



**Universidad del Desarrollo**  
Facultad de Ingeniería

PRONÓSTICO HÍBRIDO ESPACIO-TEMPORAL DE PRECIPITACIONES EN CHILE:  
INTEGRANDO REDES NEURONALES Y MÉTODOS ESTADÍSTICOS AVANZADOS

POR: CARLOS SAQUEL DEPAOLI

Proyecto de grado presentado a la Facultad de Ingeniería de la Universidad del  
Desarrollo para optar al grado académico de Magíster en Data Science

PROFESOR GUÍA:

Dr. MAURICIO HERRERA MARÍN

Enero 2026

Santiago

## AGRADECIMIENTO

A mi familia, por su apoyo inquebrantable en los momentos más difíciles de este proceso. Por la paciencia, la comprensión y la fuerza que me entregaron cuando las cosas se hicieron cuesta arriba. Sin su respaldo, este trabajo simplemente no habría sido posible.

# TABLA DE CONTENIDO

<b>RESUMEN .....</b>	<b>1</b>
<b>1. INTRODUCCIÓN.....</b>	<b>2</b>
<b>2. TRABAJO RELACIONADO .....</b>	<b>5</b>
2.1. PREDICCIÓN NUMÉRICA DEL CLIMA (NWP) .....	5
2.2. MODELOS GLOBALES BASADOS EN INTELIGENCIA ARTIFICIAL .....	6
2.3. MODELOS HÍBRIDOS E INTERPRETABILIDAD DINÁMICA.....	7
<b>3. HIPÓTESIS Y OBJETIVOS .....</b>	<b>10</b>
3.1. OBJETIVO GENERAL .....	11
3.2. OBJETIVOS ESPECÍFICOS.....	11
<b>4. DATOS Y METODOLOGÍA.....</b>	<b>12</b>
4.1. DATOS .....	12
4.1.1. FUENTE Y DOMINIO DE ESTUDIO.....	12
4.1.2. PREPROCESAMIENTO.....	14
4.1.3. DATOS PARA EL BASELINE DEEPAR.....	16
4.2. METODOLOGÍA.....	17
4.2.1. MODELO HÍBRIDO AUTOENCODER 3D + DMD.....	17
4.2.2. MODELO BASELINE .....	21
4.2.3. ESQUEMA DE EVALUACIÓN .....	23
<b>5. RESULTADOS.....</b>	<b>25</b>
5.1. MÉTRICAS DE RECONSTRUCCIÓN.....	25
5.2. DESEMPEÑO DE PRONÓSTICO DEL AE+DMD .....	31
5.2.1. MÉTRICAS AGREGADAS POR HORIZONTE .....	31
5.2.2. EVOLUCIÓN TEMPORAL DE PRECIPITACIÓN .....	32

5.2.3.	COMPARACIÓN CAMPO MEDIO OBSERVADO VS FORECAST .....	38
5.3.	COMPARACIÓN AE+DMD VS DEEPAR.....	45
5.3.1.	COMPARACIÓN TEMPORAL DEL PROMEDIO ESPACIAL Y SU ERROR (AE+DMD, DEEPAR Y PERSISTENCIA) .....	47
5.3.2.	COMPARACIÓN ESPACIAL DEL PROMEDIO TEMPORAL Y SU ERROR (AE+DMD Y DEEPAR) .....	56
5.3.3.	EVALUACIÓN POR INTENSIDAD: DETECCIÓN DE EVENTOS, ERROR EN CONDICIONES HÚMEDAS Y HABILIDAD ESPACIAL	60
5.3.4.	CASOS DE ESTUDIO DE EVENTOS DE PRECIPITACIÓN .....	67
5.3.5.	COSTO COMPUTACIONAL (ENTRENAMIENTO E INFERENCIA).....	74
5.4.	ANÁLISIS ESPECTRAL Y MODOS DMD .....	75
<b>6.</b>	<b>CONCLUSIONES .....</b>	<b>87</b>
6.1.	CUMPLIMIENTO DE OBJETIVOS .....	87
6.2.	PRINCIPALES HALLAZGOS.....	89
6.3.	LIMITACIONES.....	90
6.4.	TRABAJO FUTURO .....	92
<b>7.</b>	<b>ANEXOS .....</b>	<b>93</b>
7.1.	ANEXO A. CONFIGURACIÓN EXPERIMENTAL E HIPERPARÁMETROS .....	93
<b>8.</b>	<b>REFERENCIAS .....</b>	<b>99</b>

## **RESUMEN**

Este trabajo desarrolla un sistema de pronóstico espacio-temporal de precipitaciones en Chile basado en un modelo híbrido que combina un Autoencoder convolucional 3D (AE) con Descomposición Modal Dinámica (DMD). El desarrollo se centra en las regiones del Maule y Ñuble, utilizando precipitación diaria CR2MET (1980–2021) en alta resolución para capturar la variabilidad regional relevante. El AE reduce la dimensionalidad de cubos espacio-temporal y genera un espacio latente compacto, sobre el cual la DMD ajusta un operador lineal encargado de propagar la dinámica y producir pronósticos a horizontes de 1 a 5 días.

El desempeño del modelo AE+DMD se evalúa para horizontes de 1 a 5 días y se compara con un baseline de referencia basado en DeepAR aplicado a múltiples series univariadas, mediante métricas como MAE y RMSE. Los resultados muestran que el enfoque híbrido captura adecuadamente los patrones espacio-temporal dominantes y ofrece pronósticos competitivos, en particular para 1–3 días, con un costo computacional moderado y modos DMD que facilitan la interpretación física de la dinámica de precipitación regional.

# 1. INTRODUCCIÓN

La predicción de precipitaciones es crítica para la gestión de recursos hídricos, la agricultura y la planificación territorial, especialmente en un contexto de sequía y mayor variabilidad climática. En Chile, el fuerte gradiente norte–sur y costa–cordillera, sumado a la topografía andina, genera una distribución de lluvia altamente heterogénea, lo que dificulta obtener pronósticos locales precisos a partir de modelos globales de predicción numérica o climática. Estos modelos representan de manera explícita la física atmosférica, pero requieren alto costo computacional y no siempre entregan resolución suficiente ni buen desempeño en dominios complejos como la zona centro–sur.

En paralelo, los modelos de pronóstico basados en inteligencia artificial han mostrado que es posible aproximar la dinámica atmosférica directamente desde datos de reanálisis, reduciendo drásticamente el costo computacional. Sin embargo, su adopción directa a escala regional no es trivial y requiere infraestructura y capacidades de ingeniería que muchas instituciones no poseen. Por otro lado, enfoques puramente estadísticos o basados en series temporales agregadas pierden información espacial clave y no representan explícitamente la dinámica espacio–temporal de los campos de precipitación, lo que limita su utilidad en regiones con alta heterogeneidad espacial.

En este contexto, la Descomposición Modal Dinámica (Dynamic Mode Decomposition, DMD) surge como una alternativa intermedia. El DMD descompone datos espacio–

temporales en modos asociados a autovalores complejos, permitiendo aproximar la evolución del sistema mediante un operador lineal ajustado directamente desde observaciones. Al combinar DMD con técnicas de reducción de dimensionalidad no lineal, como Autoencoders convolucionales, es posible construir modelos híbridos que capturen patrones espacio-temporales relevantes, manteniendo a la vez interpretabilidad y bajo costo computacional.

Este proyecto se inserta en esa línea y desarrolla un modelo híbrido Autoencoder 3D + DMD (AE+DMD) para el pronóstico espacio-temporal de precipitación diaria en la zona centro-sur de Chile, con foco en las regiones del Maule y Ñuble. Se utiliza el conjunto de datos CR2MET de precipitación diaria sobre una grilla en alta resolución de  $0,05^\circ$  de latitud y longitud para el período 1980–2021, recortado al dominio  $34^\circ\text{S}$ – $38^\circ\text{S}$  y  $73^\circ\text{O}$ – $70^\circ\text{O}$ . El Autoencoder convolucional tridimensional se encarga de comprimir cubos espacio-temporales de precipitación en un espacio latente de baja dimensión, sobre el cual se ajusta un operador lineal mediante DMD para generar pronósticos a horizontes de 1 a 5 días. Se opta por capas convoluciones 3D debido a su eficiencia computacional: permiten capturar correlaciones espacio-temporales locales mediante operaciones altamente paralelizables en GPU y con un consumo de memoria más controlable que alternativas recurrentes (por ejemplo, LSTM/ConvLSTM). Esta elección es coherente con el objetivo del modelo, donde la dinámica multi-step se modela principalmente en el espacio latente a través de DMD.

Como línea de base se implementa un modelo DeepAR sobre series temporales univariadas, lo que permite comparar el esquema AE+DMD frente a un enfoque de series de tiempo más estándar. La evaluación se realiza mediante métricas MAE y RMSE. El objetivo general es determinar si un modelo centrado en DMD sobre un espacio latente aprendido puede lograr un equilibrio razonable entre precisión, interpretabilidad y eficiencia computacional para el pronóstico de precipitaciones en Maule-Ñuble, como primera etapa de un sistema de pronóstico híbrido aplicable a otros dominios de Chile.

## **2. Trabajo Relacionado**

El problema del pronóstico de precipitaciones ha sido abordado históricamente desde tres grandes líneas:

- i. Modelos numéricos de predicción del tiempo (NWP) y de clima.
- ii. Modelos estadísticos y de series de tiempo.
- iii. Enfoques data-driven basados en aprendizaje automático y profundo.

En los últimos años han aparecido además esquemas híbridos, donde se combinan técnicas de reducción de dimensionalidad y descomposición modal con modelos de aprendizaje profundo, línea en la que se inserta este trabajo.

### **2.1. Predicción numérica del clima (NWP)**

Durante gran parte del siglo XX, la predicción meteorológica se ha sustentado en los modelos numéricos de predicción del tiempo (NWP), los cuales resuelven ecuaciones diferenciales parciales que representan la dinámica atmosférica y los procesos termodinámicos del sistema climático (Kalnay, 2003). Entre los principales exponentes se encuentran el Integrated Forecasting System (IFS) del European Centre for Medium-Range Weather Forecasts (ECMWF) y el Global Forecast System (GFS) del National Centers for Environmental Prediction (NCEP).

Estos sistemas poseen gran capacidad de generalización física, pero presentan limitaciones estructurales: alto costo computacional derivado de la integración numérica tridimensional en escalas globales; dependencia de parametrizaciones físicas que introduce incertidumbre en regiones complejas; y resolución limitada, especialmente en dominios con topografía abrupta como los Andes.

En 2025, el ECMWF implementó su primer sistema de pronóstico impulsado completamente por IA, el Artificial Intelligence Forecast System (AIFS), con desempeño comparable e incluso superior al IFS físico en varios horizontes (ECMWF, 2025). Este hito marca una transición hacia modelos puramente data-driven, que emulan la física atmosférica mediante aprendizaje estadístico.

## **2.2. Modelos globales basados en inteligencia artificial**

El cambio de paradigma comenzó con el desarrollo de WeatherBench (Rasp et al., 2020), un benchmark reproducible para la evaluación de modelos de predicción climática basados en datos ERA5. Este trabajo impulsó una generación de arquitecturas que igualan o superan el rendimiento de los modelos NWP tradicionales en pronósticos de corto y mediano plazo. En esta misma línea, Reichstein et al. (2019) argumentan que el deep learning en ciencias de la Tierra no solo mejora la predicción, sino que también puede aportar comprensión de procesos cuando se combina con conocimiento físico y enfoques interpretables, respaldando el uso de modelos híbridos.

Entre los avances más relevantes destacan GraphCast (Lam et al., 2023), una red de grafos desarrollada por DeepMind capaz de predecir el estado atmosférico global a 10 días con mayor precisión que el sistema determinista HRES del ECMWF; Pangu-Weather (Bi et al., 2023), una arquitectura tridimensional jerárquica (3DEST Transformer) que representa dependencias espaciales y temporales complejas; y FourCastNet y FourCastNet-3 (Pathak et al., 2022; NVIDIA, 2025), modelos basados en transformadas de Fourier y mecanismos de atención espacial, con capacidad de ensamble y eficiencia energética sin precedentes.

Estos desarrollos demostraron que la IA puede reproducir la dinámica global con costos miles de veces inferiores al NWP, posicionando los modelos de aprendizaje profundo como alternativa viable para la predicción global a mediano plazo.

### **2.3. Modelos híbridos e interpretabilidad dinámica**

Paralelamente, ha emergido una corriente de investigación centrada en combinar aprendizaje profundo y teoría de sistemas dinámicos. El punto de partida lo constituye la Descomposición Modal Dinámica (DMD) propuesta por Schmid (2010), que permite descomponer señales espacio-temporales en modos oscilatorios lineales asociados a frecuencias y tasas de crecimiento. DMD se ha convertido en una herramienta fundamental para el análisis de sistemas no lineales complejos, incluyendo fluidodinámica y clima (Tu et al., 2014). Como referencia base y exposición sistemática del método

(formulación, variantes e interpretación), se adopta el texto de Kutz et al. (2016), ampliamente utilizado como marco estándar para sustentar el uso de DMD en sistemas complejos.

Posteriormente, Williams, Kevrekidis y Rowley (2015) generalizaron este enfoque mediante la Extended DMD (EDMD), introduciendo funciones base no lineales que permiten aproximar el operador de Koopman —un operador lineal infinito-dimensional que describe la evolución temporal de observables no lineales (Koopman, 1931; Mezić, 2005). Esta formulación sentó las bases para una familia de métodos híbridos basados en el aprendizaje de representaciones latentes que capturan la dinámica esencial de sistemas no lineales. En particular, Lusch, Kutz y Brunton (2018) proponen un enfoque de Deep Koopman que aprende, mediante un autoencoder modificado, coordenadas no lineales donde la dinámica es aproximadamente lineal, conectando explícitamente aprendizaje profundo con embeddings tipo Koopman y motivando arquitecturas como AE+DMD.

El desarrollo posterior dio origen a los Koopman Autoencoders (KoAE), que combinan redes autoencoder con la estructura del operador de Koopman para separar componentes dinámicas lentas y rápidas (Berman et al., 2023). De forma paralela, se introdujeron modelos physics-informed, como el Physics-Informed DMD (piDMD) (Baddoo et al., 2023) y el enfoque probabilístico de Pan y Duraisamy (2020), que integran información física y restricciones de estabilidad dentro del marco de Koopman, asegurando predicciones coherentes y físicamente plausibles.

Estas líneas convergen en una tendencia hacia modelos híbridos parsimoniosos e interpretables, capaces de capturar la estructura dinámica subyacente sin recurrir a ecuaciones explícitas.

### 3. Hipótesis y Objetivos

La presente investigación parte de la hipótesis de que:

*“La integración de un Autoencoder 3D con Descomposición Modal Dinámica (DMD) permite capturar de manera parsimoniosa los patrones espacio-temporales de precipitación en la zona Maule-Ñuble, ofreciendo un equilibrio entre precisión, interpretabilidad y eficiencia computacional superior al de los modelos tradicionales basados en series de tiempo.”*

Esta hipótesis se fundamenta en los resultados de investigaciones previas (Schmid, 2010; Williams et al., 2015; Berman et al., 2023; Koopman, 1931; Mezić, 2005, Lusch et al., 2018) que demuestran la capacidad de los modelos híbridos basados en el operador de Koopman para representar sistemas no lineales en espacios latentes de dinámica lineal. Al aplicar este enfoque a campos meteorológicos de alta dimensión, se espera que el modelo propuesto logre representar la evolución temporal de la precipitación con menor pérdida de información y mayor estabilidad que los métodos puramente neuronales.

### **3.1. Objetivo General**

Desarrollar un modelo híbrido de pronóstico climático basado en Autoencoders tridimensionales y Descomposición Modal Dinámica (AE+DMD), capaz de predecir la evolución espacio-temporal de la precipitación diaria en la zona Maule-Ñuble utilizando datos CR2MET (1980-2021), maximizando la interpretabilidad física y minimizando el costo computacional.

### **3.2. Objetivos Específicos**

1. Implementar y entrenar un Autoencoder 3D para la compresión espacio-temporal de los campos de precipitación diaria, identificando representaciones latentes robustas.
2. Aplicar la Descomposición Modal Dinámica (DMD) sobre el espacio latente del Autoencoder para modelar la evolución temporal mediante un operador lineal.
3. Comparar el desempeño del modelo híbrido AE+DMD frente a un baseline de referencia (DeepAR) utilizando métricas de error MAE y RMSE.
4. Evaluar la interpretabilidad física de los modos DMD, relacionándolos con patrones climáticos relevantes (ENSO, estacionalidad, topografía).
5. Analizar la eficiencia y estabilidad del modelo híbrido, considerando su escalabilidad y aplicabilidad para pronósticos regionales de bajo costo.

## 4. Datos y Metodología

### 4.1. Datos

#### 4.1.1. Fuente y dominio de estudio

El conjunto de datos utilizado corresponde a CR2MET, un producto meteorológico de alta resolución desarrollado por el Centro de Ciencia del Clima y la Resiliencia (CR2), que combina observaciones locales con reanálisis globales para estimar variables climáticas en Chile continental.

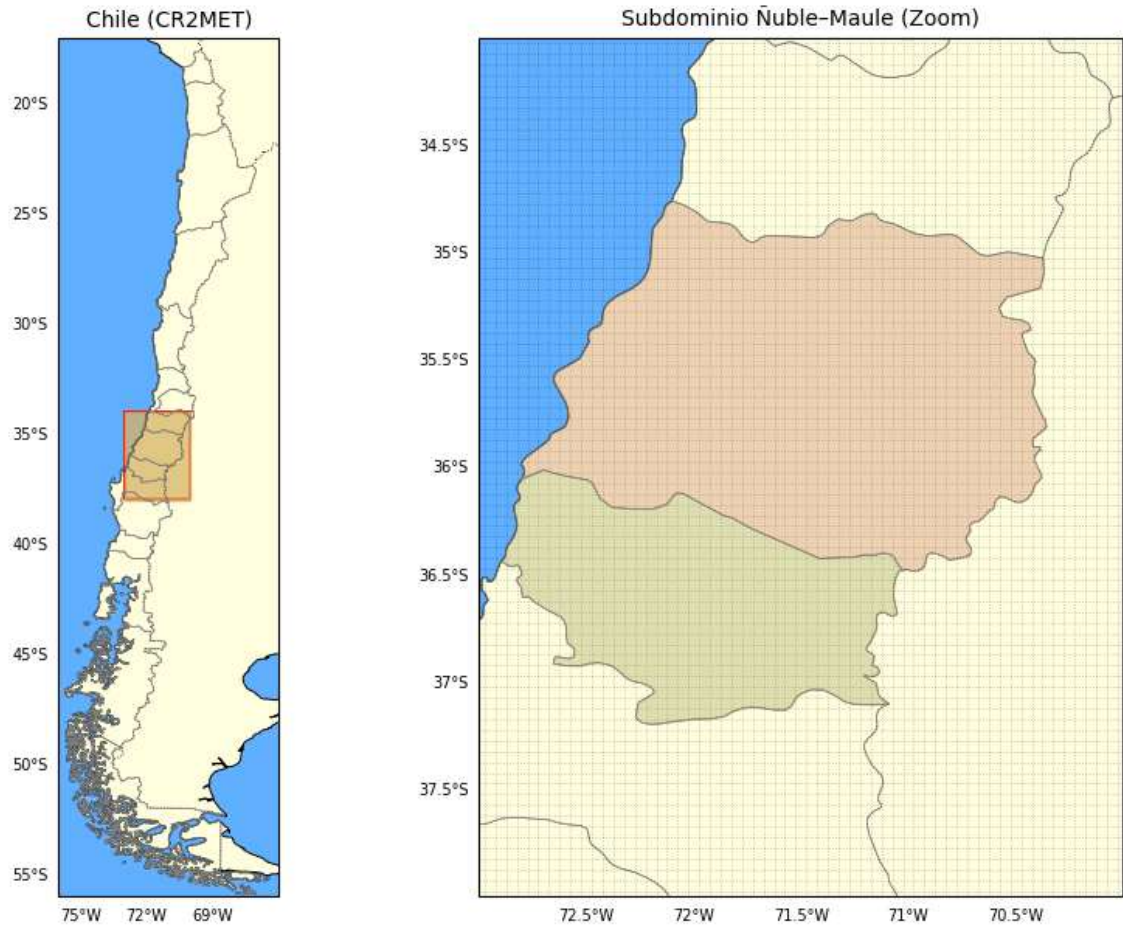
El dataset incluye campos distribuidos espacialmente de precipitación diaria, temperatura máxima y temperatura mínima, cubriendo el período 1980–2021 sobre una cuadrícula regular de resolución espacial  $0.05^\circ \times 0.05^\circ$ . Para el presente estudio, se seleccionó el subdominio geográfico comprendido entre:

- **Latitud:**  $-38^\circ$  a  $-34^\circ$
- **Longitud:**  $-73^\circ$  a  $-70^\circ$

correspondiente a las regiones del Maule y Ñuble (Figura 4.1), resultando en una grilla de  $H \times W = 80 \times 60$ . Este tamaño espacial determina directamente la carga computacional del modelo Conv3D, dado que cada ventana temporal de entrada se representa como un tensor  $X \in \mathbb{R}^{T_{IN} \times H \times W \times 1}$ . En este subdominio, la dinámica de la precipitación presenta alta variabilidad espacial y relevancia para el sector agroindustrial.

## Figura 4.1

*Dominio espacial del dataset CR2MET y subdominio de estudio Ñuble–Maule.*



*Nota:* El panel izquierdo muestra la ubicación del subdominio sobre Chile; el panel derecho presenta el detalle con las regiones de Maule y Ñuble destacadas en naranja y verde respectivamente. La malla representa la grilla del dataset (resolución  $0.05^\circ \times 0.05^\circ$ ).

## 4.1.2. Preprocesamiento

El preprocesamiento sigue los pasos generales:

### 1. Lectura y recorte espacial

- Se leen los archivos NetCDF diarios.
- Se recortan las coordenadas latitud y longitud al dominio Maule–Ñuble.
- El resultado es un campo diario  $X_t(x, y)$  de dimensión  $(H \times W)$  para cada día  $t$ .

### 2. Precisión numérica

Con el objetivo de reducir el consumo de memoria y mejorar el rendimiento en GPU, los tensores se almacenan en float16 durante el entrenamiento.

### 3. Normalización

Se normaliza utilizando la media y desviación estándar calculadas solo en el conjunto de entrenamiento:

$$\tilde{X}_t(x, y) = \frac{X_t(x, y) - \mu_{\text{train}}}{\sigma_{\text{train}}}$$

Estos parámetros  $(\mu_{\text{train}}, \sigma_{\text{train}})$  se guardan para desnormalizar al momento de evaluar y visualizar resultados.

Cabe notar que la precipitación diaria presenta una distribución altamente asimétrica, con fuerte concentración en cero y eventos extremos. En este trabajo

se utilizó una estandarización global (parámetros reportados en el Anexo A) principalmente para estabilizar la optimización del autoencoder; esta elección no busca ‘gaussianizar’ la variable ni modelar explícitamente la cero-inflación. Un análisis descriptivo más detallado (porcentaje de ceros y cuantiles por estación/temporada) y la evaluación de transformaciones y/o pérdidas acordes a distribuciones sesgadas (por ejemplo log1p o Tweedie) se plantea como extensión.

#### 4. Construcción de ventanas espacio-temporales

Para alimentar el modelo AE+DMD se construyen ventanas deslizantes de longitud  $T_{in} = 64$  días como entrada al Autoencoder.

#### 5. Partición temporal de datos

La partición se realiza respetando el orden temporal para evitar fuga de información. El entrenamiento se realiza utilizando exclusivamente el conjunto de entrenamiento y de validación.

- Entrenamiento: 1980–2015
- Validación: 2016–2018
- Prueba: 2019–2021

### **4.1.3. Datos para el baseline DeepAR**

Para el modelo baseline DeepAR se construye un conjunto de series temporales a partir de CR2MET. Cada serie se almacena como un registro independiente en el dataset de DeepAR, con frecuencia diaria y el mismo rango temporal que el usado en el modelo AE+DMD, garantizando comparabilidad entre ambos enfoques.

## 4.2. Metodología

La metodología se organiza en dos bloques principales:

1. Modelo híbrido Autoencoder 3D + DMD (AE+DMD) para pronóstico espacio-temporal.
2. Baseline de Persistencia para comparación.
3. Baseline DeepAR para comparación.

### 4.2.1. Modelo híbrido Autoencoder 3D + DMD

a) Formulación general

Sea  $X_t \in \mathbb{R}^{H \times W}$ , el campo de precipitación a día  $t$ , preprocesado como se describió antes.

Para una ventana de entrada de longitud  $T_{in}$ , definimos el bloque espacio-temporal:

$$X_{0:T_{in}-1} = \{X_0, X_1, \dots, X_{T_{in}-1}\} \in \mathbb{R}^{T_{in} \times H \times W \times 1}.$$

El Autoencoder 3D consta de:

- Encoder  $E_\theta$ :

$$z_k = E_\theta(X_k), \quad k = 0, \dots, T_{in} - 1,$$

donde  $z_k \in \mathbb{R}^r$  es el vector latente en el tiempo  $k$ , y  $r \ll H \times W \times T_{in}$ . En la práctica, el encoder es una red convolucional 3D que reduce progresivamente la dimensión espacial y temporal.

Se adopta Conv3D en lugar de arquitecturas recurrentes (por ejemplo LSTM/ConvLSTM) por eficiencia computacional: las convoluciones 3D son

altamente paralelizables en GPU y evitan el costo asociado al procesamiento secuencial en el eje temporal, manteniendo un consumo de memoria más estable al trabajar con ventanas de entrada largas. En este esquema, la dinámica multi-paso se delega al operador lineal estimado por DMD en el espacio latente.

- Decoder  $D_\theta$ :

$$\hat{X}_k = D_\theta(z_k), \quad k = 0, \dots, T_{in} - 1,$$

que reconstruye el campo original a partir del vector latente.

Reuniendo las representaciones latentes en una matriz:

$$Z = [\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_{T_{in}-1}] \in \mathbb{R}^{r \times T_{in}}.$$

La Descomposición Modal Dinámica (DMD) busca un operador lineal  $A \in \mathbb{R}^{r \times r}$  tal que:

$$\mathbf{z}_{k+1} \approx A\mathbf{z}_k, \quad k = 0, \dots, T_{in} - 2.$$

En forma matricial, definimos:

$$Z_- = [\mathbf{z}_0, \dots, \mathbf{z}_{T_{in}-2}], \quad Z_+ = [\mathbf{z}_1, \dots, \mathbf{z}_{T_{in}-1}],$$

y estimamos  $A$  mediante mínimos cuadrados:

$$A = Z_+ Z_-^\dagger,$$

donde  $Z_-^\dagger$  es la pseudoinversa.

Desde un punto de vista operator-theoretic, DMD puede interpretarse como una aproximación finito-dimensional del operador de Koopman actuando sobre observables;

en EDMD esto se formaliza mediante una base de funciones no lineales (Koopman, 1931; Mezić, 2005; Williams et al., 2015).

Una vez estimado  $A$ , el pronóstico latente a  $h$  pasos se obtiene como:

$$z_{T_{\text{in}}+h} \approx A^h z_{T_{\text{in}}-1}, \quad h = 1, \dots, H_{\text{max}}$$

Finalmente, cada estado latente pronosticado se decodifica a espacio físico:

$$\hat{X}_{T_{\text{in}}+h} = D_{\theta}(z_{T_{\text{in}}+h}).$$

#### b) Arquitectura y entrenamiento del Autoencoder

El encoder (Figura 4.2) está compuesto por 2 capas Conv3D con funciones de activación ReLU. Esto permite capturar patrones locales espacio-temporales y reducir progresivamente la resolución del cubo de entrada al DMD. Esta elección responde además a restricciones de cómputo, ya que Conv3D permite entrenar y evaluar múltiples configuraciones con menor costo que variantes recurrentes, manteniendo capacidad de modelar dependencias espacio-temporales locales.

La capa personalizada DMD se encarga de estimar, en cada batch, el operador lineal  $A$  en el espacio latente mediante un ajuste por mínimos cuadrados (con regularización ridge). A partir de la secuencia latente  $z_k$ , la capa genera los estados pronosticados  $\hat{z}_{k+1} \dots \hat{z}_{k+h}$ , los cuales se entregan al decoder compartido para producir la reconstrucción y predicción. Aunque la capa está integrada en el grafo del modelo y contribuye a la función de pérdida, el cálculo de  $A$  es analítico dentro de cada batch y no corresponde a un peso fijo aprendido

por backpropagation. El decoder refleja de forma aproximada la estructura del encoder, utilizando capas Conv3DTranspose para realizar upsampling y reconstruir la salida del espacio latente  $z_k$  para reconstrucción y  $\hat{z}_{k+1} \dots \hat{z}_{k+h}$  para predicción, a la resolución original. La capa de salida utiliza activación lineal, de modo que la reconstrucción pueda abarcar el rango de valores normalizados de precipitación.

**Figura 4.2**

*Arquitectura AE 3D + DMD.*



El entrenamiento del Autoencoder se realiza minimizando la pérdida de reconstrucción y las pérdidas de predicción sobre el conjunto de entrenamiento:

$$\mathcal{L}_{recon} = \frac{1}{N} \sum_{i=1}^N \left\| \mathbf{X}_{recon}^{(i)} - \widehat{\mathbf{X}}_{recon}^{(i)} \right\|^2, \quad X_{recon} = \{X_0, X_1, \dots, X_{T_{in}-1-h_{max}}\}$$

$$\mathcal{L}_{pred} = \frac{1}{N} \sum_{i=1}^N \left\| \mathbf{X}_{pred}^{(i)} - \widehat{\mathbf{X}}_{pred}^{(i)} \right\|^2, \quad X_{pred} = \{X_{T_{in}-h_{max}}, \dots, X_{T_{in}-1}\}$$

$$\mathcal{L} = \lambda_{recon} * \mathcal{L}_{recon} + \lambda_{pred} * \mathcal{L}_{pred}$$

donde  $\mathcal{L}_{\text{pred}}$  mide el error entre la precipitación futura real  $\mathbf{X}_{\text{pred}}^{(i)}$  y la precipitación pronosticada  $\hat{\mathbf{X}}_{\text{pred}}^{(i)}$  obtenida vía  $A^h$  y  $\mathcal{L}_{\text{recon}}$  mide el error de reconstrucción entre  $\mathbf{X}_{\text{recon}}^{(i)}$  y  $\hat{\mathbf{X}}_{\text{recon}}^{(i)}$ , y  $\lambda_{\text{recon}}, \lambda_{\text{pred}}$  controlan el peso relativo de cada término.

Se utiliza validación temprana (*early stopping*) basada en la pérdida de validación para evitar sobreajuste y seleccionar la mejor versión del modelo.

Adicionalmente, se incorporan términos de regularización sobre  $A$  estimada mediante mínimos cuadrados regularizados (ridge) para favorecer estabilidad de las trayectorias latentes.

#### 4.2.2. Modelo baseline

Se utiliza un modelo de referencia determinístico basado en persistencia, donde el pronóstico a  $h$  días corresponde al último campo observado en la ventana de entrada:  $\hat{X}_{t+h} = X_t$ . Este baseline establece un umbral mínimo de desempeño y es especialmente relevante en precipitación diaria debido a su autocorrelación de corto plazo.

Además, como modelo de referencia adicional se implementa DeepAR, un modelo de series de tiempo probabilístico basado en redes recurrentes (LSTM/GRU).

- Entrada: 4800 series temporales diarias univariadas de precipitación, correspondiente a una grilla de 80x60.
- Salida: distribución condicional de la precipitación futura para horizontes de 1 a 5 días.
- Entrenamiento: se usa el mismo período de entrenamiento/validación que para AE+DMD, con ventanas deslizantes de 120 días y esquema de validación temporal consistente.

DeepAR sirve como baseline “moderno” de series de tiempo: captura dependencias temporales y cierta no linealidad, pero no explota directamente la estructura espacial de CR2MET. Dado que DeepAR produce una distribución condicional de la precipitación futura, para comparar con AE+DMD y persistencia se requiere un pronóstico puntual. En este trabajo se utiliza como estimador puntual la media de la distribución predictiva (equivalente al promedio de múltiples muestras generadas), con la cual se calculan las métricas determinísticas reportadas en la Sección de Resultados.

### 4.2.3. Esquema de Evaluación

La evaluación de ambos modelos se realiza sobre el conjunto de prueba y contempla:

1. Métricas escalares por horizonte:

Para cada horizonte  $h = \{1 \dots 5\}$  se calculan:

- MAE (error absoluto medio)
- RMSE (raíz del error cuadrático medio)

En el caso de AE+DMD, las métricas se calculan tanto sobre el campo completo (promediando en espacio y tiempo) como sobre promedios regionales. Para DeepAR, se calculan sobre las series que el modelo pronostica.

2. Análisis temporal agregado:

- Series temporales de error (por ejemplo, MAE diario) para evaluar estabilidad del modelo en el tiempo.

3. Análisis espacial:

- Mapas de MAE/RMSE por celda de grilla para cada horizonte.
- Comparación de sesgos espaciales sistemáticos entre AE+DMD y DeepAR.

Este esquema de evaluación permite cuantificar de manera consistente el desempeño del modelo híbrido AE+DMD frente al baseline DeepAR, tanto desde el punto de vista temporal como espacial, y analizar hasta qué punto la

incorporación de DMD sobre un espacio latente aprendido entrega beneficios reales en el contexto de precipitación regional en Maule–Ñuble.

#### 4. Métricas escalares por horizonte:

Dado el carácter intermitente de la precipitación diaria, además de MAE/RMSE se incorporan métricas orientadas a eventos y habilidad espacial, evaluadas sobre el conjunto de prueba:

- Métricas de detección: se define ocurrencia de evento por umbral  $u \in \{1,10\}$  mm/día mediante  $1[X \geq u]$  y  $1[\hat{X} \geq u]$ . Con ello se computan POD, FAR, CSI y BIAS, agregando sobre el dominio espacial y el periodo de evaluación.
- Métricas condicionadas a celdas húmedas (wet-only): para  $u = 1$  mm/día, se calculan  $MAE_{wet}$  y  $RMSE_{wet}$  considerando solo celdas donde el observado cumple  $X \geq 1$ . Esto permite evaluar el error en magnitud cuando efectivamente ocurre precipitación.
- Habilidad espacial (FSS): se evalúa Fractions Skill Score para  $u \in \{1,10\}$  mm/día usando ventanas  $3 \times 3$ ,  $5 \times 5$ ,  $9 \times 9$  y  $17 \times 17$ , con el fin de medir consistencia espacial a distintas escalas (tolerancia a errores de localización).

Para facilitar la interpretación, en el cuerpo del informe se reportan resultados representativos para  $h \in \{1,3,5\}$  (y para FSS se consideran  $h \in \{3,5\}$ ).

## 5. Resultados

Antes de evaluar la capacidad de pronóstico del modelo AE+DMD, es necesario revisar qué tan bien el Autoencoder 3D reconstruye las ventanas espacio-temporales de precipitación en el conjunto de prueba, ya que cualquier error en esta etapa actúa como piso para el desempeño del bloque DMD.

### 5.1. Métricas de reconstrucción

La Tabla 5.1 muestra que el Autoencoder 3D alcanza un MAE de 0,69 y un RMSE de 1,05 en el conjunto de prueba, lo que indica que la reconstrucción de las ventanas espacio-temporales es razonablemente precisa en relación con la variabilidad típica de la precipitación diaria en el dominio Maule-Ñuble. En promedio, el modelo es capaz de reproducir la mayor parte de la señal sin incurrir en errores sistemáticos de gran magnitud.

**Tabla 5.1**

*MAE y RMSE en mm/día de reconstrucción en el conjunto de prueba.*

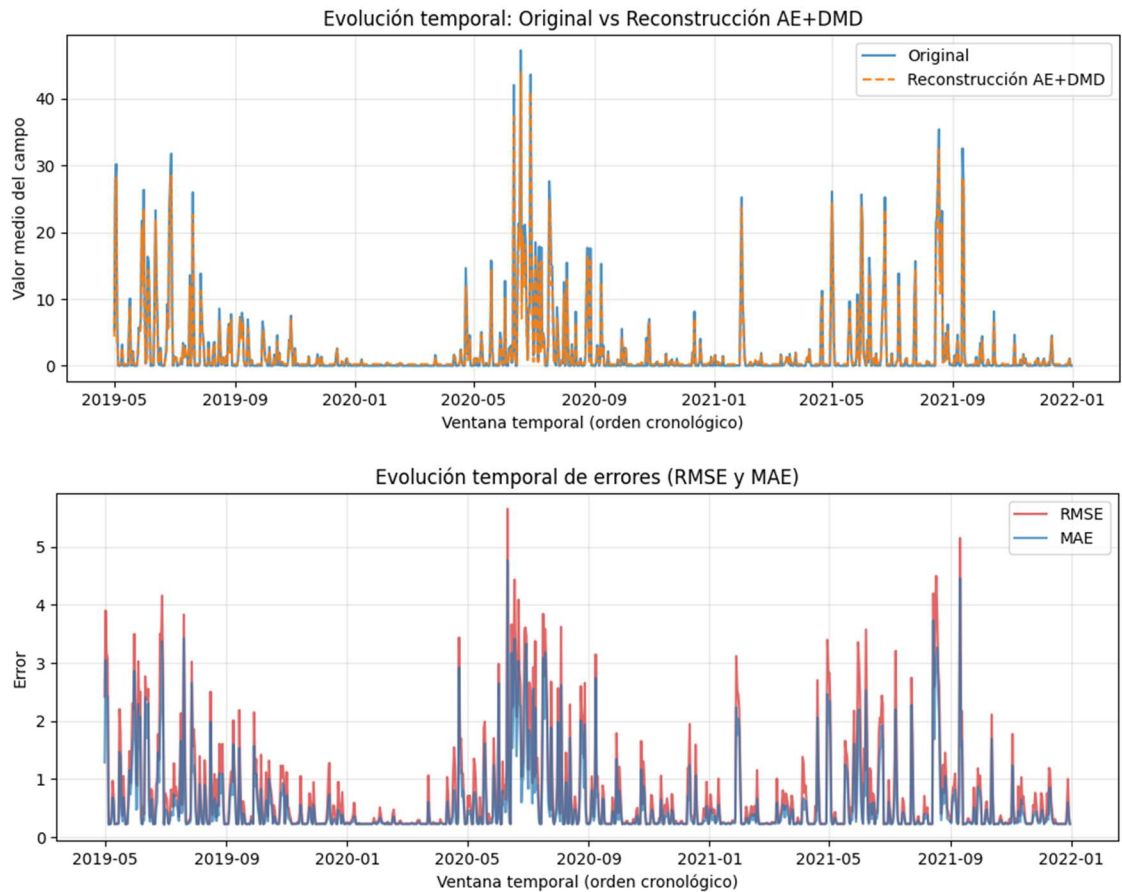
Conjunto	MAE	RMSE
Prueba	0,69	1,05

La Figura 5.1 permite analizar este comportamiento en el tiempo. La serie del promedio espacial de precipitación observada y reconstruida muestra que el Autoencoder sigue de forma adecuada la estacionalidad y la ocurrencia de eventos de lluvia, incluyendo los pulsos más relevantes, mientras que las diferencias más visibles se concentran en la altura exacta de algunos picos. De forma consistente, las curvas de MAE y RMSE en el tiempo presentan máximos coincidentes con los eventos de mayor precipitación, lo que sugiere que el AE tiende a subestimar parcialmente los episodios más intensos, pero mantiene errores bajos en condiciones secas o de lluvia débil.

Desde el punto de vista espacial, la Figura 5.2 muestra que el patrón medio de precipitación reconstruida reproduce el gradiente costa–cordillera y las zonas de mayor acumulación en el sector cordillerano. El mapa de errores evidencia que las discrepancias se concentran precisamente en estas áreas de mayor intensidad y variabilidad espacial, mientras que el valle central y sectores costeros presentan errores promedio sensiblemente menores

**Figura 5.1**

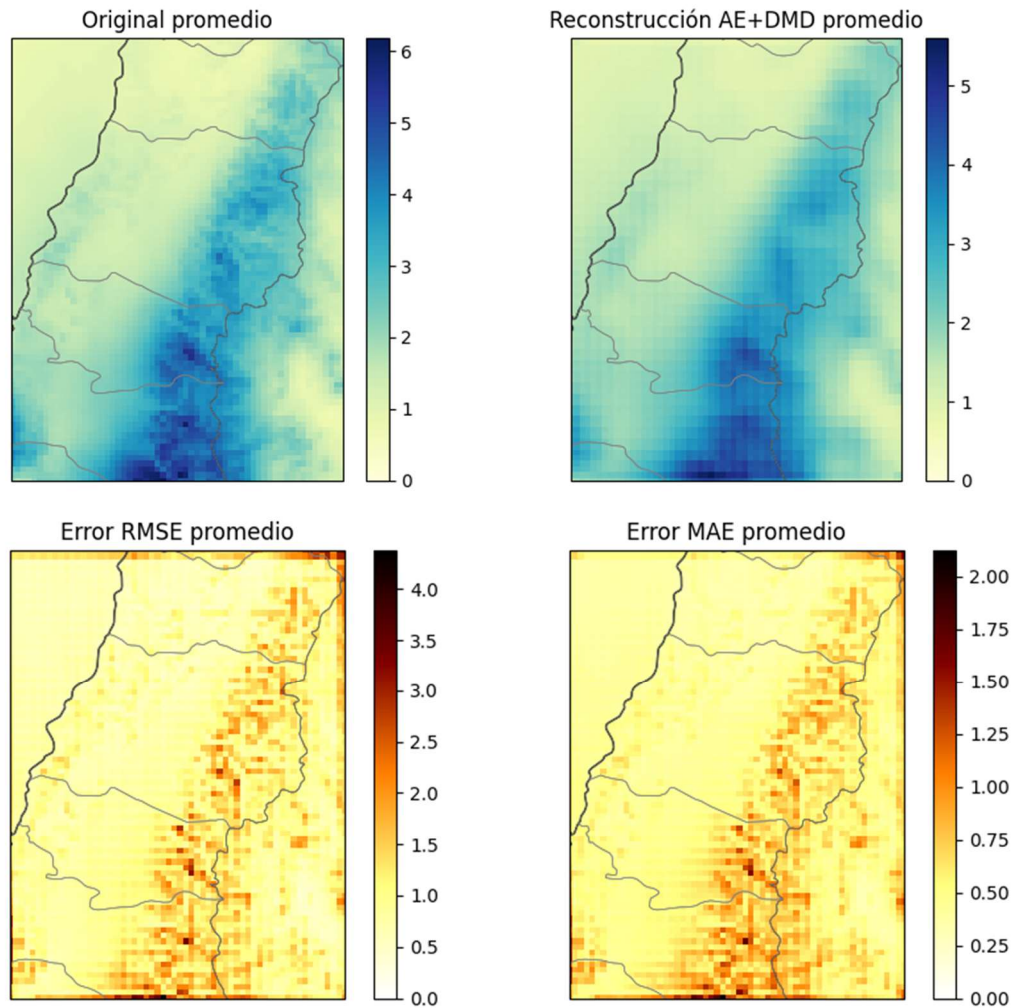
*Reconstrucción con AE+DMD de promedio espacial de precipitaciones en el tiempo.*



*Nota:* Arriba, comparación entre promedio espacial de precipitaciones y su reconstrucción en el tiempo. Abajo, promedio espacial de RMSE y MAE en el tiempo.

**Figura 5.2**

*Reconstrucción con AE+DMD de promedio temporal de precipitaciones en el espacio.*

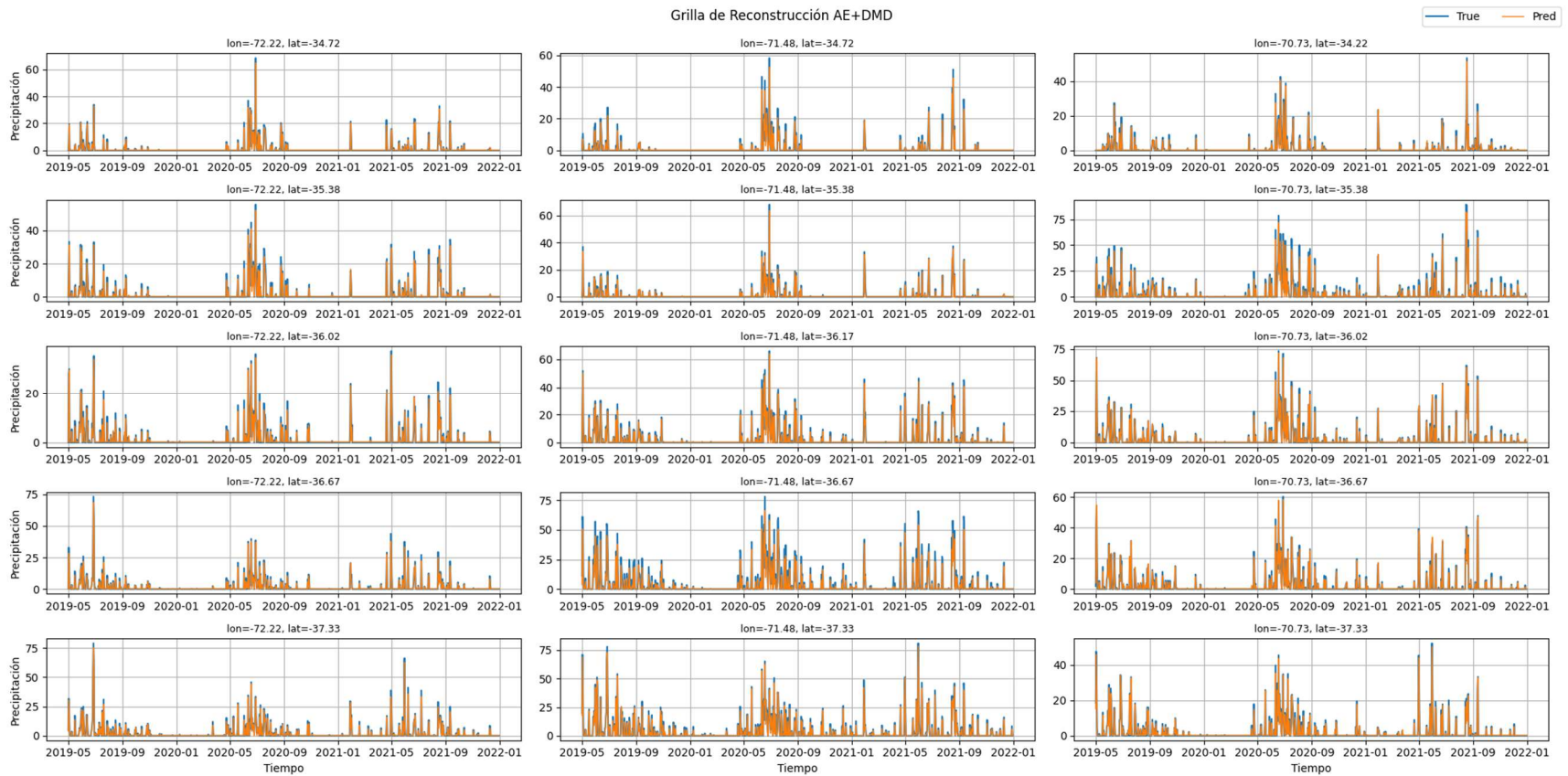


*Nota:* Arriba, comparación entre promedio temporal de precipitaciones y su reconstrucción en el espacio. Abajo, promedio temporal de RMSE y MAE en el espacio.

Finalmente, la grilla de series mostrada en la Figura 5.3 confirma que, en distintos puntos de la malla, el Autoencoder captura bien la frecuencia y localización temporal de los eventos, aunque suaviza la magnitud de algunos máximos diarios. En conjunto, estos resultados indican que el espacio latente aprendido es suficientemente representativo para que DMD modele la dinámica, aun cuando exista una pérdida de fidelidad en extremos muy localizados.

**Figura 5.3**

*Grilla espacial de reconstrucción temporal utilizando AE+DMD.*



## 5.2. Desempeño de pronóstico del AE+DMD

### 5.2.1. Métricas agregadas por horizonte

La Tabla 5.2 resume el desempeño del modelo AE+DMD para horizontes de 1 a 5 días. El MAE aumenta desde 2,00 mm/día para el pronóstico a 1 día hasta 2,86 mm/día para 4 días, mientras que el RMSE crece desde 5,13 a 7,24 mm/día en el mismo rango. Este comportamiento es coherente con la acumulación progresiva de error a medida que se propaga la dinámica en el espacio latente mediante el operador lineal. El horizonte de 5 días presenta un ligero descenso del MAE y RMSE respecto a 4 días, lo que puede explicarse por un efecto de suavizado: a medida que se promedia sobre ventanas más largas y se pierden estructuras de alta frecuencia, los errores extremos disminuyen levemente, aunque el campo pronosticado sea más difuso.

**Tabla 5.2**

*MAE y RMSE del modelo AE+DMD para diferentes horizontes.*

<b>Horizonte</b>	<b>MAE</b>	<b>RMSE</b>
Forecast 1 día	2,00	5,13
Forecast 2 días	2,32	5,69
Forecast 3 días	2,70	6,41
Forecast 4 días	2,86	7,24
Forecast 5 días	2,59	6,62

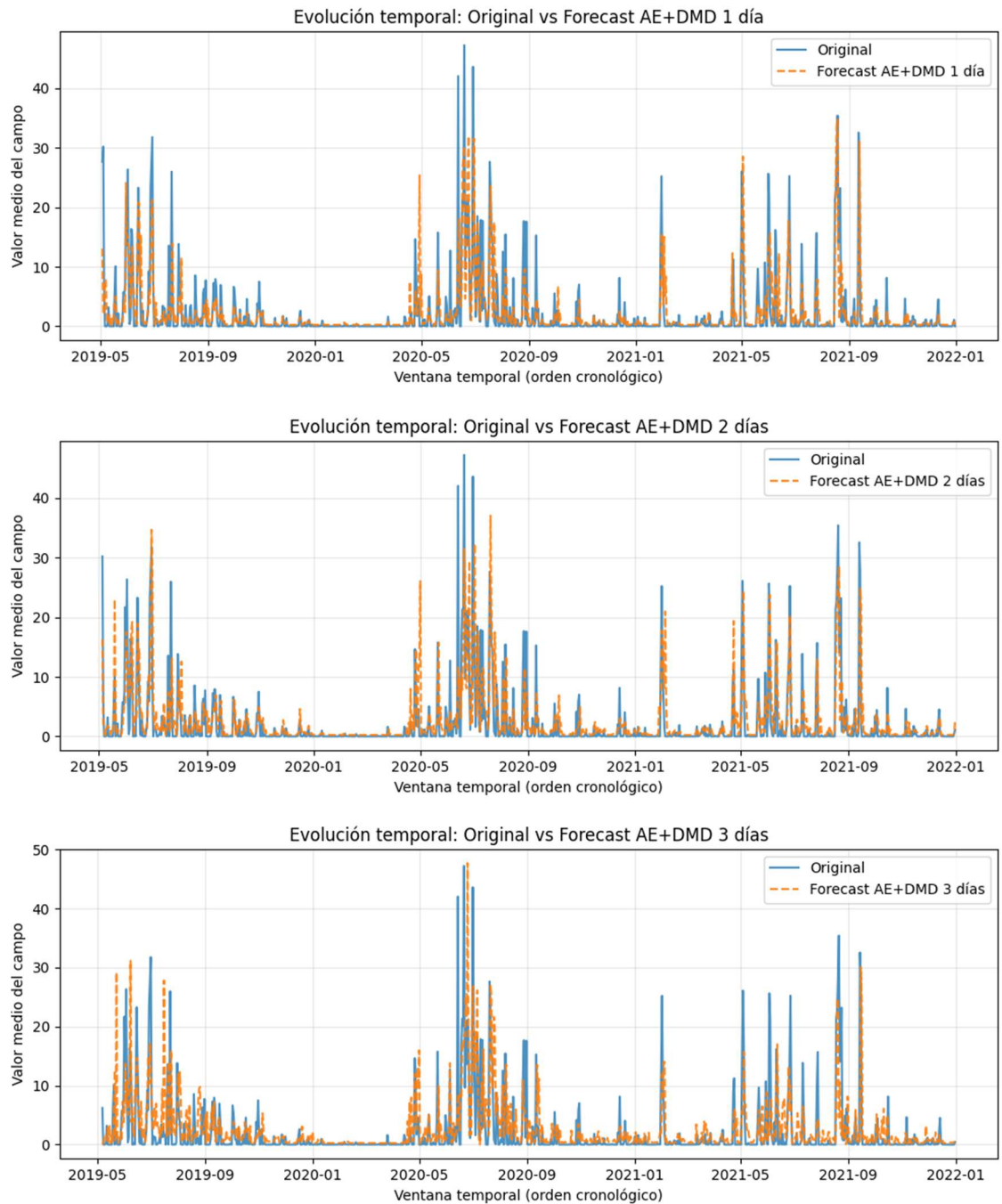
En términos prácticos, estos valores sitúan al modelo AE+DMD en un rango de error aceptable para precipitación diaria regional. El incremento entre  $h=1$  y  $h=3$  es gradual y no explosivo, lo que indica que el operador DMD mantiene una dinámica relativamente estable en horizontes cortos. A partir de  $h=4$  el error crece de forma más marcada, lo que sugiere que la aproximación lineal en el espacio latente resulta menos fiable cuando se intenta extrapolar la trayectoria más allá de unos pocos días.

### **5.2.2. Evolución temporal de precipitación**

Las Figuras 5.4 y 5.6 muestran la evolución temporal del promedio espacial de precipitación observada y pronosticada por el modelo AE+DMD para horizontes de 1 a 5 días. Para  $h=1$ , la serie pronosticada reproduce con buena precisión tanto la estacionalidad como la secuencia de eventos, capturando bien la ocurrencia y el orden de magnitud de los pulsos principales. A medida que el horizonte aumenta, la forma general de la señal se mantiene, pero los picos se vuelven más atenuados y algunos eventos intensos aparecen retrasados o suavizados, en especial para 4 y 5 días.

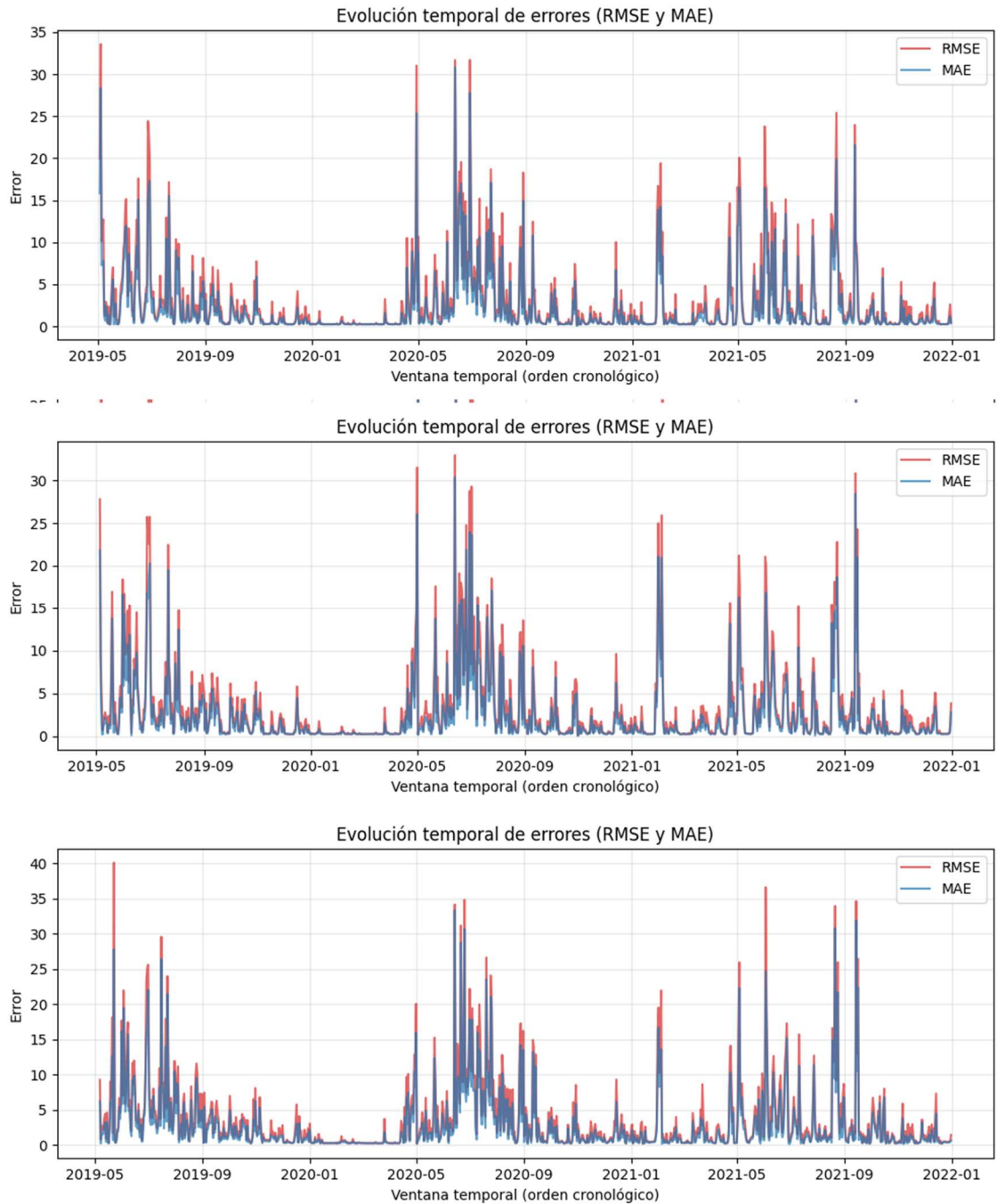
**Figura 5.4**

*Comparación entre promedio espacial de precipitaciones y su predicción de 1 a 3 días en el tiempo.*



**Figura 5.5**

*Promedio espacial de RMSE y MAE en el tiempo para horizonte de pronóstico de 1 a 3 día (visto desde arriba hacia abajo).*

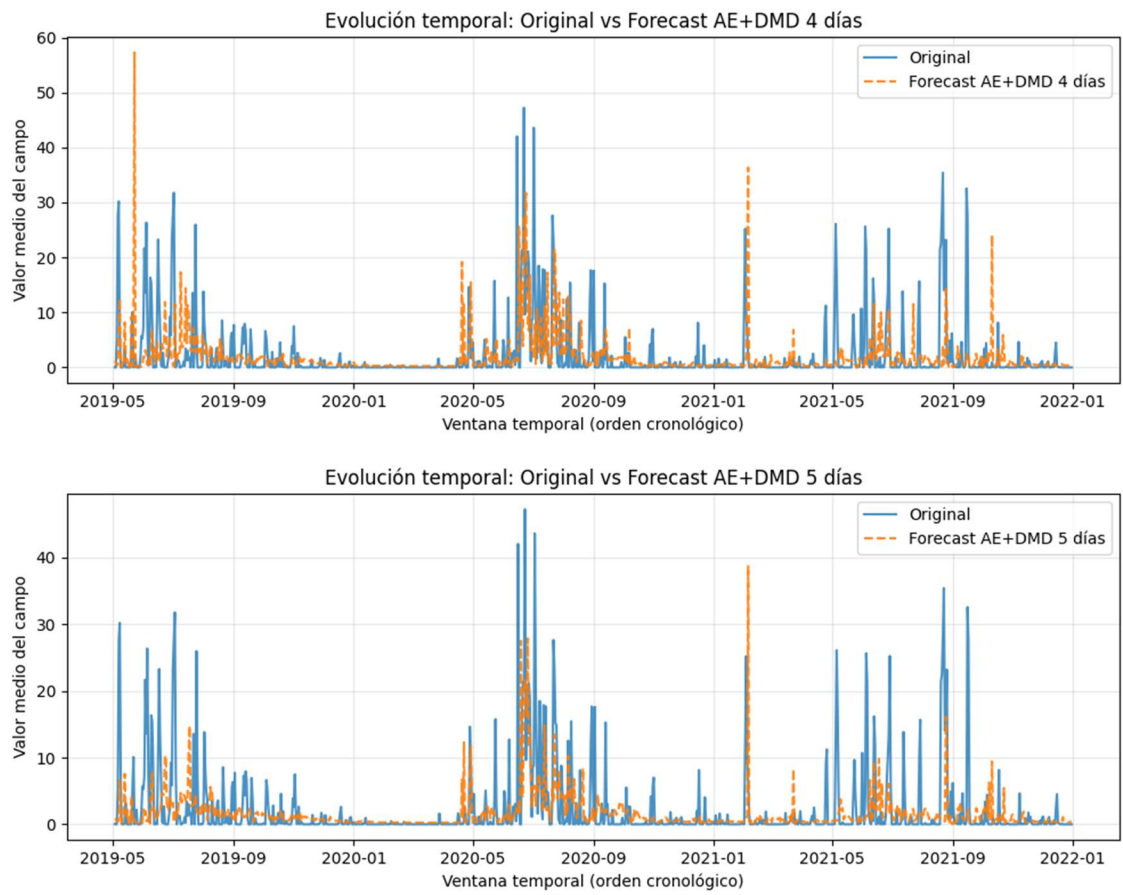


Esta impresión cualitativa se confirma con las Figuras 5.5 y 5.7, donde se presenta el MAE y el RMSE promedio en el tiempo para cada horizonte. En todos los casos, los errores se concentran alrededor de episodios de lluvia moderada a intensa, mientras que durante los períodos secos los valores de MAE y RMSE son cercanos a cero. Para  $h=1-3$  los picos de error, aunque altos en algunos eventos particulares, se mantienen acotados y el modelo muestra un comportamiento estable a lo largo de todo el período de prueba. Para  $h=4-5$ , en cambio, los máximos de error aumentan y se hacen más frecuentes, reflejando la dificultad del modelo para anticipar con precisión la intensidad de eventos que están más lejos en el tiempo.

En conjunto, la evolución temporal del error sugiere que el AE+DMD es especialmente útil en horizontes cortos, donde consigue seguir de manera razonable la secuencia de eventos y su magnitud promedio. A medida que aumenta el horizonte, el pronóstico conserva la estructura general, pero pierde fidelidad en el detalle de los picos diarios.

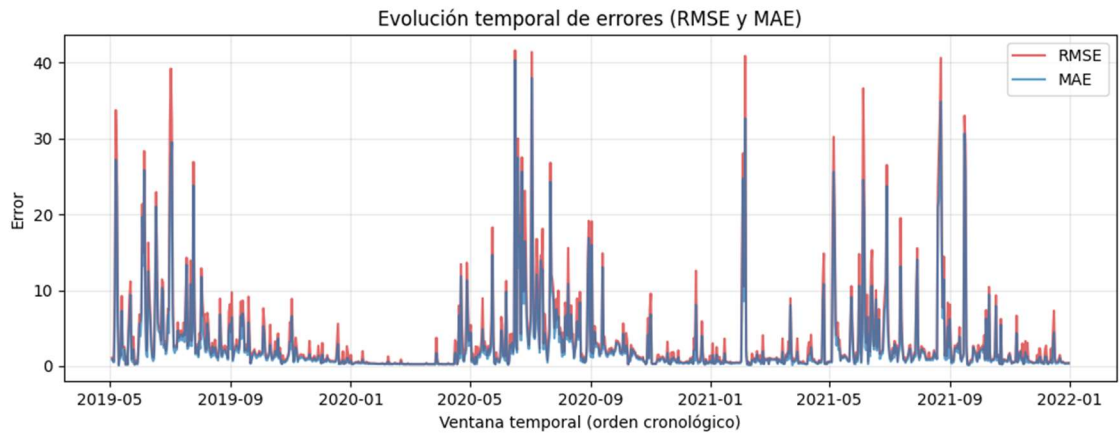
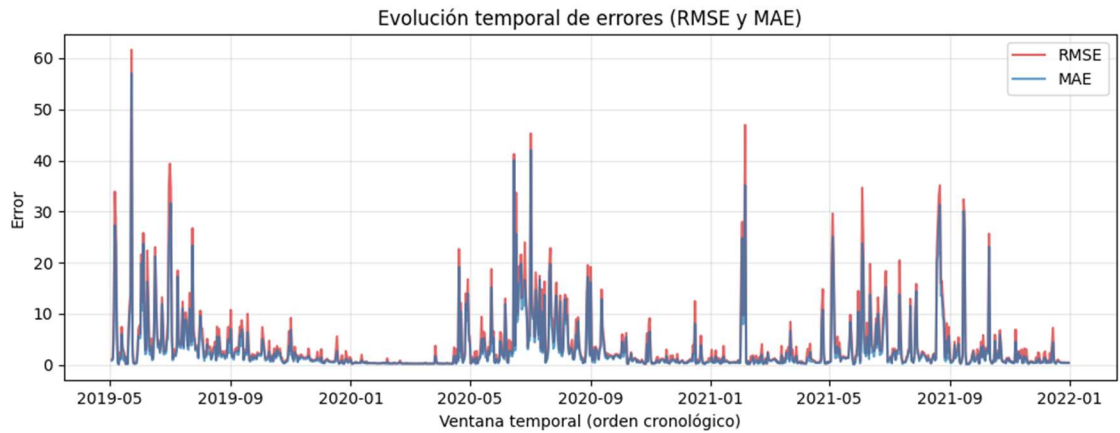
**Figura 5.6**

*Comparación entre promedio espacial de precipitaciones y su predicción de 4 a 5 días en el tiempo.*



**Figura 5.7**

*Promedio espacial de RMSE y MAE en el tiempo para horizonte de pronóstico de 4 a 5 días ( visto desde arriba hacia abajo).*



### **5.2.3. Comparación campo medio observado vs forecast**

Las Figuras 5.8 a 5.12 permiten analizar el comportamiento espacial del pronóstico del modelo AE+DMD para horizontes de 1 a 5 días, comparando el promedio temporal de la precipitación observada (desplazada según el horizonte) con el promedio del forecast y los mapas correspondientes de MAE y RMSE.

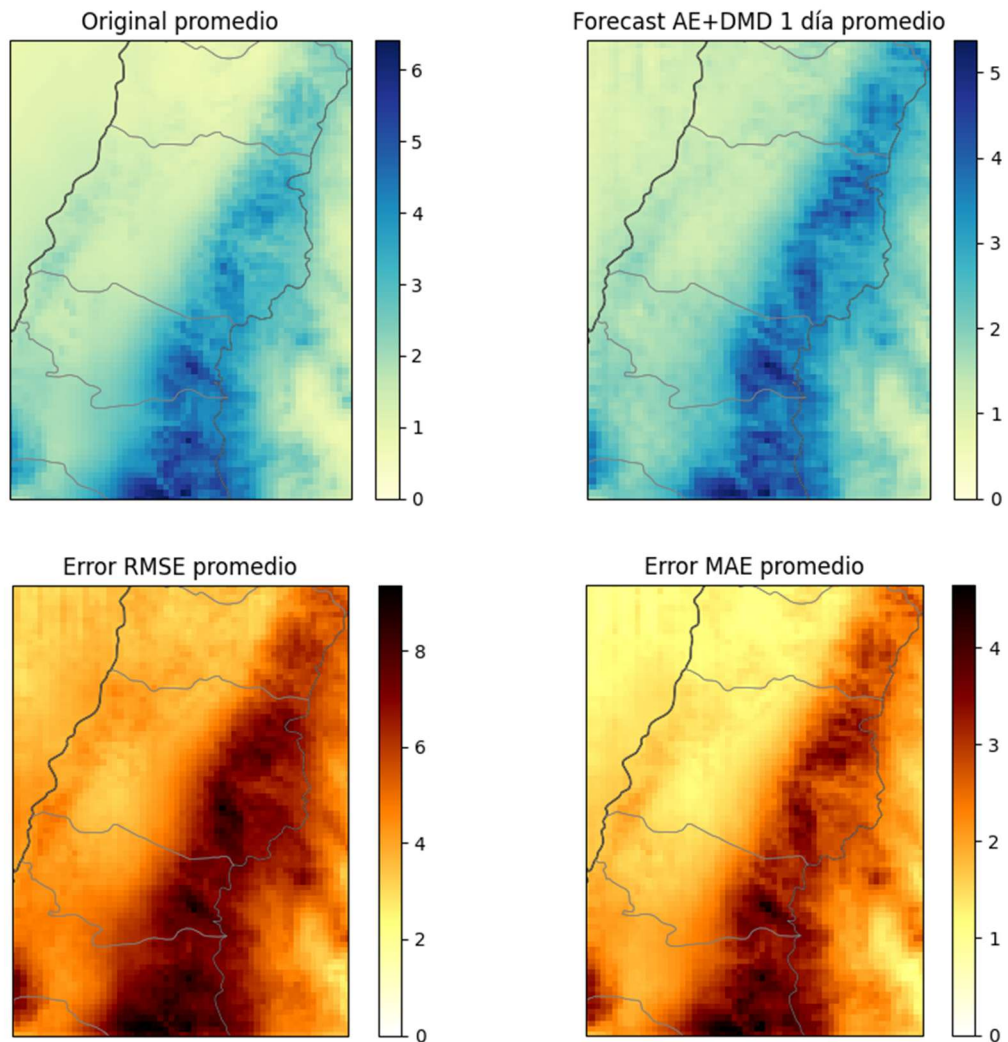
Para el horizonte de 1 día (Figura 5.8), el campo medio pronosticado reproduce de manera clara el gradiente costa–cordillera, así como la franja de máximos en el sector cordillerano del dominio. Las diferencias se concentran en la intensidad exacta de estos máximos: el modelo tiende a subestimar ligeramente la precipitación en las celdas más lluviosas y a sobreestimar marginalmente algunas zonas adyacentes, lo que se refleja en valores de RMSE más elevados en la cordillera y errores moderados en el valle central. El patrón general, sin embargo, se mantiene bien alineado con la climatología observada.

A medida que aumenta el horizonte (Figuras 5.9 a 5.12), el campo medio pronosticado conserva la estructura básica del gradiente espacial, pero la banda de máximos se vuelve progresivamente más difusa y desplazada. El mapa de errores muestra un incremento de MAE y RMSE en el sector cordillerano y precordillerano, donde la combinación de topografía compleja y mayor variabilidad temporal de la precipitación hace más difícil la anticipación de la intensidad media. En el valle central y zonas costeras los errores crecen menos, lo que indica que el modelo mantiene un desempeño razonable en áreas con menor variabilidad espacial.

En síntesis, el análisis del campo medio confirma que el modelo AE+DMD captura adecuadamente la estructura espacial dominante del dominio para todos los horizontes, con una degradación gradual de la nitidez de los máximos y un aumento del error medio a medida que se extiende el pronóstico. Las zonas más afectadas corresponden a la cordillera, donde se concentran tanto los mayores acumulados de precipitación como los errores más altos.

**Figura 5.8**

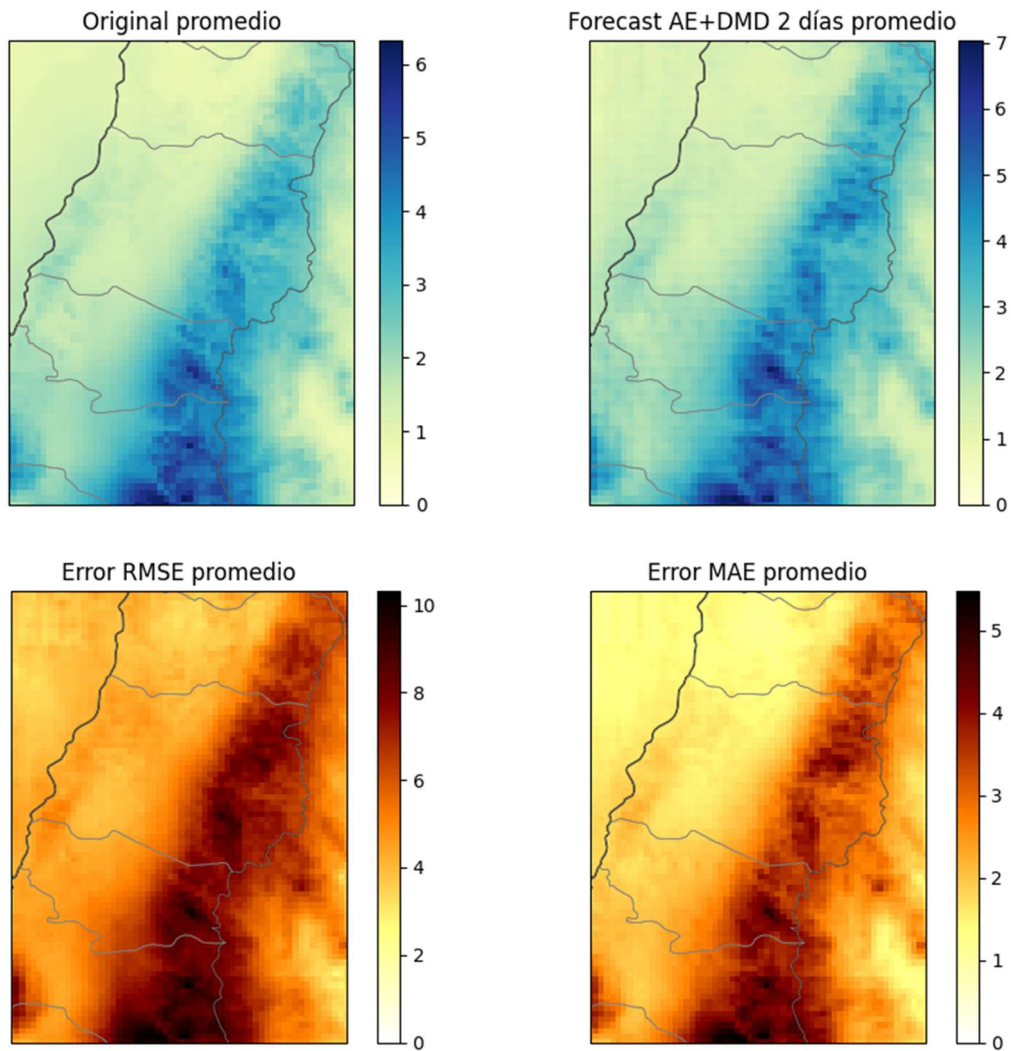
*Forecast de 1 día del promedio temporal de precipitaciones en el espacio.*



*Nota:* Arriba, comparación entre promedio temporal de precipitaciones (desplazada 1 día) y el promedio de predicción de 1 día en el espacio. Abajo, promedio temporal de RMSE y MAE en el espacio.

**Figura 5.9**

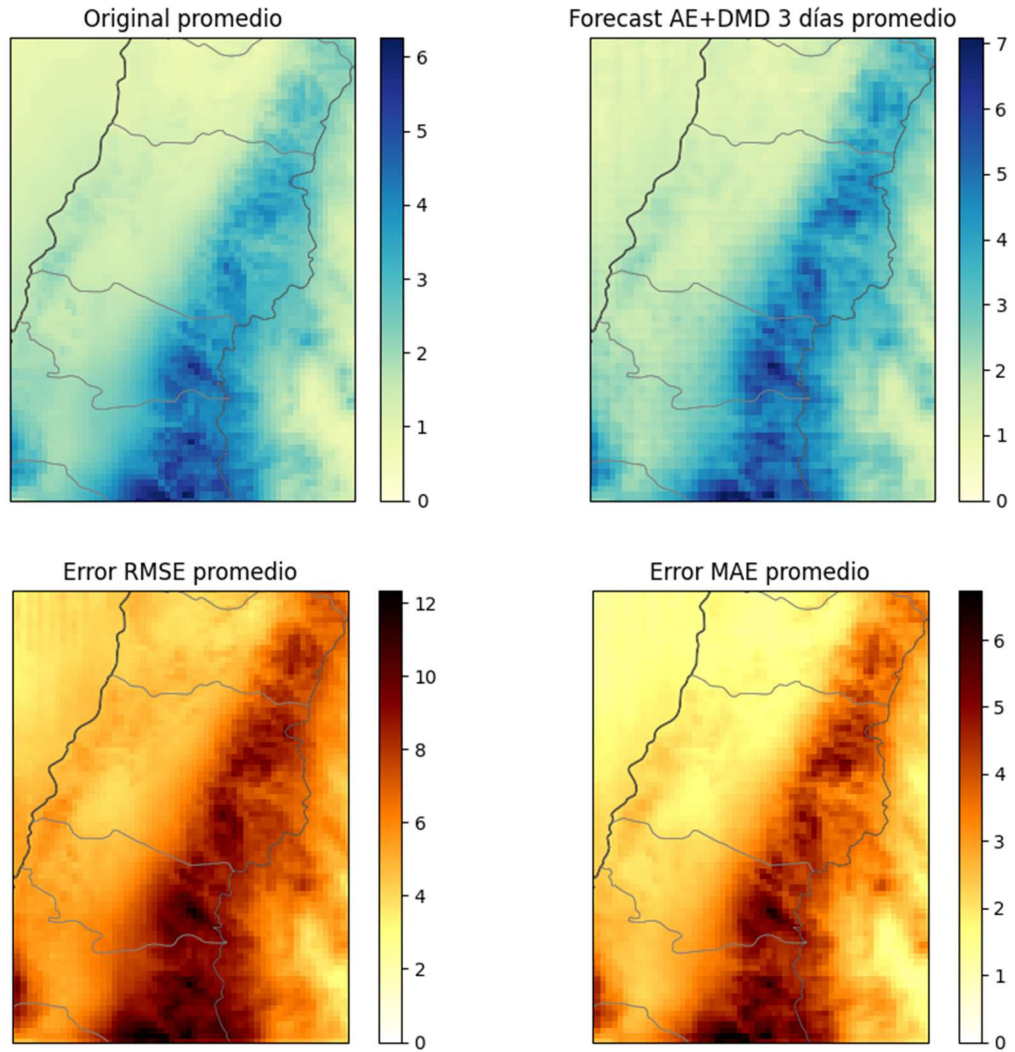
*Forecast de 2 días del promedio temporal de precipitaciones en el espacio.*



*Nota:* Arriba, comparación entre promedio temporal de precipitaciones (desplazada 2 días) y el promedio de predicción de 2 días en el espacio. Abajo, promedio temporal de RMSE y MAE en el espacio.

**Figura 5.10**

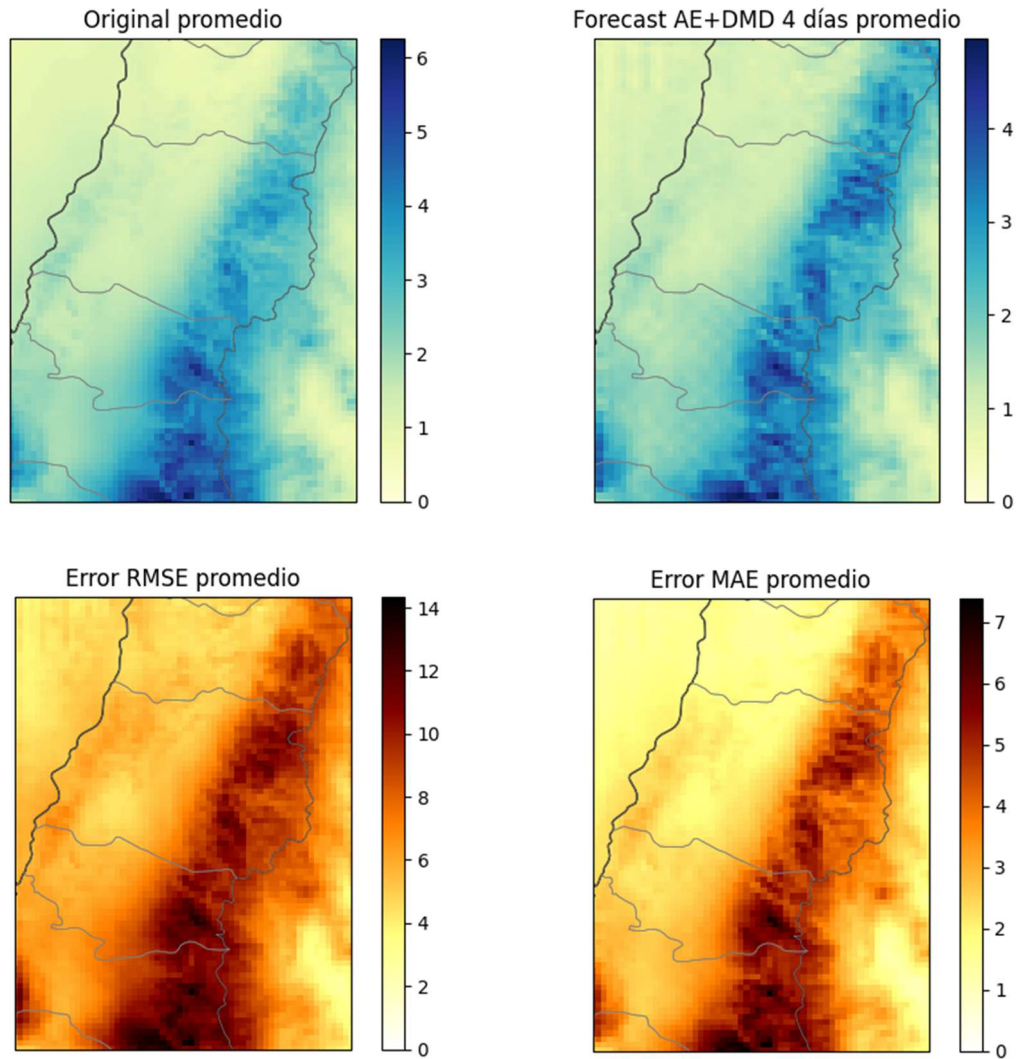
*Forecast de 3 días del promedio temporal de precipitaciones en el espacio.*



*Nota:* Arriba, comparación entre promedio temporal de precipitaciones (desplazada 3 días) y el promedio de predicción de 3 días en el espacio. Abajo, promedio temporal de RMSE y MAE en el espacio.

**Figura 5.11**

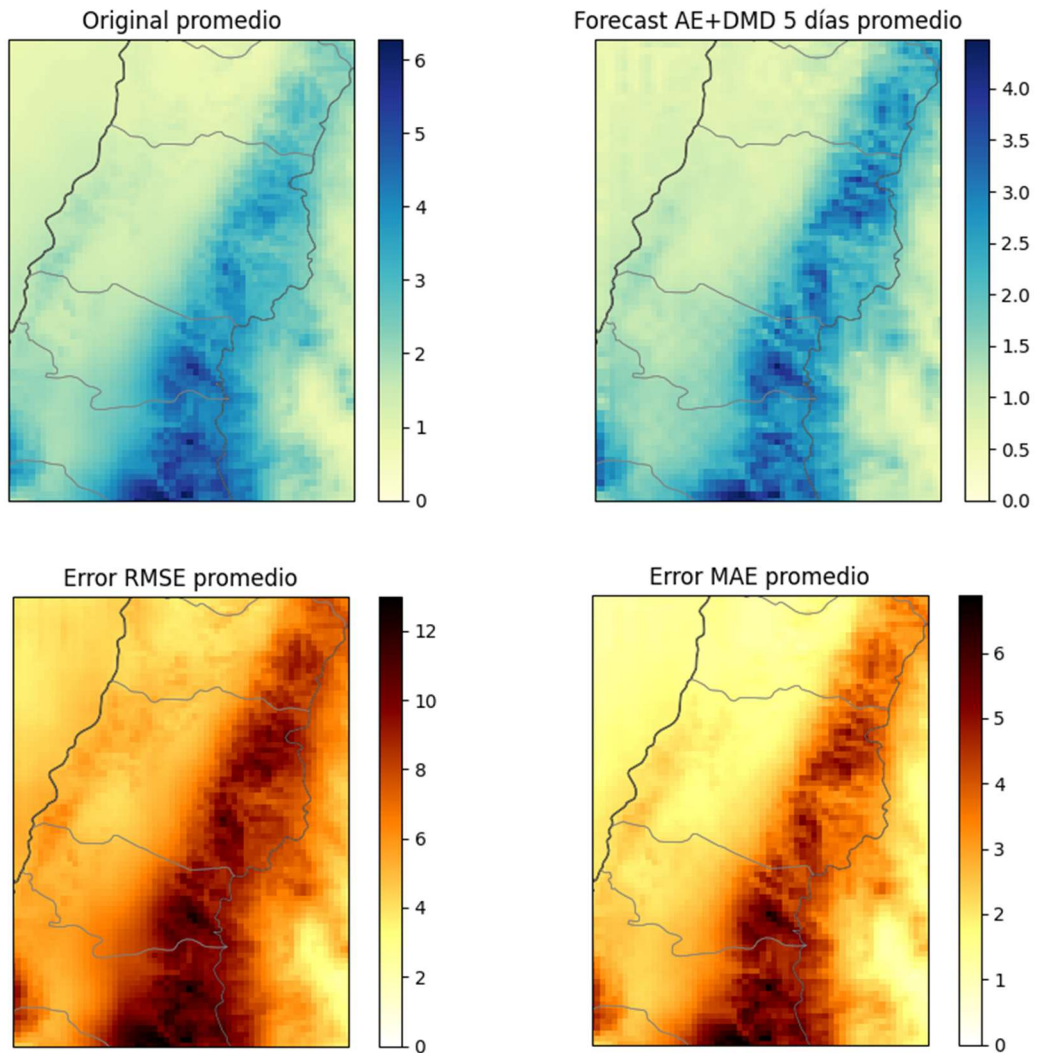
*Forecast de 4 días del promedio temporal de precipitaciones en el espacio.*



*Nota:* Arriba, comparación entre promedio temporal de precipitaciones (desplazada 4 días) y el promedio de predicción de 4 días en el espacio. Abajo, promedio temporal de RMSE y MAE en el espacio.

**Figura 5.12**

*Forecast de 5 días del promedio temporal de precipitaciones en el espacio.*



*Nota:* Arriba, comparación entre promedio temporal de precipitaciones (desplazada 5 días) y el promedio de predicción de 5 días en el espacio. Abajo, promedio temporal de RMSE y MAE en el espacio.

### 5.3. Comparación AE+DMD vs DeepAR

La Tabla 5.3 incorpora el baseline de persistencia como umbral mínimo de desempeño y una comparación directa entre el modelo híbrido AE+DMD y el baseline DeepAR. Como era esperable en precipitación diaria, la persistencia solo resulta relativamente competitiva a 1 día (MAE = 2,32; RMSE = 6,69), pero su error aumenta rápidamente con el horizonte (MAE  $\approx$  3,29 y RMSE  $\approx$  8,73–8,74 a 4–5 días), reflejando la pérdida de predictibilidad y la alta intermitencia del fenómeno.

**Tabla 5.3**

*Comparación MAE y RMSE entre modelos de Persistencia, DeepAR y AE+DMD.*

Modelo	Métrica	Forecast 1 día	Forecast 2 días	Forecast 3 días	Forecast 4 días	Forecast 5 días
Persistencia	MAE	2,32	2,95	3,19	3,29	3,29
	RMSE	6,69	8,14	8,61	8,74	8,73
DeepAR	MAE	2,40	2,35	2,30	2,28	2,28
	RMSE	7,13	7,20	7,14	7,14	7,14
AE + DMD	$\Delta$ MAE	-3,45%	20,34%	27,90%	30,70%	30,70%
	$\Delta$ RMSE	-6,58%	11,55%	17,07%	18,31%	18,21%
AE + DMD	MAE	2,00	2,32	2,70	2,86	2,59
	RMSE	5,13	5,69	6,41	7,24	6,62
AE + DMD	$\Delta$ MAE	13,79%	21,36%	15,36%	13,07%	21,28%
	$\Delta$ RMSE	23,32%	30,10%	25,55%	17,16%	24,17%

*Nota.* MAE y RMSE calculados sobre el conjunto de prueba en mm/día.  $\Delta$  MAE y  $\Delta$  RMSE representan la variación porcentual respecto a Persistencia y se calcularon como  $(\text{modelo} - \text{baseline})/\text{baseline} \times 100\%$ . Valores negativos indican mejora.

Al comparar contra persistencia mediante las métricas relativas  $\Delta$ , se observa que DeepAR presenta un comportamiento mixto: a 1 día es levemente peor que persistencia ( $\Delta\text{MAE} = -3,45\%$ ;  $\Delta\text{RMSE} = -6,58\%$ ), lo que sugiere que a muy corto plazo la inercia temporal domina y el modelo no logra superarla consistentemente. Sin embargo, desde 2 a 5 días DeepAR mejora de forma sostenida, especialmente en MAE ( $\Delta\text{MAE} \approx 20\text{--}31\%$ ), lo que indica una reducción del error promedio conforme aumenta el horizonte.

Por su parte, el modelo AE + DMD supera a persistencia en todos los horizontes y, en particular, muestra mejoras claras en RMSE ( $\Delta\text{RMSE} \approx 17\text{--}30\%$  entre 1 y 5 días), alcanzando los menores RMSE en todos los casos (por ejemplo, RMSE = 5,13 a 1 día y 6,62 a 5 días). Esto es relevante porque RMSE penaliza más los errores grandes, por lo que el modelo híbrido parece manejar mejor las discrepancias de mayor magnitud (como ejemplo, eventos más intensos o patrones espaciales que generan errores altos). En cambio, en MAE DeepAR tiende a ser más bajo que AE+DMD en horizontes intermedios (3–4 días), lo que sugiere que DeepAR reduce mejor el error promedio, mientras que AE+DMD entrega mayor robustez frente a errores grandes (menor RMSE).

En síntesis, la persistencia establece un piso exigente a 1 día, pero se degrada rápidamente; desde 2 días en adelante ambos modelos aportan mejora real. DeepAR destaca en MAE (error promedio) a horizontes mayores, mientras que AE+DMD destaca en RMSE (errores grandes) de manera consistente a lo largo de 1–5 días.

### **5.3.1. Comparación temporal del promedio espacial y su error (AE+DMD, DeepAR y Persistencia)**

Las Figuras 5.13–5.18 presentan una evaluación temporal basada en el promedio espacial del campo de precipitación. Para cada horizonte  $h \in \{1,3,5\}$ , se compara la serie temporal original con la serie pronosticada por AE+DMD, DeepAR y el baseline de Persistencia (Fig. 5.13, 5.15 y 5.17). Complementariamente, se muestra la evolución temporal del error (MAE y RMSE) calculado sobre el dominio para cada ventana temporal (Fig. 5.14, 5.16 y 5.18).

#### **Horizonte 1 día**

En  $h = 1$  (Fig. 5.13), la persistencia resulta un baseline competitivo, ya que la autocorrelación de corto plazo permite aproximar razonablemente el nivel medio del campo. No obstante, al revisar la serie temporal completa, se observa que los eventos intensos (picos) siguen siendo el principal foco de discrepancias, especialmente cuando existe un cambio abrupto entre días consecutivos. La Fig. 5.14 confirma este patrón: los mayores valores de MAE/RMSE se concentran en periodos con mayor actividad de precipitación, donde los modelos tienden a diferenciarse por su capacidad de capturar o amortiguar esos extremos.

### **Horizonte 3 días**

En  $h = 3$  (Fig. 5.15), la degradación de la persistencia se vuelve evidente: al repetir el último estado observado, el baseline pierde capacidad para anticipar cambios de régimen y su pronóstico se vuelve progresivamente menos coherente con la serie original. En contraste, los modelos entrenados mantienen mayor consistencia temporal, aunque con tendencia a subestimar picos en algunos periodos. La evolución de errores (Fig. 5.16) vuelve a mostrar que el error se dispara en ventanas asociadas a precipitación intensa, lo cual es consistente con la naturaleza intermitente del fenómeno.

### **Horizonte 5 días**

En  $h = 5$  (Fig. 5.17), la persistencia se consolida como el peor referente en términos de coherencia temporal: su supuesto base (“mantener el último estado”) ya no representa adecuadamente la dinámica a este horizonte, en especial durante semanas con alta variabilidad. La Fig. 5.18 refuerza esta lectura al mostrar errores más pronunciados y frecuentes en comparación con horizontes menores. En este escenario, las diferencias entre modelos se vuelven más visibles principalmente durante eventos extremos, donde el error aumenta fuertemente y el RMSE actúa como un indicador más sensible a discrepancias de gran magnitud.

Estos resultados cualitativos son consistentes con la comparación global: la persistencia es razonable a  $h = 1$ , pero su desempeño cae rápidamente a medida que aumenta el horizonte. En términos generales, DeepAR tiende a reducir el error promedio (MAE) a horizontes mayores, mientras que AE+DMD destaca cuando se penalizan más los errores grandes (RMSE), lo que sugiere diferencias en la forma de manejar eventos intensos.

**Nota general para Figuras 5.13, 5.15 y 5.17**

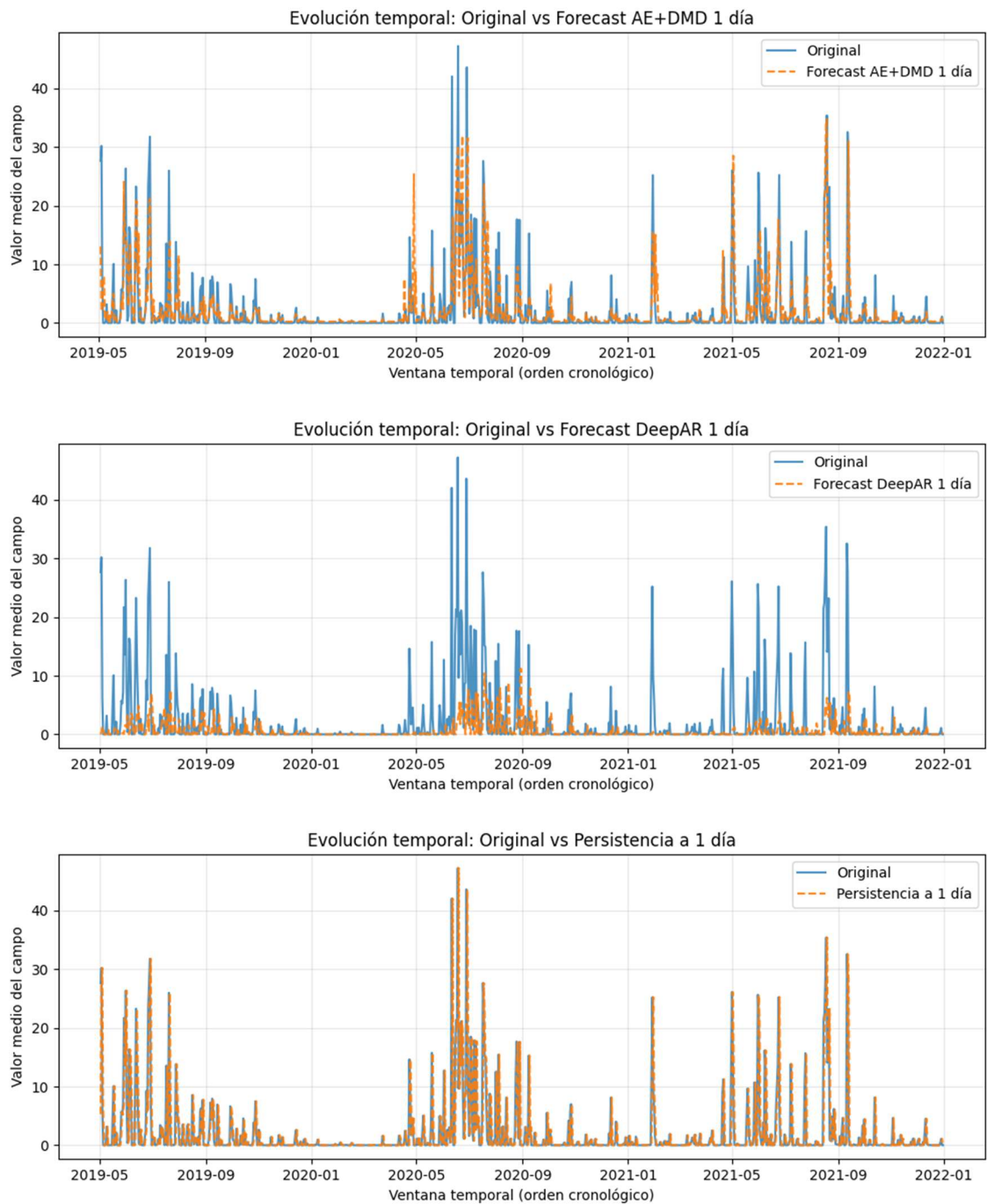
*Nota:* Arriba, comparación entre promedio espacial y predicción del AE+DMD. En medio, comparación entre promedio espacial y predicción del DeepAR. Abajo, comparación entre promedio espacial y predicción de Persistencia.

**Nota general para Figuras 5.14, 5.16 y 5.18**

*Nota:* Promedio espacial de RMSE y MAE en el tiempo, arriba AE+DMD, en medio DeepAR y abajo Persistencia.

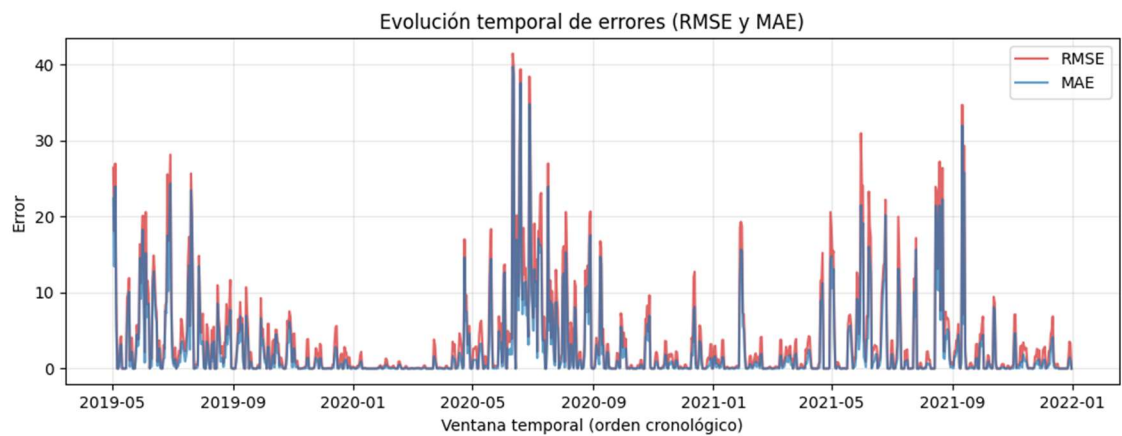
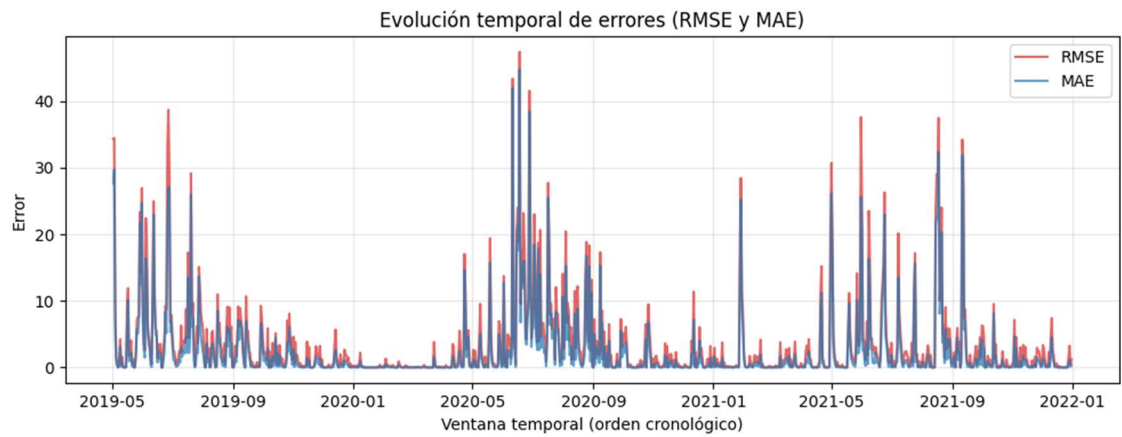
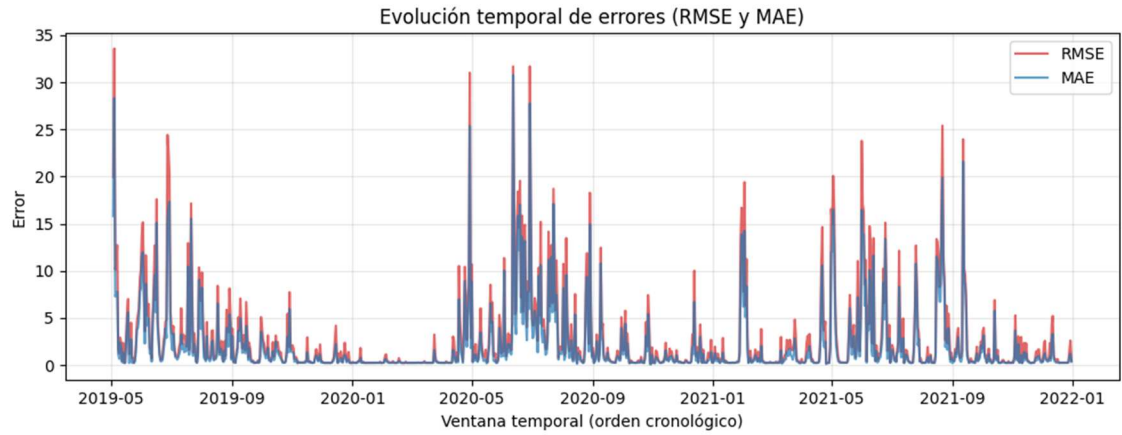
**Figura 5.13**

*Forecast de 1 día del promedio espacial de precipitaciones en el tiempo.*



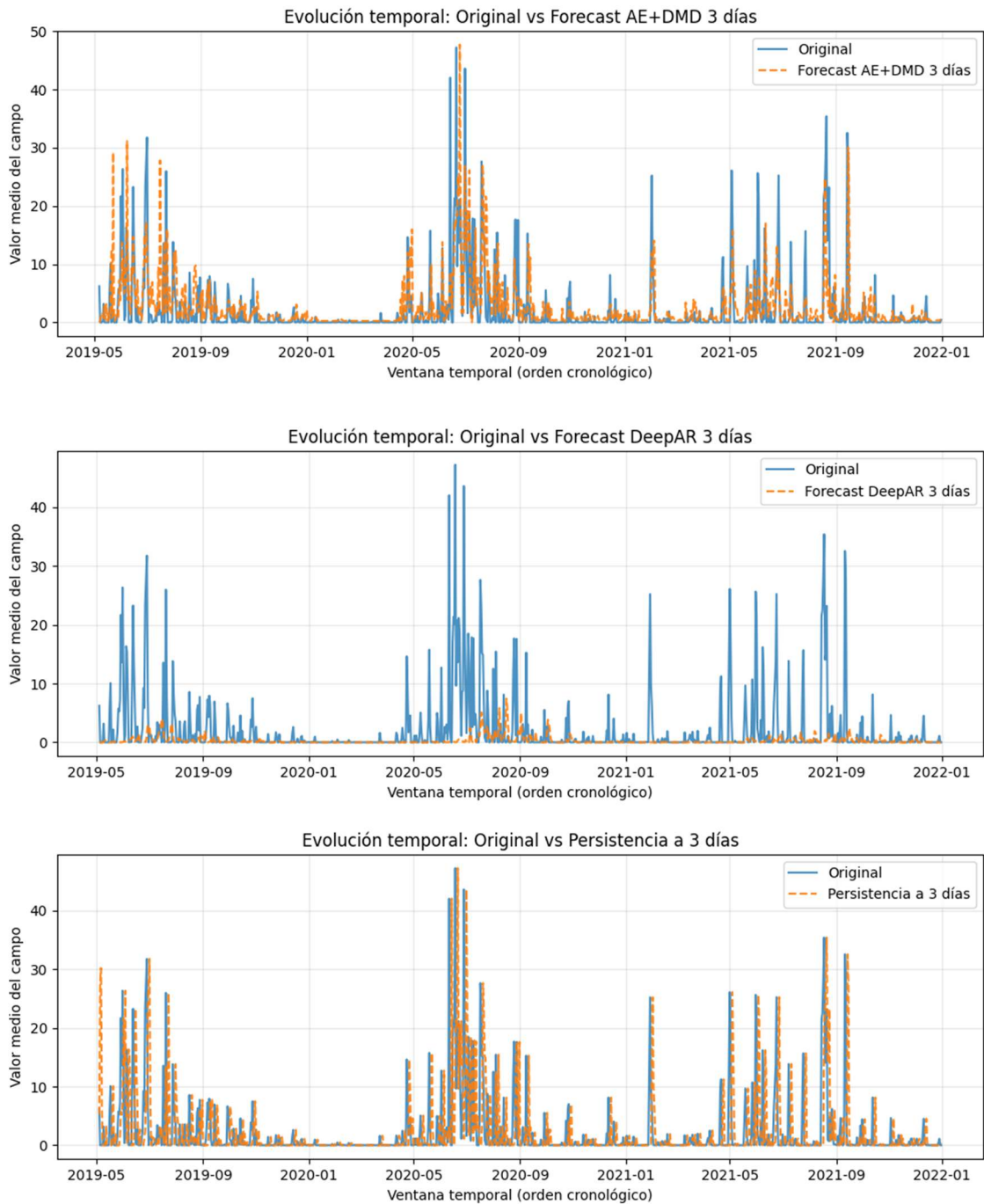
**Figura 5.14**

*Forecast del día del promedio espacial de precipitaciones en el tiempo.*



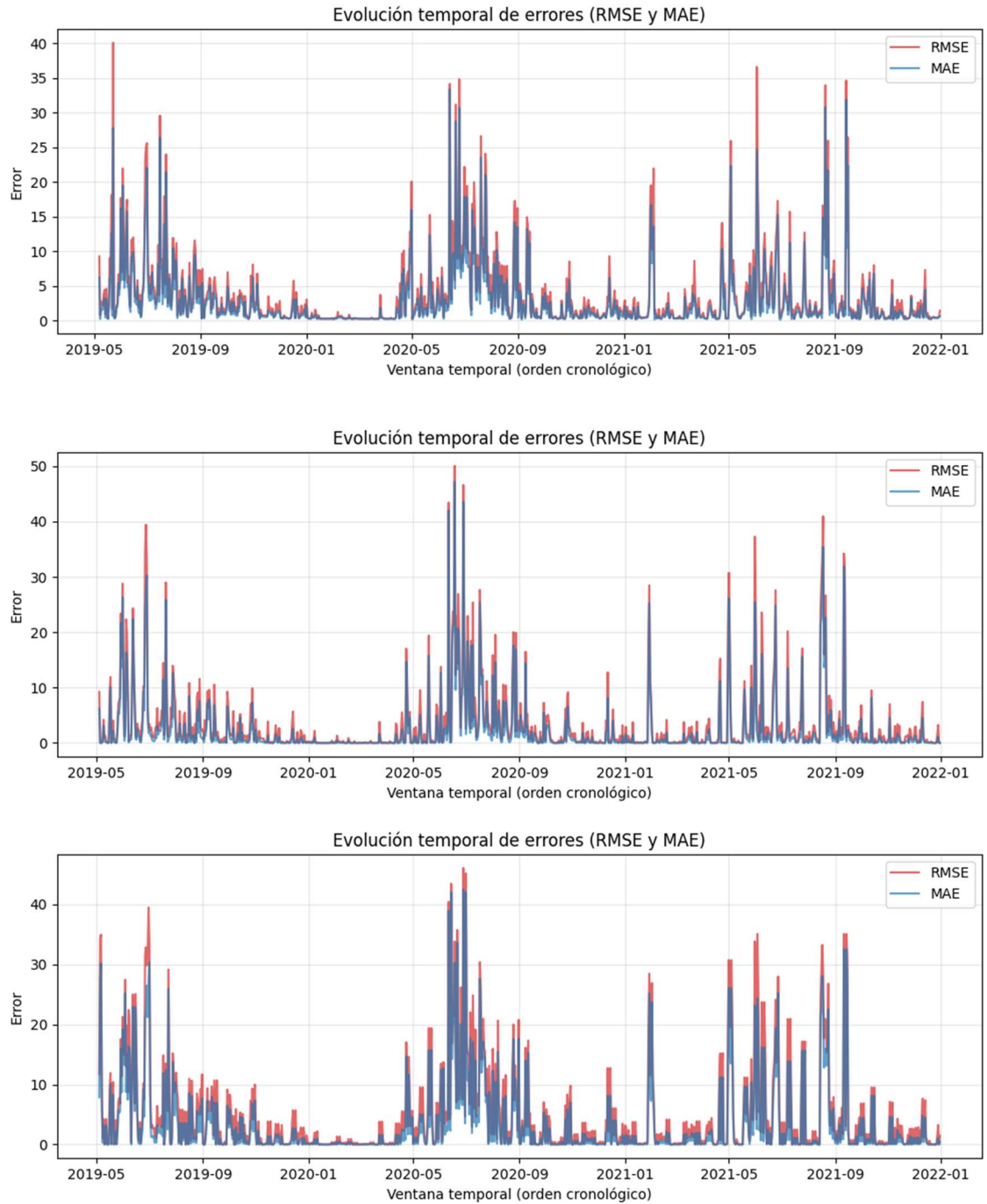
**Figura 5.15**

*Forecast de 3 días del promedio espacial de precipitaciones en el tiempo.*



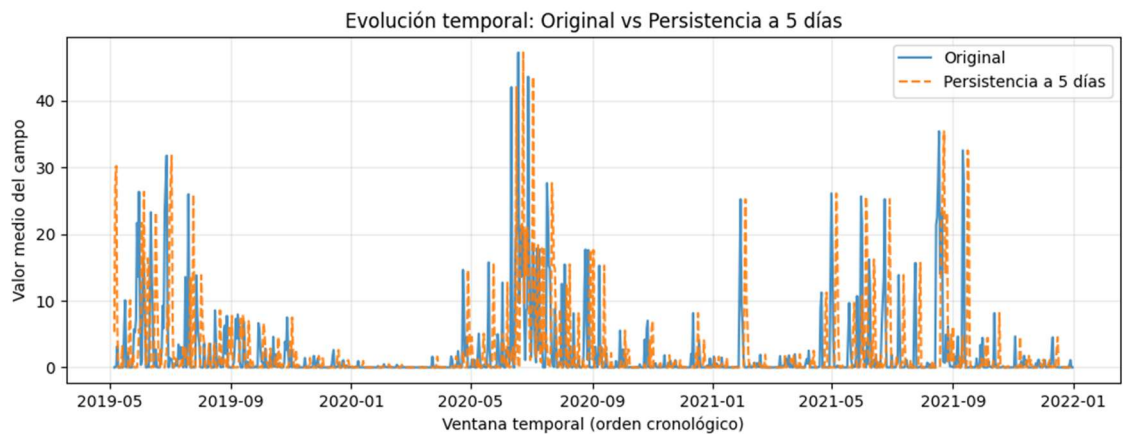
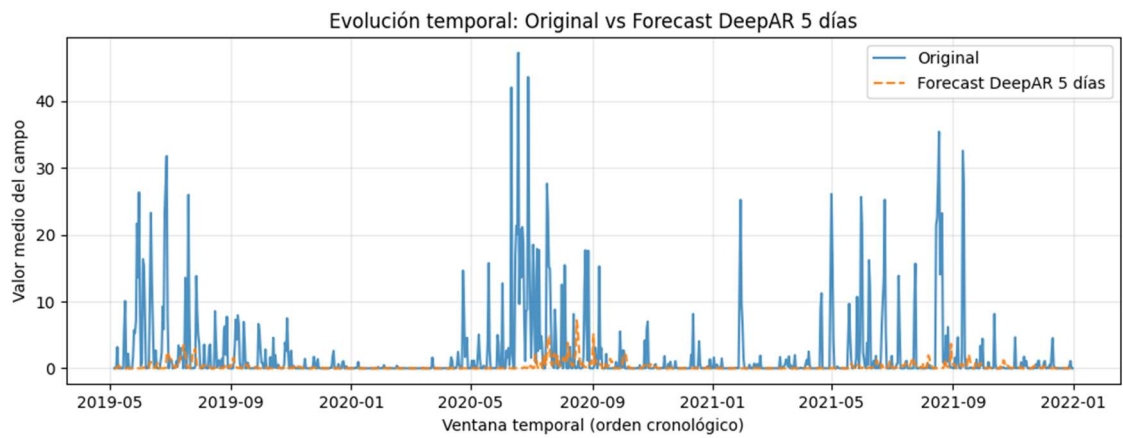
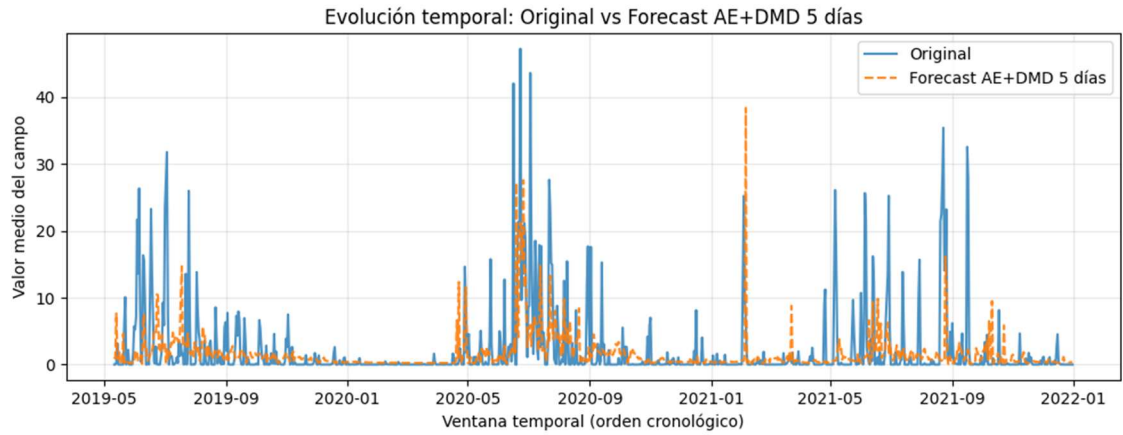
**Figura 5.16**

*Forecast de 3 días del promedio espacial de precipitaciones en el tiempo.*



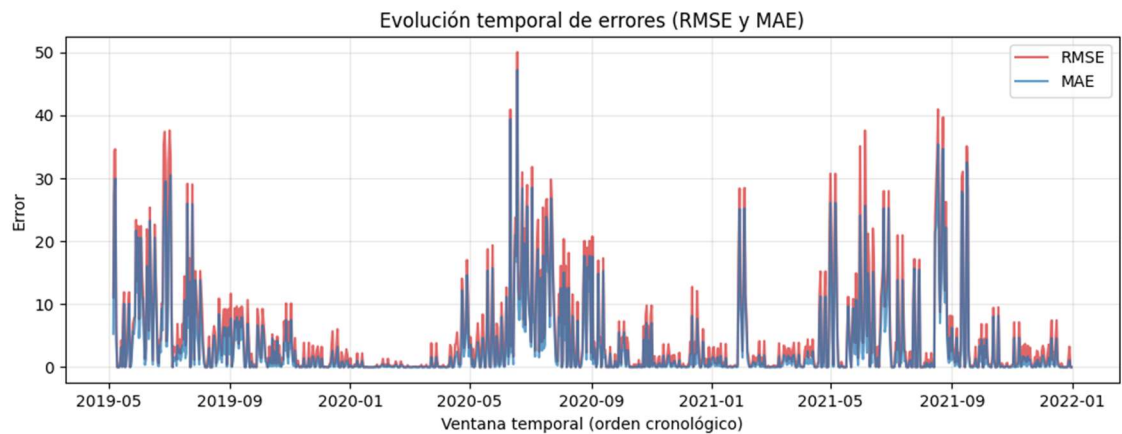
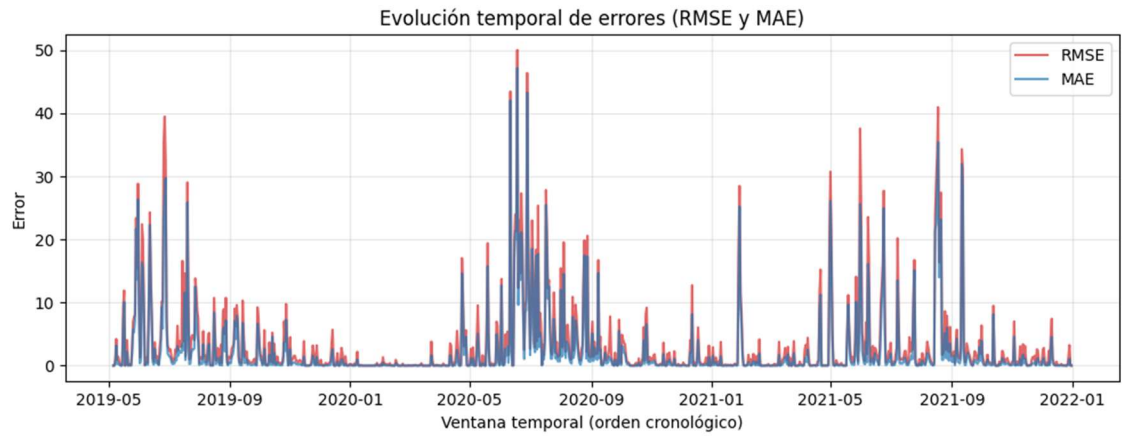
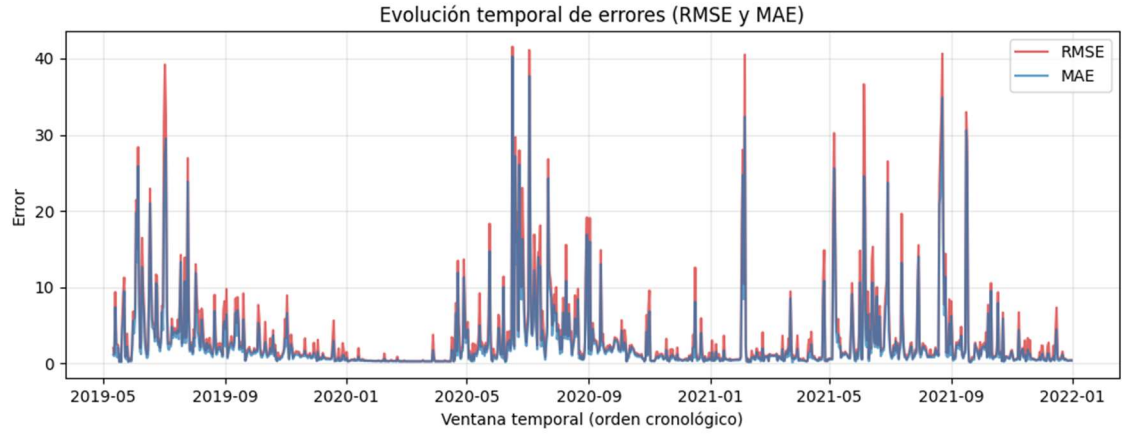
**Figura 5.17**

*Forecast de 5 días del promedio espacial de precipitaciones en el tiempo.*



**Figura 5.18**

*Forecast de 5 días del promedio espacial de precipitaciones en el tiempo.*



### **5.3.2. Comparación espacial del promedio temporal y su error (AE+DMD y DeepAR)**

La Figura 5.19 compara la distribución espacial del error para el horizonte de 1 día. Ambos modelos muestran errores más elevados en el sector cordillerano y precordillerano, donde la precipitación es más intensa y variable. Sin embargo, el mapa de errores revela que AE+DMD presenta menores errores que DeepAR en buena parte de la franja cordillerana, mientras que en el valle central y zonas costeras las diferencias son pequeñas o ligeramente favorables a DeepAR.

Esto sugiere que la representación espacio-temporal de AE+DMD aporta especialmente en regiones donde el gradiente costa-cordillera es fuerte, mientras que en áreas más homogéneas el beneficio respecto a un enfoque de series agregadas es limitado.

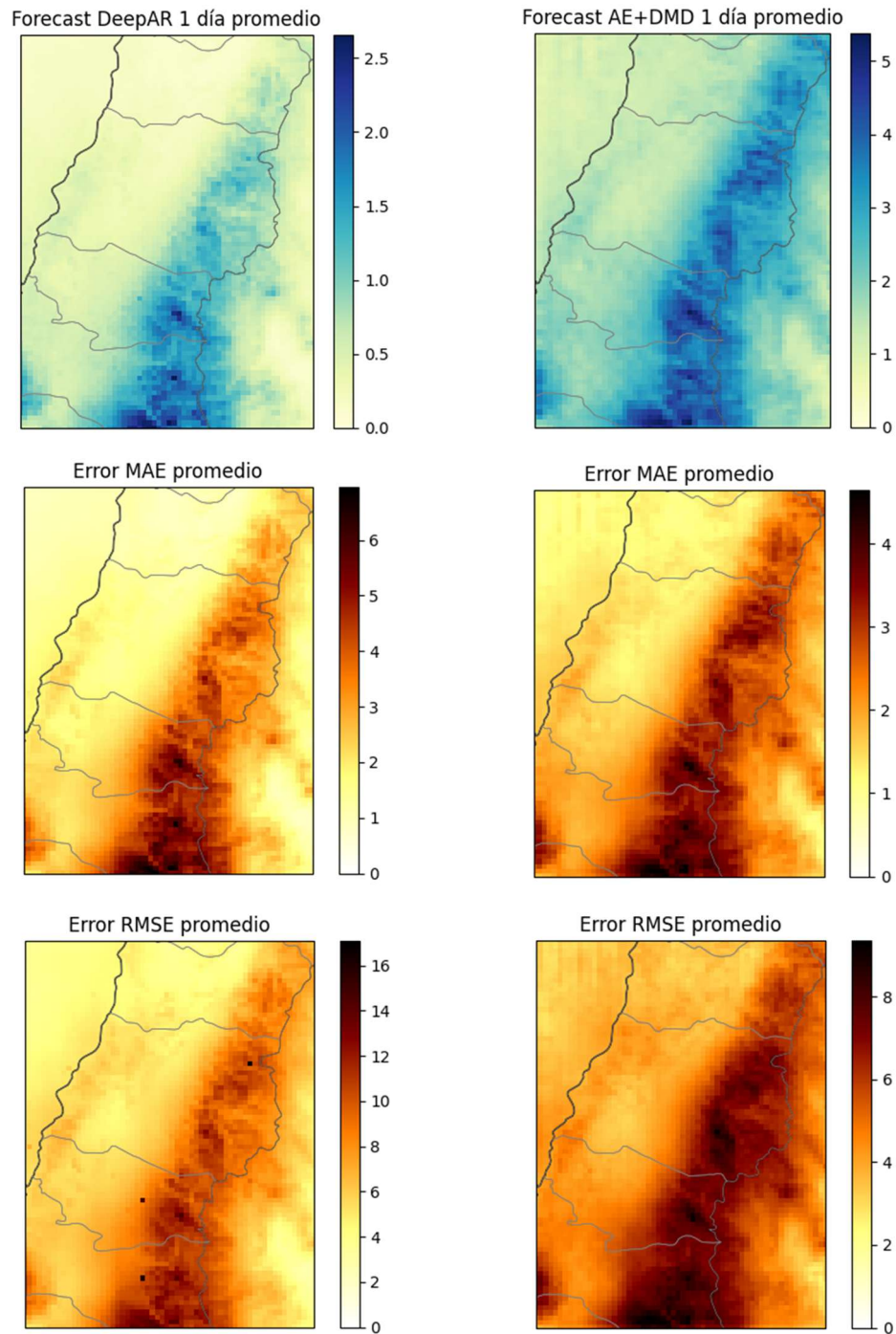
Para  $h=3$  y  $h=5$  días se observa un patrón espacial similar, aunque con valores de RMSE más altos en todos los puntos de la malla (Figuras 5.20 y 5.21). AE+DMD mantiene una ventaja clara en buena parte de la cordillera, pero la diferencia con DeepAR se reduce en sectores interiores y costeros.

#### **Nota general para Figuras 5.19 - 5.21**

*Nota:* Izquierda, promedio de predicción DeepAR en el espacio y sus errores. Derecha, promedio de predicción AE+DMD en el espacio y sus errores.

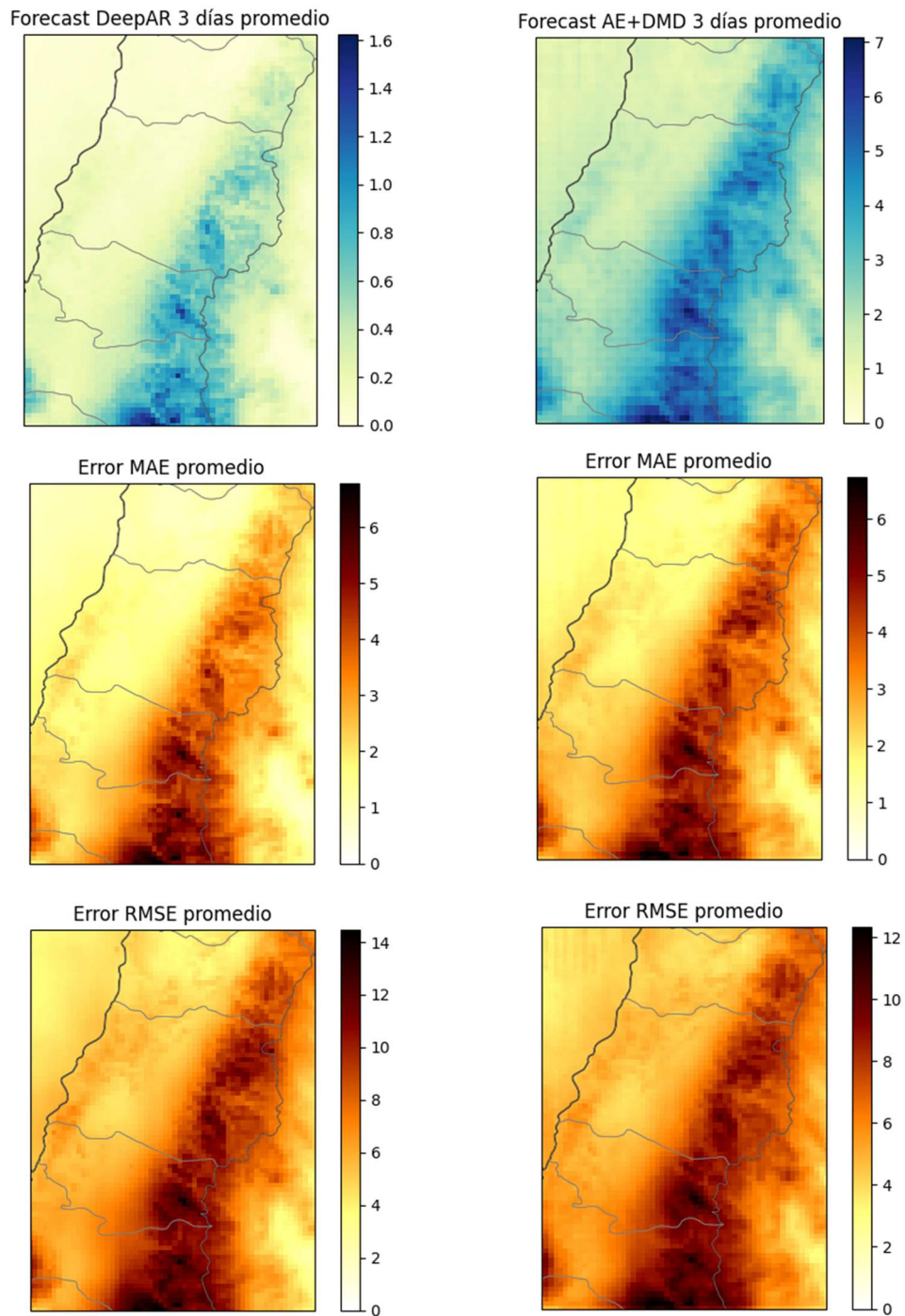
**Figura 5.19**

*Comparativo forecast de 1 día del promedio temporal de precipitaciones en el espacio.*



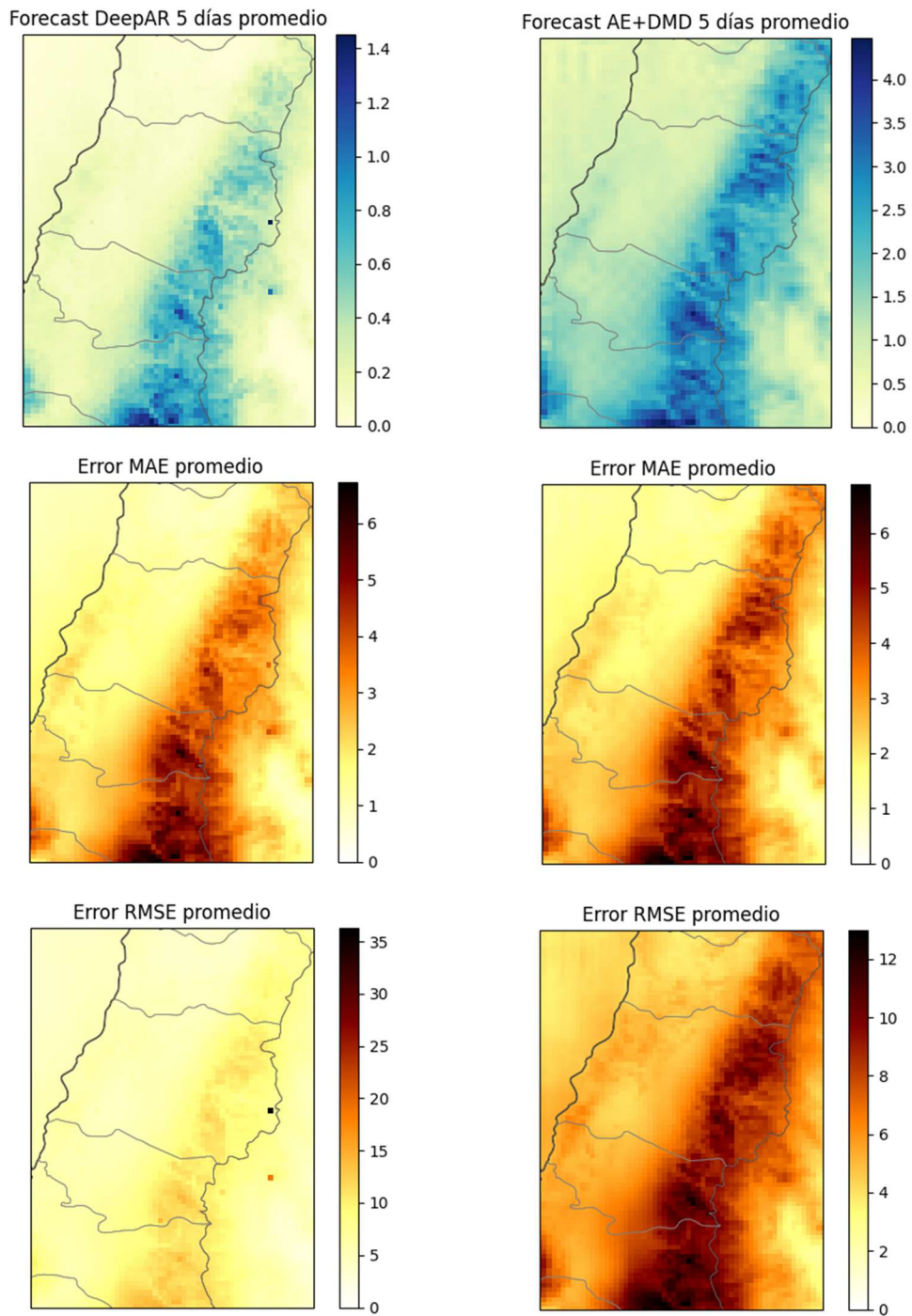
**Figura 5.20**

*Comparativo forecast de 3 días del promedio temporal de precipitaciones en el espacio.*



**Figura 5.21**

*Comparativo forecast de 5 días del promedio temporal de precipitaciones en el espacio.*



### **5.3.3. Evaluación por intensidad: detección de eventos, error en condiciones húmedas y habilidad espacial**

Las métricas globales (MAE/RMSE) entregan una visión promedio del desempeño, pero en precipitación diaria pueden sub-representar el comportamiento frente a eventos debido a la alta intermitencia y a que los periodos secos dominan el conteo. Por esta razón, se evalúa el desempeño por intensidad mediante tres enfoques complementarios:

- métricas de detección basadas en tabla de contingencia,
- error condicionado a precipitación (wet), y
- habilidad espacial usando el Fractions Skill Score (FSS).

#### **Configuración de evaluación**

Se consideran horizontes representativos  $h = \{1,3,5\}$  y umbrales de intensidad  $u = \{1,10\}$  mm/día, donde  $u = 1$  representa precipitación moderada y  $u = 10$  eventos más intensos. Para cada umbral, el campo observado  $X$  y el pronosticado  $\hat{X}$  se binarizan por píxel mediante  $1[X \geq u]$  y  $1[\hat{X} \geq u]$ , permitiendo calcular POD, FAR, CSI y BIAS sobre el dominio.

Para la evaluación wet se utiliza  $u = 1$  y se reportan MAE\_wet y RMSE\_wet, calculadas únicamente sobre celdas donde el observado supera el umbral, con el fin de medir la magnitud del error cuando efectivamente ocurre precipitación.

Finalmente, la coherencia espacial se evalúa con FSS usando ventanas  $3 \times 3$ ,  $5 \times 5$ ,  $9 \times 9$  y  $17 \times 17$ , reportando resultados para  $h = \{3, 5\}$  y  $u = \{1, 10\}$ .

### **Detección de eventos**

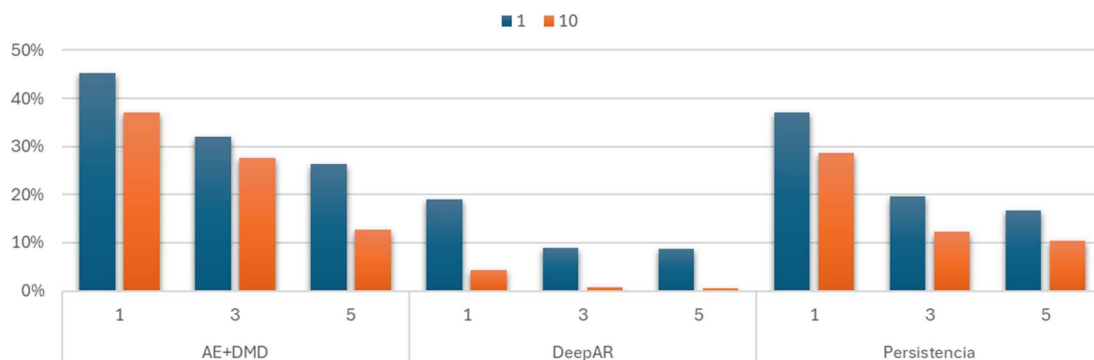
La Figura 5.22 muestra que AE+DMD obtiene la mayor habilidad de detección en ambos umbrales y horizontes, con la degradación esperable a medida que aumenta  $h$ . Por ejemplo, para  $u = 1$  se observa  $CSI \approx 0.45$  en  $h = 1$ , disminuye en  $CSI \approx 0.26$  para  $h = 5$ . En contraste, DeepAR presenta valores sensiblemente menores ( $CSI \approx 0.19$  en  $h = 1$  y  $CSI \approx 0.09$  en  $h = 5$  para  $u = 1$ ) y un colapso claro para  $u = 10$  ( $CSI \approx 0.04$  en  $h = 1$  y  $CSI$  cercano a 0.006 en  $h = 5$ ). El baseline de Persistencia mantiene un desempeño competitivo a corto plazo ( $CSI \approx 0.37$  en  $h = 1$  para  $u = 1$ ), pero cae con el horizonte ( $CSI \approx 0.17$  en  $h = 5$  para  $u = 1$ ), lo que es consistente con la limitada predictibilidad de la precipitación diaria a varios días.

Para interpretar estas diferencias se utiliza BIAS (Figura 5.23), que cuantifica el sesgo en la frecuencia de eventos ( $BIAS > 1$ : sobrepredicción;  $BIAS < 1$ : subpredicción). Persistencia mantiene BIAS cercano a 1 en ambos umbrales, como era esperable al copiar el último estado. En cambio, AE+DMD sobrepredice la ocurrencia para  $u = 1$  ( $BIAS \approx 1.56 - 2.05$ ), mientras que para  $u = 10$  tiende a subestimar eventos intensos ( $BIAS \approx 0.79$  en  $h = 1$ , llegando a  $BIAS \approx 0.31$  en  $h = 5$ ). DeepAR muestra subpredicción sistemática ( $BIAS < 1$ ) ya en  $u = 1$  ( $BIAS \approx 0.65$  en  $h = 1$ ) y se vuelve extrema en  $u =$

10 ( $BIAS \approx 0.17$  en  $h = 1$  y  $BIAS \approx 0.035$  en  $h = 5$ ), lo que ayuda a explicar su baja habilidad de detección para eventos intensos.

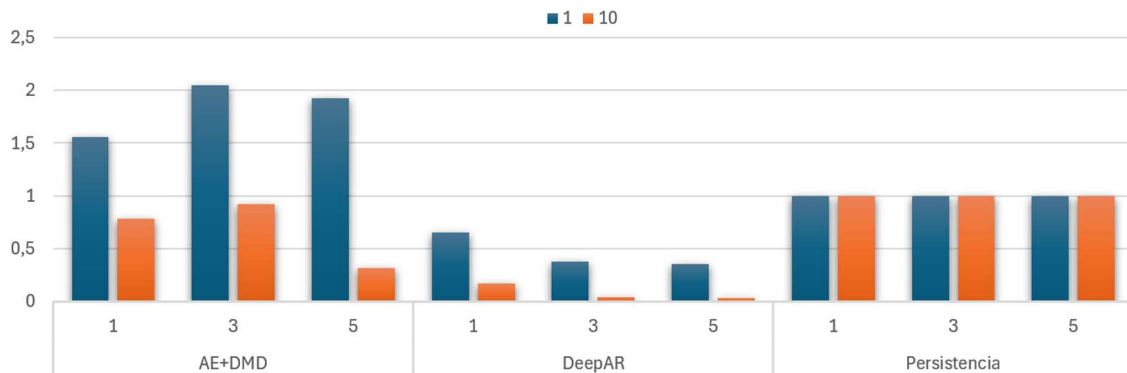
**Figura 5.22**

*Comparación de CSI para cada modelo por horizonte de predicción y umbral de precipitaciones (1mm en azul y 10mm en naranja).*



**Figura 5.23**

*Comparación de BIAS para cada modelo por horizonte de predicción y umbral de precipitaciones (1mm en azul y 10mm en naranja).*



La Tabla 5.4 refuerza el diagnóstico. Para  $u = 10$  y  $h = 1$ , AE+DMD logra  $POD \approx 48.3\%$  con  $FAR \approx 38.5\%$ , mientras que DeepAR obtiene  $POD \approx 4.8\%$  y  $FAR \approx 71.8\%$ , es decir, detecta muy pocos eventos intensos y además genera una fracción elevada de falsas alarmas. Persistencia mantiene un nivel competitivo en  $h = 1$  con  $POD \approx 4.8\%$ , pero con FAR más alto ( $FAR \approx 55.5\%$ ) y una degradación marcada con el horizonte. En conjunto, CSI, BIAS y POD/FAR indican que AE+DMD es el modelo con mejor capacidad de detección, mientras que DeepAR presenta una limitación importante al representar la ocurrencia de eventos intensos.

**Tabla 5.4**

*POD y FAR por modelo para horizontes de 1, 3 y 5 días.*

Modelo	Forecast 1 día		Forecast 3 días		Forecast 5 días	
	POD	FAR	POD	FAR	POD	FAR
AE+DMD	48,32%	38,47%	41,60%	54,66%	14,81%	52,65%
DeepAR	4,82%	71,81%	0,78%	81,76%	0,58%	83,52%
Persistencia	44,46%	55,54%	21,85%	78,14%	18,85%	81,14%

*Nota.* POD y FAR se calcularon para detección de precipitación diaria superior a 10 mm/día, sobre el conjunto de prueba.  $POD = \frac{TP}{TP+F}$  y  $FAR = \frac{FP}{TP+F}$ , donde TP, FP y FN corresponden a verdaderos positivos, falsos positivos y falsos negativos. Valores de POD más altos y FAR más bajos indican mejor desempeño.

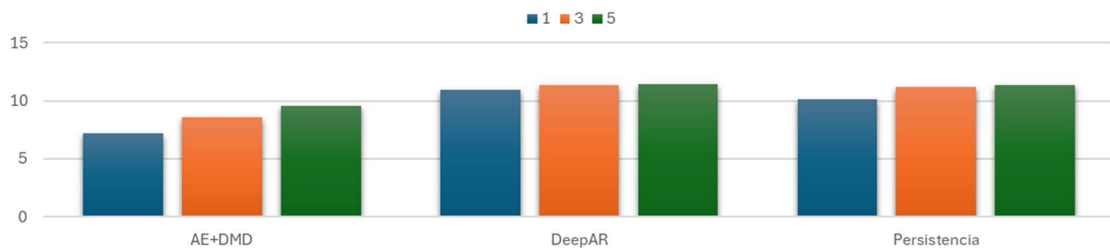
### Error condicionado a precipitación

Las Figuras 5.24 y 5.25 muestran el error únicamente en celdas húmedas ( $X \geq 1$  mm/día), evitando que los ceros dominen la evaluación. Bajo esta condición, AE+DMD presenta consistentemente menor error que DeepAR y Persistencia en los tres horizontes. Por ejemplo, en  $h = 1$  se observa  $MAE_{wet} \approx 7,21$ ,  $RMSE_{wet} \approx 10,80$  para AE+DMD, frente a valores cercanos a  $MAE_{wet} \approx 10,94$ ,  $RMSE_{wet} \approx 16,05$  en DeepAR y  $MAE_{wet} \approx 10,09$ ,  $RMSE_{wet} \approx 14,67$  en Persistencia. A  $h = 5$ , AE+DMD mantiene la ventaja ( $MAE_{wet} \approx 9,60$ ;  $RMSE_{wet} \approx 14,88$ ) frente a DeepAR ( $MAE_{wet} \approx 11,41$ ;  $RMSE_{wet} \approx 16,51$ ) y Persistencia ( $MAE_{wet} \approx 11,36$ ;  $RMSE_{wet} \approx 16,63$ ).

Esto sugiere que la mejora de AE+DMD no se debe solo a predecir periodos secos, sino también a reducir error cuando efectivamente llueve, que es donde el problema es más relevante.

**Figura 5.24**

*Comparación de MAE para cada modelo por horizonte de predicción (1 día en azul, 3 días en naranja y 5 días en verde) para umbral de intensidad de 1 mm/día.*



**Figura 5.25**

*Comparación de RMSE para cada modelo por horizonte de predicción (1 día en azul, 3 días en naranja y 5 días en verde) para umbral de intensidad de 1 mm/día.*



### **Habilidad espacial (FSS)**

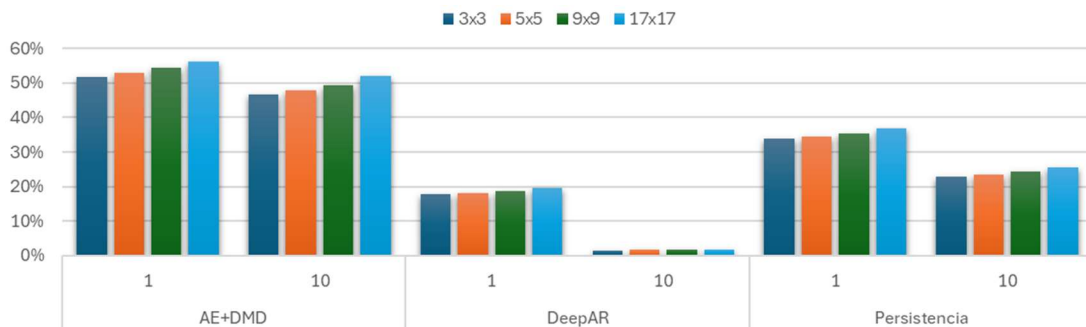
Las Figuras 5.26 y 5.27 reportan FSS en función del tamaño de ventana. En los tres modelos se observa el patrón esperado: el FSS aumenta al crecer la ventana (mayor tolerancia a desplazamientos espaciales), por lo que esta métrica es útil para evaluar si el modelo reproduce correctamente la estructura espacial, aunque falle la localización exacta píxel a píxel.

Para  $u = 1$ , AE+DMD obtiene la mayor habilidad espacial en todas las escalas. En el caso de  $h = 3$  y ventana  $17 \times 17$ , AE+DMD alcanza  $FSS \approx 0,56$ , frente a Persistencia con  $FSS \approx 0,37$  y DeepAR  $FSS \approx 0,20$ . Para  $u = 10$ , las diferencias se vuelven más marcadas: DeepAR muestra valores muy bajos incluso en ventanas grandes ( $FSS \approx 0,17$  en  $h = 3$  y  $FSS \approx 0,012$  en  $h = 5$ , ambos en ventana  $17 \times 17$ ), lo que indica escasa capacidad para reproducir la estructura espacial asociada a eventos intensos. En contraste,

AE+DMD mantiene habilidad espacial apreciable (para  $h = 5$ ,  $u = 10$ , ventana  $17 \times 17$ ),  $FSS \approx 0,26$  y supera a Persistencia  $FSS \approx 0,22$ .

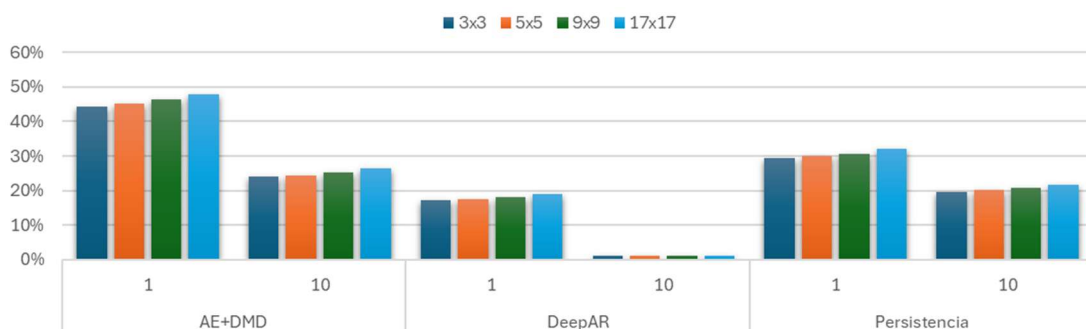
**Figura 5.26**

*Comparación de FSS para cada modelo por grilla (3x3 en azul, 5x5 en naranja, 9x9 en verde y 17x17 en celeste) para horizonte de predicción de 3 días.*



**Figura 5.27**

*Comparación de FSS para cada modelo por grilla (3x3 en azul, 5x5 en naranja, 9x9 en verde y 17x17 en celeste) para horizonte de predicción de 5 días.*



La evaluación por intensidad confirma que la precipitación diaria requiere métricas orientadas a eventos para caracterizar desempeño real. En detección ( $CSI$ ), AE+DMD muestra la mayor habilidad a  $u = 1$  y  $u = 10$ , mientras que DeepAR presenta baja habilidad, especialmente para eventos intensos, coherente con su fuerte subpredicción ( $BIAS \ll 1$ ) y su combinación de  $POD$  bajo y  $FAR$  alto en  $u = 10$ .

En condiciones húmedas, AE+DMD reduce significativamente  $MAE_{wet}$  y  $RMSE_{wet}$  en comparación con DeepAR y Persistencia. Finalmente, el análisis  $FSS$  evidencia que AE+DMD preserva mejor la coherencia espacial, con ventajas más claras al aumentar el umbral y al evaluar escalas espaciales mayores.

#### **5.3.4. Casos de estudio de eventos de precipitación**

Las Figuras 5.28 y 5.29 muestran la evolución del promedio espacial de precipitación y los campos medios para un evento con lluvias  $>20$  mm/día, comparando AE+DMD y DeepAR a 1 día y 3 días de horizonte.

- Horizonte 1 día:

AE+DMD: Sigue razonablemente la forma del evento, levanta la señal desde el 16-08, alcanza un máximo cercano al observado y luego decae.

Hay algo de suavizado y pequeño desfase, pero el pulso principal aparece y con magnitud comparable. En el mapa medio mantiene la banda lluviosa en la cordillera, aunque con intensidades algo menores.

DeepAR: La serie casi no reacciona: se mantiene muy baja durante toda la semana.

El campo medio aparece prácticamente seco. En la práctica, el evento intenso queda casi “invisible” para el modelo.

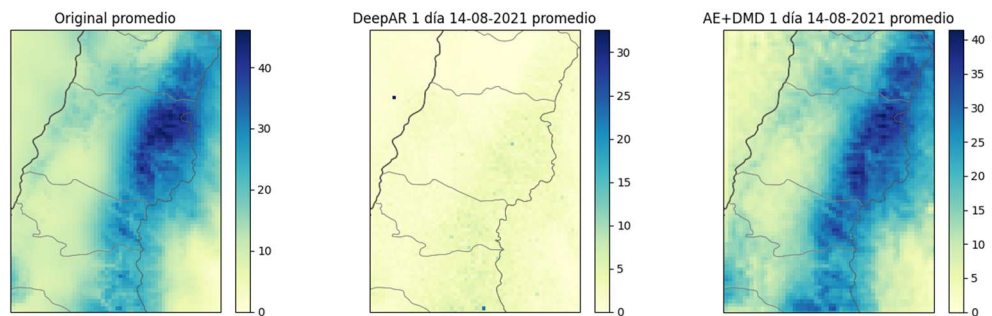
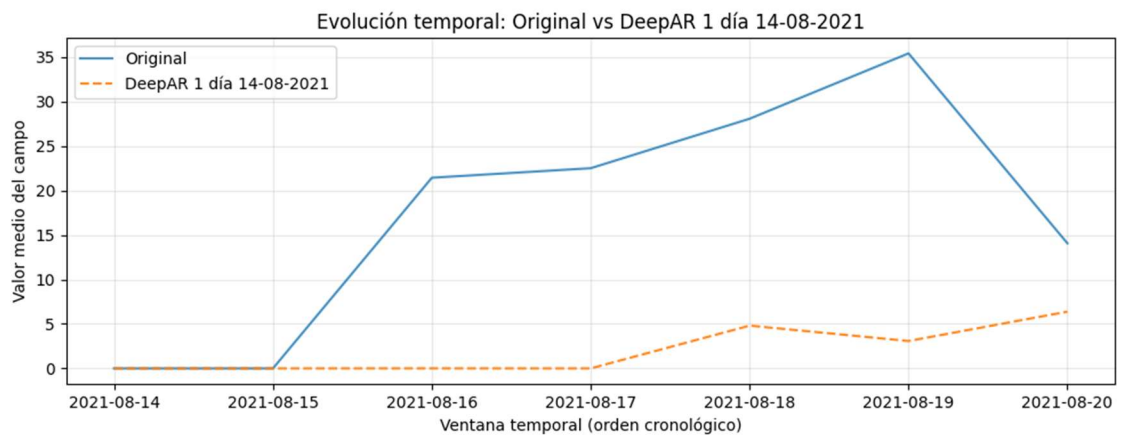
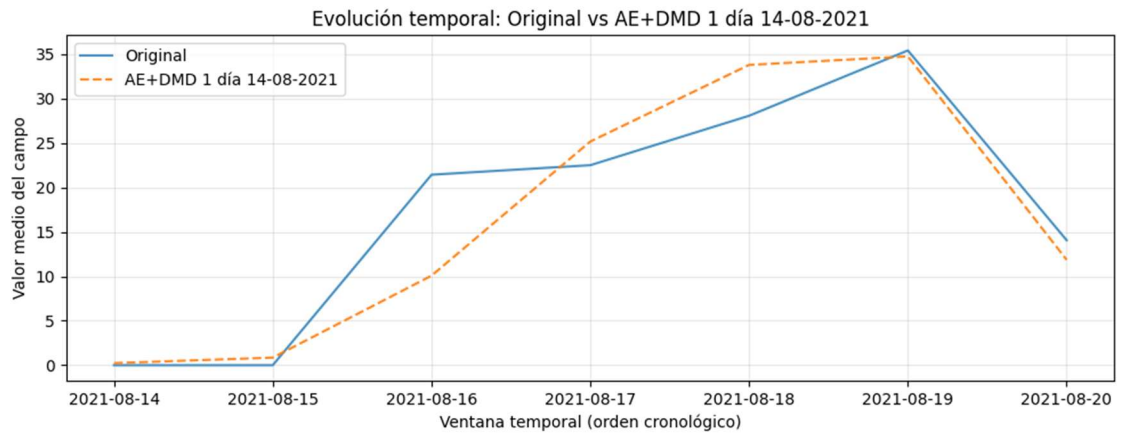
- Horizonte 3 días:

AE+DMD: Detecta que viene un episodio lluvioso, pero lo reparte en el tiempo y subestimando el máximo. El mapa medio conserva el gradiente costa–cordillera, pero con valores mucho menores que los observados.

DeepAR: La serie sigue pegada a valores muy bajos. El campo medio apenas muestra señal de precipitación, el evento no se representa.

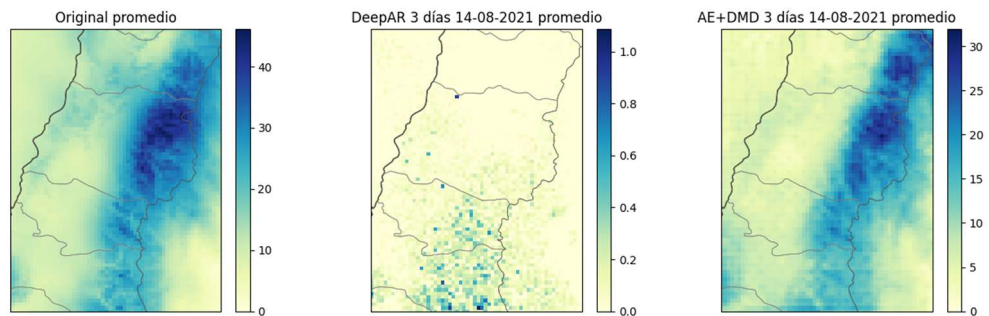
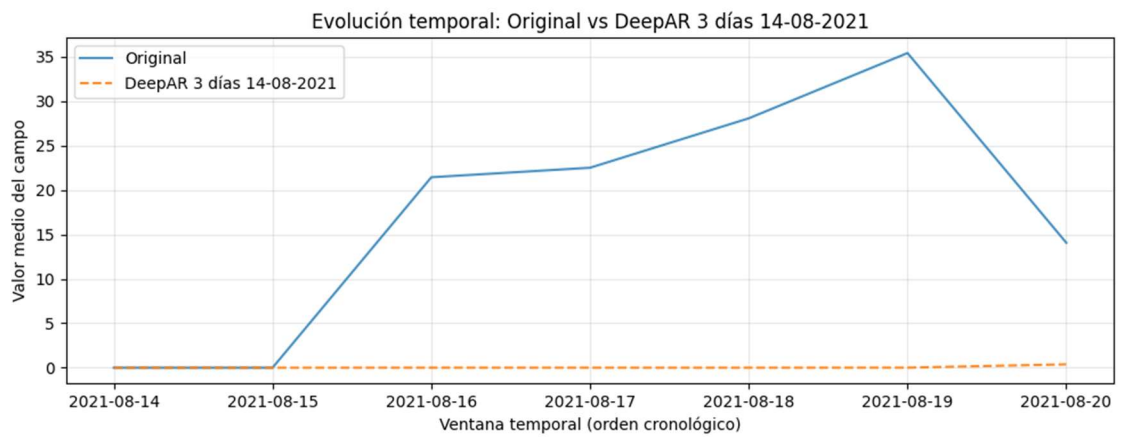
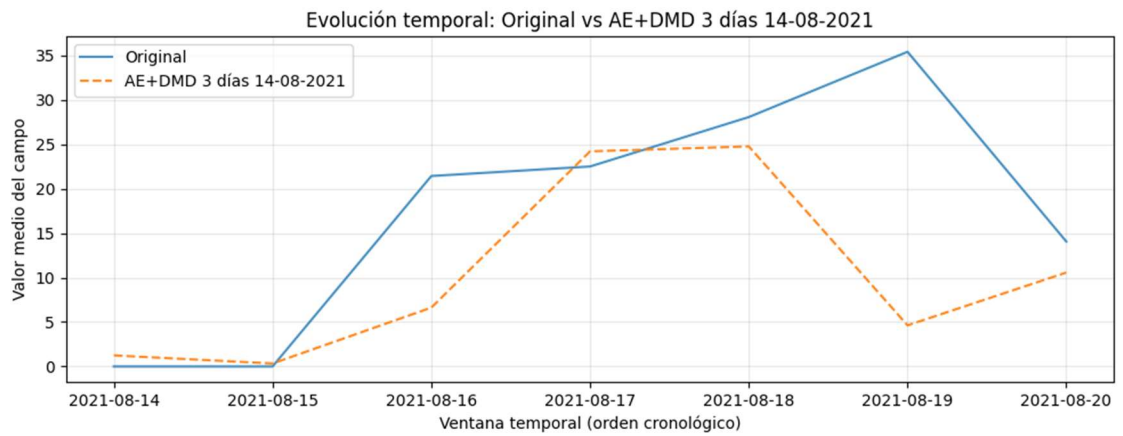
**Figura 5.28**

*Comparación de predicciones a 1 día de ambos modelos para la semana del 14-08-2021.*



**Figura 5.29**

*Comparación de predicciones a 3 días de ambos modelos para la semana del 14-08-2021.*



Las Figuras 5.30 y 5.31 muestran la evolución del promedio espacial de precipitación y los campos medios para un evento con lluvias  $<5$  mm/día, comparando AE+DMD y DeepAR a 1 día y 3 días de horizonte.

- Horizonte 1 día:

AE+DMD: La serie pronosticada se mantiene por encima de la observada casi toda la semana, sobreestima levemente la precipitación y la hace más persistente. El campo medio muestra una banda lluviosa más amplia y húmeda que la observada, extendiendo el núcleo real a lo largo de la cordillera.

DeepAR: Presenta un pico marcado el 24-09-2021 superior al observado y valores positivos cuando no llueve. En el mapa medio aparece señal de lluvia en zonas donde el campo observado es prácticamente seco.

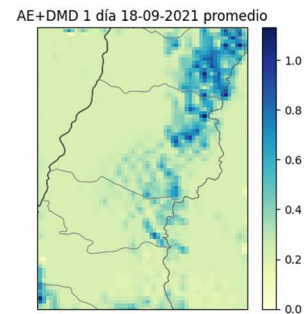
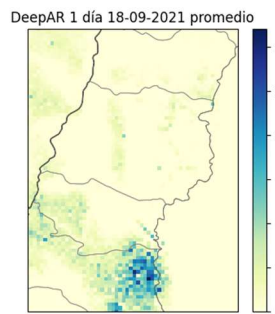
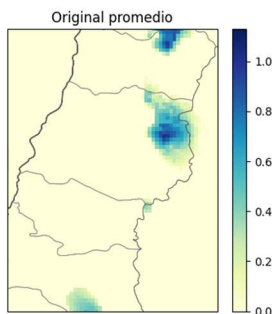
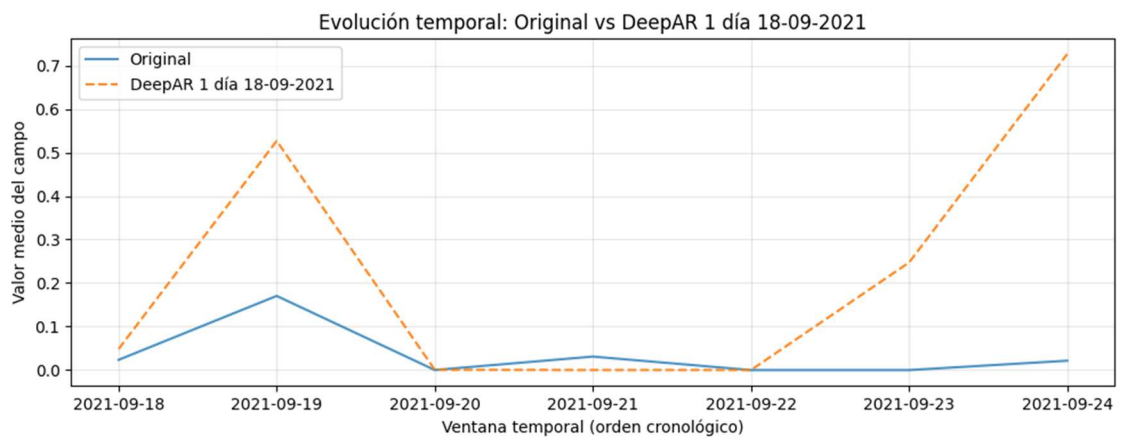
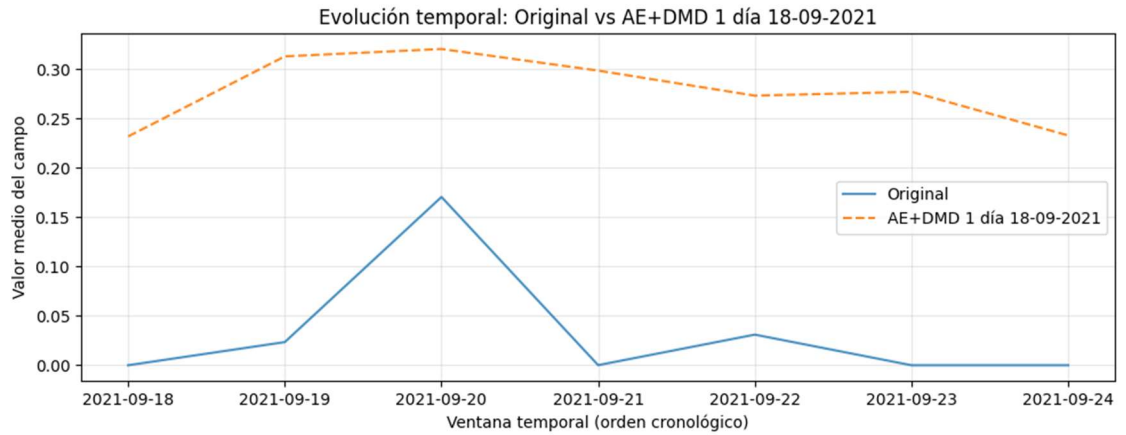
- Horizonte 3 días:

AE+DMD: Amplifica el sesgo húmedo, la serie pronosticada multiplica la magnitud de la lluvia observada y mantiene valores elevados varios días seguidos. El mapa medio muestra una franja lluviosa continua a lo largo de la cordillera, por encima del patrón real, que es más acotado y de menor intensidad.

DeepAR: También sobreestima la precipitación, con valores medios superiores a los observados y una estructura espacial demasiado extendida.

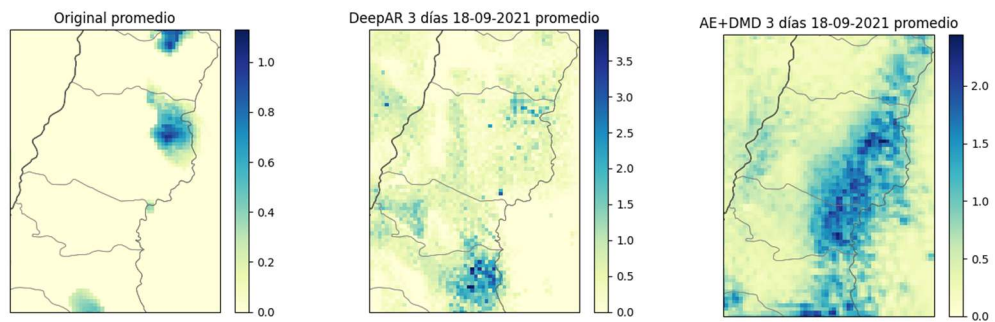
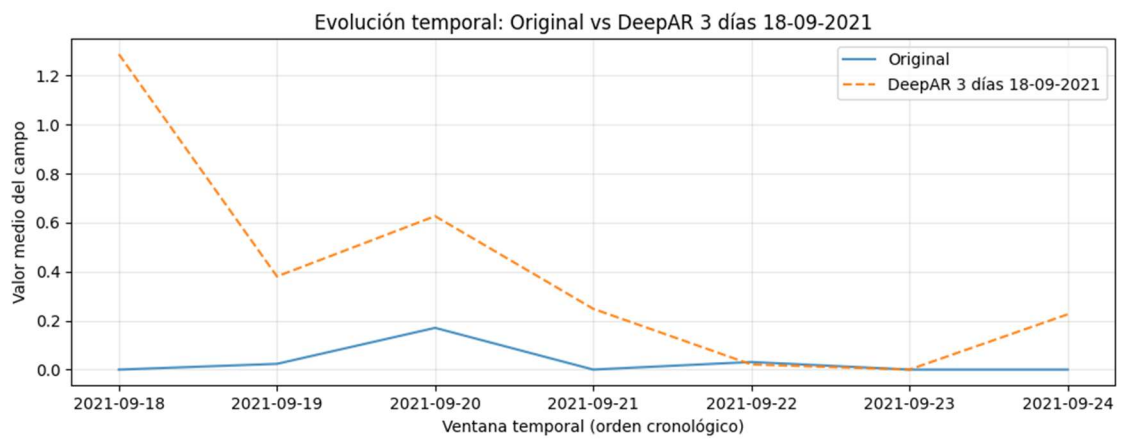
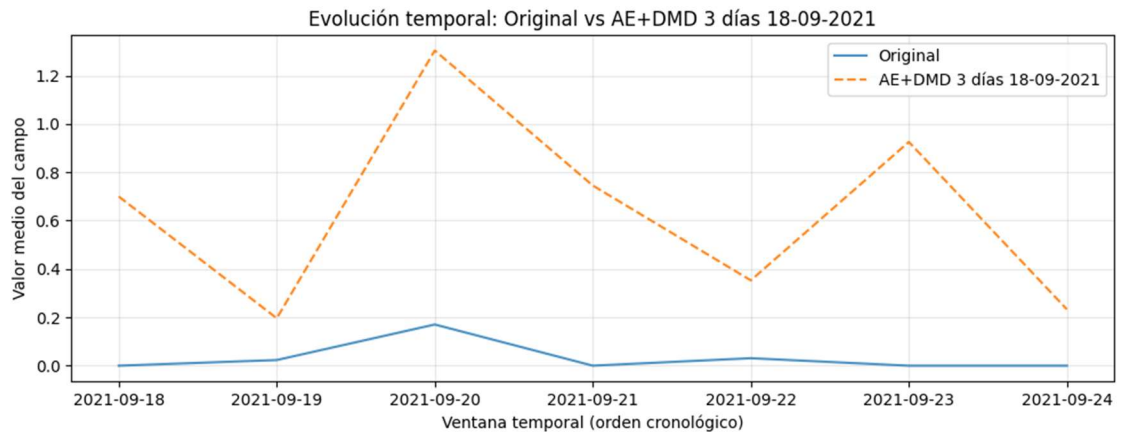
**Figura 5.30**

*Comparación de predicciones a 1 día de ambos modelos para la semana del 18-09-2021.*



**Figura 5.31**

*Comparación de predicciones a 3 días de ambos modelos para la semana del 18-09-2021.*



### 5.3.5. Costo computacional (entrenamiento e inferencia)

Además del desempeño predictivo, se cuantificó el costo computacional de cada enfoque, considerando tanto el tiempo de entrenamiento como el tiempo de inferencia sobre el conjunto de prueba (Tabla 5.5). Para asegurar comparabilidad, los tiempos corresponden a mediciones wall-clock realizadas bajo el mismo entorno de ejecución y utilizando el mismo pipeline de evaluación. La inferencia se reporta como tiempo total y como tiempo promedio por unidad evaluada (muestra/ventana) del set de prueba.

**Tabla 5.5**

*Costos computacionales de entrenamiento e inferencia.*

<b>Modelo</b>	<b>#epochs</b>	<b>Train total (min)</b>	<b>Train medio (s/epoch)</b>	<b>Inferencia total (s)</b>	<b>Inferencia media (s/día)</b>
AE+DMD	60	532,3	532,3	292,4	0,3
DeepAR	100	18,6	11,2	5905,2	6,1
Persistencia	-	-	-	0,3	< 0,1

En entrenamiento, AE+DMD es mucho más costoso que DeepAR (532,3 vs 18,6 min; ~28,6×). En cambio, en inferencia AE+DMD es considerablemente más rápido (0,3 vs 6,1 ms/unidad; ~20,3×) y también en tiempo total (292,5 vs 5905 s; ~20.1×). La persistencia

tiene un costo de inferencia prácticamente nulo, ya que corresponde a una operación determinística simple.

## 5.4. Análisis espectral y modos DMD

Esta sección analiza el espectro del operador dinámico estimado mediante Dynamic Mode Decomposition (DMD) en el espacio latente del modelo híbrido AE+DMD. El objetivo es doble: (i) caracterizar la dinámica dominante del sistema en términos de estabilidad y escalas temporales, y (ii) evaluar si emergen patrones espaciales recurrentes —en particular asociados a la zona cordillerana— más allá de un caso puntual.

Para cada ventana temporal del conjunto de prueba se estima un operador lineal  $A$  que aproxima la evolución del estado latente, y se obtiene su descomposición espectral.

Dado un autovalor discreto  $\lambda_k$ , se define la tasa continua

$$u_k = \frac{\log(\lambda_k)}{\Delta t},$$

donde  $\Delta t = 1$  día. A partir de  $\lambda_k$  se derivan: (i) el módulo  $|\lambda_k|$ , indicador de estabilidad en tiempo discreto; (ii) la frecuencia

$$f_k = \frac{\angle(\lambda_k)}{2\pi\Delta t} \text{ [ciclos/día]},$$

y (iii) el crecimiento/decadencia  $g_k = \Re(u_k)$  (valores negativos indican decaimiento). En particular,  $|\lambda_k| < 1$  corresponde a modos que decaen en el tiempo,  $|\lambda_k| \approx 1$  a componentes persistentes y  $|\lambda_k| > 1$  a crecimiento.

### **Configuración del análisis**

El análisis se realiza sobre ventanas deslizantes de longitud  $T = 64$  días del conjunto de prueba (stride unitario), procesadas en batches de tamaño 8. Esta elección busca un balance práctico: ventanas más largas tienden a estabilizar la estimación espectral, mientras que ventanas demasiado extensas diluyen cambios estacionales y reducen el número efectivo de muestras independientes.

En cada ventana se estima el operador  $A$  en el espacio latente del autoencoder (según la implementación del modelo híbrido). Para priorizar modos relevantes y reducir sensibilidad a componentes de baja contribución, se restringe el análisis espacial detallado a los  $K$  modos con mayor amplitud relativa (por ejemplo,  $K = 12$ ). La amplitud relativa  $a_k$  se computa a partir de los coeficientes temporales estimados en el espacio latente y se utiliza como indicador de importancia de cada modo en la reconstrucción temporal de la ventana.

### **Espectro en el plano complejo y estabilidad por estación**

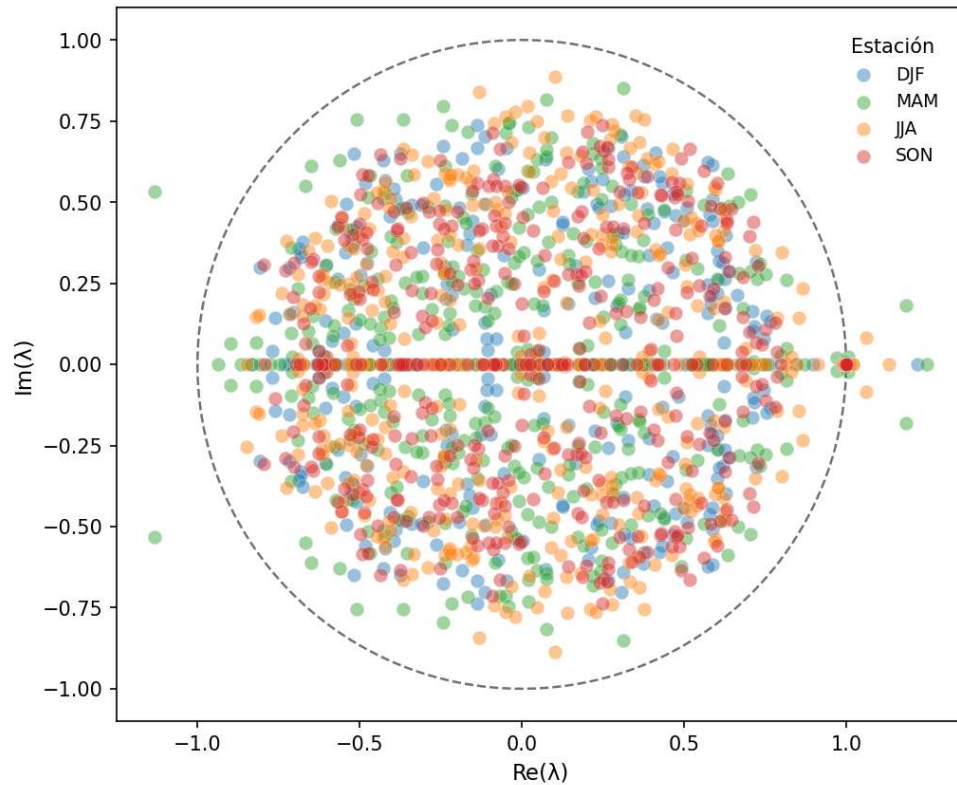
La Figura 5.32 muestra la distribución de autovalores  $\lambda_k$  en el plano complejo, separada por estación, donde cada estación está compuesta por tres meses agrupados:

- DJF: Verano
- MAM: Otoño
- JJA: Invierno
- SON: Primavera

El círculo unidad actúa como referencia directa de estabilidad: autovalores dentro del círculo corresponden a componentes que decaen en el tiempo. En conjunto, el espectro observado se concentra mayoritariamente dentro del círculo unidad, sugiriendo una dinámica global estable/decadente en la escala diaria, con componentes oscilatorias identificables mediante el ángulo de  $\lambda_k$ .

**Figura 5.32**

*Autovalores DMD (Top-K por batch) agrupados por estación.*



La Tabla 5.6 resume estadísticas del espectro por estación (número de ventanas evaluadas, mediana y dispersión de  $|\lambda_k|$ , frecuencia y tasa de crecimiento/decadencia, además del porcentaje de modos estables). En general, el porcentaje de estabilidad es alto en todas las estaciones, con valores de crecimiento predominantemente negativos, coherentes con la ausencia de modos sistemáticamente explosivos. Las frecuencias muestran una mezcla de componentes lentas (cerca de cero) y componentes oscilatorias, lo que sugiere la

coexistencia de variabilidad de baja frecuencia con oscilaciones de escala sinóptica dentro de las ventanas consideradas.

**Tabla 5.6**

*Resumen del espectro DMD por estación (n modos).*

season	días	$ \lambda_k $		growth		freq		% stable
		MED	IQR	MED	IQR	MED	IQR	
DJF	324	0,622	0,260	-0,475	0,455	0,017	0,458	99,07%
JJA	420	0,614	0,249	-0,489	0,428	0,000	0,500	98,10%
MAM	408	0,556	0,344	-0,587	0,662	0,000	0,502	97,55%
SON	408	0,557	0,320	-0,585	0,614	0,017	0,460	99,51%

### **Interpretación espacial de modos y definición de recurrencia regional**

Además de su caracterización espectral, los modos DMD proporcionan patrones espaciales interpretables. En este trabajo, el modo  $k$  se visualiza como un patrón espacial decodificado desde el espacio latente hacia el dominio espacial mediante el decoder del autoencoder. Importante: estos mapas representan patrones normalizados (unitless), no precipitación en mm/día, debido a que la escala de los autovectores de DMD es intrínsecamente arbitraria (el sistema es invariante ante reescalamiento de autovectores y coeficientes). Por ello, para visualización se normaliza el patrón y se emplea una escala cromática robusta y consistente dentro de cada figura.

Dado que una crítica común al análisis cualitativo de mapas es su dependencia de casos puntuales, se propone una medida cuantitativa para evaluar si un modo concentra su energía espacial sobre una región de interés (en este caso, la cordillera). Se define una máscara binaria  $M$  sobre el dominio espacial (Figura 5.33), construida de forma determinística a partir de dos líneas en coordenadas lon-lat que delimitan una banda asociada a la zona cordillerana dentro del subdominio de estudio. Sea  $w(i) \geq 0$  un peso espacial del patrón del modo (por ejemplo  $w(i) = |\text{campo}(i)|$ , donde *campo* es el patrón decodificado). Se define el score de concentración:

$$\text{score} = \frac{\sum_{i \in M} w(i)}{\sum_i w(i)},$$

y el área relativa de la máscara  $\text{area} = \frac{|M|}{N}$ , donde  $N$  es el número total de celdas.

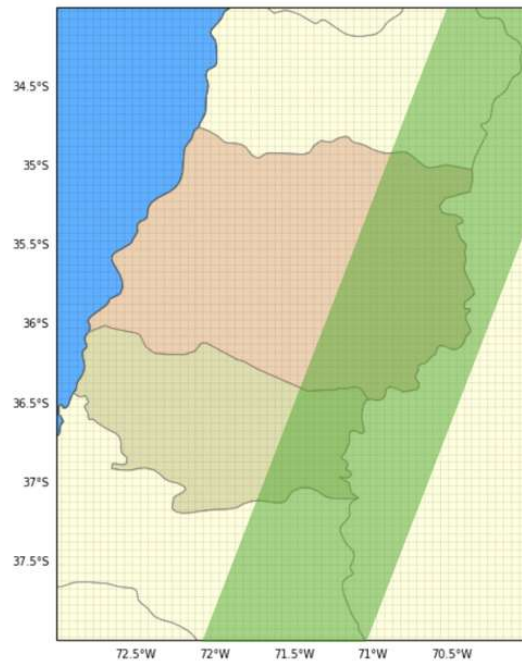
Finalmente, se define el índice de enriquecimiento:

$$\text{enrich} = \frac{\text{score}}{\text{area}}$$

Esta razón observada/esperada permite interpretar de manera directa la concentración espacial:  $\text{enrich} = 1$  corresponde a una distribución consistente con el tamaño geométrico de la región;  $\text{enrich} > 1$  indica concentración superior a la esperada por área; y  $\text{enrich} < 1$  sugiere evitación relativa.

**Figura 5.33**

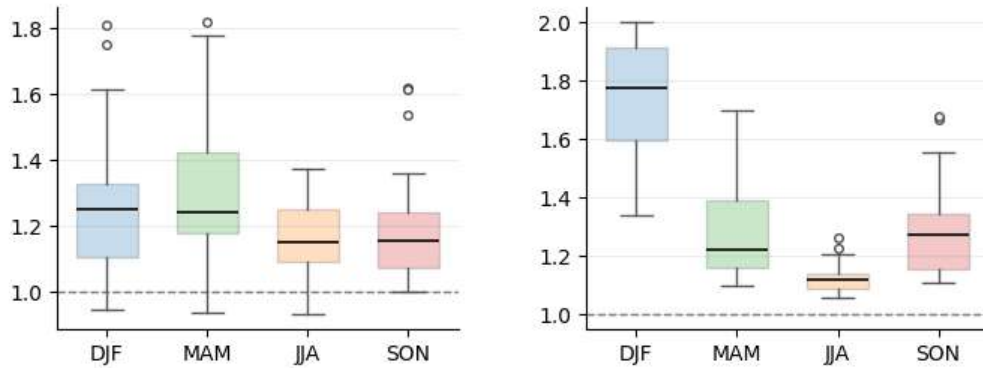
*Máscara binaria sobre zona cordillerana.*



Para cada ventana del conjunto de prueba se identifica, entre los  $K$  modos de mayor amplitud, el modo con mayor concentración en la región cordillerana (modo “más cordillerano” de la ventana). La Figura 5.34 presenta la distribución de *enrich* por estación (boxplot), evidenciando que el enriquecimiento del modo más concentrado en la cordillera se mantiene de manera consistente por encima de 1 en la mayoría de ventanas, con variabilidad estacional moderada.

**Figura 5.34**

*Distribución de enrich para conjunto de pruebas.*



*Nota:* Izquierda, distribución de enrich de los datos por estación. Derecha, distribución de enrich de los modos más representativos por estación.

Complementariamente, la Tabla 5.7 resume por estación la mediana de *enrich* y un indicador de recurrencia definido como el porcentaje de ventanas donde el enriquecimiento supera un umbral  $\tau = 1.1$ , escogido como criterio “moderado” en función de la distribución observada (Figura 5.34). Este análisis permite afirmar que la presencia de modos con patrón cordillerano no corresponde a un caso puntual, sino a un fenómeno recurrente en el conjunto de prueba y en distintas estaciones.

**Tabla 5.7**

*Resumen por estación del enriquecimiento y recurrencia por ventana.*

<i>season</i>	<i>enrich data</i>	<i>enrich modo</i>	<i>% recurrencia</i>
DJF	1,77	1.25	77,77%
JJA	1,12	1,15	74,29%
MAM	1,22	1,24	88,23%
SON	1,27	1,15	64,71%

Adicionalmente, puede compararse *enrich* del modo con un análogo computado directamente desde los datos de la ventana (a partir de la frecuencia de días húmedos en cada celda para un umbral de 1 mm). Esta comparación (Tabla 5.7) entrega contexto: en algunas estaciones la señal cordillerana es más marcada en los datos, mientras que en otras el patrón modal resalta estructuras coherentes aun cuando la estadística simple (frecuencia de días húmedos) sea más homogénea.

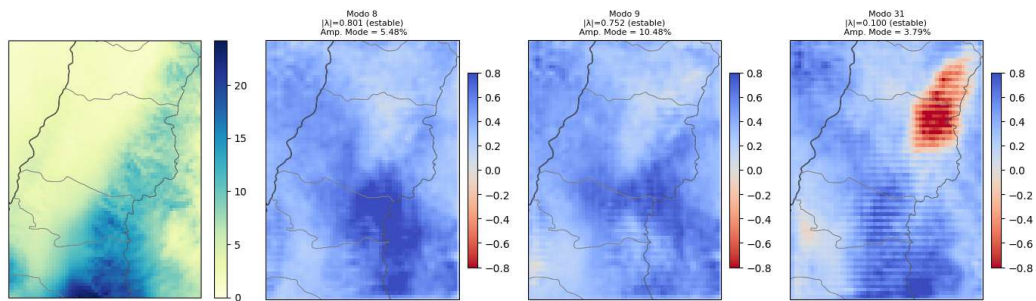
Estos resultados aportan evidencia de estructura espacial recurrente (por ejemplo, concentración relativa en zona orográfica), pero no constituyen una atribución física causal. La vinculación con topografía explícita e índices climáticos (ENSO/AAO) se deja como trabajo futuro.

### **Ejemplos representativos de patrones modales por estación**

Para facilitar la interpretación cualitativa sin caer en un catálogo de casos, se seleccionan ventanas representativas por estación (cercanas a la mediana de *enrich*) y se visualizan los Top-3 modos con mayor concentración regional dentro de la ventana. Las Figuras 5.35 y 5.36 muestran ejemplos para dos estaciones, incluyendo en cada subpanel los metadatos espectrales del modo ( $|\lambda_k|$ , frecuencia y crecimiento) junto a su amplitud relativa y su *enrich*. En estos ejemplos se observan patrones coherentes con estructuras esperables del dominio (por ejemplo, lóbulos y dipolos espaciales que resaltan diferencias costa–interior y/o concentraciones hacia el borde Este del dominio, consistente con influencia orográfica). El hecho de que patrones similares aparezcan en múltiples fechas dentro de la misma estación, y que la estadística global de recurrencia sea consistente (Figura X, Tabla X), refuerza la interpretación de que estos modos capturan componentes estructuradas de la variabilidad espacio–temporal.

**Figura 5.35**

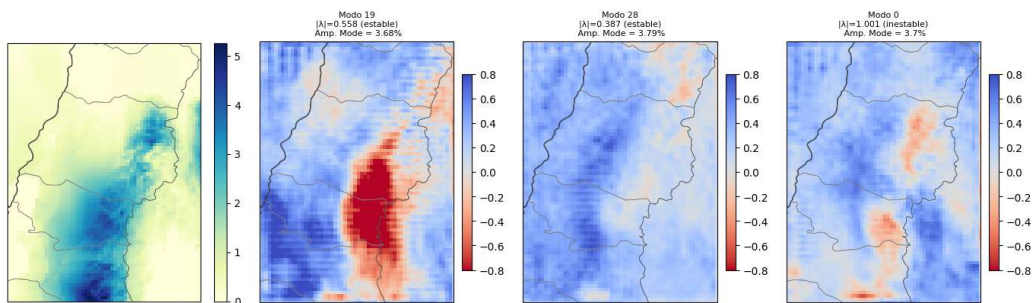
*Ventana representativa de invierno (JJA): señal del bloque y modos DMD.*



*Nota:* A la izquierda se muestra promedio de precipitaciones. A la derecha, los tres modos DMD seleccionados.

**Figura 5.36**

*Ventana representativa de verano (DJF): señal del bloque y modos DMD.*



*Nota:* A la izquierda se muestra promedio de precipitaciones. A la derecha, los tres modos DMD seleccionados.

### **Limitaciones del análisis**

Este análisis es principalmente interpretativo y depende de decisiones explícitas: longitud de ventana, preselección Top-K por amplitud y definición de la máscara regional. Además, el índice del modo  $k$  no debe interpretarse como una “identidad física” invariante entre ventanas (el orden de modos puede permutar), por lo que los análisis de repetición se entienden como frecuencia de aparición en el ranking bajo un criterio fijo (amplitud y/o concentración). Finalmente, los mapas de modos se reportan como patrones normalizados; su utilidad es describir estructura espacial coherente y su recurrencia, más que cuantificar directamente precipitación en unidades físicas.

## **6. Conclusiones**

Este trabajo desarrolló y evaluó un sistema híbrido Autoencoder 3D + Descomposición Modal Dinámica (AE+DMD) para el pronóstico espacio–temporal de precipitación diaria en el subdominio Maule–Ñuble, buscando un equilibrio entre precisión, interpretabilidad y eficiencia computacional. El enfoque propuesto se apoya en (i) compresión no lineal de ventanas espacio–temporales y (ii) modelación de la dinámica en el espacio latente mediante un operador lineal estimado con DMD. El desempeño se evaluó comparando con un baseline temporal estándar (DeepAR) y un baseline simple (persistencia), utilizando métricas escalares (MAE/RMSE), análisis temporal y métricas orientadas a eventos y habilidad espacial.

### **6.1. Cumplimiento de objetivos**

En relación con los objetivos específicos planteados en la Sección 3.2, los resultados presentados permiten verificar su cumplimiento de la siguiente manera. Primero, se entrenó un Autoencoder 3D con capacidad de compresión espacio–temporal, lo cual queda respaldado por los resultados de reconstrucción presentados en la Sección 5.1, donde se observa que el modelo preserva la estructura espacial y temporal relevante del campo de precipitación, habilitando una representación latente de menor dimensión.

Segundo, se implementó el pronóstico AE+DMD aplicando DMD en el espacio latente, y su desempeño se evaluó en la Sección 5.2. Los resultados de esa sección muestran que la dinámica lineal estimada en el espacio latente permite generar pronósticos multi-step de manera estable bajo el esquema de evaluación definido.

Tercero, se realizó la comparación con baselines, particularmente DeepAR y persistencia, utilizando métricas determinísticas (MAE/RMSE y métricas complementarias según corresponda). En la Sección 5.3 se reporta el desempeño comparativo del enfoque híbrido frente a DeepAR, definido como baseline temporal moderno, y persistencia como referencia mínima, permitiendo situar el rendimiento del modelo propuesto en un marco estándar de comparación.

Cuarto, se abordó la interpretabilidad del método mediante un análisis espectral y de modos DMD. En la Sección 5.4 se presentan resultados asociados a autovalores y modos (magnitud  $|\lambda|$ , frecuencia, tasa de crecimiento/decaimiento y amplitud relativa), lo que permite caracterizar patrones dominantes y recurrencia temporal en distintas ventanas y estaciones, aportando una lectura dinámica del comportamiento del sistema. Adicionalmente, la interpretación espacial de los modos se discute en términos de su coherencia con la configuración del dominio (incluyendo diferencias entre zonas), lo que entrega una primera lectura física cualitativa. Sin embargo, la vinculación cuantitativa con patrones climáticos externos de gran escala (como ejemplo ENSO) y una validación

orográfica explícita requieren integrar índices externos y definir un esquema de análisis con rezagos, por lo que se dejan como trabajo futuro.

Finalmente, se evaluó la eficiencia computacional del enfoque en términos de entrenamiento e inferencia. Los tiempos reportados en la sección de comparación evidencian un trade-off consistente: AE+DMD requiere mayor costo de entrenamiento, pero puede alcanzar inferencias significativamente más rápidas que DeepAR bajo el pipeline de evaluación utilizado, lo cual favorece su uso en escenarios operacionales donde el pronóstico se ejecuta repetidamente y el entrenamiento se realiza con menor frecuencia.

## **6.2. Principales hallazgos**

- El AE 3D cumple su rol como compresor espacio-temporal, capturando estructuras relevantes del campo de precipitación en una representación latente robusta, lo que hace viable aplicar un modelo dinámico posterior en un espacio de menor dimensión.
- La incorporación de DMD sobre el latente entrega un mecanismo de pronóstico parsimonioso, con un componente explícitamente interpretable: los autovalores y modos permiten caracterizar escalas temporales dominantes y patrones espaciales recurrentes (Sección 5.4), evitando que el pronóstico sea una “caja negra” puramente neuronal.

- La comparación con DeepAR es informativa pero no “simétrica” en estructura: DeepAR se utiliza como baseline temporal moderno, pero no explota explícitamente la estructura espacial del campo. En este trabajo, para comparabilidad con AE+DMD y persistencia, se utiliza un pronóstico puntual dado por la media de su distribución predictiva.
- Costo computacional (resultado clave para la hipótesis de eficiencia): los resultados de tiempos muestran un trade-off claro. AE+DMD presenta un costo de entrenamiento mayor, pero una inferencia sustancialmente más rápida que DeepAR (Tabla de costos incorporada en Resultados). Este comportamiento es coherente con la idea operacional del enfoque: entrenar con menor frecuencia y ejecutar pronósticos repetidamente con baja latencia.

### **6.3. Limitaciones**

- Naturaleza intermitente y colas pesadas de precipitación: pese a incorporar métricas orientadas a eventos, la precipitación extrema sigue siendo el caso más desafiante; por diseño, el enfoque se evaluó principalmente con métricas determinísticas y de ocurrencia/habilidad espacial, lo que no garantiza calibración probabilística de extremos.
- Interpretabilidad física incompleta: aunque el análisis espectral y de modos permite interpretar recurrencia, escalas temporales y patrones espaciales (Sección 5.4), no se realizó un contraste cuantitativo sistemático con forzantes externos

(como índices ENSO/AAO por ejemplo) ni una validación orográfica explícita mediante topografía, lo que limita el grado de atribución física.

- Además, el espacio latente es aprendido de forma puramente data-driven; no hay restricciones físicas que aseguren separaciones ‘limpias’ de procesos, por lo que un modo DMD puede representar combinaciones de mecanismos. En consecuencia, la interpretación debe entenderse como aproximación empírica.
- Baseline DeepAR evaluado en modo puntual: DeepAR es un modelo probabilístico; en este trabajo se usa la media como estimador puntual para comparación directa. Esto es correcto para MAE/RMSE, pero deja fuera la evaluación de incertidumbre (intervalos/quantiles), que podría enriquecer la comparación.

## 6.4. Trabajo futuro

- Evaluación probabilística: extender la comparación incorporando cuantiles/intervalos de DeepAR y/o un componente probabilístico en el enfoque híbrido para evaluar calibración y desempeño en colas.
- Interpretabilidad física cuantitativa: correlacionar métricas de amplitud/recurrencia de modos con índices climáticos (ENSO/AAO) y contrastar patrones espaciales con topografía/orografía para respaldar la atribución física.
- Comparación con modelos espacio–temporales puros: incorporar baselines que modelen explícitamente estructura espacial (ConvLSTM, UNet-3D, Transformers espacio–temporales, entre otros) para una comparación más “pareja” en inductive bias, manteniendo controlado el costo computacional.

## 7. Anexos

### 7.1. Anexo A. Configuración experimental e hiperparámetros

**Tabla A.1**

*Hiperparámetros de preprocesamiento.*

<b>Parámetro</b>	<b>Valor</b>	<b>Unidad</b>
Ventana de entrenamiento ( $T_{IN}$ )	64	días
Forecast ( $h$ )	1 a 5	días
Stride	1	día
Media ( $\mu$ )	3.50	mm
Desviación estándar ( $sd$ )	10.51	mm

**Tabla A.2***Arquitectura del modelo híbrido AE 3D + DMD*

<b>Bloque</b>	<b>Capa (nombre)</b>	<b>Tipo</b>	<b>#Filtros</b>	<b>Kernel (t,h,w)</b>	<b>Stride (t,h,w)</b>	<b>Padd ing</b>	<b>Activación</b>	<b>Output shape*</b>
Input	recon_target	Input	–	–	–	–	–	(64, 80, 60, 1)
Encoder	conv3d_1	Conv3D	32	(3,3,3)	(1,1,1)	same	ReLU	(64, 80, 60, 32)
Latente	Latent_Space	Conv3D	64	(3,3,3)	(2,2,2)	same	ReLU	(32, 40, 30, 64)
Dinámica	dmd_loss	DMDLossLayer	–	–	–	–	–	(32, 40, 30, 64) + $\hat{z}_{t+h}$
Decoder	conv3d_transpose	Conv3DTranspose	32	(3,3,3)	(2,2,2)	same	ReLU	(64, 80, 60, 32)
Output	conv3d_transpose_1	Conv3DTranspose	1	(3,3,3)	(1,1,1)	same	Linear	(64, 80, 60, 1)

**Tabla A.3***Configuración de entrenamiento del modelo híbrido AE+DMD.*

Categoría	Parámetro	Símbolo / nombre	Valor
Datos	Tamaño batch	batch_size	8
Optimización	Optimizador	optimizer	Adam
Optimización	Learning rate	$lr$	$1 \times 10^{-3}$
Optimización	Épocas máximas	epochs_max	60
Optimización	Early stopping	patience	20
Optimización	Criterio best model	restore_best_weights	True
Pérdida	Reconstrucción	$\lambda_{recon}$	MSE
Pérdida	Predicción 1–5 días	$\lambda_{pred}$	MSE
Pérdida	Peso recon	$W_{recon}$	1
Pérdida	Peso pred	$W_{pred}$	2
DMD (estabilidad)	Ridge	$\lambda_{ridge}$	$1 \times 10^{-4}$
DMD (estabilidad)	Umbral espectral	$\rho_{th}$	0.99
DMD (estabilidad)	Peso penalización espectral	$W_{\rho}$	$5 \times 10^{-6}$
DMD (estabilidad)	Proyección de rango	$r_{proj}$	32
Reproducibilidad	Seed	seed	42

*Nota A.* Cabezas de salida. El decoder es compartido para reconstrucción y predicción. La reconstrucción se obtiene como  $\hat{X}_{0:T_{IN}-1} = \text{Dec}(z)$ . Las predicciones a  $h \in \{1, \dots, 5\}$  se obtienen decodificando los estados latentes pronosticados por DMD:

$$\hat{X}_{t+h} = \text{Dec}(\hat{z}_{t+h}).$$

El modelo entrega seis salidas: reconstrucción (64 pasos) y predicciones multi-step (62, 60, 58, 56, 54 pasos) por alineamiento temporal (“match\_shape”).

*Nota B.* El modelo se entrena optimizando una pérdida total que combina reconstrucción y predicción decodificada, ponderadas por  $\lambda_{\text{recon}}$  y  $\lambda_{\text{pred}}$ . Durante entrenamiento, el operador A se estima por batch en el espacio latente y se controla su estabilidad mediante regularización ridge y penalización/esquemas de control espectral según los parámetros reportados.

**Tabla A.***Configuración de entrenamiento del baseline DeepAR.*

<b>Categoría</b>	<b>Parámetro</b>	<b>Símbolo / nombre</b>	<b>Valor</b>
Datos	Frecuencia temporal	freq	"D"
Datos	Estructura de series	–	Univariado N=HxW
Modelo	Estimador	DeepAREstimator	DeepAR (GluonTS Torch)
Modelo	Ventana de contexto	context_length	120
Modelo	Horizonte predicción	prediction_length	5
Modelo	Tamaño estado oculto	hidden_size	64
Modelo	Nº capas RNN	num_layers	5
Modelo	Dropout	dropout_rate	0.1
Optimización	Learning rate	lr	1,00E-02
Entrenamiento	Batch size	batch_size	64
Entrenamiento	Batches por epoch	num_batches_per_epoch	50
Entrenamiento	Épocas máximas	max_epochs	100
Entrenamiento	Aceleración	accelerator	"gpu"
Reproducibilidad	Seed	np.random.seed(42)	42

**Tabla A.5***Entorno de ejecución.*

<b>Categoría</b>	<b>Elemento</b>	<b>Nombre</b>	<b>Valor</b>
Hardware	GPU	gpu_model	NVIDIA GeForce RTX 4070 Laptop GPU
Hardware	VRAM (dedicada)	gpu_vram_gb	8 GB
Hardware	CPU	cpu_model	AMD Ryzen 9 7940HS (con Radeon 780M Graphics)
Hardware	RAM	ram_gb	32 GB
Sistema	SO / runtime	runtime	Windows (host) + Docker
GPU stack	CUDA	cuda_version	11.8
GPU stack	cuDNN	cuda_version	8
Frameworks	PyTorch	torch_version	2.2.0
Frameworks	GluonTS	gluonts_version	0.14.3
Ejecución	Aceleración	accelerator	GPU

## 8. Referencias

- Absil, P.-A., Mahony, R., & Sepulchre, R. (2007). Optimization algorithms on matrix manifolds. Princeton University Press.
- Baddoo, P. J., Herrmann, B., McKeon, B. J., Kutz, N. J., & Brunton, S. L. (2023). Physics-informed dynamic mode decomposition. *Proceedings of the Royal Society A: Mathematical, Physical, and Engineering Sciences* 479(2271), 20220576. <https://doi.org/10.1098/rspa.2022.0576>
- Berman, N., Naiman, I., & Azencot, O. (2023). Multifactor sequential disentanglement via structured Koopman autoencoders. arXiv. <https://arxiv.org/abs/2303.17264>
- Boisier, J. P., & Álvarez-Garretón, C. (2023). CR2MET: Un conjunto de datos de precipitación y temperatura de alta resolución para el período 1960–2021 en Chile continental [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7529682>
- Chandra, R., Goyal, S., & Gupta, R. (2021). Evaluation of deep learning models for multi-step ahead time series prediction. arXiv. <https://doi.org/10.48550/arXiv.2103.14250>
- Gower, J. C., & Dijksterhuis, G. B. (2004). Procrustes problems. Oxford University Press.
- Koopman, B. O. (1931). Hamiltonian systems and transformation in Hilbert space. *Proceedings of the National Academy of Sciences*, 17(5), 315–318. <https://doi.org/10.1073/pnas.17.5.315>

- Kutz, J. N., Brunton, S. L., Brunton, B. W., & Proctor, J. L. (2016). Dynamic mode decomposition: Data-driven modeling of complex systems. SIAM.  
<https://doi.org/10.1137/1.9781611974508>
- Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirmsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., & Battaglia, P. (2023). GraphCast: Learning skillful medium-range global weather forecasting. *Science*, 382(6677), 1469–1474. <https://doi.org/10.1126/science.adi2336>
- Lusch, B., Kutz, J. N., & Brunton, S. L. (2018). Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, 9, Article 4950.  
<https://doi.org/10.1038/s41467-018-07210-0>
- Marchant, M., & Silva, S. (2024). Pronóstico espaciotemporal de precipitaciones en Chile utilizando redes neuronales y métodos estadísticos avanzados. [Tesis de magíster, Universidad del Desarrollo].
- Mezić, I. (2005). Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41(1–3), 309–325.  
<https://doi.org/10.1007/s11071-005-2824-x>
- Pan, S., & Duraisamy, K. (2020). Physics-informed probabilistic learning of linear embeddings of nonlinear dynamics with guaranteed stability. arXiv.  
<https://doi.org/10.48550/arXiv.1906.03663>
- Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., & Thuerey, N. (2020). WeatherBench: A benchmark data set for data-driven weather forecasting.

- Journal of Advances in Modeling Earth Systems, 12(11), e2020MS002203.  
<https://doi.org/10.1029/2020MS002203>
- Reichstein, M., Camps-Valls, G., & Stevens, B. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195–204.  
<https://doi.org/10.1038/s41586-019-0912-1>
- Rowley, C. W., Mezić, I., Bagheri, S., Schlatter, P., & Henningson, D. S. (2009). Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics*, 641, 115–127.  
<https://doi.org/10.1017/S0022112009992059>
- Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181–1191.  
<https://doi.org/10.1016/j.ijforecast.2019.07.001>
- Schmid, P. J. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656, 5–28.  
<https://doi.org/10.1017/S0022112010001217>
- Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3D convolutional networks. *arXiv*.  
<https://doi.org/10.48550/arXiv.1412.0767>
- Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., & Kutz, J. N. (2014). On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1(2), 391–421.

Williams, M. O., Kevrekidis, I. G., & Rowley, C. W. (2015). A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6), 1307–1346.  
<https://doi.org/10.1007/s00332-015-9258-5>

Zavala, L. A., & Pérez, A. E. (2023). Predicción espacio-temporal de datos climáticos en Chile con el uso de EOFs y Deep Learning. [Tesis de magíster, Universidad del Desarrollo].