



Facultad de Ciencias

Idoneidad de técnicas univariantes y
multivariantes para la corrección de sesgos
en modelos climáticos regionales:
Aplicación para el cálculo de índices
agroclimáticos

Suitability of uni- and multi-variate bias correction
techniques in regional climate models: Applicability to the
computation of agroclimatic indices

TRABAJO FIN DE MÁSTER
Máster Universitario en Ciencia de Datos

Presentado por

JUAN JOSÉ VELASCO HORCAJADA

Dirigido por

RODRIGO GARCÍA MANZANAS

ANA CASANUEVA VICENTE

SEPTIEMBRE 2023

Resumen

Este estudio se centra en el análisis de las ventajas y limitaciones de distintas técnicas estadísticas que permiten corregir los sesgos (errores con respecto a la realidad observada) que presentan las simulaciones de un conjunto representativo de modelos regionales del clima en la Península Ibérica. En concreto, se consideran métodos de corrección univariados y multivariados y se evalúa cuál es su efecto sobre el cálculo de varios índices agroclimáticos multivariados que dependen de la temperatura y la precipitación. Estos índices resultan de especial interés para el sector vitivinícola, con el objetivo final de generar proyecciones futuras.

Los resultados obtenidos revelan que, mientras para el caso de la temperatura no existen diferencias apreciables entre las distintas técnicas de corrección de sesgos consideradas; los métodos univariados y los multivariados dan lugar a discrepancias importantes en el caso de la precipitación. En particular, el enfoque multivariado presenta ciertas deficiencias que merecen un estudio más en profundidad.

Además, las proyecciones futuras de los índices agroclimáticos analizados permiten identificar zonas en las que es esperable que el cambio climático tenga un impacto importante sobre el cultivo de la vid como consecuencia de un aumento en las temperaturas y una disminución en la precipitación. Por tanto, los resultados de este trabajo pueden ser de gran utilidad para la toma de decisiones y la implementación de medidas eficientes de adaptación al cambio climático.

Palabras clave: *modelos regionales del clima, índices agroclimáticos, técnicas de corrección de sesgos, cambio climático, viticultura*

Abstract

This study focuses on the analysis of the advantages and limitations of various statistical techniques that allow for the correction of biases (deviations from observed reality) present in simulations from a representative set of regional climate models in the Iberian Peninsula. Specifically, univariate and multivariate correction methods are considered, and their effect on the calculation of several multivariable agroclimatic indices dependent on temperature and precipitation is evaluated. These indices are of particular interest to the viticulture sector, with the ultimate goal of generating future projections.

The results obtained reveal that, while for temperature, there are no appreciable differences among the various bias correction techniques considered, both univariate and multivariate methods lead to significant discrepancies in the case of precipitation. In particular, the multivariate approach exhibits certain deficiencies that warrant further in-depth investigation.

Furthermore, the future projections of the analyzed agroclimatic indices allow the identification of regions where climate change is expected to have a significant impact on grapevine cultivation due to rising temperatures and decreased precipitation. Therefore, the findings of this study can be of great utility for decision-making and the implementation of effective measures for adapting to climate change.

Keywords: *Regional climate models, agroclimatic indices, bias correction techniques, climate change, viticulture.*

Agradecimientos

Quiero dar las gracias a mis directores, el Dr. Rodrigo García Manzanas y la Dra. Ana Casanueva Vicente, por su dedicación, esfuerzo, tiempo y apoyo a lo largo de todo este trabajo de investigación. Les agradezco enormemente haberme dado la oportunidad de participar en este proyecto y su confianza. Finalmente, quiero agradecer de todo corazón el apoyo incondicional de mi familia, amigos y mi pareja. Han sido un pilar fundamental y estoy eternamente agradecido por tenerlos a mi lado en cada paso que doy.

Índice general

1. Introducción	1
1.1. Motivación	5
1.2. Objetivos	5
1.3. Estructura	6
2. Datos	9
2.1. Región de estudio: La Península Ibérica	9
2.2. Observaciones	10
2.3. Modelos Regionales del Clima	11
2.4. Índices agroclimáticos	13
3. Métodos	17
3.1. Métodos de Corrección de Sesgo (<i>BC</i>)	17
3.1.1. Métodos Univariados	18
3.1.2. Métodos Multivariados	21
3.2. Marco de Validación	22
3.2.1. Validación Cruzada	22
3.2.2. Métricas de Validación	23
3.3. Aplicación a condiciones de cambio climático	25
4. Resultados y Discusión	27
4.1. Indicadores agroclimáticos observados en clima presente	27
4.2. Evaluación de los métodos de corrección de sesgos	28
4.2.1. Validación cruzada con años aleatorios	28
4.2.2. Capacidad de los métodos para representar años cálidos	39

4.3. Proyecciones de cambio climático de los índices agroclimáticos	42
5. Conclusiones Principales y Trabajo Futuro	47
5.1. Conclusiones Principales	47
5.2. Trabajo Futuro	48
5.3. Reproducibilidad de los Resultados	49
5.4. Contexto de la Investigación	50
 Bibliografía	 51

CAPÍTULO 1

Introducción

El aumento de la frecuencia y la intensidad de los eventos climáticos extremos se ha convertido en una creciente preocupación y un tema central de investigación en la actualidad. La ocurrencia de estos eventos tienen repercusiones globales que afectan a la mayoría de sectores socio-económicos, incluyendo la salud, el turismo, la energía y, en particular, la agricultura (Casanueva, 2016).

El cambio climático puede manifestarse a través de fenómenos como aumentos en la temperatura, en la variabilidad en los patrones de lluvia y en la aparición de sequías prolongadas, lo cual plantea importantes desafíos en el manejo de los cultivos agrícolas. En este contexto es particularmente relevante tratar de comprender y caracterizar mejor la ocurrencia de fenómenos extremos que pueden dar lugar a riesgos importantes (Serrano-Notivoli et al., 2020).

La región mediterránea, particularmente España, Francia e Italia, representan aproximadamente el 50% (<https://www.oiv.int/>) del total de la superficie vitivinícola del mundo (Cardell et al., 2019), por lo que su economía se ve afectada en gran medida por los cambios en las condiciones meteorológicas y climáticas (Jones, 2006), que juegan un papel crítico en el crecimiento de la uva y, por ende, en la calidad y cantidad del vino producido (Lorenzo et al., 2013; van Leeuwen et al., 2004; Santos et al., 2011).

Ante esta realidad, comprender cómo evolucionarán las condiciones climáticas en estas regiones en el futuro resulta crucial para anticipar y mitigar los riesgos potenciales, al tiempo que puede ayudar a guiar la toma de decisiones en el sector

vitivinícola que fomenten el cultivo de variedades de vid más resilientes y promuevan el uso de prácticas sostenibles y adaptativas (Liu et al., 2014).

Durante las últimas décadas se han desarrollado diferentes índices agroclimáticos (Huglin, 1978; Gladstones et al., 1992; Bois, 2016; Roell et al., 2020) que permiten evaluar la idoneidad de una región determinada para distintos tipos de cultivo. Estos índices tienen en cuenta diversas variables meteorológicas como la temperatura, la velocidad del viento, la precipitación o las horas de insolación recibidas. En el caso concreto de la vid destacan tres índices: el índice de BEDD (*Biologically Effective Degree-Days*) y el índice de Huglin, que tienen en cuenta el calor acumulado durante la temporada de crecimiento de la uva; y el índice de Branas, Bernon y Levadoux, que evalúa la probabilidad de que la vid sea afectada por el moho (Lorenzo et al., 2013).

Para obtener resultados precisos y fiables en estos índices, es esencial contar con datos de alta calidad que se basen en observaciones meteorológicas precisas y/o modelos climáticos confiables. Estos datos son la base sobre la cual se calculan los valores de los índices y, por lo tanto, resultan cruciales para el desarrollo de estrategias eficaces en el sector agrícola. Los modelos climáticos representan, de una forma simplificada, los procesos físicos que conforman el sistema climático de la Tierra, abarcando desde la dinámica atmosférica hasta los flujos oceánicos (Cannon, 2016) (entre otros). A medida que se va mejorando la modelización numérica que describe los procesos físicos (Meehl, 1995) y se incrementa la capacidad de cómputo, junto con la incorporación de observaciones más precisas, la fidelidad de las simulaciones provenientes de estos modelos también aumenta.

Los Modelos Climáticos Globales o GCMs (*Global Climate Models*) presentan una resolución espacial grosera (en torno a las decenas de km en el caso de los más avanzados), por lo que, a pesar de describir de manera fidedigna la circulación global, no pueden describir adecuadamente los procesos que ocurren en escalas espaciales regionales/locales, especialmente en zonas con una orografía compleja como la Península Ibérica (ver Figura 1.1). En consecuencia, en las últimas décadas se han ido desarrollando diferentes técnicas de reducción de escala (Hewitson and Crane, 1996; Kidson and Thompson, 1998; Maraun et al., 2010) (o *downscaling* en inglés) con el objetivo de trasladar la información climática a gran escala de los GCMs a la escala regional requerida para la evaluación de los impactos derivados del clima.

Por un lado, los Modelos del Clima Regionales o RCMs (*Regional Climate Models*) son la herramienta utilizada para un *downscaling* de tipo dinámico. En concreto, mediante técnicas de interpolación y cálculo numérico, los RCMs resuelven las ecuaciones que rigen la dinámica de la atmósfera en una rejilla tridimensional a

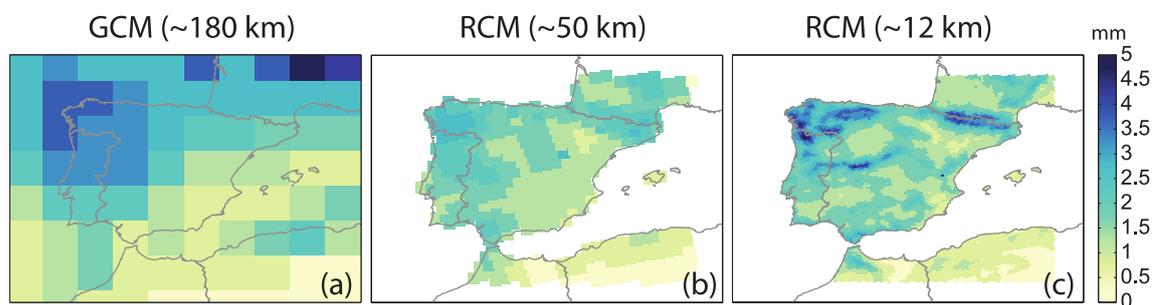


Figura 1.1: Ejemplo de diferentes resoluciones espaciales (tamaño de píxel) en modelos del clima y técnicas de regionalización. Se muestra la precipitación media anual (en mm) para el año 1989 representada por (a) un GCM (ECHAM 5, escenario histórico), un RCM (COSMO CLM, anidado a reanálisis) en sus versiones de resolución gruesa (b) y alta (c). Fuente: Casanueva (2016)

mayor resolución espacial que los GCMs y en un dominio espacial limitado, incorporando condiciones iniciales y de contorno de los GCMs (Feser et al., 2011; Casanueva, 2016). La realización de dichas simulaciones requiere la utilización de importantes recursos computacionales y, a menudo, se debe buscar un equilibrio entre la resolución espacial, el tamaño del dominio de estudio y el alcance temporal. Por ejemplo, la resolución espacial típica de los RCMs en simulaciones multi-decadales es de aproximadamente 50km en dominios continentales pudiendo llegar a 12km en dominios más pequeños (Figura 1.1) o incluso a 2-3km en el caso de simulaciones de una década (Ban et al., 2014). Además del núcleo dinámico, los RCMs representan a través de parametrizaciones físicas los procesos que ocurren a una escala más pequeña que la resolución espacial de la rejilla en que se desarrollan. Algunas parametrizaciones comunes son la de la capa límite planetaria, convección, microfísica, radiación y suelo y vegetación.

Aunque los RCMs son una herramienta importante para comprender el cambio climático y sus efectos en diferentes regiones del mundo y presentan un valor añadido frente a los GCMs (Rummukainen, 2016; Sørland et al., 2018), a menudo las variables climáticas simuladas por dichos modelos no concuerdan con las observaciones (Teutschbein and Seibert, 2012) y, por lo tanto, podrían no ser útiles directamente. A pesar de los esfuerzos continuos por mejorar la precisión de los RCMs, estos continúan presentando importantes sesgos sistemáticos como consecuencia, entre otros factores, de la imposibilidad para modelizar adecuadamente ciertos fenómenos físicos que ocurren en escalas espaciales más pequeñas que la resolución nativa (Stevens and Bony, 2013; Teutschbein and Seibert, 2012).

Como se ha ido señalando en estudios de la última década (Teutschbein and Seibert, 2012; Maraun et al., 2015; Cannon, 2018), los errores sistemáticos presentes en los RCMs pueden, en principio, corregirse mediante técnicas de corrección de sesgos (BC, por sus siglas en inglés), que suelen resultar baratos y rápidos desde el punto de vista computacional. En líneas generales, las técnicas de BC tratan de modificar la distribución estadística de los datos simulados para hacerla compatible con la de las observaciones que se toman como referencia. Aunque en la actualidad existen una gran variedad de algoritmos que persiguen este mismo fin, el *mapeo cuantil-cuantil* es uno de los más utilizados en estudios de impacto del cambio climático y pronósticos meteorológicos (Cannon, 2016). Típicamente, el mapeo cuantil a cuantil y la gran mayoría de métodos de corrección de sesgos se aplican de manera individual a cada variables climática simulada por el RCM, sin considerar la dependencia espacio-temporal que existe entre distintas variables, lo cual puede constituir una limitación clave en aplicaciones que requieren del uso simultáneo de más de una variable climática (Rocheta et al., 2014); por ejemplo, el cálculo de algunos índices agroclimáticos (Vicente-Serrano et al., 2010) o índices de incendios (Wang et al., 2015).

Como alternativa a los métodos univariados de corrección de sesgos, en los últimos años han aparecido algoritmos multivariante o MBC (*Multivariate Bias Correction*), que corrigen los sesgos en múltiples variables simultáneamente. Como ejemplo, Cannon (2016) ha presentado dos técnicas de corrección de sesgo multivariante, MBCp y MBCr, que pretenden preservar la correlación de Pearson y Spearman, respectivamente, entre dos o más variables corregidas. En 2018, el mismo autor (Cannon, 2018) introdujo un nuevo algoritmo multivariante conocido como MBCn. Este método se describe como una modificación del algoritmo *N-pdf*, originalmente empleado en visión artificial, adaptado para un uso en aplicaciones en las que intervienen múltiples variables de modelos climáticos. Aunque estos algoritmos han supuesto un avance en la corrección de sesgos, se siguen analizando sus limitaciones.

En el marco de este Trabajo Fin de Máster se han empleado técnicas estadísticas de corrección de sesgos para abordar las limitaciones presentes en una serie de RCMs centrándonos en el cultivo de la vid en la Península Ibérica. El objetivo principal de este estudio es evaluar la efectividad de estas técnicas, y para lograrlo, se ha puesto el foco en el análisis de los índices agroclimáticos y multivariados de Huglin, BEDD y Branäs, Bernon y Levadoux.

1.1. *Motivación*

La motivación detrás de este Trabajo de Fin de Máster se deriva de la creciente preocupación sobre los eventos climáticos extremos que está experimentando la Península Ibérica en los últimos años, y en particular, su impacto en los cultivos. En este sentido, hay que destacar que la economía de nuestra región depende en buena medida del sector agrícola, y en particular del vitivinícola, lo que la coloca en una posición especialmente vulnerable ante el cambio climático: el aumento de las temperaturas, los cambios en los patrones de precipitación y la aparición de fenómenos extremos pueden ejercer un impacto significativo en la productividad.

La necesidad de comprender y mitigar los efectos del cambio climático en esta región ha impulsado el estudio de los índices agroclimáticos, herramientas fundamentales para evaluar y comprender cómo estas variaciones climáticas afectan a la agricultura y a la biodiversidad en la Península Ibérica. No obstante, para utilizar estos índices de manera efectiva, es esencial abordar los sesgos que a menudo están presentes en los modelos climáticos regionales utilizados para simular las variables climáticas. Estos sesgos pueden distorsionar la representación de las condiciones climáticas simuladas y, por lo tanto, afectar las evaluaciones de riesgo y las estrategias de adaptación.

En este contexto, la motivación principal de este trabajo es contribuir al conocimiento y la capacidad de respuesta frente a los impactos del cambio climático en el sector vitivinícola en la Península Ibérica. En concreto, al alcanzar los objetivos propuestos (ver Sección 1.2), se pretende mejorar la precisión y la confiabilidad de las proyecciones futuras dadas por los modelos climáticos, lo que a su vez facilitará una evaluación más precisa de cómo las condiciones climáticas futuras pueden afectar las condiciones para el cultivo de la vid en la región. Además, se pretende también proporcionar herramientas de software para la comunidad de investigadores del clima que faciliten el estudio de los impactos del clima en la agricultura.

1.2. *Objetivos*

Basándose en las consideraciones previas, este trabajo se centra en el estudio de índices agroclimáticos multivariados, una aproximación crucial para la evaluación del impacto del cambio climático en los sistemas naturales y agrícolas que abarcan la región de la Península Ibérica. Para ello, se aplican diferentes técnicas estadísticas para la corrección de los sesgos sistemáticos de modelos regionales del clima. Se plantean tres objetivos fundamentales:

- Realizar un estudio detallado del cálculo de índices agroclimáticos, basado en un conjunto representativo de RCMs, incluyendo la aplicación y evaluación de diversas técnicas de corrección de sesgos, tanto univariadas como multivariadas. Este análisis se sustenta en las simulaciones proporcionadas por modelos climáticos regionales (RCMs), que requieren correcciones de sesgo para su utilización en el cálculo de índices agroclimáticos.
- Desarrollar *software* que permita la automatización en el cálculo de índices agroclimáticos y aplicación de técnicas de corrección de sesgo multivariantes. Para ello se plantea crear dos *wrappers* en R (R Core Team, 2021) e integrarlos en `climate4R` (Iturbide et al., 2019a), un entorno de trabajo de creciente utilización desarrollado por el Grupo de Meteorología de la Universidad de Cantabria. `climate4R` busca aglutinar bajo un marco común los diferentes procedimientos que se requieren típicamente en estudios del clima, desde el acceso a los datos hasta su procesamiento y visualización, todo ello pensando en las necesidades de un amplio abanico de usuarios. El primer *wrappers* implica la integración en `climate4R` de la librería `agroclim` (Serrano-Notivoli et al., 2020), utilizada para el cálculo de índices agroclimáticos. El segundo consiste en permitir que la librería `downscaleR` de `climate4R` pueda ejecutar los métodos de corrección de sesgo multivariantes definidos en el paquete `MBC` (Cannon, 2018). La creación de estos *wrappers* tiene como objetivo principal lograr una integración fluida y eficiente entre `climate4R` y las dos librerías mencionadas.

1.3. Estructura

La presente memoria se estructura de manera secuencial, con cada capítulo apoyándose sobre el anterior para proporcionar una visión completa y coherente de la investigación realizada. La disposición de los capítulos es la siguiente:

El Capítulo 2 presenta una descripción detallada de los datos observacionales que se han seleccionado dentro del contexto geográfico de la Península Ibérica. Además, se proporciona una visión de los modelos climáticos que se han empleado, detallando sus características claves y la razón detrás de su elección para abordar con éxito los desafíos y metas que definen el trabajo. La última sección de capítulo se dedica a describir los índices agroclimáticos seleccionados, su relevancia y cómo se calculan. Este capítulo, en su totalidad, sienta las bases necesarias para comprender la estructura y la calidad de los datos subyacentes en nuestra investigación.

A continuación, el Capítulo 3 establece el marco teórico y metodológico seguido. En concreto, se introducen las técnicas de corrección de sesgo utilizadas, tanto las univariadas como las multivariadas. Además, se presenta también el marco de validación cruzada que se ha empleado para evaluar la utilidad de las técnicas de corrección de sesgo y garantizar la robustez de los resultados.

Los resultados de este estudio se presentan y discuten exhaustivamente en el Capítulo 4. Este capítulo se centra en el análisis y la interpretación de los índices agroclimáticos obtenidos, que constituyen el resultado principal y fundamental de esta investigación. A lo largo de este trabajo, se ha llevado a cabo una evaluación integral de diversos métodos de corrección de sesgos, se han analizado las proyecciones de cambio climático para estos índices y se han identificado las limitaciones y desafíos asociados con dichas proyecciones.

Finalmente, en el Capítulo 5 se exponen las conclusiones principales y los resultados más significativos. Además, se señala las posibles direcciones para futuras investigaciones y se proporciona toda la información necesaria para poder reproducir el trabajo, contribuyendo a la transparencia y continuidad de la investigación.

CAPÍTULO 2

Datos

2.1. *Región de estudio: La Península Ibérica*

La Península Ibérica se caracteriza por su posición geográfica en la zona de transición entre las regiones extratropicales y subtropicales y está influenciada principalmente por los climas atlántico y mediterráneo (Lorenzo et al., 2013; Casanueva et al., 2020). En el norte, prevalecen las temperaturas templadas con precipitaciones regulares a lo largo del año, mientras que en el sureste se encuentran áreas secas y semiáridas en las que las precipitaciones anuales son escasas, lo que plantea serios desafíos distintos para el sector agrícola. Esta amplia gama de condiciones climáticas y la riqueza de cultivos presentes hacen de la Península Ibérica un lugar idóneo para el uso de índices agroclimáticos que permitan comprender cómo los sucesos climáticos extremos pueden influir en la productividad agrícola y en la capacidad de adaptación del sector a este tipo de desafíos.

Esta alta variabilidad climática, junto con una destacable complejidad orográfica, hacen de la Península Ibérica un escenario ideal para evaluar el rendimiento de distintas técnicas estadísticas para la corrección de sesgos en modelos climáticos, lo que resulta crucial en un contexto global donde la comprensión y predicción de eventos climáticos extremos se vuelven cada vez más relevantes para la toma de decisiones y la mitigación de riesgos.

2.2. Observaciones

La disponibilidad de datos climáticos precisos y detallados desempeña un papel fundamental en la comprensión de los fenómenos climáticos, desde escalas globales hasta regionales (Sun et al., 2018). En Europa, el primer conjunto de datos de observaciones en rejilla (originalmente con una resolución espacial de 25×25 km) a nivel continental, E-OBS, se generó como parte del proyecto europeo ENSEMBLES, perteneciente al 7^o Programa Marco de Investigación. Este conjunto de datos incluía información detallada acerca de la precipitación diaria, así como las temperaturas máxima, mínima y media, además de datos sobre la presión a nivel del mar (Cornes et al., 2018).

En el contexto de la Península Ibérica, hubo dos iniciativas pioneras que se llevaron a cabo para la creación de conjuntos de datos de precipitación en rejillas de alta resolución. Estas iniciativas, lideradas por Herrera et al. (2012) en el caso de España, y por Belo-Pereira et al. (2011) en el de Portugal, aplicaron una metodología similar a la utilizada en la elaboración de E-OBS. En años posteriores, Herrera et al. (2019) desarrollaron Iberia01, la base de datos observacionales que se utiliza en este trabajo, que cubre toda la Península Ibérica con una resolución espacial de 0.1° (aproximadamente 10×10 km). Esta rejilla es el resultado de aplicar técnicas de interpolación (*thin plate splines* sobre los valores mensuales considerando la orografía como covariable y *ordinary kriging* para interpolar las anomalías diarias) sobre una red observaciones de calidad de 3487 y 276 estaciones de precipitación y temperatura, respectivamente, de la Agencia Estatal de Meteorología (AEMET), el Instituto Portugués del Mar y la Atmósfera (IPMA) y la Agencia Portuguesa de Medio Ambiente (APA). Fue el primer conjunto de datos climáticos a alta resolución espacial, regular, sin discontinuidades fronterizas, que ofreció valores diarios de precipitación acumulada y temperatura (máxima, mínima y media) para la Península Ibérica, abarcando el periodo comprendido entre 1971 y 2015. Estas características resultan fundamentales para llevar a cabo una evaluación precisa de las simulaciones de modelos regionales del clima.

En el presente trabajo, hemos optado por utilizar el periodo 1986-2005 como el rango temporal de interés. Esta selección se basa en la relevancia y disponibilidad de datos de los modelos regionales, buscando un equilibrio entre la disponibilidad de datos y la inclusión de un periodo lo suficientemente representativo en términos de eventos climáticos y tendencias.

2.3. Modelos Regionales del Clima

En la evolución de la investigación del clima regional en Europa, la iniciativa EURO-CORDEX ha asumido un papel fundamental como continuación y ampliación de los esfuerzos iniciados por el proyecto europeo ENSEMBLES. EURO-CORDEX es la rama europea de la iniciativa CORDEX (*Coordinated Regional Downscaling Experiment*), un programa respaldado por el *World Climate Research Program* (WCRP) con el propósito de generar proyecciones de cambio climático a nivel regional en todo el mundo (Ramos-Calzado et al., 2017). Dentro del proyecto CORDEX, EURO-CORDEX desempeña la función de generar simulaciones de cambio climático en Europa mediante la combinación de múltiples RCMs anidados a diversos GCMs de CMIP5¹ (*Coupled Model Intercomparison Project Phase 5*, Taylor et al., 2011). Los RCMs incluidos en EURO-CORDEX presentan dos resoluciones espaciales diferentes, la general de CORDEX de 0.44° (EUR-44, aproximadamente 50 km) y, adicionalmente, una resolución más fina, de 0.11° (EUR-11, aproximadamente 12 km).

En este estudio se han utilizado simulaciones de RCMs anidadas a dos GCMs de CMIP5: MPI y HadGEM. La elección de estos dos modelos se basa en las notables diferencias encontradas en la señal de cambio climático de la temperatura media en Europa proyectada para finales del siglo XXI según las simulaciones de EURO-CORDEX (Figura 2.1). Como se observa, las simulaciones anidadas al modelo global MPI tienden a simular un clima más frío, mientras que las anidadas a HadGEM muestran proyecciones más cálidas. Esta selección de GCMs permite una evaluación más completa al cubrir un mayor rango de incertidumbre en lo relativo a posibles condiciones futuras.

Los RCMs que se anidan a los modelos globales MPI y HadGEM y han sido utilizados en este trabajo son RCA y CCLM, los cuales fueron desarrollados por el SMHI (*Swedish Meteorological and Hydrological Institute*) y CLMcom (*Climate Limited-area Modelling Community*), respectivamente. Las simulaciones históricas de ambos RCMs abarcaron el periodo de 1986 a 2005. Por otro lado, las simulaciones futuras de escenarios climáticos de ambos modelos regionales se llevaron a cabo desde el año 2006 hasta el 2099. En el presente estudio se consideró el escenario de altas emisiones de gases de efecto invernadero (RCP 8.5), que proporciona un marco para explorar diferentes escenarios climáticos futuros y sirve de base para

¹El experimento de intercomparación de modelos acoplados (CMIP) representa la iniciativa más ambiciosa en estudios de modelización climática a nivel mundial. CMIP5 corresponde a la quinta fase/versión de CMIP

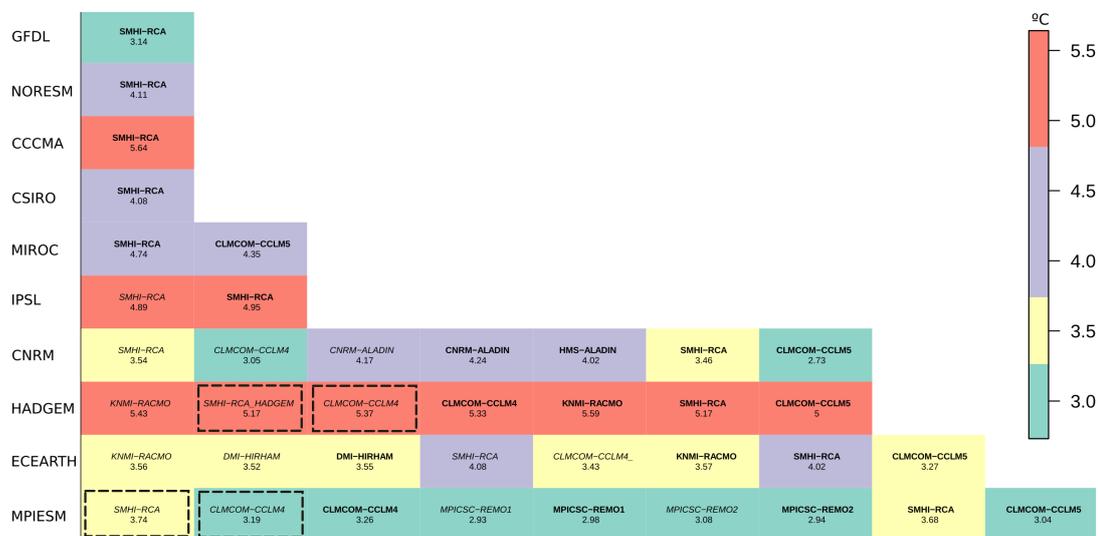


Figura 2.1: Señal de cambio climático en la temperatura media de verano (junio, julio, agosto), promediada en toda Europa durante el período de 2070-2099 en comparación con 1981-2010. Los colores representan la distribución en cuartiles de las proyecciones. Cada fila corresponde a un GCM, y en cada celda se representa una simulación de RCM. Las simulaciones de EUR-44 se resaltan en negrita, mientras que las EUR-11 se indican en cursiva. Las cuatro simulaciones utilizadas en el presente estudio aparecen en rectángulos de línea discontinua. Fuente: García-León et al. (2021)

el estudio de los impactos asociados. En concreto, el RCP8.5 describe un escenario extremo de emisiones en el que se seguirían liberando a la atmósfera gases de efecto invernadero sin restricción alguna hasta 2100, alcanzando para ese momento una presión radiativa promedio de 8.5 W/m^2 .

Se utilizaron los datos de temperatura máxima, mínima y media diaria, así como de precipitación acumulada diaria de los modelos regionales mencionados. Estos datos del proyecto CORDEX (véase la Tabla 2.1) se encuentran disponibles a través de una red internacional de nodos, ESGF (Earth System Grid Federation), y el Grupo de Meteorología de Santander dispone del primer nodo español que facilita su acceso. Concretamente, se accedió a estos datos utilizando `climate4R` y su paquete `loader`, diseñado para la lectura remota de datos proporcionados por el Servicio Climático de Datos de Santander (<https://scds.es/en>).

Es importante destacar que las cuatro simulaciones consideradas (Tabla 2.1) presentan un calendario de 365 días, por lo que para armonizarlos con las observaciones de Iberia01, se incluyeron valores faltantes (NAs) a lo largo de los años bisiestos. Además, todos los análisis se realizaron sobre la rejilla regular de las observaciones

($0.1^\circ \times 0.1^\circ$). Por esta razón, todas las simulaciones de los RCMs, que originalmente tenían una resolución espacial de 0.11° (EUR-11) en una rejilla regular rotada, fueron interpoladas usando el método del vecino más cercano para que se adecuaran a la rejilla de Iberia01.

EUR-11			
GCM	RCM	Versión (del RCM)	Institución que corre el RCM
MPI	RCA	SMHI-RCA4_v1a	SMHI
	CCLM	CLMcom-CCLM4-8-17_v1	CLMcom
HadGEM	RCA	SMHI-RCA4_v1	SMHI
	CCLM	CLMcom-CCLM4-8-17_v1	CLMcom

Tabla 2.1: Simulaciones de Modelos Climáticos Regionales (RCMs) de EURO-CORDEX que se han utilizado sobre el dominio EUR-11, caracterizado por una resolución espacial aproximada de 12 km. De todos ellos se han considerado simulaciones históricas (periodo 1986-2005) y de cambio climático bajo el escenario de altas emisiones de gases de efecto invernadero (RCP 8.5).

2.4. Índices agroclimáticos

Los índices agroclimáticos son de gran utilidad para la gestión agrícola y en la toma de decisiones relacionadas con la agricultura puesto que permiten cuantificar y comunicar de forma sencilla los riesgos climáticos que pueden afectar a los cultivos. Estos índices se basan en umbrales críticos que son vitales para el desarrollo de las plantas. Algunos de estos índices pueden adaptarse a ciertos cultivos de manera sencilla; por ejemplo teniendo en cuenta las fechas clave en el ciclo de vida de la planta en cuestión, desde la siembra hasta la madurez (Serrano-Notivoli et al., 2020). En el caso del ciclo fenológico de la vid, podemos identificar varias etapas claramente definidas. La primera, conocida como período inactivo o latente, abarca desde octubre hasta abril, durante la temporada de invierno (en la Península Ibérica). Durante este período, la vid permanece en gran medida indiferente a los cambios de temperatura, excepto en casos extremos de heladas intensas. Le sigue la etapa de brotación, generalmente comprendida entre abril y junio, marcando el inicio del crecimiento activo de la vid. En esta fase, se observan los primeros signos de brotes y crecimiento vegetativo. A continuación, llega la etapa de floración, que se extiende desde junio hasta mediados de agosto, durante la cual comienzan a aparecer los

Código	Nombre	VARIABLES	Periodo	Referencia
HI	Índice de Huglin	Temperatura media (T_{mean}) Temperatura máxima (T_{max})	1 de abril hasta el 30 de septiembre	Huglin (1978)
BEDD	Índice de grados-días biológicamente efectivos	Temperatura mínima (T_{min}) Temperatura máxima (T_{max})	1 de abril hasta el 31 de octubre	Honorio Guisado (2016)
BBLI	Índice hidrotérmico de Branas, Bernon y Levadoux	Temperatura media (T_{mean}) Precipitación (P)	1 de abril hasta el 31 de agosto	Branas et al. (1946)

Tabla 2.3: Descripción detallada de los índices agroclimáticos multivariados empleados en este estudio. Las fechas de aplicación de estos índices siguen el ciclo agrícola del hemisferio norte.

primeros frutos, aún inmaduros, de la planta. Finalmente, la etapa de maduración ocurre desde mediados de agosto hasta finales de septiembre. En esta fase, los frutos comienzan a agrandarse, alcanzando un grado de azúcar mayor y experimentan un cambio de color. Las uvas blancas pasan de verde a amarillo, mientras que las uvas negras adquieren tonalidades púrpuras (Jones et al., 2005; Lorenzo et al., 2013).

En este estudio se han empleado tres índices diferentes, todos ellos de naturaleza multivariante (véase la Tabla 2.3): el índice de Huglin (HI), el índice de días-grados biológicamente efectivos (BEDD) y el índice hidrotérmico de Branas, Bernon y Levadoux (BBLI). Una vez que se interpolan las variables climáticas simuladas por los RCMs para que coincidieran con la rejilla de observaciones y se corrigieron los sesgos presentes, se procedió a llevar a cabo el análisis de los tres índices mencionados.

El índice de Huglin (HI, Ecuación 2.1) se utiliza para determinar el potencial de una región para el cultivo de ciertas variedades de uva en términos de la temperatura. Este índice se calcula como el producto del coeficiente K (Ecuación 2.2) y el promedio de las temperaturas diarias medias y máximas en relación con la temperatura base de 10°C . Este cálculo tiene en cuenta todos los días comprendidos en la temporada de desarrollo y maduración de la vid, que generalmente abarca desde el 1 de abril hasta el 30 de septiembre en el hemisferio norte. Para regiones vitivinícolas localizadas en el hemisferio sur, el periodo que ocupa este índice va del 1 de octubre al 30 de marzo (Huglin, 1978).

$$HI = \sum_{1 \text{ de abril}}^{30 \text{ de septiembre}} K \cdot \text{máx} \left[\frac{T_{mean} + T_{max}}{2}, 10 \right] \quad (2.1)$$

El coeficiente K es un factor de corrección que tiene en cuenta el período promedio de luz diurna de la región en la que se está evaluando el índice. El coeficiente K se basa en la latitud (ϕ) de la zona vitivinícola y se emplea para incluir el aumento de la luz del día en latitudes más elevadas (Tonietto and Carbonneau, 2004):

$$K = \begin{cases} \text{Interpolación lineal} \\ \text{entre 1.02 y 1.06} & \text{si } 40 \leq \phi \leq 50 \\ 1 & \text{si } \phi < 40 \text{ ó } \phi > 50 \end{cases} \quad (2.2)$$

El índice de grados-días biológicamente efectivos o BEDD (*Biologically Effective Degree-Days*, Ecuación 2.3), al igual que el índice de Huglin, ofrece información acerca de la acumulación de calor. Se basa en el cálculo de los grados-día de crecimiento, que se obtienen sumando las temperaturas diarias medias (calculadas como la media de la mínima y la máxima) desde 10°C hasta 19°C. El índice BEDD se calcula durante la temporada de crecimiento de la vid, que por lo general abarca desde el 1 de abril hasta el 31 de octubre en el hemisferio norte, mientras que en el hemisferio sur se extiende desde el 1 de octubre hasta el 30 de abril (Honorio Guisado, 2016).

$$BEDD = \sum_{1 \text{ de abril}}^{31 \text{ de octubre}} DTR_{adj} \cdot \text{mín} \left[K \cdot \text{máx} \left[\frac{T_{max} + T_{min}}{2}, 10 \right], 19 \right] \quad (2.3)$$

donde K es el factor de corrección relacionado con la duración del día (Ecuación 2.2), y DTR_{adj} es el término de ajuste que tiene en cuenta el rango de temperatura diurna DTR (diferencia de temperatura mínima y máxima diaria):

$$DTR_{adj} = \begin{cases} 0,25 (DTR - 13) & \text{si } DTR > 13 \\ 0 & \text{si } 10 \leq DTR \leq 13 \\ 0,25 (DTR - 10) & \text{si } DTR < 10 \end{cases} \quad (2.4)$$

El índice hidrotérmico de Branas, Bernon y Levadoux (BBLI, Ecuación 2.5) se utiliza para examinar el impacto de la temperatura y la precipitación en el rendimiento de la uva y la calidad del vino, además de evaluar las posibilidades hidrotérmicas del suelo de cultivo (Branas et al., 1946). Este índice se calcula sumando mensualmente los productos de la temperatura media (en grados centígrados) y las precipitaciones (en milímetros) durante el período que abarca desde abril hasta agosto.

$$\text{BBLI} = \sum_{\text{1 de abril}}^{\text{31 de agosto}} T_{\text{mean, monthly}} \cdot P_{\text{monthly}} \quad (2.5)$$

El valor de este índice es un indicador relacionado con la probabilidad de que las vides sean afectadas por el moho, una de las enfermedades más comunes y devastadoras en la viticultura (Lorenzo et al., 2013; Cardell et al., 2019).

Para el cálculo de todos los índices mencionados, se ha empleado el paquete de R `agroclimR` desarrollado por Serrano-Notivoli et al. (2020). Como parte del presente trabajo, se ha incorporado esta librería dentro del framework de `climate4R` para facilitar y automatizar el análisis climático y la evaluación de impactos en el sector agrario. Esta integración ha permitido aprovechar las herramientas y funcionalidades de `climate4R` para un análisis más completo y eficiente de los datos climáticos en relación con la viticultura.

CAPÍTULO 3

Métodos

La primera sección de este capítulo se dedica a una detallada descripción de los algoritmos de corrección de sesgo que han sido aplicados en esta investigación. Se abordan tanto los enfoques univariados, que se centran en mitigar el sesgo en variables individuales, como los enfoques multivariados, que consideran las relaciones de dependencia entre múltiples variables. Esta parte del capítulo tiene como objetivo principal facilitar la comprensión de su funcionamiento y su aplicabilidad en el contexto de este estudio.

En la segunda parte del capítulo se presenta el marco de validación cruzada que ha sido implementado para evaluar la eficacia de las técnicas de corrección de sesgo aplicadas. Este proceso de validación cruzada desempeña un papel crucial, ya que garantiza la robustez de los resultados obtenidos.

3.1. Métodos de Corrección de Sesgo (BC)

A pesar de su valiosa contribución para entender el cambio climático a escala regional, las simulaciones provenientes de RCMs suelen presentar sesgos sistemáticos (diferencias con respecto a las observaciones de referencia), lo cual impide su uso directo en estudios de impactos.

Para limitar esta problemática se han desarrollado diversos métodos de corrección/ajuste de sesgo durante las últimas décadas. Estos métodos pueden ser tanto univariados (es decir, que actúan sobre una única variable climática) como multivariantes (que tratan de preservar la estructura de dependencia entre múltiples varia-

bles climáticas). Sin embargo, todos estos métodos tiene una característica común: ajustan los datos simulados por el RCM para que su distribución estadística sea coherente con la de las observaciones utilizadas como referencia, lo que les permite alinear de manera más precisa las predicciones del modelo con la realidad observada. Hay que tener en cuenta que, al ajustar los valores del modelo (sesgado) hacia las observaciones correspondientes, puede afectar indirectamente las tendencias y la señal de cambio climático resultante (Maraun, 2013; Casanueva et al., 2020).

La fundamentación subyacente de estas técnicas reside en el proceso de "entrenamiento/calibración" que se llevan a cabo, que habitualmente se realiza en un período histórico reciente con registros observacionales disponibles, como, por ejemplo, en el intervalo de 1986-2005. Posteriormente, los factores de corrección determinados se aplican para mejorar las simulaciones del RCM en un período futuro, como el comprendido entre 2080-2099 (Teutschbein and Seibert, 2012).

Los métodos de corrección de sesgo asumen que los sesgos sistemáticos identificados en el modelo en su periodo de entrenamiento/calibración son estacionarios, es decir, que también aparecerán en otro periodo de tiempo. Esto significa que el algoritmo de corrección y sus parámetros, diseñados para abordar sesgos en las condiciones climáticas del periodo de entrenamiento/calibración (normalmente el periodo histórico), también se aplican para condiciones climáticas futuras o en otro periodo de tiempo. A pesar de su utilidad, esta asunción plantea una posible limitación en su aplicación en condiciones de cambio climático.

A continuación se describen los tres métodos univariados y los dos multivariados que se han utilizado en este trabajo para la corrección de sesgos en los RCMs (véase la Tabla 3.2). Es importante mencionar que, en un principio, se llevó a cabo la calibración de estos métodos utilizando la serie temporal anual completa. Sin embargo, se observaron sesgos significativos después de aplicar los métodos de corrección de sesgo. Por este motivo, se optó por realizar la calibración y aplicación de los métodos de forma mensual. Esta estrategia considera no solo que los sesgos varían en función de la magnitud de la variable (lo que implica correcciones diferentes para distintos cuantiles en todos los métodos, excepto en PQM) sino también que dependen de la época del año.

3.1.1. Métodos Univariados

Todos estos métodos actúan sobre cada variable de forma independiente, que se combinan posteriormente para el cálculo de los índices agroclimáticos.

- **Mapeo cuantil-cuantil empírico (EQM):** Método empírico que se basa en

Código	Nombre	P/E	U/M	Paquete	Referencia
EQM	Mapeo cuantil empírico	E	U	climate4R	Déqué (2007)
PQM	Mapeo cuantil paramétrico	P	U	climate4R	Piani et al. (2010)
QDM	Mapeo cuantil delta	E	U	climate4R	Cannon et al. (2015)
MBCr	Corrección de sesgo multivariado (correlación de rango de Spearman)	E	M	MBC	Cannon (2016)
MBCn	Corrección de sesgo multivariado (N-pdf)	E	M	MBC	Cannon (2018)

Tabla 3.2: Métodos de corrección de sesgo utilizados en el estudio, indicando el código, el nombre, el carácter paramétrico/empírico (P/E), el carácter univariado/multivariado (U/M), el paquete de software y la referencia del método.

el mapeo de las funciones de distribución acumulativa (CDF) entre los datos simulados y los datos observados. Se emplea una función de transferencia que se ajusta durante el periodo de control para mapear los cuantiles de la función de distribución acumulativa empírica de los datos del modelo sobre la distribución observada. Este enfoque garantiza que la función de distribución acumulativa de los datos simulados corregidos (después de aplicar el mapeo de cuantiles) sea idéntica a las funciones de distribución acumulativa de los valores observados si los datos de entrenamiento/calibración son los mismos que se utilizan para aplicar la corrección. (Beyer et al., 2020). El método EQM empleado en este estudio se lleva a cabo utilizando la librería de `climate4R` (Iturbide et al., 2019b), utilizando la configuración que ajusta todos los cuantiles. Además, emplea extrapolación constante, lo que implica que en situaciones donde los valores a corregir se ubiquen fuera del intervalo de calibración, se aplicarán la primera y última corrección para valores por debajo y por encima del rango de calibración, respectivamente. Asimismo, para el caso de precipitación ajusta la frecuencia de días húmedos (días con precipitación ≥ 1) utilizando un umbral de precipitación que hace coincidir las frecuencias observada y simulada de días húmedos. En el caso de que haya más días secos en el modelo que en la observación se aplica una adaptación de frecuencias, que consiste en generar valores aleatorios de precipitación muestreados de una distribución Gamma.

- **Mapeo cuantil paramétrico (PQM):** En contraste con el mapeo cuantil

empírico, el mapeo cuantil paramétrico utiliza una función de distribución acumulativa teórica/paramétrica para ajustar los datos simulados del modelo a los datos observados. Esto implica realizar una suposición acerca la forma de la distribución de probabilidad de los datos, lo que puede ser útil en situaciones en las que se conoce que los datos siguen una distribución específica, como la distribución normal en el caso de la temperatura o la distribución Gamma en el caso de la precipitación. A diferencia del enfoque EQM, no se incluyen ajustes específicos para tendencias, y la corrección de valores extremos se lleva a cabo mediante la distribución paramétrica ajustada. La implementación utilizada es la de la librería `climate4R`, en la cual se aplican las mismas correcciones en cuanto a la frecuencia de días húmedos y secos que para el EQM.

- **Mapeo cuantil delta (QDM):** Método empírico que se divide en tres etapas clave para ajustar proyecciones de modelos climáticos. El objetivo principal del QDM es preservar los cambios relativos proyectados por el modelo en todos los cuantiles, al mismo tiempo que corrige cualquier sesgo sistemático presente en estos cuantiles en comparación con los valores observados. Al lograr esto, el método QDM consigue preservar la tendencia presente en todos los cuantiles. Las tres etapas del QDM son las siguientes:
 - (I) Eliminar las tendencias presentes en las salidas del modelo para el futuro. Esto se hace por cuantil, lo que significa que se ajustan las tendencias de cada nivel de probabilidad por separado.
 - (II) Aplicar el mapeo cuantil a todos los cuantiles empíricos a los que se han eliminado la tendencia en la etapa anterior. El mapeo cuantil se encarga de ajustar estos cuantiles de manera que se alineen de manera más precisa con los valores observados históricos.
 - (III) Se vuelven a incorporar las tendencias proyectadas por el modelo a los cuantiles corregidos por sesgo en la etapa anterior. Esto asegura que los cambios relativos previstos por el modelo se mantengan en todas las partes de la distribución de probabilidad.

El método QDM empleado en este estudio se lleva a cabo utilizando la librería de `climate4R`, que cuenta con la implementación de Cannon et al. (2015).

3.1.2. *Métodos Multivariados*

Los sistemas climáticos se caracterizan por interacciones complejas entre múltiples variables, como la temperatura, la humedad, la presión atmosférica y otras. Los métodos de corrección de sesgo univariados (Subsección 3.1.1), aunque son efectivos en la corrección de sesgos individuales, a menudo pasan por alto las estructuras de dependencia que existen entre diferentes variables climáticas. A pesar de eso, son capaces de preservar o no distorsionar las relaciones entre las variables simuladas por el modelo (Wilcke et al., 2013; Casanueva et al., 2019) pero no son capaces de ajustar sesgos en esas relaciones. Ignorar esta dependencia intervariable puede tener un impacto significativo en el posterior cálculo de índices (por ejemplo, agroclimáticos) que se basan en múltiples variables climáticas (Cannon, 2016).

En esta sección se describe el algoritmo MBC, que combina el ajuste de distribuciones marginales a través del mapeo de cuantiles QDM (incluyendo la capacidad de preservar los cambios proyectados en cuantiles simulados) con el ajuste de la estructura de dependencia a través del reescalado multivariante. Se presentan dos variantes: MBCr, que preserva la estructura de correlación de Spearman entre dos o más variables y MBCn, que corrige todas las características estadísticas de la distribución multivariante continua. Es importante destacar que, debido a que estos métodos requieren muchos más recursos computacionales (almacenamiento en memoria de varias variables, procesos iterativos, ...) se aplicaron a parejas de variables en lugar de corregir las cuatro variables a la vez. Algunas pruebas mostraron resultados similares al introducir 2 o 4 variables, de tal forma que se puede asumir que incluir menos variables no distorsiona los resultados. Ambos métodos se han llevado a cabo utilizando la librería MBC de Cannon (2016, 2018), que como parte del presente trabajo, se ha integrado en `climate4R`.

- **Corrección de sesgo multivariado basada en la correlación de rango de Spearman (MBCr):** Es un algoritmo de corrección de sesgo multivariado que combina mapeo cuantil y el reescalado lineal multivariado. Se caracteriza por considerar que la estructura de dependencia entre dos o más variables se basa en la correlación de rango de Spearman. Para garantizar que las proyecciones ajustadas por MBCr sean coherentes con las observaciones, se utiliza el mapeo cuantil delta univariado (QDM) para ajustar las distribuciones marginales de cada variable individual. Es un procedimiento iterativo que sucesivamente emplea un reescalado lineal multivariado para adecuar la estructura de dependencia según la correlación de Spearman, con el objetivo de que esta se asemeje de manera más precisa a los valores observados. En

esencia, el algoritmo MBC altera la secuencia temporal inherente al modelo climático subyacente para corregir la estructura de dependencia, pero procura preservar la secuencia temporal resultante de la salida del modelo.

- **Corrección de sesgo multivariado basado en el algoritmo N -pdft:** Es una adaptación del algoritmo N-pdft, que se utiliza en campos como visión artificial y procesamiento de imágenes (Pitie et al., 2005). El algoritmo consta de 3 pasos que se repiten iterativamente hasta que la distribución multivariante coincide con la distribución de los datos observados:

- (I) Aplicar una rotación ortogonal aleatoria a los conjuntos de datos.
- (II) Corregir las distribuciones marginales de los datos rotados mediante un mapeo cuantil delta (QDM). Al aplicar la rotación en el paso previo, se consigue crear combinaciones lineales de las variables originales, evitando así corregir cada variable por separado con el método univariado QDM.
- (III) Aplicar la rotación inversa a los datos resultantes.

MBCn es una generalización multivariada del mapeo cuantil, que permite transferir todas las características estadísticas de una distribución multivariante continua observada a la correspondiente distribución multivariante de las variables simuladas. A diferencia de otros algoritmos multivariados de corrección de sesgo, como el método MBCr. propuestos por el mismo autor, Cannon (2016), MBCn no se limita a corregir una única medida específica de la dependencia conjunta, como la correlación de rango de Spearman, ni realiza suposiciones fuertes acerca de la estacionariedad de la secuencia temporal en el modelo climático. Al igual que el algoritmo MBCr, MBCn preserva las tendencias.

3.2. Marco de Validación

3.2.1. Validación Cruzada

El proceso de validación de los métodos de corrección de sesgo se ha realizado a través de una validación cruzada de tipo k -fold con $k = 2$. En este enfoque, se divide el conjunto de datos en dos subconjuntos (cada uno de 10 años) que se alternan para ser usados como conjunto de entrenamiento del método (*train*) y como conjunto de prueba (*test*). Al ser dos conjuntos de datos independientes, este tipo de validación permite evaluar la capacidad de generalización del método y evitar el sobreajuste,

asegurando una evaluación justa del rendimiento de las técnicas de corrección de sesgo.

Con el objetivo de asegurar una evaluación más rigurosa, se han realizado dos experimentos diferentes, cada uno enfocado en un aspecto específico. En el primer experimento, se dividió el período histórico en dos *fold* de manera aleatoria, lo que permitió evaluar el rendimiento de las técnicas en segmentos temporales diferentes. Por otro lado, en el segundo experimento, la división de los conjuntos de datos se basó en la calificación de los años en fríos y cálidos, en base a las temperaturas promediadas en la Península Ibérica. Para llevar a cabo esta clasificación, se utilizó el conjunto de datos históricos de observaciones y se identificaron los años más fríos y más cálidos según su temperatura media (se calculó la temperatura media de la serie en cada punto y su promedio espacial). Esta estrategia brindó la oportunidad de evaluar cómo las técnicas de corrección de sesgo responden en condiciones climáticas más cálidas, lo que resulta esencial para comprender su robustez y su aplicabilidad en distintos contextos climáticos de calentamiento global. La idea detrás de esta aproximación consiste en validar si estas técnicas serán fiables para corregir las simulaciones futuras de los RCM, que se desarrollan bajo condiciones más cálidas, utilizando como conjunto de calibración los datos observados que, en comparación, representarán condiciones climáticas más frías.

3.2.2. Métricas de Validación

Como paso previo a examinar la habilidad de los modelos corregidos para representar los índices agroclimáticos, se evaluaron las variables climáticas corregidas, que son la base para calcular dichos índices:

$$\text{BIAS}_{abs} = \overline{Y_{sim}} - \overline{Y_{obs}} \quad (3.1)$$

$$\text{BIAS}_{rel} = 100 \cdot \frac{\overline{Y_{sim}} - \overline{Y_{obs}}}{\overline{Y_{obs}}} \quad (3.2)$$

$$r_{\text{Pearson}}(Y, Z) = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(Z_i - \bar{Z})}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (Z_i - \bar{Z})^2}} \quad (3.3)$$

$$\rho_{\text{Spearman}}(Y, Z) = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (3.4)$$

donde \bar{Y} corresponde a la media (temporal) de la variable climática, d son las diferencias entre los rangos de los valores de las dos variables climáticas, Y y Z , y n es el número de datos empleado.

El sesgo absoluto ($BIAS_{abs}$, Ecuación 3.1), se aplica en nuestro análisis a las variables de temperatura (media, máxima y mínima). Esta métrica cuantifica la diferencia entre la media de la variable climática simulada y la observada. Para la precipitación, se utiliza habitualmente el sesgo relativo ($BIAS_{rel}$, Ecuación 3.2), expresado en porcentaje. Para evaluar la relación entre pares de variables climáticas, ya sean ambas simuladas por el modelo RCM u observadas, se ha utilizado la correlación de Spearman (ρ , Ecuación 3.4) para la combinación de precipitación y temperatura media, y la correlación de Pearson (r , Ecuación 3.3) para analizar la relación entre temperaturas. Esta métrica nos ayuda a comprender si la relación de dependencia entre las variables se mantiene después de aplicar los métodos de corrección de sesgo.

Para evaluar la calidad de los índices agroclimáticos simulados hemos aplicado un conjunto de métricas de validación. Estas métricas proporcionan una comprensión detallada del rendimiento de nuestros métodos de BC en comparación con los datos observados. En este contexto, se destacan dos métricas clave: el sesgo relativo ($BIAS_{rel}$) y el sesgo relativo en los 10 años más cálidos ($BIAS_{rel}^{warm}$). A continuación, presentaremos estas métricas junto con sus respectivas ecuaciones:

$$BIAS_{rel} = 100 \cdot \frac{\overline{X_{sim}} - \overline{X_{obs}}}{\overline{X_{obs}}} \quad (3.5)$$

$$BIAS_{rel}^{warm} = 100 \cdot \frac{\overline{X_{sim}^{warm}} - \overline{X_{obs}^{warm}}}{\overline{X_{obs}^{warm}}} \quad (3.6)$$

donde \overline{X} representa el promedio temporal del índice agroclimático, el subíndice *sim* se refiere al modelo GCM-RCM, y el subíndice *obs* hace referencia a las observaciones.

Las métricas de validación aplicadas a los índices agroclimáticos comprenden tres componentes fundamentales. En primer lugar, se encuentra el sesgo relativo (Ecuación 3.5), el cual determina la diferencia porcentual entre la media del índice agroclimático calculado a partir de las variables climáticas simuladas por el modelo RCM y la observación, en relación con la observación misma. Esta métrica proporciona una medida de la tendencia general del modelo a subestimar o sobreestimar los valores del índice. En segundo lugar, se emplea el sesgo relativo en los años cálidos (Ecuación 3.6), que evalúa el rendimiento del modelo en situaciones específicamente cálidas, imitando un contexto de calentamiento global en el clima futuro.

3.3. *Aplicación a condiciones de cambio climático*

En este estudio, los métodos de corrección de sesgo también se han aplicado a un periodo futuro de 20 años bajo el escenario RCP8.5, el cual representa un nivel de calentamiento global estimado de aproximadamente 4°C. Se calcula la señal de cambio climático, utilizando la Ecuación (3.7), para cuantificar el impacto del cambio climático según las proyecciones del modelo RCP8.5 en relación con las observaciones:

$$\text{Señal de cambio climático (\%)} = 100 \cdot \frac{\overline{X_{RCP8,5}} - \overline{X_{HIST}}}{\overline{X_{HIST}}} \quad (3.7)$$

donde \overline{X} representa el promedio temporal del índice agroclimático y el subíndice *HIST* hace referencia a las observaciones.

CAPÍTULO 4

Resultados y Discusión

4.1. Indicadores agroclimáticos observados en clima presente

La Figura 4.1 presenta dos representaciones de los índices HI, BEDD y BBLI en la Península Ibérica. En la fila superior, se muestra el valor medio estos índices calculados para el período de 1986 a 2005. En la fila inferior, los índices se calcularon específicamente utilizando los 10 años más cálidos dentro de ese período. Estas representaciones ofrecen una visión detallada de cómo estos índices varían espacialmente en función de las condiciones climáticas durante el período completo y durante los años más cálidos. Cada uno de estos índices ofrece información clave sobre las condiciones climáticas que afectan directamente a la agricultura y la producción de uva en la región.

En primer lugar, se observa que tanto el HI como el BEDD muestran un patrón mayoritariamente orográfico en la Península Ibérica. Estos índices, que están estrechamente relacionados con las condiciones de temperatura (véase la Tabla 2.3), muestran valores más altos en el sur de la península y en las zonas costeras del Mediterráneo. Además, se destacan las cuencas de los ríos Guadalquivir y Ebro, donde también se observan valores notablemente elevados. Este patrón refleja la influencia de la temperatura en estas áreas, que tienden a experimentar climas más cálidos y secos en comparación con el norte y las regiones montañosas de la Península Ibérica.

Por otro lado, el índice BBLI muestra un patrón diferente (Figura 4.1). Este índice, que está influenciado por la precipitación, presenta valores significativamente más elevados en la vertiente Atlántica, en el norte de la Península Ibérica y en la

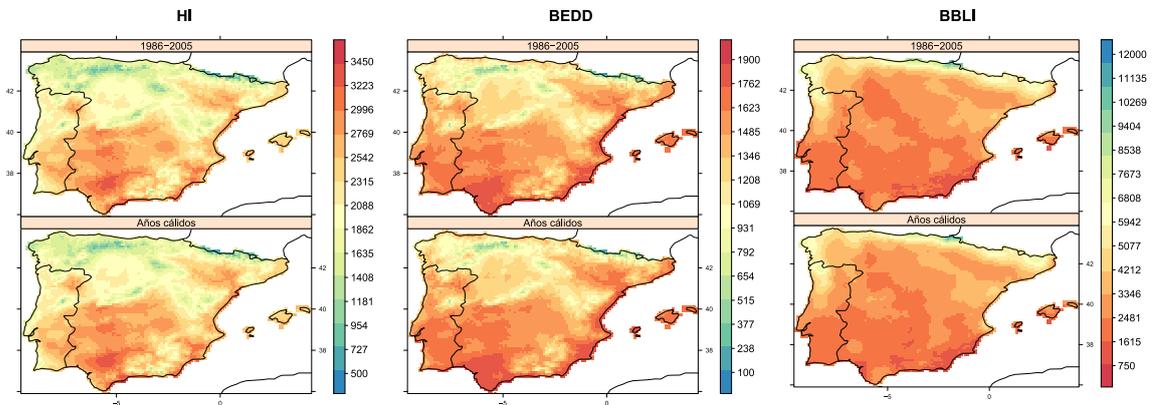


Figura 4.1: Índices agroclimáticos (en columnas) calculados con los datos observacionales de Iberia01. La primera fila muestra los valores promedio de los índices para el período de tiempo 1986-2005, mientras que la segunda fila se basa en los 10 años más cálidos según dichas observaciones.

región de los Pirineos, zonas conocidas por su mayor acumulación de lluvias a lo largo del año. Esta relación directa entre el BBLI y la cantidad de precipitación destaca en la evaluación del riesgo de desarrollo de enfermedades fúngicas, como el moho. Si bien una mayor precipitación puede ser beneficiosa para el crecimiento de los cultivos, es crucial tener en cuenta que también puede desencadenar un exceso de humedad, lo que a su vez aumenta el riesgo de patologías que afectan negativamente a la producción agrícola. En el norte de la península, especialmente la zona de las Rías Baixas, que se caracteriza por ser una importante región vitícola, se observan valores del índice BBLI fluctuando en un rango entre 6800 y 4000. Esto indica la presencia de condiciones propicias para las infestaciones de mildiu en la región, y este riesgo podría aumentar si no se implementan medidas preventivas para evitar su aparición. Es relevante resaltar que, en términos generales, la Península Ibérica tiene un promedio de alrededor de 1600 en este índice (Figura 4.1) y se establece 1500 como el límite a partir del cual pueden producirse ataques de mildiu (Aranguren González et al., 2015).

4.2. Evaluación de los métodos de corrección de sesgos

4.2.1. Validación cruzada con años aleatorios

En esta sección se explora el impacto de los métodos de corrección de sesgo en el experimento de *2-fold* con años aleatorios, en particular, cómo mejoran la

representación de las variables climáticas (temperatura mínima, máxima y media y precipitación) y el efecto en los índices agroclimáticos calculados a partir de éstas. El análisis detallado de estos efectos permite proporcionar una comprensión más profunda de cómo la corrección de sesgos en las variables climáticas se reflejan en los índices agroclimáticos, contribuyendo así a evaluar la robustez y la fiabilidad de nuestro experimento.

Efecto de los métodos BC sobre las variables climáticas

En las Figuras 4.2, 4.2, 4.2 y 4.2 se muestra, mediante boxplots, el sesgo de los RCMs para las distintas variables climáticas (temperatura media, temperatura máxima, temperatura mínima y precipitación, respectivamente) antes y después de aplicarles las técnicas de corrección/ajuste de sesgo consideradas (véase la leyenda de colores).

Antes de aplicar los métodos BC, todos los modelos presentan sesgos importantes (con respecto a las observaciones de Iberia01) en las variables climáticas. La magnitud de los errores varía en función de la variable. Por ejemplo, en el caso de las temperaturas, los sesgos previos a la corrección oscilan dentro de un rango de -3°C a 3°C . Por otro lado, en lo que respecta a la precipitación, en la Figura 4.5 se observan diferencias hasta del 100% en comparación con los valores observados.

Sin embargo, tras la aplicación de la corrección de sesgos, se observa una reducción sistemática de estos errores. Los valores de sesgo se reducen considerablemente en los modelos corregidos en comparación con sus versiones no corregidas. Esto es especialmente evidente en el caso de la temperatura media (Figura 4.2), donde el error después de la corrección prácticamente se reduce a cero. Este resultado es coherente con el hecho de que, a pesar de tratarse de una validación cruzada que utiliza datos independientes para el entrenamiento y la prueba, se están ajustando de manera efectiva todos los cuantiles en el caso de EQM, QDM, MBCr, MBCn (y en cierta medida, también la media) y, de manera explícita, la media en el caso de PQM.

En lo que respecta a la precipitación (Figura 4.5), se observa una reducción significativa de sesgos en los métodos univariantes después de la corrección. Sin embargo, este fenómeno no se reproduce de manera igual de efectiva en los métodos MBC. Los sesgos sistemáticos que persisten en los métodos MBC pueden deberse a varias razones.

Una de las posibles explicaciones radica en la naturaleza iterativa de los métodos MBC, lo que implica la necesidad de ajustar con mayor precisión el número de ite-

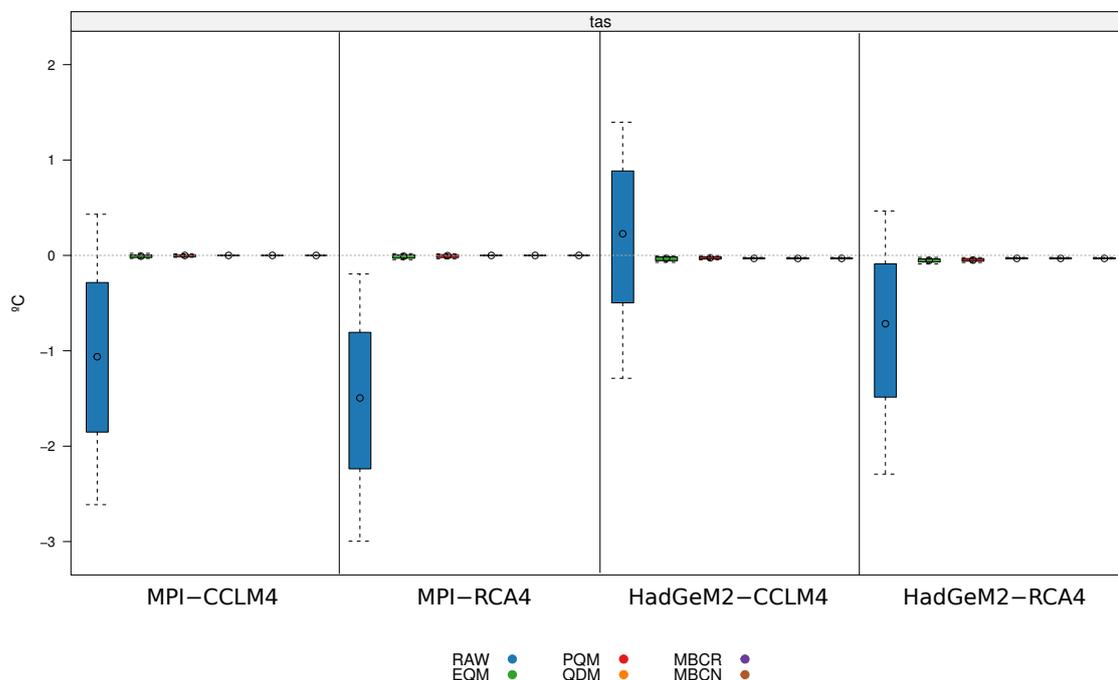


Figura 4.2: Sesgo absoluto (en $^{\circ}\text{C}$) de la temperatura media (tas) para cada par GCM-RCM (separados entre sí por líneas verticales) y método de ajuste de sesgo empleado para corregirlo. Dentro de cada par GCM-RCM, el primer boxplot azul (raw) representa el sesgo antes de aplicar ninguna corrección. El sesgo absoluto se determina tomando como referencia los datos observados (Iberia01) y centrándose en los meses de abril a octubre del periodo de tiempo 1986-2005.

raciones para lograr una corrección óptima. Como avisa el autor de ambas técnicas MBC, Cannon (2018), estos métodos requieren un elevado número de iteraciones para converger hacia la distribución objetivo lo que puede generar un costo computacional excesivo. Además, si el tamaño de la muestra es considerablemente grande en relación con la cantidad de variables), existe el riesgo de sobreajuste.

Por otro lado, es importante tener en cuenta que, en el caso de la precipitación, puede que se requieran más variables además de la temperatura para lograr un mejor ajuste.

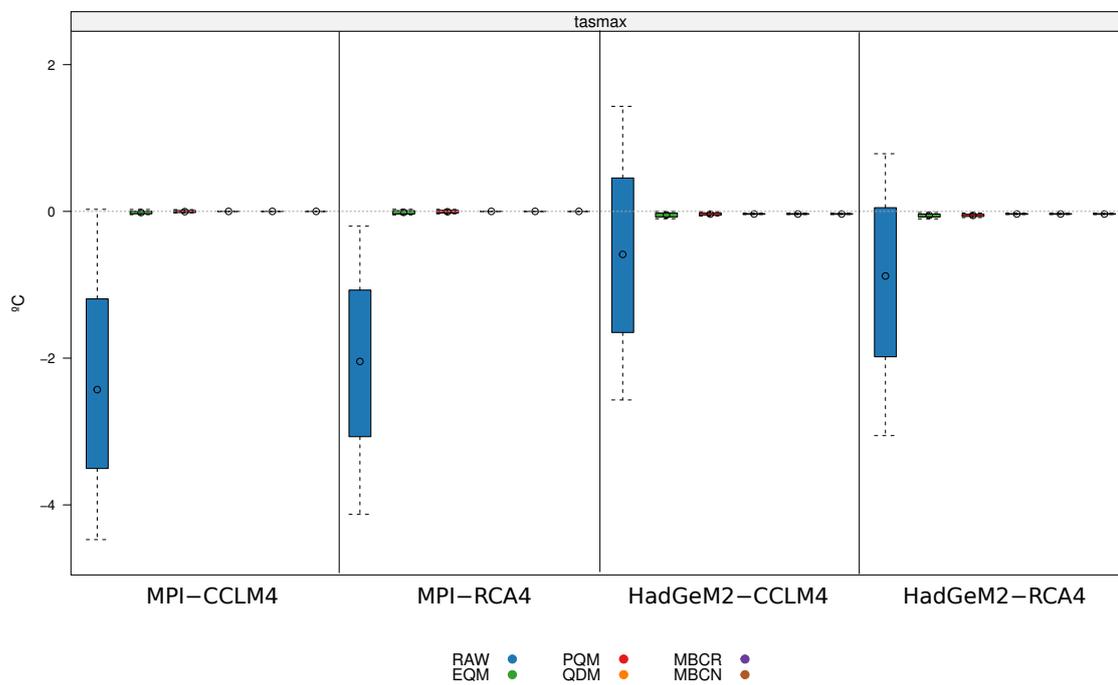


Figura 4.3: Como la Figura 4.2 pero para la temperatura máxima (*tasmax*).

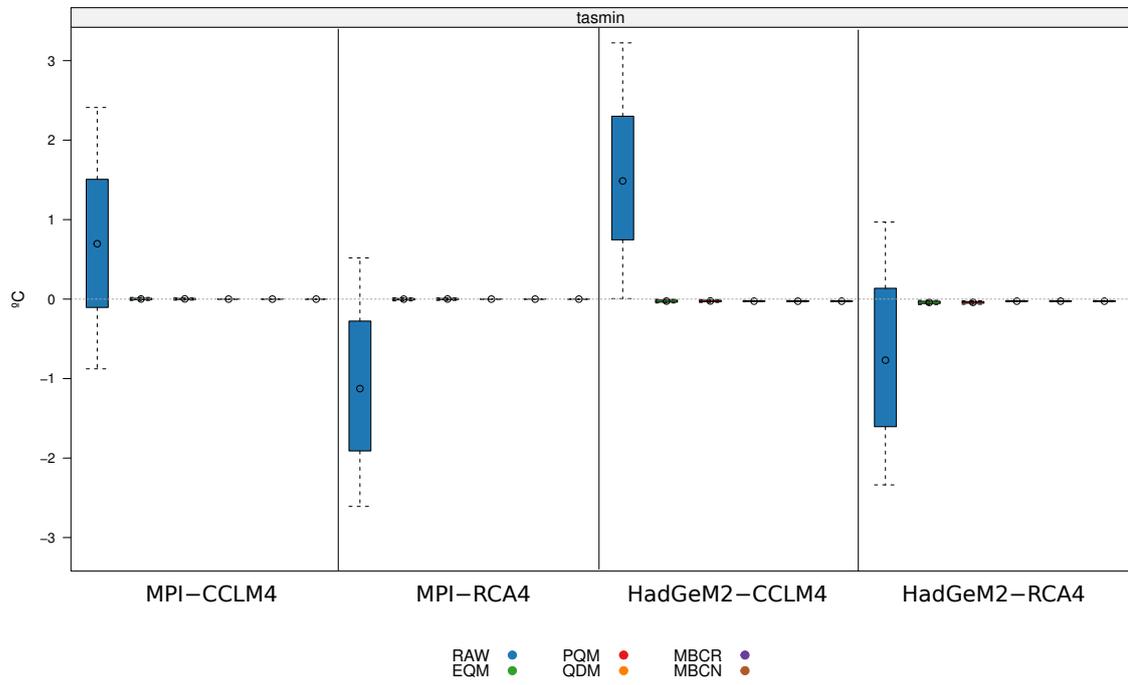


Figura 4.4: Como la Figura 4.2 pero para la temperatura mínima (*tasmin*).

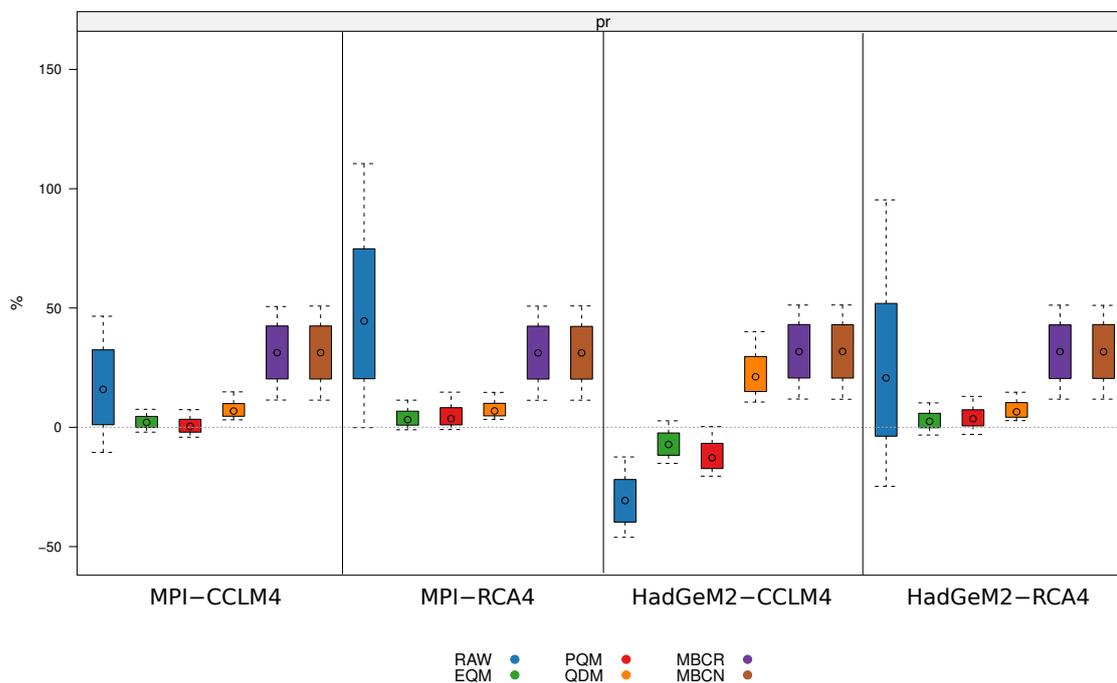


Figura 4.5: Como la Figura 4.2 pero para la precipitación. En este caso los errores se expresan en porcentaje (respecto a Iberia01).

En relación con la estructura de dependencia entre las variables climáticas, se ha estudiado las correlación entre la temperatura media y la temperatura máxima (Figura 4.6), entre la temperatura mínima y la temperatura máxima (Figura 4.7), y entre la precipitación y la temperatura media (Figura 4.8).

Los métodos MBC representan de manera efectiva la alta correlación presente en las observaciones, que suele ser de alrededor de $+0,90$ entre la temperatura media y la temperatura máxima, y de aproximadamente $+0,75$ entre la temperatura máxima y la temperatura mínima. Sin embargo, es importante destacar que la correlación entre la temperatura media y la temperatura máxima empeora en los MBC, a pesar de que estos métodos están diseñados para preservar las correlaciones observadas. Este aspecto deberá ser objeto de un análisis adicional y considerado como un requisito importante en futuras investigaciones. Por otro lado, los métodos univariantes no distorsionan estas correlaciones, lo que es un resultado positivo en términos de preservación de relaciones climáticas fundamentales.

Sin embargo, en lo que concierne al vínculo entre la temperatura media y la precipitación, se observa que la correlación en las observaciones es baja, alrededor de $-0,3$. Sorprendentemente, los métodos MBC degradan esta correlación, arrojando

do valores cercanos a cero. Este hallazgo plantea interrogantes sobre la capacidad de los métodos MBC para mantener las relaciones entre la temperatura media y la precipitación, y sugiere la necesidad de investigar en mayor profundidad este comportamiento.

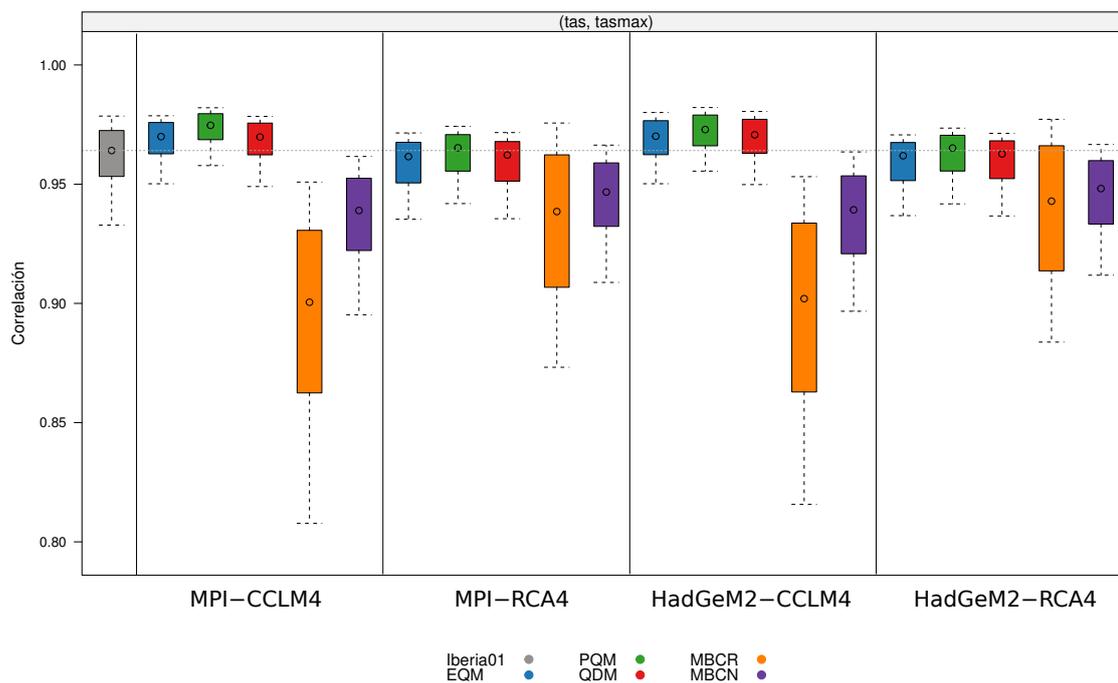


Figura 4.6: Correlación de Pearson entre la temperatura media (tas) y la temperatura máxima ($tasmax$) para cada par GCM-RCM (separados entre sí por líneas verticales) y método de corrección de sesgo empleado. El primer boxplot gris representa la correlación observada de acuerdo a Iberia01. Todos los resultados corresponden al periodo 1986-2005 centrándose en los meses de abril a octubre.

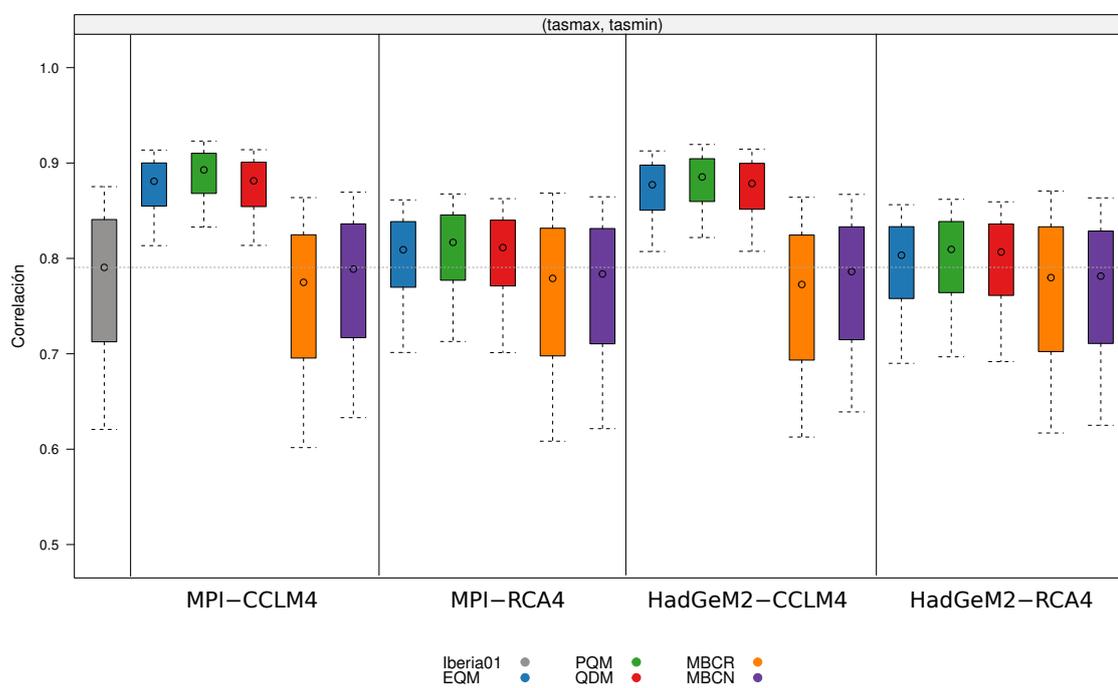


Figura 4.7: Como la Figura 4.6 la temperatura máxima (*tasmax*) y la temperatura mínima (*tasmin*).

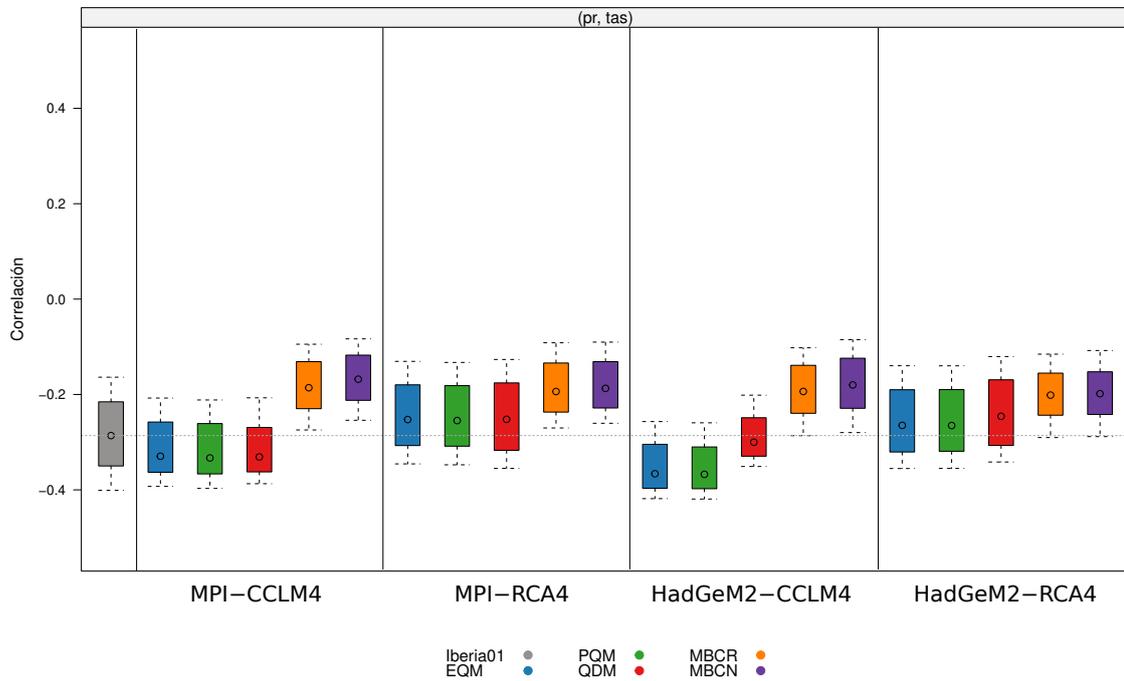


Figura 4.8: Como la Figura 4.6 pero para la precipitación (*pr*) y la temperatura media (*tas*). En este caso se ha utilizado la correlación de Spearman en lugar de la de Pearson.

Efecto de los métodos BC sobre los índices agroclimáticos

En esta sección se explora el impacto de los métodos de corrección de sesgo sobre los índices agroclimáticos, incluyendo el índice Huglin (HI), los Grados-Días Biológicamente Efectivos (BEDD) y el índice de Branas, Bernon y Levadoux (BBLI).

En la Figura 4.9 se ha observado que la corrección de sesgos logró reducir de manera efectiva el error medio en el índice de Huglin, manteniéndolo generalmente por debajo del 5%. Sin embargo, se han identificado ciertas variaciones regionales que merece la pena destacar. En particular, se observaron sesgos ligeramente más altos en el noroeste de la Península Ibérica y en áreas montañosas. Estas diferencias resaltan la importancia de considerar las particularidades climáticas de cada región, ya que pueden estar sujetas a condiciones climáticas especiales que los modelos climáticos y las correcciones pueden tener dificultades para capturar con precisión.

Además, cabe resaltar que el patrón espacial del sesgo en el índice de Huglin parece depender tanto del modelo climático específico que se corrige como del método de corrección utilizado. En particular, se ha observado que, en el caso de los métodos EQM y PQM, el patrón del sesgo parece estar más influenciado por la elección del

modelo climático. Esto subraya la importancia de la elección del modelo climático en los estudios de cambio climático y su influencia en la calidad de las predicciones para la viticultura en la región.

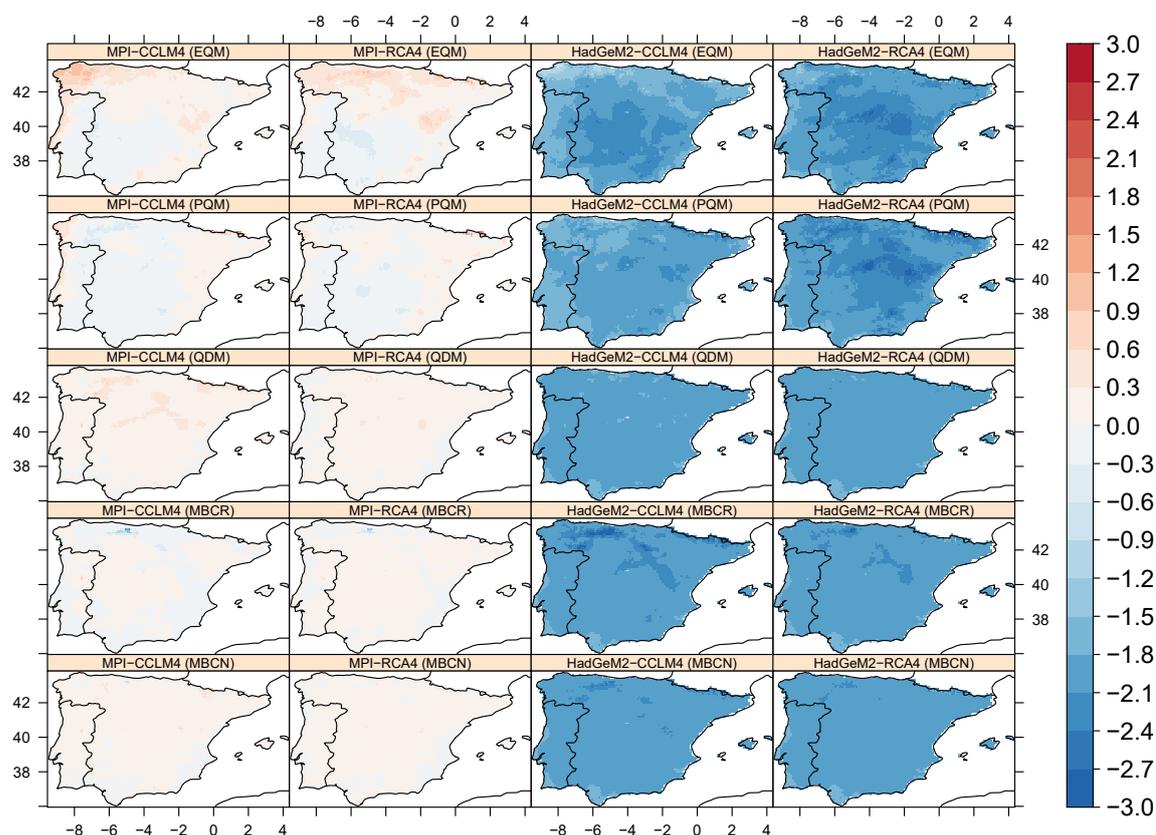


Figura 4.9: Sesgo relativo (en %) del índice HI para cada modelo climático (en columnas) y método de corrección de sesgo (en filas). El sesgo relativo, como se define en la Ecuación 3.4, se determina tomando como referencia los datos observados (Iberia01). Las variables climáticas corregidas con las que se calcula el índice corresponden a aquellas obtenidas para el experimento de validación cruzada que dividió el conjunto de datos en *2-fold* de forma aleatoria.

En relación al índice BEDD, como se muestra en la Figura 4.10, se ha observado que los sesgos en el BEDD son generalmente bajos, con valores inferiores al 4%, en la mayor parte de la Península Ibérica. No obstante, es importante señalar que las zonas montañosas muestran un sesgo más definido en comparación con el índice de Huglin. Esto se debe a que el índice de BEDD se calcula teniendo en cuenta las temperaturas máximas y mínimas diarias. Las temperaturas mínimas tienden a ser más extremas, lo que significa que sus valores se sitúan en la cola de la distribución.

Por este motivo, los métodos de BC pueden tener dificultades para ajustarlos con precisión, lo que se refleja en valores más altos del sesgo en el índice BEDD.

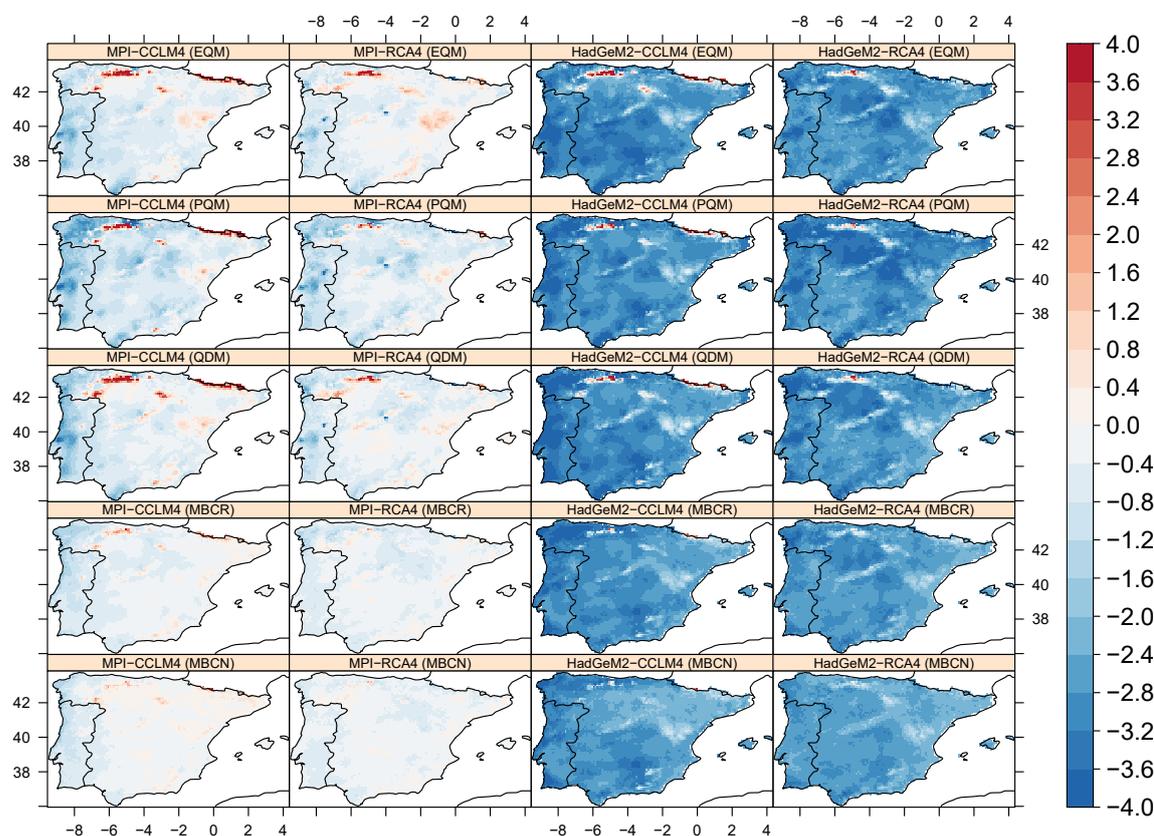


Figura 4.10: Como la Figura 4.9, pero para el índice BEDD.

En lo que respecta al índice BBLI, el patrón espacial de sesgo presenta diferencias significativas entre los distintos métodos de corrección considerados, especialmente entre los univariados y los multivariados. Estas variaciones pueden atribuirse posiblemente a problemas relacionados con la corrección de la precipitación utilizando los métodos MBC, como se discutió previamente en la sección anterior. La corrección de la precipitación parece desempeñar un papel crucial en la precisión de la corrección del índice BBLI.

Sin embargo, en el caso de los métodos de corrección univariados, se observan sesgos bajos en el índice BBLI, inferiores al 20%. Esto sugiere que estos métodos logran corregir de manera efectiva los sesgos relacionados con la precipitación, lo que es fundamental para evaluar las condiciones climáticas relevantes para la viticultura.

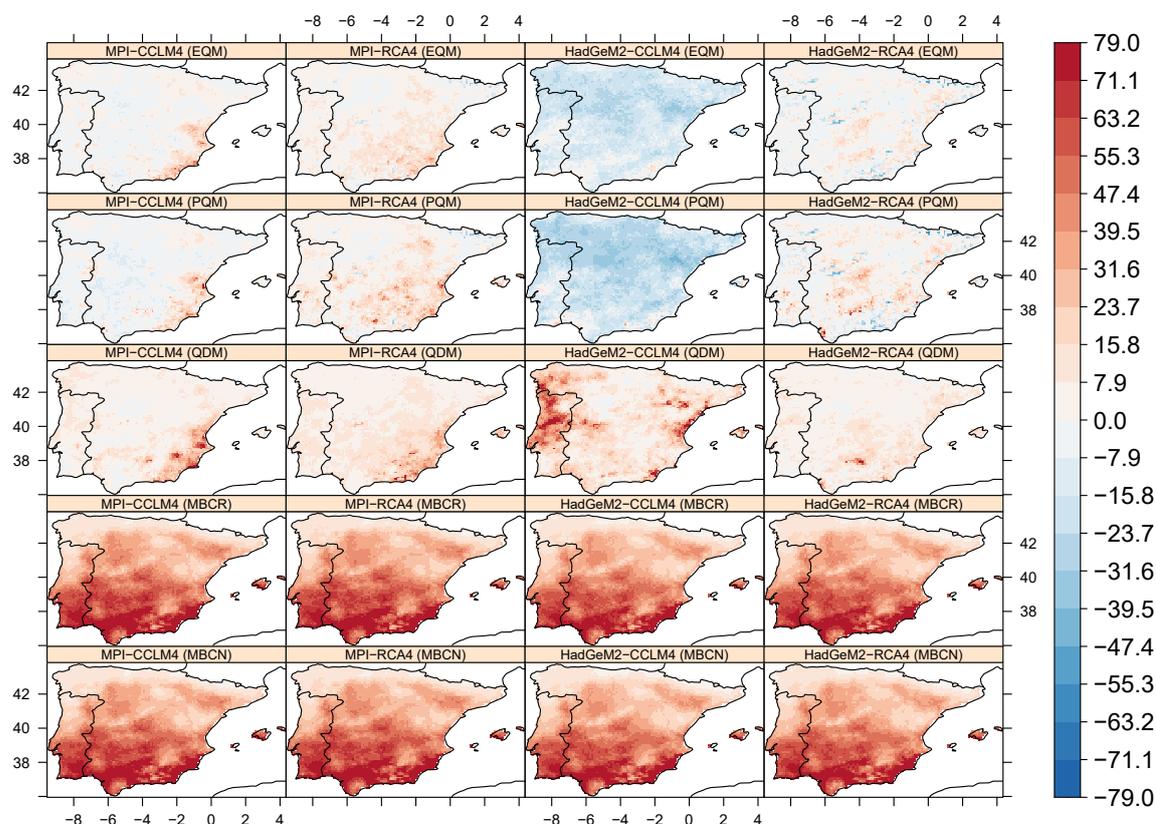


Figura 4.11: Como la Figura 4.9, pero para el índice BBI.

4.2.2. Capacidad de los métodos para representar años cálidos

En esta sección nos adentraremos en el análisis de la influencia de los métodos de corrección de sesgo en los índices agroclimáticos, centrándonos específicamente en un conjunto de datos que incluye los 10 años más cálidos en la Península Ibérica en el periodo 1986-2005 (entrenando con los 10 años más fríos de ese periodo, véase Sección 3.2.1). Tras haber explorado previamente cómo estos métodos afectan a las variables climáticas y a los índices agroclimáticos en un experimento de *2-fold* con años aleatorios, esta etapa del estudio nos permitirá comprender más a fondo la robustez de los métodos de corrección en un contexto de condiciones climáticas más cálidas.

Como muestran las Figuras 4.12 y 4.13, se observa que se intensifican los sesgos negativos en el índice de Hugin (HI) y el índice BEDD en comparación a los representados en las Figuras 4.9 y 4.10. Estos sesgos negativos están asociados principalmente a la subestimación de las temperaturas, lo que acaba teniendo un impacto

importante sobre los índices.

En concreto, el método EQM realiza una extrapolación constante, es decir, aplica el factor de corrección del último cuantil corregido a todos los valores por encima del mayor valor del periodo de calibración. Sin embargo, parece que esta extrapolación no resulta suficiente en este caso. Por otro lado, el método PQM, que no realiza una extrapolación explícita, sino que se basa en la media y la desviación estándar ajustadas en el periodo de calibración (para el caso de las temperaturas), también enfrenta dificultades en la corrección de sesgos en condiciones anómalamente cálidas.

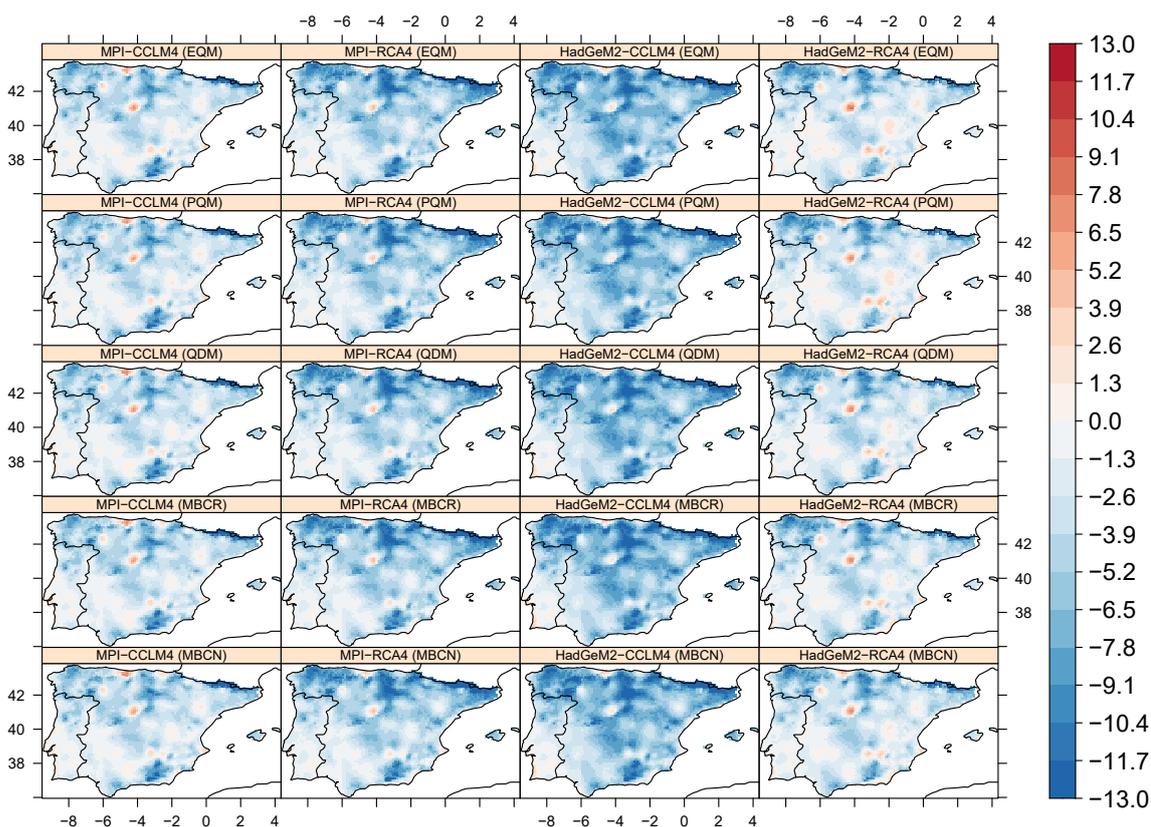


Figura 4.12: Como Figura 4.9 pero las variables utilizadas para calcular el índice corresponden a las obtenidas en el experimento de validación cruzada, donde el modelo se entrenó con años fríos y se aplicó la corrección a años cálidos. Se muestra el sesgo relativo para el subconjunto de los 10 años más cálidos en el periodo 1986-2005.

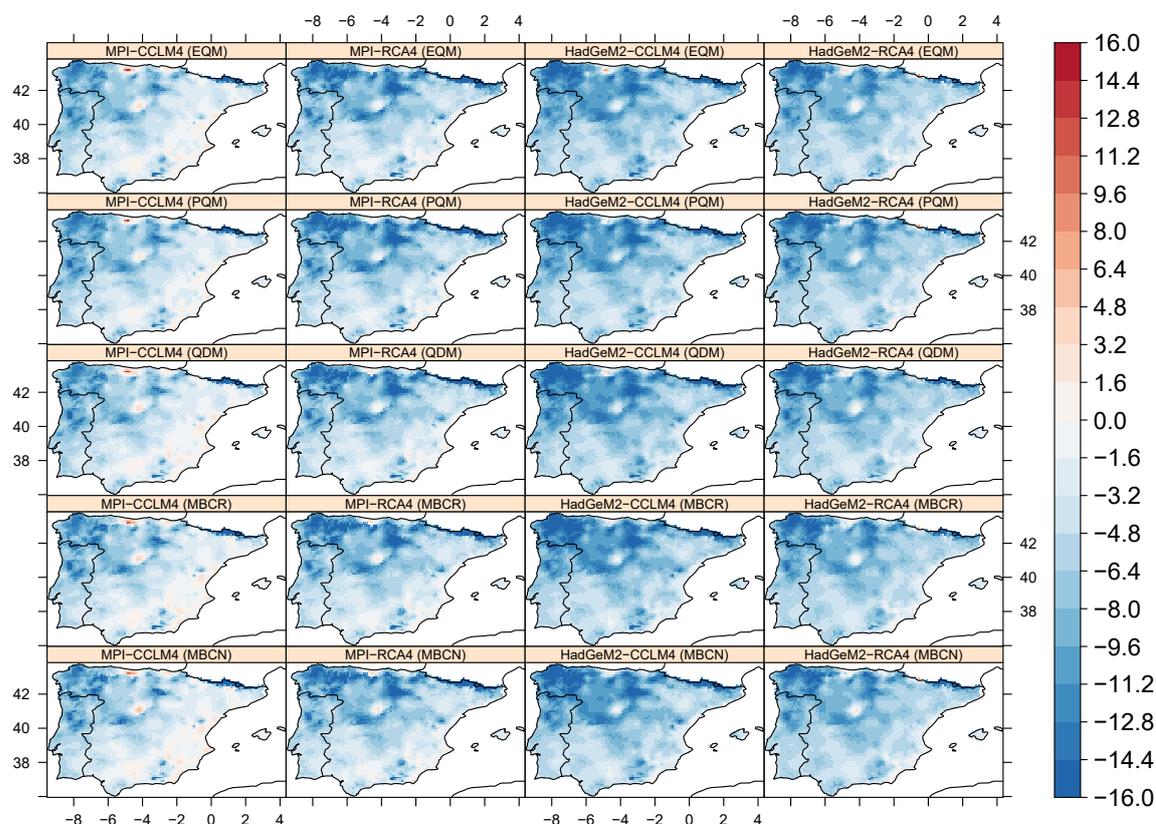


Figura 4.13: Como la Figura 4.12, pero para el índice índice BEDD.

Otro resultado relevante es que no se observan diferencias significativas entre los métodos de corrección univariados y multivariados, con la excepción del índice BBLI. En el caso de este último, la Figura 4.14 muestra que los métodos MBC no logran el rendimiento esperado, lo que podría deberse a las dificultades relacionadas con la corrección de la precipitación, como se mencionó anteriormente.

En conclusión, a partir de los resultados obtenidos en la validación cruzada con años cálidos y años fríos, se destaca una limitación importante en todos los métodos de corrección de sesgos cuando se aplican en un contexto de cambio climático con expectativas de aumento de las temperaturas. Esta limitación es evidente al comparar las representaciones de los índices agroclimáticos entre el experimento de validación cruzada con años aleatorios (Sección 4.2.1) y el realizado con la división de años cálidos y fríos (sección actual).

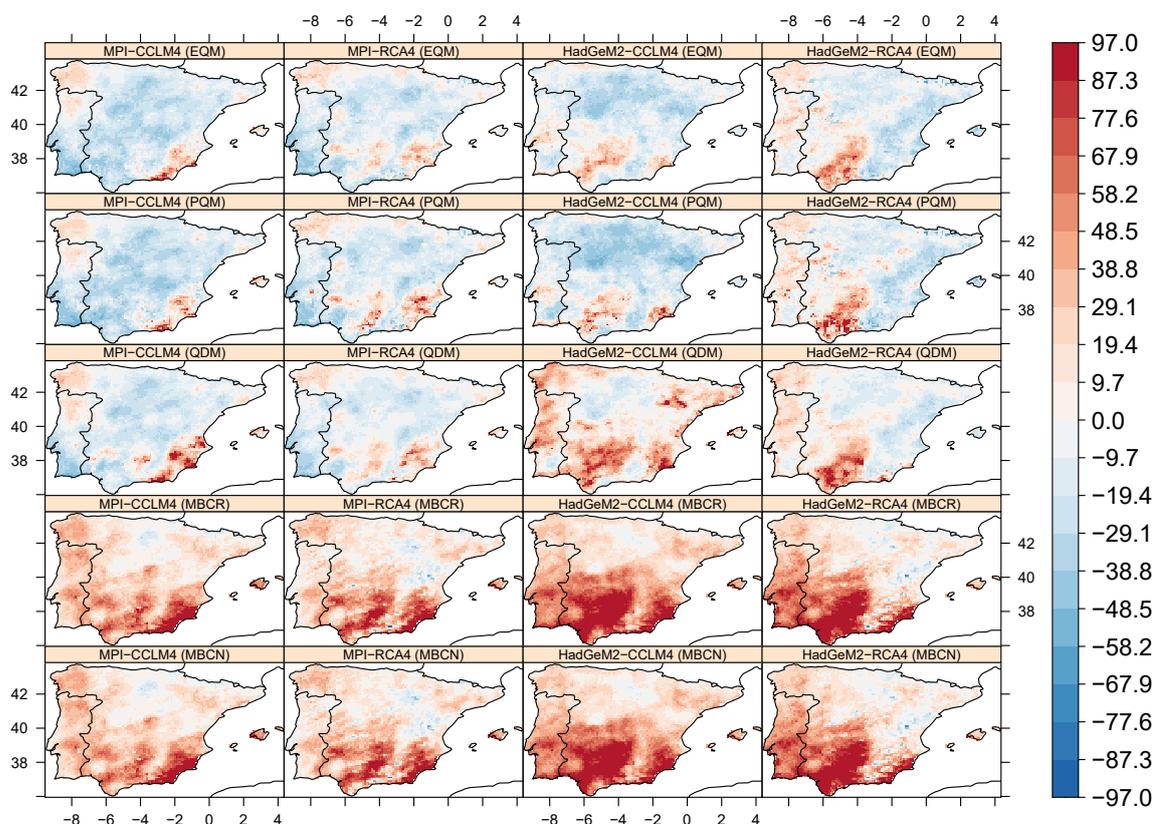


Figura 4.14: Como la Figura 4.12, pero para el índice índice BBLI.

4.3. Proyecciones de cambio climático de los índices agroclimáticos

En esta sección, nos sumergimos en el análisis de las proyecciones climáticas futuras de los cuatro modelos climáticos GCM-RCM, previamente ajustados mediante técnicas de corrección de sesgo. El objetivo principal es caracterizar, a través de los índices considerados, el impacto del cambio climático sobre el cultivo de la vid en la Península Ibérica y evaluar cómo estas correcciones afectan las señales de cambio climático proyectadas en comparación con los datos observados (Iberia01). Esta elección nos permite analizar cómo las predicciones climáticas se desvían de las condiciones históricas que la sociedad y los ecosistemas han experimentado en las últimas dos décadas.

Para llevar a cabo este estudio, las proyecciones climáticas se realizan en el contexto de un escenario de cambio climático severo como el representado por el RCP8.5. En concreto, nos centramos para cada GCM en el período de 20 años

para el cual se alcanza un calentamiento de 4°C (con respecto a los niveles preindustriales) a nivel global. En el caso del modelo HadGEM, el período considerado abarca desde el año 2062 hasta el 2081, mientras que el modelo MPI va del 2071 hasta al 2091.

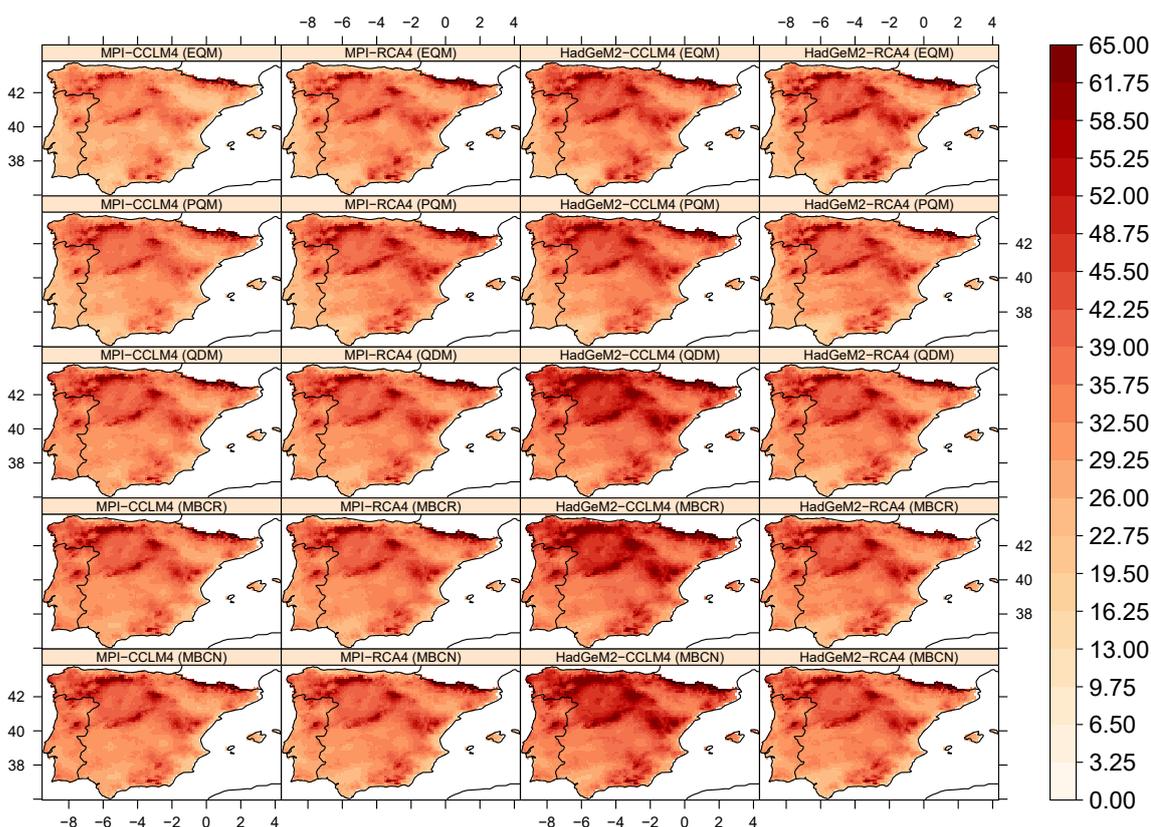


Figura 4.15: Señal de cambio climático (en %) del índice HI bajo un escenario RCP8.5 para un nivel de calentamiento global de 4°C para cada modelo climático (en columnas) y método de corrección/ajuste de sesgo (en filas). El sesgo relativo, como se define en la Ecuación 3.5, en el intervalo de tiempo comprendido entre 1986 y 2005. Las variables climáticas de los modelos han sido corregidas por los diferentes métodos de corrección de sesgo.

En primer lugar, observamos que tanto el índice de Huglin como el índice BEDD, que dependen exclusivamente de las temperaturas, muestran un aumento generalizado en toda la Península Ibérica para todos los pares GCM-RCM. Este aumento es especialmente notable en áreas como la meseta norte y las zonas montañosas, precisamente aquellas regiones que originalmente tenían valores más bajos según las observaciones (como se ilustra en la Figura 4.1). Los incrementos en estos índices agroclimáticos oscilan entre un 45 % y un 65 %, lo que refleja claramente el incre-

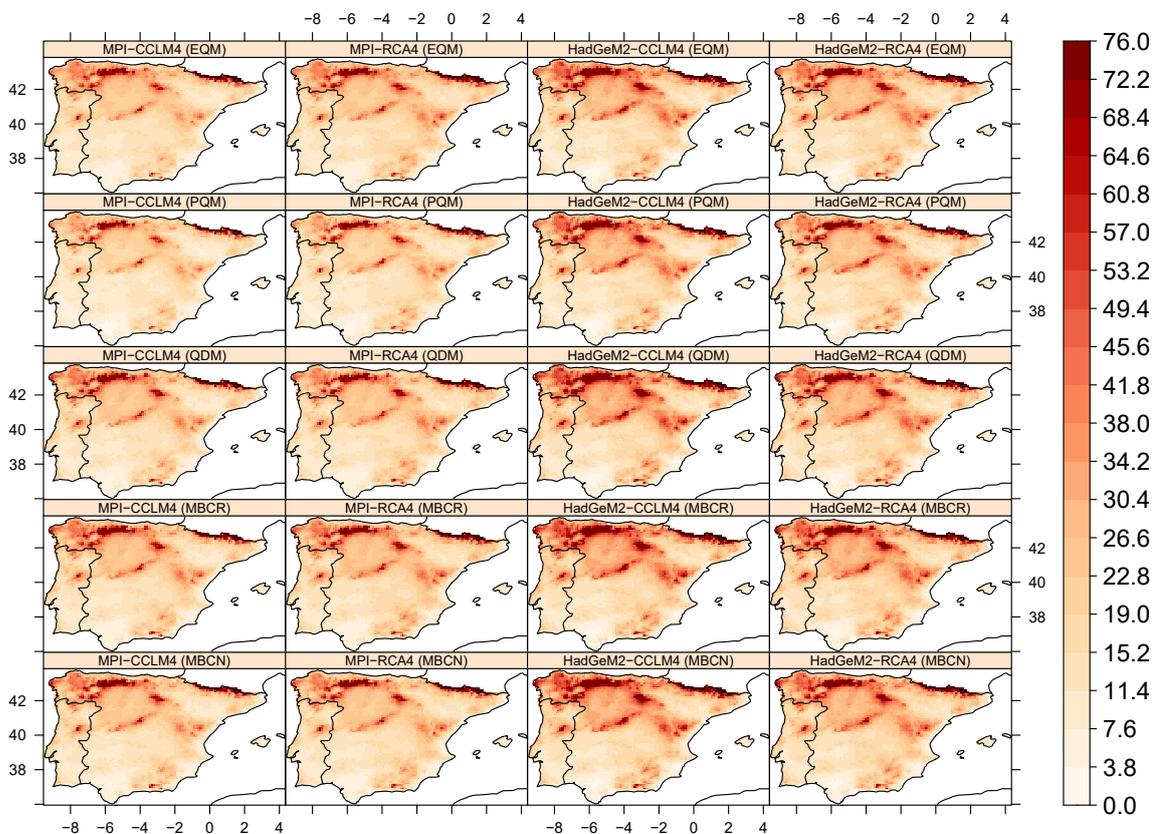


Figura 4.16: Como la Figura 4.15 pero para el índice BEDD.

mento de las temperatura y condiciones climáticas más cálidas.

Es importante destacar que, en línea con lo esperado debido a las diferencias en la sensibilidad climática de los GCM utilizados, las simulaciones anidadas al modelo HadGEM, del cual se conoce que es más cálido que el MPI, muestran un aumento mayor en estos índices. Este resultado resalta la influencia del modelo climático subyacente (especialmente del GCM) en la proyección de las condiciones agroclimáticas futuras.

Otro resultado relevante que se observa en las Figuras 4.15 y 4.16 es que, para un par GCM-RCM concreto, no hay grandes diferencias entre los distintos métodos de corrección de sesgos aplicados.

En el caso del índice BBLI (Figura 4.17), todas las proyecciones de modelos climáticos sugieren una disminución en amplias partes de la Península Ibérica, siendo esta especialmente notable en las simulaciones de CCLM4, independientemente del GCM utilizado. Esta disminución podría estar relacionada con una reducción en la precipitación en estas áreas, lo que tendría un impacto significativo en las condiciones

agroclimáticas futuras.

Es importante destacar que los resultados obtenidos con el uso de los métodos MBC para el índice BBLI no son fiables, lo que se fundamenta en los malos resultados obtenidos en el experimento de evaluación previo (como se discutió en la Sección 4.2). Esta falta de fiabilidad en las simulaciones con MBC subraya la importancia de elegir métodos de corrección adecuados y de evaluar rigurosamente su desempeño antes de aplicarlos en estudios de cambio climático y proyecciones agroclimáticas.

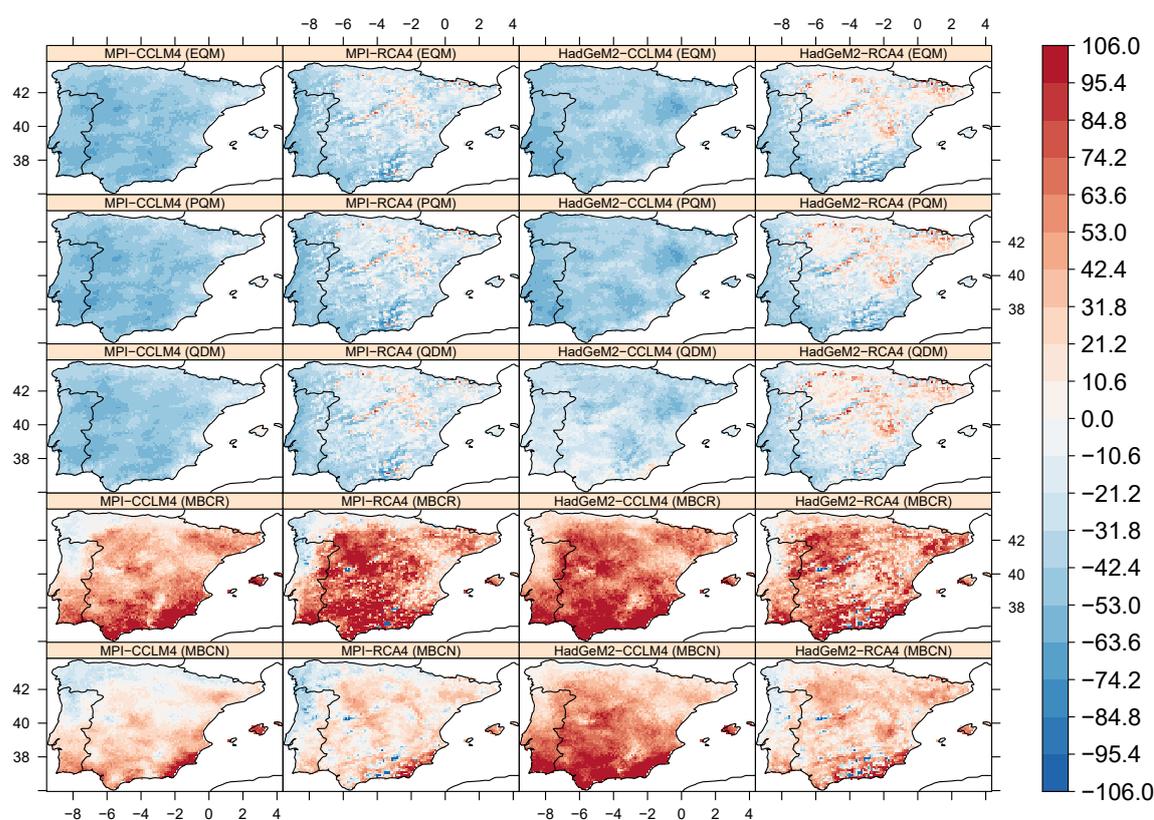


Figura 4.17: Como la Figura 4.15 pero para el índice BEDD.

Como se ha explicado, todos los resultados presentados en esta sección se basan en el escenario RCP8.5, que representa un contexto de altas emisiones de gases de efecto invernadero. Es importante destacar que, para proporcionar una visión más completa y precisa de las condiciones agroclimáticas futuras en la Península Ibérica, es necesario considerar y evaluar también otros escenarios de emisión. Esto permitiría ofrecer información más diversa y valiosa a los organismos encargados de elaborar estrategias de adaptación al cambio climático.

CAPÍTULO 5

Conclusiones Principales y Trabajo Futuro

5.1. Conclusiones Principales

En este estudio se llevó a cabo un análisis exhaustivo sobre la idoneidad de distintas técnicas estadísticas para la corrección de sesgos en los modelos regionales del clima, utilizando tanto un enfoque univariado como otro multivariado, y impacto sobre los índices agroclimáticos relevantes para la viticultura en la Península Ibérica. A continuación, exponemos brevemente las principales conclusiones que se han obtenido, las cuales están alineadas con los objetivos presentados en la Sección 1.2 y arrojan luz sobre diversos aspectos:

La evaluación de los métodos de corrección de sesgos reveló que, en términos generales, no existen grandes diferencias entre las distintas implementaciones analizadas. Además, la incertidumbre asociada a la elección del método de corrección de sesgos resultó ser menor que la asociada al modelo climático (en particular, en el caso de los modelos globales). Aun así, los resultados obtenidos por medio de distintos métodos de corrección de sesgos presentan cierta variabilidad.

A pesar de la alta correlación observada entre las variables de temperatura, como temperatura media y temperatura máxima, así como entre la temperatura mínima y temperatura máxima, no se encontró una mejora significativa en la aplicación de métodos multivariados en comparación con los enfoques univariados para la corrección de sesgos. Por tanto, en el caso de estas variables, no se necesita recurrir a técnicas multivariadas, que son más complejas que las univariadas. Por otro lado, la corrección de la precipitación presentó dificultades más notables en su ajuste

utilizando métodos multivariados. Esto puede estar relacionado con la intrincada estructura de dependencia que existe entre la precipitación y otras variables climáticas. La precipitación está influenciada por múltiples factores climáticos además de la temperatura, como la humedad, la presión atmosférica, etc. Es posible que estos factores deban considerarse y abordarse de manera simultánea en la corrección de la precipitación usando métodos multivariados.

A partir de los resultados obtenidos en la validación cruzada con años cálidos y años fríos, se destaca una limitación importante en todos los métodos de corrección de sesgos cuando se aplican en un contexto de cambio climático con expectativas de aumento de las temperaturas. Esta limitación es evidente al comparar las representaciones de los índices agroclimáticos entre el experimento de validación cruzada con años aleatorios (como se discutió en la Sección 4.2.1) y el realizado con la división de años cálidos y fríos (Sección 4.2.2).

Las proyecciones de cambio climático realizadas bajo el escenario RCP8.5 y a un nivel de calentamiento global de 4°C, que representa uno de los escenarios futuros más extremos, indican un aumento generalizado en los índices agroclimáticos de Huglin (HI) y BEDD en toda la Península Ibérica. Esta tendencia es especialmente pronunciada en áreas que inicialmente presentaban los valores observados más bajos, como la meseta norte y las zonas montañosas. En contraste, el índice BBLI muestra una disminución en amplias partes de la Península, especialmente de acuerdo al RCM CCLM4, independientemente del modelo climático GCM subyacente. Estos resultados podrían estar asociados a una reducción en la precipitación, lo que sugiere que las proyecciones futuras podrían llevar a condiciones climáticas menos favorables para la viticultura en ciertas áreas.

5.2. Trabajo Futuro

A continuación, delineamos algunas actividades futuras que constituyen la continuación natural de este trabajo. Estas actividades abordan cuestiones clave identificadas durante la realización del mismo y tienen como objetivo ampliar nuestra comprensión y mejorar la fiabilidad de las proyecciones climáticas.

- Profundizar hacia una evaluación más detallada de los métodos de corrección multivariados, especialmente en lo que respecta a la precipitación. Se buscará ajustar el número de iteraciones, incorporar un conjunto más amplio de variables climáticas en el proceso de corrección y realizar un análisis más exhaustivo para comprender limitaciones y posibles mejoras en el ajuste de la precipitación.

- Extender el análisis a otros escenarios de emisiones (RCPs) y/o períodos de calentamiento global para proporcionar una visión más integral del impacto que se podría esperar del cambio climático sobre la producción vitivinícola.

5.3. *Reproducibilidad de los Resultados*

Se ha realizado un esfuerzo importante para garantizar la reproducibilidad de todos los resultados presentados en este trabajo, adoptando en la medida de lo posible los principios FAIR para la gestión científica de datos (Wilkinson et al., 2016).

`climate4R` (Iturbide et al., 2019b) es un conjunto de paquetes de *R* para el acceso transparente a los datos climáticos, su posprocesamiento (incluida la corrección de sesgos/regionalización en el paquete `downscaleR`, Bedia et al. (2020)) y su visualización (Frías et al., 2018). `climate4R` se basa en el uso de estructuras de datos en rejilla y datos puntuales procedentes de observaciones, reanálisis (pseudo-observaciones), predicciones estacionales y proyecciones climáticas. Por otro lado, `climate4R` está conectado de forma transparente (y remota) al Santander Climate Data Service (<https://scds.es>), que ofrece varios conjuntos de datos de última generación (incluidos los subconjuntos de modelos de CMIP5 y CORDEX). Recientemente `climate4R` ha sentado el marco del desarrollo del Atlas interactivo (<https://interactive-atlas.ipcc.ch/>) del IPCC (Panel Intergubernamental de Expertos sobre el Cambio Climático), que por primera vez proporciona un soporte en forma digital al último informe de evaluación de dicho organismo y sigue la aplicación de principios de buenas prácticas FAIR de encontrabilidad, accesibilidad, interoperabilidad y reutilización (Iturbide et al., 2022).

Para llevar a cabo las implementaciones de las librerías `MBC` y `agroclimR` en `climate4R`, se han modificado las librerías `downscaleR` y `climate4R.indices`, respectivamente. Ambas librerías, con las modificaciones pertinentes que han permitido desarrollar este TFM, se pueden encontrar en los siguientes enlaces de GitHub:

-  `downscaleR`
-  `climate4R.indices`

Estas dos librerías modificadas suponen una mejora notable en `climate4R`, puesto que extienden el conjunto de funcionalidades disponibles en este framework.

Todos los recursos necesarios para reproducir los resultados presentados en este trabajo se encuentran en un repositorio público en GitHub:  TFM. Los archivos de

código están organizados en el directorio *Scripts*, junto con los datos necesarios en el directorio *Datos*. También se ha incluido archivos R que generan las figuras presentadas en este documento, ubicados en *Scripts/Create_Figures*. Todo el proyecto se desarrolló utilizando el lenguaje de programación gratuito R.

5.4. Contexto de la Investigación

Quiero expresar mi profundo agradecimiento a todas las instituciones y entidades/personas colaboradoras que han hecho posible la realización de este trabajo de investigación.

En primer lugar, al grupo de trabajo en clima regional del *World Climate Research Program*, organismo coordinador de CORDEX y panel responsable de CMIP5, gracias a los cuales se han desarrollado las simulaciones de modelos del clima usadas en este trabajo.

En segundo lugar, a los grupos de modelización climática por producir y poner a disposición sus simulaciones (listadas en la Tabla 2.1) y a la *Earth System Grid Federation*, un esfuerzo internacional dirigido por el Programa de Diagnóstico e Intercomparación de Modelos Climáticos del Departamento de Energía de EE.UU., la Red Europea para la Modelización del Sistema Terrestre y otros socios de la red.

Me gustaría asimismo mencionar que el acceso a los datos que se utilizan en este trabajo se realizó a través del Servicio Climático de Datos de Santander (<https://scds.es/en>), mantenido por el Grupo de Meteorología de la Universidad de Cantabria, formado por investigadores del Instituto de Física de Cantabria y del Departamento de Matemática Aplicada y Ciencias de la Computación de la Universidad de Cantabria.

Por último, me gustaría también destacar que este trabajo se enmarca en las actividades del proyecto nacional de I+D “Eventos extremos compuestos para la evaluación de los impactos del cambio climático en la agricultura” (COMPOUND: TED2021-131334A-I00), financiado por MCIN/AEI/10.13039/501100011033 y por los fondos NextGenerationEU/PRTR de la Unión Europea.

Bibliografía

- ARANGUREN GONZÁLEZ, M., RODRÍGUEZ, J., and ACEBEDO, Y. (2015). Determinación de los índices bioclimáticos y tipo de clima para la vid en las condiciones de jagüey grande, matanzas, cuba.
- BAN, N., SCHMIDLI, J., and SCHÄR, C. (2014). Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *Journal of Geophysical Research: Atmospheres*, 119(13):7889–7907. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD021478>.
- BEDIA, J., BAÑO MEDINA, J., LEGASA, M. N., ITURBIDE, M., MANZANAS, R., HERRERA, S., CASANUEVA, A., SAN-MARTÍN, D., COFIÑO, A. S., and GUTIÉRREZ, J. M. (2020). Statistical downscaling with the downscaleR package (v3.1.0): contribution to the value intercomparison experiment. *Geoscientific Model Development*, 13(3):1711–1735. URL <https://gmd.copernicus.org/articles/13/1711/2020/>.
- BELO-PEREIRA, M., DUTRA, E., and VITERBO, P. (2011). Evaluation of global precipitation data sets over the iberian peninsula. *Journal of Geophysical Research: Atmospheres*, 116(D20). URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD015481>.
- BEYER, R., KRAPP, M., and MANICA, A. (2020). An empirical evaluation of bias correction methods for palaeoclimate simulations. *Climate of the Past*, 16(4):1493–1508. URL <https://cp.copernicus.org/articles/16/1493/2020/>.

- BOIS, M. (2016). A transnational friendship in the age of extremes: Leon trotsky and the pfemferts. *Twentieth Century Communism*, 10(10):9–29.
- BRANAS, J., BERNON, G., and LEVADOUX, L. (1946). *Eléments de viticulture générale*.
- CANNON, A. J. (2016). Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure. *Journal of Climate*, 29(19):7045–7064. URL <https://journals.ametsoc.org/view/journals/clim/29/19/jcli-d-15-0679.1.xml>.
- CANNON, A. J. (2018). Multivariate quantile mapping bias correction: an n-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics*, 50(1):31–49. URL <https://doi.org/10.1007/s00382-017-3580-6>.
- 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? *Journal of Climate*, 28(17):6938 – 6959. URL <https://journals.ametsoc.org/view/journals/clim/28/17/jcli-d-14-00754.1.xml>.
- CARDELL, M. F., AMENGUAL, A., and ROMERO, R. (2019). Future effects of climate change on the suitability of wine grape production across europe. *Regional Environmental Change*, 19(8):2299–2310. URL <https://doi.org/10.1007/s10113-019-01502-x>.
- CASANUEVA, A. (2016). *Comparison of statistical and dynamical climate downscaling techniques: screening of methods for their use in impact studies*. Ph.D. thesis. URL <http://hdl.handle.net/10902/8321>.
- CASANUEVA, A., HERRERA, S., ITURBIDE, M., LANGE, S., JURY, M., DOSIO, A., MARAUN, D., and GUTIÉRREZ, J. M. (2020). Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. *Atmospheric Science Letters*, 21(7):e978. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/asl.978>.
- CASANUEVA, A., KOTLARSKI, S., HERRERA, S., FISCHER, A. M., KJELLSTROM, T., and SCHWIERZ, C. (2019). Climate projections of a multivariate heat stress

- index: the role of downscaling and bias correction. *Geoscientific Model Development*, 12(8):3419–3438. URL <https://gmd.copernicus.org/articles/12/3419/2019/>.
- CORNES, R. C., VAN DER SCHRIER, G., VAN DEN BESSELAAR, E. J., and JONES, P. D. (2018). An ensemble version of the e-obs temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123(17):9391–9409.
- DÉQUÉ, M. (2007). Frequency of precipitation and temperature extremes over france in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global and Planetary Change*, 57(1):16–26. URL <https://www.sciencedirect.com/science/article/pii/S0921818106002748>. Extreme Climatic Events.
- FESER, F., ROCKEL, B., VON STORCH, H., WINTERFELDT, J., and ZAHN, M. (2011). Regional climate models add value to global model data: a review and selected examples. *Bulletin of the American Meteorological Society*, 92(9):1181–1192.
- FRÍAS, M., ITURBIDE, M., MANZANAS, R., BEDIA, J., FERNÁNDEZ, J., HERRERA, S., COFIÑO, A., and GUTIÉRREZ, J. (2018). An r package to visualize and communicate uncertainty in seasonal climate prediction. *Environmental Modelling Software*, 99:101–110. URL <https://www.sciencedirect.com/science/article/pii/S1364815217305157>.
- GARCÍA-LEÓN, D., CASANUEVA, A., STANDARDI, G., BURGSTALL, A., FLOURIS, A. D., and NYBO, L. (2021). Current and projected regional economic impacts of heatwaves in europe. *Nature Communications*, 12(1):5807. URL <https://doi.org/10.1038/s41467-021-26050-z>.
- GLADSTONES, J. ET AL. (1992). *Viticulture and environment..* Winetitles.
- HERRERA, S., CARDOSO, R. M., SOARES, P. M., ESPÍRITO-SANTO, F., VITERBO, P., and GUTIÉRREZ, J. M. (2019). Iberia01: a new gridded dataset of daily precipitation and temperatures over iberia. *Earth System Science Data*, 11(4):1947–1956. URL <https://essd.copernicus.org/articles/11/1947/2019/>.
- HERRERA, S., GUTIÉRREZ, J. M., ANCELL, R., PONS, M. R., FRÍAS, M. D., and FERNÁNDEZ, J. (2012). Development and analysis of a 50-year high-resolution

- daily gridded precipitation dataset over spain (spain02). *International Journal of Climatology*, 32(1):74–85. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.2256>.
- HEWITSON, B. C. and CRANE, R. G. (1996). Climate downscaling: techniques and application. *Climate Research*, 7(2):85–95.
- HONORIO GUIADO, F. (2016). *Caracterización y zonificación vitícola de España mediante análisis multivariante de variables bioclimáticas*. Ph.D. thesis. URL <http://hdl.handle.net/10662/4090>.
- HUGLIN, M. (1978). Nouveau mode d'évaluation des possibilités héliothermiques d'un milieu viticole.
- ITURBIDE, M., BEDIA, J., HERRERA, S., BAÑO-MEDINA, J., FERNÁNDEZ, J., FRÍAS, M., MANZANAS, R., SAN-MARTÍN, D., CIMADEVILLA, E., COFIÑO, A., and GUTIÉRREZ, J. (2019a). The r-based climate4r open framework for reproducible climate data access and post-processing. *Environmental Modelling Software*, 111:42–54. URL <https://www.sciencedirect.com/science/article/pii/S1364815218303049>.
- ITURBIDE, M., BEDIA, J., HERRERA, S., BAÑO-MEDINA, J., FERNÁNDEZ, J., FRÍAS, M., MANZANAS, R., SAN-MARTÍN, D., CIMADEVILLA, E., COFIÑO, A., and GUTIÉRREZ, J. (2019b). The r-based climate4r open framework for reproducible climate data access and post-processing. *Environmental Modelling Software*, 111:42–54. URL <https://www.sciencedirect.com/science/article/pii/S1364815218303049>.
- ITURBIDE, M., FERNÁNDEZ, J., GUTIÉRREZ, J. M., PIRANI, A., HUARD, D., AL KHOURDAJIE, A., BAÑO-MEDINA, J., BEDIA, J., CASANUEVA, A., CIMADEVILLA, E., COFIÑO, A. S., DE FELICE, M., DIEZ-SIERRA, J., GARCÍA-DÍEZ, M., GOLDIE, J., HERRERA, D. A., HERRERA, S., MANZANAS, R., MILOVAC, J., RADHAKRISHNAN, A., SAN-MARTÍN, D., SPINUSO, A., THYNG, K. M., TRENHAM, C., and YELEKÇI, Ö. (2022). Implementation of fair principles in the ipcc: the wgi ar6 atlas repository. *Scientific Data*, 9(1):629. URL <https://doi.org/10.1038/s41597-022-01739-y>.
- JONES, G. (2006). Climate and terroir: Impacts of climate variability and change on wine. *Geoscience Canada Reprint Series*.

- JONES, G., DUCHÊNE, E., TOMASI, D., YUSTE, J., BRASLAVSKA, O., SCHULTZ, H., MARTINEZ, C., BOSO, S., LANGELLIER, F., PERRUCHOT, C., ET AL. (2005). Changes in european winegrape phenology and relationships with climate. In *XIV International GESCO Viticulture Congress, Geisenheim, Germany, 23-27 August, 2005*, pp. 54–61. Groupe d’Etude des Systemes de COnduite de la vigne (GESCO).
- KIDSON, J. W. and THOMPSON, C. S. (1998). A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate*, 11(4):735–753.
- LIU, M., RAJAGOPALAN, K., CHUNG, S. H., JIANG, X., HARRISON, J., NERGUI, T., GUENTHER, A., MILLER, C., REYES, J., TAGUE, C., CHOATE, J., SALATHÉ, E. P., STÖCKLE, C. O., and ADAM, J. C. (2014). What is the importance of climate model bias when projecting the impacts of climate change on land surface processes? *Biogeosciences*, 11(10):2601–2622. URL <https://bg.copernicus.org/articles/11/2601/2014/>.
- LORENZO, M. N., TABOADA, J. J., LORENZO, J. F., and RAMOS, A. M. (2013). Influence of climate on grape production and wine quality in the rías baixas, north-western spain. *Regional Environmental Change*, 13(4):887–896. URL <https://doi.org/10.1007/s10113-012-0387-1>.
- MARAUN, D. (2013). Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *Journal of Climate*, 26(6):2137 – 2143. URL <https://journals.ametsoc.org/view/journals/clim/26/6/jcli-d-12-00821.1.xml>.
- MARAUN, D., WETTERHALL, F., IRESON, A., CHANDLER, R., KENDON, E., WIDMANN, M., BRIENEN, S., RUST, H., SAUTER, T., THEMESL, M., ET AL. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of geophysics*, 48(3).
- MARAUN, D., WIDMANN, M., GUTIÉRREZ, J. M., KOTLARSKI, S., CHANDLER, R. E., HERTIG, E., WIBIG, J., HUTH, R., and WILCKE, R. A. (2015). Value: A framework to validate downscaling approaches for climate change studies. *Earth’s Future*, 3(1):1–14. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014EF000259>.

- MEEHL, G. A. (1995). Global coupled general circulation models. *Bulletin of the American Meteorological Society*, 76(6):951–957.
- PIANI, C., HAERTER, J. O., and COPPOLA, E. (2010). Statistical bias correction for daily precipitation in regional climate models over europe. *Theoretical and Applied Climatology*, 99(1):187–192. URL <https://doi.org/10.1007/s00704-009-0134-9>.
- PITIE, F., KOKARAM, A., and DAHYOT, R. (2005). N-dimensional probability density function transfer and its application to color transfer. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 2, pp. 1434–1439 Vol. 2.
- R CORE TEAM (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- RAMOS-CALZADO, P., AMBAR-FRANCÉS, P., CASADO-CALLE, M., PASTOR-SAAVEDRA, A., and RODRIGUEZ-CAMINO, E. (2017). *Guía de escenarios regionalizados de cambio climático sobre España a partir de los resultados del IPCC-AR5*.
- ROCHETA, E., EVANS, J. P., and SHARMA, A. (2014). Assessing atmospheric bias correction for dynamical consistency using potential vorticity. *Environmental Research Letters*, 9(12):124010. URL <https://dx.doi.org/10.1088/1748-9326/9/12/124010>.
- ROELL, Y. E., PENG, Y., BEUCHER, A., GREVE, M. B., and GREVE, M. H. (2020). Development of hierarchical terron workflow based on gridded data—a case study in denmark. *Computers & Geosciences*, 138:104454.
- RUMMUKAINEN, M. (2016). Added value in regional climate modeling. *WIRES Climate Change*, 7(1):145–159. URL <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcc.378>.
- SANTOS, J. A., MALHEIRO, A. C., KARREMANN, M. K., and PINTO, J. G. (2011). Statistical modelling of grapevine yield in the port wine region under present and future climate conditions. *International Journal of Biometeorology*, 55(2):119–131. URL <https://doi.org/10.1007/s00484-010-0318-0>.

- SERRANO-NOTIVOLI, R., TOMÁS-BURGUERA, M., MARTÍ, A., and BEGUERÍA, S. (2020). An integrated package to evaluate climatic suitability for agriculture. *Computers and Electronics in Agriculture*, 176:105473. URL <https://www.sciencedirect.com/science/article/pii/S0168169920306530>.
- STEVENS, B. and BONY, S. (2013). What are climate models missing? *Science*, 340(6136):1053–1054. URL <https://www.science.org/doi/abs/10.1126/science.1237554>.
- SUN, Q., MIAO, C., DUAN, Q., ASHOURI, H., SOROOSHIAN, S., and HSU, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1):79–107. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017RG000574>.
- SØRLAND, S. L., SCHÄR, C., LÜTHI, D., and KJELLSTRÖM, E. (2018). Bias patterns and climate change signals in gcm-rcm model chains. *Environmental Research Letters*, 13(7):074017. URL <https://dx.doi.org/10.1088/1748-9326/aacc77>.
- TAYLOR, K., RONALD, S., and MEEHL, G. (2011). An overview of cmip5 and the experiment design. *Bulletin of the American Meteorological Society*, 93:485–498.
- TEUTSCHBEIN, C. and SEIBERT, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456-457:12–29. URL <https://www.sciencedirect.com/science/article/pii/S0022169412004556>.
- TONIETTO, J. and CARBONNEAU, A. (2004). A multicriteria climatic classification system for grape-growing regions worldwide. *Agricultural and forest meteorology*, 124(1-2):81–97.
- VAN LEEUWEN, C., FRIANT, P., CHONÉ, X., TREGOAT, O., KOUNDOURAS, S., and DUBOURDIEU, D. (2004). Influence of climate, soil, and cultivar on terroir. *American Journal of Enology and Viticulture*, 55(3):207–217. URL <https://www.ajevonline.org/content/55/3/207>.
- VICENTE-SERRANO, S. M., BEGUERÍA, S., and LÓPEZ-MORENO, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7):1696 –

1718. URL <https://journals.ametsoc.org/view/journals/clim/23/7/2009jcli2909.1.xml>.

WANG, X., THOMPSON, D. K., MARSHALL, G. A., TYMSTRA, C., CARR, R., and FLANNIGAN, M. D. (2015). Increasing frequency of extreme fire weather in canada with climate change. *Climatic Change*, 130(4):573–586. URL <https://doi.org/10.1007/s10584-015-1375-5>.

WILCKE, R. A. I., MENDLIK, T., and GOBIET, A. (2013). Multi-variable error correction of regional climate models. *Climatic Change*, 120(4):871–887. URL <https://doi.org/10.1007/s10584-013-0845-x>.

WILKINSON, M. D., DUMONTIER, M., AALBERSBERG, I. J., APPLETON, G., AXTON, M., BAAK, A., BLOMBERG, N., BOITEN, J.-W., DA SILVA SANTOS, L. B., BOURNE, P. E., BOUWMAN, J., BROOKES, A. J., CLARK, T., CROSAS, M., DILLO, I., DUMON, O., EDMUNDS, S., EVELO, C. T., FINKERS, R., GONZALEZ-BELTRAN, A., GRAY, A. J., GROTH, P., GOBLE, C., GRETHE, J. S., HERINGA, J., 'T HOEN, P. A., HOOFT, R., KUHN, T., KOK, R., KOK, J., LUSHER, S. J., MARTONE, M. E., MONS, A., PACKER, A. L., PERSSON, B., ROCCA-SERRA, P., ROOS, M., VAN SCHAIK, R., SANSONE, S.-A., SCHULTES, E., SENGSTAG, T., SLATER, T., STRAWN, G., SWERTZ, M. A., THOMPSON, M., VAN DER LEI, J., VAN MULLIGEN, E., VELTEROP, J., WAAGMEESTER, A., WITTENBURG, P., WOLSTENCROFT, K., ZHAO, J., and MONS, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018. URL <http://www.nature.com/articles/sdata201618>.