

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
MAGÍSTER EN GEOFÍSICA



Estimación automática de deslizamientos en la interfase de subducción usando datos GNSS y
Machine Learning

Valentina Inzunza García

Pre-proyecto de Tesis para en el marco del Magíster de Geofísica (**Programa Articulación**)

2022



I. DATOS GENERALES DEL PROYECTO

PROGRAMA:			
TÍTULO DEL PROYECTO: Estimación automática de deslizamientos en la interfase de subducción usando datos GNSS y Machine Learning			
TÍTULO DEL PROYECTO (INGLÉS): Automatic estimation of slip on the interface of subduction using GNSS data and Machine Learning			
NOMBRE DE LA/EL ESTUDIANTE: Valentina Inzunza		AÑO DE INGRESO: 2022	
LÍNEA DE INVESTIGACIÓN: Modelación		DISCIPLINA PRINCIPAL:	
FECHA DE RECEPCIÓN EN EL PROGRAMA:	FECHA DE DEFENSA DE PROYECTO DE TESIS:	FECHA DE LA INSCRIPCIÓN EN LA DIR. DE POSTGRADO	FECHA ESTIMADA DEL EXAMEN DE GRADO:
COMITÉ DE TESIS			

Profesor(a) Guía Nombre y Apellidos Depto- Universidad			
Miembro del Comité Nombre y Apellidos Depto.-Universidad	Miembro del Comité Nombre y Apellidos Depto. -Universidad	Miembro del Comité Nombre y Apellidos Depto.-Universidad	

Nombre y Apellidos Directora Magíster en Geofísica Departamento de Geofísica Universidad de Concepción			



II. FORMULACIÓN DEL PROYECTO

a. Fundamento teórico y estado del arte

Chile es un país caracterizado por su alta actividad sísmica debido a su ubicación sobre la zona de subducción donde la placa de Nazca subducta bajo la Sudamericana con una velocidad de convergencia de 66 milímetros por año (e.g Altamimi et al., 2016), por lo que está constantemente expuesto a la ocurrencia de eventos sísmicos tales como los terremotos de Maule 2010 (e.g Moreno et al., 2010), Iquique 2014 (e.g Ruiz et al., 2014), Illapel 2015 (Shrivastava et al. 2016). Es por esto, que resulta importante monitorear y estudiar los terremotos para tener un mejor entendimiento de ellos.

En un sistema de subducción las placas tectónicas se mueven en direcciones contrarias a cada lado de la zona de contacto. Con el tiempo, la tensión que se genera por el roce entre estas placas se va acumulando hasta un momento en donde la tensión supera el límite de fricción. En ese momento, la falla se rompe provocando una ruptura, produciéndose un desplazamiento rápido y liberando la tensión acumulada, generando un terremoto. Para la ocurrencia de terremotos, existe un periodo en donde ocurre un proceso de acumulación de energía, y posteriormente un periodo donde se libera esa energía.

El ciclo sísmico se refiere a este proceso de acumulación y liberación de energía en la corteza terrestre. Este se divide en tres periodos: intersísmico, cosísmico y postsísmico

El período intersísmico corresponde al intervalo de tiempo donde se acumula la deformación en la zona de contacto entre un terremoto y otro, tiene una duración de decenas a cientos de años). En un ambiente de subducción, se refiere al tiempo en que la placa oceánica se mantiene acoplada con la placa continental, es decir, sin poder deslizar libremente. En el período cosísmico se produce la dislocación en la zona de contacto produciéndose un terremoto, donde se libera una gran cantidad de energía en un breve lapso, tiene una duración de segundos a minutos. En este período se produce el deslizamiento cosísmico del terremoto, es decir la cantidad de desplazamiento que se produce a lo largo de la falla. Finalmente, el periodo postsísmico comprende la evolución de la corteza justo después de haber culminado el terremoto. En este periodo se genera un deslizamiento en la zona de contacto que se prolonga por días, meses y años luego de ocurrido el evento principal.

El uso de estaciones GNSS (*Sistema Global de Navegación por Satélite*) han permitido un aumento en la calidad y cantidad de observaciones de desplazamientos en la superficie de la corteza terrestre, ayudando al progreso del estudio de los terremotos y el ciclo sísmico (e.g Baez et al., 2018). Cuando las observaciones GNSS se miden a lo largo del tiempo, los cambios en las posiciones de las estaciones entregan valiosa información sobre la acumulación de tensiones en la corteza.

A partir de las observaciones de estaciones GNSS, se puede obtener valiosa información con relación al estudio de los terremotos, como el deslizamiento cosísmico o *slip* de un terremoto.

Determinar el deslizamiento es importante ya que permite comprender la mecánica de los terremotos, como la forma en que se libera la energía, como se transfiere la tensión y como se deslizan las fallas. Además, el deslizamiento de un terremoto sirve como referencia en el estudio del tamaño potencial de futuros terremotos a lo largo de la misma falla.

La estimación del deslizamiento cosismico se hace a partir de velocidades cosísmicas obtenidas de estaciones GNSS, las cuales se utilizan para resolver el problema inverso de estimar un deslizamiento en profundidad a partir de movimientos en superficie. Esto se ha realizado en múltiples trabajos para distintos



terremotos como Tohoku 2011 (e.g Ozawa et al., 2012), Maule 2010 (e.g Pollitz et al., 2011) e Illapel 2015 (e.g Shrivastava et al., 2016)

Grandes terremotos como el de Maule de 2010 (e.g Moreno et al., 2010) y el de Tohoku de 2011 (e.g Loveless et al., 2011) tienen registro que las rupturas de segmento de falla estaban al menos, parcialmente acopladas antes de la ocurrencia del terremoto. Por esto, parece importante evaluar el acoplamiento intersísmico y su variación espacial para evaluar el potencial sísmico (Bejar-Pizarro et al., 2013).

El acoplamiento se define como el grado de bloqueo de una falla durante el periodo de acumulación de tensiones entre eventos sísmicos (en el periodo intersísmico). A partir de mediciones geodésicas, el acoplamiento indica un estado de ausencia o bajo slip (Métois et al., 2012). Si una falla de límite de placa se desliza al ritmo de convergencia de placas a largo plazo, podemos decir que está desacoplada. Un fuerte acoplamiento entre placas significa que la falla está bloqueada y es capaz de acumular tensiones, mientras que un acoplamiento débil significa que la falla está desbloqueada o sólo es capaz de acumular tensiones bajas (Eos Transactions, 2011)

El patrón de acoplamiento intersísmico suele ser heterogéneo y, hasta cierto punto, coincide con el patrón de los grandes terremotos (Kaneko et al., 2010). La determinación del acoplamiento intersísmico puede proporcionar valiosa información sobre la dinámica de la falla, incluyendo la cantidad de energía que se está acumulando y la tasa a la que se está acumulando. De esta forma, puede ser utilizado para estimar la probabilidad de un terremoto en una falla determinada y, por lo tanto, para evaluar el riesgo sísmico en una zona. Es por esto que es importante generar mapas de acoplamiento para tener una aproximación de la ubicación y la magnitud de la acumulación de energía que podría liberarse en un terremoto futuro.

Se han realizado distintos estudios en referencia al acoplamiento intersísmico; buscando relacionar la acumulación del stress intersísmico y la liberación cosísmica durante el terremoto del Maule 2010 (e.g Moreno et al., 2010); modelación de acoplamiento intersísmico en la zona de subducción de los Andes Central encontrando que las áreas con un gran slip coinciden con las áreas bloqueadas en el periodo intersísmico. (e.g Chlieh et al., 2011); determinación de acoplamiento intersísmico en el Himalaya a partir de mediciones geodésicas (e.g Stevens et al., 2015).

El cálculo del acoplamiento intersísmico y del deslizamiento cosísmico se realiza resolviendo el problema inverso dado movimientos en la superficie medidos por estaciones GNSS. Para resolver el problema inverso se necesitan un modelo físico para calcular la respuesta del medio a una dislocación y las observaciones. Este problema de estimación es conocido por ser un problema mal condicionado ya que puede haber varias soluciones posibles. Dentro de las metodologías para resolver el problema inverso están los enfoques; optimal (e.g Yañez, 2020) y Bayesiano (e.g Duputel et al., 2015)

Para el enfoque optimal se requiere introducir un término de regularización para estimar el mejor modelo. Uno de los métodos más utilizados es el de los mínimos cuadrados regularizado. El método de mínimos cuadrados regularizado se utiliza para controlar la complejidad del modelo y prevenir el sobreajuste. El método se basa en minimizar la siguiente expresión:

$$\min_m \|W_x(G_m - d)\| - \xi^2 \|w_h(H_m - h_0)\|$$

Donde:

- m es el conjunto de datos.



- $\|W_x(G_m - d)\|$ es la norma del producto de la matriz de pesos w_x y la diferencia entre la salida del modelo G_m y la salida que se busca d . Representa la regularización del ajuste de los datos de entrada.
- ξ es el parámetro de regularización, que se utiliza para controlar la complejidad del modelo.
- $\|w_h(H_m - h_0)\|$ es la norma del producto de la matriz de pesos w_h y la diferencia entre los valores de salida del modelo para las características de entrada H_m y los valores de salida que se busca h_0 . Representa la regularización de las características del modelo.

En la técnica de mínimos cuadrados regularizado, se agrega un término de penalización (ξ) adicional al método de mínimos cuadrados clásico para controlar la complejidad del modelo. El término de penalización se calcula como la suma de los valores absolutos o cuadrados de los coeficientes del modelo, multiplicados por un parámetro de regularización que controla la importancia de la penalización en la función de coste.

El parámetro de regularización se ajusta para encontrar el equilibrio correcto entre el ajuste del modelo y la complejidad.

En este trabajo se buscará implementar un método alternativo al problema inverso para determinar deslizamientos de terremotos y mapas de acoplamiento, sin la necesidad de resolver el problema inverso, el cual es un problema que se considera mal condicionado.

Machine Learning es un subcampo de las ciencias de la computación y una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas que permitan que los computadores aprendan por sí mismos, sin la necesidad de ser explícitamente programados. Hoy en día, los algoritmos de Machine Learning se utilizan para diversas tareas como: clasificación, regresión, reconocimiento de imágenes y del habla, asistentes virtuales, detección de fraudes, filtrado de correo electrónico, entre otros.

Una de las herramientas más populares en Machine Learning son las redes neuronales, que se basan en la estructura y funcionamiento del cerebro humano para procesar información. Las redes neuronales son un conjunto de algoritmos que simulan la forma en que el cerebro humano procesa información a través de neuronas interconectadas.

En los últimos años, la disponibilidad de grandes conjuntos de datos combinada con la mejora de los algoritmos y el crecimiento exponencial de la tecnología ha provocado un aumento del interés por temas involucrados con Machine Learning. De esta misma manera, se está convirtiendo en una de las herramientas más interesantes para la ciencia. Ha demostrado ser capaz de acelerar considerablemente tanto la investigación básica como la aplicada. En la actualidad, se está teniendo un *boom* de trabajos que desarrollan y aplican Machine Learning a distintos problemas científicos (Schdmit et al., 2018).

Con la aparición de los algoritmos de Machine Learning, la comunidad científica dispone de un nuevo conjunto de herramientas para abordar el problema relacionado a terremotos (Johnson et al., 2019).

Los avances en esta materia incluyen: la predicción de terremotos de laboratorio (e.g Rouet-Leduc et al., 2017), modelación sismo-tectónica analógica que permite la simulación de múltiples ciclos sísmicos en un tiempo y escala experimentales convenientes (e.g Corbi et al., 2018), la estimación de la fricción de las fallas a escala de laboratorio (e.g Rouet-Leduc et al., 2018), la predicción de las tasas de desplazamiento del GPS asociadas a eventos de slip lento (e.g Rouet-Leduc et al., 2018), y la predicción de localizaciones de réplicas de terremotos (e.g DeVries et al., 2018).



Debido a las ventajas que traen los algoritmos de Machine Learning para el avance científico, se le dará provecho a esta tecnología para proponer un nuevo método de estimación de deslizamientos cosísmicos y de acoplamiento intersísmico.

b. Objetivos e hipótesis

Objetivo general: Estimar deslizamientos en la interfase de subducción usando Machine Learning para determinar movimientos cosísmicos e intersísmicos

Objetivos específicos:

- Generar datos sintéticos de desplazamiento
- Generar set de datos para entrenar el modelo
- Evaluación comparativa de distintas estrategias de Machine Learning
- Evaluar performance del modelo respecto a inversiones estándares

Hipótesis:

El modelo permite estimar distribuciones de deslizamiento en la interfase de subducción de forma robusta.

Esto permitirá su aplicación a estimación de magnitud y ubicación de grandes terremotos de forma rápida, dando paso a una posible utilización para un sistema de alerta temprana.

Otra aplicación de este método permitirá estimar distribuciones de deslizamiento cosísmico



c. Metodología

1. Generación de datos sintéticos

La generación de datos sintéticos se hizo a partir de un código Python de generación de deslizamientos de terremotos sintéticos escrito por F. Ortega.

Las funciones de Green se generaron usando la geometría de *Slab 2.0* (Hayes et al., 2018). Se usa una geometría de coordenadas cartesianas desde los 18°S hasta los 45°S de Chile.

La malla utilizada está discretizada en 1.667 triángulos y los desplazamientos de la superficie se estimaron utilizando *TDdispHS*, el cual es un código analítico de dislocación triangular (Nikkhoo et al., 2015).

Los deslizamientos de los terremotos sintéticos se hacen a partir de fuentes elípticas, la base de datos se construye a partir de una generación aleatoria en un espacio donde varía el tamaño de la elipse, su grado de rotación, la ubicación, además del deslizamiento del terremoto sintético.

2. Redes neuronales

Las redes neuronales son un tipo de modelo de Machine Learning inspirado en la estructura y funcionamiento del cerebro humano, este imita la forma en que las neuronas se comunican entre sí. El objetivo de la red neuronal es resolver los problemas de la misma manera que lo hace el cerebro humano (Wang et al., 2003).

Las redes neuronales para este trabajo se utilizarán en primera instancia, para estimar deslizamientos cosísmico de estaciones GNSS a partir de sus desplazamientos cosísmicos. Esto se hará a partir de la base de estaciones GNSS sintéticas. Tras entrenar la red para estaciones sintéticas y tener un modelo consolidado, se entrenará nuevamente la red para datos de estaciones GNSS reales para evaluar el desempeño del modelo.

Las redes neuronales se organizan en varias capas llenas de neuronas. Estas capas normalmente son tres: una capa de entrada, donde se ingresan los datos de entrada; una o más capas ocultas; y una capa de salida, que entregará el resultado final de la red. Los datos de entrada de la primera capa se propagan desde cada neurona hasta cada neurona de la capa siguiente, para que al final se envíe un resultado desde la capa de salida.

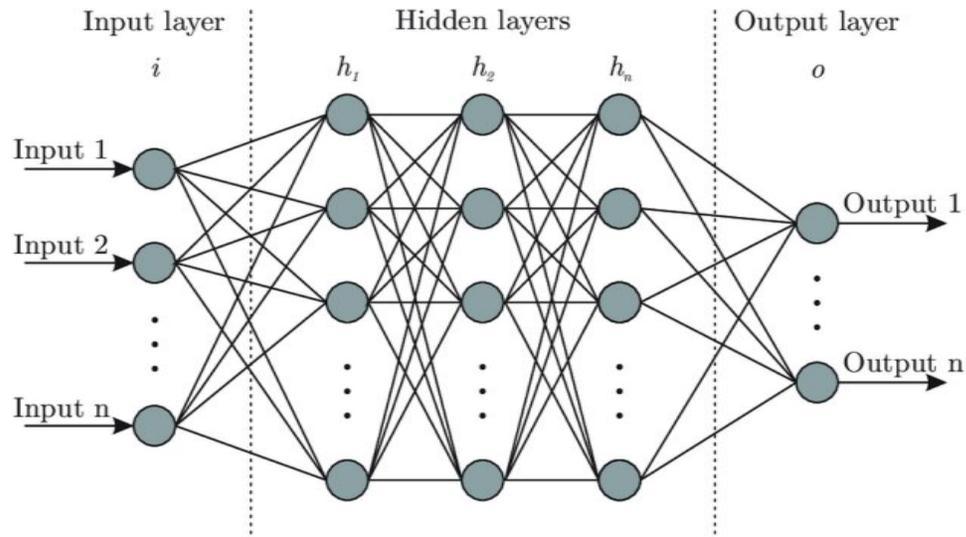


Figura 1: Ejemplo de estructura de una red neuronal (Sharma, 2020)

Los datos de entrada en este caso serán los desplazamientos cosísmicos y como salida se obtendrán los deslizamientos generados o *slip* por el terremoto. De esta forma, a partir de la información entregada, la red realizará estimaciones de deslizamiento a partir de los desplazamientos entregadas. Cabe mencionar que a modo de que la red aprenda de manera más fácil, antes de ingresar los datos se escalaron a un rango entre 0 y 1, de la siguiente forma:

$$\frac{x - x_{max}}{x_{max} - x_{min}}$$

Donde:

- x es el conjunto de datos.
- x_{max} el valor máximo del conjunto de datos.
- x_{min} el valor mínimo del conjunto de datos.

La red aprende a través del entrenamiento, examinando los registros individuales, generando una predicción para cada época (un ciclo completo de entrenamiento) y realizando ajustes a los pesos cuando realiza una predicción incorrecta. En un problema supervisado, se van presentando los datos a modo de ejemplo, donde el resultado ya es conocido (estos datos serían el conjunto de entrenamiento), y se comparan las predicciones obtenidas con los resultados ya conocidos. La información obtenida de esta comparación se envía hacia atrás a través de la red, cambiando los pesos de las neuronas gradualmente (lo que se conoce como *backpropagation*). A medida que progresa el entrenamiento, la red se va haciendo cada vez más precisa. En general, mientras más entrenamiento tenga la red, mejor serán las predicciones. Una vez entrenada, la red se puede aplicar a casos futuros en los que se desconoce el resultado (Lapedes et al., 1998).

Las neuronas realizan cálculos matemáticos y ajustan los pesos de las conexiones entre ellas a medida que se les presenta un conjunto de datos de entrenamiento. Los pesos son valores numéricos asociados con las conexiones entre las neuronas.



Para entrenar el modelo, se utilizará el 80% de los datos como conjunto de entrenamiento y un 20% de los datos como conjunto de prueba o test para evaluar el desempeño con datos que no ha visto. Además, se considera el 10% de los datos como conjunto de validación, es decir, un conjunto de datos separado del conjunto de entrenamiento y de test para evaluar el rendimiento del modelo durante el proceso de entrenamiento.

El tipo de capas que se usarán en este caso serán las capas densas. En este tipo de capas, cada neurona de la capa está conectada a todas las neuronas de la capa anterior y de la capa siguiente.

La ecuación matemática que define una capa densa:

$$y = f(Wx + b)$$

Donde:

- y es el vector de salida de la capa.
- x es el vector de entrada de la capa.
- W es la matriz de pesos que conecta las neuronas de la capa anterior con las neuronas de la capa actual.
- b es el vector de sesgo que se suma al producto de Wx .
- f es la función de activación aplicada a la salida de la capa.

La función de activación es una función matemática que se aplica a la salida de una neurona o de una capa de neuronas con el propósito de introducir no linealidad en el modelo. Las funciones de activación permiten a la red modelar funciones complejas y no lineales (Sharma et al., 2020).

Existen varias funciones de activación utilizadas en redes neuronales, como la función sigmoide, la función ReLU, la función tangente hiperbólica, la función softmax, entre otras. Cada función de activación tiene sus propias propiedades y es adecuada para diferentes tipos de problemas. En este caso se usará la función ReLU y ELU

La función de activación ReLU (*Rectified Linear Unit*) es una de las funciones de activación más utilizadas debido a su simplicidad y eficiencia computacional. Es una función no lineal que se aplica a la salida de cada neurona en una capa de la red neuronal.

La función ReLU se define como:

$$f(x) = \max(0, x)$$

Esto significa que si la entrada (x) es positiva, la función de activación ReLU devuelve la misma entrada. Si la entrada es negativa, la función devuelve cero.

Por su lado, la función de activación ELU (*Exponential Linear Unit*) es una función no lineal, es similar a la función ReLU, pero con una curva suave para valores negativos, lo que la hace más apropiada para problemas en los que las entradas pueden tomar valores negativos.

La función ELU se define como:

$$f(x) = x, \quad \text{si } x \geq 0$$

$$f(x) = \alpha(\exp(x) - 1) \quad \text{si } x < 0$$

Donde α es un parámetro controla la tasa de crecimiento de la función para valores negativos.

En la figura 2 se puede apreciar la diferencia entre la función de activación ReLU y ELU:

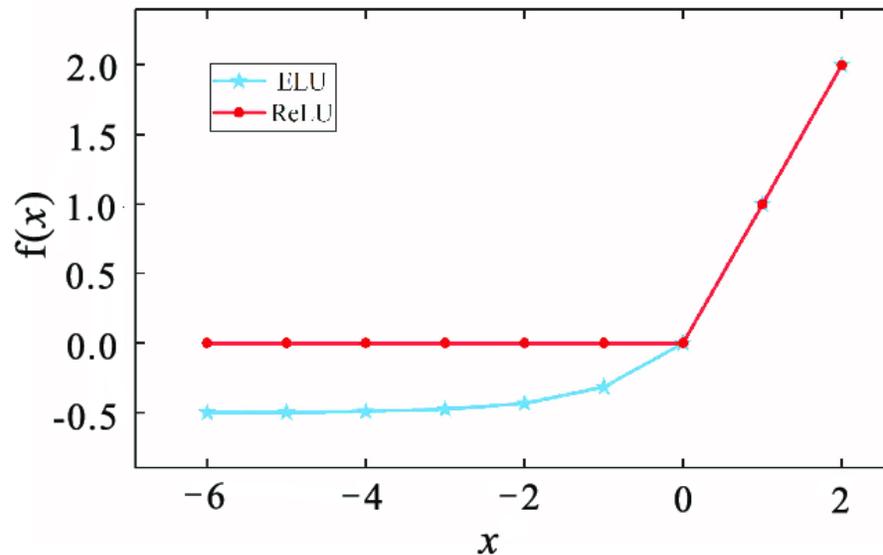


Figura 2: Función de activación ReLU y ELU. (Hashimoto et al., 2020).

Por otra parte, para poder medir el desempeño del modelo se utilizan funciones de pérdida. Estas funciones sirven para entregar una medida de qué tan bien un modelo está realizando una tarea específica durante el entrenamiento, en este caso se utiliza el error cuadrático medio (MSE). El objetivo de cualquier modelo será minimizar la función de pérdida

Como métricas, para evaluar el desempeño de la red y su evolución se utilizan error cuadrático medio (MSE) y error absoluto medio (MAE).

El error cuadrático medio (MSE) se utiliza para medir la diferencia cuadrática entre los valores reales y los valores estimados en un modelo. Está definido por:

$$MSE = \frac{1}{N} \sum (y_i - y'_i)^2$$

Donde:

- N es el número de observaciones.
- y_i es el valor real de la i -ésima observación.
- y'_i es el valor estimado de la i -ésima observación.

El error absoluto medio (MAE) por su parte se utiliza para medir la diferencia absoluta entre los valores reales y los valores estimados en un modelo. Está dado por:

$$MAE = \frac{1}{N} \sum |y_i - y'_i|$$



Donde:

- N es el número de observaciones.
- y_i es el valor real de la i -ésima observación.
- y'_i es el valor estimado de la i -ésima observación.

Por último, para entrenar de manera efectiva cualquier modelo, es fundamental el uso de un optimizador. Un optimizador es un algoritmo utilizado durante el entrenamiento del modelo para ajustar los parámetros del modelo con el objetivo de minimizar la función de pérdida ya que permite ajustar los pesos y sesgos de la red en función de los datos de entrenamiento y minimizar la función de pérdida. En este modelo, se utilizará el optimizador Adam.

Adam (*Adaptive Moment Estimation*) es uno de los optimizadores más usados debido a su eficiencia de convergencia, velocidad de convergencia y robustez. Adam utiliza diferentes tamaños de paso para actualizar los pesos según su importancia.

2.1. Capas de Dropout

Con el objetivo de mejorar el rendimiento de la red neuronal, se hizo uso de lo que se llaman capas de Dropout. Las capas de Dropout son una forma de regularización que ayuda a reducir el sobreajuste en las redes neuronales. Durante el entrenamiento, algunas de las neuronas en la capa de dropout se desactivan de forma aleatoria con una probabilidad determinada (en este caso se utilizó un 20% de probabilidad), lo que significa que no contribuyen a la propagación hacia adelante de la señal de entrada. Esto tiene el efecto de reducir la capacidad de la red y de obligarla a aprender características más robustas y útiles (Srivastava et al., 2014)

Estas capas de dropout fueron implementadas en la capa de entrada de la red, en dos capas ocultas y en la capa de salida.

2.2 Generadores.

Con la finalidad de generar más datos y de forma más rápida para así entrenar mejores modelos, se hará uso de generadores de Tensorflow.

Los generadores son una técnica de procesamiento de datos que se utiliza para alimentar datos a modelos. En lugar de cargar todos los datos en la memoria, los generadores permiten cargar datos de manera gradual a medida que el modelo los necesita durante el entrenamiento. El uso de generadores es muy útil cuando se trabaja con grandes conjuntos de datos que ocupan mucho espacio en la memoria.



d. Plan de Trabajo

Actividad	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept	Oct	Nov	Dic
Mejoramiento del performance del modelo para datos reales										
Evaluación comparativa con métodos clásicos de inversión										
Creación de datos sintéticos para acoplamiento intersísmico										
Implementación del modelo para movimientos intersísmicos										
Evaluación del modelo con datos reales de acoplamiento										
Evaluación performance a métodos clásicos de inversión										
Elaboración de manuscrito										
Preparación de la defensa										

e. Trabajo avanzado

En primer lugar, se generó la base de datos de terremotos sintéticos. Los datos generados quedan almacenados en un dataset, el cual está conformado por los desplazamientos cosísmicos y los deslizamientos del terremoto sintético.

La distribución de estaciones sintéticas utilizada es la siguiente:

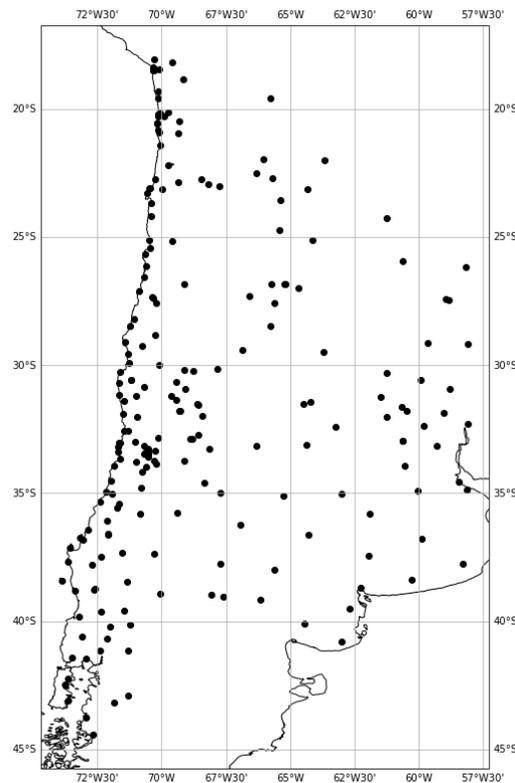


Figura 3: Distribución de estaciones sintéticas

En cuanto a los datos generados sintéticamente, algunas características de estos son las siguientes:

- Los desplazamientos generados tienen un mínimo de -26 cm y un máximo de 18 cm
- Los deslizamientos tienen un mínimo de 20 centímetros y un máximo de 60 centímetros
- Las elipses generadas tienen un rango de 40 a 150 metros de tamaño

En general, se están generando datasets con alrededor de 200.000 terremotos sintéticos para entrenar el modelo. Cabe mencionar que, para los primeros casos a estudiar por motivos de simpleza, se considera el valor absoluto de las velocidades y de los deslizamientos.

Tras tener los datos sintéticos generados, se procedió a utilizar un modelo de redes neuronales de la librería *Keras* ejecutada sobre *Tensorflow* en Python.



La arquitectura de red neuronal utilizada para los cuatro primeros casos es la siguiente:

- 1° capa de entrada: 234 neuronas (651 en el primer caso)

Capa de Dropout

- 2° capa oculta: 500 neuronas
- 3° capa oculta: 100 neuronas
- 4° capa oculta: 500 neuronas

Capa de Dropout

- 5° capa oculta: 800 neuronas
- 6° capa oculta: 1000 neuronas

Capa de Dropout

- 7° capa oculta: 1200 neuronas
- 8° capa de salida: 1667 neuronas

Capa de Dropout

Todas estas capas tienen función de activación ReLU.

Como primera prueba, se entrenó el modelo con 150.000 eventos puramente sintéticos. Es decir, con la distribución sintética de estaciones. Tras graficar las predicciones entregadas por el modelo, se obtuvo lo siguiente:

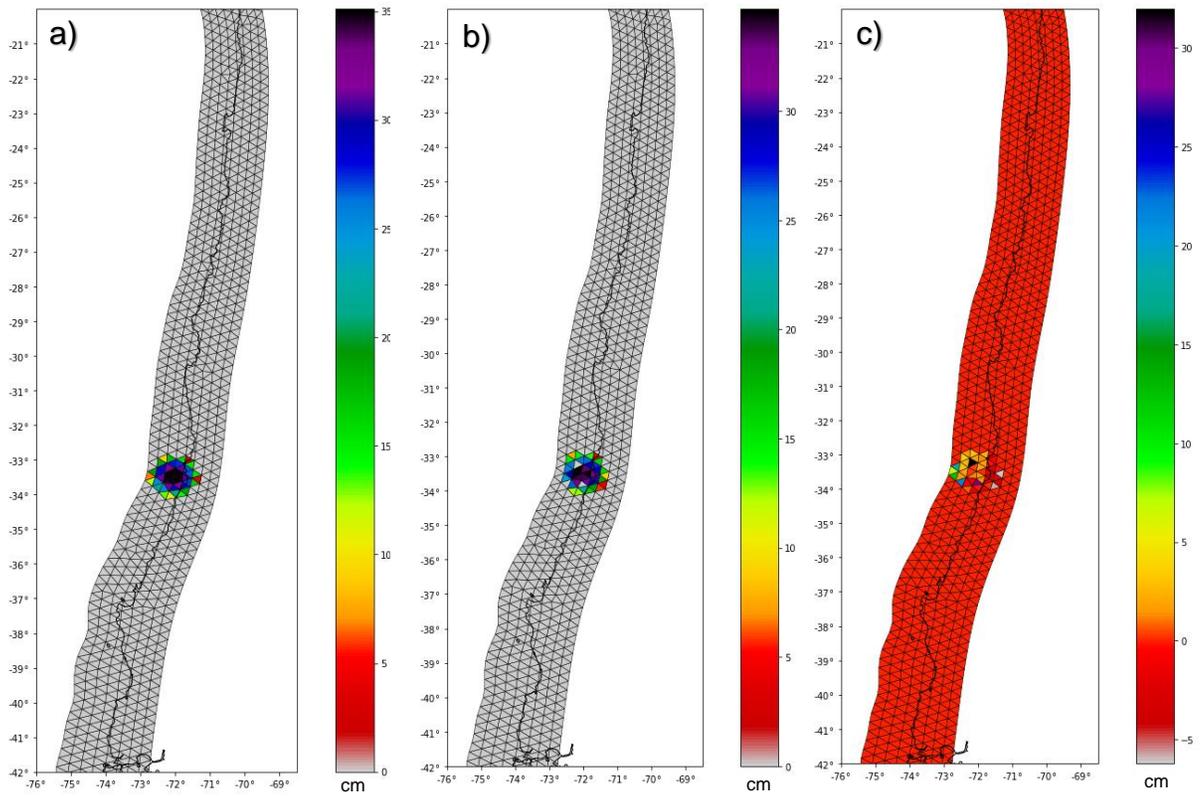


Figura 4: a) Deslizamiento de un terremoto generado sintéticamente b) Estimación del deslizamiento del mismo terremoto usando el modelo entrenado c) Residuales de la estimación del modelo

Ahora, acercando la zona del terremoto sintético:

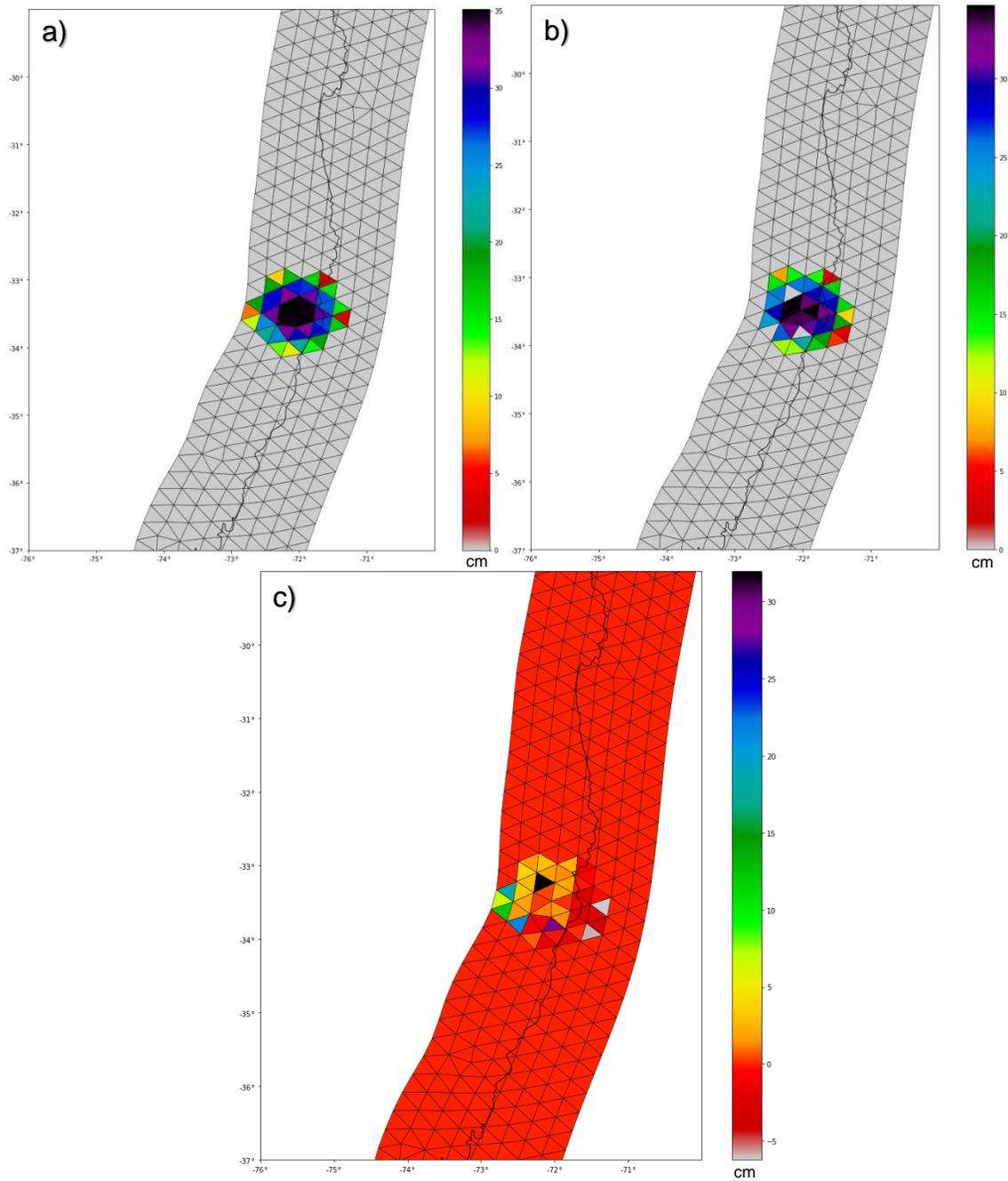


Figura 5: a) Deslizamiento de un terremoto generado sinteticamente b) Estimación del deslizamiento del mismo terremoto usando el modelo entrenado c) Residuales de la estimación del modelo

Por otro lado, los vectores de desplazamiento del deslizamiento estimado es el siguiente:

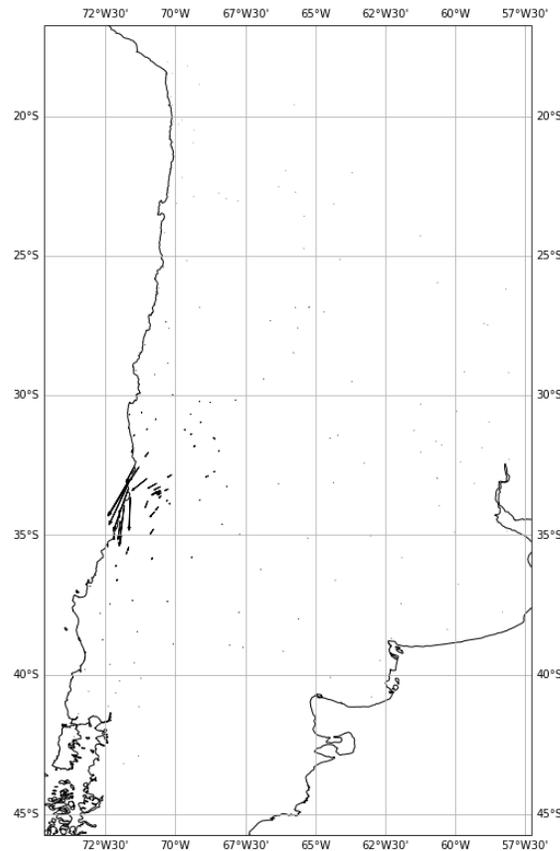


Figura 6: Vectores de desplazamiento para el terremoto sintético estimados por el modelo

De la figura 4 y 5, se aprecia que el modelo estima de manera correcta la ubicación del terremoto. La distribución del deslizamiento y la magnitud del deslizamiento son similares, pero no idénticas. De la figura 4 y 5 c), se puede ver nuevamente en los residuales que existen algunas diferencias en la distribución del deslizamiento.

En cuanto al rendimiento de la red neuronal, tenemos que el MSE y el MAE evolucionaron de la siguiente manera a través del tiempo:

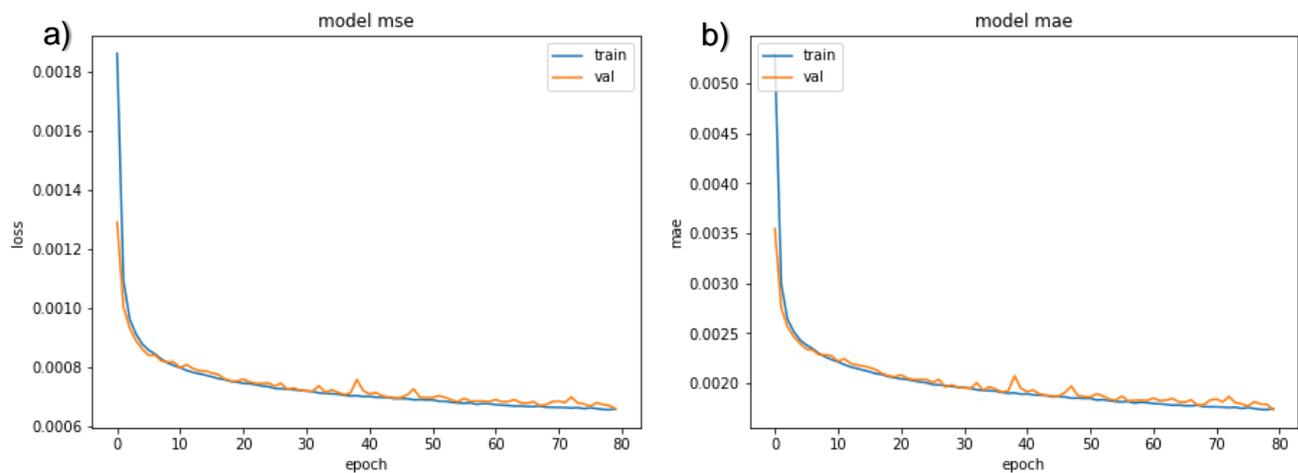


Figura 7: Evolución de las métricas MSE y MAE para el modelo. En azul se indica la evolución para el conjunto de entrenamiento y en naranja para el conjunto de validación a) Evolución para MSE b) Evolución para MAE

El MSE alcanzó un valor mínimo de 0.0006 aproximadamente a lo largo de 80 épocas y el MAE un mínimo de 0.0020. Se puede ver que a lo largo de las épocas ambas métricas van disminuyendo

Tras entrenar el modelo con datos sintéticos, se continuó a evaluar el modelo con datos reales. En específico, se utilizaron desplazamientos cosísmicos del terremoto de Illapel de 2015 de Mw 8.4.

Para trabajar con los datos reales, primero, se cambió la distribución de estaciones sintéticas por las estaciones en las que se registraron los desplazamientos del terremoto. La distribución de estaciones es la siguiente:

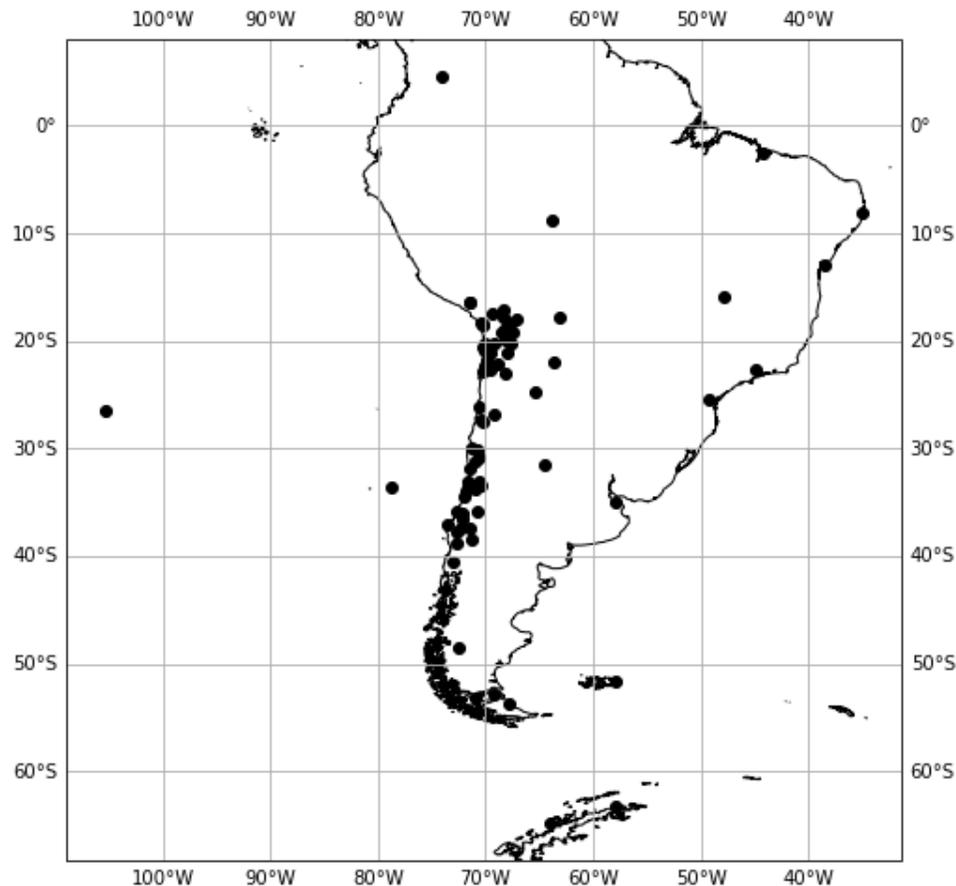


Figura 8: Distribución de estaciones

Tras cambiar las estaciones, se entrenó nuevamente el modelo con redes neuronales, nuevamente con 150.000 terremotos sintéticos. Tras graficar los resultados, tenemos lo siguiente:

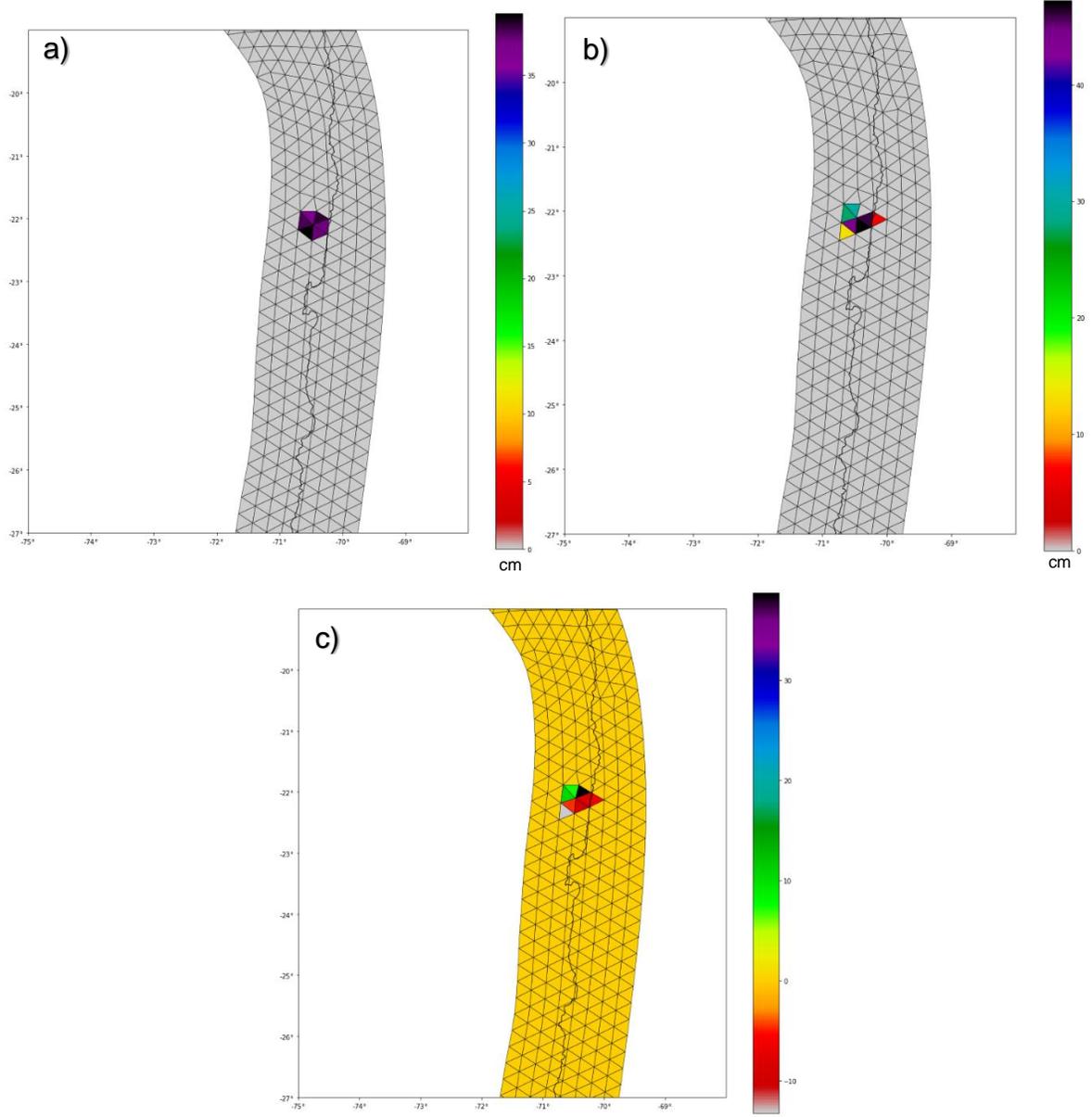


Figura 9: a) Deslizamiento de un terremoto generado sinteticamente b) Estimación del deslizamiento del mismo terremoto usando el modelo entrenado c) Residuales de la estimación del modelo

Los vectores de desplazamiento asociados a este terremoto sintético:

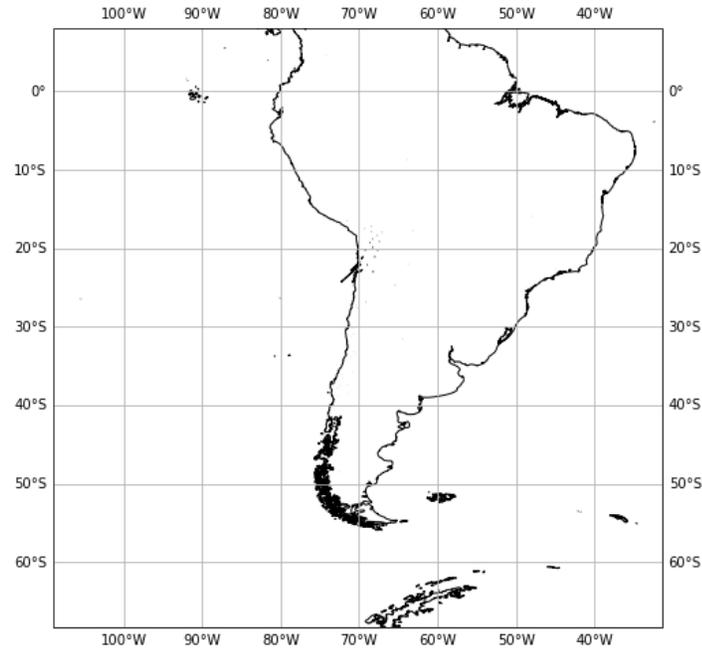


Figura 10: Vectores de desplazamiento para el terremoto sintético estimados por el modelo

De la figura 9, podemos notar que al igual que en el modelo anterior, la ubicación es estimada de manera correcta. La forma del deslizamiento y la magnitud del deslizamiento nuevamente son similares, pero no idénticas al terremoto sintético.

En cuanto al rendimiento del modelo tenemos la siguiente evolución de las métricas:

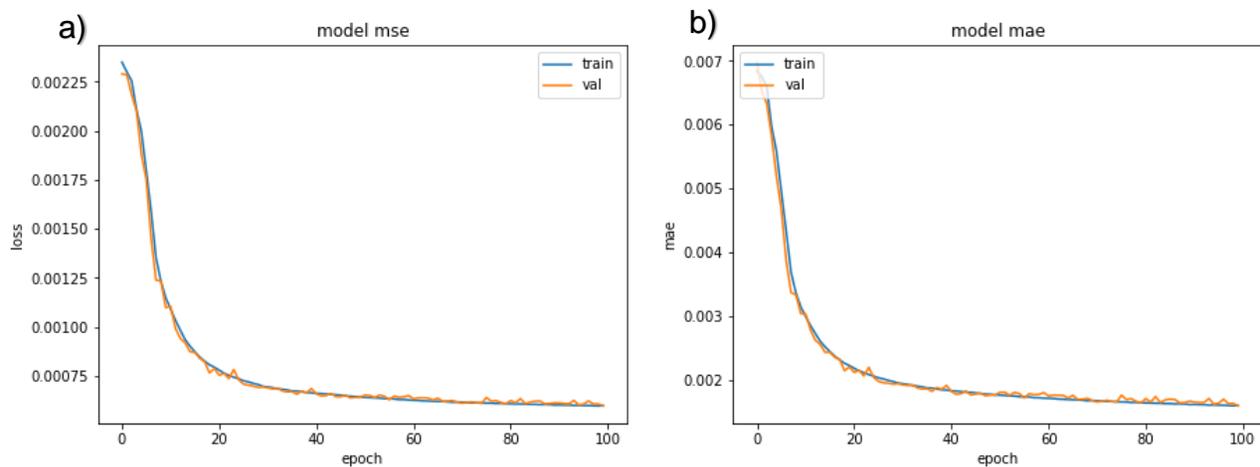


Figura 11: Evolución de las métricas MSE y MAE para el modelo. En azul se indica la evolución para el conjunto de entrenamiento y en naranja para el conjunto de validación a) Evolución para MSE b) Evolución para MAE

El MSE para este modelo alcanza un valor mínimo de 0.00075 aproximadamente a lo largo de 100 épocas y el MAE un mínimo de 0.002, siendo valores muy similares al del primer modelo entrenado.

Luego de tener el modelo entrenado, se quiso utilizar como conjunto de prueba los desplazamientos cosísmicos del terremoto de Illapel de 2015 para probar el modelo con datos reales. Únicamente con

los desplazamientos, el modelo estimará un deslizamiento. Se obtiene el siguiente deslizamiento cosísmico:

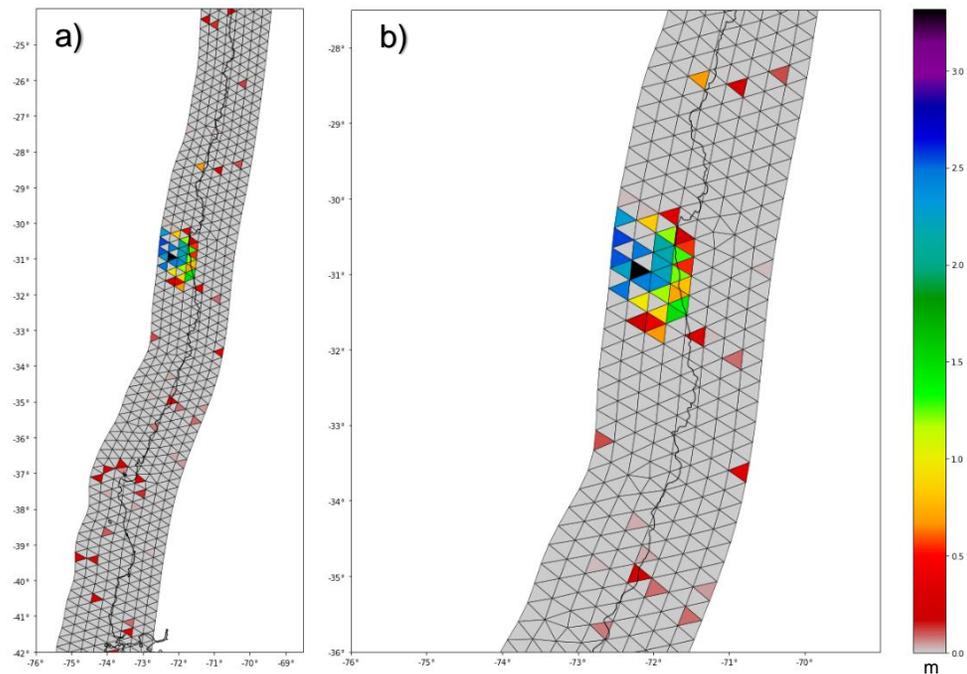


Figura 12: a) Deslizamiento cosísmico estimado a partir del modelo entrenado b) Acercamiento a la zona del terremoto

Ahora, si comparamos este resultado con un deslizamiento modelado (con una distribución diferente de estaciones) a partir de una inversión con análisis de componentes principales (Shrivstava et al., 2016):

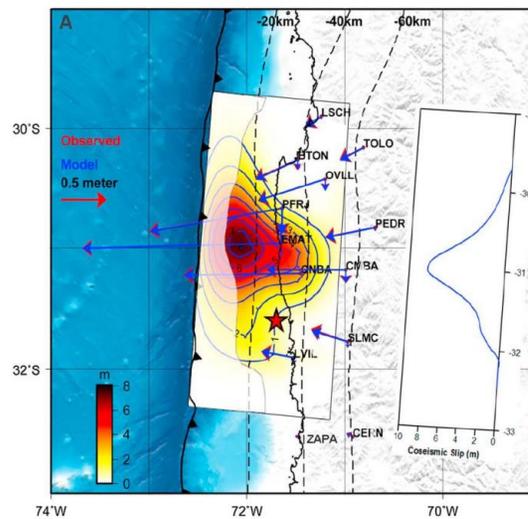


Figura 13: Modelamiento del deslizamiento de terremoto de Illapel de 2015 (Shrivastava et al., 2016)

De la figura 12 y 13, se puede notar que el modelo estima de manera correcta la ubicación del terremoto, en cuanto a la escala la estimada tiene un máximo de 3 metros, en cambio el modelado tiene un máximo de 8 metros. En cuanto a la forma del deslizamiento, en términos generales se asemeja sin embargo, existen espacios vacíos en la zona del terremoto estimado, además de pequeños deslizamientos (segmentos en rojo) fuera de la zona del terremoto.

Como primera medida para mejorar la estimación del deslizamiento de este terremoto, se agregó ruido a los datos generados sintéticamente. Se agregó un ruido base de media 0 y una distribución normal, este ruido se multiplica por la desviación estándar de los datos.

Tras agregar el ruido y entrenar con 300.000 terremotos sintéticos con ruido, se obtiene el siguiente deslizamiento:

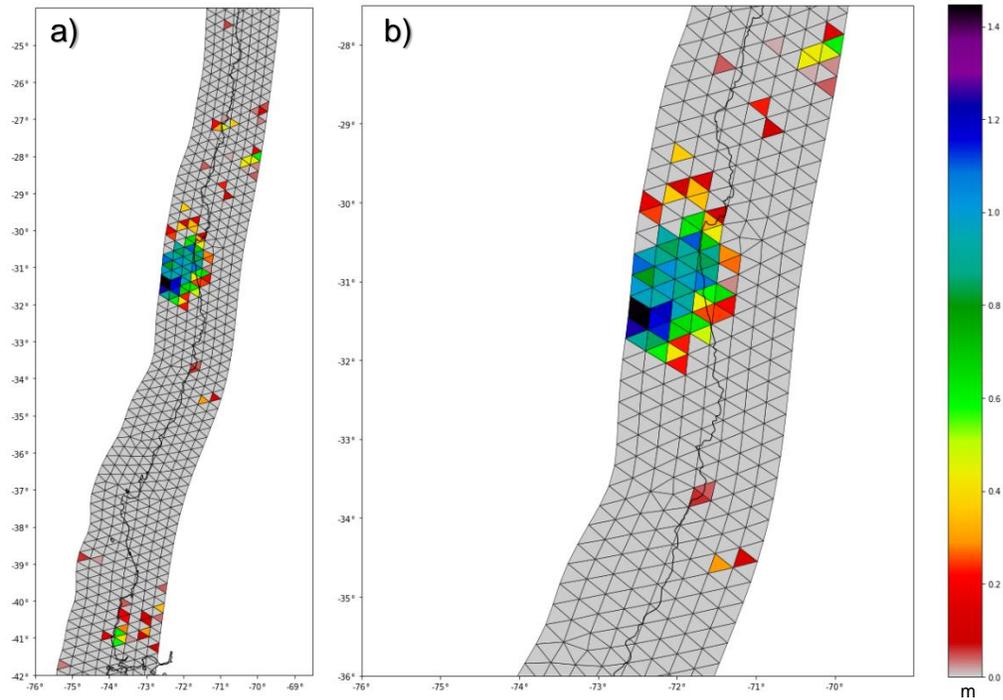


Figura 14: a) Deslizamiento cosísmico del terremoto de Illapel estimado por el modelo b) Acercamiento a la zona del terremoto

De la figura 14, es posible notar que al entrenar el modelo nuevamente, pero agregando ruido a los datos, estima una forma de slip más uniforme, sin espacios de deslizamiento vacíos. También se puede notar una disminución en los deslizamientos fuera de la zona del terremoto. La escala sigue siendo diferente y la ubicación sigue siendo correcta.

La evolución de las métricas para este modelo:

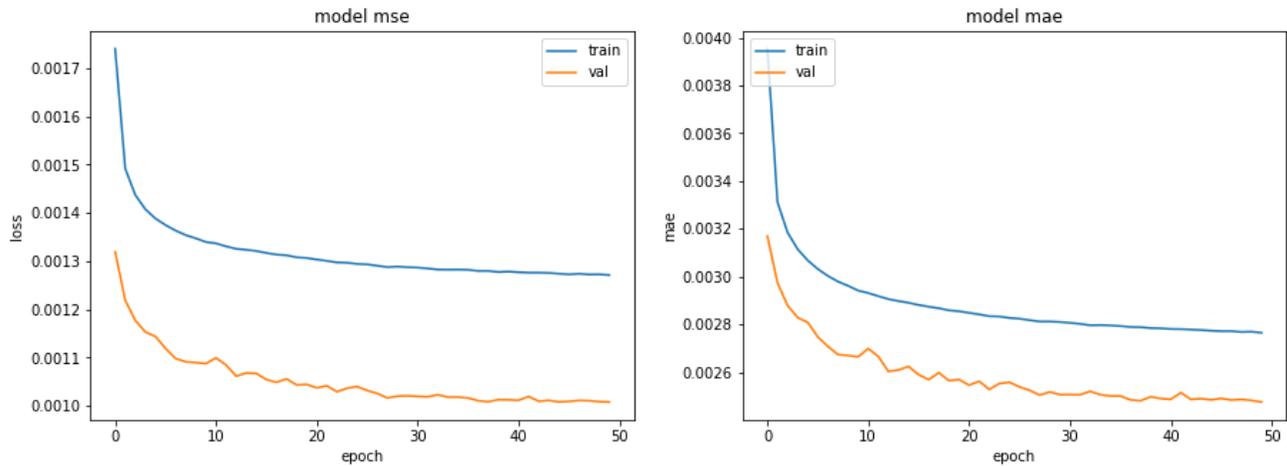


Figura 15: Evolución de las métricas MSE y MAE para el modelo. En azul se indica la evolución para el conjunto de entrenamiento y en naranja para el conjunto de validación a) Evolución para MSE b) Evolución para MAE

Se alcanzó un valor mínimo de MSE para el conjunto de entrenamiento de 0.0013 y de 0.001 para el conjunto de validación. Para el MAE se alcanzaron valores mínimos de 0.0028 y 0.0026 para MSE y MAE respectivamente. Se puede ver que las métricas del modelo disminuyeron en comparación a los primeros modelos entrenados, esto puede deberse a que el modelo se entrenó con ruido.

Con la finalidad de seguir mejorando el modelo, ahora se consideran tanto los valores positivos como negativos, además de utilizar un conjunto de entrenamiento con más datos (450.000 terremotos sintéticos) gracias al uso de generadores de Tensorflow.

El modelo utilizado en este caso es:

- 1° capa de entrada: 234 neuronas
- 2°, 3°, 4° y 5° capa oculta: 100 neuronas
- 6° capa de salida: 1667 neuronas

Capa de Dropout

Todas estas capas tienen función de activación ELU para tomar en cuenta tanto los valores positivos como los negativos

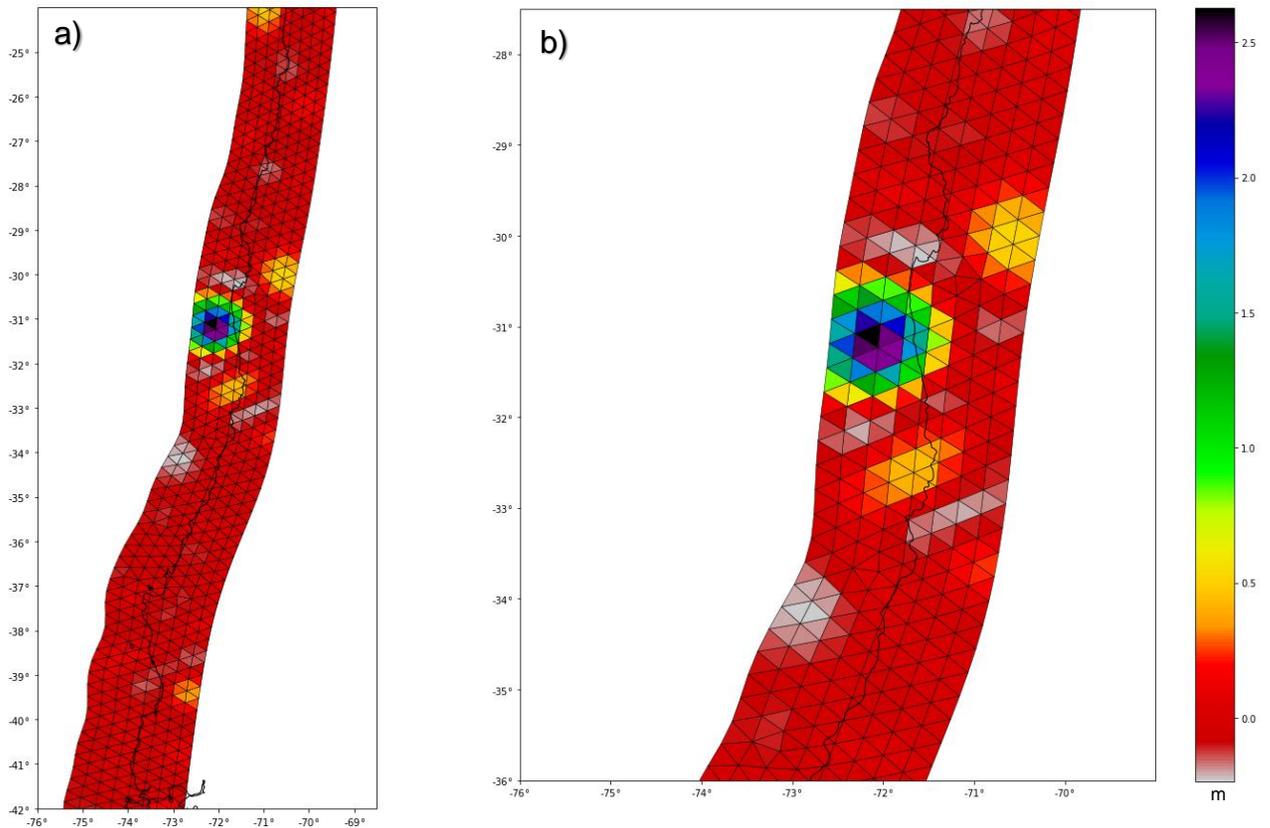


Figura 16: a) Deslizamiento cosísmico del terremoto de Illapel estimado por el modelo b) Acercamiento a la zona del terremoto

De la figura 16, se puede ver que el modelo estima de manera correcta la ubicación y la forma del deslizamiento del terremoto. Sin embargo, en zonas fuera del área del terremoto, estima deslizamientos negativos o pequeños deslizamientos donde no deberían presentarse.

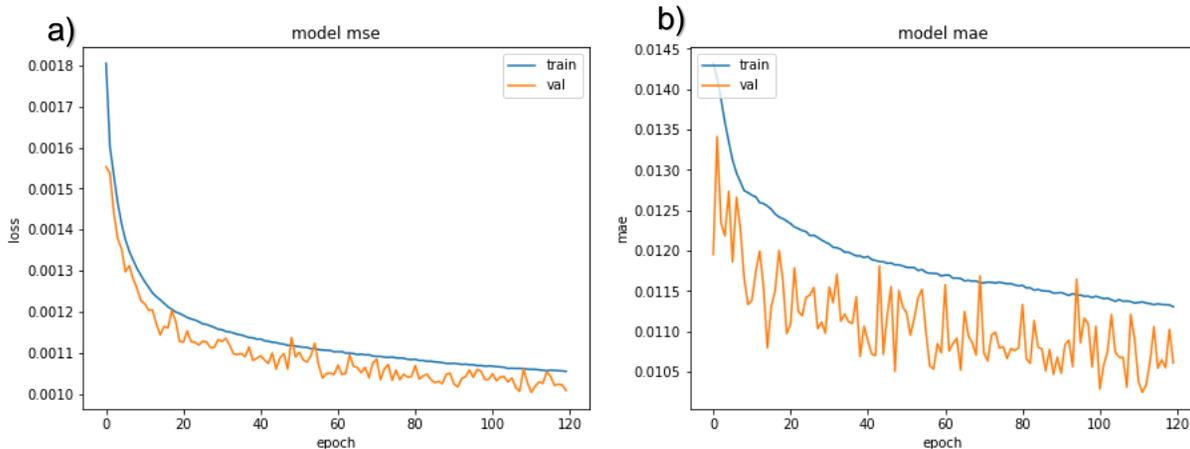


Figura 17: Evolución de las métricas MSE y MAE para el modelo. En azul se indica la evolución para el conjunto de entrenamiento y en naranja para el conjunto de validación a) Evolución para MSE b) Evolución para MAE

El MSE alcanzó un valor mínimo de 0.001 aproximadamente, por su lado el MAE alcanzó un valor mínimo de 0.0105, este desempeño es similar al modelo anterior

En general, el modelo estima de manera correcta la ubicación y la distribución del deslizamiento, sin embargo, presenta problemas en todos los casos con la escala del terremoto de Illapel. Esto puede deberse que al generar los datos no se generaron los terremotos sintéticos con un deslizamiento tan



alto como lo es el caso real presentado. Esto puede solucionarse generando nuevos datos con terremotos con mayores deslizamientos.

}



f. REFERENCIAS BIBLIOGRÁFICAS:

- Altamimi, Z., P. Rebischung, L. Métivier, and X. Collilieux (2016) ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions, *J. Geophys. Res. Solid Earth*, 121, 6109–6131, doi:10.1002/2016JB013098
- Baez, J. C., Ruiz, S., Leyton, F., and Del-Campo, F. (2018) Slow slip event preceding the Valparaíso Mw 6.9 2017 earthquake in Central Chile, captured by GPS observations” p. 11312
- Béjar-Pizarro, M., Socquet, A., Armijo, R. (2013) Andean structural control on interseismic coupling in the North Chile subduction zone. *Nature Geosci* 6, 462–467. <https://doi.org/10.1038/ngeo1802>
- Chlieh, M., Perfettini, H., Tavera, H., Avouac, J.-P., Remy, D., Nocquet, J.-M., Rolandone, F., Bondoux, F., Gabalda, G., and Bonvalot, S. (2011), Interseismic coupling and seismic potential along the Central Andes subduction zone, *J. Geophys. Res.*, 116, B12405, doi:10.1029/2010JB008166.
- Corbi, F., Sandri, L., Bedford, J., Funicello, F., Brizzi, S., Rosenau, M., & Lallemand, S. (2019). Machine learning can predict the timing and size of analog earthquakes. *Geophysical Research Letters*, 46, 1303–1311. <https://doi.org/10.1029/2018GL081251>
- DeVries, P., Viegas, F., Wattenberg, M., & Meade, B. (2018) Deep learning of aftershock patterns following large earthquakes. *Nature*, 560.
- Duputel, Z., et al. (2015), The Iquique earthquake sequence of April 2014: Bayesian modeling accounting for prediction uncertainty, *Geophys. Res. Lett.*, 42, 7949–7957, doi:10.1002/2015GL065402
- Eos Transactions (2011) "Coupling" Semantics and Science in Earthquake Research
- Johnson P. A., Rouet-Leduc B. (2019) Laboratory earthquake forecasting: A machine learning competitions.
- Kaneko, Y., Avouac, J.P., Lapusta, N. (2010) Towards inferring earthquake patterns from geodetic observations of interseismic coupling. 3. 363-U24. 10.1038/NGEO84
- Hashimoto, F., Ote, K., Oida, T., Teramoto, A., Ouchi, Y. (2020). Compressed-Sensing Magnetic Resonance Image Reconstruction Using an Iterative Convolutional Neural Network Approach. *Applied Sciences*. 10. 1902. 10.3390/app10061902.
- Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., et al. (2018). Slab 2, a Comprehensive Subduction Zone Geometry Model. *Science* 362 (6410), 58–61. doi:10.1126/4723
- Lapedes, A., Farber R., (1988) How Neural Net Works
- Loveless, J. P. & Meade, B. J. (2011) Spatial correlation of interseismic coupling and coseismic rupture extent of the 2011 Mw = 9.0 Tohoku-oki earthquake. *Geophys. Res. Lett.* 38, L17306
- Metóis, M., Socque, A., Vigny, C. (2012) Interseismic coupling, segmentation and mechanical behavior of the Central Chile subduction zone
- Métois, M., Vigny, C., Socquet, A. (2016) Interseismic Coupling, Megathrust Earthquakes and Seismic Swarms Along the Chilean Subduction Zone
- Métois, M., Vigny, C., Socquet, A. (2013) Revisiting the North Chile seismic gap segmentation using GPS-derived interseismic coupling



- Moreno, M., Rosenau, M. & Oncken, O. (2010) Maule earthquake slip correlates with pre-seismic locking of Andean subduction zone. *Nature* 467, 198–202
- Nikkhoo, M., and Walter, T. R. (2015) Triangular Dislocation: an Analytical, Artefact-free Solution. *Geophys. J. Int.* 201 (2), 1119–1141. doi:10.1093/gji/ggv035
- Ozawa, S., T. Nishimura, H. Munekane, H. Suito, T. Kobayashi, M. Tobita, & T. Imakiire. (2012) Preceding, coseismic, and postseismic slips of the 2011 Tohoku earthquake, Japan, *J. Geophys. Res.*, 117, B07404, doi:10.1029/2011JB009120..
- Pollitz, F. F., Bürgmann, R., and Banerjee, P. (2011) Geodetic slip model of the 2011 M9.0 Tohoku earthquake, *Geophys. Res. Lett.*, 38, L00G08, doi:10.1029/2011GL048632.
- Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A., (2017) Machine learning predicts laboratory earthquakes.
- Rouet-Leduc, B., Hulbert, C., Bolton, D. C., Ren, C. X., Riviere, J., Marone, C., Guyer, R. A. (2018) Estimating fault friction from seismic signals in the laboratory.
- Rouet-Leduc, B., Hulbert, C & Johnson, P. A., (2018) Breaking Cascadia's silence: Machine learning reveals the constant chatter of the megathrust.
- Ruiz, S., Metois, M., Fuenzalida, A., Ruiz, J., Leyton, F., Grandin, R., Campos, J. (2014) Intense foreshocks and a slow slip event preceded the 2014 Iquique Mw 8.1 earthquake. *Science*, 345(6201), 1165–1169. doi:10.1126/science.1256074
- Schmidt, J., Marques, M.R.G., Botti, S. (2019) Recent advances and applications of machine learning in solid-state materials science. *npj Comput Mater* 5, 83, doi: 10.1038/s41524-019-0221-0
- Sharma S., Sharma S., Athaiy, A. (2020) International Journal of Engineering Applied Sciences and Technology. 2020 Vol. 4, Issue 12, ISSN No. 2455-2143, Pages 310-316
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research.* 15. 1929-1958.
- Shrivastava, M. N., González, G., Moreno, M., Chlieh, M., Salazar, P., Reddy, C. D., Báez, J. C., Yáñez, G., González, J., and de la Llera, J. C. (2016) Coseismic slip and afterslip of the 2015 Mw 8.3 Illapel (Chile) earthquake determined from continuous GPS data, *Geophys. Res. Lett.*, 43, 10,710– 10,719, doi: 10.1002/2016GL070684.
- Stevens, V. L., and J. P. Avouac (2015) Interseismic coupling on the main Himalayan thrust, *Geophys. Res. Lett.*, 42, 5828–5837, doi:10.1002/2015GL064845.
- Wang, Sun-Chong. (2003) *Interdisciplinary Computing in Java Programming*
- Yáñez, V (2020) Deformación intersísmica y acoplamiento interplaca a partir de datos GNSS en la brecha sísmica de la Región de Atacama

