

Evaluación de métodos de regionalización estadística para la generación de proyecciones climáticas en el marco del PNACC-2 2021-2030

Nota técnica 41 de AEMET

Carlos Correa Guinea

Alfonso Hernanz Lázaro

Esteban Rodríguez Guisado





Aviso Legal: los contenidos de esta publicación podrán ser reutilizados, citando la fuente y la fecha, en su caso, de la última actualización



Edita:

© Ministerio para la Transición Ecológica y el Reto Demográfico

Agencia Estatal de Meteorología

Madrid, 2023

Catálogo de Publicaciones de la Administración General del Estado:

<https://cpage.mpr.gob.es>

NIPO: 666-23-009-0

<https://doi.org/10.31978/666-23-009-0>

Agencia Estatal de Meteorología (AEMET)

C/ Leonardo Prieto Castro, 8

28040 Madrid

<http://www.aemet.es/>



@Aemet_Esp



<https://www.facebook.com/AgenciaEstataldeMeteorologia>

ÍNDICE

LISTA DE ACRÓNIMOS	4
1. INTRODUCCIÓN.....	7
2. DESCRIPCIÓN DE MÉTODOS ESD	8
2.1. <i>Perfect prognosis</i>	9
2.1.a. Función de Transferencia	10
2.1.b. Tipos de Tiempo y Análogos	17
2.2. <i>Model Output Statistics</i>	18
2.2.a. <i>Delta Change</i>	19
2.2.b. <i>Quantile Mapping</i> y <i>Quantile Delta Mapping</i>	20
3. EVALUACIÓN Y SELECCIÓN DE MODELOS CLIMÁTICOS GLOBALES	20
3.1. Modelos evaluables en GCMeval y disponibilidad de variables en ESGF ..	21
3.2. Evaluación de modelos junto a la selección de EURO-CORDEX.....	21
3.3. Selección definitiva de modelos	24
4. EVALUACIÓN DE MÉTODOS ESD	26
4.1. Introducción.....	26
4.2. Evaluación en Península y Baleares	29
4.2.a. Ciclo anual.....	29
4.2.b. Sesgo	31
4.2.c. Correlación	34
4.2.d. Varianza	35
4.2.e. Correlación espacial	37
4.2.f. Raíz del error cuadrático medio	38
4.2.g. Tendencias a futuro	39
4.3. Evaluación en Canarias	43
4.3.a. Ciclo anual.....	43
4.3.b. Sesgo	44
4.3.c. Correlación	48

4.3.d. Varianza	49
4.3.e. Correlación espacial	50
4.3.f. Raíz del error cuadrático medio	52
4.3.g. Tendencias a futuro	52
5. CONCLUSIONES	52
6. REFERENCIAS	53
ANEXO A. SESGOS Y TENDENCIAS ESTACIONALES EN PENÍNSULA Y BALEARES	57
ANEXO B. SESGOS Y TENDENCIAS ESTACIONALES EN CANARIAS	73

EVALUACIÓN DE MÉTODOS DE REGIONALIZACIÓN ESTADÍSTICA PARA LA GENERACIÓN DE PROYECCIONES CLIMÁTICAS EN EL MARCO DEL PNACC-2 2021-2030.

Carlos CORREA GUINEA, Alfonso HERNANZ LÁZARO, Esteban RODRÍGUEZ
GUISADO

Agencia Estatal de Meteorología, Área de Modelización y Evaluación del Clima
ccorreag@aemet.es, ahernanzl@aemet.es, erodriguezg@aemet.es

RESUMEN

La Agencia Estatal de Meteorología (AEMET) es responsable de la elaboración de las proyecciones climáticas regionalizadas sobre el territorio de España según lo establecido en el segundo Plan Nacional de Adaptación al Cambio Climático (PNACC-2). Las proyecciones climáticas regionalizadas suponen un aumento de la resolución espacial con respecto a las proyecciones climáticas proporcionadas por los modelos climáticos globales. AEMET elabora las proyecciones climáticas regionalizadas sobre España aplicando métodos de regionalización empírico-estadísticos a las proyecciones climáticas de un conjunto de modelos climáticos globales. Las proyecciones climáticas regionalizadas son necesarias para la realización de estudios de impacto y vulnerabilidad que requieren datos en alta resolución espacial. La presente nota técnica tiene dos objetivos: 1) evaluar y seleccionar un conjunto de modelos climáticos globales cuyas proyecciones serán regionalizadas; 2) comparar diferentes métodos empírico-estadísticos de regionalización a fin de elegir un método para regionalizar los datos diarios de las siguientes variables climáticas de interés (temperatura máxima, temperatura mínima y precipitación acumulada). En cuanto al primer objetivo, los resultados han permitido seleccionar un conjunto de once modelos climáticos globales que serán regionalizados. Respecto al segundo objetivo, a partir de la comparativa de métodos se establece que la temperatura máxima y la temperatura mínima se regionalizarán utilizando el método híbrido de regresión-análogos (MLR-ANA), mientras que la precipitación acumulada se regionalizará aplicando el método *eXtreme Gradient Boost* (XGB). Además, tras la regionalización, tanto en temperaturas como en precipitación, se establece la conveniencia de aplicar una corrección de sesgos a posteriori mediante *Quantile Delta Mapping* (QDM) para mejorar los índices de comparación.

Palabras clave: cambio climático; modelos climáticos globales; proyecciones climáticas; regionalización; métodos empírico-estadísticos.

LISTA DE ACRÓNIMOS

ANA: *Analogs* (análogos).

ANA-LOC-kNN: subtipo del método de análogos basado en analogía sinóptica combinada con analogía local.

ANN: *Artificial Neural Networks* (redes neuronales artificiales).

BC: *Bias Correction* (corrección de sesgos).

CART: *Classification And Regression Tree* (árbol de clasificación y regresión).

CDD: máximo anual de la temperatura máxima diaria (índice CLIMDEX).

CSDI: duración de ola de frío: número anual de días con al menos 6 días consecutivos con temperatura mínima < percentil 10 (índice CLIMDEX).

CLIMDEX: *Climate extremes índices* (índices climáticos de valores extremos).

CMIP5: *Coupled Model Intercomparison Project – Phase 5* (fase 5 del proyecto de intercomparación de modelos climáticos acoplados).

CMIP6: *Coupled Model Intercomparison Project – Phase 5* (fase 6 del proyecto de intercomparación de modelos climáticos acoplados).

CORDEX: *Coordinated Regional Climate Downscaling Experiment* (Experimento coordinado de regionalización climática).

CWD: duración máxima del periodo seco: número máximo de días consecutivos con precipitación < 1 mm (índice CLIMDEX).

DEF: periodo estacional Diciembre – Enero – Febrero.

ECDF: *Empirical Cumulative Distribution Function* (función de distribución empírica).

ERA5: *ECMWF Reanalysis v5* (versión 5 del reanálisis del centro europeo de predicción a medio plazo o CEPPM).

ESD: *Empirical Statistical Downscaling* (regionalización empírico-estadística).

ESGF: *Earth System Grid Federation* (grupo de colaboración dedicado al desarrollo y gestión de un sistema Peer-to-Peer, P2P, difundir y analizar datos climáticos de gran volumen).

EURO-CORDEX: EURO-CORDEX es la rama europea de la iniciativa internacional CORDEX.

GCM: *Global Climate Model* (modelo climático global).

GCMeval: herramienta interactiva de evaluación de GCM (Parding et al., 2020; <https://gcmeval.met.no/>).

GLM: *Generalized Linear Model* (modelos lineales generalizados).

GLM-EXP: *Generalized Linear Model* (modelos lineales generalizados) con función de enlace de tipo exponencial.

HPC: *High-Performance Computing* (Supercomputación).

hur: humedad relativa por niveles de presión (dato diario).

hus: humedad específica por niveles de presión (dato diario).

IPCC: *The Intergovernmental Panel on Climate Change* (Panel Intergubernamental sobre el Cambio Climático).

JJA: periodo estacional Junio - Julio – Agosto.

LIN: modelos lineales simples.

MAM: periodo estacional Marzo - Abril – Mayo.

ML: *Machine Learning* (aprendizaje automático).

MLP: *MultiLayer Perceptron* (*perceptrón multicapa*).

MLR: *Multiple Linear Regression* (regresión lineal múltiple).

MLR-ANA: método híbrido que combina regresión lineal múltiple con análogos.

MOS: *Model Output Statistics*.

OBS: observaciones.

PRCPTOT: precipitación total anual en la que se consideran solo los días lluviosos (aquellos con precipitación ≥ 1 mm) (índice CLIMDEX).

PNACC-2: segunda fase del Plan Nacional de Adaptación al cambio climático.

PP: *Perfect Prognosis* (pronóstico perfecto).

PR/pr: precipitación acumulada diaria.

psl: presión a nivel del mar (dato diario).

pyClim-SDM: *software* de regionalización climática mediante métodos ESD desarrollado en AEMET.

QDM: *Quantile Delta Mapping*.

QDMs *Quantile Delta Mapping by season*.

QM: *Quantile Mapping*.

R01: número de días en el año con precipitación ≥ 1 mm, el número anual de días lluviosos (índice CLIMDEX).

R95p: precipitación total anual en la que se consideran solo los días muy lluviosos, aquellos en que se supera el percentil 95 de los días con precipitación ≥ 1 mm (índice CLIMDEX).

RAW: datos en bruto de los GCM interpolados por el método del vecino más cercano.

RELU: *Rectified Linear Unit* (unidad lineal rectificada).

RF: *Random Forest* (bosques aleatorios).

RCM: *Regional Climate Model* (modelo climático regional).

SDII: índice simple de intensidad de precipitación en días lluviosos (índice CLIMDEX).

SLURM: *Simple Linux Utility for Resources Management*.

SON: periodo estacional Septiembre - Octubre – Noviembre.

SSP: *Shared Socioeconomic Pathway* o trayectoria socioeconómica compartida.

SSP5-8.5: SSP con desarrollo impulsado por combustibles fósiles y forzamiento radiativo en el año 2100 de 8.5 W/m^2 .

SVM: *Support Vectors Machine* (máquinas de vector de soporte).

ta: temperatura por niveles de presión (dato diario).

tas: temperatura en superficie (dato diario).

TASMAX/tasmax: temperatura máxima diaria.

TASMIN/tasmin: temperatura mínima diaria.

TF: *Transfer Function* (función de transferencia).

TNm: media anual de la temperatura mínima diaria (índice CLIMDEX).

TNn: mínimo anual de la temperatura mínima diaria (índice CLIMDEX).

TXx: máximo anual de la temperatura mínima diaria (índice CLIMDEX).

TXm: media anual de la temperatura máxima diaria (índice CLIMDEX).

TXn: mínimo anual de la temperatura máxima diaria (índice CLIMDEX).

TXx: máximo anual de la temperatura máxima diaria (índice CLIMDEX).

ua: viento zona por niveles de presión 1 (dato diario).

UNFCCC: *United Nations Framework Convention on Climate Change* (convención marco de las Naciones Unidas sobre el cambio climático).

va: viento meridional por niveles de presión (dato diario).

WCRP: *World Climate Research Programm* (programa mundial de investigaciones climáticas).

WG: *Weather Generators* (generadores de tiempo).

WSDI: duración de ola de calor: número anual de días con al menos 6 días consecutivos con temperatura máxima > percentil 90 (índice CLIMDEX).

WT: *Weather Types* (Tipos de Tiempo).

XGB: *eXtreme Gradient Boost* (método de refuerzo de gradientes extremo).

1. INTRODUCCIÓN

Los modelos climáticos globales o GCM (*Global Climate Models*) son la principal herramienta para predecir el clima futuro, sin embargo, es sabido que sus proyecciones pueden presentar sesgos y que su resolución espacial es habitualmente insuficiente tanto para gestionar el impacto del cambio climático a escala regional como para desarrollar estrategias eficaces de adaptación, lo cual es esencial en muchas regiones vulnerables del planeta. Para solventar esta problemática surgen las técnicas de regionalización climática enfocadas a aumentar, en una región determinada, la resolución de las proyecciones climáticas globales proporcionadas por los GCM. Las técnicas de regionalización se pueden dividir en dos categorías principales: dinámicas y empírico-estadísticas.

La regionalización dinámica consiste fundamentalmente en anidar un modelo climático regional o RCM (*Regional Climate Model*) de alta resolución dentro de un GCM. La fortaleza de la regionalización dinámica radica en que se fundamenta en representaciones explícitas de principios físicos (por ejemplo, las leyes de la termodinámica y la mecánica de fluidos, que se espera que se mantengan bajo condiciones de cambio climático) y, también, en que proporciona simulaciones con consistencia física (por ejemplo, entre las distintas variables) y espacial. Las principales desventajas de los RCM son que pueden dar lugar a sesgos a gran escala y que sus simulaciones son computacionalmente costosas.

El enfoque alternativo para obtener información en alta resolución sobre el cambio climático a escala local o regional es la regionalización empírico-estadística de las proyecciones climáticas globales. Este enfoque, conocido como metodología ESD (*Empirical Statistical Downscaling*), se basa en la existencia de relaciones empírico-estadísticas entre variables a escala global (predictores \mathbf{x}) y variables a escala regional (predictandos \mathbf{y}) formalizadas a través de un modelo estadístico $F()$:

$$\mathbf{y} \sim F(\mathbf{x}) \quad (1.1)$$

La obtención de estas relaciones empírico-estadísticas permite proporcionar proyecciones con mayor resolución y una representación más precisa de eventos extremos localizados que respaldan una evaluación más adecuada del impacto del cambio climático y una planificación más detallada de las estrategias de adaptación. Los métodos estadísticos de regionalización se han aplicado ampliamente en estudios de impacto y adaptación debido a que son métodos de bajo coste computacional en comparación con los métodos dinámicos de regionalización y a que son capaces de regionalizar hasta la escala puntual. No obstante, la metodología ESD también tiene limitaciones: 1) la hipótesis de estacionariedad; 2) la necesidad de observaciones (series largas y de calidad); 3) la herencia de los sesgos de los GCM.

El [Acuerdo de París](#) alcanzado en 2015 por la Convención Marco de las Naciones Unidas sobre el Cambio Climático (UNFCCC, 2015) establece en su artículo 7 que las Partes deben tomar las medidas de adaptación pertinentes en sus respectivos territorios. Para esta tarea se necesita información a escala regional sobre el cambio climático. En este sentido, se está realizando un esfuerzo cada vez mayor para proporcionar proyecciones climáticas de alta resolución en todo el mundo. El experimento internacional [CORDEX](#) (*Coordinated Regional Climate Downscaling Experiment*), patrocinado por el [WCRP](#) (*World Climate Research Programme*), es la fuente más utilizada de proyecciones climáticas de alta resolución en la que se apoyan tanto los estudios de impacto climático regionales como la planificación de las líneas estratégicas de adaptación. Sin embargo, las proyecciones regionalizadas de CORDEX carecen

de suficiente resolución para determinadas aplicaciones y tienen diferente cobertura en distintas regiones, pudiendo ser demasiado escaso el número de simulaciones en algunas de ellas. Por esta razón, en ocasiones, las Partes necesitan generar adicionalmente sus propias proyecciones climáticas regionalizadas, generalmente utilizando métodos ESD.

En esta nota técnica se presenta la evaluación de los métodos ESD realizada en AEMET como paso previo a la próxima regionalización de los modelos climáticos globales del CMIP6 (*Coupled Model Intercomparison Project – Phase 6*) en España y que será complementada por las simulaciones de EURO-CORDEX. CMIP proporciona proyecciones climáticas fundamentales para las actividades esenciales del WCRP y la climatología en todo el mundo, así como para tomadores de decisiones políticas, en su objetivo de comprender los cambios climáticos pasados, presentes y futuros. El CMIP y su infraestructura de datos asociada se han vuelto esenciales para el Panel Intergubernamental sobre el Cambio Climático (IPCC). La regionalización de los modelos climáticos globales del CMIP6 en el territorio nacional tiene como objetivo la generación de proyecciones climáticas en las distintas trayectorias socioeconómicas compartidas (SSP, véase Riahi *et al.*, 2017) en el marco de la segunda fase del Plan Nacional de Adaptación al Cambio Climático ([PNACC-2](#)) para el periodo 2021-2030. El PNACC-2, elaborado por la Oficina Española de Cambio Climático del Ministerio para la Transición Ecológica y el Reto Demográfico, constituye el instrumento de planificación básico para promover la acción coordinada frente a los efectos del cambio climático en España. Tiene como principal objetivo evitar o reducir los daños presentes y futuros derivados del cambio climático y construir una economía y una sociedad más resilientes.

2. DESCRIPCIÓN DE MÉTODOS ESD

Hay una gran variedad de métodos ESD que pertenecen a diferentes familias y que asumen diferentes hipótesis, pero todos ellos tienen en común que se calibran o se entrenan utilizando un conjunto de datos para buscar patrones y relaciones entre las entradas (predictores) y las salidas (predictandos). El objetivo no es otro que ajustar los parámetros las relaciones empírico-estadísticas de manera que se puedan realizar predicciones precisas de las salidas correspondientes a nuevos datos de entrada no pertenecientes al conjunto de datos de calibración. Los métodos ESD se pueden clasificar según su estrategia de calibración en *Perfect Prognosis* (PP) y *Model Output Statistics* (MOS). En los métodos ESD del tipo PP la calibración del modelo se realiza completamente en el mundo real, es decir, predictores y predictandos se toman o bien de observaciones, o bien de datos procedentes de reanálisis. En el caso de los predictores es frecuente utilizar datos de reanálisis, mientras que para los predictandos se suelen usar rejillas de observación a alta resolución o datos observacionales de estaciones meteorológicas. Por el contrario, en los métodos tipo MOS, la calibración se basa en encontrar relaciones entre predictores que son la salida de los GCM en el periodo de calibración y predictandos de alta resolución observados en dicho periodo. Tanto en PP como en MOS, el modelo calibrado se aplica a \mathbf{x}_{GCM} (los predictores en el futuro a baja resolución proporcionados por las salidas de los GCM) para estimar \mathbf{y}_{sim} (los predictandos simulados en el futuro a alta resolución).

$$\mathbf{y}_{sim} \sim F(\mathbf{x}_{GCM}) \quad (2.1)$$

Adicionalmente, también se aplica el modelo sobre las series históricas de los GCM para permitir así realizar tareas de evaluación, identificación y corrección de sesgos.

Hay que destacar que las diferencias en la metodología de calibración tienen implicaciones en la simulación posterior. En PP existe correspondencia temporal o sincronía entre predictores y predictandos observados en el proceso de calibración (calibración por pares) y el modelo calibrado propaga el sesgo del GCM en la simulación, mientras que en MOS no existe correspondencia temporal entre predictores modelizados y predictandos observados (calibración por distribución), y el modelo calibrado elimina el sesgo de los GCM automáticamente en la simulación (Maraun y Widmann, 2018). Esto último es una ventaja de MOS frente a PP, sin embargo, la calibración por distribución solo tiene sentido en la práctica cuando predictor y predictando tienen la misma dimensión física (por ejemplo, cuando el predictando es la precipitación en alta resolución y el predictor es la precipitación en baja resolución), es decir, cuando el método es homogéneo. Los dos principales inconvenientes de los métodos MOS son la suposición de la estacionariedad de los sesgos del modelo en el futuro y el impacto que pueden tener los diferentes ajustes en las tendencias dadas por los GCM. Sin embargo, algunos métodos MOS son capaces de preservar las tendencias, mientras que otros ni siquiera asumen la estacionariedad.

2.1. Perfect Prognosis

Los métodos ESD tipo PP asumen las tres hipótesis siguientes:

- 1) Los predictores deben poder simularse de forma realista y sin sesgo en el clima presente. Además, en el contexto del cambio climático, los predictores deben responder a los forzamientos radiativos externos.
- 2) Los predictores deben contener suficiente información capaz de explicar una gran parte de la variabilidad local en todas las escalas temporales de interés.
- 3) El modelo estadístico debe tener una estructura adecuada, es decir, la influencia del predictor en el predictando debe estar razonablemente bien incluida en el modelo. La estructura debe permitir al menos extrapolaciones moderadas con respecto a las condiciones en que se haya calibrado.

Hay que tener en cuenta que las hipótesis 2) y 3) ya incluyen la hipótesis de invarianza temporal o estacionariedad: si se incluyen todos los predictores relevantes para la caracterización del cambio climático y su influencia en el predictando está adecuadamente modelizada, también más allá de los estados observados, el modelo será válido para simular un clima futuro.

Dentro de PP existen tres familias principales de métodos: función de transferencia (TF), análogos/tipos de tiempo (ANA/WT) y generadores de tiempo (WG).

La familia de métodos TF se basa en la existencia de relaciones estadísticas entre predictores a gran escala y predictandos locales (Sailor y Li, 1999; Wilby *et al.*, 2002), que se detectan y calibran en el presente y se aplican en simulaciones futuras. Estas relaciones pueden ir desde modelos lineales simples (LIN) como la Regresión Lineal Múltiple (MLR) y los Modelos Lineales Generalizados (GLM) (Wilby *et al.*, 2002) hasta relaciones no lineales complejas basadas en algoritmos de *machine learning* como las *Artificial Neural Networks* (ANN; McCulloch y Pitts, 1943; Rosenblatt, 1958), *Support Vector Machines* (SVM; Boser *et al.*, 1992; Cortes y Vapnik, 1995; Vapnik, 1995), *Randon Forest* (RF; Breiman, 2001) o *eXtreme Gradient Boost* (Chen y Guestrin, 2016).

La familia de análogos/tipos de tiempo (ANA/WT) se basa en la suposición de condiciones locales similares en situaciones sinópticas similares (Lorenz, 1969; Zorita y von Storch, 1999).

Estos métodos buscan condiciones sinópticas análogas que ocurrieron en el pasado, anteriormente a las proyectadas por los GCM, y su mayor limitación radica en su incapacidad para predecir valores fuera del rango observado. Los métodos ANA/WT han de aplicarse con mayor cautela, pues se ha comprobado que la aplicación de estos métodos subestima profundamente las tendencias de las proyecciones a futuro en la Península Ibérica (Gutiérrez et al., 2013). Una forma de solventar este problema es usar métodos híbridos que combinan métodos ANA/WT con métodos de regresión de la familia TF. Para una descripción más detallada de los métodos híbridos, consultar (Maraun y Widmann, 2018).

Por último, los Generadores de Tiempo (WG) son modelos estocásticos capaces de producir series sintéticas emparejando sus aspectos marginales y temporales con estadísticas climatológicas condicionadas a propiedades dadas por simulaciones de los GCM (Wilks y Wilby, 1999).

2.1.a. Función de transferencia

En este apartado se describen los siguientes modelos de función de transferencia (TF): modelos lineales (LIN) y modelos basados en *Machine Learning* (ML).

Modelos lineales

Dentro de los modelos lineales se distinguen los modelos de regresión lineal múltiple (MLR) y los modelos lineales generalizados (GLM).

Modelos de regresión lineal múltiple

La opción más sencilla para vincular estadísticamente el predictando con los predictores es formular un modelo de regresión lineal múltiple. Se relaciona un conjunto de M predictores, x_1, x_2, \dots, x_M , con el predictando, Y , y se supone que el modelo tiene la forma:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM} + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.2)$$

donde Y_i es la observación i -ésima del predictando para un conjunto de valores fijos de los predictores, $x_{i1}, x_{i2}, \dots, x_{iM}$; ε_i es el error aleatorio no observable asociado a Y_i y $\beta_0, \beta_1, \beta_2, \dots, \beta_M$ son $M+1$ parámetros desconocidos. Se supone que la distribución del error aleatorio es una normal de media nula $\varepsilon_i \sim N(0, \sigma^2)$, por tanto:

$$\begin{aligned} E(Y_i) &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM} \\ \text{Var}(Y_i) &= \sigma^2 \quad i = 1, 2, \dots, n \end{aligned} \quad (2.3)$$

Las n ecuaciones

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_M x_{1M} + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_M x_{2M} + \varepsilon_2 \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_M x_{nM} + \varepsilon_n \end{aligned} \quad (2.4)$$

pueden expresarse matricialmente del siguiente modo:

$$Y = X\beta + \varepsilon \quad (2.5)$$

donde

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1M} \\ 1 & x_{21} & x_{22} & \cdots & x_{2M} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nM} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_M \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.6)$$

con

$$\begin{aligned} Y &\sim N(X\beta, \sigma^2 I) \\ \varepsilon &\sim N(0, \sigma^2 I) \end{aligned} \quad (2.7)$$

Los parámetros desconocidos $\beta_0, \beta_1, \beta_2, \dots, \beta_M$ pueden estimarse mediante el método de mínimos cuadrados:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_M \end{bmatrix} = (X^t X)^{-1} X^t Y \quad (2.8)$$

Como $\varepsilon_i \sim N(0, \sigma^2)$, a los mismos parámetros estimados contenidos en el vector se puede llegar igualmente por el método de máxima verosimilitud. Por su parte, los valores estimados del predictando se pueden calcular del siguiente modo, definiéndose así el modelo estadístico F :

$$\hat{Y} = X \hat{\beta} = F(X) \quad (2.9)$$

Modelos lineales generalizados

También son modelos de regresión en los que existe una relación lineal entre predictores y predictando a través de una función de transferencia. El método de regresión GLM constituye una extensión de los modelos de regresión lineal tradicionales al permitir que el predictando tenga una distribución no normal y al relajar la suposición de varianza constante. A menudo se utilizan para analizar datos en los que la variable de salida es binaria (por ejemplo, éxito/fracaso), un conteo (por ejemplo, número de ocurrencias) o continua con una distribución sesgada. La estructura básica de un GLM involucra tres componentes: el predictor lineal, la función de enlace y la función de distribución. El predictor lineal es una función de una o más variables predictoras, que luego se transforma mediante la función de enlace. La función de enlace sirve para conectar el predictor lineal a la función de distribución, que determina la distribución de probabilidad de la variable de salida. Ejemplos de funciones de enlace comúnmente utilizadas en el método GLM son la función *logit* para datos binarios, la función logarítmica para datos de conteo y la función identidad para datos continuos. La elección de la función de enlace depende de la naturaleza de los datos y de la pregunta de investigación abordada. El método GLM se puede ajustar utilizando la estimación por máxima verosimilitud, que consiste en encontrar los valores de los parámetros que maximizan la verosimilitud de los datos observados dado el modelo. La calidad del ajuste del modelo se puede evaluar utilizando medidas como la desviación, los gráficos de residuos y las pruebas de bondad de ajuste. Los

GLM tienen muchas aplicaciones en diversos campos y ofrecen un marco flexible para analizar datos con diferentes tipos de variables de respuesta.

La esperanza en estos modelos toma la expresión:

$$g(E(Y_i)) = g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM} \quad i = 1, 2, \dots, n \quad (2.10)$$

donde $g()$ es la función de enlace; μ_i es el valor esperado del predictando, condicionado al vector de predictores $x_{i1}, x_{i2}, \dots, x_{iM}$.

La función de enlace $g()$ transfiere el valor esperado de tal forma que la influencia de los predictores puede suponerse lineal. La distribución de Y , condicionada a $x_{i1}, x_{i2}, \dots, x_{iM}$, se supone que es de tipo exponencial. La varianza del predictando es, por tanto, una función de la media predicha:

$$Var(Y_i) = V(\mu_i) \quad (2.12)$$

Dependiendo del predictando a simular, se elige una función de enlace u otra. Por ejemplo, la ocurrencia de precipitación puede modelizarse a través de una regresión logística, que usa un predictor continuo para predecir un predictando discreto y binario. El predictando toma valores 0 o 1, correspondientes a “seco” o “húmedo”, y los predictores usualmente describen la circulación y el transporte de humedad a gran escala. La media predicha es simplemente la probabilidad de día húmedo p_i :

$$p_i = \mu = g^{-1}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM}) = \frac{e^{(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM})}}{1 + e^{(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM})}} \quad (2.12)$$

La función *logit*

$$g(p_i) = \log\left(\frac{p_i}{1-p_i}\right) \quad (2.13)$$

asegura que p_i esté comprendida entre 0 y 1.

Pero también la influencia no lineal de los predictores puede modelizarse. A menudo, la intensidad de precipitación se supone que sigue una distribución de tipo gamma. La media de la distribución gamma puede calcularse mediante la siguiente función exponencial (dando lugar al método GLM-EXP):

$$\mu = g^{-1}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM}) = e^{(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_M x_{iM})} \quad (2.14)$$

La función de enlace de tipo exponencial permite añadir efectos multiplicativos de los predictores en los predictandos y asegura un valor esperado positivo de tal forma que la influencia de los predictores puede asumirse como lineal. El parámetro de forma γ se modeliza independientemente de los predictores. La varianza es simplemente:

$$\sigma^2 = \frac{\mu^2}{\gamma} \quad (2.15)$$

Esto es lo mismo que decir que el coeficiente de variación σ/μ se modeliza como una constante.

Modelos de regresión no lineal basados en *machine learning*

Existen multitud de técnicas de *machine learning* aplicables como métodos ESD. Aquí mostramos los dos ejemplos siguientes: las redes neuronales artificiales o *Artificial Neural Networks* (ANN) y *eXtreme Gradient Boosting* (XGB).

Artificial Neural Networks

En relación al funcionamiento y a la estructura de las ANN, hay que referirse necesariamente al perceptrón como unidad fundamental en el aprendizaje profundo. El perceptrón es una máquina de aprendizaje simple que puede tener múltiples entradas, cada una con un peso que indica su importancia, y genera una salida binaria (0 o 1). Cuando varios perceptrones se combinan, forman una ANN cuya capacidad predictiva depende principalmente de que tenga suficientes datos de calidad para el entrenamiento y capacidad de cálculo. La estructura de una ANN se define por la distribución de las neuronas. Una estructura comúnmente utilizada es la distribución en capas, conocida como MLP (*MultiLayer Perceptron*). En la figura 2.1 se muestra un ejemplo de una ANN de tres capas. La primera capa (i) es la capa de entrada, donde cada neurona maneja la información de una variable de entrada.

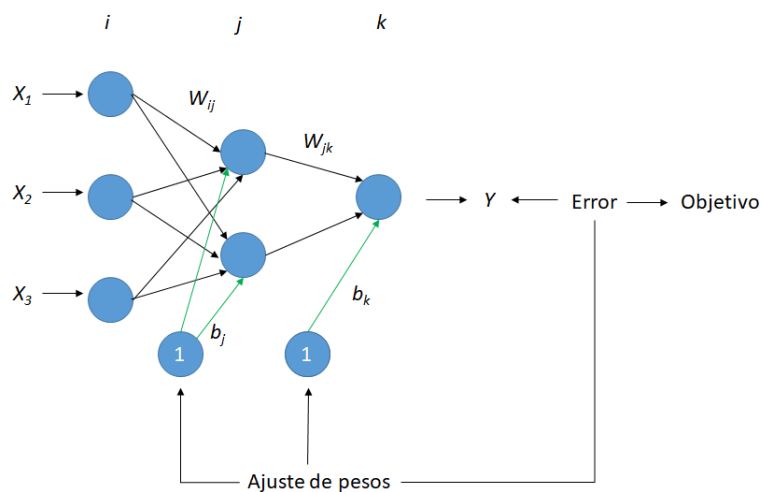


Figura 2.1: Esquema de una red neuronal artificial de tres capas.

En una aplicación de regionalización estadística, esta capa recibiría información de los predictores. La última capa (k) es la capa de salida, que puede tener una o varias neuronas para gestionar la salida. Si el objetivo es entrenar la ANN para determinar la relación entre una variable predictando y sus predictores, esta capa de salida tendría una sola neurona. La capa intermedia (j), llamada capa oculta, recibe información de la capa de entrada, la procesa y la dirige hacia la capa de salida. Por lo tanto, la información fluye desde la entrada hacia la salida, lo que hace que este tipo de MLP se denomine *feedforward*.

Generalizando, la salida de una ANN se obtiene (Haykin, 2009; García-Valero, 2021) de la siguiente manera:

$$Y_j = f(I_j) = f(\sum_{i=1}^n (W_{ij}A_i) + b_j) \quad (2.16)$$

donde W_{ij} es el peso de una determinada neurona en la capa anterior i con dicha neurona localizada en la capa j , A_i es la salida de una neurona determinada en la capa anterior, n representa el número de neuronas de la capa anterior, f es la función de activación y b_j es el peso de la neurona *bias* con la neurona j que recibe su señal. Una característica indispensable de la función de activación es que sea continua y diferenciable. Las funciones de activación pueden ser de tipo escalón, lineales y no lineales. Sin embargo, son estas últimas las que se utilizan principalmente para entrenar ANN que abordan problemas complejos. Entre las funciones no lineales más ampliamente empleadas se encuentran: la función logística, la tangente hiperbólica, la unidad lineal rectificadora (también conocida como RELU) y sus modificaciones.

Una forma común de entrenar una ANN es a través de la supervisión, utilizando conjuntos de datos que contienen valores de entrada y salida conocidos. En este proceso, las salidas generadas por la red neuronal se comparan con los valores conocidos, lo que permite calcular el error de la salida. Por lo general, se utiliza el error cuadrático medio, E , como medida de dicho error:

$$E = \frac{1}{n} \sum_{t=1}^n (T_t - A_t)^2 \quad (2.17)$$

n es el número de etapas temporales que se emplearán en la calibración de los pesos, A_t es la activación o salida proporcionada por la ANN para un determinado instante de tiempo t y T_t es la respuesta conocida en ese mismo instante de tiempo t . E está influenciado por los valores de los pesos, el conjunto de entrenamiento y la función de activación. El objetivo del entrenamiento es minimizar E , lo cual requiere ajustar cuidadosamente los pesos. Uno de los métodos comúnmente utilizados para este propósito es la propagación hacia atrás, también conocida como *backpropagation* en inglés. Este método recibe su nombre porque ajusta primero los pesos de las neuronas más cercanas a la capa de salida, y luego avanza hacia las capas superiores, ajustando los pesos en el camino hasta llegar a los pesos entre las neuronas de entrada y la primera capa oculta (figura 2.1). En este proceso de ajuste, se utiliza la técnica del gradiente descendente, que sigue la dirección negativa del gradiente de la función de error para alcanzar su valor mínimo. Por consiguiente, la función de error depende de cada uno de los pesos que definen las conexiones entre las neuronas. Siguiendo el esquema representado en la figura 2.1, esta dependencia puede expresarse como $E = f(W_{ij}, W_{jk}, b_j, b_k)$, de manera que el mínimo de la función de error se alcanza cuando se obtienen los valores mínimos de sus diferentes derivadas parciales.

$$\frac{dE}{dW} = \sum_{j=1}^{n_j} \sum_{i=1}^{n_i} \frac{\partial E}{\partial W_{ij}} + \sum_{k=1}^{n_k} \sum_{j=1}^{n_j} \frac{\partial E}{\partial W_{jk}} + \sum_{j=1}^{n_j} \frac{\partial E}{\partial b_j} + \sum_{k=1}^{n_k} \frac{\partial E}{\partial b_k} \quad (2.18)$$

En la ecuación anterior, n_i , n_j y n_k representan el número de neuronas en las capas i , j y k , respectivamente. W_{ij} denota los pesos entre una neurona específica en la capa i y otra neurona en la capa j , mientras que b_j y b_k se refieren a los pesos de sesgo aplicados a una neurona

determinada en la capa j y k , respectivamente. El método de *backpropagation* utiliza la regla de la cadena para calcular los diferentes gradientes establecidos en la ecuación. En este proceso, se considera que el peso entre dos neuronas depende de la señal de entrada recibida por la neurona en la capa anterior.

Para una descripción más detallada del método ANN y su aplicación en problemas de regionalización climática, véase la Nota Técnica 34 de AEMET (García-Valero, 2021).

eXtreme Gradient Boosting

XGB (Chen y Guestrin, 2016) consiste en la construcción de un modelo que combina múltiples árboles de decisión débiles (*weak learners*) para mejorar la precisión en la predicción de valores numéricos. XGB se basa en la optimización iterativa, donde se agregan árboles de decisión de forma secuencial. En cada iteración, se ajustan los pesos de los árboles anteriores en función de los errores residuales, permitiendo que el modelo se enfoque en los patrones más difíciles de capturar. Esto se logra mediante el uso del gradiente descendente, que actualiza los pesos de los árboles existentes y agrega nuevos árboles para minimizar el error residual. XGB utiliza una función de pérdida específica para problemas de regresión, como el error cuadrático medio o el error absoluto medio, a fin de medir la discrepancia entre las predicciones del modelo y las observaciones. Durante el entrenamiento, se optimiza esta función de pérdida para encontrar los mejores valores de los hiperparámetros del modelo, como la profundidad del árbol, la tasa de aprendizaje y la cantidad de árboles a utilizar. Además, XGB incluye técnicas de regularización para evitar el sobreajuste, como la limitación de la profundidad del árbol, la asignación de pesos a las hojas y la penalización de la complejidad del modelo. En resumen, XGB en problemas de regresión consiste en construir un modelo de árboles de decisión mediante la optimización iterativa de una función de pérdida específica, utilizando el gradiente descendente y las técnicas de regularización para mejorar la precisión de las predicciones y evitar el sobreajuste.

El método XGB tiene similitudes y diferencias con respecto a RF (*Random Forest*). Al igual que RF, XGB ajusta un conjunto de árboles. En RF se parte de árboles complejos y se reduce la varianza mediante regularización y agregación de árboles independientes, aspirando al óptimo sesgo/varianza reduciendo la varianza. Por el contrario, el método XGB realiza el camino opuesto: se fuerzan árboles deliberadamente sencillos con mucho sesgo y poca varianza, generándose un conjunto de árboles que no son independientes, sino secuenciales (su orden importa), aspirando al óptimo sesgo/varianza reduciendo el sesgo en vez de la varianza. El método XGB suele ser más ventajoso que el método RF debido a su mejor rendimiento, que se explica gracias a su enfoque de optimización iterativa, a la regularización (reducción de la complejidad de los árboles y la limitación de la profundidad, lo que ayuda a evitar el sobreajuste y mejora la generalización del modelo), a su control del óptimo sesgo/varianza y a su eficiencia computacional. Estas características lo convierten en una opción sólida para abordar problemas tanto de clasificación como de regresión.

A continuación se detalla la formulación matemática del método XGB (Chen y Guestrin, 2016; Osman *et al.*, 2021).

Sea un conjunto de datos D dado con n ejemplos y m características:

$$D = \{(\mathbf{x}_i, Y_i)\} \quad i = 1, 2, \dots, n \quad \mathbf{x}_i \in \mathbb{R}^m \quad Y_i \in \mathbb{R} \quad (2.19)$$

Un modelo de conjunto de árboles utiliza K funciones aditivas (el número de árboles) para predecir la salida:

$$\hat{Y}_k = \phi(\mathbf{x}_i) = \sum_{k=1}^K f_k(\mathbf{x}_i) \quad f_k \in \mathcal{F} \quad (2.20)$$

$$\mathcal{F} = \{f(\mathbf{x}) = w_{q(\mathbf{x})}\} \quad q : \mathbb{R}^m \rightarrow T, w \in \mathbb{R}^T \quad (2.21)$$

\mathcal{F} es el espacio de árboles de regresión (también conocido como CART). Aquí, q representa la estructura de cada árbol que asigna un ejemplo al índice de la hoja correspondiente. T es el número de hojas en el árbol. Cada f_k corresponde a una estructura de árbol independiente q y a pesos de las hojas w . A diferencia de los árboles de decisión, cada árbol de regresión contiene una puntuación continua en cada una de las hojas; se utiliza w_i para representar dicha puntuación en la i -ésima hoja.

Para resolver la ecuación (2.20) necesitamos minimizar la siguiente función objetivo:

$$\mathcal{L}(\phi) = \sum_i l(\hat{Y}_i, Y_i) + \sum_k \Omega(f_k) \quad \text{donde} \quad \Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (2.22)$$

l es una función de pérdida convexa diferenciable que mide la diferencia entre la predicción \hat{Y}_i y el objetivo Y_i . El segundo término penaliza la complejidad del modelo (es decir, las funciones del árbol de regresión). El término adicional de regularización ayuda a suavizar los pesos aprendidos finales para evitar el sobreajuste. λ es el hiperparámetro de regularización. El aumento de λ reduce fuertemente la influencia de las hojas pequeñas (las que tienen pocas observaciones), mientras que solo tiene un impacto reducido en las hojas más grandes (las que tienen muchas observaciones). El hiperparámetro γ permite podar los nodos que tienen una ganancia mínima. La poda se da cuando la diferencia entre la ganancia y γ es negativa. T representa el número de hojas del árbol y w es el peso de cada hoja. En los árboles de decisión, se utiliza el *boosting* para minimizar la función objetivo durante el entrenamiento del modelo, lo cual se logra agregando una nueva función, f , a medida que el modelo continúa entrenando. Entonces, en la t -ésima iteración se agrega una nueva función (árbol) de la siguiente manera:

$$\mathcal{L}^{(t)} = \sum_i^n l(Y_i, \hat{Y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (2.23)$$

$$\mathcal{L}_{split} = \frac{1}{2} \left(\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} + \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right) - \gamma \quad (2.24)$$

$$g_i = \partial_{\hat{Y}_i^{(t-1)}} l(Y_i, \hat{Y}_i^{(t-1)}) \quad (2.25)$$

$$h_i = \partial_{\hat{Y}_i^{(t-1)}}^2 l(Y_i, \hat{Y}_i^{(t-1)}) \quad (2.26)$$

El algoritmo comienza desde una única hoja y agrega iterativamente ramas al árbol. Siendo $I_j = \{i | q(x_i) = j\}$ el conjunto de instancias de la hoja j , denominamos I_L e I_R a los conjuntos de instancias de los nodos a izquierda y derecha, respectivamente, después de la división (*split*). Si $I = I_L \cup I_R$, entonces la reducción de la pérdida después de la división viene dada por (2.24).

Cabe destacar que en el método XGB, al ser de tipo TF (también aplicable a los métodos ya descritos: GLM y ANN), se calibra una relación estadística para cada punto del mapa, y que en el caso de precipitación se calibran dos relaciones: una para la ocurrencia de precipitación (usando 0.1 mm como umbral de precipitación) y otra para la intensidad de precipitación.

2.1.b. Tipos de Tiempo y Análogos

En esta sección se abordan los fundamentos de la familia de métodos de WT (tipos de tiempo) y ANA (análogos: Lorenz, 1969; Zorita y von Storch, 1999; para más información: Maraun y Widmann, 2018).

Tipos de Tiempo

Los métodos de tipos de tiempo (WT, del inglés *Weather Types*) se han usado ampliamente para caracterizar la evolución meteorológica espacial y temporal a escala sinóptica. La relación entre el flujo a gran escala y el tiempo a escala local es a menudo bastante compleja y no se modeliza de forma adecuada simplemente vinculando predictores y predictandos a través de un modelo lineal. La idea de la regionalización basada en los métodos WT es, por tanto, categorizar el tiempo local en un conjunto de tipos de tiempo X_k , $k = 1, \dots, K$, que incorporen los patrones de circulación sinóptica más importantes.

El método más simple consistiría en considerar la media μ de un tiempo local y para un tipo de tiempo dado. Para un día determinado i

$$\mu_i = \mu(X_{k(i)}) \quad (2.27)$$

donde k_i es un número que denota al tipo de tiempo en el día i . $\mu(X_k)$ se estima promediando el tiempo observado a lo largo de todos los días en los que el tipo de tiempo X_k tiene lugar. El método, sin embargo, es bastante básico en tanto que no tiene en cuenta toda la variabilidad existente dentro de cada tipo de tiempo en superficie. Así pues, esto solamente tiene aplicación para estimar la precipitación acumulada anual o mensual integrando sobre la precipitación media en la secuencia de los tipos de tiempo modelizados.

Algunas técnicas de regionalización más avanzadas predicen, por lo tanto, una distribución completa G_θ condicionada al tipo de tiempo:

$$y_i \sim G_\theta \text{ con } \Theta_i = \Theta(X_{k(i)}) \quad (2.28)$$

donde $\Theta(X_k)$ es un vector de parámetros de la distribución, estimado independientemente para todos los días donde el tipo de tiempo X_k tiene lugar. No obstante, en la mayoría de los casos, estos modelos incluyen un componente específico de Markov para dar cuenta de la memoria a corto plazo: lo que se conoce como la familia de métodos basada en generadores de tiempo y que no se aborda en la evaluación del presente documento.

Análogos

El método ANA también tiene su origen en la evaluación de la predecibilidad del tiempo y para realizar predicciones meteorológicas a corto plazo (Lorenz, 1969). Posteriormente, comenzó a utilizarse en estudios de regionalización. La regionalización por el método de ANA se basa en suponer que dos estados atmosféricos a gran escala muy parecidos darán lugar a estados de tiempo similares en superficie. El tiempo local en un día i , dado un determinado estado sinóptico X_i , se regionaliza del siguiente modo: primero se busca un análogo del estado

sinóptico $X_j = \text{análogo}(X_i)$ en un día i entre las observaciones y_j del tiempo local disponibles. Estas observaciones son usadas como predicción del tiempo local en el día i :

$$y_i = y_j \text{ con } j : X_j = \text{análogo}(X_i) \quad (2.29)$$

Los análogos se definen para un conjunto de predictores X , que normalmente se filtra para eliminar el ruido a pequeña escala, minimizando una función de coste como, por ejemplo, la distancia euclídea:

$$\text{análogo}(X_i) = X_j | \min(X_i - X_j)^2 \quad (2.30)$$

Un factor clave en el éxito en la aplicación del método de análogos, aunque trivial, es la existencia de análogos, esto es, estados donde $(X_i - X_j)^2$ no solo sea minimizado, sino también lo suficientemente pequeño. El método de análogos cuenta con dos características importantes: a) utiliza los patrones o tipos de tiempo y, por tanto, la estructura de covarianza espacial del tiempo a escala local se mantiene en los campos simulados, característica importante en los estudios hidrológicos y b) no hace ninguna suposición sobre la forma de la distribución de probabilidad de las variables regionalizadas, facilitando la construcción de escenarios para variables no distribuidas normalmente como sucede con la precipitación diaria. Por tanto, la habilidad del proceso de regionalización va a depender de la medida de analogía específica utilizada y, de forma más general, del proceso de selección global (para más detalle, véase Petisco, 2008 a, b; Petisco *et al.*, 2012).

El método ANA-LOC-kNN, empleado en la evaluación de métodos ESD descrita en el apartado 4, es un subtipo muy específico de análogos implementado en el *software* pyClim-SDM (Hernanz *et al.*, 2023a; Hernanz *et al.*, 2023b). El subtipo ANA-LOC-kNN está basado en analogía sinóptica (kNN: k análogos más cercanos (Hernanz *et al.*, 2021)) combinada con analogía local (LOC). La analogía local viene dada por la similitud (medida por la distancia euclídea) de un conjunto de predictores significativos para cada punto espacial y tipo de tiempo (Petisco de Lara, 2008a; Amblar-Francés *et al.*, 2017; Hernanz *et al.*, 2021).

A partir de análogos también se pueden generar métodos híbridos tales como MLR-ANA, que realiza regresión lineal múltiple basada en los días análogos previamente obtenidos. En primer lugar, los días análogos son seleccionados usando la distancia euclídea de los campos análogos como métrica de similitud y, a continuación, para cada punto, se calibra una regresión múltiple utilizando los días análogos seleccionados. Para más detalle, véase Petisco de Lara (2008b). El método MLR-ANA también está implementado en pyClim-SDM y se emplea en la evaluación de métodos ESD en el apartado 4.

2.2. Model Output Statistics

Se define como MOS cualquier método ESD que establezca una función de transferencia estadística entre el modelo GCM y los datos observados para aplicar luego esta función de transferencia en el post-proceso de los datos del modelo (Maraun y Widmann, 2018). Ya se apuntó que en la modelización del cambio climático, los resultados de los GCM y las observaciones no se encuentran generalmente en sincronía, y que la calibración solo puede ser por distribución, dando lugar a una serie de implicaciones:

- La calibración por distribución solo es posible entre variables dimensionalmente idénticas.
- Sin sincronía entre la simulación y la observación, los indicadores típicos de verificación de los modelos meteorológicos (como correlaciones, errores cuadráticos medios y otros scores) no son aplicables. Es decir, la verificación está lejos de ser trivial para los métodos MOS en climatología.

La mayoría de los métodos MOS consisten en esencia en hacer una corrección de sesgos, por tanto, en el texto se usará el término “corrección de sesgos” (*Bias Correction* o BC) como sinónimo de MOS homogéneos. Se define el sesgo en un instante t como la diferencia sistemática entre una propiedad modelizada, $\theta_{mod}(t)$, y una propiedad observada, $\theta_{obs}(t)$:

$$Bias_{\theta}(t) = \theta_{mod}(t) - \theta_{obs}(t) \quad (2.31)$$

Algunos métodos MOS se pueden aplicar como métodos de regionalización, pero también como métodos de corrección de sesgos tras haber realizado la regionalización con otros métodos.

Las principales hipótesis que asumen los métodos de corrección de sesgos son:

- 1) Los predictores requieren ser simulados verazmente. Para el clima presente, las simulaciones tienen que ser realistas más allá de los sesgos corregibles. En el contexto del cambio climático, los cambios en los predictores deben ser simulados con verosimilitud.
- 2) Los predictores han de ser representativos del predictando.
- 3) La función de transferencia tiene que tener una estructura apropiada, que sea aplicable bajo condiciones de cambio climático.

2.2.a. Delta change

El método *Delta Change* perturba las series temporales mediante la señal de cambio climático simulada. Hay dos variantes principales dependiendo de si la corrección, delta (Δ), se aplica de manera absoluta (aditiva, $y_{corregido} = y + \Delta$) o relativa (multiplicativa, $y_{corregido} = y \cdot \Delta$). La elección depende de la naturaleza intrínseca de la variable a corregir. De acuerdo con la literatura (véase Maraun y Widmann, 2018), la temperatura se corrige aditivamente, mientras que la precipitación se corrige de forma multiplicativa.

La proyección de temperatura y_{i+T}^f en un día futuro $i + T$ viene dada por la temperatura observada y_i^p en el momento i del registro observacional, pero perturbado o modificado por la señal aditiva de cambio climático de la temperatura media, que es la diferencia entre la media de la temperatura futura simulada $\overline{x^f}$ y la media de la temperatura histórica simulada $\overline{x^p}$:

$$y_{i+T}^f = y_i^p + \Delta = y_i^p + (\overline{x^f} - \overline{x^p}) \quad (2.32)$$

De forma análoga, la proyección de la precipitación y_{i+T}^f en $i + T$ en el futuro viene dada por la precipitación observada y_i^p en el día i del registro de observaciones, pero perturbada por la señal de cambio climático relativa de la precipitación media:

$$y_{i+T}^f = y_i^p \cdot \Delta = y_i^p \frac{\overline{x^f}}{x^p} \quad (2.33)$$

El método *Delta Change* se limita a utilizar los cambios simulados por un modelo dinámico. Este método, más que aplicar una corrección de sesgos a los datos de los modelos, modifica los datos observacionales usando una señal de cambio climático. Por ello, para distinguir entre *Delta Change* y los métodos de corrección de sesgos propiamente dichos, nos referimos a estos últimos como métodos directos.

2.2.b. *Quantile Mapping y Quantile Delta Mapping*

El método de corrección de sesgos denominado *Quantile Mapping* (QM) es uno de los métodos directos más comunes dentro de la familia MOS. El método ajusta diferentes cuantiles individualmente. Partiendo de valores históricos simulados y observados, x_i^p e y_i^p , en el periodo de calibración, se estiman las correspondientes funciones de distribución acumulada (FDA), $F_x(x)$ y $F_y(y)$. Los valores futuros simulados x_{i+T}^f , es decir, los cuantiles específicos de la distribución simulada, se mapean sobre los cuantiles observados correspondientes aplicando la siguiente fórmula:

$$y_{i+T}^f = F_y^{-1}(F_x(x_{i+T}^f)) \quad (2.34)$$

El método *Quantile Delta Mapping* (QDM, véase Cannon *et al.*, 2015) es una de las muchas variantes del método QM y es el método escogido para realizar las correcciones de sesgos en la evaluación descrita en el apartado 4. En el método QDM se aplica *Delta Change* a cada cuantil de las series observadas y simuladas de tal manera que las tendencias se conservan en todos los cuantiles, sin asumir en ningún momento la hipótesis de transferibilidad de los sesgos. En el caso aditivo, la corrección Δ será la diferencia entre los valores futuros simulados x_{i+T}^f y los cuantiles de los valores históricos simulados x_i^p . En el caso multiplicativo, Δ será el cociente entre los valores futuros simulados x_{i+T}^f y los cuantiles de los valores históricos simulados x_i^p . Los cuantiles de los valores históricos simulados x_i^p se pueden determinar fácilmente evaluando la función ECDF (*Empirical Cumulative Distribution Function*) para calcular los cuantiles acumulativos de x_{i+T}^f y luego utilizando esos cuantiles para calcular los cuantiles correspondientes en el conjunto de datos x_i^p . Como se ha descrito anteriormente, en el caso de la temperatura (caso aditivo) la corrección Δ se suma a cada cuantil de las observaciones históricas y_i^p , mientras que en el caso de la precipitación (caso multiplicativo) se multiplica por un factor Δ a cada cuantil de y_i^p .

3. EVALUACIÓN Y SELECCIÓN DE MODELOS CLIMÁTICOS GLOBALES

Existen multitud de modelos climáticos globales, lo cual resulta muy adecuado para poder caracterizar las incertidumbres en las proyecciones futuras. Sin embargo, el manejo de tantos datos puede resultar costoso para elaborar estudios de impactos y adaptación. Por lo tanto, se hace necesario realizar una selección de un subconjunto de modelos que permita al mismo tiempo caracterizar la incertidumbre futura y reducir el volumen de datos a explorar. Además, es bien sabido que la selección de modelos puede condicionar notablemente las conclusiones

alcanzadas, por lo que una buena evaluación y selección de modelos es un paso fundamental para obtener proyecciones de cambio climático que sean fiables y robustas.

El objetivo de la evaluación de los GCM llevada a cabo fue obtener un conjunto o *ensemble* asumible de modelos del CMIP6 que, disponiendo de las variables que se han estimado necesarias para aplicar los métodos ESD, ofreciesen buenos índices comparativos en el periodo histórico y conservasen de manera aproximada la señal de cambio (y su incertidumbre) en las proyecciones futuras para el escenario más extremo, el SSP5-8.5.

Para la evaluación se ha utilizado la herramienta interactiva [GCMeval](https://gcmeval.met.no/) (Parding *et al.*, 2020; <https://gcmeval.met.no/>) siguiendo los siguientes pasos.

3.1. Modelos evaluables en GCMeval y disponibilidad de variables en ESGF

GCMeval contempla 101 GCM del CMIP6 en su listado de modelos evaluables. Estos modelos se examinaron en el repositorio *Earth System Grid Federation* ([ESGF](#)) para comprobar si disponían o no de las variables con dato diario que se habían considerado necesarias (se listan en la tabla 3.1) para poder aplicar los distintos métodos ESD. El resultado de la comprobación en ESGF es que de los 101 GCM evaluables en GCMeval tan solo 57 disponen de los variables necesarias.

Variable necesaria (nomenclatura en ESGF)	Variable (valores diarios)
tasmax & tasmin / tas	temperatura máxima y mínima en superficie / temperatura en superficie
pr	precipitación acumulada
hus / hur	humedad específica / humedad relativa
psl	presión a nivel del mar
ua	viento zonal por niveles de presión
va	viento meridional por niveles de presión
ta	temperatura por niveles de presión

Tabla 3.1: Variables consideradas como necesarias para la aplicación de los métodos ESD. Si no están disponibles las temperaturas máxima y mínima superficiales, entonces será necesaria la temperatura en superficie. Si no está disponible la humedad específica, entonces será necesaria la humedad relativa. En cuanto a los predictores por niveles de presión (hus / hur, ua, va y ta), se requieren únicamente los niveles 850 hPa, 700 hPa y 500 hPa.

3.2. Evaluación de modelos junto a la selección de EURO-CORDEX

Tras haber identificado los modelos disponibles (“*ensemble* de oportunidad”), el objetivo es reducir ese *ensemble* a un subconjunto representativo del mismo en términos de cambio futuro. No obstante, como acción previa a esta subselección es conveniente descartar aquellos modelos que no consigan reproducir de forma satisfactoria ni siquiera el clima presente en base a unos criterios determinados. Con este objetivo, en este segundo paso se evaluaron los 101 modelos contemplados en GCMeval y se descartaron los 51 modelos con peores resultados. Para ello, se han usado los pesos de la tabla 3.2 para los índices comparativos, las variables (temperatura y precipitación), la región geográfica y los parámetros estacionales seleccionables.

Parámetro	Peso
Región principal: Sur de Europa/Mediterráneo [MED:13]	2
Región secundaria: Global	0
Temperatura	1
Precipitación	1
Año completo	1
DEF	1
MAM	1
JJA	1
SON	1
<i>Bias</i>	1
Correlación espacial	1
<i>Spatial sd ratio</i>	1
RMSE del ciclo anual	1

Tabla 3.2: Pesos seleccionados para los parámetros a evaluar con GCMeval

Los índices comparativos que emplea GCMeval subrayan diferentes aspectos sobre cómo se comportan los GCM y están calculados en el periodo de referencia 1981-2010. El sesgo o *bias* absoluto se calcula como la diferencia entre promedios del GCM particular y el conjunto de datos observacionales de referencia: ERA5 (Hersbach H. *et al.*, 2020) para temperatura y GPCP versión 2.3 (Adler *et al.*, 2003) para precipitación. La correlación espacial indica si el modelo y las observaciones tienen una estructura espacial similar. El ratio de las desviaciones típicas espaciales del GCM y de los datos de referencia ($\sigma_{GCM}/\sigma_{ref}$) establece una comparación entre las desviaciones típicas espaciales del GCM y las de los datos observacionales de referencia. La raíz del error cuadrático medio (RMSE) es la del ciclo medio anual del GCM con respecto al ciclo medio anual de los datos de referencia. Una vez calculados los índices comparativos, GCMeval establece una clasificación de modelos en base a los pesos definidos por el usuario tanto para los índices comparativos como para la región geográfica, las variables, la región geográfica y las estaciones. GCMeval calcula la clasificación por etapas: 1) cálculo de la clasificación para cada índice comparativo en cada región, variable y estación del año; 2) las clasificaciones se multiplican por sus respectivos pesos y se calcula la suma para cada GCM; 3) se calcula una nueva clasificación en base a la suma de los ranking con peso.

Al resultado de la búsqueda de los 50 mejores GCM con GCMeval se añadió el conjunto de los 6 modelos (CMCC-CM2-SR5_r1i1p1f1, CNRM-ESM2-1_r1i1p1f2, EC-Earth3-Veg_r1i1p1f1, MIROC6_r1i1p1f1, MPI-ESM1-2-HR_r1i1p1f1 y NorESM2-MM_r1i1p1f1) que se seleccionaron en el proyecto EURO-CORDEX como resultado del proceso de evaluación exhaustiva al que sometieron a todos los modelos del CMIP6. En el *ensemble* resultaron un total de 53 modelos (3 de los modelos de EURO-CORDEX se habían quedado

fuera de los 50 mejores según las métricas de GCMeval y se añadieron posteriormente) y se enumeran en la tabla 3.3 con su posición en la clasificación.

Modelo	Clasificación
ACCESS-CM2_r1i1p1f1	34
ACCESS-CM2_r2i1p1f1	23
ACCESS-CM2_r3i1p1f1	28
ACCESS-ESM1-5_r1i1p1f1	48.5
ACCESS-ESM1-5_r7i1p1f1	45.5
BCC-CSM2-MR_r1i1p1f1	22
CESM2-WACCM_r1i1p1f1	52
CMCC-CM2-SR5_r1i1p1f1	84
CNRM-CM6-1_r1i1p1f2	50.5
CNRM-CM6-1_r4i1p1f2	36
CNRM-CM6-1_r6i1p1f2	21
CNRM-ESM2-1_r1i1p1f2	35
CNRM-ESM2-1_r2i1p1f2	25
CNRM-ESM2-1_r3i1p1f2	20
CNRM-ESM2-1_r4i1p1f2	14
CNRM-ESM2-1_r5i1p1f2	17
EC-Earth3-Veg_r1i1p1f1	9
EC-Earth3-Veg_r2i1p1f1	10
EC-Earth3-Veg_r3i1p1f1	12
EC-Earth3-Veg_r4i1p1f1	31
EC-Earth3-Veg_r6i1p1f1	27
EC-Earth3_r1i1p1f1	16
EC-Earth3_r4i1p1f1	11
GFDL-ESM4_r1i1p1f1	15
GISS-E2-1-G_r1i1p1f2	29
GISS-E2-1-G_r1i1p3f1	19
GISS-E2-1-G_r1i1p5f1	32
GISS-E2-1-G_r2i1p3f1	26
GISS-E2-1-G_r3i1p3f1	30
GISS-E2-1-G_r4i1p3f1	40
GISS-E2-1-G_r5i1p3f1	24
HadGEM3-GC31-LL_r1i1p1f3	33
HadGEM3-GC31-LL_r2i1p1f3	13
HadGEM3-GC31-LL_r3i1p1f3	18
HadGEM3-GC31-LL_r4i1p1f3	39
HadGEM3-GC31-MM_r1i1p1f3	6
HadGEM3-GC31-MM_r2i1p1f3	5
HadGEM3-GC31-MM_r3i1p1f3	3
HadGEM3-GC31-MM_r4i1p1f3	8
IITM-ESM_r1i1p1f1	37
KACE-1-0-G_r1i1p1f1	38

MIROC6_rli1p1f1	78
MPI-ESM1-2-HR_rli1p1f1	4
MPI-ESM1-2-HR_r2i1p1f1	7
MPI-ESM1-2-LR_r2i1p1f1	41
MPI-ESM1-2-LR_r8i1p1f1	47
MRI-ESM2-0_rli1p1f1	2
MRI-ESM2-0_rli2p1f1	1
NorESM2-MM_rli1p1f1	79
UKESM1-0-LL_r2i1p1f2	43
UKESM1_0_LL_r3i1p1f2	42
UKESM1_0_LL_r4i1p1f2	44
UKESM1_0_LL_r8i1p1f2	53

Tabla 3.3: *Ensemble* de modelos que incluye a los 50 mejores de GCMeval y la selección de EURO-CORDEX. Se incluye la posición de cada uno de ellos en la clasificación sobre los 101 modelos contemplados en GCMeval de acuerdo a los pesos de la tabla 3.2.

3.3. Selección definitiva de modelos

El *ensemble* de la tabla 3.3 se tomó como referencia para cuantificar la señal de cambio a futuro en el escenario SSP5-8.5, así como su incertidumbre (dada por la dispersión del *ensemble*), que se pretende mantener en la selección final de modelos. A continuación, se escogieron los modelos pertenecientes al *ensemble* de la tabla 3.3 que, además, contenían todas las variables necesarias e indicadas en la tabla 3.1. El resultado es la selección de la tabla 3.4. No obstante, se filtró todavía más esta selección escogiendo únicamente un modelo por familia a fin de que no hubiera familias con una representación excesiva en la selección final. Así, la selección definitiva con el mejor modelo de cada familia se incluye en la tabla 3.5. La comparativa de su señal de cambio (y su incertidumbre) a futuro con respecto al *ensemble* de la tabla 3.3 (el *ensemble* de referencia) se muestra en la figura 3.1. En ella se observa cómo los *boxplots* asociados a la selección de la tabla 3.5, son capaces de recoger la incertidumbre de los *boxplots* azules correspondientes al *ensemble* de referencia, tanto para temperatura como para precipitación.

Modelo	Clasificación
ACCESS-CM2_rli1p1f1	34
ACCESS-ESM1-5_rli1p1f1	40
ACCESS-ESM1-5_r7i1p1f1	41
CMCC-CM2-SR5_rli1p1f1	53
CNRM-CM6-1_rli1p1f2	44
CNRM-ESM2-1_rli1p1f2	35
EC-Earth3-Veg_rli1p1f1	9
EC-Earth3-Veg_r2i1p1f1	10
EC-Earth3-Veg_r3i1p1f1	12
EC-Earth3-Veg_r4i1p1f1	37
EC-Earth3_rli1p1f1	19
EC-Earth3_r4i1p1f1	11
IITM-ESM_rli1p1f1	32
KACE-1-0-G_rli1p1f1	43
MIROC6_rli1p1f1	51

MPI-ESM1-2-HR_r1i1p1f1	6
MPI-ESM1-2-HR_r2i1p1f1	8
MPI-ESM1-2-LR_r2i1p1f1	47
MPI-ESM1-2-LR_r8i1p1f1	50
MRI-ESM2-0_r1i1p1f1	2.5
NorESM2-MM_r1i1p1f1	52
UKESM1-0-LL_r2i1p1f2	42
UKESM1-0-LL_r3i1p1f2	46
UKESM1-0-LL_r4i1p1f2	49
UKESM1-0-LL_r8i1p1f2	48

Tabla 3.4: Modelos del *ensemble* de la tabla 3.3 que, además, contienen todas las variables necesarias para aplicar los métodos ESD.

Modelo	Clasificación	Desarrollador del GCM
ACCESS-CM2_r1i1p1f1	34	CSIRO y Bureau of Meteorology (BoM), Australia
CMCC-CM2-SR5_r1i1p1f1	53	Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italia
CNRM-ESM2-1_r1i1p1f2	35	Centre National de Recherches Meteorologiques, Francia
EC-Earth3-Veg_r1i1p1f1	9	EC-Earth Consortium, Europa
IITM-ESM_r1i1p1f1	32	Centre for Climate Change Research, Indian Institute of Tropical Meteorology, India
KACE-1-0-G_r1i1p1f1	43	National Institute of Meteorological Sciences/Korea Meteorological Administration, Climate Research Division, Corea del Sur
MIROC6_r1i1p1f1	51	JAMSTEC (Japan Agency for Marine-Earth Science and Technology), Japón
MPI-ESM1-2-HR_r1i1p1f1	6	Max Planck Institute (MPI) for Meteorology, Alemania
MRI-ESM2-0_r1i1p1f1	2.5	Meteorological Research Institute, Tsukuba, Japón
NorESM2-MM_r1i1p1f1	52	Consortio (CICERO, MET-Norway, NERSC, NILU, UiB, UiO y UNI), Noruega
UKESM1-0-LL_r2i1p1f2	42	Met Office Hadley Centre, Reino Unido

Tabla 3.5: *Ensemble* definitivo de modelos tras seleccionar el mejor modelo por familia según su posición en la clasificación.

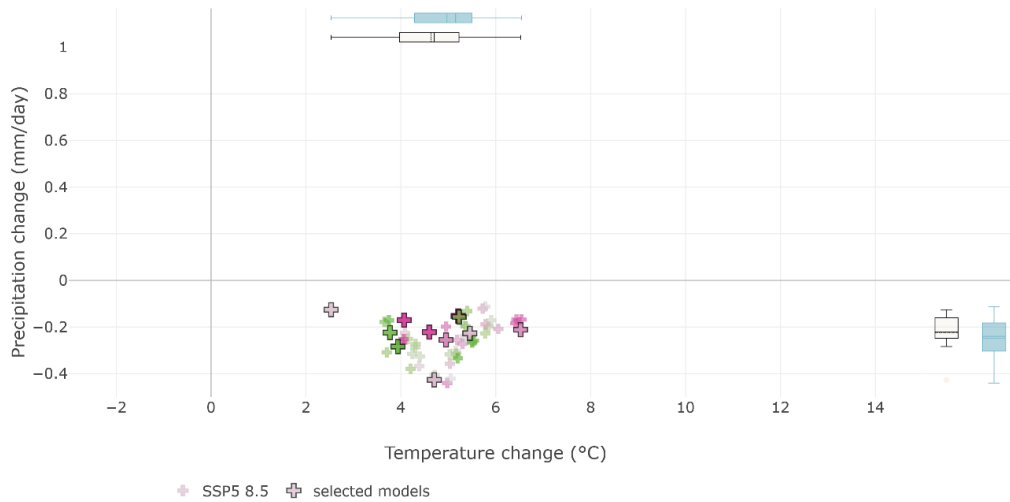


Figura 3.1: Los *boxplots* azules son los correspondientes al *ensemble* de referencia (tabla 3.3), mientras que los *boxplots* negros se corresponden con la selección de modelos de la tabla 3.5. Las cruces representan a los modelos individuales del *ensemble* de referencia, con colores verdes para aquellos mejor posicionados en el ranking y magentas para los peores. Los modelos de la selección de la tabla 3.5 tienen su cruz correspondiente con el borde grueso.

4. EVALUACIÓN DE MÉTODOS ESD

4.1. Introducción

Los métodos ESD usados en AEMET en la regionalización de los GCM de CMIP5 en el marco del Quinto Informe de Evaluación del IPCC fueron los siguientes:

Métodos ESD usados en la regionalización de los modelos de CMIP5 para cada predictando	
Temperatura máxima diaria	MLR, MLR-ANA y ANN
Temperatura mínima diaria	MLR, MLR-ANA y ANN
Precipitación acumulada diaria	GLM-EXP, ANA-LOC-kNN y ANN

Tabla 4.1, métodos ESD usados en la regionalización de los modelos de CMIP5.

Para la regionalización de la precipitación de los GCM de CMIP6 en el marco del Sexto Informe de Evaluación del IPCC se propone el método de *machine learning* XGB como nuevo método candidato a sustituir a los métodos anteriores en la regionalización de los GCM de CMIP6 por parte de AEMET. Para las temperaturas, máxima y mínima, no se propone ningún método alternativo a los anteriormente aplicados en la regionalización de los GCM de CMIP5 en el marco del Quinto Informe de Evaluación del IPCC. Nuevas metodologías de *machine learning* como XGB y de *deep learning* como CNN (*Convolutional Neural Network*), a pesar del buen rendimiento que ofrecen dentro de su rango de calibración, al extrapolar pueden conducir a grandes errores y para variables con una señal importante de cambio y una fuerte relación lineal con predictores, como en el caso de la temperatura superficial (véase Hernanz *et al.*, 2023c),

métodos lineales más simples pueden ser más apropiados y capaces de conducir a conclusiones más realistas.

En la evaluación se aplicó a cada uno de los métodos una corrección de sesgos tras la regionalización de tipo QDMs (el método de corrección de sesgos fue aplicado por estación del año, por lo que se designa añadiendo una “s” final a la sigla). Esta corrección de sesgos, QDMs, es una versión mejorada respecto a la versión de (Cannon *et al.*, 2015), pues se le ha añadido el factor de corrección de las tendencias de los valores medios propuesto por (Pierce *et al.*, 2015).

En la tabla 4.2 se listan los métodos de regionalización evaluados finalmente para cada predictando (temperatura máxima diaria, temperatura mínima diaria y precipitación acumulada diaria).

Métodos ESD (con y sin corrección de sesgos estacional) evaluados para la regionalización de los modelos de CMIP6	
Temperatura máxima diaria	MLR
	MLR+QDMs
	MLR-ANA
	MLR-ANA+QDMs
	ANN
	ANN+QDMs
Temperatura mínima diaria	MLR
	MLR+QDMs
	MLR-ANA
	MLR-ANA+QDMs
	ANN
	ANN+QDMs
Precipitación acumulada diaria	GLM-EXP
	GLM-EXP+QDMs
	ANA-LOC-kNN
	ANA-LOC-kNN+QDMs
	ANN
	ANN+QDMs
	XGB
	XGB+QDMs

Tabla 4.2: Métodos ESD evaluados para la regionalización de los modelos de CMIP6

Para realizar la evaluación se utilizó el *software* pyClim-SDM (Hernanz *et al.*, 2023a; Hernanz *et al.*, 2023b) desarrollado en el Área de Evaluación y Modelización del Clima de AEMET. Este *software* cuenta con un amplio abanico de métodos de regionalización estadística de última generación de diferentes familias y una interfaz gráfica fácil de usar, sencilla e intuitiva. pyClim-SDM ha sido desarrollado con la opción para ejecutarse paralelizando los cálculos, en computadoras de alto rendimiento (HPC) bajo el popular administrador de carga de trabajo SLURM (Yoo , Jette y Grondona, 2003). pyClim-SDM está preparado para regionalizar los datos diarios de las siguientes variables de superficie: temperatura media, máxima y mínima, precipitación, componentes de viento zonal y meridional, humedad relativa y cobertura de nubes. Tanto su código como su manual de usuario están disponibles en <https://github.com/ahernanz/pyClim-SDM>. Los datos de reanálisis utilizados en la evaluación fueron los de ERA5 (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>). Los datos observacionales utilizados fueron las series homogeneizadas asociadas a una selección de la red de estaciones de observación de AEMET. En Península y Baleares se

utilizaron datos de 3460 estaciones de temperatura y 5803 estaciones de precipitación, mientras que en Canarias se usaron datos de 244 estaciones de temperatura y 460 de precipitación. Además de los predictandos, se han obtenido resultados a alta resolución de una serie de índices CLIMDEX (<https://www.climdex.org/>).

El periodo utilizado tanto para la calibración como para la evaluación es de 1979 a 2020. La independencia entre los datos de calibración (*training*) y evaluación (*testing*) se logra dividiendo el periodo completo en cinco hojas o subconjuntos del mismo tamaño aproximadamente. Luego, el modelo se entrena 5 veces, utilizando cada vez 4 hojas como conjunto de entrenamiento y la hoja restante como conjunto de evaluación. El proceso se repite 5 veces, de modo que cada hoja se utiliza como conjunto de evaluación solamente una vez. Después de las 5 iteraciones, se calcula el rendimiento promedio del método ESD, lo que proporciona una evaluación más robusta y menos sesgada del desempeño del método en comparación a la realizada con una única división de los datos en dos conjuntos: calibración y evaluación. Este esquema conocido como *k-folds* ($k = 5$ en nuestro caso) o validación cruzada *k-fold* es una técnica comúnmente utilizada en *machine learning* para evaluar el rendimiento de métodos ESD y evitar el sobreajuste. La validación cruzada *k-fold* es particularmente útil cuando el tamaño del conjunto de datos es limitado, ya que permite aprovechar al máximo la información disponible para entrenar y evaluar el modelo. Además, ayuda a proporcionar una estimación más precisa de la capacidad predictiva del modelo partiendo de datos con los que no ha sido entrenado.

Los métodos ESD han sido entrenados siguiendo el esquema *5-folds* en el periodo 1979–2020 sobre datos pseudo-observacionales de reanálisis para realizar la evaluación de los métodos. También se utilizó el esquema *5-folds* para regionalizar los GCM seleccionados, calculando así las proyecciones climáticas a alta resolución para el periodo 2015-2100 bajo el escenario SSP5-8.5 y analizar sus tendencias. Se ha utilizado el periodo de referencia 1981-2010 para estandarizar los predictores y calcular la señal de cambio.

En los siguientes apartados se muestran y se analizan los resultados más destacables de la evaluación (ciclo anual, sesgo, correlación, sesgo de la varianza, correlación espacial, raíz del error cuadrático medio) tanto de los predictandos regionalizados como de algunos de los índices CLIMDEX seleccionados (TXm, TXn, TXx, TNm, TNn, TNx, WSDI, CSDI, PRCPTOT, R01, R95p, SDII, CDD y CWD; véase tabla 4.3).

Índice CLIMDEX	Descripción
TXm	Media anual de la temperatura máxima diaria
TXn	Mínimo anual de la temperatura máxima diaria
TXx	Máximo anual de la temperatura máxima diaria
TNm	Media anual de la temperatura mínima diaria
TNn	Mínimo anual de la temperatura mínima diaria
TNx	Máximo anual de la temperatura mínima diaria
WSDI	Duración de ola de calor: número anual de días con al menos 6 días consecutivos con TX > percentil 90
CSDI	Duración de ola de frío: número anual de días con al menos 6 días consecutivos con TN < percentil 10

PRCPTOT	Precipitación total anual en la que se consideran solo los días lluviosos (aquellos con precipitación ≥ 1 mm)
R01	Número de días en el año con precipitación ≥ 1 mm, el número anual de días lluviosos
R95p	Precipitación total anual en la que se consideran solo los días muy lluviosos (aquellos en que se supera el percentil 95 de los días con precipitación ≥ 1 mm)
SDII	Índice simple de intensidad de precipitación en días lluviosos (precipitación media en los días lluviosos)
CDD	Duración máxima del periodo seco: número máximo de días consecutivos con precipitación < 1 mm
CWD	Duración máxima del periodo húmedo: número máximo de días consecutivos con precipitación ≥ 1 mm

Tabla 4.3: Índices CLIMDEX empleados en la evaluación.

4.2. Evaluación en Península y Baleares

4.2.a. Ciclo anual

Las figuras 4.1 y 4.2 muestran la comparativa de los métodos de regionalización estadística y las observaciones (OBS) en las estaciones de observación de AEMET en Península y Baleares para el ciclo anual promediado espacialmente de la temperatura máxima diaria y de la temperatura mínima diaria, respectivamente. Se constata que todos los métodos evaluados reproducen el ciclo anual observado en ambas temperaturas con gran aproximación.

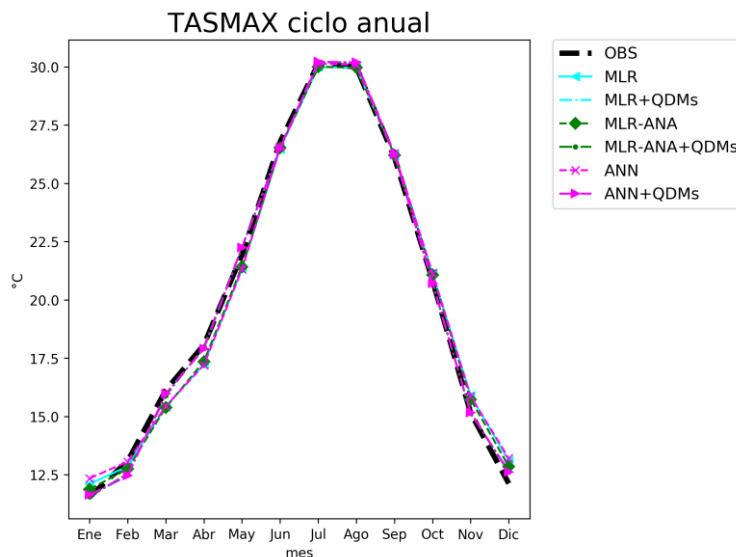


Figura 4.1: Ciclo anual de la temperatura máxima diaria.

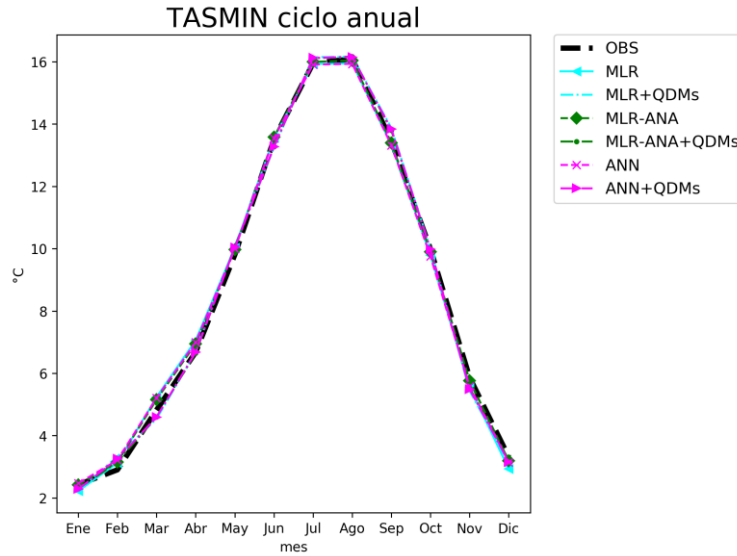


Figura 4.2: Ciclo anual de la temperatura mínima diaria.

La figura 4.3 muestra la misma comparativa de los métodos de regionalización estadística, pero, en este caso, para el ciclo anual de la precipitación acumulada diaria. El nuevo método candidato con corrección de sesgos, XGB+QDMs, presenta una aproximación al ciclo anual de precipitación observada notablemente superior a la que ofrecen los demás métodos que, a diferencia de lo que sucedía con las temperaturas, presentan desviaciones importantes respecto a las observaciones. XGB y XGB+QDMs reproducen el ciclo anual de forma muy similar, sin embargo, XGB+QDMs se aproxima más a las observaciones en la segunda mitad del año y sobre todo en otoño, el periodo estacional en el que XGB se distancia más de las observaciones.

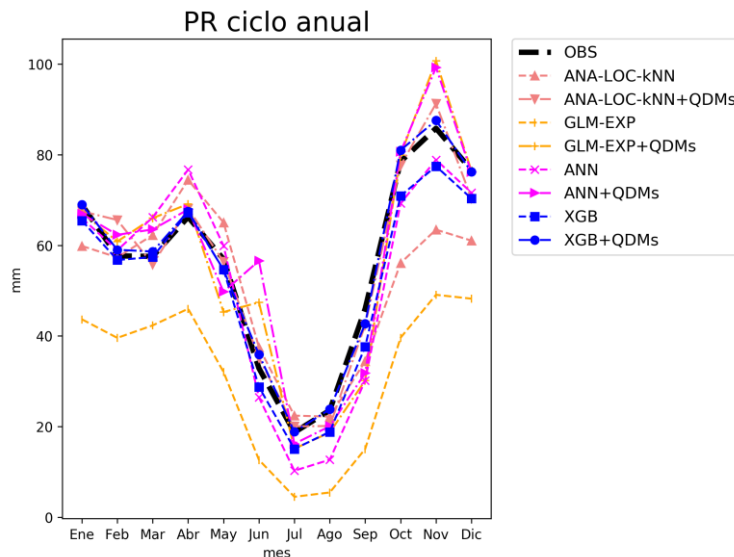


Figura 4.3: Ciclo anual de la precipitación acumulada diaria.

4.2.b. Sesgo

En este apartado se analizan comparativamente los sesgos anuales obtenidos para los índices CLIMDEX asociados a las temperaturas y a las precipitaciones con los métodos de regionalización considerados (los sesgos estacionales se han incluido en el anexo A). La figura 4.4 representa mediante *boxplots* los sesgos asociados a los índices TXm y TNm. En ambos índices, los diferentes métodos ofrecen sesgos muy pequeños (valor medio en general por debajo de 0.1 °C). Tras aplicar QDMs todos los métodos proporcionan los sesgos muy reducidos.

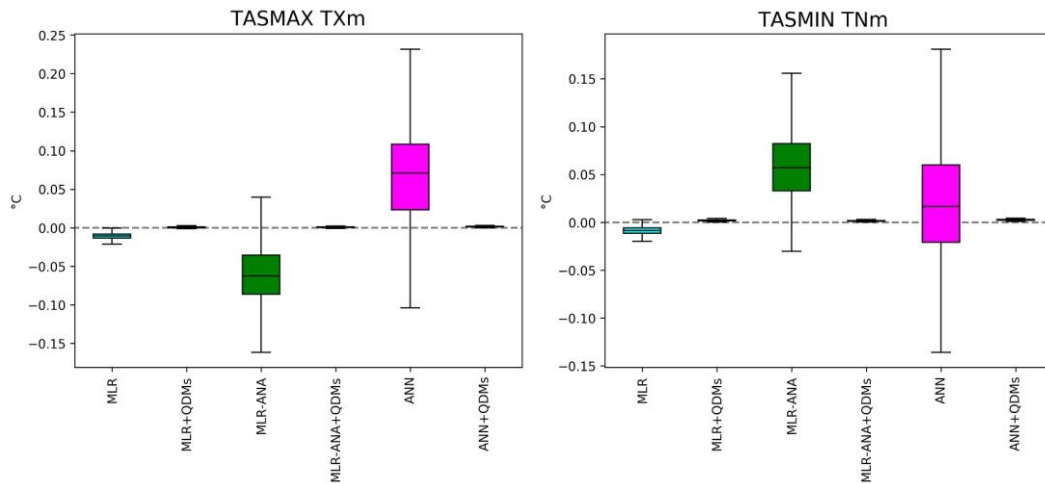


Figura 4.4. Sesgos para los índices CLIMDEX: TXm y TNm. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

La figura 4.5 muestra los sesgos asociados a los índices: TXn y TNn.

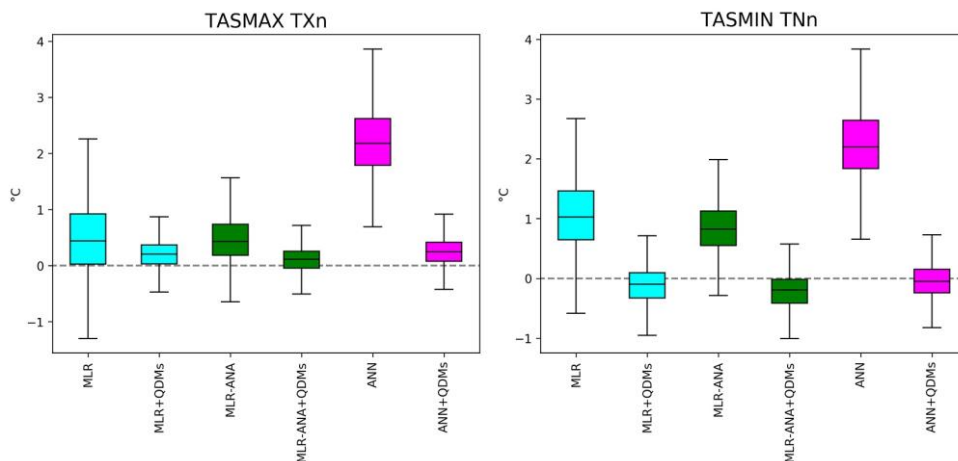


Figura 4.5, sesgos para los índices CLIMDEX: TXn y TNn. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

La figura 4.6 muestra los sesgos asociados a los índices: TXx y TNx.

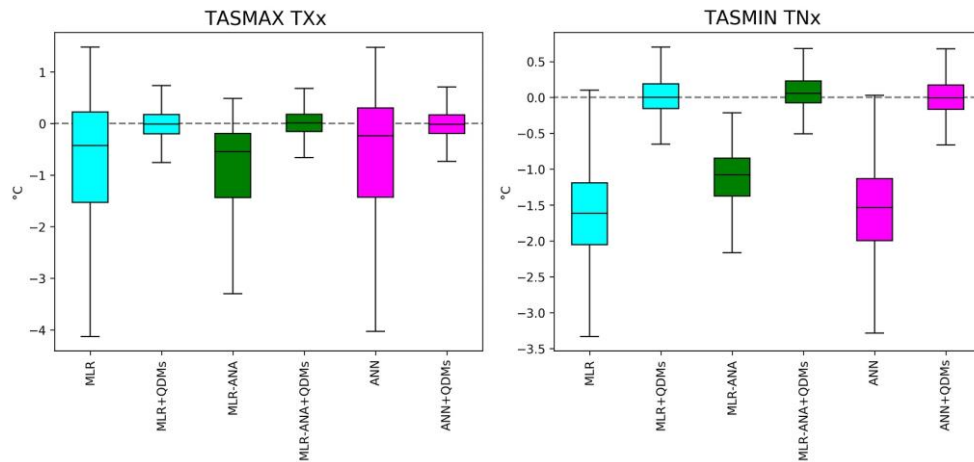


Figura 4.6, sesgos para los índices CLIMDEX: TXx y TNx. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

Al comparar los sesgos en los valores medios (TXm y TNm, figura 4.4) y en los valores extremos (TXn, TNn, TXx y TNx, figuras 4.5 y 4.6) destaca que, mientras que en los valores medios todos los métodos ofrecen sesgos muy pequeños, en los valores extremos ya aparecen sesgos importantes ($\sim 1\text{ }^{\circ}\text{C}$, un orden de magnitud superior) y diferencias notables entre los métodos. En los índices de las figuras 4.5 y 4.6, los métodos ofrecen los mejores resultados tras aplicar QDMs. También se pone de manifiesto que MLR-ANA+QDMs mejora ligeramente los resultados de MLR+QDMs y ANN+QDMs en TXn y TXx. No sucede así en TNn y TNx, pero los sesgos resultantes son similares.

La figura 4.7 muestra los sesgos asociados a los índices correspondientes a los aspectos temporales de las temperaturas: WSDI y CSDI.

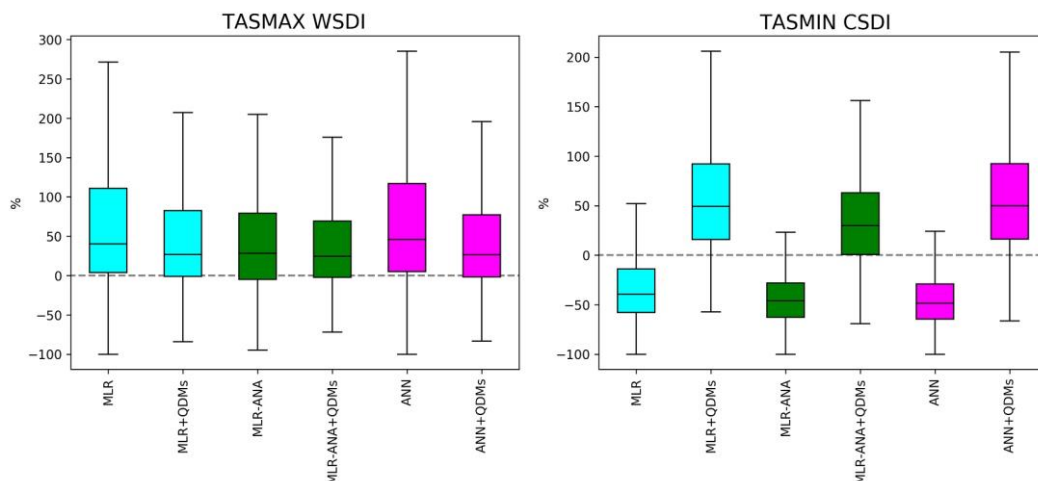


Figura 4.7: Sesgos para los índices CLIMDEX: WSDI y CSDI. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

En WSDI, las reducciones de sesgos son moderadas tras aplicar QDMs. El menor sesgo viene dado por MLR-ANA+QDMs. En CSDI, la aplicación de QDMs convierte los sesgos negativos en positivos y aumenta el tamaño de los *boxplots*. MLR-ANA+QDMs mejora a MLR+QDMs y ANN+QDMs.

La figura 4.8 representa los sesgos asociados a los índices CLIMDEX de precipitación: PRCPTOT, R01, R95p y SDII.

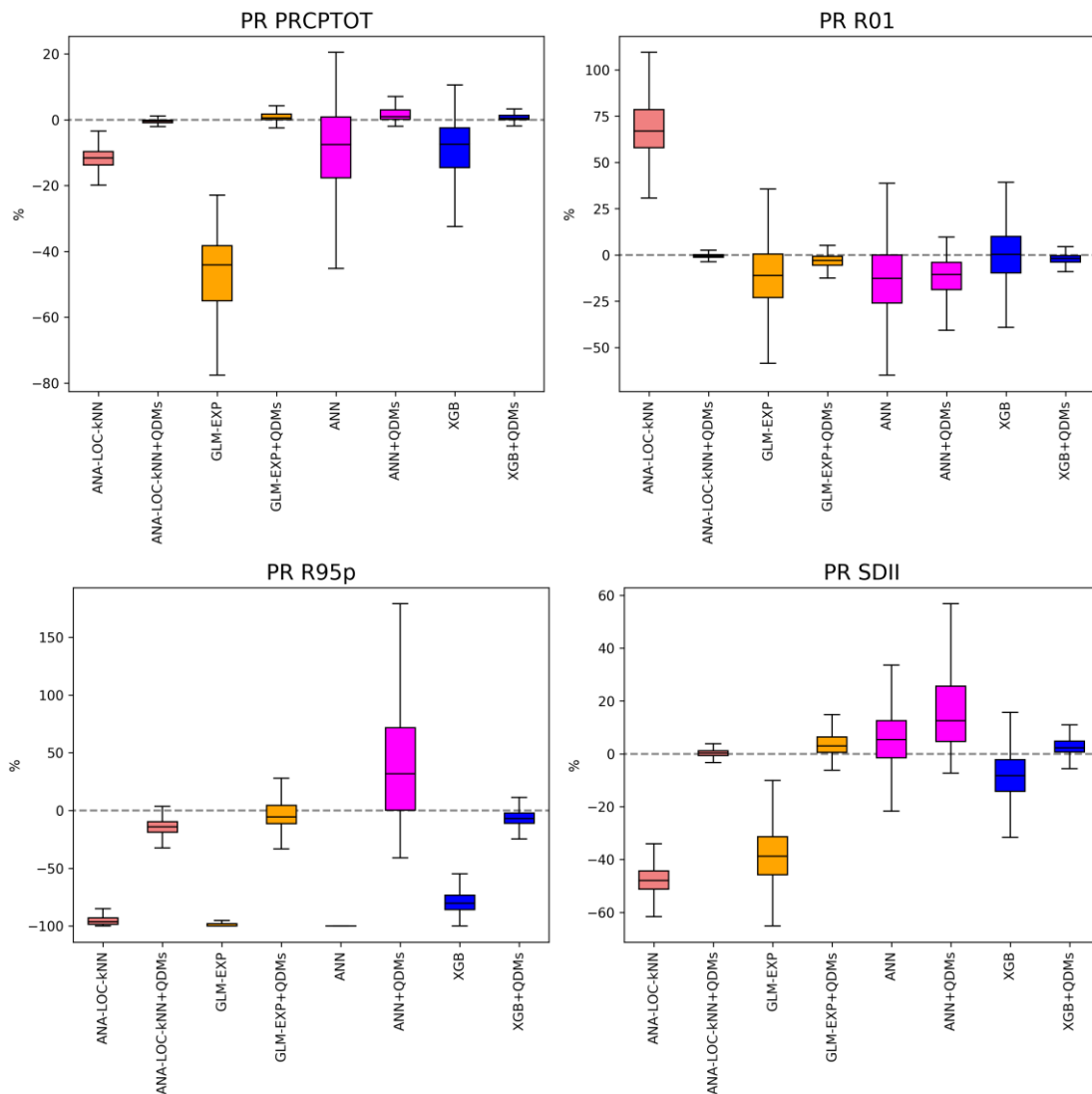


Figura 4.8: Sesgos para los índices CLIMDEX: PRCPTOT, R01, R95p y SDII. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

En todos los índices de la figura 4.8, ANA-LOC-kNN+QDMs y XGB+QDMs son los métodos que reducen más los sesgos. Es destacable la mejora de XGB+QDMs respecto a ANA-LOC-kNN+QDMs en R95p, es decir, en la simulación de las precipitaciones extremas. En los demás índices ambos métodos dan lugar a sesgos parecidos.

La figura 4.9 representa los sesgos asociados a los índices CLIMDEX de precipitación: CDD y CWD. Como sucedía en los índices de la figura 4.8, ANA-LOC-kNN+QDMs y XGB+QDMs son los métodos que más reducen los sesgos en CDD y CWD.

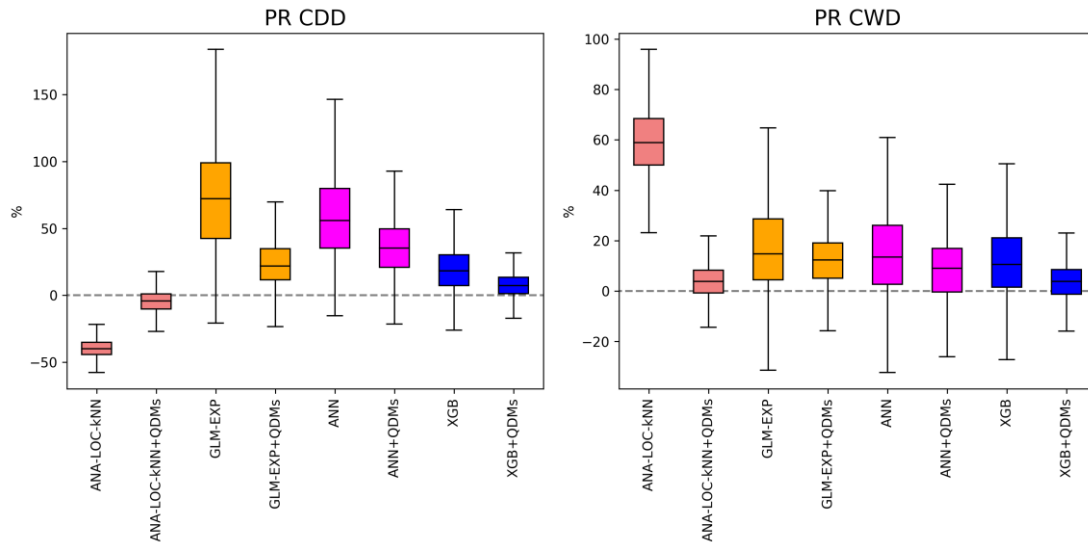


Figura 4.9: Sesgos para los índices: CDD y CWD. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

4.2.c. Correlación

La figura 4.10 representa el coeficiente de correlación obtenido por cada método en la temperatura máxima diaria y la temperatura mínima diaria.

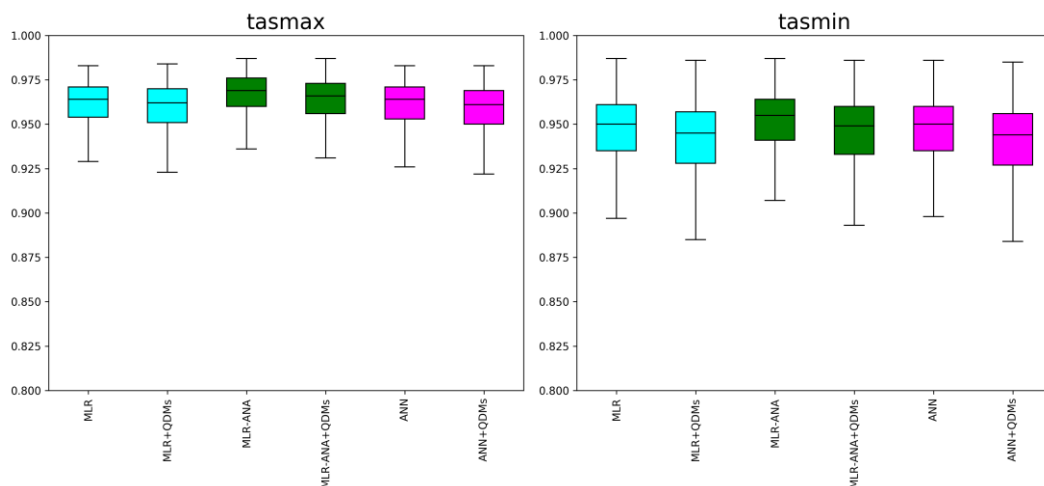


Figura 4.10: Comparativa del coeficiente de correlación obtenido con los diferentes métodos en temperaturas. Cada *boxplot* se refiere a un método ESD determinado y contiene los coeficientes de correlación de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

En temperaturas, a todos los métodos les corresponde un coeficiente de correlación elevado, ligeramente por debajo de 1. En temperatura mínima diaria las correlaciones son ligeramente inferiores respecto a las de la temperatura máxima diaria. En cualquier caso, estas se sitúan por encima de 0.875.

La figura 4.11 representa el coeficiente de correlación obtenido por cada método en la precipitación acumulada diaria.

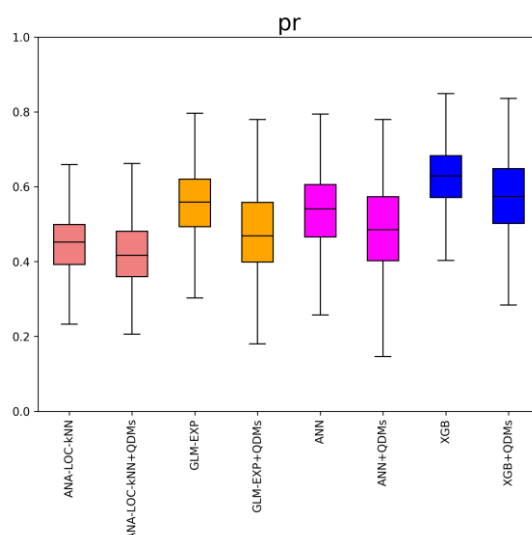


Figura 4.11: Comparativa del coeficiente de correlación obtenido con los diferentes métodos en la precipitación acumulada diaria. Cada *boxplot* se refiere a un método ESD determinado y contiene los coeficientes de correlación de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

En precipitación, las correlaciones son menores que en las temperaturas. Además, en este caso, se aprecia que tanto XGB como XGB+QDMs proporcionan unos resultados ligeramente mejores que los demás métodos.

4.2.d. Varianza

La figura 4.12 representa el sesgo de la varianza obtenida por cada método en la temperatura máxima diaria y la temperatura mínima diaria. Para estas temperaturas, todos los métodos con QDMs reducen el sesgo de la varianza de manera similar.

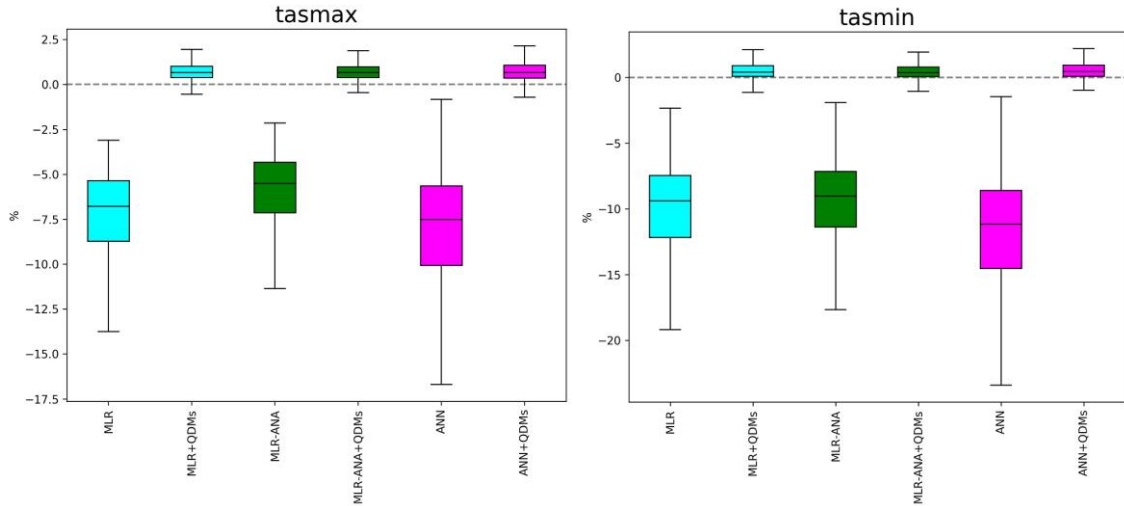


Figura 4.12: Comparativa del sesgo de la varianza obtenida con los diferentes métodos en temperaturas. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de la varianza de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

La figura 4.13 representa el sesgo de la varianza obtenida por cada método en la precipitación acumulada diaria.

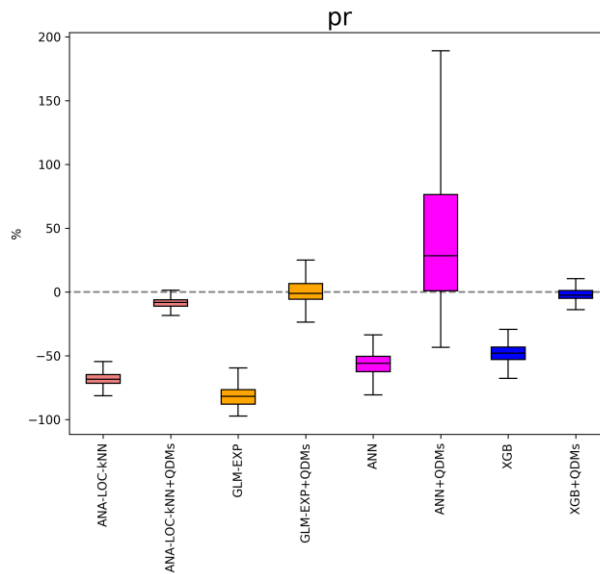


Figura 4.13: Comparativa del sesgo de la varianza obtenida con los diferentes métodos en precipitación. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de la varianza de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

En el caso de la precipitación, puede observarse que XGB+QDMs es el método que más reduce el sesgo de la varianza. Respecto al llamativo aumento de tamaño del *boxplot* asociado al sesgo de la varianza para la precipitación tras la aplicación de la corrección de sesgos al método de redes neuronales ANN+QDMs, hay que aclarar que los métodos de corrección de sesgos por percentiles, como QDM, corrigen automáticamente los casos con excesiva frecuencia de días húmedos, pero pueden presentar problemas en el caso contrario, cuando la serie a corregir presenta demasiados días secos. Los datos generados por ANN contienen muchos puntos que corresponden a este segundo caso, y esto hace que tras la corrección de sesgos puedan aparecer algunos valores muy elevados de la precipitación, que se traducen en una gran varianza frente a la varianza observada.

4.2.e. Correlación espacial

La correlación espacial para la temperatura máxima diaria y la temperatura mínima diaria está representada en la figura 4.14. En ambas temperaturas todos los métodos proporcionan resultados parecidos, aunque la aplicación de QDMs reduce ligeramente la correlación espacial.

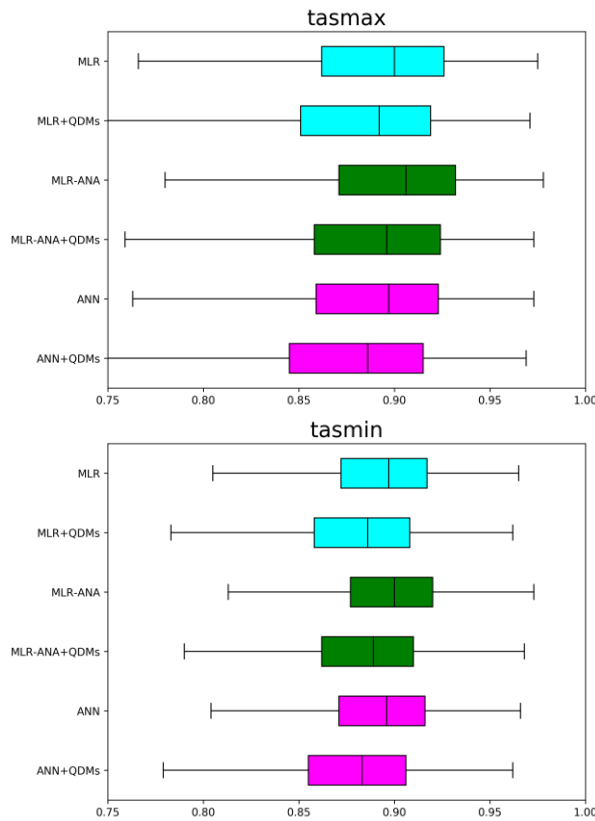


Figura 4.14: Correlación espacial para la temperatura máxima diaria y la temperatura mínima diaria. Cada *boxplot* contiene un valor por cada día del periodo de evaluación, correspondiente al coeficiente de correlación entre los patrones espaciales simulados y observados.

Por su parte, la correlación espacial para la precipitación acumulada diaria está representada en la figura 4.15. En este caso, el método XGB es el que proporciona correlaciones espaciales más

elevadas. El método XGB+QDMs proporciona una correlación espacial más baja que XGB, pero que no empeora significativamente las obtenidas con los métodos antiguos. En la figura se observa también la existencia de correlación espacial negativa en algunos puntos. Esto significa que si en dichos puntos se recoge una cantidad de precipitación acumulada diaria por encima del promedio, es más probable que los puntos cercanos recojan valores por debajo del promedio y viceversa. Este patrón indica una oposición en los patrones de precipitación entre áreas cercanas. Puede haber diversas razones para la existencia de una correlación espacial negativa en la precipitación, como factores topográficos, patrones climáticos regionales o interacciones atmosféricas y oceanográficas. Además, la correlación espacial negativa no implica causalidad, sino simplemente una relación estadística en los datos observados.

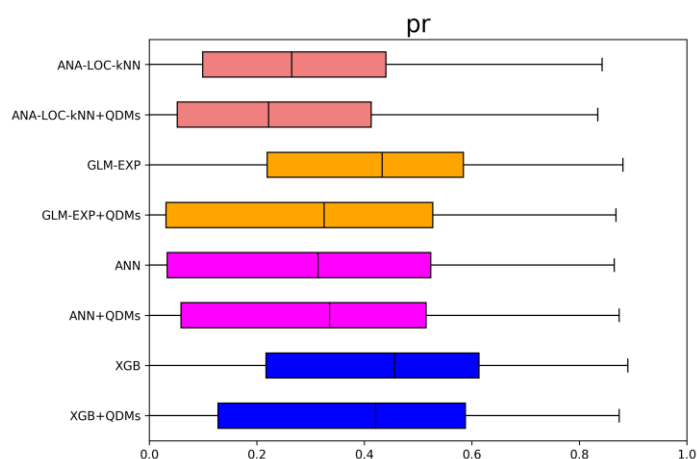


Figura 4.15: Correlación espacial para la precipitación acumulada diaria. Cada *boxplot* contiene un valor por cada día del periodo de evaluación, correspondiente al coeficiente de correlación entre los patrones espaciales simulados y observados.

4.2.f. Raíz del error cuadrático medio (RMSE)

La figura 4.16 representa el RMSE para las temperaturas máxima y mínima diarias. El método MLR-ANA es el que genera menor RMSE para ambas temperaturas. La aplicación de QDMs aumenta ligeramente el RMSE en todos los métodos evaluados para ambas temperaturas.

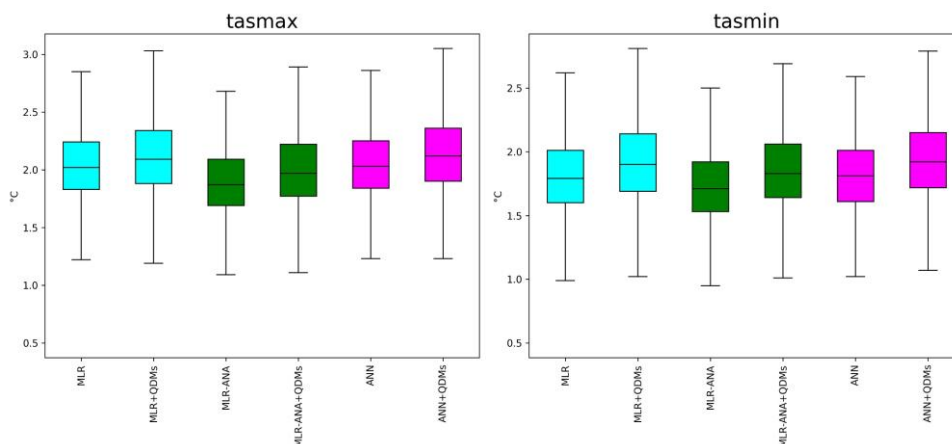


Figura 4.16: RMSE para Las temperaturas máxima y mínima diarias. Cada *boxplot* se refiere a un método ESD determinado y contiene el RMSE de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

4.2.g. Tendencias a futuro

En este apartado se comparan resultados de las proyecciones regionalizadas para los distintos métodos evaluados con la tendencia a futuro de los GCM. Los GCM operan en una escala espacial mayor y distinta a los métodos ESD y se trata de comprobar si se conservan las tendencias en la escala en la que operan los GCM. Por ello, esta comprobación solo es válida promediando en dominios espaciales amplios, nunca en dominios subrejilla. Para realizar la comprobación, se regionalizaron los datos del escenario SSP5-8.5 de los GCM seleccionados en el apartado 3.3 y listados en la tabla 3.5 (CMCC-CM2-SR5_r1i1p1f1 y IITM-ESM_r1i1p1f1 no se utilizaron porque ni tasmax ni tasmin están disponibles en ESGF) y se representaron sus tendencias a futuro junto a las tendencias de los GCM sin regionalizar, pero interpolados por el vecino más cercano (llamaremos RAW a estos últimos en las figuras). Las tendencias se representan para índices CLIMDEX asociados a las temperaturas medias, TXm y TNm, y a la precipitación total, PRCPTOT. Las gráficas de las tendencias muestran la señal de cambio respecto al periodo de referencia 1981-2010.

La figura 4.17 muestra cómo todos los métodos (con y sin QDMs) mantienen las tendencias de los GCM en TXm.

La figura 4.18 muestra cómo los métodos mantienen las tendencias relativamente bien de los GCM en TNm, aunque subestiman ligeramente los valores de la señal de cambio. En este sentido, MLR-ANA+QDMs es el método que la subestima menos.

Por último, la figura 4.19 compara las tendencias que ofrecen los métodos de regionalización con las de los GCM en PRCPTOT. ANA-LOC-kNN y ANA-LOC-kNN+QDMs sobreestiman los valores de la señal de cambio PRCPTOT de los GCM. Por el contrario, GLM-EXP, GLM-EXP+QDMs, ANN y ANN+QDMs subestiman los valores de la señal de cambio PRCPTOT. Solamente XGB y XGB+QDMs son capaces de ajustarse correctamente a la señal de cambio.

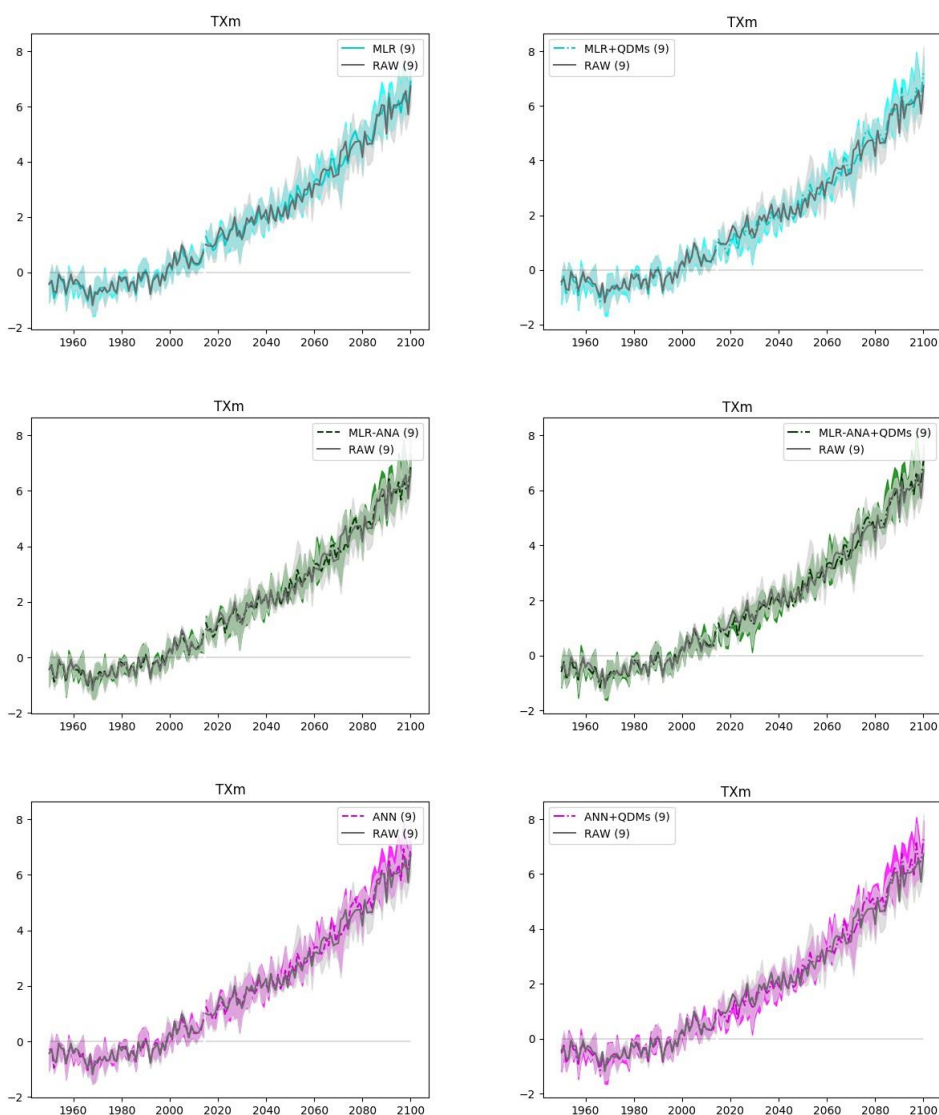


Figura 4.17: Comparativa para TXm entre proyecciones con distintos métodos (en colores) y los GCM (RAW, en gris) en el escenario SSP5-8.5 (se representan los cambios en °C respecto a la media en el periodo de referencia 1981-2020).

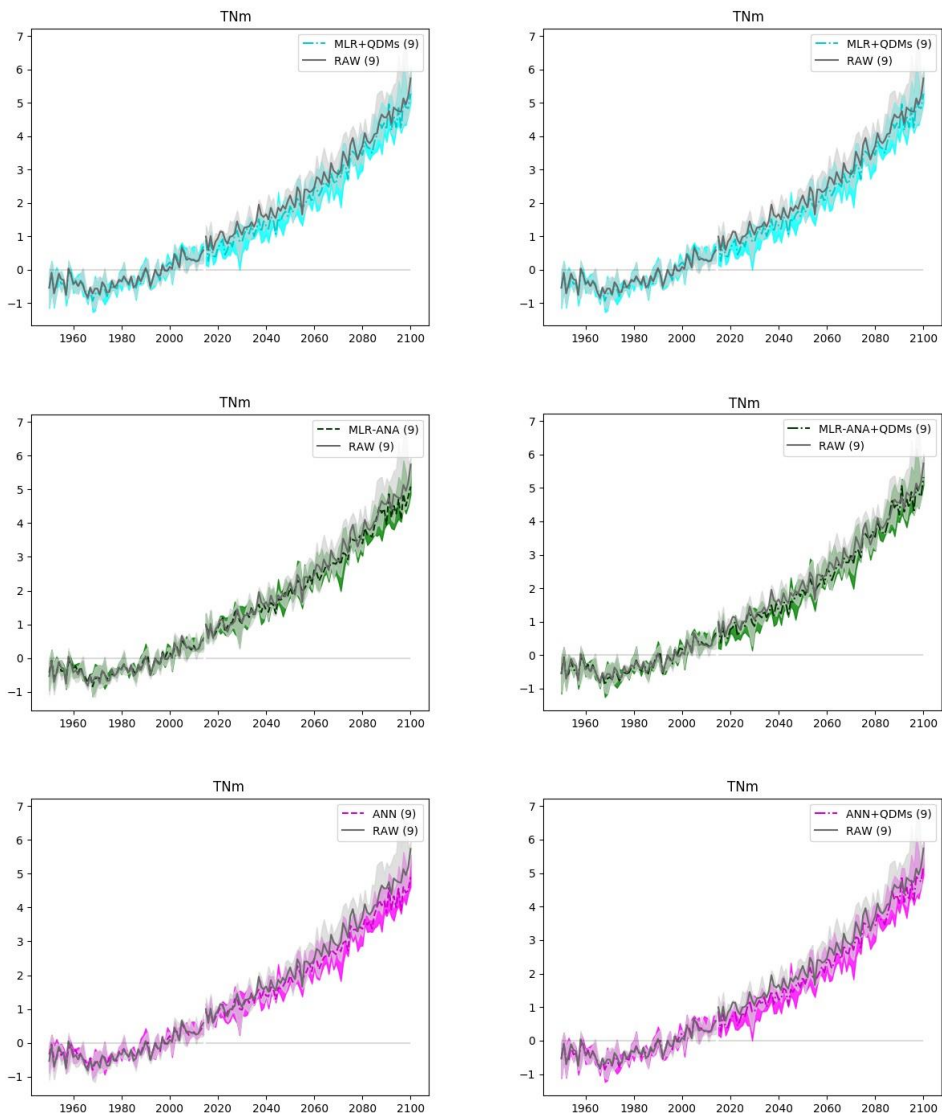


Figura 4.18: Comparativa para TNm entre proyecciones con distintos métodos (en colores) y los GCM (RAW, en gris) en el escenario SSP5-8.5 (se representan los cambios en °C respecto a la media en el periodo de referencia 1981-2020).

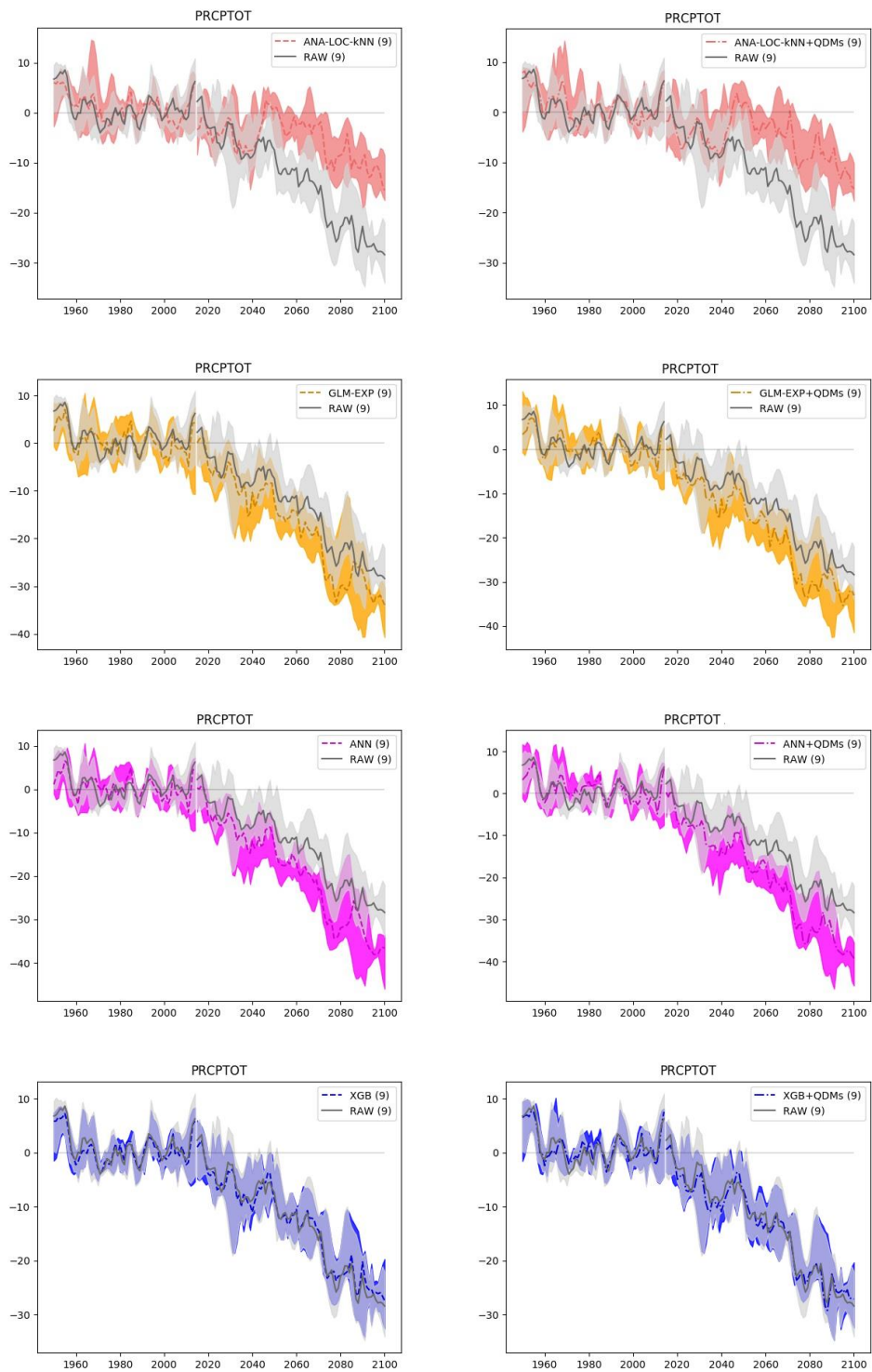


Figura 4.19, comparativa para PRCPTOT entre proyecciones con distintos métodos (en colores) y los GCM (RAW, en gris) en el escenario SSP5-8.5 (se representan los cambios en % respecto a la media en el periodo de referencia 1981-2020).

4.3. Evaluación en Canarias

En el presente apartado se muestra la evaluación de los métodos ESD en las estaciones de AEMET en Canarias de manera análoga a la realizada en la Península y Baleares en el apartado 4.2.

4.3.a. Ciclo anual

Las figuras 4.20 y 4.21 muestran la comparativa de los métodos de regionalización estadística en las estaciones de observación (OBS) que AEMET tiene distribuidas en Canarias para el ciclo anual de la temperatura máxima diaria y de la temperatura mínima diaria, respectivamente.

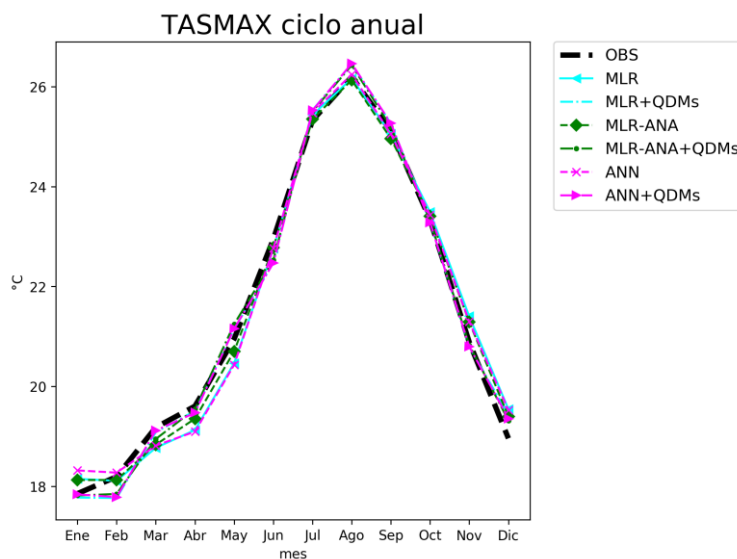


Figura 4.20: Ciclo anual de la temperatura máxima diaria.

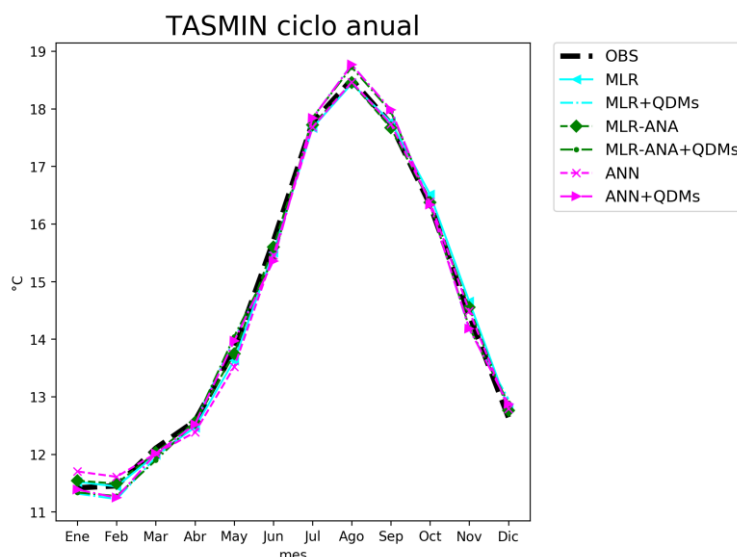


Figura 4.21: Ciclo anual de la temperatura mínima diaria.

Se constata que todos los métodos evaluados reproducen el ciclo anual observado en ambas temperaturas con gran aproximación.

La figura 4.22 muestra la misma comparativa de los métodos de regionalización estadística, pero referida al ciclo anual de la precipitación acumulada. En este caso, XGB+QDMs, presenta una aproximación al ciclo anual de precipitación observada notablemente superior a la que ofrecen los demás métodos, los cuales, a diferencia de lo que sucedía con las temperaturas, no son capaces de reproducir correctamente el ciclo anual de la precipitación.

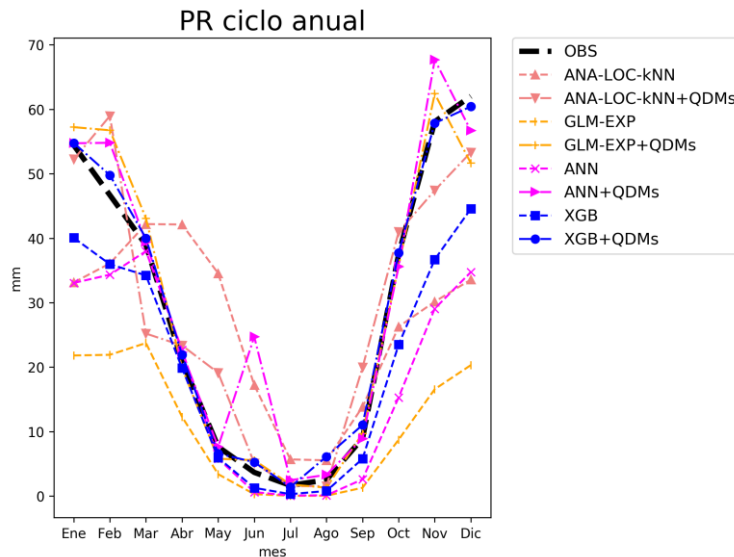


Figura 4.22: Ciclo anual de la precipitación acumulada diaria.

4.3.b. Sesgo

Se analizan comparativamente los sesgos anuales obtenidos para los índices CLIMDEX asociados a las temperaturas y a las precipitaciones con los métodos de regionalización considerados (los sesgos estacionales se encuentran en el Anexo B).

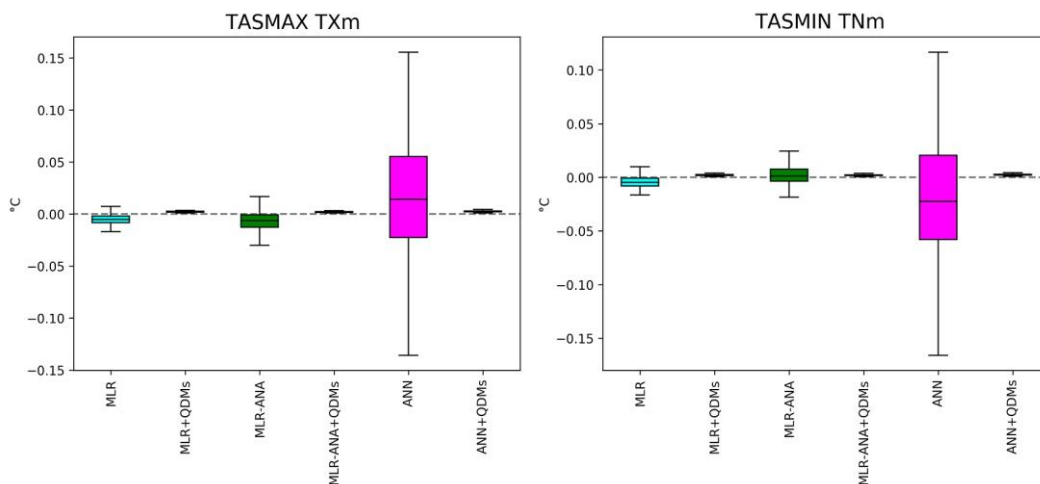


Figura 4.23: sesgos para los índices CLIMDEX: TXm y TNm. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

La figura 4.23 representa los sesgos asociados a los índices TXm y TNm. En ambos índices, la aplicación de QDMs a cada uno de los métodos evaluados proporciona los sesgos muy reducidos, llevando al rango intercuartílico a niveles insignificantes, inferiores a 0.01 °C.

La figura 4.24 muestra los sesgos asociados a los índices de los valores extremos de temperatura: TXn, TNn, TXx y TNx. En los cuatro índices, la aplicación de QDMs tras la regionalización reduce significativamente los sesgos. En estos índices, al igual que ocurría al analizarlos en Península y Baleares, los sesgos son mayores que en los índices que contienen los promedios anuales, TXm y TNm, (ver figura 4.23).

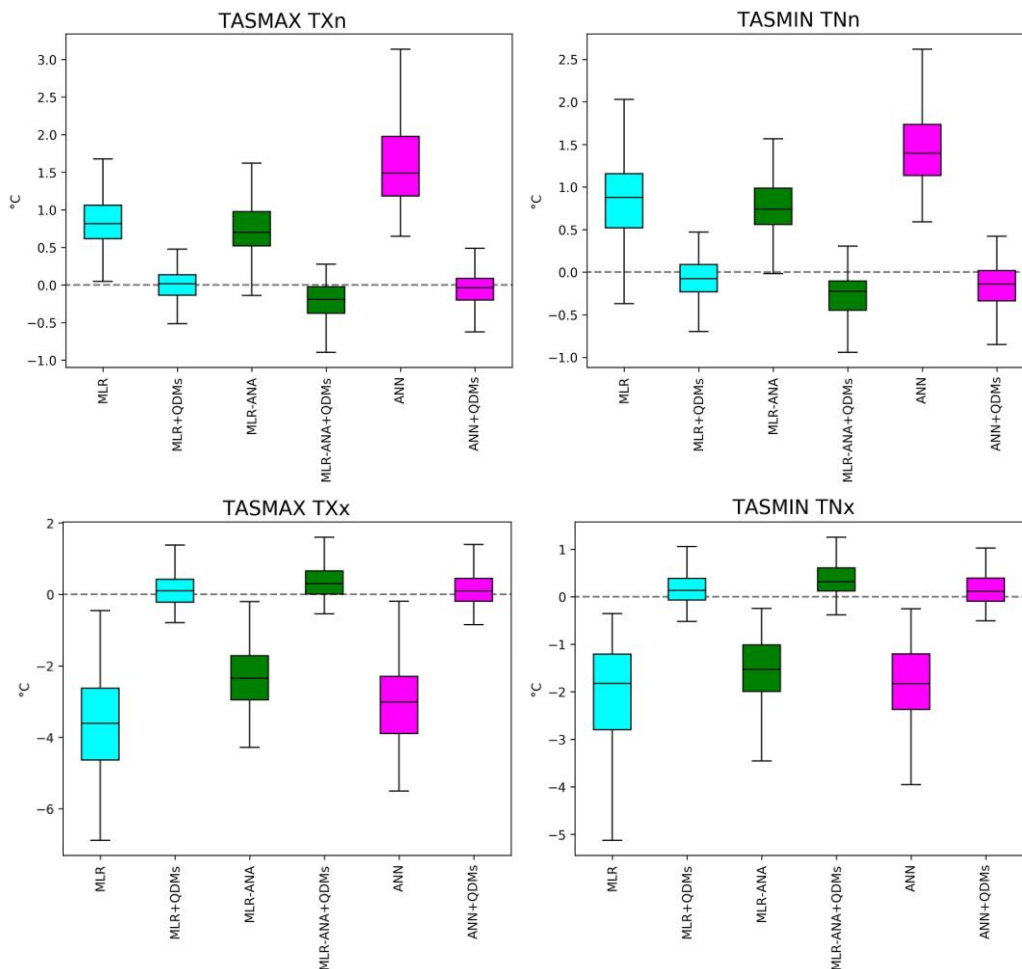


Figura 4.24: Sesgos para los índices CLIMDEX: TXn, TNn, TXx y TNx. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

La figura 4.25 muestra los sesgos asociados a los índices: WSDI y CSDI. En ambos índices, la aplicación de QDMs tras la regionalización cambia el signo del sesgo de la mediana (de negativo a positivo). A diferencia de lo que sucede en WSDI, en CSDI la aplicación de QDMs aumenta el tamaño de los *boxplots* respecto al resto de métodos, pero al menos consigue centrar la mediana en las cercanías del 0 con el método MLR-ANA+QDMs, superando en este aspecto a los métodos MLR+QDMs y ANN+QDMs. Se constata, por tanto, que en los índices WSDI y CSDI, tanto en Península y Baleares como en Canarias, la aplicación de QDMs no proporciona

las mejoras significativas que sí ofrece en otros índices CLIMDEX de temperaturas (véase figuras 4.23 y 4.24).

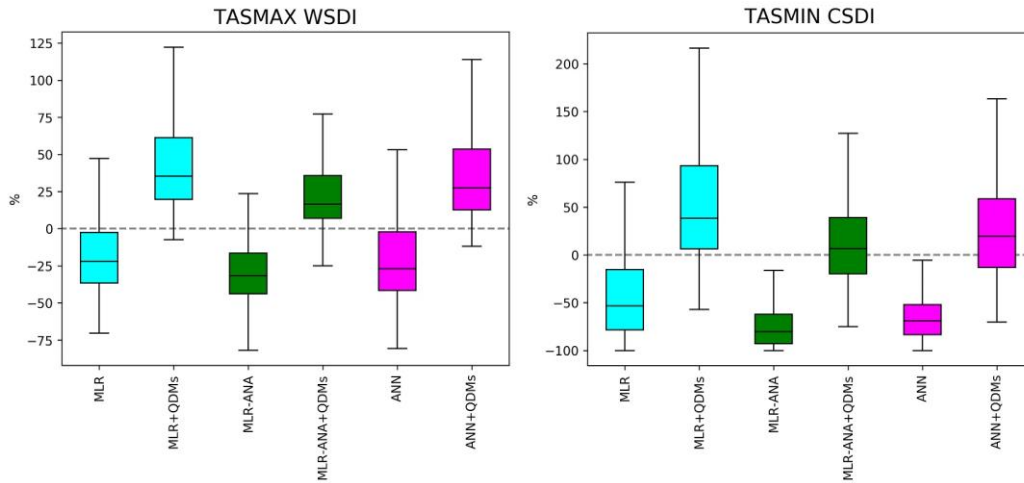


Figura 4.25: Sesgos para los índices CLIMDEX: WSDI y CSDI. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

La figura 4.26 representa los sesgos asociados a los índices CLIMDEX de precipitación: PRCPTOT, R01, R95p, SDII, CDD y CWD.

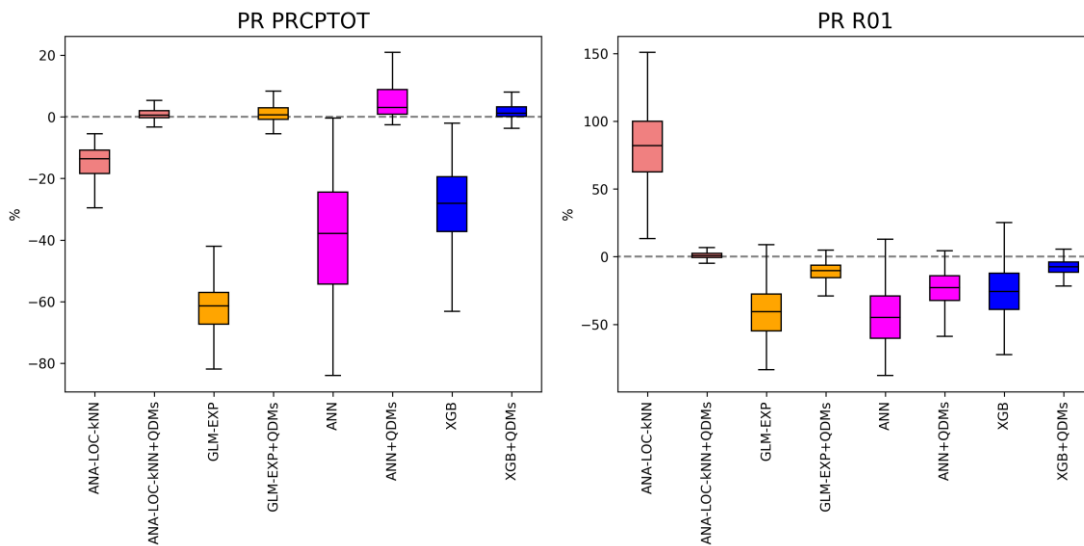


Figura 4.26: Sesgos para los índices CLIMDEX: PRCPTOT y R01.

La figura 4.27 representa los sesgos asociados a los índices CLIMDEX de precipitación: R95p y SDII.

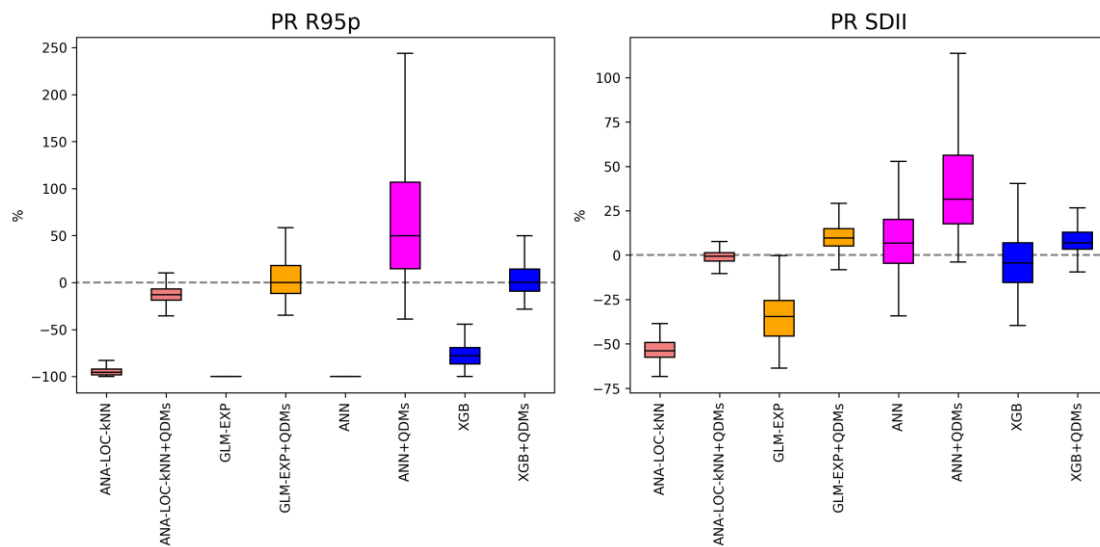


Figura 4.27: Sesgos para los índices CLIMDEX: R95p y SDII.

En PRCPTOT, R01, R95p y SDII, ANA-LOC-kNN+QDMs es el método que proporciona menores sesgos seguido de cerca por XGB+QDMs.

La figura 4.28 representa los sesgos asociados a los índices CLIMDEX de precipitación: CDD y CWD.

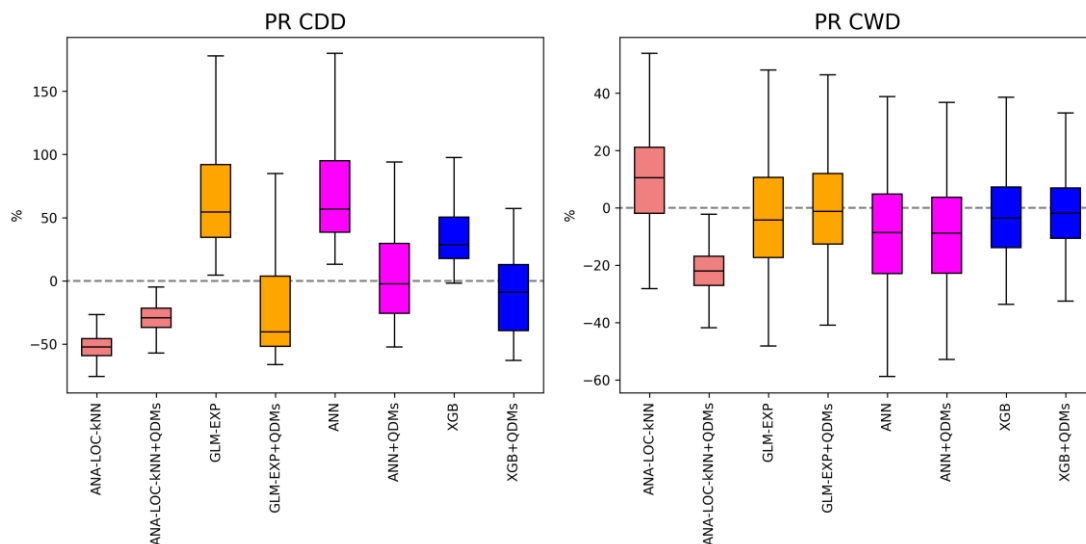


Figura 4.28: Sesgos para los índices CLIMDEX: CDD y CWD.

En CDD y CWD es XGB+QDMs es el que da lugar a mejores resultados. El comportamiento de los sesgos asociados a los índices CLIMDEX de precipitación en Canarias muestra un patrón bastante similar al que se obtiene en Península y Baleares (apartado 4.2.b).

4.3.c. Correlación

La figura 4.29 representa el coeficiente de correlación obtenido en Canarias por cada método en temperatura máxima diaria y la temperatura mínima diaria.

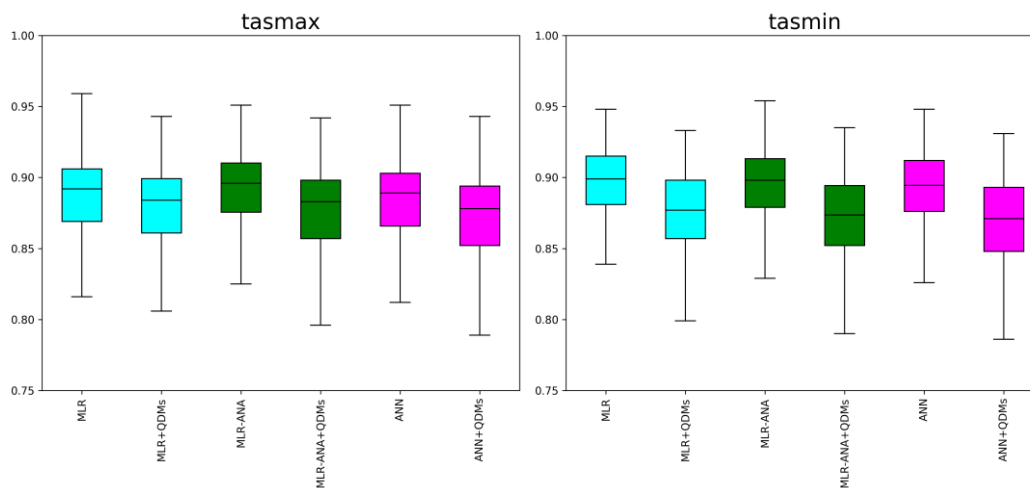


Figura 4.29: Comparativa del coeficiente de correlación obtenido con los diferentes métodos en temperaturas. Cada *boxplot* se refiere a un método ESD determinado y contiene los coeficientes de correlación de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

La figura 4.30 representa el coeficiente de correlación obtenido en Canarias por cada método en la precipitación acumulada diaria.

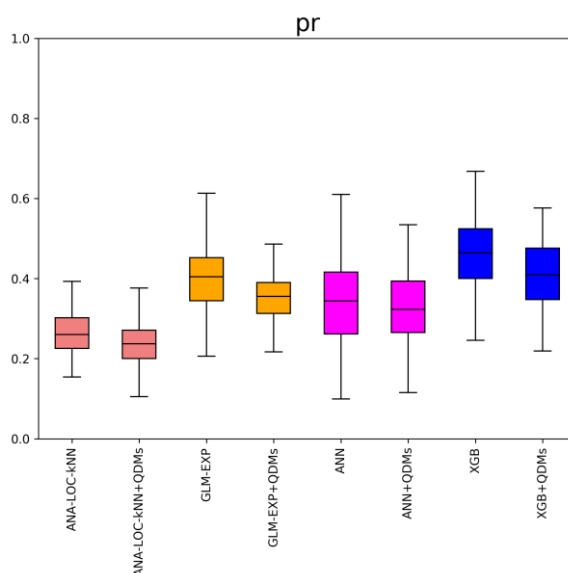


Figura 4.30: Comparativa del coeficiente de correlación obtenido con los diferentes métodos en precipitación. Cada *boxplot* se refiere a un método ESD determinado y contiene los

coeficientes de correlación de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

En temperaturas, a todos los métodos les corresponde un coeficiente de correlación elevado y cercano a 0.9. En precipitación, las correlaciones son significativamente menores que en temperaturas, situándose mayoritariamente por debajo de 0.5. En este caso sí que destacan tanto XGB como XGB+QDMs al proporcionar correlaciones significativamente mayores que los métodos antiguos. Otro aspecto importante es que los valores del coeficiente de correlación en los tres predictandos son menores en Canarias que en Península y Baleares (apartado 4.2.c). Esto se debe a las diferencias en el régimen normal de precipitaciones entre el archipiélago canario y Península y Baleares.

4.3.d. Varianza

La figura 4.31 representa los sesgos de la varianza obtenida en Canarias por cada método en temperatura máxima diaria y temperatura mínima diaria.

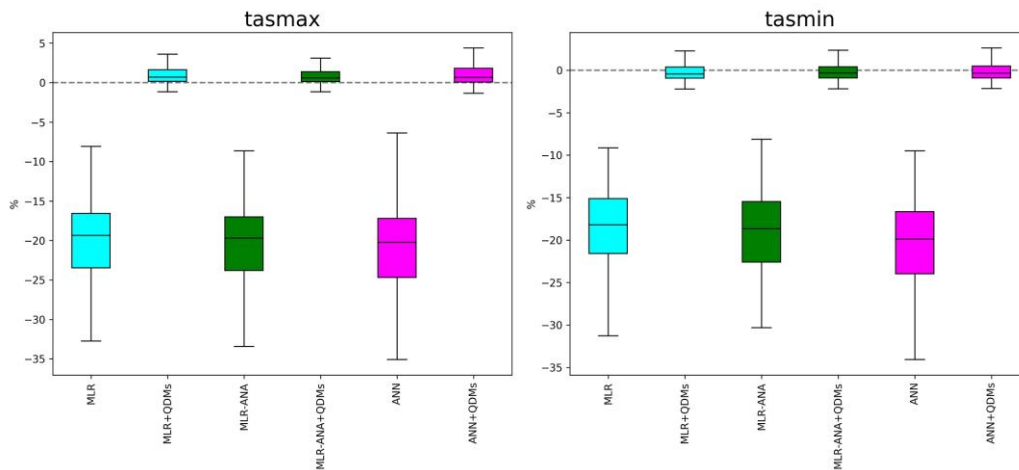


Figura 4.31: Comparativa de los sesgos de la varianza obtenidos con los diferentes métodos en temperaturas y precipitación. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de la varianza de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

En el caso de las temperaturas puede observarse cómo la aplicación de QDMs tras la regionalización reduce con eficacia el sesgo de la varianza respecto a los métodos antiguos, sin embargo, en la precipitación acumulada diaria (figura 4.32) la corrección del sesgo de la varianza no resulta tan eficaz en términos relativos; el mejor método es ANA-LOC-kNN+QDMs seguido de XGB+QDMs. Respecto al llamativo aumento de tamaño del *boxplot* asociado al sesgo de la varianza para la precipitación tras la aplicación de la corrección de sesgos al método de redes neuronales ANN+QDMs, hay que aclarar que los métodos de corrección de sesgos por percentiles, como QDM, corrigen automáticamente los casos con excesiva frecuencia de días húmedos, pero pueden presentar problemas en el caso contrario, cuando la serie a corregir presenta demasiados días secos. Los datos generados por ANN contienen muchos puntos que corresponden a este segundo caso, y esto hace que tras la corrección de sesgos puedan aparecer algunos valores muy elevados de la precipitación, que se traducen en una gran varianza frente a la varianza observada.

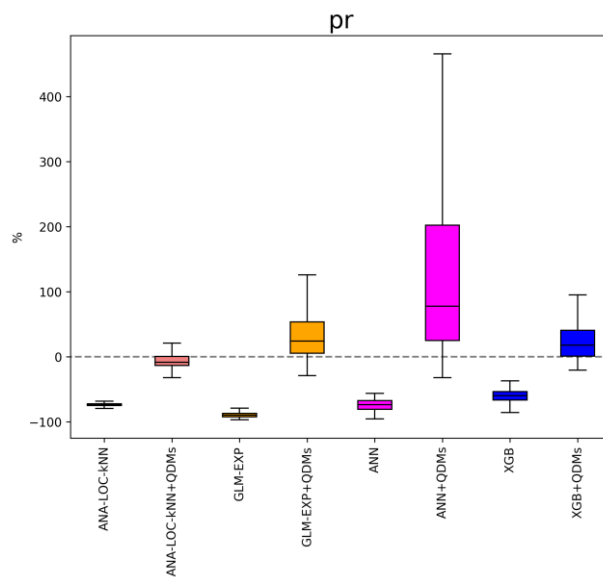


Figura 4.32: Comparativa de los sesgos de la varianza obtenidos con los diferentes métodos en la precipitación acumulada diaria. Cada *boxplot* se refiere a un método ESD determinado y contiene los sesgos de la varianza de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

4.3.e. Correlación espacial

La correlación espacial en Canarias para la temperatura máxima diaria está representada en la figura 4.33, mientras que la asociada a la temperatura mínima diaria se dibuja en la figura 4.34. Los valores de correlación espacial de las temperaturas son menores en Canarias que en Península y Baleares (comparar con figura 4.14 del apartado 4.2.e). Además, en Canarias se hace más evidente la diferencia entre las correlaciones espaciales en tasmax y en tasmin, siendo las primeras inferiores a las segundas.

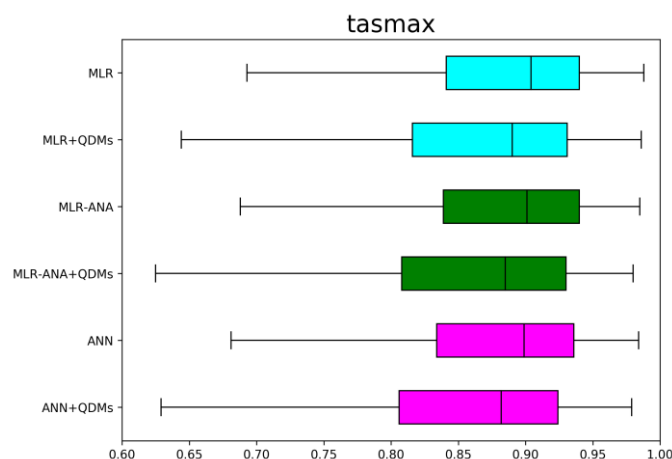


Figura 4.33: Correlación espacial para la temperatura máxima diaria. Cada *boxplot* contiene un valor por cada día del periodo de evaluación, correspondiente al coeficiente de correlación entre los patrones espaciales simulados y observados.

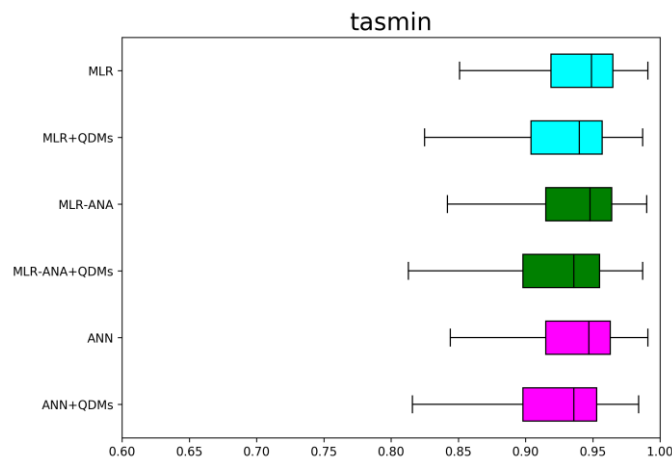


Figura 4.34: Correlación espacial para la temperatura mínima diaria. Cada *boxplot* contiene un valor por cada día del periodo de evaluación, correspondiente al coeficiente de correlación espacial entre los patrones espaciales simulados y observados.

La correlación espacial para la precipitación acumulada diaria se representa en la figura 4.35. El método XGB es el que proporciona correlaciones espaciales más elevadas. XGB+QDMs proporciona una correlación espacial más baja que XGB y que GLM-EXP, equiparable a la de ANN-LOC-kNN, pero que supera a la de ANN. Los valores de correlación espacial en precipitación son menores en Canarias en comparación con Península y Baleares (ver figura 4.15). Con respecto a la aparición de puntos con correlación espacial negativa, véase lo ya comentado en el apartado al respecto en el apartado 4.2.e.

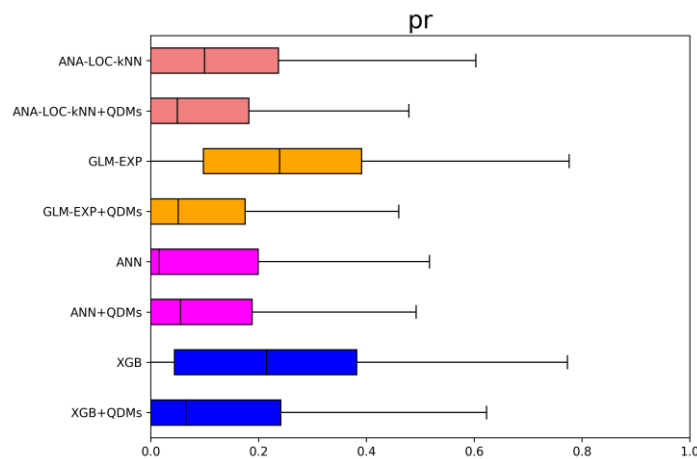


Figura 4.35: Correlación espacial para la precipitación acumulada diaria. Cada *boxplot* contiene un valor por cada día del periodo de evaluación, correspondiente al coeficiente de correlación espacial entre los patrones espaciales simulados y observados.

4.3.f. Raíz del error cuadrático medio (RMSE)

La figura 4.36 representa el RMSE en Canarias para las temperaturas máxima y mínima diarias. El método MLR-ANA es el que genera menor RMSE para ambas temperaturas. La aplicación de QDMs aumenta ligeramente el RMSE en todos los métodos evaluados para ambas temperaturas. En Canarias se obtienen resultados muy similares a los obtenidos en Península y Baleares (véase figura 4.16).

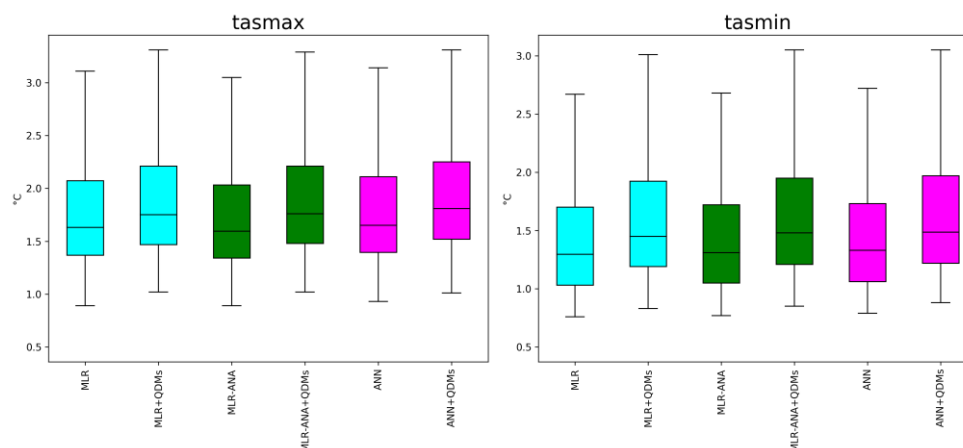


Figura 4.36: RMSE para Las temperaturas máxima y mínima diarias. Cada *boxplot* se refiere a un método ESD determinado y contiene el RMSE de todos los puntos analizados, es decir, donde se localizan las estaciones de observación.

4.3.g. Tendencias a futuro

Este apartado (que debería ser análogo al apartado 4.2.g, pero analizando las tendencias en Canarias en vez de en Península y Baleares) no procede. Como ya se explicó en el apartado 4.2.g, la comparativa de las tendencias solo es válida promediando en dominios espaciales amplios, nunca en dominios subrejilla. Las Islas Canarias no presentan un área lo suficientemente extensa como para poder realizar la comparativa entre las tendencias de los GCM y la de sus proyecciones regionalizadas por métodos ESD.

5. CONCLUSIONES

En el presente informe se establece la estrategia de AEMET para la generación de escenarios de cambio climático de alta resolución sobre el territorio español en el marco del PNACC-2 (2021-2030). Dichos escenarios serán generados por métodos estadísticos, utilizando los modelos climáticos del Sexto Informe de Evaluación del IPCC, y complementados por metodologías dinámicas procedentes de EURO-CORDEX cuando estén disponibles. El objetivo de este informe es evaluar y seleccionar tanto los modelos climáticos globales a utilizar como los métodos estadísticos de regionalización.

Como resultado de la evaluación de modelos climáticos globales se ha seleccionado un *ensemble* de 11 modelos entre los que se incluyen los 6 modelos que se utilizarán en EURO-CORDEX. Los modelos seleccionados son: 1) ACCESS-CM2_r1i1p1f1; 2) CMCC-CM2-SR5_r1i1p1f1; 3) CNRM-ESM2-1_r1i1p1f2; 4) EC-Earth3-Veg_r1i1p1f1; 5) IITM-

ESM_r1i1p1f1; 6) KACE-1-0-G_r1i1p1f1; 7) MIROC6_r1i1p1f1; 8) MPI-ESM1-2-HR_r1i1p1f1; 9) MRI-ESM2-0_r1i1p1f1; 10) NorESM2-MM_r1i1p1f1; 11) UKESM1-0-LL_r2i1p1f2.

Las conclusiones extraídas de la evaluación realizada de los métodos ESD son las siguientes:

- 1) No todas las metodologías mantienen las tendencias de los GCM. En precipitación, solamente los métodos XGB y XGB+QDMs siguen con buena aproximación las tendencias de los GCM en el índice PRCPTOT. En este índice, los demás métodos presentan desviaciones significativas respecto a la tendencia de los GCM. En temperaturas, las tendencias en los índices TXm y TNm se mantienen mejor en los métodos analizados sin que se aparezcan grandes desviaciones. El método MLR-ANA+QDMs destaca ligeramente sobre los demás métodos en TNm.
- 2) El método XGB+QDMs consigue en general resultados significativamente mejores que el resto de métodos en la evaluación de los índices de precipitación, con especial énfasis al analizar los sesgos de los valores extremos. En los índices de temperatura, el método MLR-ANA+QDMs ofrece los mejores resultados en general, pero no sobresale tanto como lo hace XGB+QDMs en el caso la precipitación, mejorando significativamente al resto de métodos.
- 3) A pesar de los buenos resultados en la mayor parte de los aspectos evaluados, algunos de estos aún presentan dificultades para todos los métodos. Es el caso tanto de los patrones espaciales de precipitación, con correlaciones relativamente bajas, como de los sesgos que resultan en los índices de olas de calor y de frío, WSDI y CSDI.

Por último, y en base a las conclusiones anteriores extraídas de la evaluación, AEMET opta por MLR-ANA+QDMs y XGB+QDMs como métodos de regionalización estadística de temperaturas y precipitación, respectivamente, para calcular las proyecciones climáticas en el marco de la segunda fase del Plan Nacional de Adaptación al Cambio Climático (PNACC-2, 2021-2030).

6. REFERENCIAS

Adler, R.F., Huffman G.J., Chang A., Ferraro R., Xie Janowiak, J., Rudolf B., Schneider U., Curtis S., Bolvin D., Gruber A., Susskind J. and Arkin P. (2003): The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, 4,1147-1167.

Amblar-Francés, P., Casado-Calle, M.J., Pastor-Saavedra, M.A., Ramos-Calzado, P., and Rodríguez-Camino, E. (2017), Guía de escenarios regionalizados de cambio climático sobre España a partir de los resultados del IPCC-AR5. Disponible en: https://www.aemet.es/documentos/es/conocerlas/recursos_en_linea/publicaciones_y_estudios/publicaciones/Guia_escenarios_AR5/Guia_escenarios_AR5.pdf

Boser, B., Guyon, I. and Vapnik, V. (1992), A Training Algorithm for Optimal Margin Classifier. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory* 5. <https://doi.org/10.1145/130385.130401>

Breiman, L. (2001), Random forests. *Machine learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>

Cannon, A.J. Sobie, S.R. and Murdock, T.Q. (2015) Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *J. Climate*, 28, 6938–6959, <https://doi.org/10.1175/JCLI-D-14-00754.1>

Chazarra A., Lorenzo B., Romero R., Moreno J.V. (2022), Evolución de los climas de Köppen en España en el periodo 1951-2020. Nota Técnica 37 del Área de Climatología y Aplicaciones Operativas Agencia Estatal de Meteorología. https://www.aemet.es/documentos/es/conocerlas/recursos_en_linea/publicaciones_y_estudios/publicaciones/NT_37_AEMET/NT_37_AEMET.pdf

Chen, T. and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM. <https://doi.org/10.1145/2939672.2939785>

Cortes, C. and Vapnik, V. (1995), Support-vector networks. *Machine Learning*, 20, 273–297. <https://doi.org/10.1007/BF00994018>

García-Valero, J.A. (2021) Redes neuronales artificiales. Aplicación a la regionalización de la precipitación y temperaturas diarias. Nota Técnica 34 de AEMET. Disponible en: https://www.aemet.es/documentos/es/conocerlas/recursos_en_linea/publicaciones_y_estudios/publicaciones/NT_34_Redес_neuronales_artificiales/NT_34_Redес_neuronales_artificiales.pdf

Gutiérrez, J. M., San-Martín, D., Brands, S., Manzanas, R. and Herrera, S. (2013), ‘Reassessing statistical downscaling techniques for their robust application under climate change conditions’, *J. Climate* 26(1), 171–188. <https://doi.org/10.1175/JCLI-D-11-00687.1>

Haykin, S. (2009), *Neural Networks and Learning Machines*. Third Edition, Pearson Education, Inc., McMaster University, Hamilton. <http://dai.fmph.uniba.sk/courses/NN/haykin.neural-networks.3ed.2009.pdf>

Hernanz, A., García-Valero, J. A., Domínguez, M., Ramos-Calzado, P., Pastor-Saavedra, M. A. and Rodríguez-Camino, E. (2021). Evaluation of statistical downscaling methods for climate change projections over Spain: present conditions with perf predictors. *International Journal of Climatology*, 42(2), 762– 776. <https://doi.org/10.1002/joc.7271>

Hernanz, A., Correa, C., Domínguez, M., Rodríguez-Guisado, E. and Rodríguez-Camino, E. (2023a). Statistical downscaling in the Tropics and Mid-latitudes: a comparative assessment over two representative regions. *Journal of Applied Meteorology and Climatology*, 62(6), 737-753. <https://doi.org/10.1175/JAMC-D-22-0164.1>

Hernanz, A., Correa, C., García-Valero, J. A., Domínguez, M., Rodríguez-Guisado, E., and Rodríguez-Camino, E. (2023b). pyClim-SDM: Service for generation of statistically downscaled climate change projections supporting national adaptation strategies. *Climate Services*, Volume 32, December 2023, 100408. <https://doi.org/10.1016/j.cliser.2023.100408>

Hernanz, A., Correa, C., Sánchez-Perrino, J.C., Prieto-Rico, I., Rodríguez-Guisado, E., Domínguez, M. and Rodríguez-Camino, E. (2023c). On the limitations of deep learning for statistical downscaling of climate change projections: classical and hybrid approaches. *Atmospheric Science Letters*, e1195. <http://doi.org/10.1002/asl.1195>.

Hersbach, H, Bell, B, Berrisford, P, et al. (2020) The ERA5 global reanalysis. *Q J R Meteorol Soc.*; 146: 1999– 2049. <https://doi.org/10.1002/qj.3803>

Lorenz, E.N. (1969) Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*, 26(4), 636–646. [https://doi.org/10.1175/1520-0469\(1969\)26<636:APARBN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2)

Maraun, D. and Widmann, M. (2018). *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781107588783>

McCulloch, W.S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115–133. <https://doi.org/10.1007/BF02478259>

Osman, A. I. A., Ahmed, A. N., Chow, M. F., Huang, Y. F. and El-Shafie, A. (2021). Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal*, 12(2), 1545-1556. <https://doi.org/10.1016/j.asej.2020.11.011>

Parding, K. M., Dobler, A., McSweeney, C. F., Landgren, O. A., Benestad, R., Erlandsen, H. B., Mezghani, A., Gregow, H., Rätty, O., Viktor, E., El Zohbi, J., Christensen, O. B., Loukos, H. (2020), GCMeval – An interactive tool for evaluation and selection of climate model ensembles. *Climate Services*, Volume 18, 100167. <https://doi.org/10.1016/j.cliser.2020.100167>

Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., & Hegewisch, K. C. (2015). Improved Bias Correction Techniques for Hydrological Simulations of Climate Change, *Journal of Hydrometeorology*, 16(6), 2421-2442. <https://doi.org/10.1175/JHM-D-14-0236.1>

Petisco de Lara, S.E. (2008a) Método de regionalización de precipitación basado en análogos. Explicación y Validación. Nota Técnica 3A del Área de Evaluación y Modelización del Cambio Climático de AEMET. <http://hdl.handle.net/20.500.11765/5220>

Petisco de Lara, S.E. (2008b) Método de regionalización de temperatura basado en análogos. Explicación y Validación. Nota Técnica 3B del Área de Evaluación y Modelización del Cambio Climático de AEMET. <http://hdl.handle.net/20.500.11765/5221>

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O’neill, B. C., Fujimori, S., ... & Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change*, 42, 153-168. <https://doi.org/10.1016/j.gloenvcha.2016.05.009>

Rosenblatt, F. (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386– 408. <https://doi.org/10.1037/h0042519>

Sailor, D.J. and Li, X. (1999) A semiempirical downscaling approach for predicting regional temperature impacts associated with climatic change. *Journal of Climate*, 12(1), 103–114. <https://doi.org/10.1175/1520-0442-12.1.103>

Vapnik, V.N. (1995) *The Nature of Statistical Learning Theory*. Berlin, Heidelberg: Springer-Verlag.

Wilby, R., Dawson, C. and Barrow, E.M. (2002) SDSM—a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling & Software*, 17, 145–157. [https://doi.org/10.1016/S1364-8152\(01\)00060-3](https://doi.org/10.1016/S1364-8152(01)00060-3)

Wilks, D. S., and Wilby, R. L. (1999). The weather generation game: A review of stochastic weather models. *Progress in Physical Geography*, 23, 329–357. <https://doi.org/10.1177/030913339902300302>

Zorita, E. and von Storch, H. (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of Climate*, 12, 2474–2489. [https://doi.org/10.1175/1520-0442\(1999\)012<2474:TAMAAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2)

ANEXO A: SESGOS Y TENDENCIAS ESTACIONALES EN PENÍNSULA Y BALEARES

En el presente anexo se proporcionan figuras adicionales por estaciones del año que complementan a los apartados 4.2.b y 4.2.g referidos a sesgos y tendencias a futuro, respectivamente.

Las figuras desde A.1 hasta A.14 (ambas incluidas) muestran mapas con los índices CLIMDEX analizados en el apartado 4.2 en las estaciones de AEMET en la Península y Baleares en los cuatro periodos estacionales del año considerados (DEF: diciembre, enero y febrero; MAM: marzo, abril y mayo; JJA: junio, julio y agosto; SON: septiembre, octubre y noviembre). Para cada índice CLIMDEX se muestran mapas tanto de los sesgos de los valores regionalizados (índices de temperaturas mediante el método MLR-ANA+QDMs e índices de precipitación mediante el método XGB+QDMs) como del promedio de las observaciones en el periodo de referencia (1981-2010). Los mapas con los valores observados tienen especial relevancia para la interpretación de los índices CLIMDEX, cuyo sesgo se expresa en términos relativos, ya que si el valor observado es pequeño, el sesgo relativo puede ser alto (incluso cuando sea pequeño en términos absolutos).

La figura A.15 muestra las gráficas de las tendencias de las proyecciones de TXm, TNm y PRCPTOT (los índices analizados en el apartado 4.2.g), regionalizadas mediante los métodos MLR-ANA+QDMs (TXm y TNm) y XGB+QDMs (PRCPTOT), comparadas con RAW (los datos en bruto de los GCM interpolados mediante el método del vecino más cercano) en el escenario SSP5-8.5, extrayendo los resultados para cada estación del año: DEF, MAM, JJA y SON. Las figuras revelan que los métodos seleccionados son capaces de seguir las tendencias RAW con precisión también estacionalmente. La única gráfica que muestra una desviación apreciable es la asociada a PRCPTOT en JJA en el futuro lejano (2070-2100), pero esta desviación queda completamente enmascarada en el cómputo anual (mostrado anteriormente en la figura 4.19) que sigue con gran aproximación la tendencia de RAW.

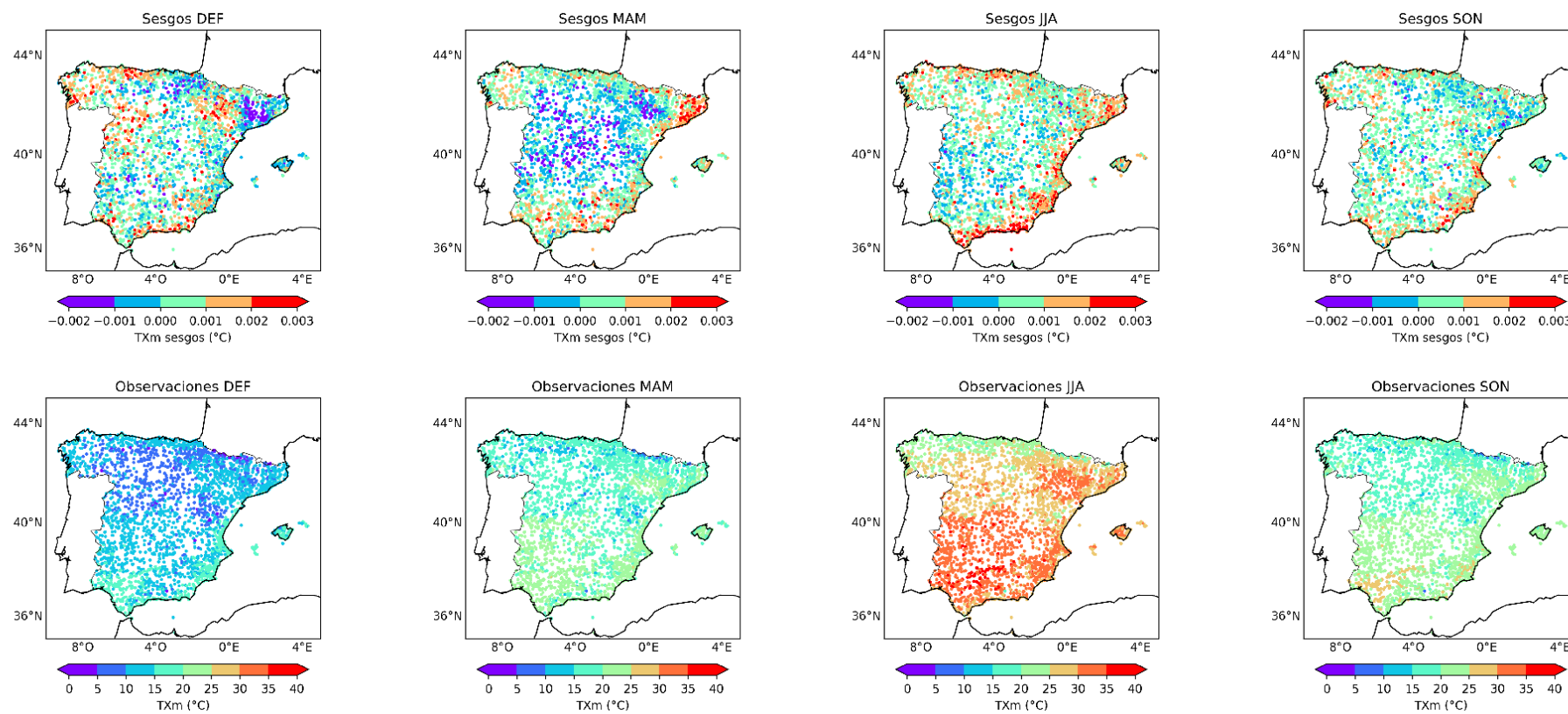


Figura A.1: Sesgos y observaciones estacionales para TXm.

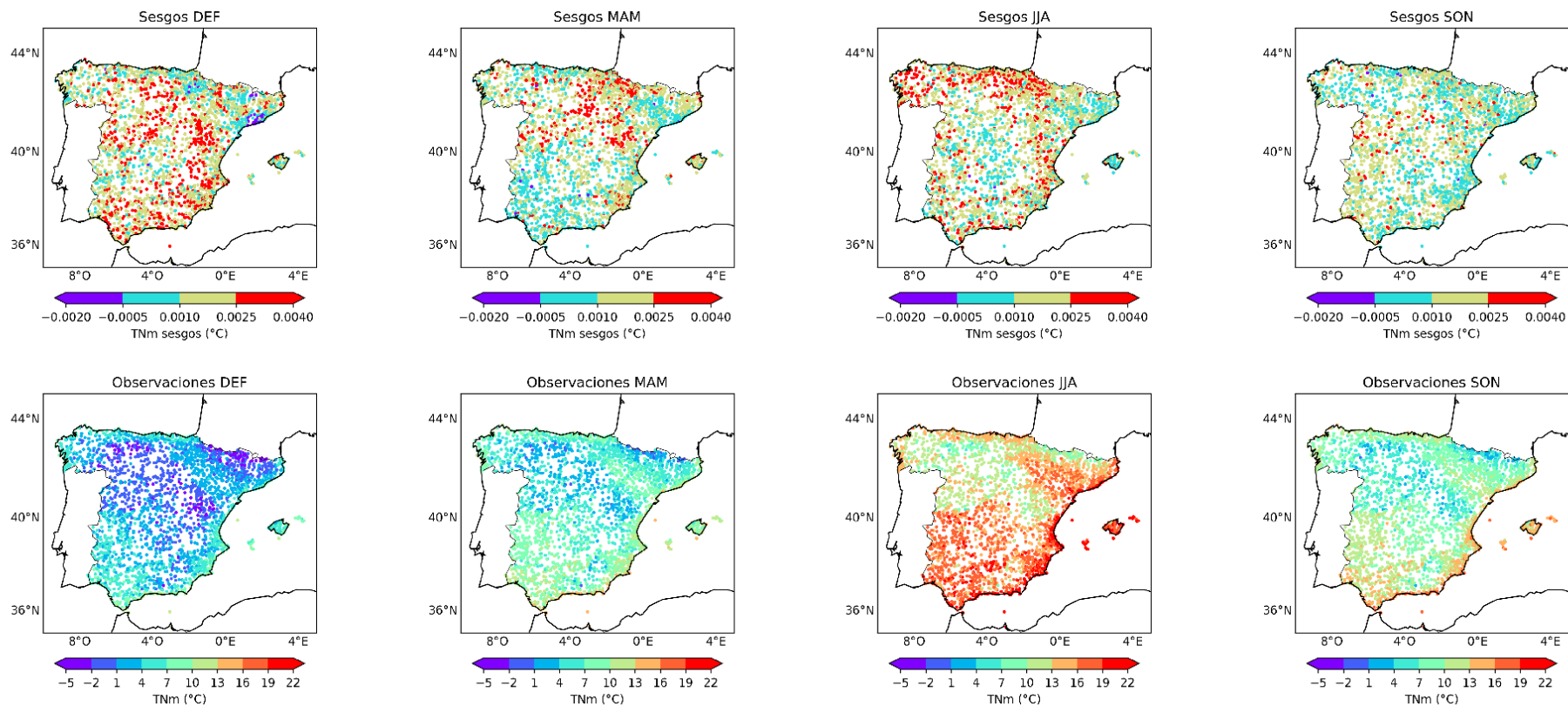


Figura A.2: Sesgos y observaciones estacionales para TNm.

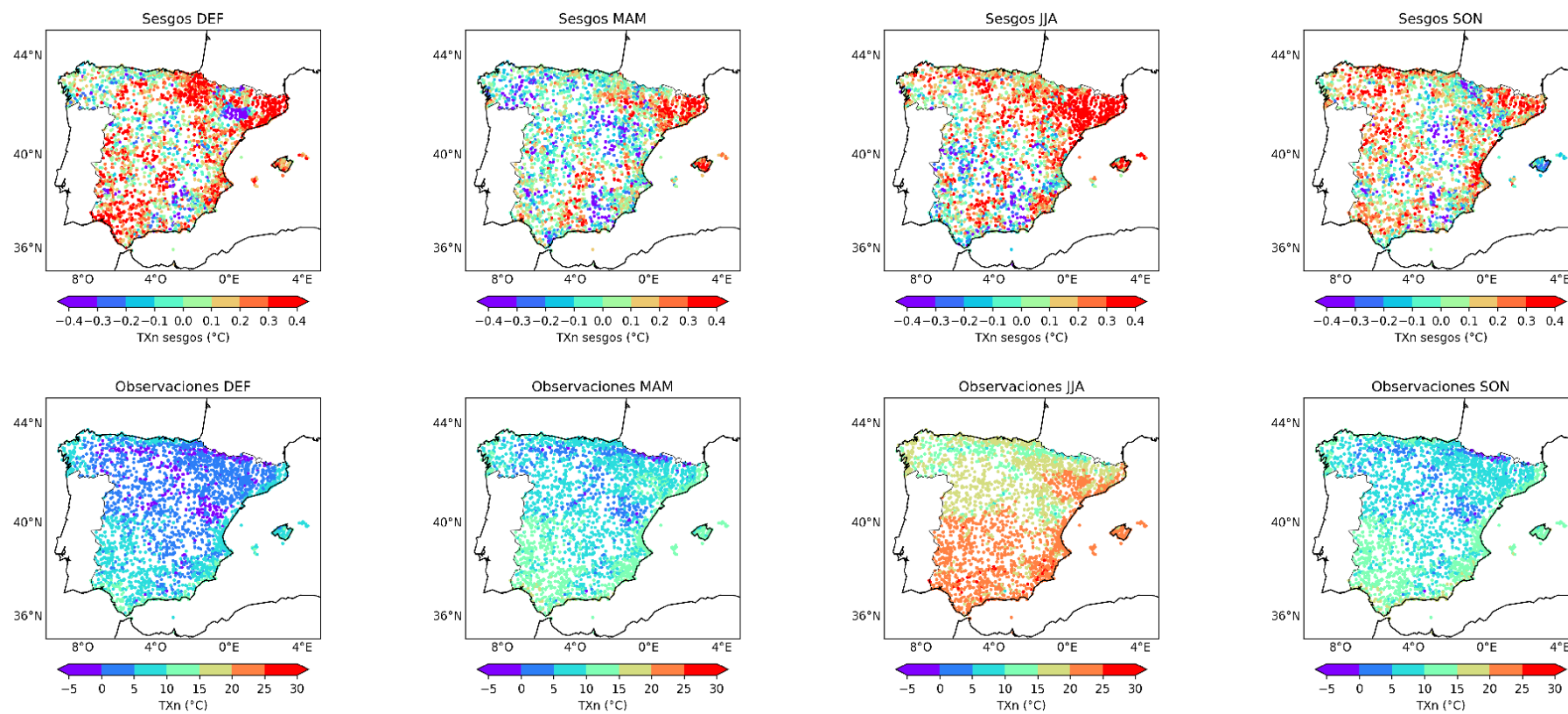


Figura A.3: Sesgos y observaciones estacionales para TXn.

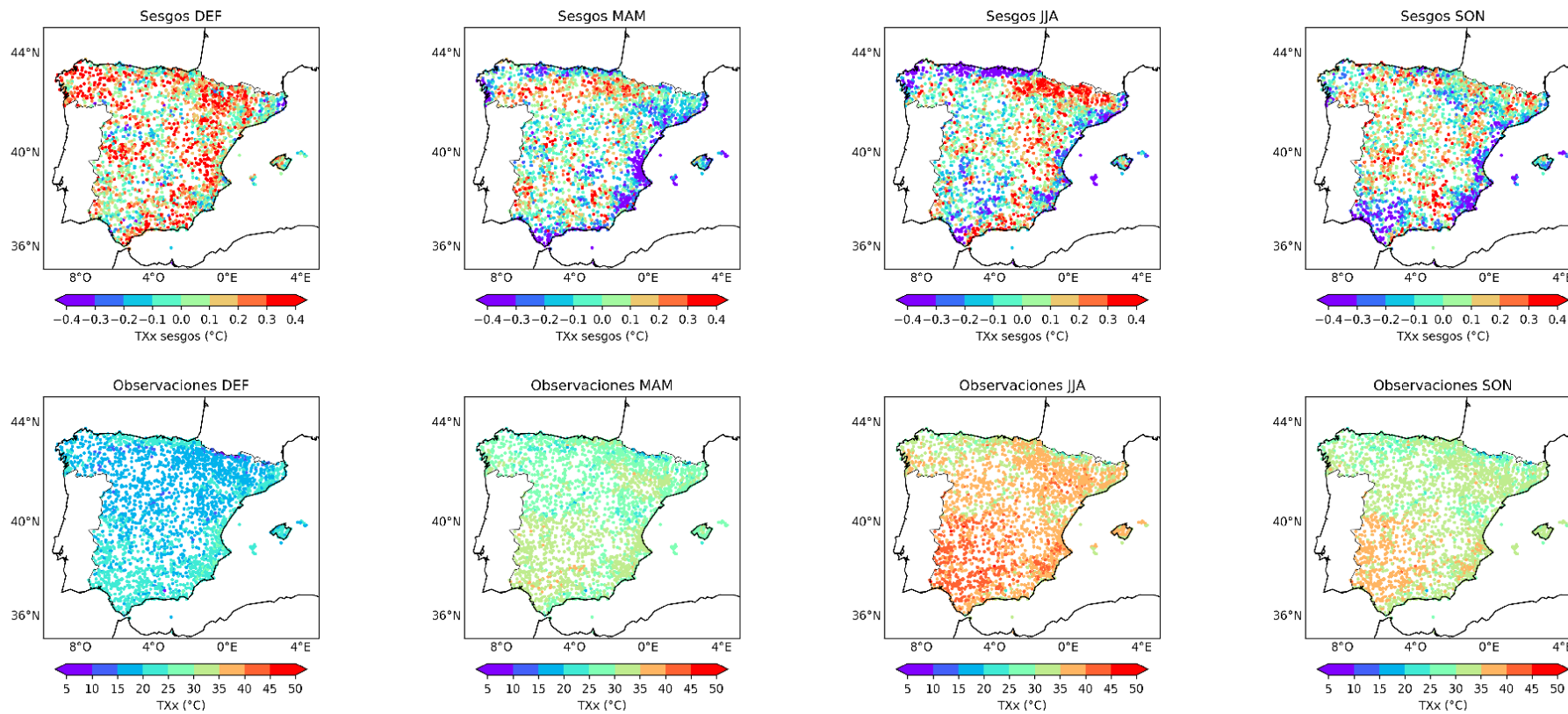


Figura A.4: Sesgos y observaciones estacionales para TXx.

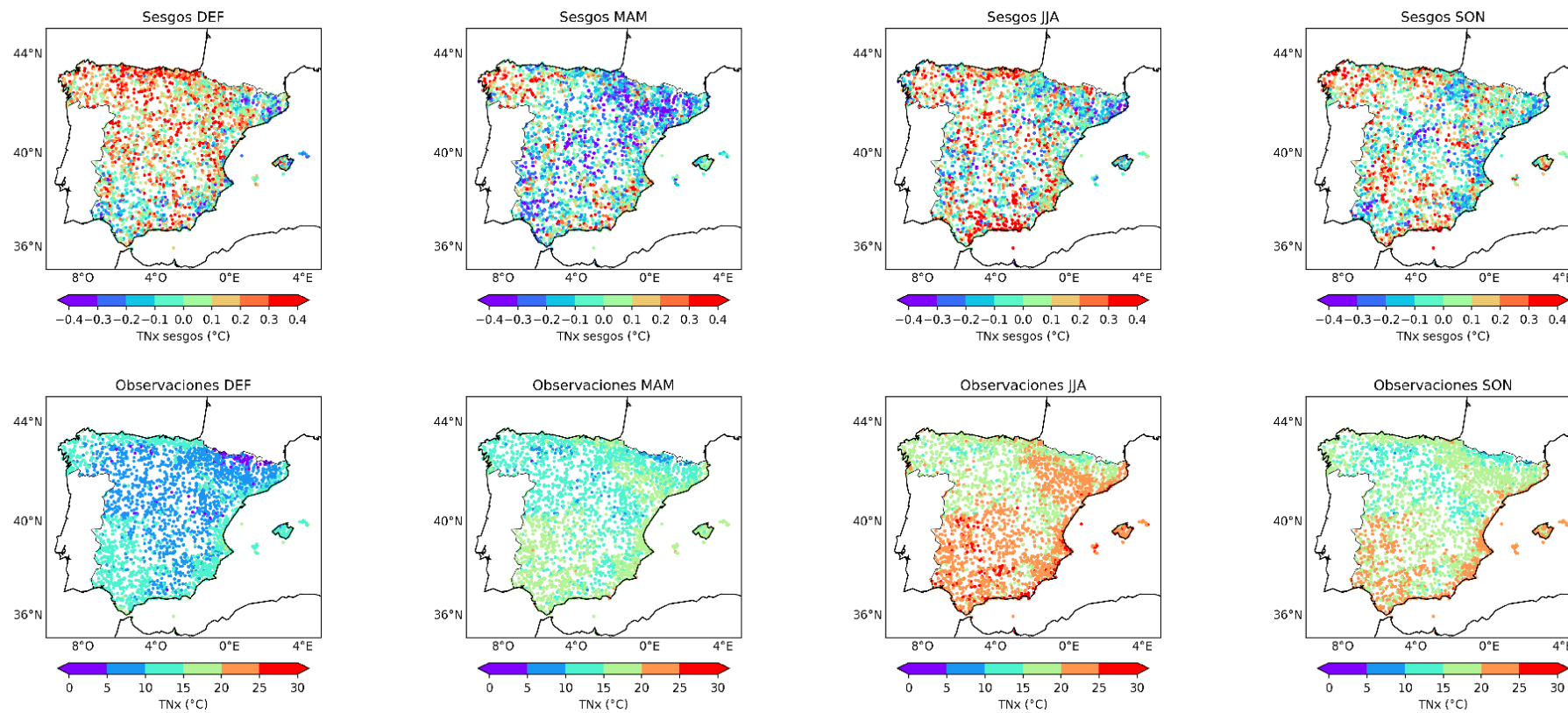


Figura A.5: Sesgos y observaciones estacionales para TNx.

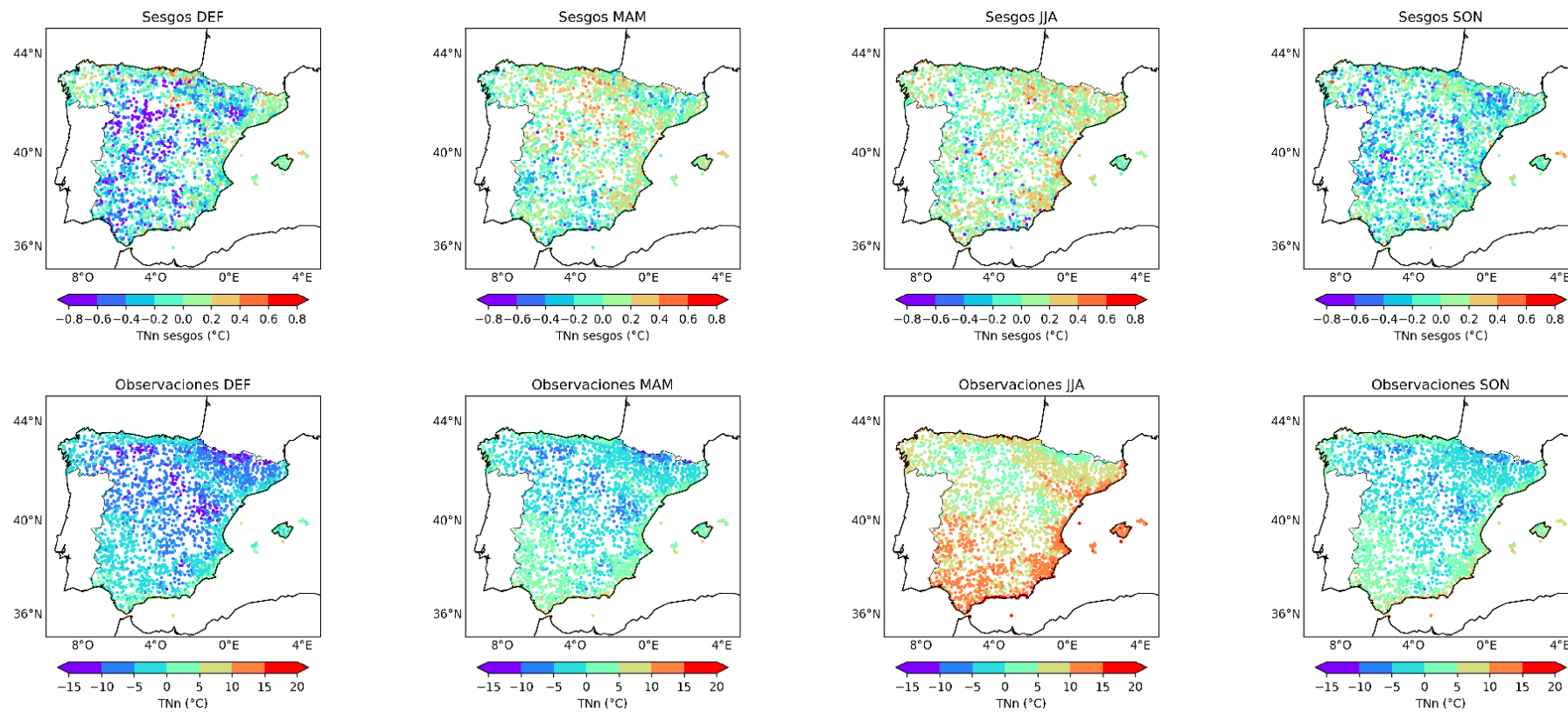


Figura A.6: Sesgos y observaciones estacionales para TNn.

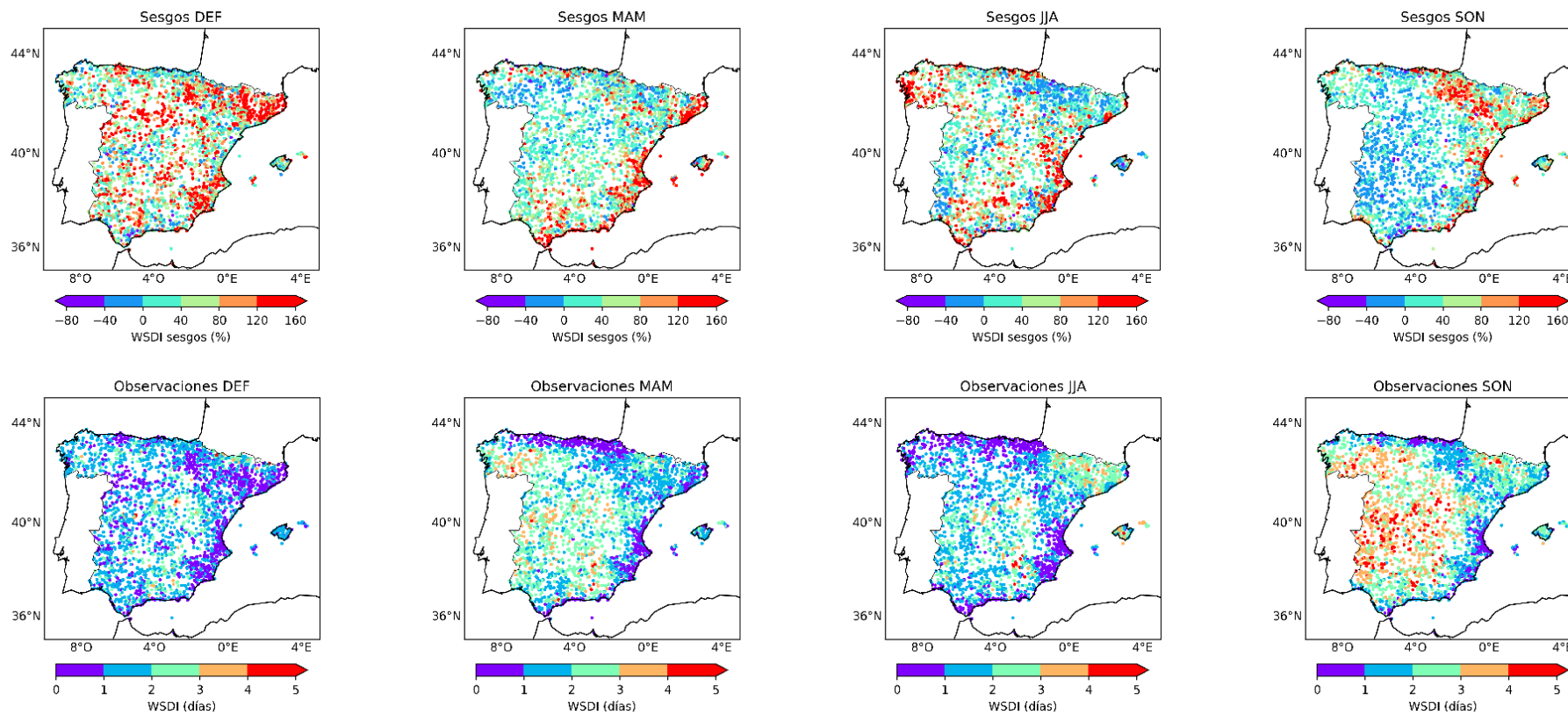


Figura A.7: Sesgos y observaciones estacionales para WSDI.

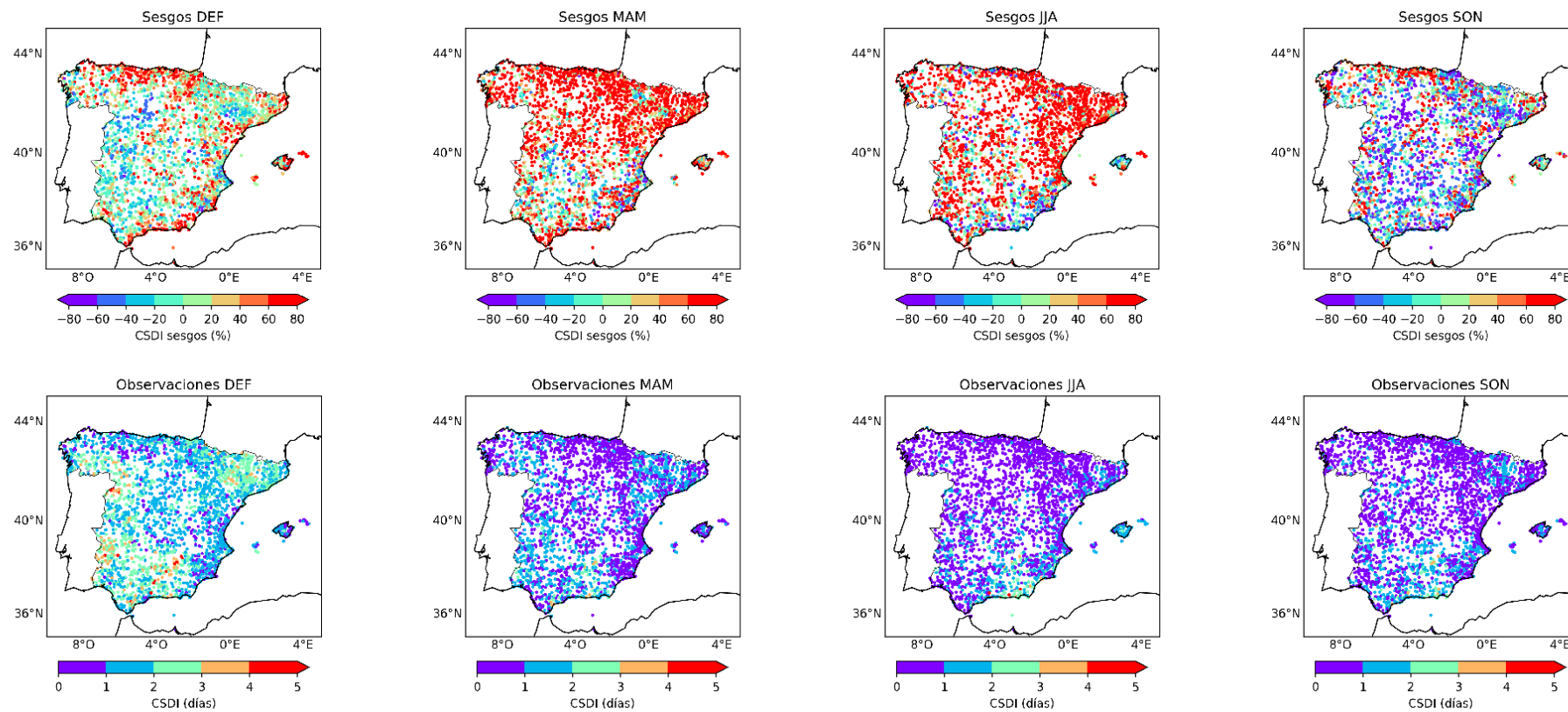


Figura A.8: Sesgos y observaciones estacionales para CSDI.

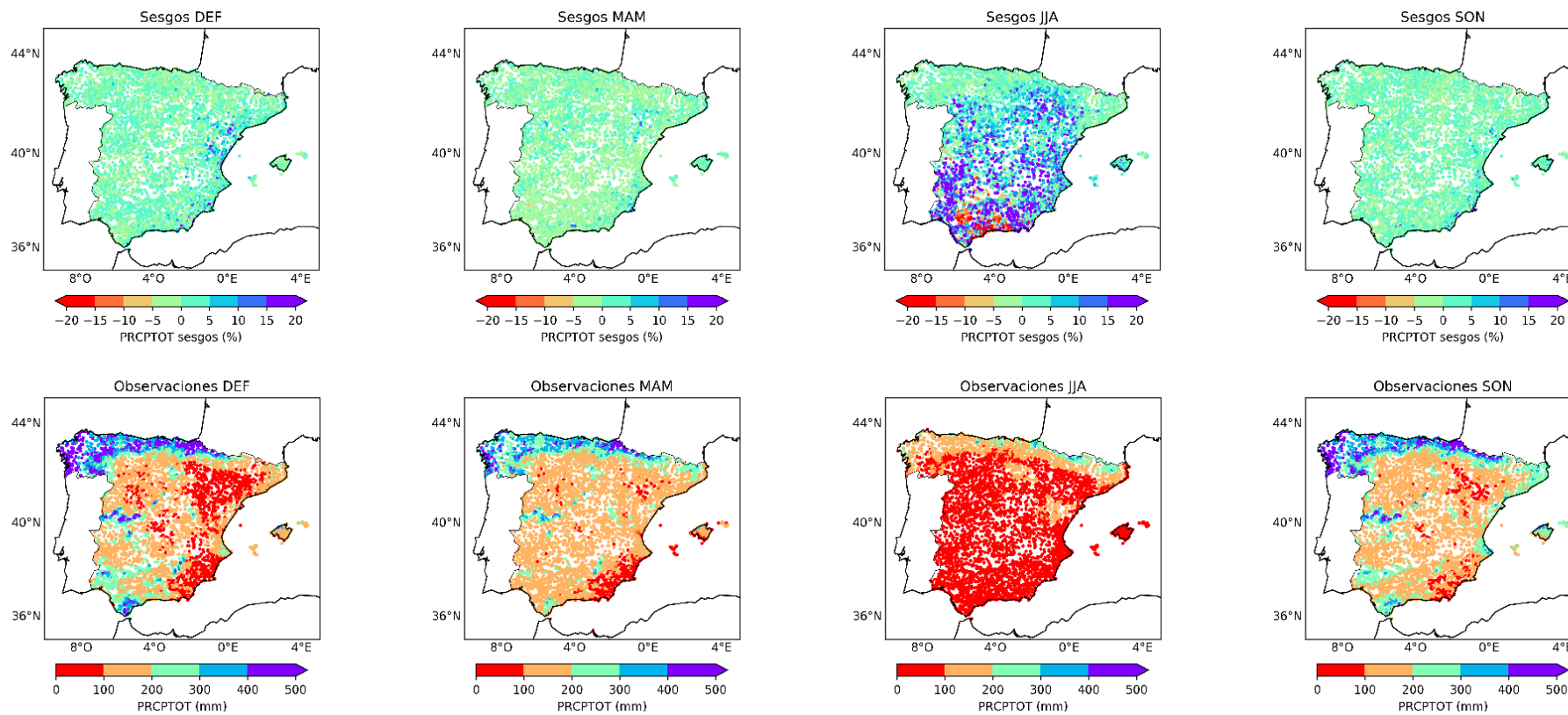


Figura A.9: Sesgos y observaciones estacionales para PRCPTOT.

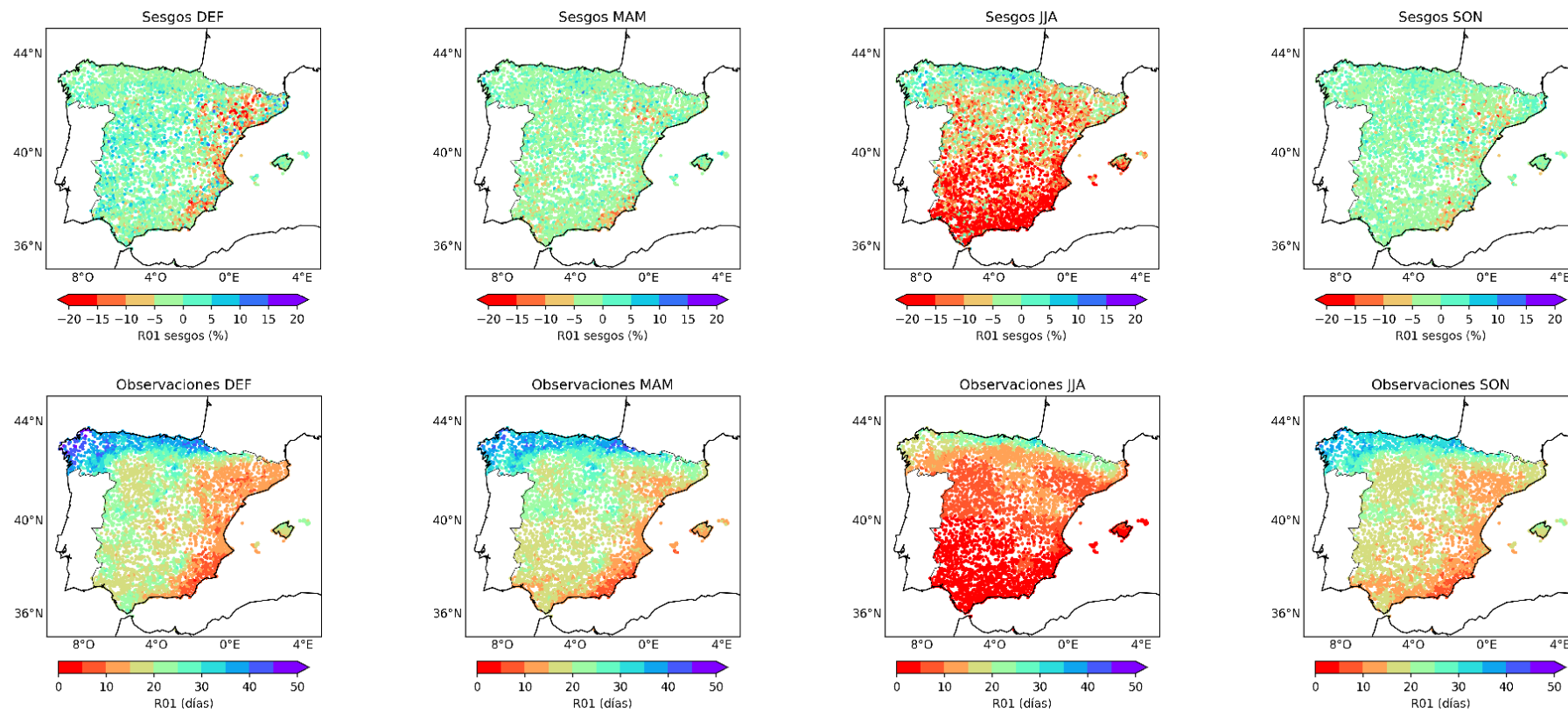


Figura A.10: Sesgos y observaciones estacionales para R01.

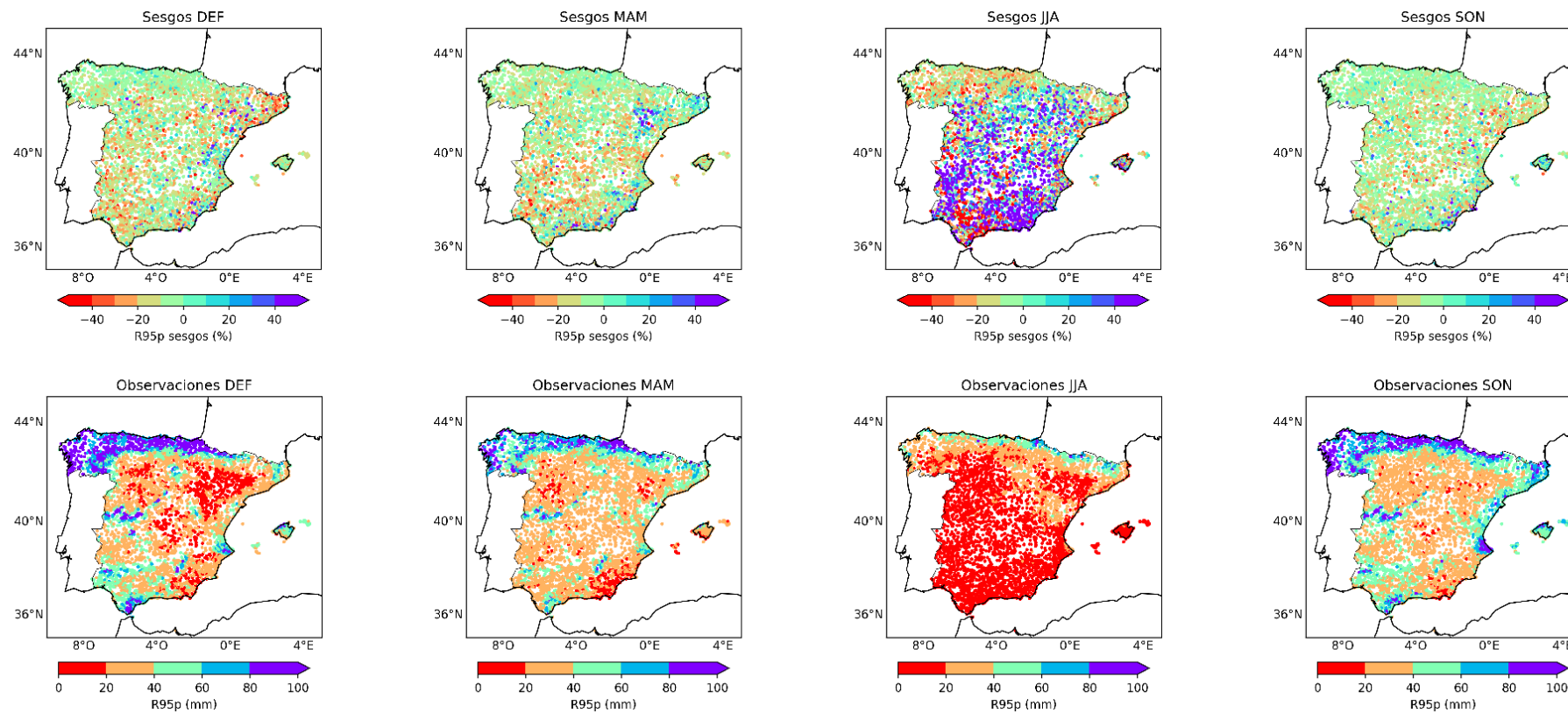


Figura A.11: Sesgos y observaciones estacionales para R95p.

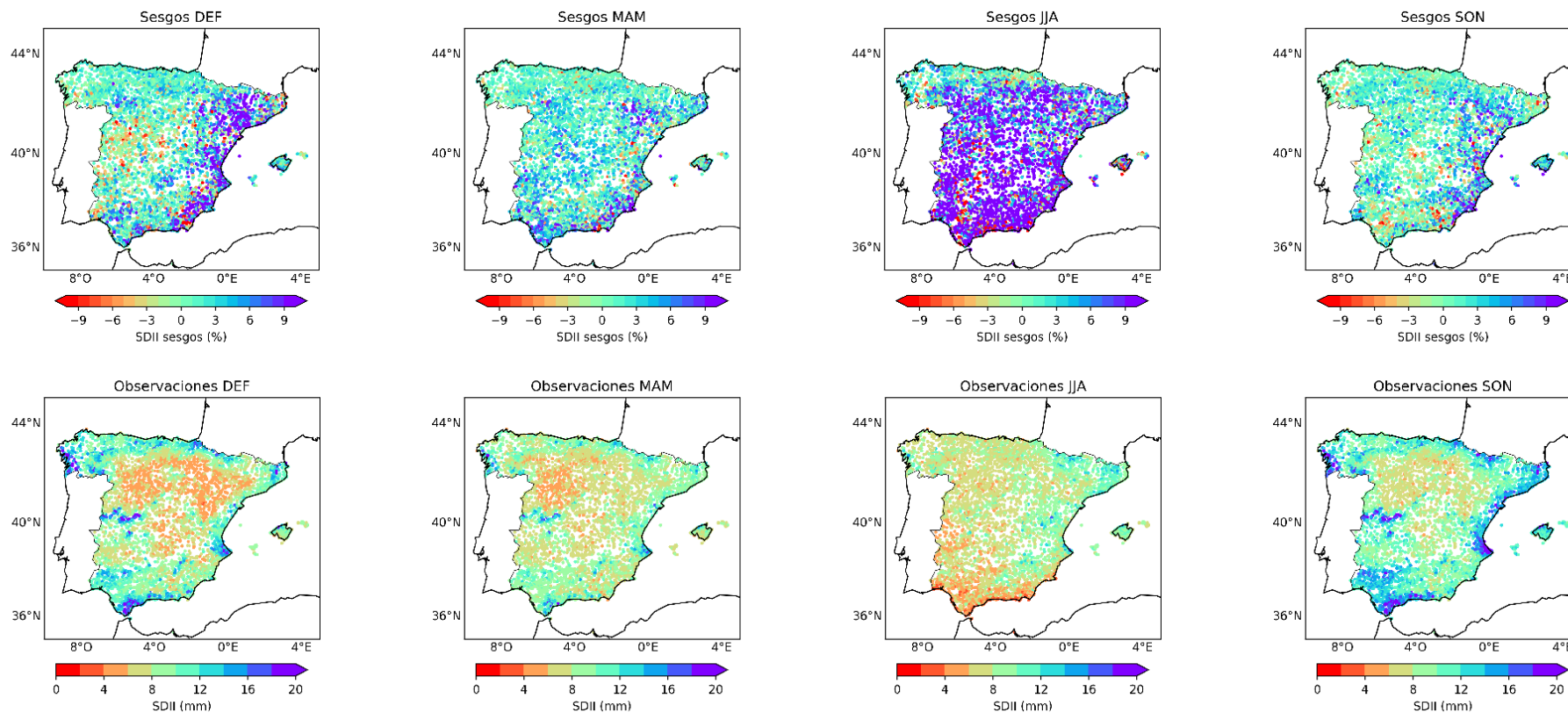


Figura A.12: Sesgos y observaciones estacionales para SDII.

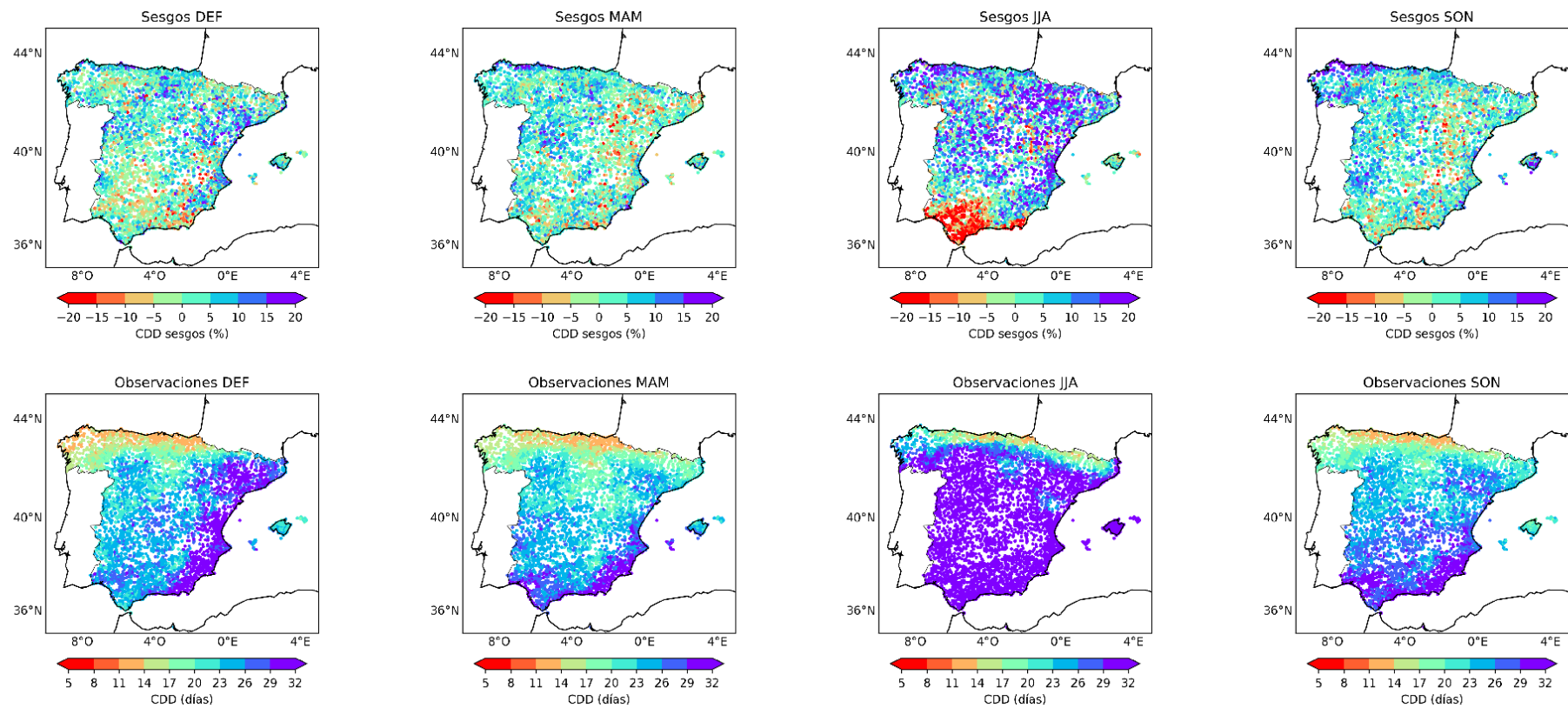


Figura A.13: Sesgos y observaciones estacionales para CDD.

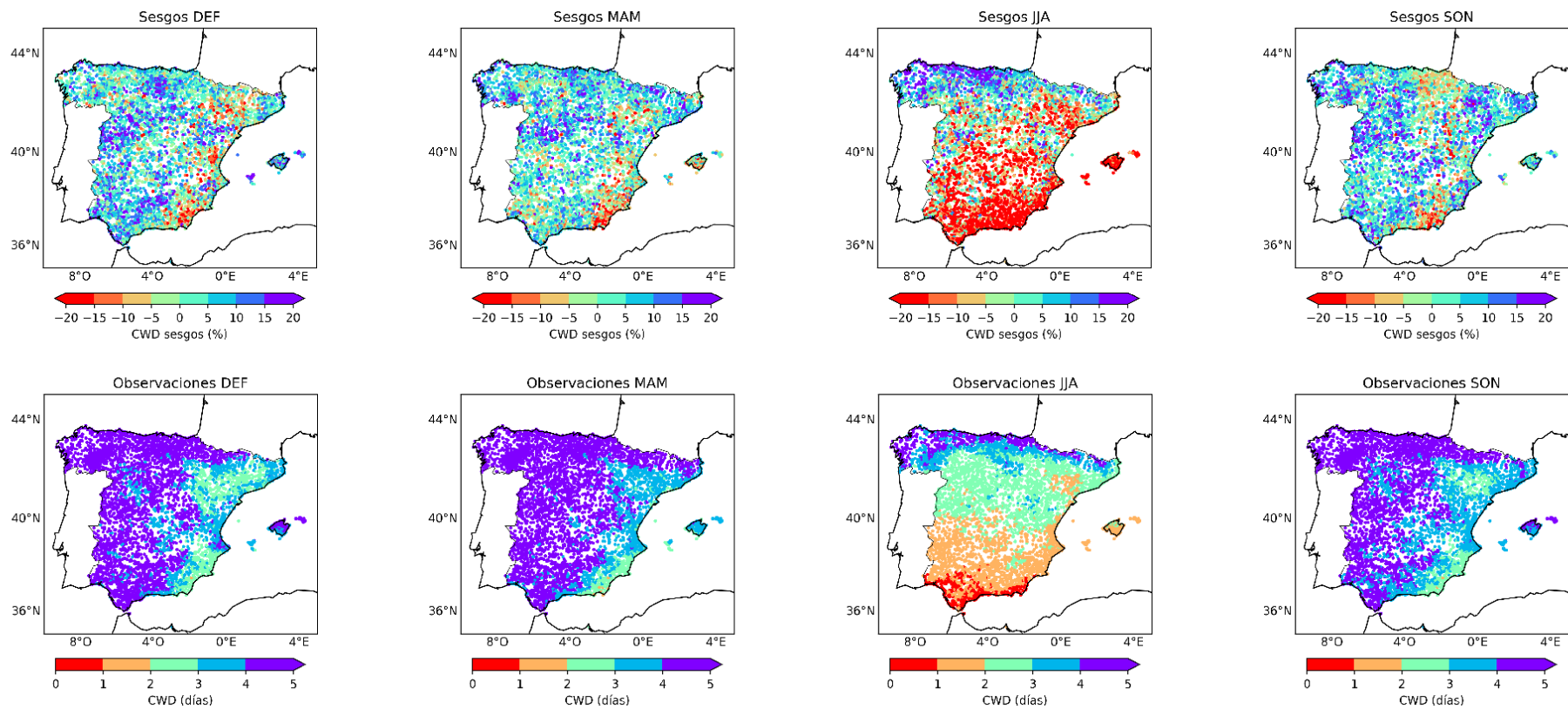


Figura A.14: Sesgos y observaciones estacionales para CWD.

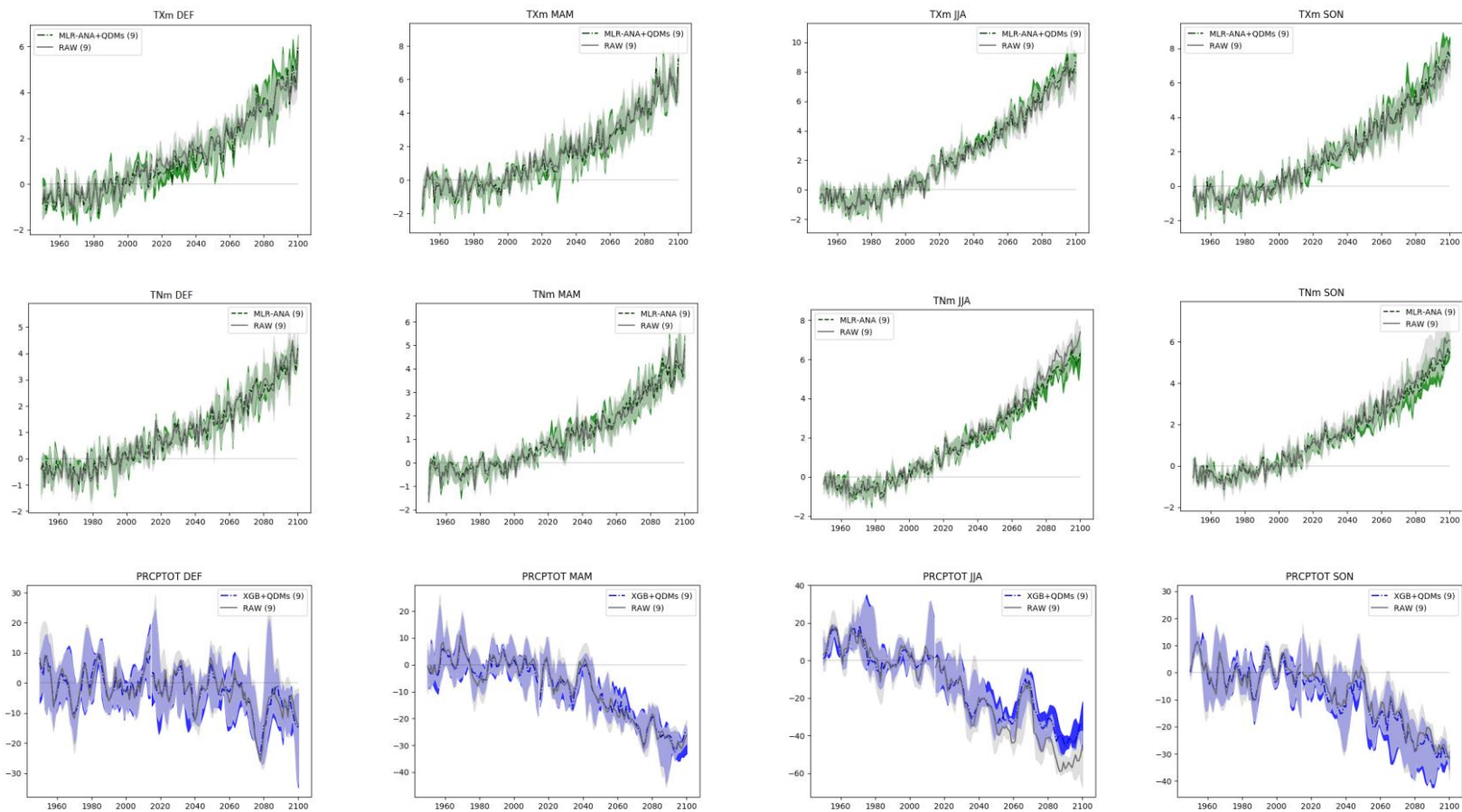


Figura A.15: Tendencias estacionales para las proyecciones regionalizadas con MLR-ANA+QDMs (color verde) en temperaturas y con XGB+QDMs (color azul) en precipitación en el escenario SSP5-8.5 comparadas con RAW (color gris). En columnas se tienen las estaciones: DEF, MAM, JJA y SON. Las filas corresponden a los índices: TXm, TNm y PRCPTOT

ANEXO B: SESGOS Y TENDENCIAS ESTACIONALES EN CANARIAS

En el presente anexo se proporcionan figuras adicionales por estaciones del año que complementan al apartado 4.3.b referido a los sesgos en Canarias.

Las figuras desde B.1 hasta B.14 (ambas incluidas) muestran mapas con los índices CLIMDEX analizados en el apartado 4.3.b en las estaciones de AEMET en Canarias de manera análoga a como se presentaron los resultados estacionales en Península y Baleares en el anexo A. Para cada índice CLIMDEX, se muestran mapas tanto de los sesgos de los valores regionalizados (índices de temperaturas mediante el método MLR-ANA+QDMs e índices de precipitación mediante el método XGB+QDMs) como del promedio de las observaciones en el periodo de referencia (1981-2010). Los mapas con los valores observados tienen especial relevancia para la interpretación de los índices CLIMDEX, cuyo sesgo se expresa en términos relativos, ya que si el valor observado es pequeño, el sesgo relativo puede ser alto (incluso cuando sea pequeño en términos absolutos).

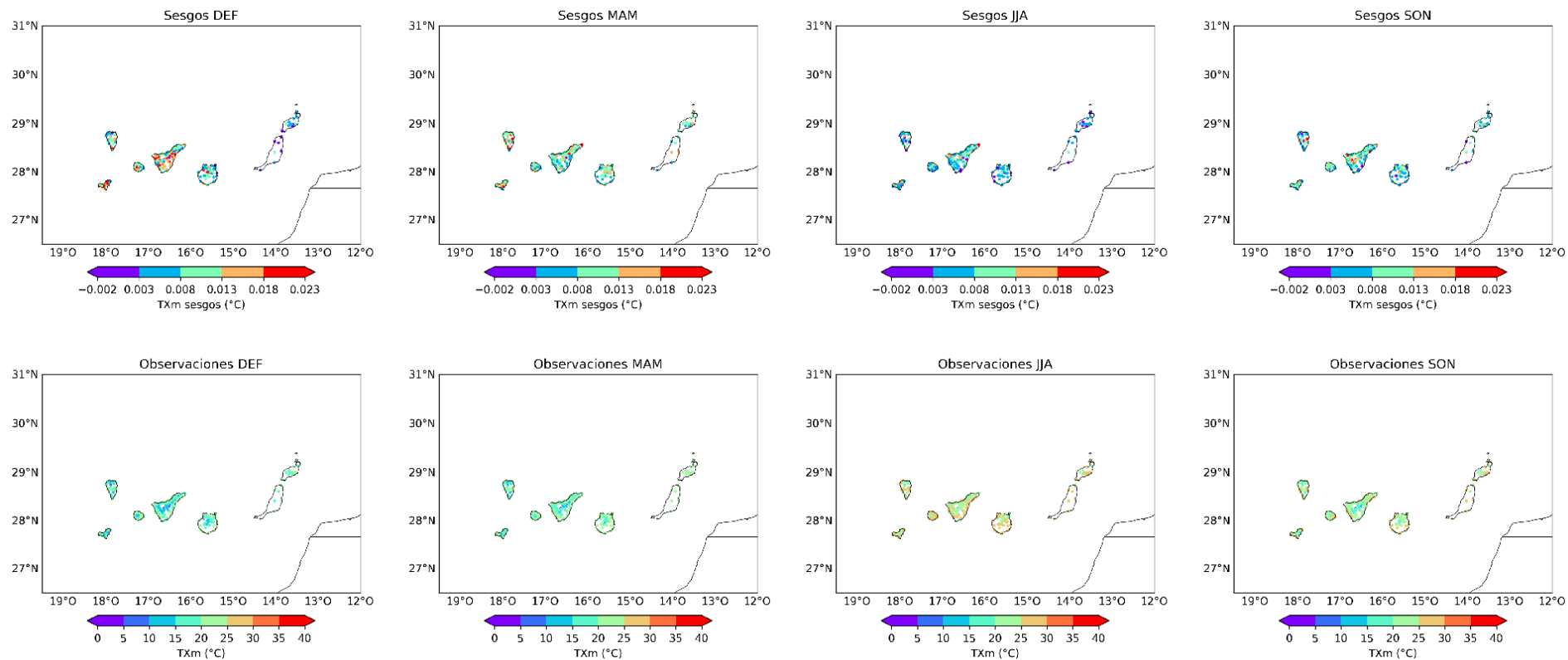


Figura B.1, sesgos y observaciones estacionales para TXm.

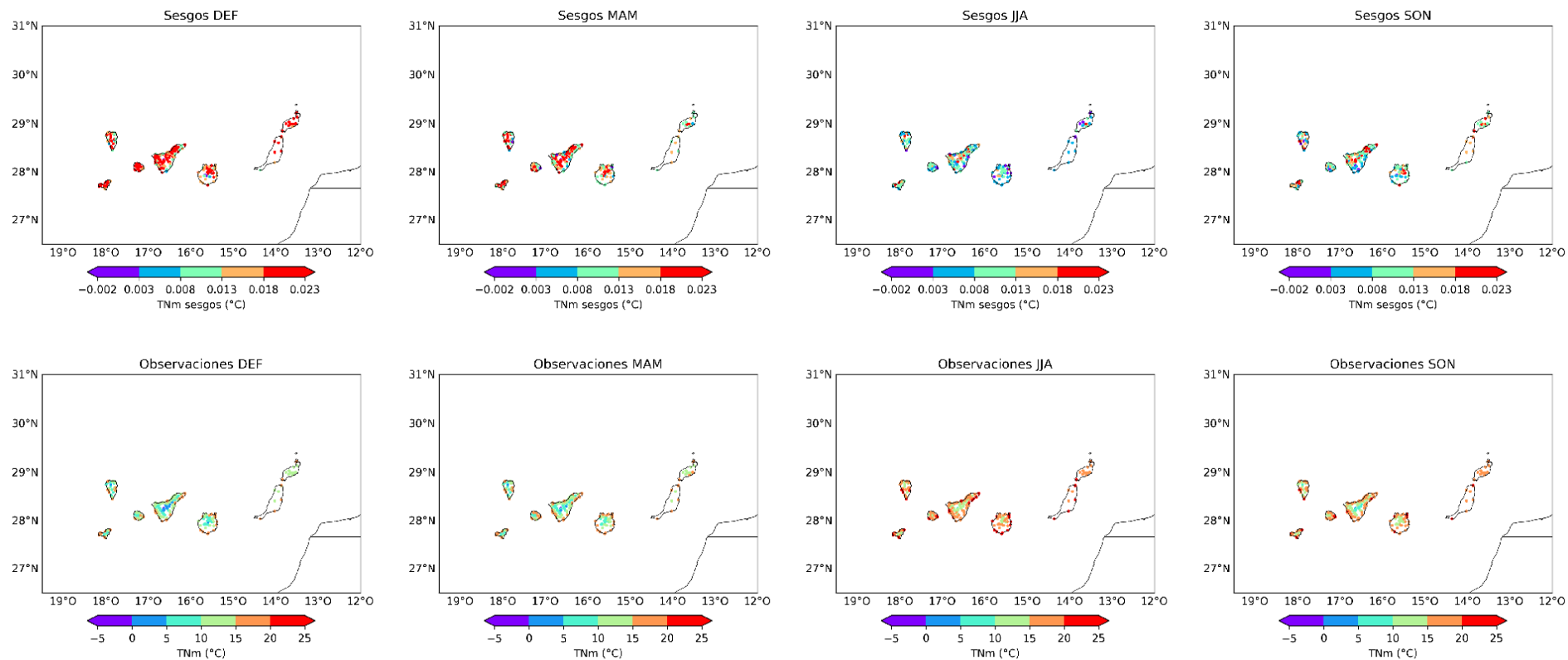


Figura B.2, sesgos y observaciones estacionales para TNm.

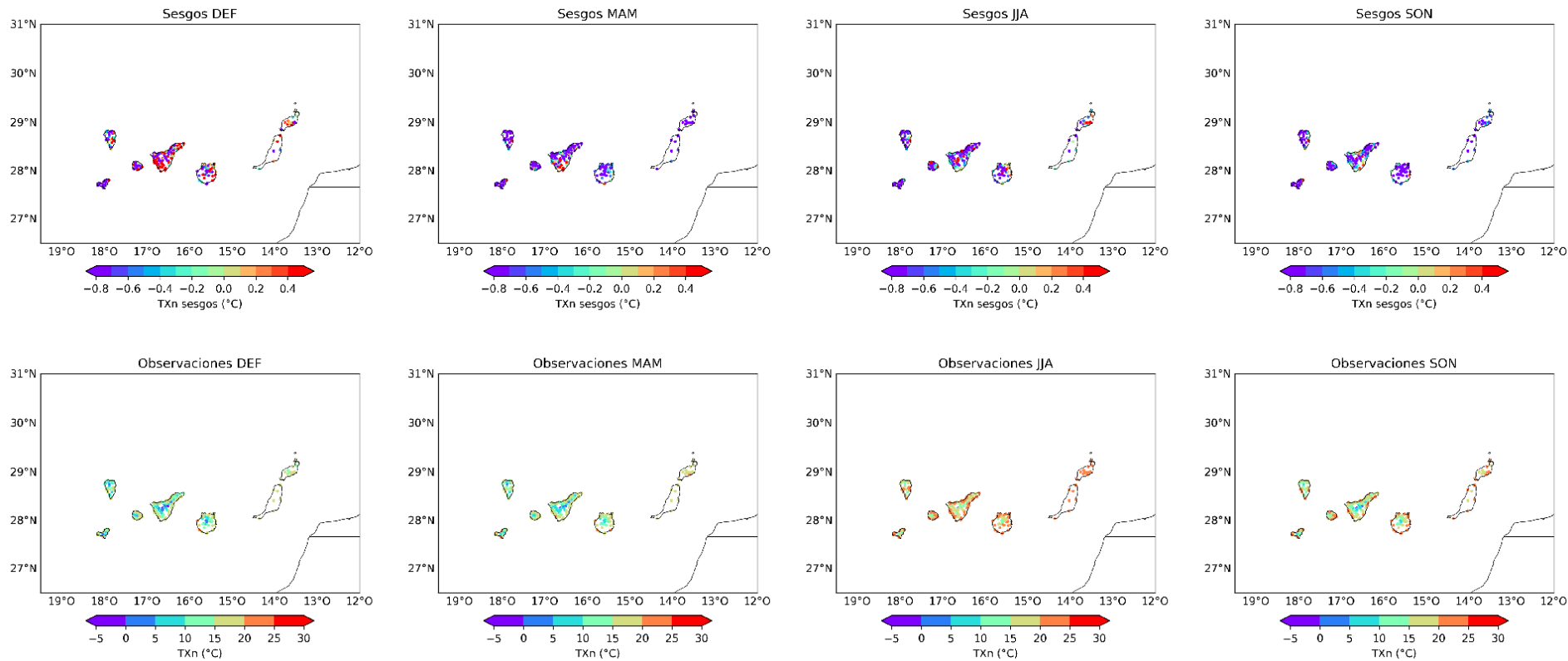


Figura B.3, sesgos y observaciones estacionales para TXn.

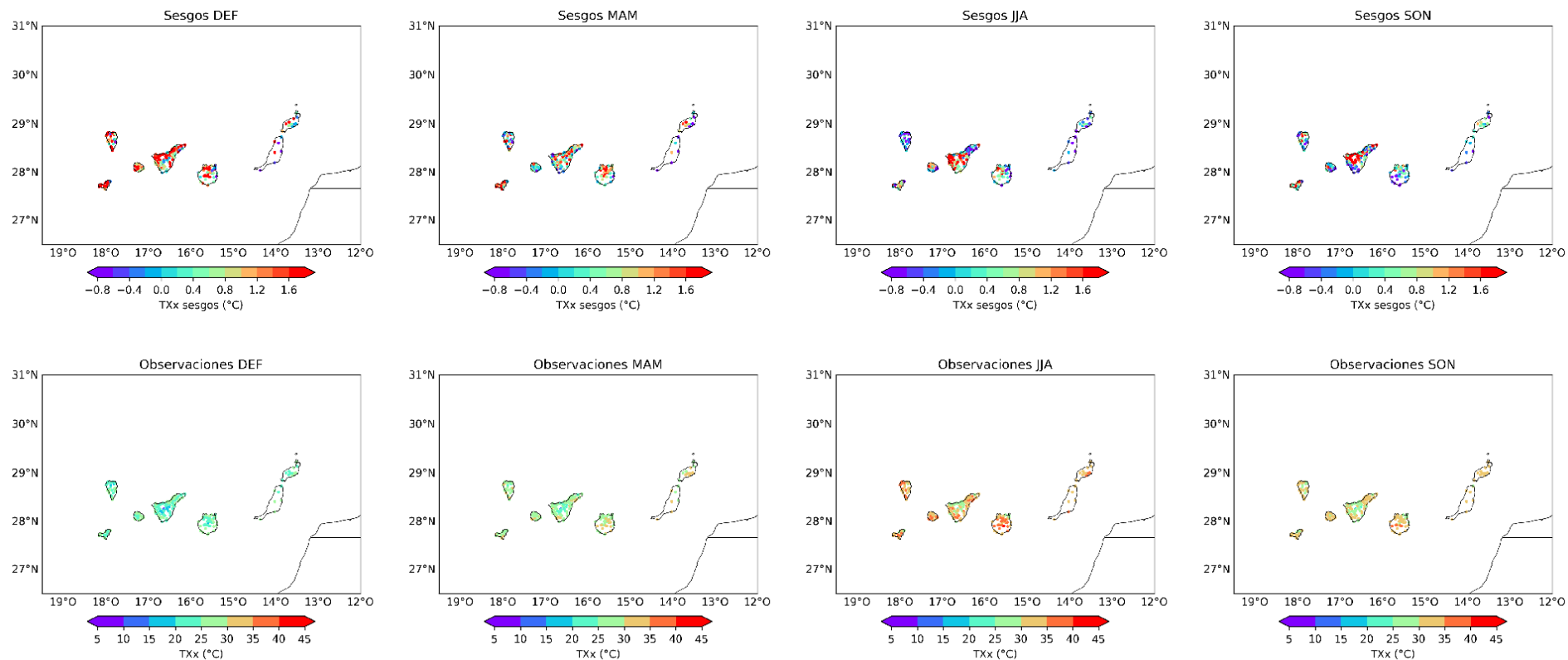


Figura B.4, sesgos y observaciones estacionales para TXx.

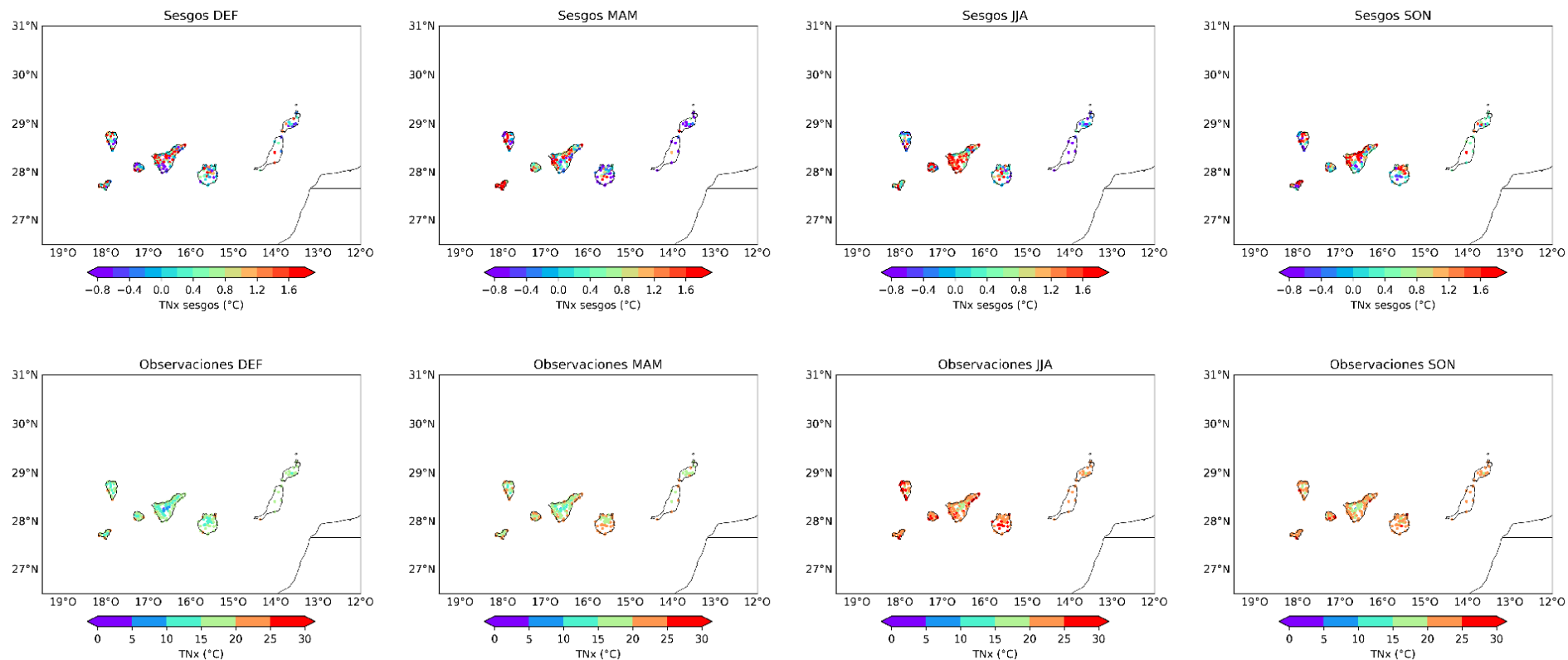


Figura B.5, sesgos y observaciones estacionales para TNx.

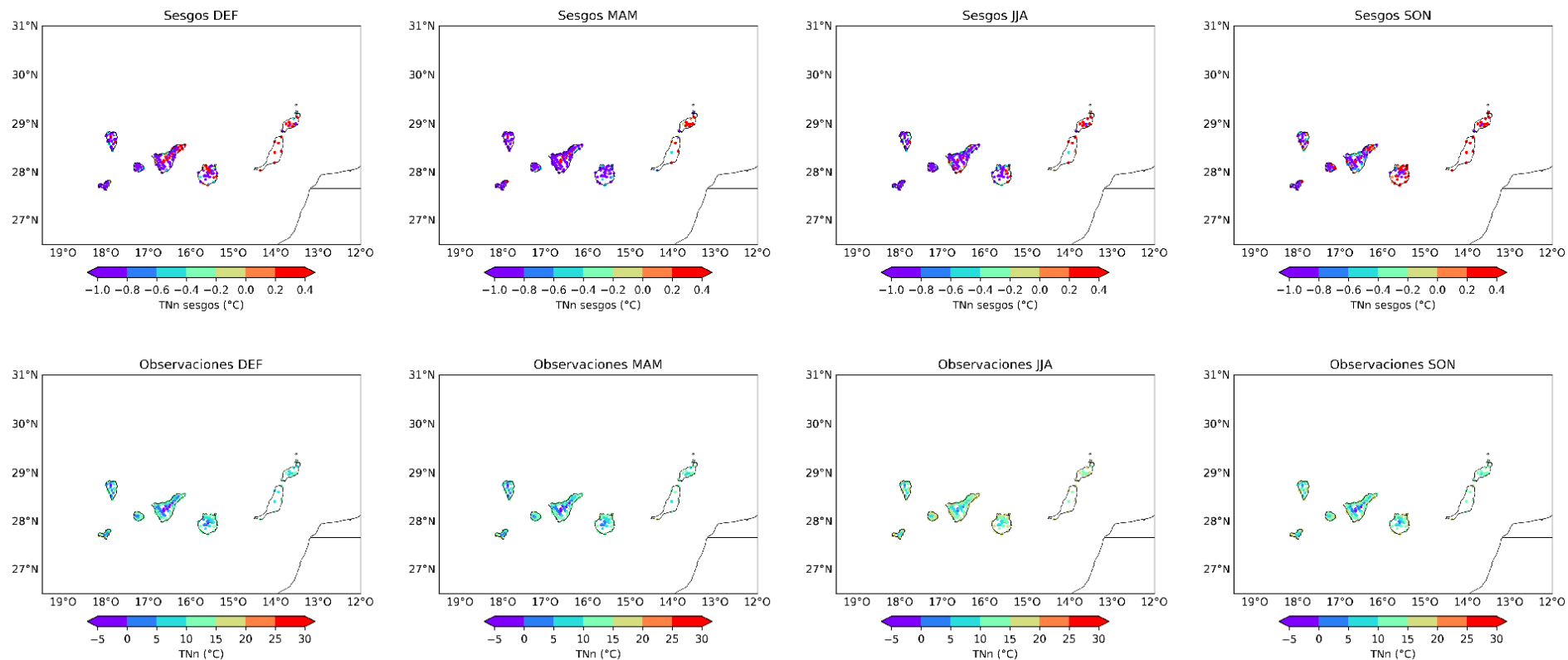


Figura B.6, sesgos y observaciones estacionales para TNn.

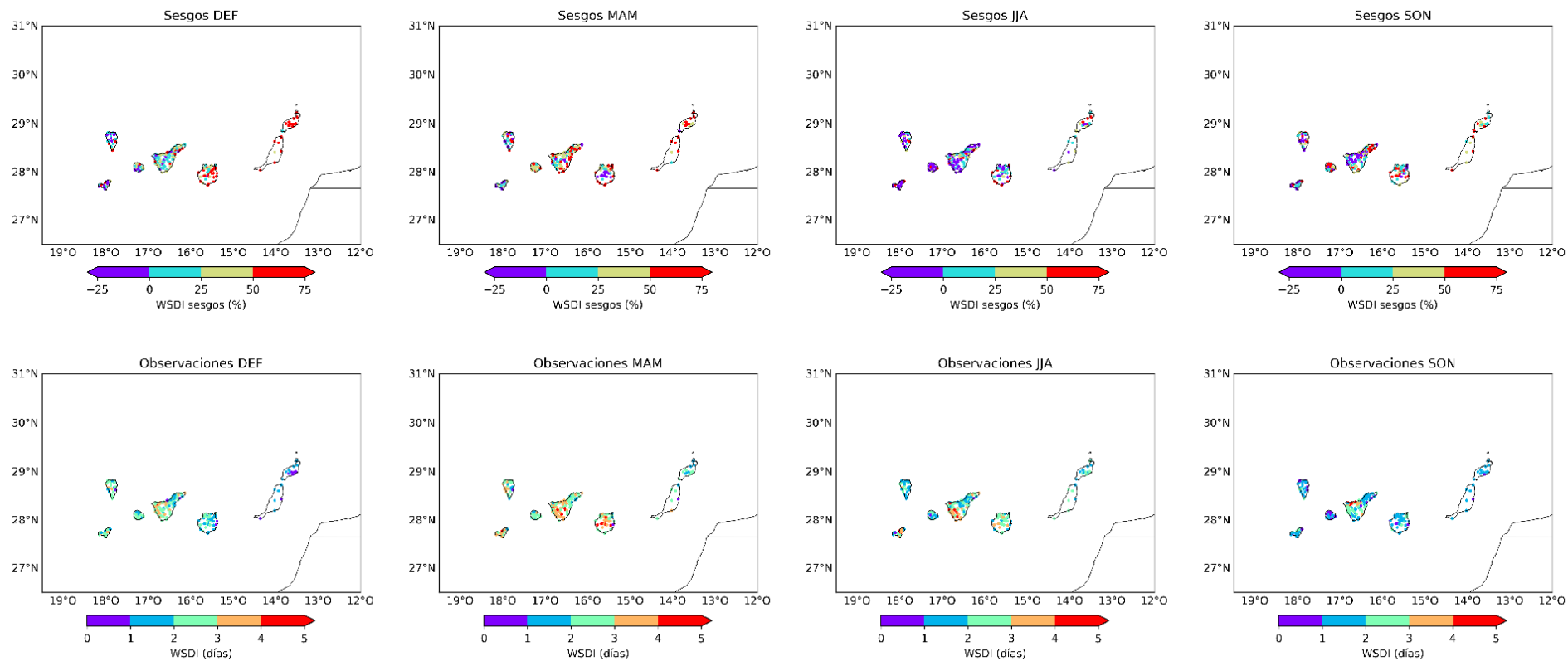


Figura B.7, sesgos y observaciones estacionales para WSDI.

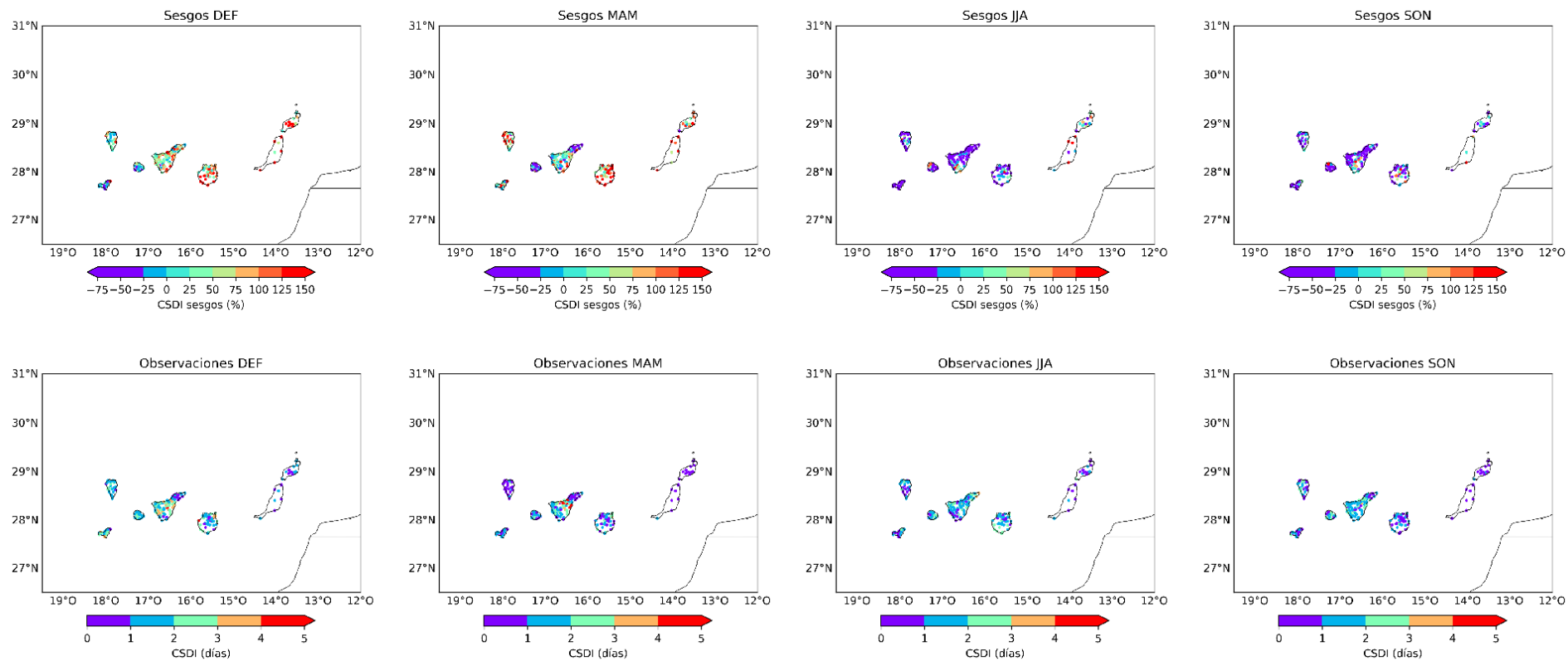


Figura B.8, sesgos y observaciones estacionales para CSDI.

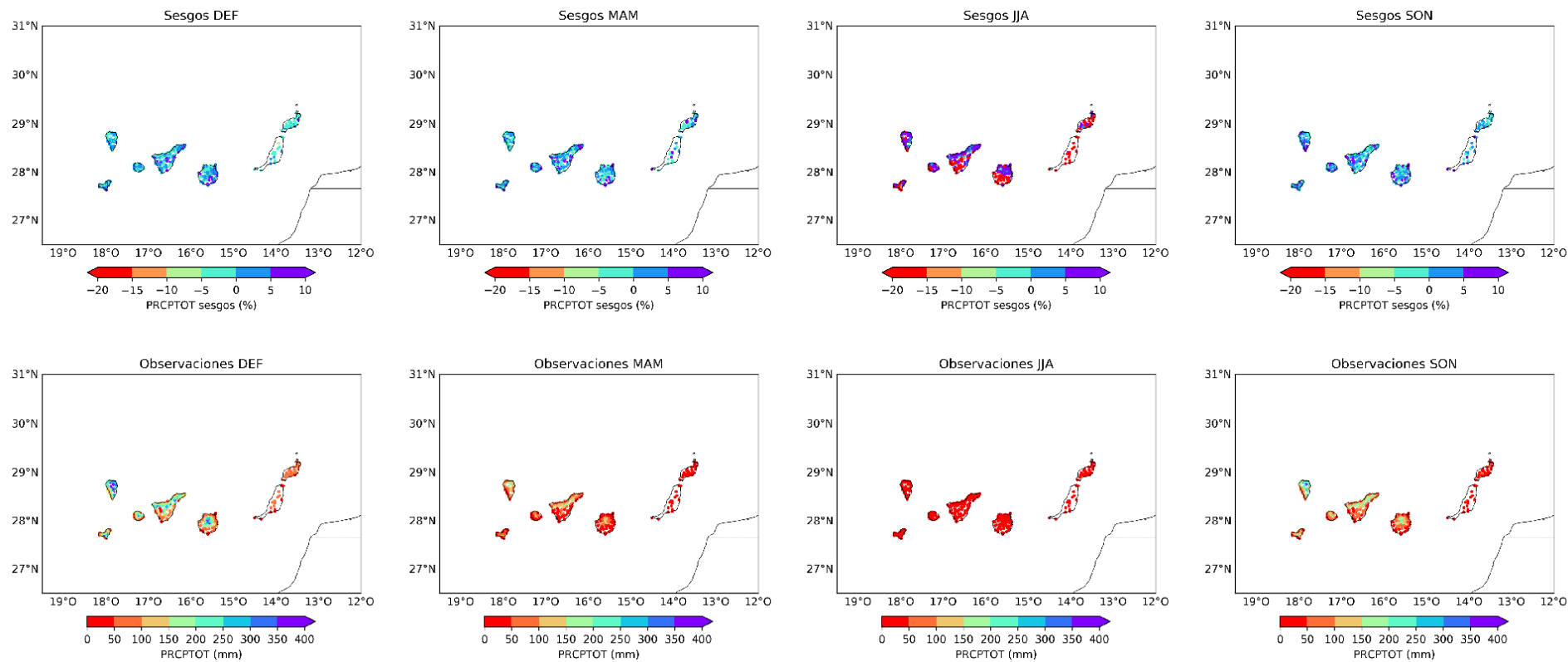


Figura B.9, sesgos y observaciones estacionales para PRCPTOT.

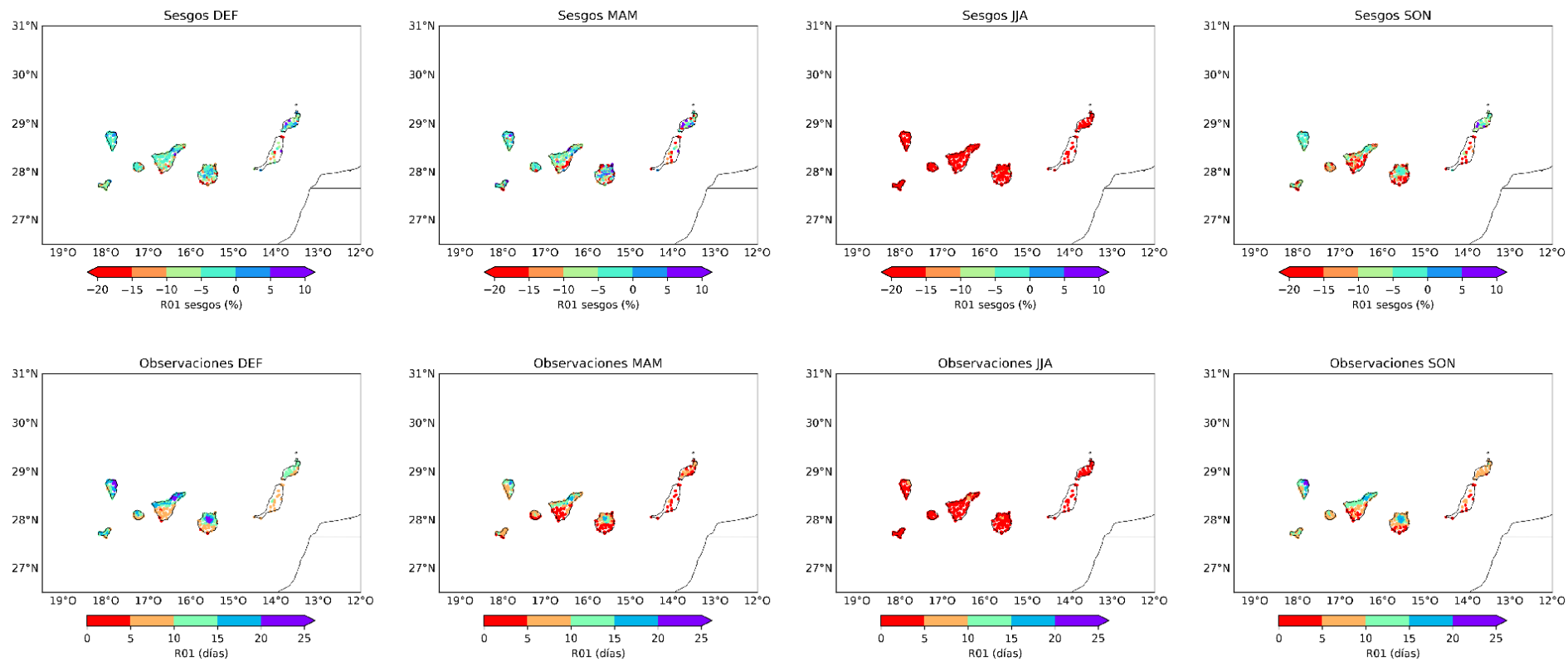


Figura B.10, sesgos y observaciones estacionales para R01.

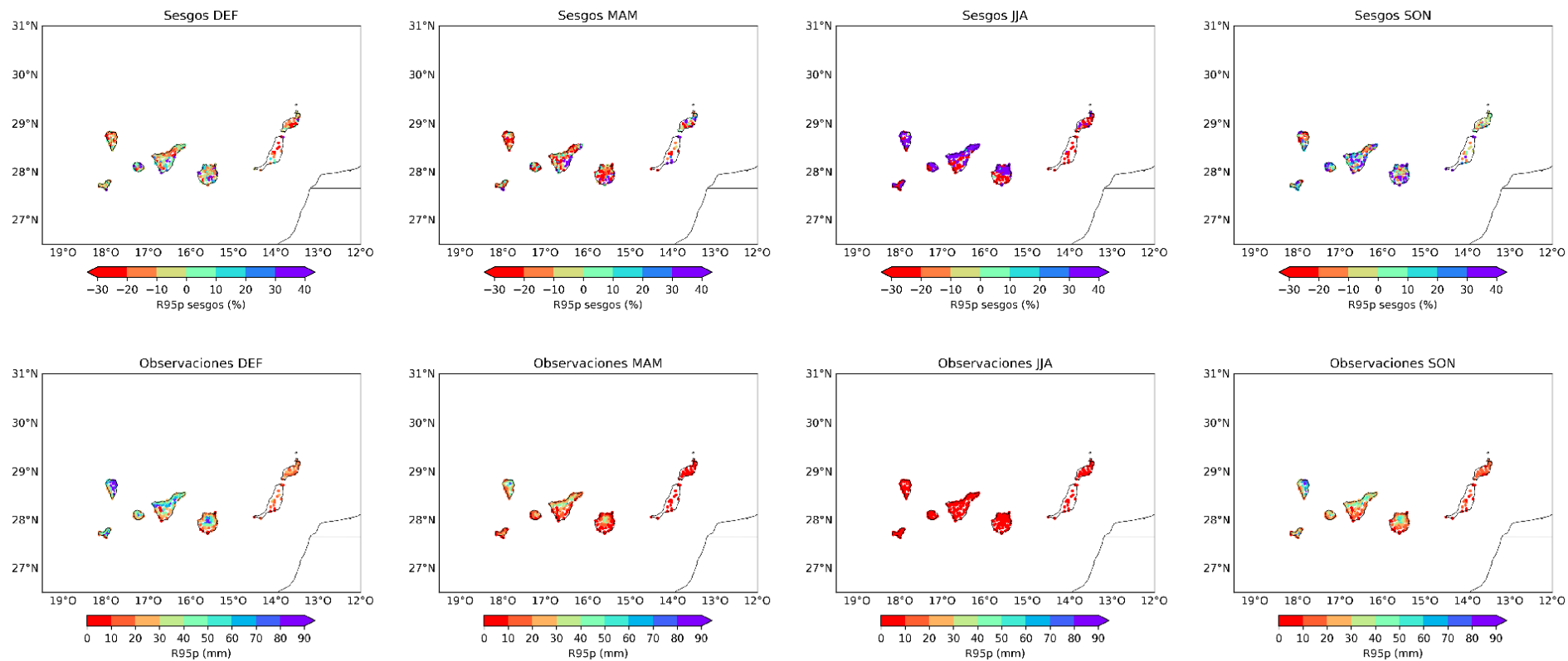


Figura B.11, sesgos y observaciones estacionales para R95p.

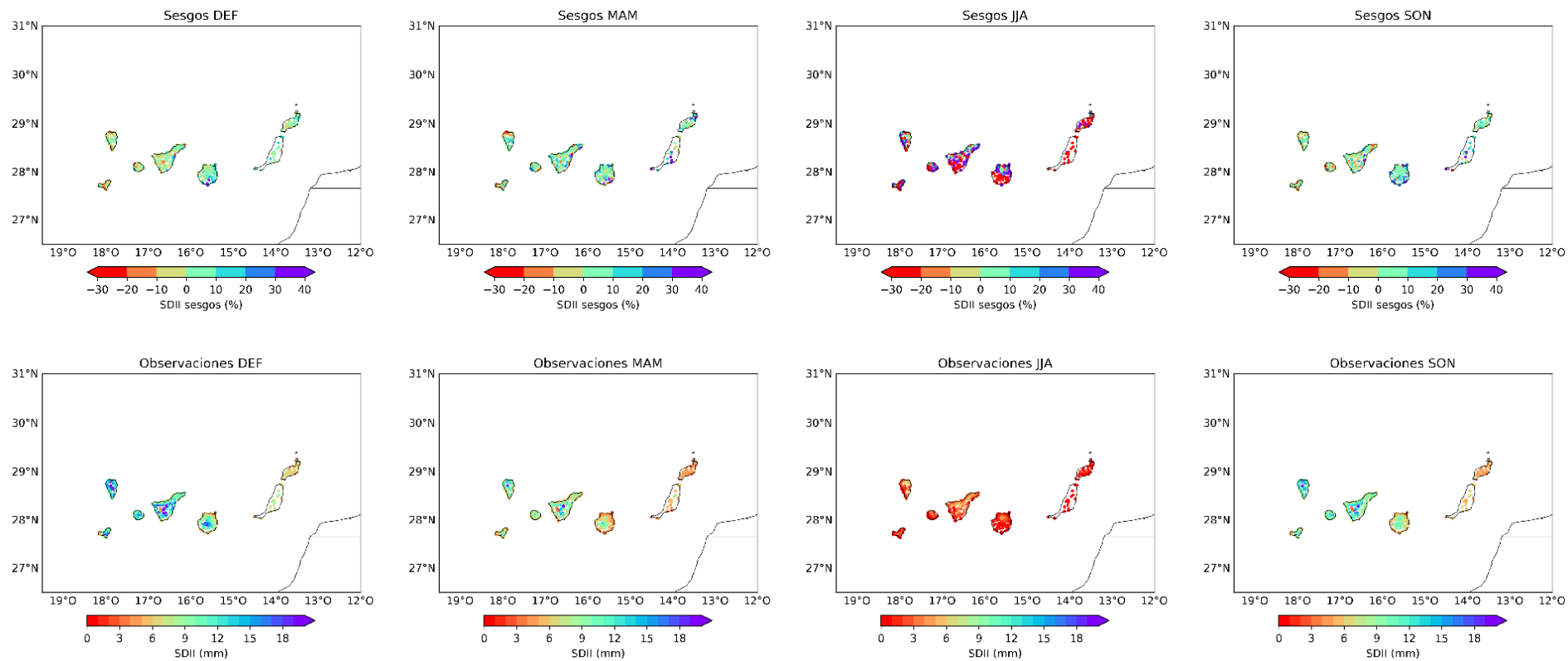


Figura B.12, sesgos y observaciones estacionales para SDII.

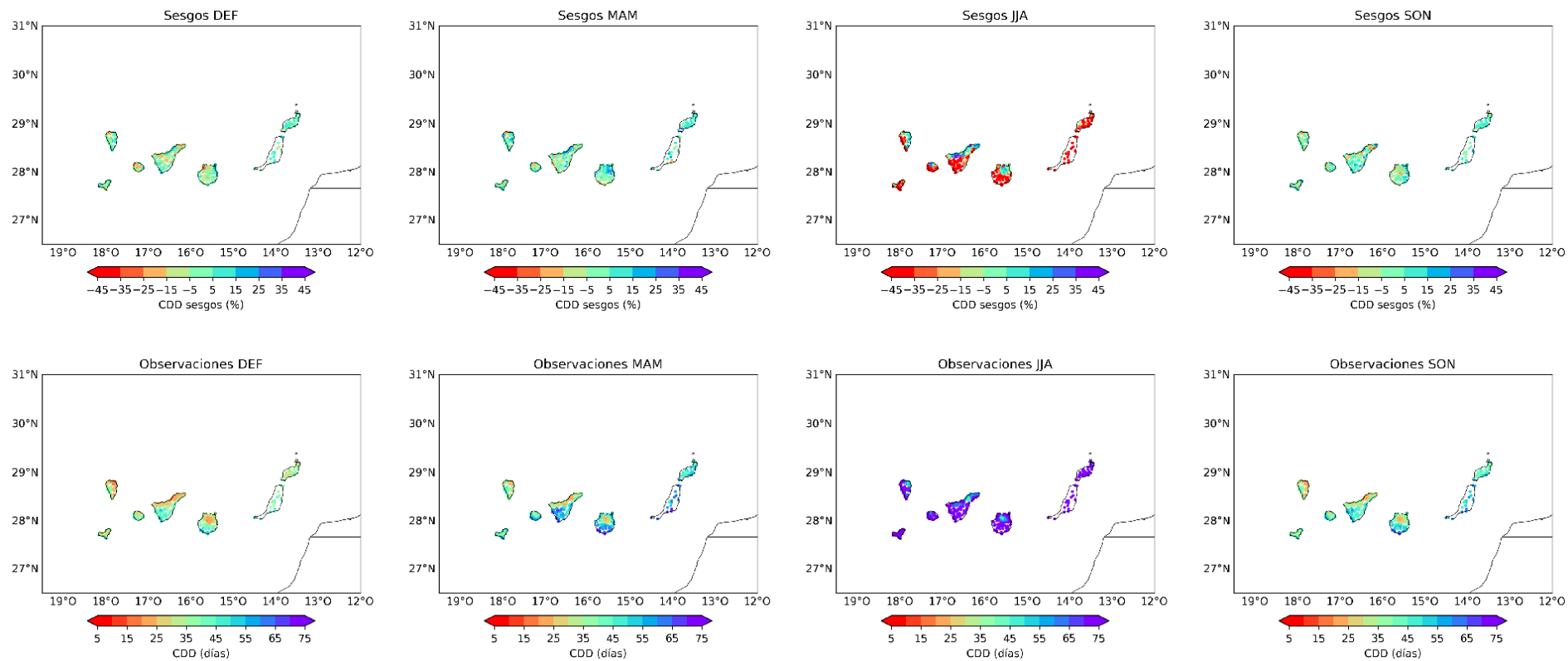


Figura B.13, sesgos y observaciones estacionales para CDD.

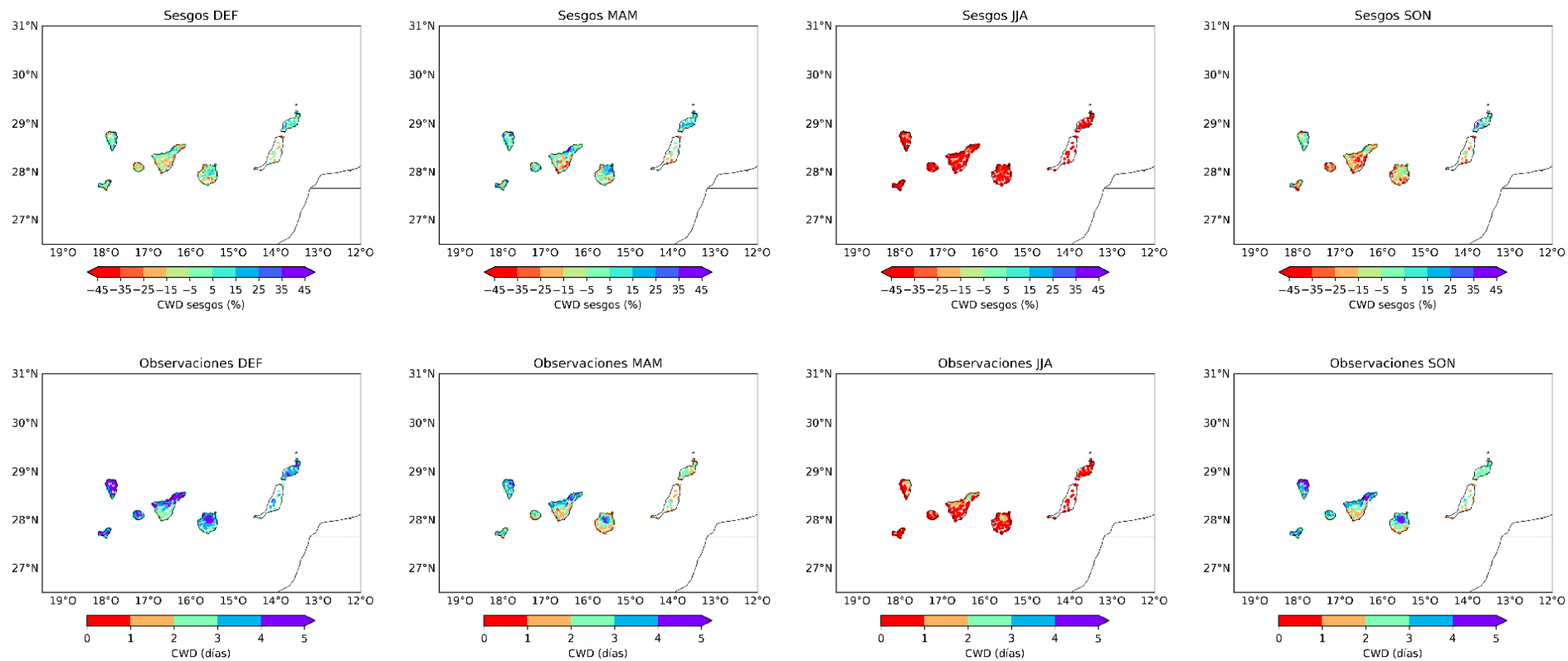


Figura B.14, sesgos y observaciones estacionales para CWD.