosísmico sigue siendo un problema inverso complejo y mal condicionado, caracterizado por soluciones no únicas. Existen dos enfoques extremos para abordar este problema inverso. El primero, conocido como el enfoque de optimización, se centra en encontrar una solución que minimice un término de ajuste de datos y un término de regularización, este último utilizado para definir información previa que estabilice la solución (e.g., Harris and Segall, 1987; Ortega-Culaciati et al., 2021). El segundo es un enfoque bayesiano, donde se muestrea un conjunto de modelos a partir de una distribución de probabilidad posterior del deslizamiento (e.g., Minson et al., 2013; Duputel et al., 2014).

Con el avance de la tecnología y la proliferación de grandes volúmenes de datos, los algoritmos de Machine Learning (ML) han tomado un papel importante en problemas sismológicos. Avances recientes incluyen: la aplicación de técnicas de Deep Learning para detectar terremotos de baja frecuencia (Münchmeyer et al., 2024), eliminar ruido de datos HR-GNSS (Thomas et al., 2023), identificar automáticamente eventos de deslizamiento lento (Donoso et al., 2021), analizar tectónica a través del agrupamiento de velocidades intersísmicas de estaciones GNSS (e.g., Yáñez-Cuadra et al., 2023) y estimar acoplamiento intersísmico mediante aprendizaje supervisado (Barra et al., 2024). En este estudio, proponemos un nuevo enfoque para estimar deslizamiento cosísmico utilizando redes neuronales artificiales, donde definimos

información previa del deslizamiento a través de las características del conjunto de entrenamiento, de manera similar a Barra et al. (2024), para mejorar la estabilidad y abordar la complejidad del problema inverso mal condicionado. Nuestro modelo se entrena utilizando un conjunto diverso de escenarios sísmicos sintéticos, diseñados para reflejar una amplia gama de comportamientos de deslizamiento en fallas.

Una vez entrenado el modelo, evaluamos su efectividad en un escenario real estimando el deslizamiento cosísmico del terremoto de Illapel, Chile, ocurrido el 16 de septiembre de 2015 (M_w 8.3). Este terremoto, ampliamente percibido en la región, afectó un amplio segmento de la zona de subducción del centro de Chile a las 22:54:31 (UTC). Restringimos la distribución del deslizamiento utilizando desplazamientos cosísmicos GNSS disponibles en Klein et al. (2017). Además, realizamos un análisis de sensibilidad sobre diferentes hiperparámetros y condiciones de datos que impactan el rendimiento de la red neuronal. Este análisis es crucial para demostrar la sensibilidad de nuestro modelo a distintas configuraciones y optimizar su desempeño. Para validar nuestra metodología, la comparamos con la técnica tradicional de inversión por mínimos cuadrados regularizados, identificando y evaluando las diferencias en el desempeño entre ambos métodos.

1.2. Propósito de la tesis

En este trabajo se propone un modelo basado en redes neuronales para estimar distribución de deslizamiento cosísmico. El modelo se entrena utilizando datos sintéticos y se valida posteriormente con datos reales del terremoto de Illapel de 2015. Además, se analiza cómo distintos hiperparámetros y condiciones de los datos influyen en el rendimiento del modelo.

1.2.1. Objetivos de investigación

1.2.1.1. Hipótesis de investigación

Hipótesis principal: El modelo propuesto de redes neuronales puede estimar correctamente distribución de deslizamiento cosísmico a partir del entrenamiento con datos sintéticos de estaciones GNSS.

Hipótesis alternativa: El modelo propuesto no logra estimar adecuadamente distribución de deslizamiento cosísmico, mostrando errores significativos en

comparación con los métodos tradicionales.

1.2.1.2. Objetivo general

Desarrollar y validar un modelo de red neuronal para estimar distribución del deslizamiento cosísmico a partir de datos sintéticos, y aplicar dicho modelo a un caso real utilizando datos del terremoto de Illapel de 2015.

1.2.1.3. Objetivos específicos

1. Desarrollar un conjunto de datos sintéticos de deslizamiento cosísmico y desplazamientos en superficie utilizando modelos de dislocación elástica.
2. Entrenar un modelo de red neuronal utilizando los datos sintéticos generados para estimar deslizamiento cosísmico.
3. Validar el modelo de red neuronal aplicándolo a datos reales del terremoto de Illapel, y comparar los resultados con otros estudios del terremoto.

Capítulo 2

Manuscrito

Application of neural networks for estimating coseismic slip distribution using synthetic GNSS data

V. Inzunza[1], M. Moreno[2], V. Yáñez-Cuadra[3], F. Ortega-Culaciati[4], I. Calisto[1], M. Miller[1]

[1]Departamento de Geofísica, Universidad de Concepción, Chile
[2]Departamento de Ingeniería Estructural y Geotécnica, Pontificia Universidad Católica de Chile
[3]TerraSur Geofísica, Concepción, Chile
[4]Departamento de Geofísica, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile

Abstract

Estimating coseismic slip accurately and rapidly is crucial for understanding seismic events and improving early warning systems. In this study, we introduce a novel methodology using neural networks to estimate coseismic slip, using synthetic GNSS data for robust training. This approach showcases remarkable computational efficiency, with the model delivering accurate slip distribution predictions in just 0.07 seconds, without the need for specialized hardware. Validated with real-world data from the Mw 8.3 Illapel earthquake, the model demonstrated its effectiveness. Additionally, our

exploration of hyperparameters and data conditions not only evaluates the model's current performance but also identifies opportunities for future improvements.

2.1. Introduction

Plate tectonics, through the accumulation and release of strain, is responsible for the most powerful earthquakes on Earth, especially within subduction zones. These zones, where one tectonic plate slides under another, have the potential to cause devastating tsunamigenic earthquakes, such as the 2004 M_w 9.1 Sumatra-Andaman earthquake (e.g., Okal and Stein, 2009; Lay et al., 2005; Chlieh et al., 2007) and the 2011 M_w 9.0 Tohoku earthquake (e.g., Fujii et al., 2011; Ozawa et al., 2012; Tajima et al., 2013). In these convergent margins, elastic strain energy builds up over decades or centuries, with a significant fraction being subsequently released during earthquakes within the shallow part of the subduction interface. Chile is a highly active seismic region (e.g., Cisternas et al., 2017), where some of the largest-magnitude earthquakes ever recorded have occurred along the subduction zone where the Nazca Plate goes beneath the South American Plate at a rate of 66 mm/year (Altamimi et al., 2016). Notable recent events at this margin include the 1960 M_w 9.5 Valdivia Earthquake (e.g., Barrientos and Ward, 2007; Fujii and Satake, 2012; Lorenzo-Martín et al., 2006), the 2010 M_w 8.8 Maule Earthquake (e.g., Moreno et al., 2010, 2012; Bedford et al., 2013; Delouis et al., 2010), the 2014 M_w 8.2 Iquique Earthquake (e.g., Duputel et al., 2015; Jara et al., 2018; Meng et al., 2015) and the 2015 M_w 8.3 Illapel Earthquake (e.g., Melgar et al., 2016; Tilmann et al., 2015; Heidarzadeh et al., 2015).

Within the past few decades, the use of Global Navigation Satellite System (GNSS) stations has significantly improved the observation of both interseismic and coseismic surface displacements, advancing our understanding of earthquake dynamics. As GNSS stations have become more widespread globally, they now provide high-resolution measurements of crustal deformation, especially in subduction zones, allowing us to identify the characteristics and behaviors of large megathrust earthquakes (e.g., Ruegg et al., 2009; Chen et al., 2015; Luo et al., 2020). Estimating coseismic slip of earthquakes is essential to comprehending the underlying mechanics, including energy release, stress redistribution, and fault dynamics. This understanding is vital for seismic risk evaluation, developing mitigation strategies, and informing rapid response efforts (Iinuma et al., 2012).

Common approaches for estimating coseismic slip involve constructing a linear forward model through the generation of Green's functions based on elastic dislocations (e.g., Okada, 1985; Nikkhoo and Walter, 2015). Despite the linearity of the forward model, estimating coseismic slip remains a complex and ill-posed inverse problem, characterized by non-unique solutions. There are two end-member approaches to deal with the inverse problem. The first, known as the optimization approach, focuses on finding a solution to the inverse problem that minimizes a data misfit term and a regularization term, the latter used to define prior information that stabilizes the solution (e.g., Harris and Segall, 1987; Ortega-Culaciati et al., 2021). The second is a Bayesian approach, where an ensemble of models is sampled from a posterior probability distribution of slip (e.g., Minson et al., 2013; Duputel et al., 2014).

With the advance of technology and the proliferation of big data, Machine Learning (ML) algorithms have taken on a significant role in seismological problems. Recent progress includes the application of deep learning techniques for detecting low-frequency earthquakes (Münchmeyer et al., 2024) and denoising HR-GNSS data (Thomas et al., 2023), as well as the automatic identification of slow slip events (Donoso et al., 2021), tectonic analysis through the clustering of interseismic velocities from GNSS stations (e.g., Yáñez-Cuadra et al., 2023) and estimation of plate interface locking using supervised ML (Barra et al., 2024). In this study, we propose a novel approach to estimate coseismic slip by employing artificial neural networks, where we define prior information on slip through the characteristics of the training set – in a similar manner as Barra et al. (2024) – to improve stability and deal with the complexities of the ill-posed inverse problem. Our model is trained using a diverse set of synthetic earthquake scenarios, designed to reflect a wide range of fault slip behaviors.

Once the model is trained, we evaluate its effectiveness in a real-world scenario by estimating the coseismic slip of the September 16, 2015 (M_w8,3) Illapel, Chile, earthquake. Widely felt across the region, this earthquake shook a broad segment of the central Chilean subduction zone at 22:54:31 (UTC). We constrained the slip distribution using GNSS co-seismic displacements available from Klein et al. (2017). Furthermore, we conduct a sensitivity analysis on different hyperparameters and data conditions that impact the performance of the neural network. This analysis is crucial for demonstrating the sensitivity of our model to different settings and optimizing its configurations. To validate our methodology, we compare it with the

traditional Regularized Least Squares inversion technique to identify and evaluate the differences in performance between these two methods.

2.2. Methods

2.2.1. Synthetic GNSS data

We used the SLAB2.0 geometry defined by Hayes et al. (2018) to represent the subduction megathrust fault interface. This geometry is discretized into a triangular mesh consisting of 1002 elements across Cartesian coordinates ranging from approximately 29°S to 33°S along the Chilean coastline, corresponding to the area impacted by the Illapel earthquake. Green's functions, for modeling the surface displacements due to subsurface slip, are derived using the *TDdispHS* triangular dislocation model by Nikkhoo and Walter (2015).

For the generation of synthetic earthquakes, we adapted the code developed by Agnew (2013), originally designed to simulate synthetic GNSS time series for slow slip events, to incorporate a representation of surface displacements associated with earthquakes slip distributions. For simulating coseismic slip, we employed a linear forward model to relate coseismic slip on the fault interface, along both the strike and dip components, to the GNSS surface displacements. The synthetic dataset was generated at the positions of the 107 continuous GNSS stations provided by Klein et al. (2017), which include daily and survey data.

Synthetic earthquake slips are modeled as elliptical sources, with randomly generated characteristics to closely mimic real earthquake behaviors. The ellipse dimensions are set within specific ranges, while slip magnitudes are assigned within predefined limits. Rake angles are based on typical values observed in Chilean seismic subduction events, and Gaussian noise is added to the calculated surface displacements to reflect the standard deviation of the errors reported by GNSS stations. These synthetic earthquakes, with magnitudes ranging from approximately M_w 7 to M_w 9, have spatial extents that fall within the mesh dimensions. The ranges used for the properties of these synthetic earthquake sources are detailed in Table 2.2.1.

After training our model, we employed Monte Carlo error propagation to estimate uncertainties arising from measurement errors, which were assumed to follow independent Gaussian distributions with standard deviations based on actual

Properties	Minimum	Maximum
Slip (m)	5	20
Rake (°)	80	100
Length (km)	30	300

Cuadro 2.2.1: Properties of synthetic earthquake slips and ellipses: This table summarizes the minimum and maximum values for slip magnitude, rake angle, and length extent.

measurements. To assess these uncertainties, we generated 50,000 data realizations, a slightly larger number than that used by Barra et al. (2024).

2.2.2. Least Squares Inversion

For the quasi-static slip estimation problem, a linear relationship between observed data ($\mathbf{d}$) and model parameters ($\mathbf{m}$) is often represented as the forward model:

$$\mathbf{d} = \mathbf{G}\mathbf{m} \tag{2.2.1}$$

where the Green's function $\mathbf{G}$, defines a mapping from model parameters $\mathbf{m}$ into observable data $\mathbf{d}$ vector spaces. Thus, allowing us to predict the data for a particular model $\mathbf{m}$. The inverse problem, deals with the estimation of values of m, given experimental observations d and their relation defined by the forward model (e.g., Menke, 1989; Tarantola, 2005; Aster et al., 2013).

Among the various inversion techniques that can be applied to estimate a slip distribution, the Least Squares method is widely utilized due to its simplicity and ease of uncertainty quantification, as has an analytical solution for the estimated model and the covariance matrix representing its uncertainties. This method aims to find the model parameters $\mathbf{m}$ that minimize the sum of the squared differences between the observed data and their prediction by the forward model (Lawson and Hanson, 1974). As slip inversion is known to be a highly ill-posed problem, a regularization term is often added to the objective function of the Least Squares problem. In most cases, slip inversion relies in solving the Least Squares problem with Tikhonov Regularization,

$$\min_{\mathbf{m}} \|\mathbf{W_d}(\mathbf{G}\mathbf{m} - \mathbf{d})\|_2^2 + \varepsilon^2 \|\mathbf{H}\mathbf{m}\|_2^2 \tag{2.2.2}$$

where $\mathbf{W_d}$ is a weight matrix such that $\mathbf{W_d}^\top \mathbf{W_d} = \mathbf{C_d^{-1}}$, $\mathbf{C_d}$ is the covariance matrix

of the observed data $\mathbf{d}$, $\mathbf{H}$ is a regularization operator, and ε is a regularization parameter that needs to be determined using a model class selection technique (e.g., Craven and Wahba, 1979; Hansen and O'Leary, 1993). The solution of 2.2.2 can be written as,

$$\widetilde{\mathbf{m}} = (\mathbf{G^T C_d^{-1} G} + \varepsilon^{\mathbf{2}} \mathbf{H}^\top \mathbf{H})^{-1} \mathbf{G^T C_d^{-1} d} \tag{2.2.3}$$

$$\widetilde{\mathbf{C_m}} = (\mathbf{G^T C_d^{-1} G} + \varepsilon^{\mathbf{2}} \mathbf{H}^\top \mathbf{H})^{-1} \tag{2.2.4}$$

where $\widetilde{\mathbf{m}}$ is the estimated model parameters and $\widetilde{\mathbf{C_m}}$ the covariance matrix representing the uncertainties of the estimated model parameters. As a typical choice (e.g., Lohman, 2004; Delouis et al., 2010; Awaluddin et al., 2012; Tung and Masterlark, 2016; Yáñez-Cuadra et al., 2023), we use $\mathbf{H} = \boldsymbol{\nabla}^{\mathbf{2}}$, a Laplacian operator, to define a slip smoothing constraint, aimed to deal with the inherent instability of the slip inversion. We refer the reader to Ortega-Culaciati et al. (2021) and references therein, for further details on the Linear Least Squares method and on the effectiveness of the variety of regularization types that can be defined to deal with the inherent instabilities of the quasi-static slip estimation problem.

2.2.3. Artificial Neural Networks

Artificial Neural Networks (ANN) are a type of Machine Learning (ML) methods inspired by the human brain structure. ANN mimic the way neurons communicate with each other, aiming to solve problems in a manner similar to human cognitive processes.

These networks are structured in layers, each filled with a number of neurons. Typically, a neural network includes three interconnected layers: an input layer, one or more hidden layers, and an output layer. The input layer receives the data, and the hidden layers process the data through a series of computations, allowing the network to extract patterns and features from the input. Finally, the output layer delivers the final result, emulating a streamlined flow of information similar to neural pathways in the brain (e.g., Wang, 2003; Goodfellow et al., 2016).

During training of the neural network, an optimization algorithm iteratively adjusts its internal parameters (neurons weights and biases) to minimize a loss function. Typically, the totality of the training data, splitted into batches, pass multiple times through the network, with each complete cycle called an epoch. During each

epoch, the network makes predictions and then adjusts the neuron connection weights and biases based on the differences between predicted and actual outcomes; this adjustment is accomplished through a method known as backpropagation (Yegnanarayana, 2009). In the context of our study, we train the model to predict earthquake slip distributions, using synthetic earthquake scenarios, where surface crustal displacements at the location of GNSS sites and their causative fault slip distribution are initially provided. This allows the network to learn the relationship between the surface crustal displacements and the causative fault slip distribution. After training, the model can predict fault slip distributions from new displacement data it has not previously seen.

The architecture of our preferred model is straightforward, consisting of just one hidden layer with 100 neurons, alongside the input and output layers. The input layer receives surface crustal dispalcements at the location of GNSS sites, and the output layer predicts quantities that will be later translated into a fault slip distribution. To prevent overfitting, a dropout layer with a 35 % rate is implemented between the hidden layer and the output layer.

We employed dense layers in which each neuron connects to every neuron in both the previous and subsequent layers. For the activation function in the hidden layer, we used the GELU (Gaussian Error Linear Unit) function (Hendrycks D., 2016), which serves as a smoother alternative to ReLU (Krizhevsky et al., 2012). This function introduces probabilistic elements derived from the Gaussian distribution, making it particularly effective for more complex models where smoother outputs are beneficial. GELU is mathematically defined as:

$$\mathrm{GELU}(x) = x\,P(X \leq x) = x\,\Phi(x). \tag{2.2.5}$$

where $\Phi\,(x)$ represents the cumulative distribution function of the Gaussian. This function effectively weights the input x by the likelihood of x being less than or equal to itself under a standard Gaussian distribution. This non-linear activation helps the function to gradually adjust the outputs. For the output layer, we used the sigmoid function, a widely used non-linear activation function. The sigmoid function transforms the values into the range from 0 to 1. It can be defined as:

$$\mathrm{Sigmoid}(x) = \frac{1}{1 + e^{-x}}. \tag{2.2.6}$$

The model was implemented and trained using the *TensorFlow* library (Abadi et al., 2015), a widely used framework for building and training machine learning models. The training process consisted of 10 epochs with 300,000 synthetic cases. We partitioned the dataset into two main subsets: 80 % for training and 20 % for testing. Furthermore, from the training subset, 10 % was allocated for validation.

Before training, the data were normalized using min-max scaling to ensure that all values fell within the 0 to 1 range, thereby optimizing the efficiency of the learning process. After training, the neural network's output was rescaled using the same parameters that were applied during the initial scaling process. To evaluate the model's effectiveness, we used the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) as performance metrics, defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{2.2.7}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{2.2.8}$$

where y_i and $\hat{y}_i$ represent the actual and predicted slip values, respectively, and n is the total number of samples. The Mean Squared Error (MSE), defined as

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{2.2.9}$$

served as the loss function, measuring the differences between the slip predicted by the network and the actual values of slip in the training set. We used the *Adam* (Kingma and Ba, 2014) optimizer to find the optimal weights and biases of the neural network model.

2.3. Results

2.3.1. Synthetic case

Initially, we tested our model on the synthetic cases we developed. The model demonstrates the capability to capture the coseismic slip, aligning closely with the

anticipated magnitudes, with minor discrepancies in the slip distribution (as shown in Fig. 2.3.1). The predicted slip tends to be smaller and more concentrated in the center. This pattern, which appears consistently across all synthetic cases, indicates a slight underestimation in the slip distribution. The mean RMSE for the GNSS displacements of the synthetic cases is 0.13 m, while the mean MAE is 0.06 m.

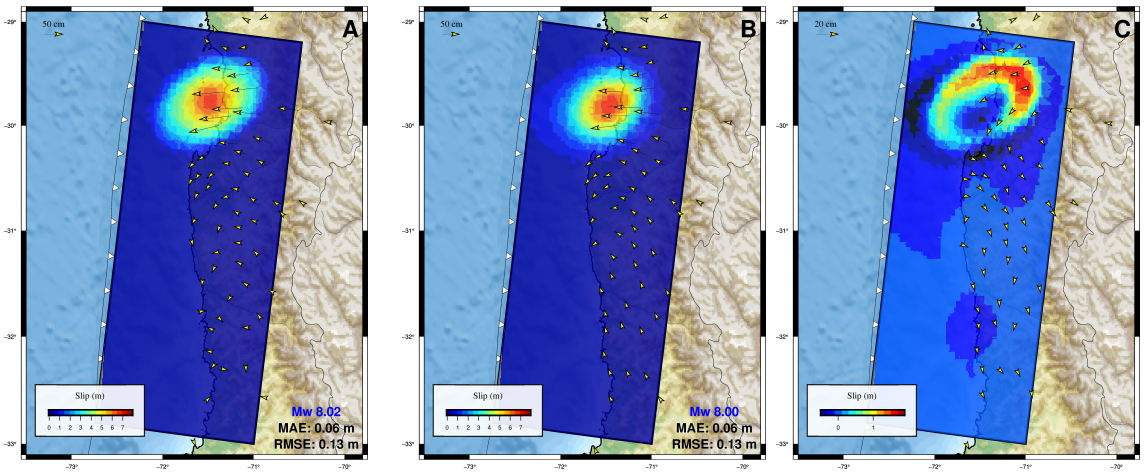


Figura 2.3.1: Example of a synthetic case: (a) Slip distribution with displacement vectors shown in yellow for a synthetic scenario; magnitude and RMSE details displayed in the bottom right; (b) Model prediction for the synthetic case; (c) Residuals and slip differences for the presented synthetic case.

2.3.2. Illapel case

Using our preferred model to estimate the coseismic slip of the Illapel earthquake (Figure 2), we found that the main event's slip distribution spans from 30.3°S to 31.9°S in latitude and from 72.6°W to 71.1°W in longitude. The peak coseismic slip reached 9 meters, with an along-strike rupture length of approximately 177 km (Figure 2.3.2). The slip distribution yields a geodetic seismic moment of $M_o = 3{,}35 \times 10^{21}$ Nm, equivalent to an $M_w = 8{,}32$ earthquake.

Regarding the GNSS displacement vectors in the horizontal components, they are generally well-recovered from the original data but are slightly more oriented towards the south and larger than expected. Similarly, the vertical components are accurately captured but also appear slightly larger than anticipated, as shown in Figure S1. Despite these discrepancies, the overall recovery of the displacement vectors is notably accurate in both components. For the horizontal components, the average residual is 3.5 cm, while for the vertical components, it is 1.6 cm, which we consider acceptable

given the maximum displacement values of approximately 2 meters.

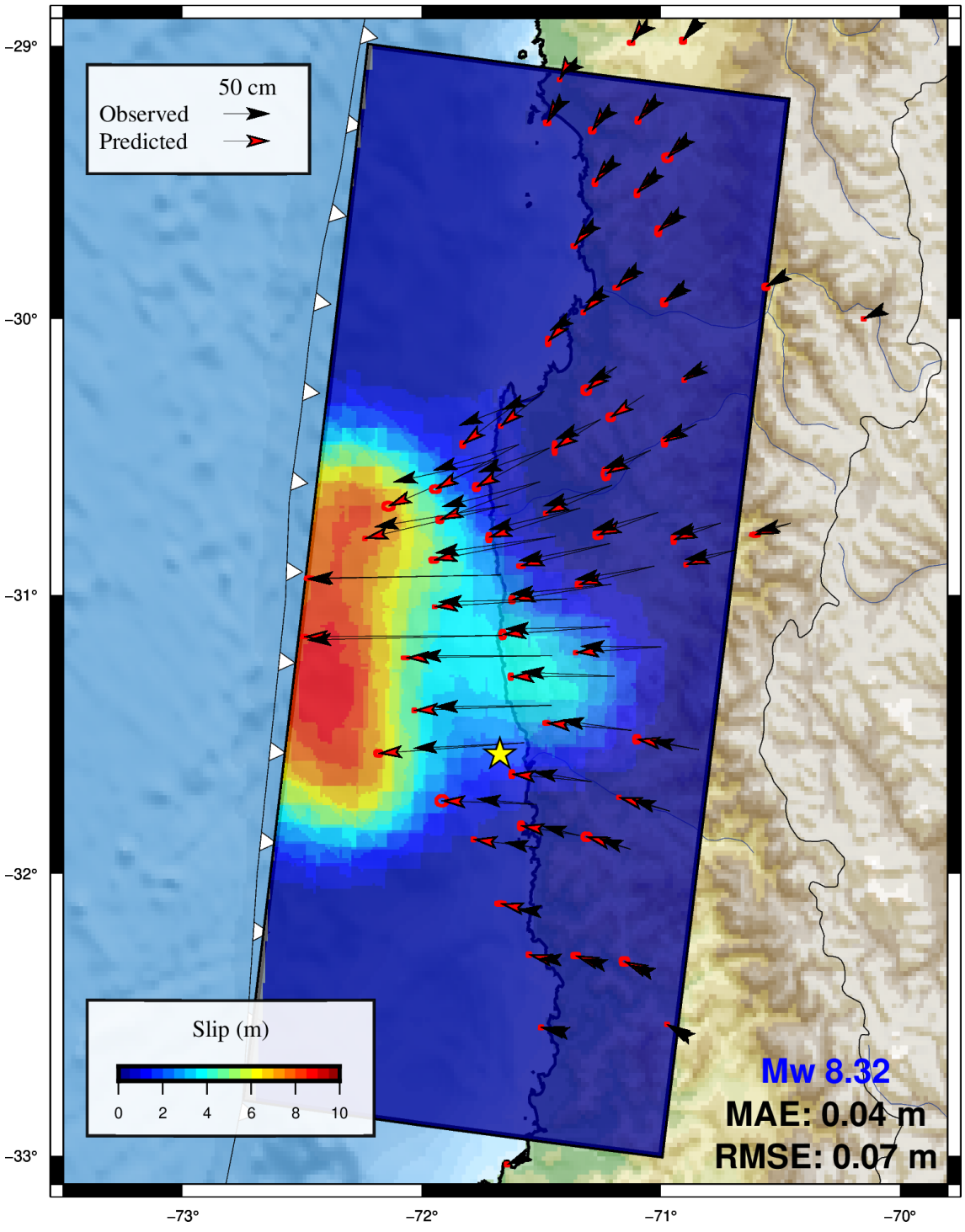


Figura 2.3.2: Estimated coseismic slip distribution for the Illapel earthquake using our preferred model, showing displacement vectors. Predicted vectors are in red, observed vectors are in black, and 95 % confidence intervals are included as red ellipses to illustrate the uncertainties in the model's predictions.

2.3.3. Analysis of impact of hyperparameters and data conditions

In this section, we demonstrate the impact of both model hyperparameters and experimental conditions on the estimation of coseismic slip for the Illapel case. To ensure a consistent comparison, we maintained the same data volume and architecture across all models, modifying only the necessary parameters or conditions. Specifically, we explored the influence of the choice of activation function for the hidden layer, the dropout rate, and the number of training epochs. Concurrently, we examined various experimental conditions critical to our study. We utilized a subset of only 13 GNSS stations, as described by Shrivastava et al. (2016), to understand the impact of the number of GNSS stations on the model's performance. We also varied the volume of synthetic cases used for training. Additionally, to assess the model's robustness, we introduced noise to the GNSS training data. Although such noise is not considered as a hyperparameter, its incorporation is crucial in defining the training environment and data characteristics that influence model training and performance. The results of this analysis are presented in Fig. 2.3.3 and additionally in Figures S2 to S9.

With respect to activation functions, Mish demonstrated the second-best performance after GELU (Figure 2.3.3a.), followed by SELU, while ReLU and Swish exhibited the poorest performance. This trend is reflected in both the metric values and residual analysis ash showed in Figures S2 and S3. Although most activation functions tend to overestimate slip, the GELU function used in the preferred model is capable of accurately recovering it. This can be attributed to GELU's smooth probabilistic curve, which preserves gradient information better than other activation functions.

The dropout rate plays a crucial role in enhancing the model's ability to generalize by mitigating overfitting. Intermediate dropout rates between 30 and 50 % perform better, as evidenced in Figures S4 and S5. In contrast, lower dropout rates, such as 10 to 20 %, lead to higher RMSE and MAE values, and higher residuals (Figure 2.3.3b.). Conversely, higher dropout rates help to smooth the results but can mask finer details in the data. A dropout rate of 35 % is considered to be the most effective in this case, achieving the best balance.

Changing the number of training epochs did not result in significant changes in RMSE, MAE, or residuals (Figures 2.3.3c, S6, and S7). The lack of a clear pattern suggests that increasing the number of epochs does not lead to better model performance.

This analysis identifies the activation function as the most critical hyperparameter for our model, followed by the dropout rate and the number of training epochs.

In terms of data conditions, the reduction in the number of GNSS stations negatively impacted the model's performance, leading to higher RMSE and MAE values. Despite this, the model was still able to reasonably estimate the slip's magnitude and distribution, although with larger uncertainties due to the sparsely-distributed stations (Figure 2.3.3d). As expected, the use of more synthetic cases resulted in better performance (Figures S8 and S9), indicating that at least 200,000 synthetic cases for training are needed for robust results (Figure 2.3.3e). Introducing noise confirmed the model's robustness, as it still accurately estimated coseismic slip and magnitude as shown in Figure 2.3.3f.

2.4. Discussion

The preferred neural network model, tested on the Illapel earthquake, is consistent with the slip distribution estimated in other studies. Our model's maximum slip (9 m) is consistent with the estimates from Melgar et al. (2016), Klein et al. (2017), Zhang et al. (2017) (10 m) and Shrivastava et al. (2016) (8 m), Williamson et al. (2017) (11 m) and Carrasco et al. (2019) (9 m). The along-strike length of the slip (177 km) is similar to other studies: Zhang et al. (2017) (170km); Klein et al. (2017), Melgar et al. (2016), Shrivastava et al. (2016) (200 km), Williamson et al. (2017) (125 km) and Carrasco et al. (2019) (180 km). Furthermore, the seismic moment in our model ($3{,}66 \times 10^{21}$ Nm, i.e. M_w 8.32) is consistent with the USGS W-phase model of $3{,}19 \times 10^{21}$ Nm, i.e. M_w 8.27.

Comparing our preferred model with an inversion using the Regularized Least Squares (RLS) method, we observed that our model exhibits higher residuals. Specifically, the root mean square (RMS) of the GNSS data residuals for our model is 6 cm, compared to 2.2 cm for the RLS inversion model. As illustrated in Fig. 3.1.1, the RLS inversion model consistently shows lower residuals across the study area. This discrepancy is particularly noticeable in the southern zone, where our model's residuals are significantly higher, as seen in Figure 3.1.1a and Figure 3.1.1b. The GNSS data residual graph (Figure 3.1.1c and Figure 3.1.1d) indicates that residuals from our model are more dispersed and include more extreme values, although their values are generally distributed around 0 cm. The vertical components exhibit lower residuals

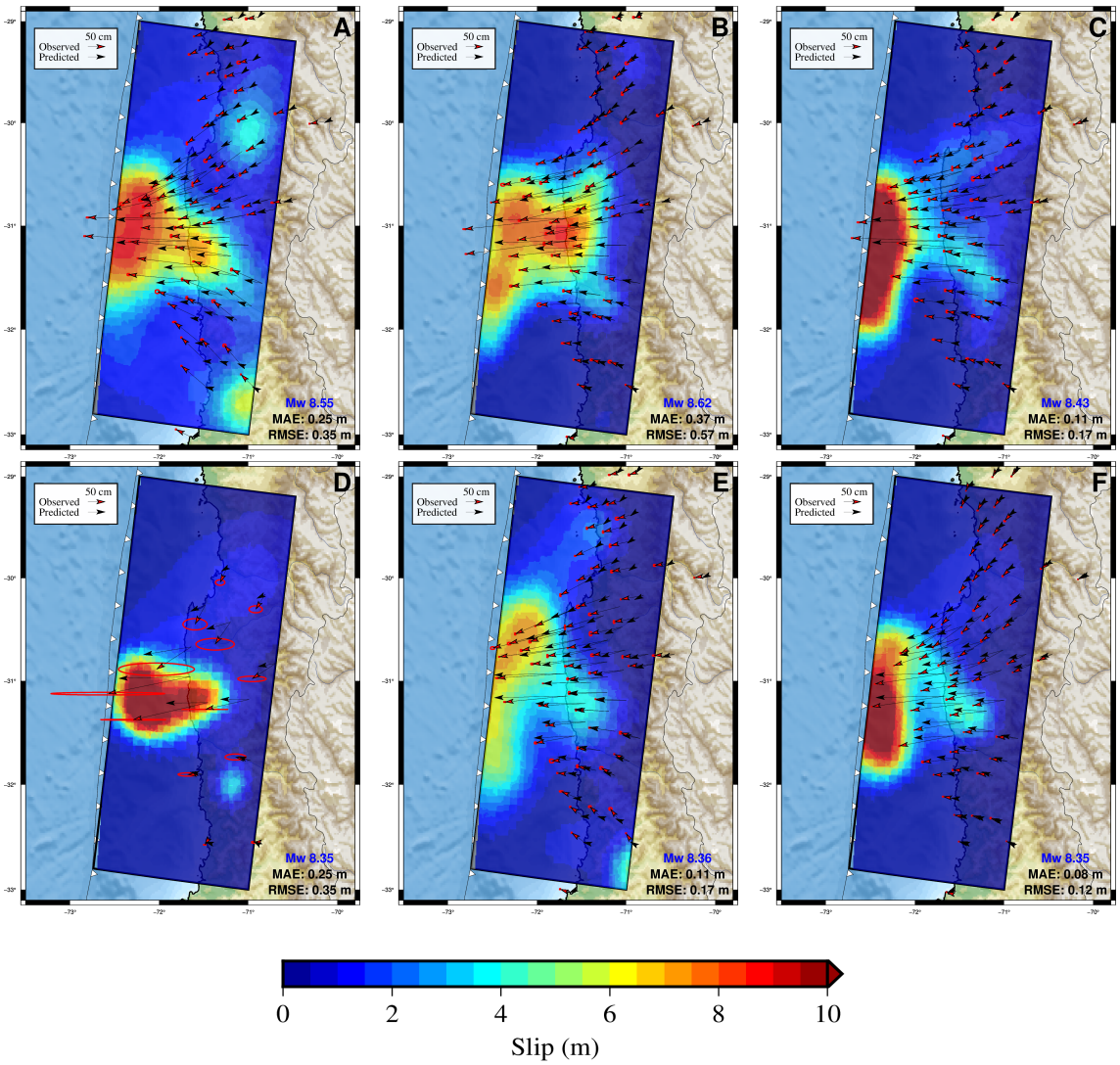


Figura 2.3.3: Impact of different hyperparameters and experimental conditions on coseismic slip estimation for the Illapel case. Each panel displays the predicted and observed displacement vectors, with predictions in red and observations in black. The panels also include 95 % confidence ellipses illustrating the uncertainty in the predictions. Subfigures show the effects of: (a) Mish activation function in the hidden layer, (b) 20 % dropout rate, (c) 50 training epochs, (d) using 13 only GNSS stations, (e) training with 200,000 synthetic cases, (f) introducing noise into the training data.

compared to the horizontal components, as detailed in Figure S10.

Furthermore, a significant observation from our model is the estimation of substantial slip close to the trench, contrasting with the inversion model's slip, which is positioned further away from the trench. This finding aligns with the slip patterns noted by Carrasco et al. (2019) and Caballero et al. (2023), where the slip is also estimated

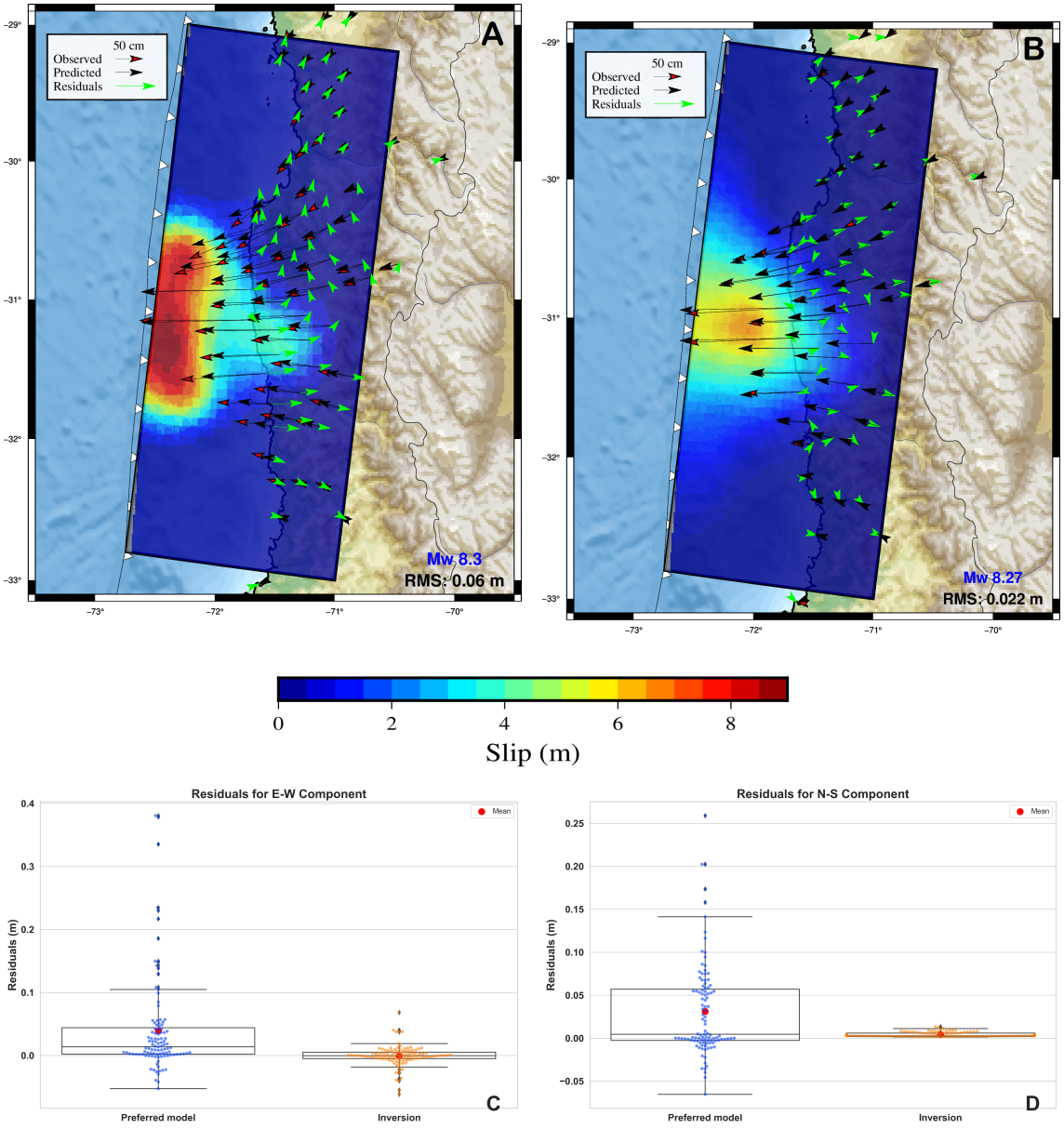


Figura 2.4.1: (a) Preferred model slip distribution with residuals highlighted in green, (b) Slip distribution from Regularized Least Squares inversion, (c) Comparison of East-West (E-W) component residuals for the preferred model (in blue) and classic inversion (in orange), (d) Comparison of North-South (N-S) component residuals across both models.

near the trench and is consistent with the tsunami observations of (Lay et al., 2016), underscoring a possible area-specific frictional behavior at the subduction megathrust that our model captures more effectively, even though is constrained only by onland GNSS observations.

2.5. Conclusions

In this study, we developed a novel neural network model that rapidly and accurately estimates coseismic slip. The model demonstrated high accuracy in predicting both the magnitude and spatial distribution of slip for synthetic cases, and was subsequently validated with the Mw 8.3 Illapel earthquake, achieving a GNSS displacement RMSE of 0.07 m and showing consistency with previously published solutions. This capability is essential for improving early warning systems and advancing real-time seismic hazard assessments.

A standout feature of our approach is its computational efficiency: once trained, the model generates slip estimations in just 0.07 seconds on a conventional computer without requiring iterative model tuning, significantly simplifying the estimation process. These computations were performed on a standard personal computer equipped with an AMD Ryzen 5 5600G processor and 16 GB of RAM, without the use of a dedicated GPU. This speed, combined with minimal hardware requirements, highlights the potential of neural networks to help with quick earthquake response and risk mitigation.

Performance analysis showed that the East-West component of slip had the highest residuals and variability, indicating more potential errors in this direction. The North-South component followed, while the vertical component showed the lowest variability, likely due to its smaller magnitude, making it easier for the neural network to learn. These patterns suggest specific areas where the model could be further improved.

Moreover, expanding the training dataset with more synthetic cases could potentially refine the model's predictive accuracy. Incorporating more realistic synthetic scenarios, particularly ellipses that closely resemble real cases, into the training process may lead to better outcomes and a more reliable model. Overall, while the model performs well, these improvements could help achieve even more precise results.

Capítulo 3

Conclusiones

3.1. Discusión General y Conclusión

El modelo de red neuronal preferido, aplicado al terremoto de Illapel, es consistente con la distribución de deslizamiento estimada en otros estudios. La magnitud máxima del deslizamiento estimada por nuestro modelo (9 m) coincide con los valores reportados por Melgar et al. (2016), Klein et al. (2017), Zhang et al. (2017) (10 m), Shrivastava et al. (2016) (8 m), Williamson et al. (2017) (11 m) y Carrasco et al. (2019) (9 m). La longitud a lo largo del rumbo estimada (177 km) también es similar a la reportada en otros estudios: Zhang et al. (2017) (170 km); Klein et al. (2017), Melgar et al. (2016), Shrivastava et al. (2016) (200 km), Williamson et al. (2017) (125 km) y Carrasco et al. (2019) (180 km). Además, el momento sísmico estimado por nuestro modelo ($3{,}66 \times 10^{21}$ Nm, equivalente a M_w 8.32) es consistente con el modelo W-phase de USGS, que reporta $3{,}19 \times 10^{21}$ Nm, equivalente a M_w 8.27.

Al comparar nuestro modelo con una inversión basada en el método de mínimos cuadrados regularizados (RLS), se observó que nuestro modelo presenta residuales mayores. En particular, la raíz del promedio cuadrático (RMS) de los residuales de datos GNSS para nuestro modelo es de 6 cm, en comparación con 2.2 cm para el modelo RLS. Como se ilustra en la Fig. 3.1.1, el modelo RLS muestra consistentemente residuales más bajos en toda el área de estudio. Esta discrepancia es especialmente notable en la zona sur, donde los residuales de nuestro modelo son significativamente mayores, como se observa en las Figuras 3.1.1a y 3.1.1b. El gráfico de residuales de datos GNSS (Figuras 3.1.1c y 3.1.1d) indica que los residuales de nuestro modelo

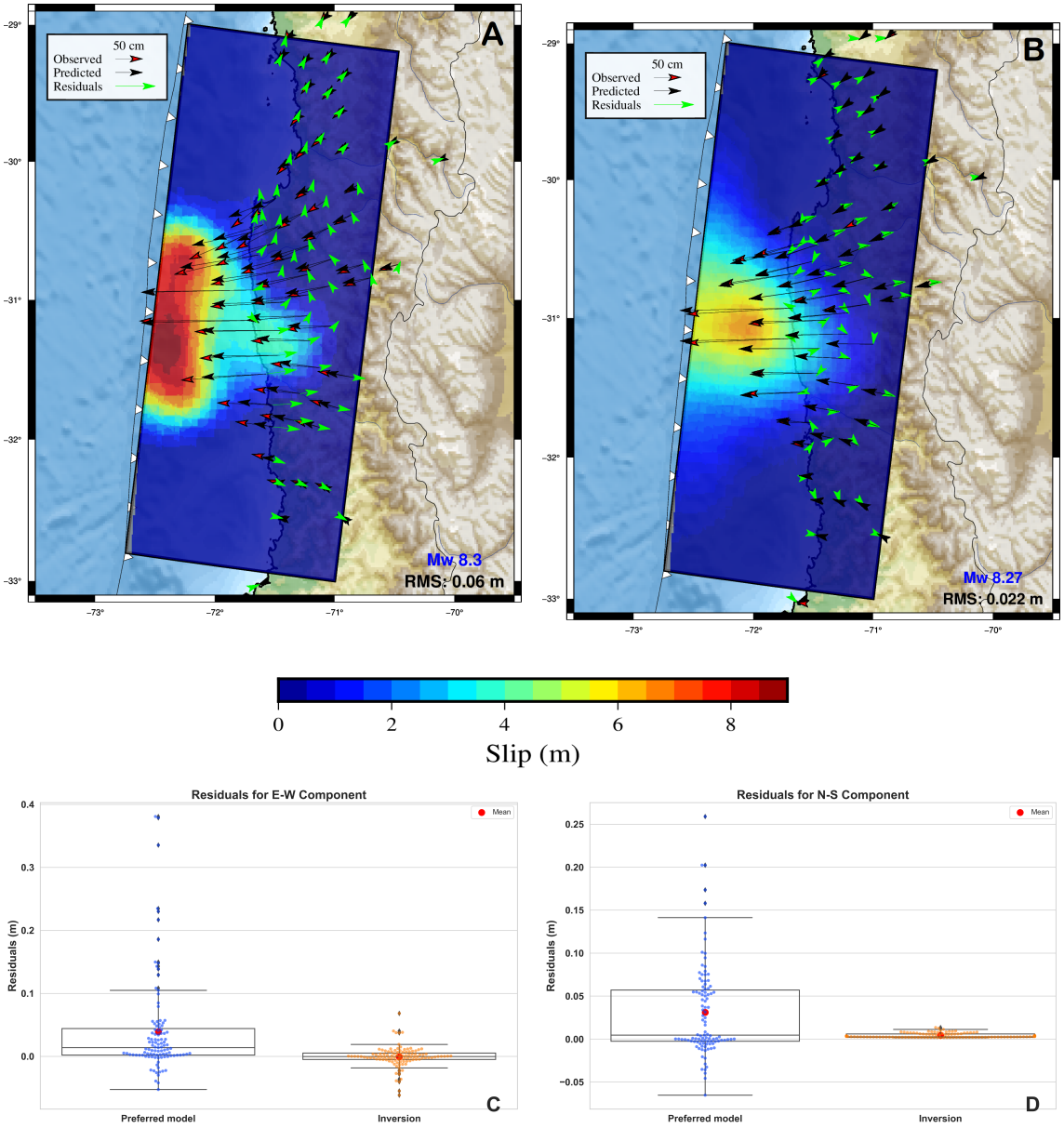


Figura 3.1.1: (a) Distribución de deslizamiento del modelo preferido con los residuales destacados en verde, (b) distribución de deslizamiento de la inversión por mínimos cuadrados regularizados, (c) comparación de los residuales en la componente Este-Oeste (E-O) entre el modelo preferido (en azul) y la inversión clásica (en naranja), (d) comparación de los residuales en la componente Norte-Sur (N-S) en ambos modelos.

están más dispersos e incluyen valores más extremos, aunque en general se distribuyen alrededor de 0 cm. Las componentes verticales presentan residuales más bajos en comparación con las componentes horizontales, como se detalla en la Figura S10.

Además, una observación significativa de nuestro modelo es la estimación de un deslizamiento sustancial cerca de la fosa, en contraste con el modelo de inversión, cuyo deslizamiento se posiciona más lejos de la fosa. Este hallazgo es consistente con los patrones de deslizamiento observados por Carrasco et al. (2019) y Caballero et al. (2023), donde también se estima el deslizamiento cerca de la fosa, concordando con las observaciones de tsunamis de Lay et al. (2016). Esto resalta un posible comportamiento friccional específico de la megafalla de subducción que nuestro modelo captura con mayor efectividad, incluso al estar restringido únicamente por observaciones GNSS en tierra.

En este estudio, desarrollamos un modelo novedoso de red neuronal capaz de estimar rápidamente y con precisión el deslizamiento cosísmico. El modelo demostró alta precisión en la predicción tanto de la magnitud como de la distribución espacial del deslizamiento en casos sintéticos, validándose posteriormente con el terremoto de Illapel de 2015 (M_w 8.3). El modelo alcanzó un RMSE de desplazamiento GNSS de 0.07 m y mostró consistencia con soluciones publicadas previamente. Esta capacidad es esencial para mejorar los sistemas de alerta temprana y avanzar en la evaluación del riesgo sísmico en tiempo real.

Una característica destacada de nuestro enfoque es su eficiencia computacional: una vez entrenado, el modelo genera estimaciones de deslizamiento en solo 0.07 segundos en un computador convencional, sin necesidad de ajustes iterativos, simplificando significativamente el proceso de estimación. Estas simulaciones se realizaron en un computador personal estándar equipado con un procesador AMD Ryzen 5 5600G y 16 GB de RAM, sin el uso de una GPU dedicada. Esta velocidad, combinada con los requisitos mínimos de hardware, resalta el potencial de las redes neuronales para asistir en la respuesta rápida a terremotos y en la mitigación de riesgos.

El análisis del rendimiento mostró que la componente Este-Oeste del deslizamiento presentó los residuales y la variabilidad más altos, indicando posibles errores mayores en esta dirección. Le siguió la componente Norte-Sur, mientras que la componente vertical mostró la menor variabilidad, probablemente debido a su menor magnitud, lo que facilita el aprendizaje de la red neuronal. Estos patrones sugieren áreas específicas

donde el modelo podría mejorarse aún más.

Asimismo, la ampliación del conjunto de entrenamiento con más casos sintéticos podría refinar la precisión predictiva del modelo. Incorporar escenarios sintéticos más realistas, particularmente elipses que se asemejen estrechamente a casos reales, en el proceso de entrenamiento puede conducir a mejores resultados y un modelo más confiable. En general, aunque el modelo muestra un buen desempeño, estas mejoras podrían ayudar a lograr resultados aún más precisos.

Bibliografía

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). Tensorflow: Large-scale machine learning on heterogeneous systems. https://www.tensorflow.org/. Software available from tensorflow.org.

Agnew, D. (2013). Realistic simulations of geodetic network data: The fakenet package. *Seismological Research Letters*, 84:426–432.

Altamimi, Z., Rebischung, P., Métivier, L., and Collilieux, X. (2016). Itrf2014: A new release of the international terrestrial reference frame modeling non-linear station motions: Itrf2014. *Journal of Geophysical Research: Solid Earth*, 121.

Aster, R. C., Borchers, B., and Thurber, C. H. (2013). *Parameter estimation and inverse problems*. Elsevier.

Awaluddin, M., Meilano, I., and Widiyantoro, S. (2012). Estimation of slip distribution of the 2007 bengkulu earthquake from gps observation using least squares inversion method. *ITB Journal of Engineering Science*, 44:186–205.

Barra, S., Moreno, M., Ortega-Culaciati, F., Benavente, R., Araya, R., Bedford, J., and Calisto, I. (2024). A supervised machine learning approach for estimating plate interface locking: Application to central chile. *Physics of the Earth and Planetary Interiors*.

Barrientos, S. and Ward, S. (2007). The 1960 chile earthquake - inversion for slip distribution from surface deformation. *Geophysical Journal International*, 103:589 – 598.

Bedford, J., Moreno, M., Baez, J., Lange, D., Tilmann, F., Rosenau, M., Heidbach, O., Oncken, O., Bartsch, M., Rietbrock, A., Tassara, A., Bevis, M., and Vigny, C. (2013). A high-resolution, time-variable afterslip model for the 2010 maule mw = 8.8, chile megathrust earthquake. *Earth and Planetary Science Letters*, 383:26–36.

Caballero, E., Duputel, Z., Twardzik, C., Rivera, L., Klein, E., Jiang, J., Liang, C., Zhu, L., Jolivet, R., Fielding, E., and Simons, M. (2023). Revisiting the 2015 m w

= 8.3 illapel earthquake: unveiling complex fault slip properties using bayesian inversion. *Geophysical Journal International*, 235:2828–2845.

Carrasco, S., Ruiz, J. A., Contreras-Reyes, E., and Ortega-Culaciati, F. (2019). Shallow intraplate seismicity related to the illapel 2015 mw 8.4 earthquake: Implications from the seismic source. *Tectonophysics*, 766:205–218.

Chen, K., Ge, M., Li, X., Babeyko, A., Ramatschi, M., and Bradke, M. (2015). Retrieving real-time precise co-seismic displacements with a standalone single-frequency gps receiver. *Advances in Space Research*, 56.

Chlieh, M., Avouac, J.-P., Hjörleifsdóttir, V., Song, T.-R., Ji, C., Sieh, K., Sladen, A., Hébert, H., Prawirodirdjo, L., Bock, Y., and Galetzka, J. (2007). Coseismic slip and afterslip of the great m_w 9.15 sumatra–andaman earthquake of 2004. *Bulletin of the Seismological Society of America*, 97.

Cisternas, M., Carvajal, M., Wesson, R., Ely, L., and Gorigoitia, N. (2017). Exploring the historical earthquakes preceding the giant 1960 chile earthquake in a time-dependent seismogenic zone. *Bulletin of the Seismological Society of America*, 107.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions - Estimating the correct degree of smoothing by the method of Generalized Cross-Validation. *Numerische Mathematik*, 31(4):377–403.

Delouis, B., Nocquet, J.-M., and Vallée, M. (2010). Slip distribution of the february 27, 2010 mw = 8.8 maule earthquake, central chile, from static and high-rate gps, insar, and broadband teleseismic data. *Geophysical Research Letters*, 37:L17305.

Donoso, F., Moreno, M., Ortega Culaciati, F., Bedford, J., and Benavente, R. (2021). Automatic detection of slow slip events using the picca: Application to chilean gnss data. *Frontiers in Earth Science*, 9.

Duputel, Z., Agram, P., Simons, M., Minson, S., and Beck, J. (2014). Accounting for prediction uncertainty when inferring subsurface fault slip. 197(1):464–482.

Duputel, Z., Jiang, J., Jolivet, R., Simons, M., Rivera, L., Ampuero, J. P., Riel, B., Owen, S., Moore, A., Samsonov, S., Ortega Culaciati, F., and Minson, S. (2015). The iquique earthquake sequence of april 2014: Bayesian modeling accounting for prediction uncertainty. *Geophysical Research Letters*, 42.

Fujii, Y. and Satake, K. (2012). Slip distribution and seismic moment of the 2010 and 1960 chilean earthquakes inferred from tsunami waveforms and coastal geodetic data. *Pure and Applied Geophysics*, 170.

Fujii, Y., Satake, K., Sakai, S., Shinohara, M., and Kanazawa, T. (2011). Tsunami source of the 2011 off the pacific coast of tohoku earthquake. *Earth, Planets, and Space*, 63:815–820.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org.

Hansen, P. C. and O'Leary, D. P. (1993). The use of the L-curve in the regularization of discrete ill-posed problems. 14(6):1487–1503.

Harris, R. A. and Segall, P. (1987). Detection of a locked zone at depth on the parkfield, california, segment of the san andreas fault. *Journal of Geophysical Research*, 92(B8):7945–7962.

Hayes, G., Moore, G., Portner, D., Hearne, M., Flamme, H., Furtney, M., and Smoczyk, G. (2018). Slab2, a comprehensive subduction zone geometry model. *Science*, 362.

Heidarzadeh, M., Murotani, S., Satake, K., Ishibe, T., and Gusman, A. (2015). Source model of the 16 september 2015 illapel, chile, mw 8.4 earthquake based on teleseismic and tsunami data. *Geophysical Research Letters*, 43:n/a–n/a.

Hendrycks D., Gimpel, K. (2016). Gaussian error linear units (gelus). *arXiv Preprint arXiv:1606.08415*.

Iinuma, T., Hino, R., Kido, M., Inazu, D., Osada, Y., Ito, Y., Ohzono, M., Tsushima, H., Suzuki, S., Fujimoto, H., and Miura, S. (2012). Coseismic slip distribution of the 2011 off the pacific coast of tohoku earthquake (m9.0) refined by means of seafloor geodetic data. *Journal of Geophysical Research (Solid Earth)*, 117:7409–.

Jara, J., Sánchez-Reyes, H., Socquet, A., Cotton, F., Virieux, J., Maksymowicz, A., Diaz-Mojica, J., Walpersdorf, A., Ruiz, J., Cotte, N., and Norabuena, E. (2018). Kinematic study of iquique 2014 m 8.1 earthquake: Understanding the segmentation of the seismogenic zone. *Earth and Planetary Science Letters*, 503:131–143.

Kingma, D. and Ba, J. (2014). Adam: A method for stochastic optimization. *International Conference on Learning Representations*.

Klein, E., Vigny, C., Fleitout, L., Grandin, R., Jolivet, R., Rivera, E., and Métois, M. (2017). A comprehensive analysis of the illapel 2015 mw8.3 earthquake from gps and insar data. *Earth and Planetary Science Letters*, 469.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C. J., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.

Lawson, C. L. and Hanson, R. J. (1974). *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, New Jersey.

Lay, T., Kanamori, H., Ammon, C., Nettles, M., Ward, S., Aster, R., Beck, S., Bilek, S., Brudzinski, M., Butler, R., Deshon, H., Ekström, G., Satake, K., and Sipkin, S. (2005). The great sumatra-andaman earthquake of 26 december 2004. *Science (New York, N.Y.)*, 308:1127–33.

Lay, T., Li, L., and Cheung, K. F. (2016). Modeling tsunami observations to evaluate a proposed late tsunami earthquake stage for the 16 september 2015 illapel, chile,

m w 8.3 earthquake: Test of tsunami earthquake scenario. *Geophysical Research Letters*, 43.

Lohman, R. (2004). The inversion of geodetic data for earthquake parameters.

Lorenzo-Martín, F., Roth, F., and Wang, R. (2006). Inversion for rheological parameters from post-seismic surface deformation associated with the 1960 valdivia earthquake, chile. *Geophysical Journal International - GEOPHYS J INT*, 164:75–87.

Luo, H., Ambrosius, B., Russo, R., Mocanu, V., Wang, K., Bevis, M., and Fernandes, R. (2020). A recent increase in megathrust locking in the southernmost rupture area of the giant 1960 chile earthquake. *Earth and Planetary Science Letters*, 537:116200.

Melgar, D., Fan, W., Riquelme, S., Geng, J., Liang, C., Fuentes, M., Easton, G., Allen, R., Shearer, P., and Fielding, E. (2016). Slip segmentation and slow rupture to the trench during the 2015, mw 8.3 illapel, chile earthquake. *Geophysical Research Letters*, 43.

Meng, L., Huang, H., Burgmann, R., Ampuero, J. P., and Strader, A. (2015). Dual megathrust slip behaviors of the 2014 iquique earthquake sequence. *Earth and Planetary Science Letters*, 411.

Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, San Diego, California.

Minson, S., Simons, M., and Beck, J. (2013). Bayesian inversion for finite fault earthquake source models i — theory and algorithm. *Geophysical Journal International*, pages 1–36.

Moreno, M., Melnick, D., Rosenau, M., Baez, J., Klotz, J., Oncken, O., Tassara, A., Chen, J., Bataille, K., Bevis, M., Socquet, A., Bolte, J., Vigny, C., Brooks, B., Ryder, I., Grund, V., Smalley, B., Carrizo, D., Bartsch, M., and Hase, H. (2012). Toward understanding tectonic control on the mw 8.8 2010 maule chile earthquake. *Earth and Planetary Science Letters*, s 321–322:152–165.

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

Münchmeyer, J., Giffard-Roisin, S., Malfante, M., Frank, W., Poli, P., Marsan, D., and Socquet, A. (2024). Deep learning detects uncataloged low-frequency earthquakes across regions. *Seismica*, 3.

Nikkhoo, M. and Walter, T. R. (2015). Triangular dislocation: an analytical, artefact-free solution. *Geophysical Journal International*, 201:1119–1141.

Okada, Y. (1985). Surface deformation to shear and tensile faults in a halfspace. *Bulletin of the Seismological Society of America*, 75.

Okal, E. and Stein, S. (2009). Observations of ultra-long period normal modes from the 2004 sumatra–andaman earthquake. *Physics of The Earth and Planetary Interiors - PHYS EARTH PLANET INTERIORS*, 175:53–62.

Ortega-Culaciati, F., Simons, M., Ruiz, J., Rivera, L., and Díaz-Salazar, N. (2021). An epic tikhonov regularization: Application to quasi-static fault slip inversion. *Journal of Geophysical Research: Solid Earth*, 126(7):e2020JB021141.

Ozawa, S., Nishimura, T., Munekane, H., Suito, H., Kobayashi, T., Tobita, M., and Imakiire, T. (2012). Preceding, coseismic, and postseismic slips of the 2011 tohoku earthquake, japan. *Journal of Geophysical Research (Solid Earth)*, 117:7404–.

Ruegg, J., Rudloff, A., Vigny, C., Madariaga, R., Chabalier, J., Campos, J., Kausel, E., Barrientos, S., and Dimitrov, D. (2009). Interseismic strain accumulation measured by gps in the seismic gap between constitución and concepción in chile. *Physics of The Earth and Planetary Interiors*, 175:78–85.

Shrivastava, M., González, G., Moreno, M., Chlieh, M., Salazar, P., Reddy, C., Baez, J., Yanez, G., González-Carrasco, J., and De la Llera, J. (2016). Coseismic slip and afterslip of the 2015 m w 8.3 illapel (chile) earthquake determined from continuous gps data: Illapel earthquake 2015. *Geophysical Research Letters*, 43.

Tajima, F., Mori, J., and Kennett, B. (2013). A review of the 2011 tohoku-oki earthquake (mw 9.0): Large-scale rupture across heterogeneous plate coupling. *Tectonophysics*, 586:15–34.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics.

Thomas, A., Melgar, D., Dybing, S., and Searcy, J. (2023). Deep learning for denoising high-rate global navigation satellite system data. *Seismica*, 2.

Tilmann, F., Zhang, Y., Moreno, M., Saul, J., Eckelmann, F., Palo, M., Deng, Z., Babeyko, A., Chen, K., Baez, J., Schurr, B., Wang, R., and Dahm, T. (2015). The 2015 illapel earthquake, central chile: A type case for a characteristic earthquake? *Geophysical Research Letters*, 43.

Tung, S. and Masterlark, T. (2016). Coseismic slip distribution of the 2015 mw7.8 gorkha, nepal, earthquake from joint inversion of gps and insar data for slip within a 3-d heterogeneous domain. *Journal of Geophysical Research: Solid Earth*, 121.

Wang, S.-C. (2003). *Artificial Neural Network*.

Williamson, A., Newman, A., and Cummins, P. (2017). Reconstruction of coseismic slip from the 2015 illapel earthquake using combined geodetic and tsunami waveform data: Reconstruction of illapel coseismic slip. *Journal of Geophysical Research: Solid Earth*, 122.

Yegnanarayana, B. (2009). *Artificial Neural Networks*. PHI Learning Pvt. Ltd., New Delhi.

Yáñez-Cuadra, V., Moreno, M., Ortega Culaciati, F., Donoso, F., Baez, J., and Tassara, A. (2023). Mosaicking andean morphostructure and seismic cycle crustal deformation patterns using gnss velocities and machine learning. *Frontiers in Earth Science*, 11.

Zhang, Y., Zhang, G., Hetland, E., Shan, X., Wen, S., and Zuo, R. (2017). *Coseismic Fault Slip of the September 16, 2015 Mw 8.3 Illapel, Chile Earthquake Estimated from InSAR Data*, pages 73–82.

Apéndice A

Material suplementario

A1. Evolución del entrenamiento del modelo

En este apartado se muestra la evolución del modelo durante 10 épocas, evaluada mediante la función de pérdida (*loss*) y la raíz del error cuadrático medio (RMSE)

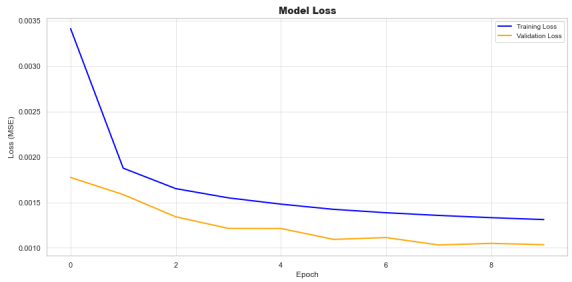


Figura A1.1: Evolución de la función de pérdida (*loss*) del modelo a lo largo de 10 épocas de entrenamiento.

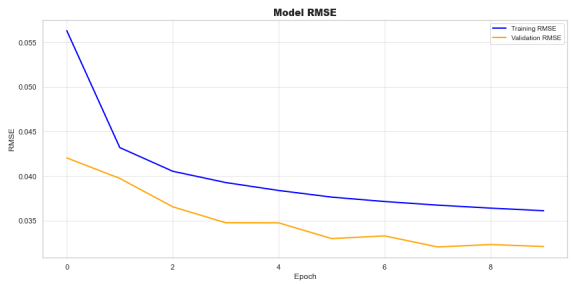


Figura A1.2: Evolución de la raíz del error cuadrático medio (RMSE) del modelo a lo largo de 10 épocas de entrenamiento.

A2. Análisis de hiperparámetros y condiciones de datos

En esta sección se evalúa cómo distintos hiperparámetros del modelo y condiciones de los datos influyen en la estimación del deslizamiento cosísmico para el caso del terremoto de Illapel. Los hiperparámetros analizados incluyen las funciones de activación, las tasas de *dropout* y el número de épocas de entrenamiento. Por otro lado, las condiciones experimentales abarcan el número de estaciones GNSS, el volumen de casos sintéticos y la incorporación de ruido en los datos de entrenamiento.

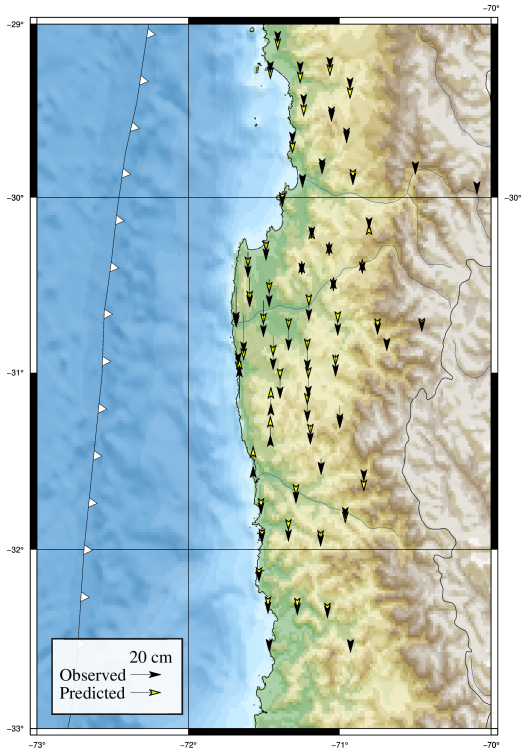
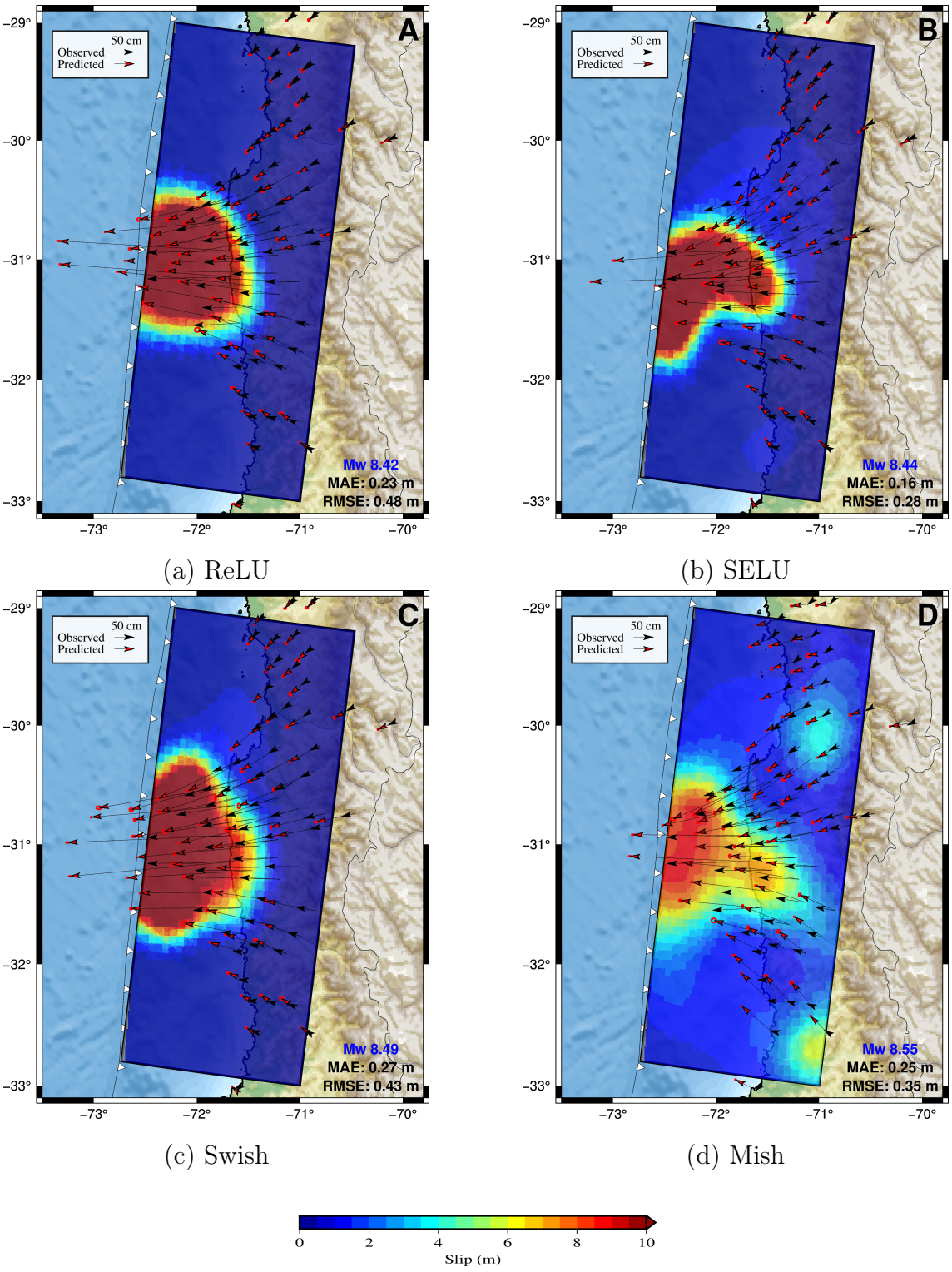


Figura A2.1: Vectores de desplazamiento vertical para Illapel: predicciones (amarillo) y observaciones (negro).



(a) ReLU (b) SELU

(c) Swish (d) Mish

Figura A2.2: Impacto de diferentes funciones de activación en la capa oculta: (a) ReLU, (b) SELU, (c) Swish, (d) Mish. Vectores rojos: predicciones, vectores negros: datos.

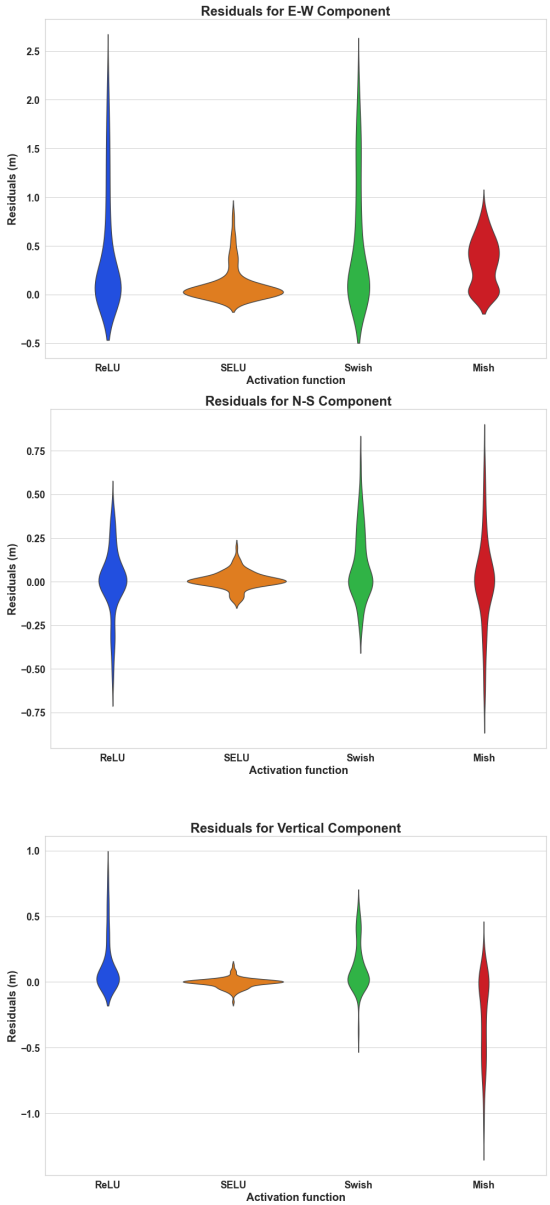


Figura A2.3: Residuales con diferentes funciones de activación: superior (componente E-W), medio (componente N-S) e inferior (componente vertical).

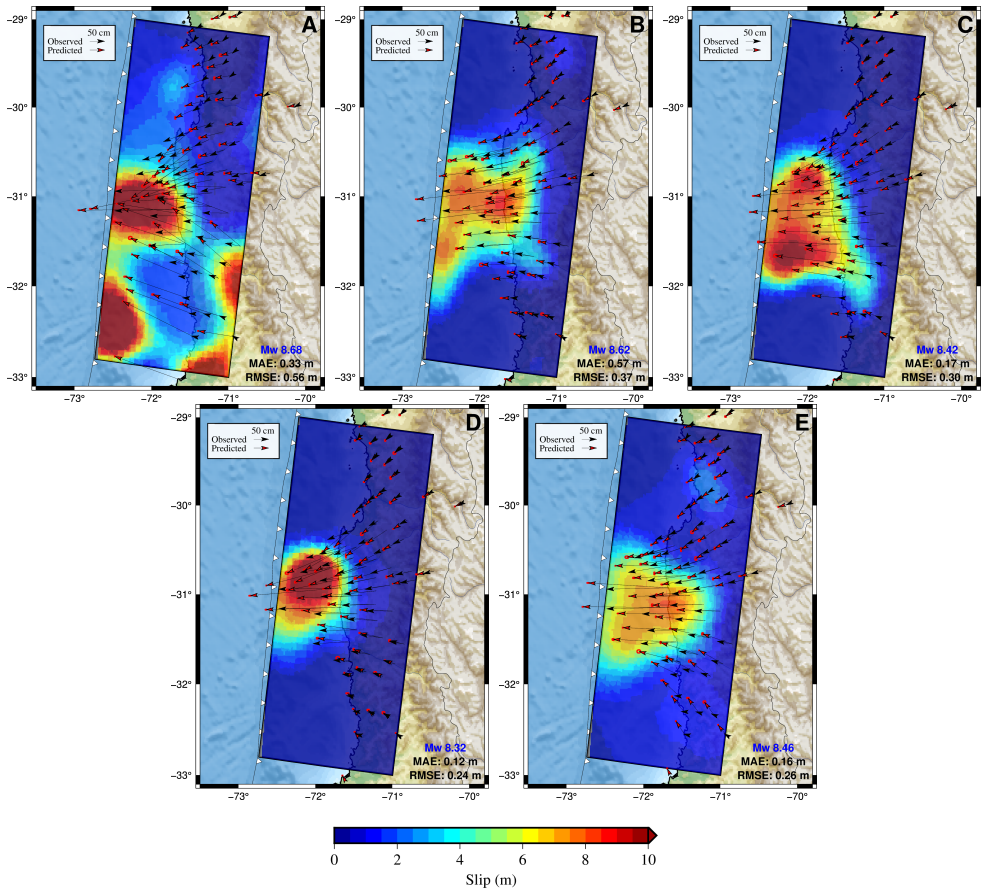


Figura A2.4: Impacto de las tasas de dropout en la distribución de deslizamiento: (a) 10 %, (b) 20 %, (c) 30 %, (d) 40 %, (e) 50 %.

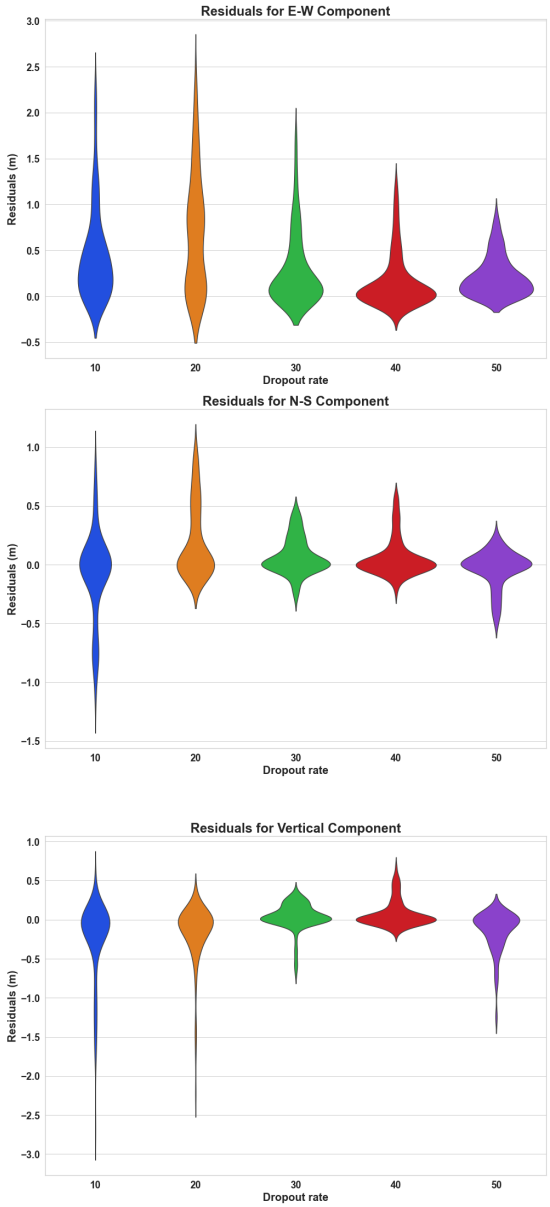


Figura A2.5: Residuales con diferentes tasas de dropout para: superior (componente E-W), medio (componente N-S) e inferior (componente vertical).

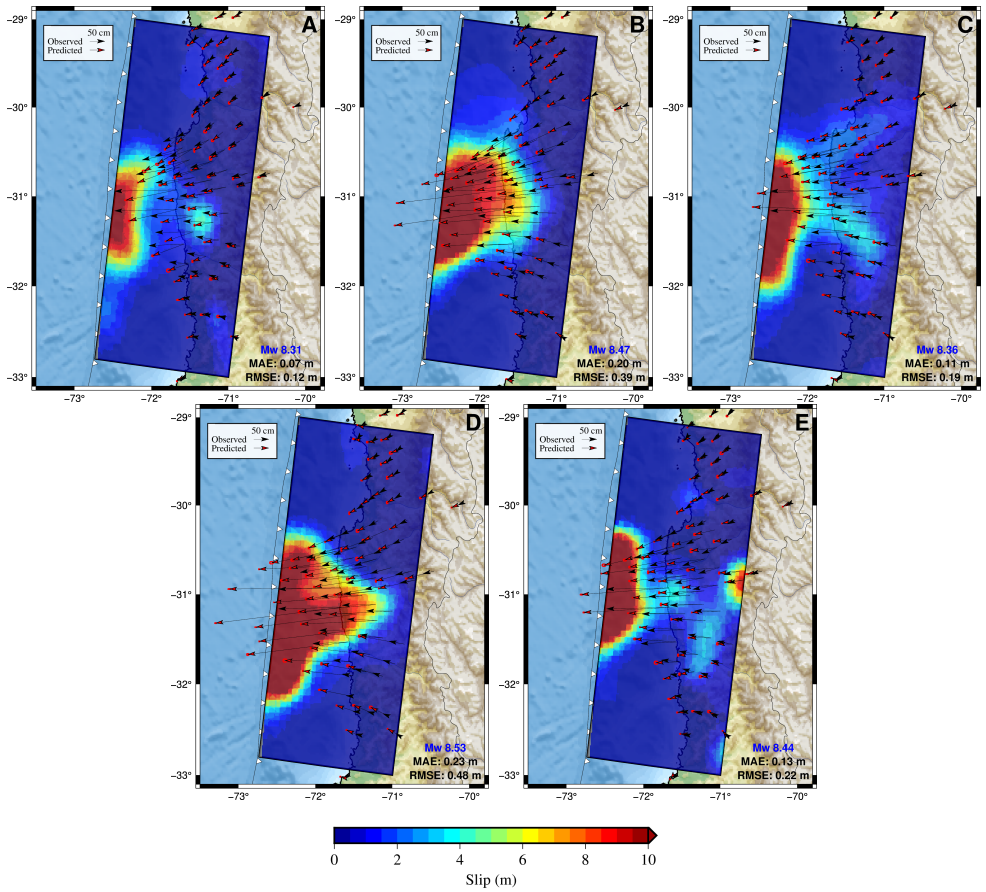


Figura A2.6: Impacto de las épocas de entrenamiento en la distribución de deslizamiento: (a) 20, (b) 30, (c) 50, (d) 80, (e) 100 épocas.

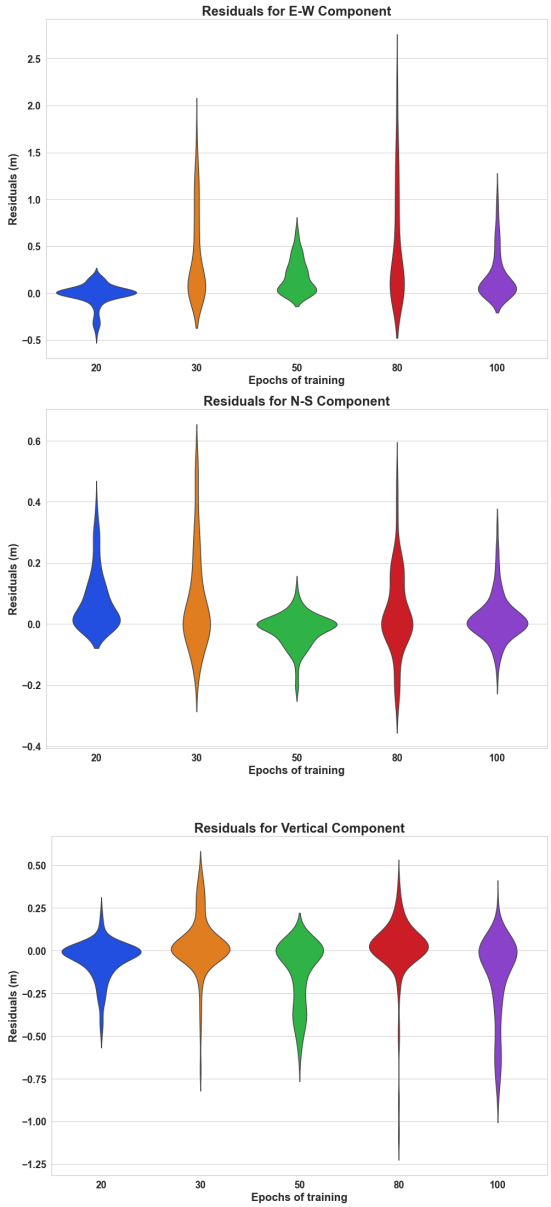


Figura A2.7: Residuales para diferentes épocas: arriba (componente E-W), en medio (componente N-S), y al fondo (componente vertical).

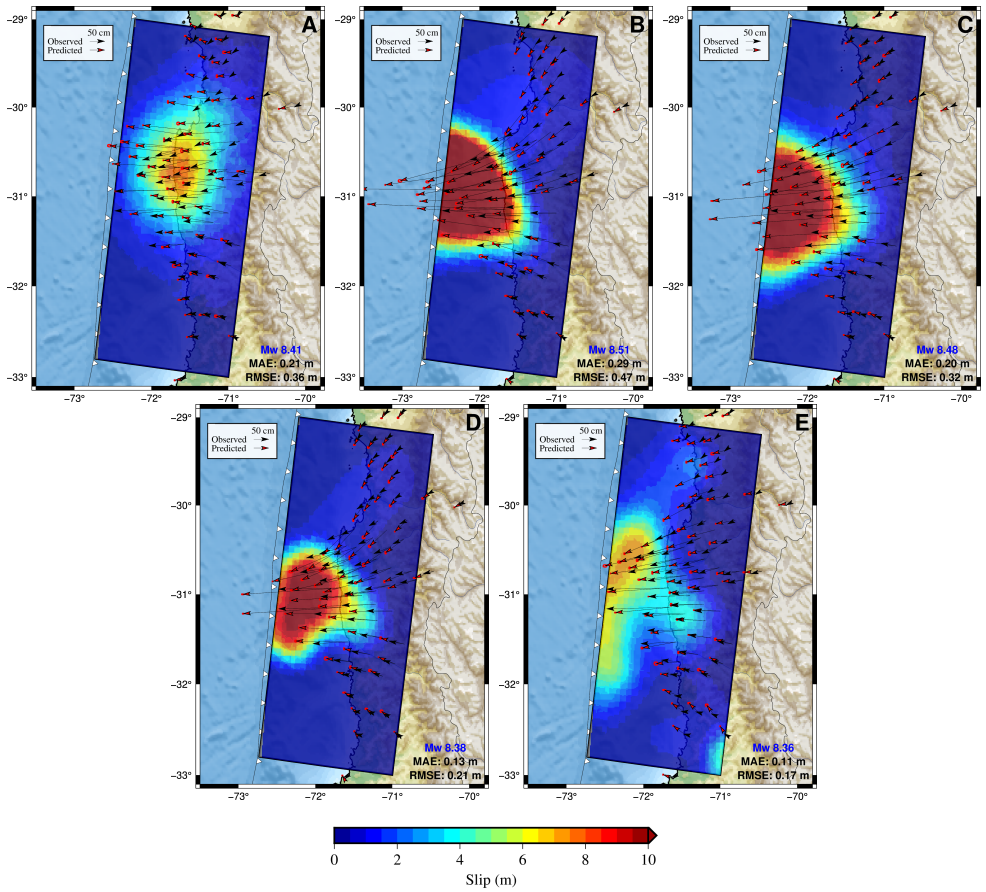


Figura A2.8: Impacto de los casos sintéticos en la distribución del deslizamiento: (a) 5,000, (b) 15,000, (c) 50,000, (d) 100,000, (e) 200,000 casos.

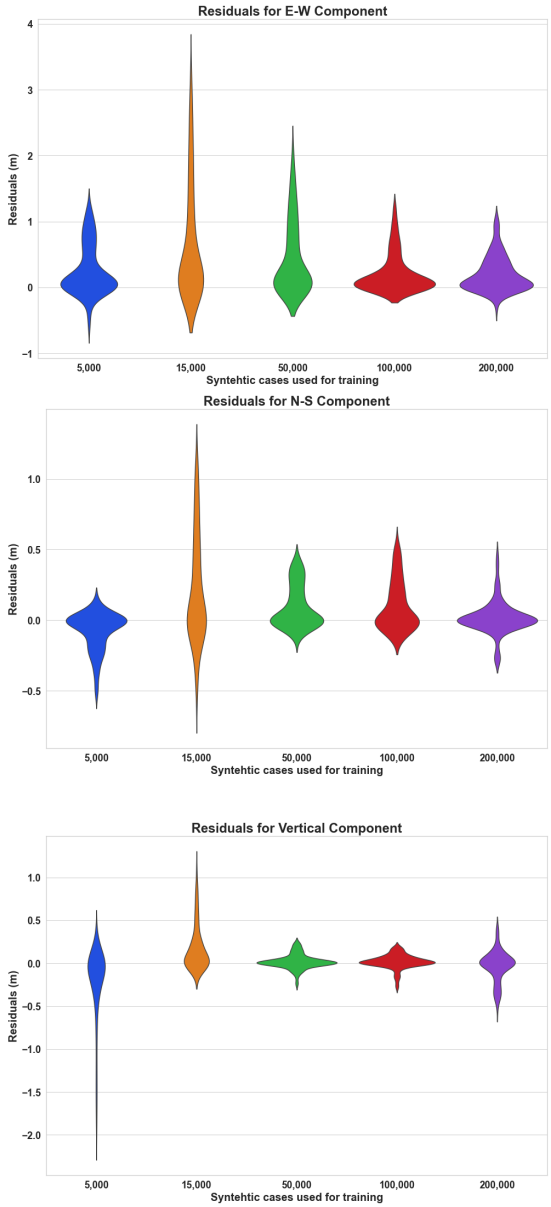


Figura A2.9: Residuales con diferentes cantidades de casos sintéticos utilizados para el entrenamiento: arriba (componente E-W), en medio (componente N-S), y al fondo (componente vertical).