

**SOCIEDAD CHILENA DE INGENIERÍA HIDRÁULICA**  
**XXVI CONGRESO CHILENO DE INGENIERÍA HIDRÁULICA**

**DIAGNÓSTICO DE MODELOS DEL CMIP6 PARA EVALUAR EL IMPACTO DEL  
CAMBIO CLIMÁTICO EN CHILE**

**FELIPE GATEÑO<sup>1</sup>**  
**PABLO MENDOZA<sup>2,3</sup>**  
**NICOLÁS VÁSQUEZ<sup>2</sup>**  
**HÉCTOR JIMÉNEZ<sup>3</sup>**  
**CATALINA JEREZ<sup>3</sup>**  
**MIGUEL LAGOS-ZÚÑIGA<sup>3,4</sup>**  
**XIMENA VARGAS<sup>2</sup>**  
**EDUARDO RUBIO-ÁLVAREZ<sup>1</sup>**  
**SANTIAGO MONTSERRAT<sup>3</sup>**

**RESUMEN**

En este trabajo, se presenta un esquema metodológico para evaluar las simulaciones mensuales de precipitación y temperatura de modelos climáticos globales (MCGs) con el fin de disminuir la dispersión en las proyecciones climáticas anuales y estacionales. Específicamente, se propone el *Past Performance Index (PPI)* con el objetivo de evaluar cinco aspectos de la climatología histórica de los modelos: (i) las medias climatológicas; (ii) la variabilidad interanual; (iii) el ciclo estacional; (iv) la distribución probabilística mensual; y (v) los patrones espaciales de las medias climatológicas. La formulación del *PPI* permite diagnosticar y clasificar los MCGs en un diagrama sencillo que muestra el rendimiento conjunto para la precipitación y la temperatura. Se demuestra la utilidad de este enfoque para evaluar 27 modelos de la Sexta Fase de Intercomparación de Modelos Acoplados (CMIP6), evidenciando que es posible restringir la dispersión de las proyecciones en cinco regiones con climas muy diferentes a lo largo de Chile continental. También examinamos el grado de correspondencia entre el conjunto de modelos clasificados como "satisfactorios" en base al *PPI* y la capacidad de los MCGs para reproducir las respuestas a teleconexiones, donde se identifica que los mejores modelos según el *PPI* no se corresponden necesariamente con los que reproducen las teleconexiones históricamente observadas, lo que sugiere que este último criterio complementa nuestro esquema de evaluación de los MCGs.

---

<sup>1</sup> ERIDANUS Ltda. felipeгатeno@eridanus.cl

<sup>2</sup> Departamento de Ingeniería Civil, Universidad de Chile.

<sup>3</sup> Advanced Mining Technology Center (AMTC), Universidad de Chile.

<sup>4</sup> Centro de Ciencia del Clima y la Resiliencia (CR2), Universidad de Chile.

## 1. INTRODUCCIÓN

La cuantificación de impactos asociados al cambio climático es crucial para diseñar estrategias de mitigación y adaptación (IPCC, 2021). En este sentido, los MCGs desempeñan un papel clave para caracterizar cómo podría evolucionar el clima de la Tierra en respuesta a influencias naturales y humanas, siendo ampliamente utilizados para evaluaciones de proyecciones hidrológicas (Wilby y Harris, 2006; Bastola et al., 2011; Vidal et al., 2016). Sin embargo, varios estudios han demostrado que la selección del MCG es una de las decisiones que más contribuye a la dispersión de resultados en la cadena de modelación (e.g., Wilby & Harris, 2006; Bastola et al., 2011; Chen et al., 2011; Vidal et al., 2016; Chegwiddden et al., 2019), destacándose la necesidad de desarrollar estrategias robustas que permitan evaluar y seleccionar modelos climáticos adecuados para reducir la dispersión en proyecciones hidroclimáticas. Un enfoque comúnmente utilizado es el cálculo de métricas que permitan identificar qué modelos representan adecuadamente la climatología observada (Zhang et al., 2018; Rivera & Arnould, 2020; Lun et al., 2021; Di Virgilio et al., 2022). Sin embargo, la aplicación de este enfoque tiene dificultades debido a que el “mejor” subconjunto de modelos depende de las características del dominio de estudio, de las variables analizadas y de las métricas de evaluación aplicadas (Knutti et al., 2010; Notz, 2015; Karmalkar et al., 2019). En cualquier caso, la elección de las métricas es un paso crucial en la clasificación y/o selección de modelos climáticos, con medidas como la media, la desviación estándar, el coeficiente de variación y las estadísticas de tendencia como opciones comunes (por ejemplo, Rupp et al., 2013; Ahmadalipour et al., 2017).

Considerando la necesidad de cuantificar y/o reducir la dispersión de proyecciones climáticas asociada a la selección de MCGs, junto con la gran diversidad hidroclimática que presenta Chile continental, este estudio propone la aplicación de un esquema metodológico flexible para diagnosticar modelos climáticos mediante (i) la evaluación de características climatológicas de la precipitación y temperatura mensual, y (ii) el análisis de la capacidad de reproducir teleconexiones observadas en la zona de estudio. Se introduce el *Past Performance Index (PPI)*, métrica construida para evaluar aspectos de la precipitación y temperatura fundamentales para posteriores aplicaciones en hidrología.

## 2. ZONA DE ESTUDIO

El dominio de estudio consiste en Chile Continental dividido en cinco macrozonas siguiendo el criterio de Alvarez-Garretón et al. (2018) (ver Tabla 1): (1) Norte Grande, desde la región de Arica y Parinacota hasta Antofagasta; (2) Norte Chico, desde Atacama hasta Coquimbo; (3) Zona Central, desde Valparaíso hasta el Maule; (4) Zona Sur, desde Ñuble hasta Los Lagos; y (5) Zona Austral, regiones de Aysén y de Magallanes.

Tabla 1: Características orográficas y climáticas por macrozona (periodo 1979-2020).

Macrozona	Elevación media (máxima) [m.s.n.m.]	Precipitación media anual [mm/año]	Temperatura media anual y extremas [°C]
Norte Grande (Far North)	2531 (6118)	66	11,0 (2,9 - 19,1)
Norte Chico (Near North)	2116 (5980)	117	10,2 (4,0 - 16,5)
Central (Central)	1070 (5575)	833	12,1 (5,8 - 18,3)
Sur (South)	534 (3062)	2057	10,5 (4,9 - 16,1)
Austral (Austral)	524 (3154)	1844	5,0 (0,8 - 0,0)

### 3. DATOS

#### 3.1. PRODUCTO OBSERVACIONAL

Se utilizó el producto grillado CR2METv2.0 (Boisier et al., 2018), el cual contiene series diarias de precipitación (de ahora en adelante Pr), temperaturas mínimas (Tmin) y máximas (Tmax) para todo Chile continental a una resolución horizontal de  $0,05^\circ \times 0,05^\circ$  para el periodo 01/01/1979-31/03/2020. A partir del promedio de Tmin y Tmax, se obtiene una estimación de las temperaturas medias diarias (Tas).

Para evaluar la capacidad de los MCG de simular las teleconexiones que afectan al clima en de Chile continental, se consideran dos patrones de gran escala: (i) El Niño Oscilación del Sur (ENSO) y el (ii) Modo Anular del Sur (SAM). Para El Niño, se considera dos índices en el periodo enero/1979-diciembre/2014: El Niño-1+2<sup>5</sup> y El Niño-3.4<sup>6</sup>, ya que ambos pueden afectar de forma diferente al clima de las distintas macrozonas analizadas. Para los valores observados de SAM<sup>7</sup> se utilizan series temporales mensuales para el periodo enero/1979 a diciembre/2011, correspondientes al periodo disponible.

#### 3.2. MODELOS DE CIRCULACIÓN GENERAL

Utilizamos los resultados de 27 MCGs (Tabla 2) asociados al CMIP6 y seleccionados en función de la disponibilidad de campos diarios de temperatura mínima y máxima (Tmin y Tmax) y precipitación (Pr). Para evaluar el rendimiento histórico de los MCGs, utilizamos un miembro del conjunto (r1i1p1f1) de cada modelo, considerando enero/1979-dic/2014 (es decir, 35 años) como el periodo de evaluación, ya que corresponde a la intersección del periodo con información de CR2METv2.0 y el periodo histórico de los MCGs (definido entre 01/01/1850 y 31/12/2014). Todos los resultados de los MCG se interpolan bilinealmente en una cuadrícula común de  $0,25^\circ \times 0,25^\circ$  para estandarizar la resolución horizontal y facilitar su comparación en términos de rendimiento histórico. Del mismo modo, se escala el conjunto de datos observacionales a la misma resolución horizontal.

Para reconstruir los índices ENSO y SAM a partir de los MCGs, se utilizan las series temporales mensuales de temperatura superficial del mar agregadas espacialmente para las zonas asociadas a los índices El Niño-1+2 y El Niño-3.4, y las series mensuales de presión a nivel del mar a  $40^\circ\text{S}$  y  $65^\circ\text{S}$ .

Tabla 2: Modelos de circulación general evaluados en este estudio. Disponibles en <https://esgf-node.llnl.gov/search/cmip6/>.

ID	Modelo	Institución (País/Región)	Resolución horizontal lat/lon (°)
1	ACCESS-ESM1-5-gn	CSIRO-ARCCSS (Australia)	1,25/1,88
2	BCC-CSM2-MR-gn	BCC (China)	1,13/1,13
3	CanESM5-gn	CCCMA (Canada)	2,80/2,80
4	CESM2-WACCM-gn	NCAR (U.S.A.)	0,94/1,25
5	CMCC-CM2-SR5-gn	CMCC (Italy)	0,94/1,25
6	CMCC-ESM2-gn		0,94/1,25

<sup>5</sup> <https://psl.noaa.gov/data/timeseries/monthly/NINO12/>

<sup>6</sup> <https://psl.noaa.gov/data/timeseries/monthly/NINO34/>

<sup>7</sup> [https://psl.noaa.gov/data/20thC\\_Rean/timeseries/monthly/SAM/](https://psl.noaa.gov/data/20thC_Rean/timeseries/monthly/SAM/)

ID	Modelo	Institución (País/Región)	Resolución horizontal lat/lon (°)
7	EC-Earth3-CC	EC-Earth consortium (Europe)	0,70/0,70
8	EC-Earth3		0,70/0,70
9	EC-Earth3-Veg		0,70/0,70
10	EC-Earth3-Veg-LR		1,13/1,13
11	FGOALS-g3-gn	CAS (China)	2,00/2,00
12	GFDL-CM4-gr1	NOAA-GFDL (U.S.A.)	1,00/1,25
13	GFDL-CM4-gr2		2,00/2,50
14	GFDL-ESM4-gr1		1,00/1,25
15	IITM-ESM-gn	CCCR-IITM (India)	1,90/1,88
16	INM-CM4-8-gr1	INM (Russia)	1,50/2,00
17	INM-CM5-0-gr1		1,50/2,00
18	IPSL-CM6A-LR-gn	IPSL (France)	1,25/2,50
19	KIOST-ESM-gr1	KIOST (Korea)	1,88/1,88
20	MIROC6-gn	JAMSTEC, AORI, NIES, RCCS (Japan)	1,40/1,40
21	MPI-ESM1-2-HR-gn	MPI-M (Germany)	0,94/0,94
22	MPI-ESM1-2-LR-gn		1,87/1,88
23	MRI-ESM2-0-gn	MRI (Japan)	1,12/1,13
24	NESM3-gn	NUIST (China)	1,87/1,88
25	NorESM2-LM-gn	NCC (Norway)	1,90/2,50
26	NorESM2-MM-gn		0,94/1,25
27	TaiESM1-gn	AS-RCEC (Taiwan)	0,94/1,25

## 4. MÉTODOS

Los campos diarios de  $0,25^\circ \times 0,25^\circ$  de precipitación y temperaturas medias de los MCGs y CR2METv2.0 se utilizan para generar series mensuales para el periodo Ene/1979-Dic/2014 en cada píxel de la grilla. Luego, se aplica en cada macrozona una metodología de diagnóstico que consta de tres pasos: (i) evaluación de los MCGs en el periodo histórico en función de cinco aspectos de la climatología; (ii) clasificación de cada MCG en función de su rendimiento histórico, y (iii) caracterización de los impactos del proceso de selección de los MCGs en la propagación de las proyecciones climáticas. Estas etapas se describen detalladamente en las siguientes subsecciones.

### 4.1. CARACTERÍSTICAS CLIMÁTICAS

Se propone una metodología de evaluación de MCGs basado en solo cinco aspectos:

1. **Ciclo estacional:** Se obtienen los promedios mensuales climatológicos del producto observacional y de cada MCG, para luego calcular la correlación de Pearson entre la serie estacional de cada modelo y la asociada a CR2METv2.0 ( $r_{estacional}$ ).
2. **Distribución probabilística de la serie mensual:** Se compara la distribución de probabilidad empírica de datos mensuales simulados (por cada MCG) con la muestra asociada a CR2METv2.0 a través del cálculo del Perkins Skill Score ( $PSS$ , Perkins et al., 2007). El  $PSS$  compara los histogramas de probabilidad simulados con el observado, otorgando  $PSS = 1$  a aquellos modelos que tengan la misma distribución de probabilidad que la obtenida para CR2METv2.0 y  $PSS = 0$  a aquellos que no se ajusten correctamente.
3. **Promedios anuales:** Se calcula la serie anual del producto observacional y de cada MCG. Luego, se calcula la razón entre el promedio de la serie anual de cada modelo ( $\mu_s$ ) y la

asociada a CR2METv2.0 ( $\mu_o$ ), esto es,  $ratio_\mu = \mu_s/\mu_o$ , siendo valores cercanos a 1 indicadores de una correcta reproducción del promedio anual.

4. **Variabilidad interanual:** Se calcula la razón entre la desviación estándar de la serie anual de cada modelo ( $\sigma_s$ ) y la asociada a CR2METv2.0 ( $\sigma_o$ ), esto es,  $ratio_\sigma = \sigma_s/\sigma_o$ . Valores cercanos a 1 indican una correcta reproducción de la variabilidad interanual.
5. **Patrones espaciales:** Para cada píxel en la zona de interés, se calcula el promedio anual para todo el periodo histórico, tanto para CR2METv2.0 como para cada modelo. Luego, se calcula la correlación ( $r_{spatial}$ ) entre el mapa observado (CR2METv2.0) y el simulado por cada modelo.

Estas 5 métricas se calculan por separado para la precipitación y temperatura, para luego calcular el *PPI* de acuerdo a la siguiente ecuación:

$$PPI_{variable,k} = 1 - \sqrt{\sum_{i=1}^{N=5} S_i \cdot (1 - M_i)^2}$$

Donde  $PPI_{variable,k}$  corresponde a la métrica obtenida por el modelo  $k$  obtenido para una variable determinada (precipitación o temperatura),  $M_i$  hace referencia a la métrica  $i$ -ésima obtenida para dicho modelo, y  $S_i$  es el peso que tiene dicha métrica en el cálculo de  $PPI_{variable,j}$ . Se decide modificar el peso de las métricas debido a que algunos aspectos logran ser adecuadamente simulados luego de un proceso de corrección de sesgo de las series simuladas por los MCGs. Para evaluar el efecto de la variación del peso  $S_i$  en la evaluación de los modelos, se decide explorar tres escenarios en los cuales los valores  $S_{ratio_\mu}$  y  $S_{ratio_\sigma}$  (i.e., los pesos asociados a las métricas de diagnóstico de montos y variabilidad interanual, respectivamente) toman valores de 1, 0,5 y 0. Que un modelo obtenga  $PPI \sim 1$  indica que el MCG reproduce adecuadamente los cinco aspectos evaluados, y con el objetivo de poder clasificarlos según su desempeño en esta métrica, se calcula la distancia euclidiana entre los valores obtenidos por cada modelo y el óptimo de la forma  $ED = \sqrt{(1 - PPI_{Pr})^2 + (1 - PPI_{Tas})^2}$ , donde  $ED < 0,5$  se considera como un modelo satisfactorio,  $0,5 < ED < 1,0$  como moderado, y  $ED \geq 1,0$  como insatisfactorio.

#### 4.2. TELECONEXIONES

Dado que las teleconexiones (i.e., una relación estadística entre patrones climáticos en diferentes regiones geográficas) son un factor clave para la variabilidad hidroclimática en nuestro país, extraemos series temporales mensuales de índices climáticos observados (Niño-1+2, Niño-3.4 y SAM) y de precipitación y temperatura mensuales de CR2METv2.0, y las agregamos temporalmente (i) por año hidrológico, (ii) por año calendario, (iii) para las estaciones de invierno (JJA) y (iv) verano (DEF). A continuación se calcula, para cada macrozona y agregación temporal, el coeficiente de correlación de Pearson entre los índices climáticos observados y las series temporales observadas de precipitación y temperatura, junto con sus respectivos intervalos de confianza del 90%, y se selecciona la ventana temporal con valores de correlación superiores a 0,35 para su posterior análisis (Tabla 3). Del mismo modo, se calcula el coeficiente de correlación de Pearson utilizando las series temporales de los índices y variables climáticas derivadas de los MCG, obteniéndose las correlaciones simuladas para cada MCG y macrozona. Así, cada MCG se clasifica en cada macrozona como (i) "Pr y Tas" si el MCG reproduce teleconexiones para ambas variables; (ii) "Only Pr" y (iii) "Only Tas", si sólo reproduce las teleconexiones asociadas a

precipitación o temperatura, respectivamente; o (iv) "Neither" si no reproduce ninguna teleconexión.

Tabla 3: Teleconexiones observadas para cada macrozona.

Zona	Variable	Índice	Agregación temporal	Coef. de correlación de Pearson observado e intervalos de confianza ( $CI_5$ , $CI_{95}$ ) [-]
Norte Grande	Pr	Niño-1+2	Invierno (JJA)	0,36 (0,09-0,58)
	Tas	Niño-3.4	Año hidrológico	0,66 (0,47-0,80)
		Niño-1+2	Verano (DEF)	0,73 (0,57-0,84)
Norte Chico	Pr	Niño-1+2	Año hidrológico	0,63 (0,42-0,78)
		Niño-3.4	Invierno (JJA)	0,56 (0,33-0,73)
	Tas	Niño-3.4	Año hidrológico	0,54 (0,31-0,72)
		Niño-1+2	Verano (DEF)	0,56 (0,33-0,73)
Central	Pr	Niño-1+2	Año hidrológico	0,51 (0,26-0,69)
		Niño-3.4	Año hidrológico	0,59 (0,37-0,75)
	Tas	Niño-3.4	Año hidrológico	0,50 (0,25-0,69)
Sur	Pr	Niño-3.4	Año hidrológico	0,58 (0,36-0,74)
	Tas	Niño-1+2	Año hidrológico	0,36 (0,09-0,58)
Austral	Pr	SAM	Invierno (JJA)	0,40 (0,12-0,62)
	Tas	SAM	Año hidrológico	0,47 (0,20-0,67)

#### 4.3. PROYECCIONES DE CAMBIO

Para probar la efectividad del enfoque de evaluación del modelo en la reducción de la dispersión de las proyecciones climáticas, consideramos simulaciones de los GCMs utilizando el escenario SSP5-8.5, enfocándonos en analizar las (i) series temporales anuales de precipitación y temperatura, y cambios proyectados en (ii) valores medios anuales y (iii) valores climatológicos mensuales. Específicamente, se compara la dispersión en los cambios proyectados obtenidos del conjunto completo de 27 miembros, con la dispersión proporcionada por los cinco GCMs con los valores de  $PPI$  más altos. Con este fin, calculamos los cambios anuales y estacionales para el período 2064-2098 con respecto a 1979-2013. Además, analizamos cómo los pesos para  $ratio_\mu$  y  $ratio_\sigma$  afectan la selección del modelo basada en el  $PPI$ .

Por último, dado que la corrección de sesgo en los MCGs es una práctica común en las evaluaciones de impacto hidrológico (por ejemplo, Wilby y Harris, 2006; Chen et al., 2011; Clark et al., 2016), se aplica el método Quantile Delta Mapping (QDM; Cannon et al., 2015) a cada uno de los 27 modelos estudiados, con el objetivo de comparar la dispersión en los cambios proyectados al excluir el sesgo en los cálculos de  $PPI$  y aplicar la corrección de sesgo al conjunto resultante de MCGs, con la dispersión obtenida a partir de la salida cruda de los modelos.

## 5. RESULTADOS

### 5.1. EVALUACIÓN HISTÓRICA

La Figura 1 muestra las cinco métricas necesarias para calcular el  $PPI$ , para cada variable y macrozona. La evaluación de la precipitación revela que no existe un MCG que se desempeñe satisfactoriamente con una métrica para todas las macrozonas simultáneamente. Por ejemplo, el

ciclo estacional ( $r_{cycle}$ ), está, en general, representado adecuadamente en las zonas Norte Grande, Central y Sur; sin embargo, se desvía del valor óptimo en las zonas Norte Chico y Austral. La correlación espacial ( $r_{spatial}$ ) se reproduce satisfactoriamente en las macrozonas Norte Grande y Austral, aunque los valores simulados se desvían de  $r_{spatial} = 1$  en las zonas Sur y Central. Los valores de  $PSS$  se aproximan a 1 en las macrozonas Central, Sur y Austral, obteniéndose  $PSS > 0,6$  en las macrozonas Norte Grande y Norte Chico. Se obtienen sesgos considerables ( $>2$ ) en las métricas  $ratio_{\mu}$  y el  $ratio_{\sigma}$ , especialmente en las zonas Norte Grande y Norte Chico. Sin embargo, estos sesgos disminuyen hacia el sur de Chile, con una subestimación de la precipitación en las macrozonas Central y Sur. Las métricas de rendimiento para la temperatura son en general mejores que para la precipitación, aunque la variabilidad interanual ( $ratio_{\sigma}$ ) está sobreestimada desde la macrozona Norte Grande hasta la macrozona Sur.

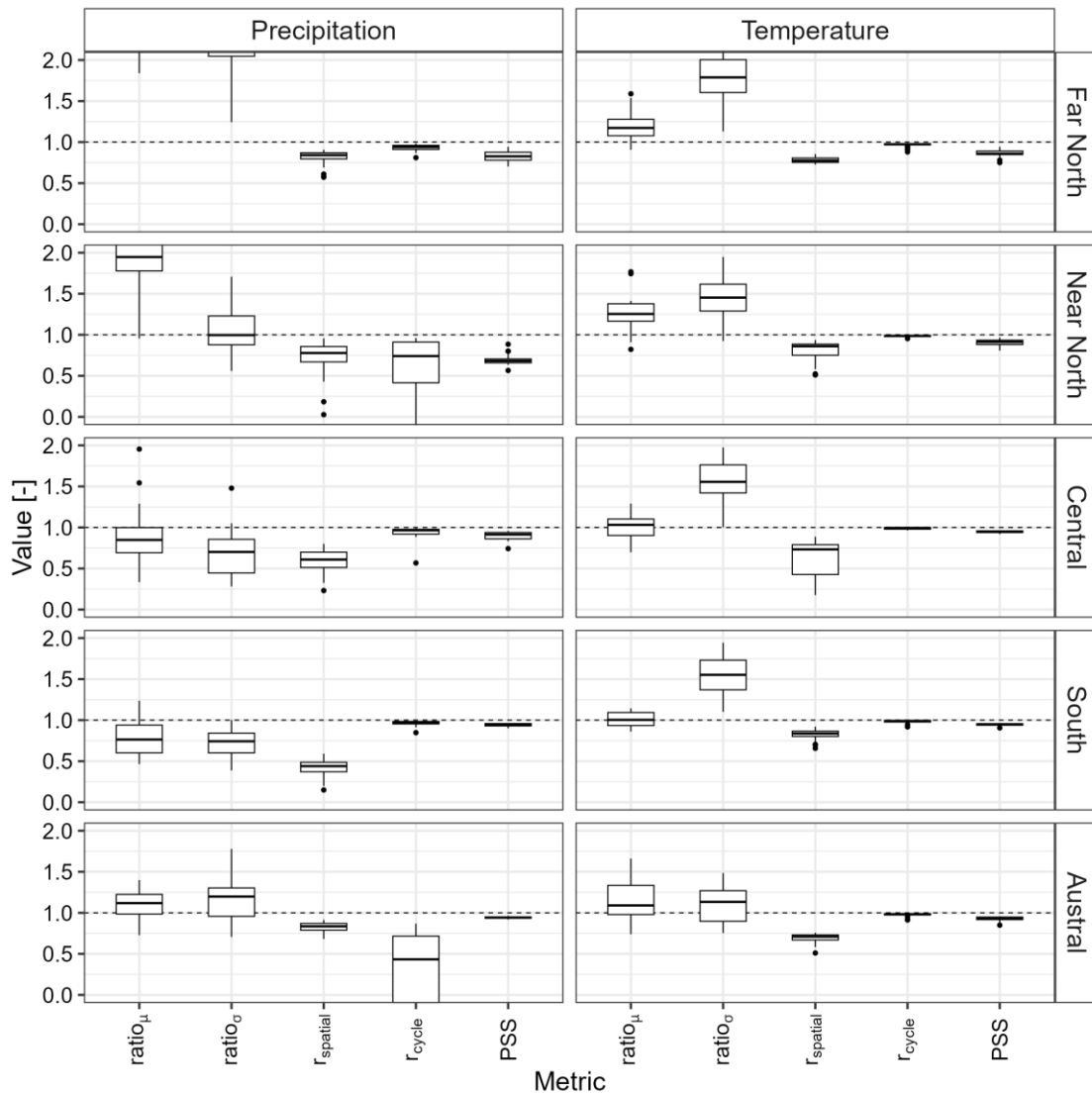


Figura 1: Métricas de evaluación de los modelos, obtenidas para la precipitación (izquierda) y la temperatura (derecha) en cada macrozona (filas). Cada gráfico de caja comprende los resultados de los 27 MCG: los recuadros corresponden al rango intercuartil (RIQ, es decir, los percentiles 25 y 75), la línea horizontal de cada recuadro es la mediana y los bigotes se extienden hasta el  $\pm 1,5$  RIQ del conjunto. El valor ideal de cada métrica está representado por la línea negra discontinua.

La Figura 2 muestra los resultados del *PPI* para cada MCG enumerado en la Tabla 2, junto con la capacidad de simular teleconexiones para la precipitación y/o la temperatura. Los resultados de la Figura 2a (con todos los pesos  $S_k = 1$ ) muestran una transición en el rendimiento de los modelos de norte a sur: a pesar de que ningún MCG se encuentra dentro de las áreas moderada y satisfactoria ( $ED < 1$ ) en el Extremo Norte, el número de modelos en estas categorías aumenta hacia el sur. En efecto, la macrozona Austral es la única en la que algunos MCGs se sitúan dentro de la zona satisfactoria para dicha combinación de pesos  $S_k$ . Solo un MCG muestra un rendimiento al menos moderado en cuatro de las cinco macrozonas (GFDL-CM4-gr1), cinco modelos lo hacen para tres macrozonas (EC-Earth3-CC-gr, EC-Earth3-gr, EC-Earth3-Veg-LR-gr, GFDL-ESM4-gr1 y MPI-ESM1-2-LR-gn), y 10 modelos arrojan un rendimiento moderado en dos macrozonas. En cuanto a las teleconexiones, los MCG muestran un mejor rendimiento en las macrozonas meridionales: en el Norte Grande 17 MCGs cumplen el criterio de teleconexiones para Pr o Tas, y sólo un MCG lo cumple para ambas variables; en la macrozona Norte Chico 14 MCGs cumplen el criterio para una de las variables; mientras que en la macrozona Central (Sur) 13 (10) MCGs representan adecuadamente una sola variable, y 5 (1) modelos consiguen representar teleconexiones para Pr y Tas simultáneamente; en la zona Austral, 18 modelos cumplen el criterio de teleconexión para Pr o Tas, y tres modelos lo superan para ambas variables. Curiosamente, los MCGs que reproducen adecuadamente las teleconexiones se clasifican, en general, como insatisfactorios en función del *PPI*, lo que sugiere que el criterio de teleconexión puede utilizarse de forma complementaria con el enfoque propuesto para diagnosticar la precipitación y la temperatura de los MCGs.

Las Figuras 2b y 2c muestran que, independientemente de la macrozona analizada, todos los MCGs presentan una mejora en su clasificación de resultados cuando se disminuyen los pesos tanto de  $ratio_\mu$  como de  $ratio_\sigma$ . Sin embargo, la mejora más importante se observa en las macrozonas Norte Grande y Norte Chico, donde los MCGs presentan, en general, grandes sesgos en los valores medios anuales: cuando se utiliza  $S_k = 1$  para todas las métricas, ninguna de estas dos macrozonas obtiene siquiera un MCG clasificado como satisfactorio; mientras que cuando  $S_{ratio_\mu} = S_{ratio_\sigma} = 0$ , 23 MCGs son clasificados como satisfactorios en el Norte Grande y 9 lo son en el Norte Chico.



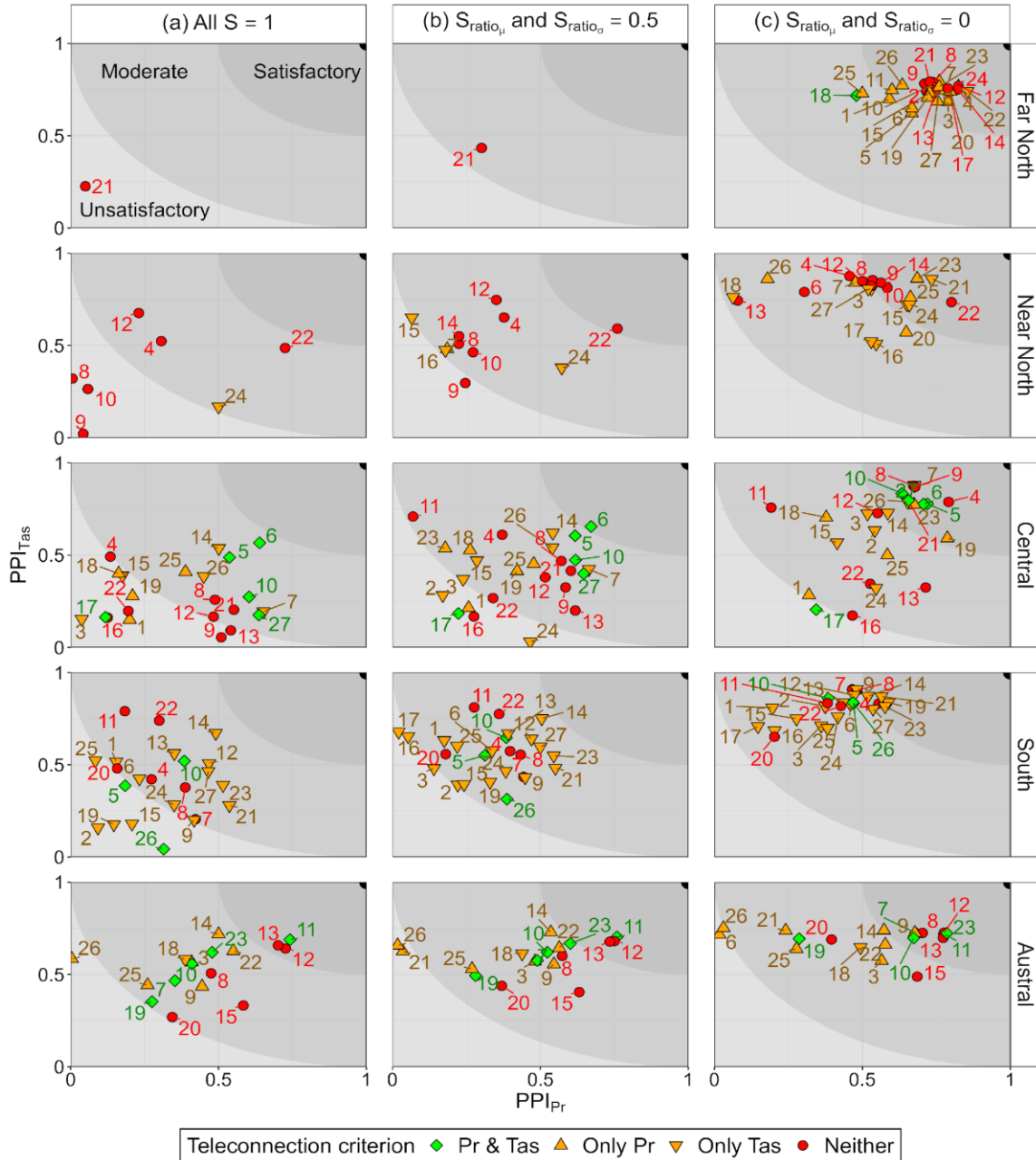


Figura 2: *Past Performance Index* para la precipitación ( $PPI_{Pr}$ , eje x) y la temperatura del aire ( $PPI_{Tas}$ , eje y) en cada macrozona (filas). Los números representan los MCGs listados en la Tabla 2, y los símbolos indican si el modelo reproduce teleconexiones tanto para Pr como para Tas (círculos verdes), sólo una de ellas (triángulos naranjas), o ninguna (diamantes rojos). Las zonas sombreadas indican si el rendimiento del MCG es satisfactorio, moderado o insatisfactorio. Los resultados del  $PPI$  se presentan utilizando (a) el mismo peso para todas las métricas ( $S_k = 1$ ); (b)  $S_k = 0,5$  para  $ratio_{\mu}$  y  $ratio_{\sigma}$ , con  $S_k = 1$  para el resto de las métricas; y (c)  $S_k = 0$  para  $ratio_{\mu}$  y  $ratio_{\sigma}$ , con  $S_k = 1$  para el resto de las métricas.

¿Hasta qué punto es eficaz el  $PPI$  para discriminar los MCGs que no pueden reproducir la estacionalidad histórica? La Figura 3 ilustra las implicancias de una selección basada en el  $PPI$  sobre los ciclos anuales de precipitación simulados considerando (i) todos los MCGs y (ii) los cinco modelos con los valores de  $PPI$  más altos. En general, los cinco mejores MCGs presentan una menor dispersión en los montos mensuales en comparación con el conjunto completo de

modelos. En particular, la reducción del peso de  $ratio_{\mu}$  y  $ratio_{\sigma}$  al calcular el  $PPI$  tiene un impacto insignificante en la dispersión de las medias mensuales de los 5 mejores modelos (compárese las Figuras 3a y 3b), aunque dicha dispersión aumenta durante el periodo estival (diciembre-enero-febrero) en la macrozona Austral. Cuando se excluyen  $ratio_{\mu}$  y  $ratio_{\sigma}$  para calcular el  $PPI$  (Figura 3c), aumenta la dispersión entre los cinco mejores MCGs. Por ejemplo, la dispersión en las simulaciones de precipitación mensual en las macrozonas Norte Chico y Sur durante el invierno (junio-julio-agosto) es comparable a la dispersión del conjunto completo de modelos. Además, el sesgo aumenta durante diciembre-marzo en el Norte Grande y durante mayo-septiembre en la macrozona Central (es decir, durante las estaciones lluviosas) con respecto al producto de referencia, aunque esto es de esperar ya que la estación húmeda contribuye a la mayor parte de la precipitación anual dentro de estas regiones, y el  $ratio_{\mu}$  se excluye de la evaluación.

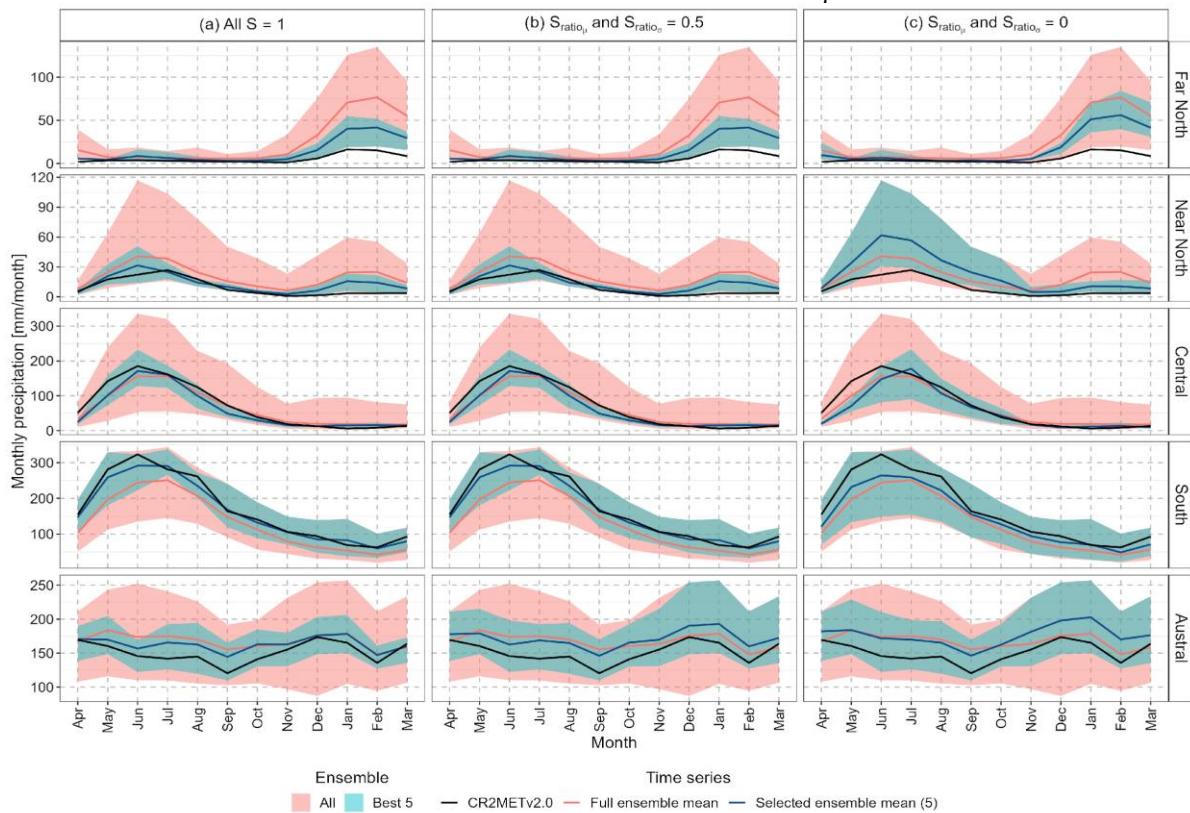


Figura 3: Estacionalidad media de las precipitaciones para el periodo Abr/1979-Mar/2014 obtenida con todos los MCGs (rosa), y los cinco mejores MCGs (turquesa) según los valores más altos del  $PPI$  calculados con las combinaciones de  $S_k$  analizadas en la Figura 2.

## 5.2. PROYECCIONES ANUALES Y ESTACIONALES

Ahora se examina el potencial del esquema de evaluación de MCGs para reducir la dispersión de las proyecciones climáticas anuales y estacionales. Sólo se presentan las proyecciones de precipitación, dada su influencia inmediata en los sistemas hidrológicos. Los resultados de la Figura 4 muestran que, independientemente de los pesos utilizados para calcular el  $PPI$ , los mejores cinco MCGs producen una menor dispersión en comparación con el conjunto completo de modelos, con excepción de la macrozona Sur para  $S_{ratio_{\mu}} = S_{ratio_{\sigma}} = 0$  (Figura 5c). Cuando se aplica la corrección de sesgo, la dispersión proyectada obtenida a partir de los mejores cinco MCGs también es menor que la dispersión del conjunto de 27 miembros en todas las macrozonas.

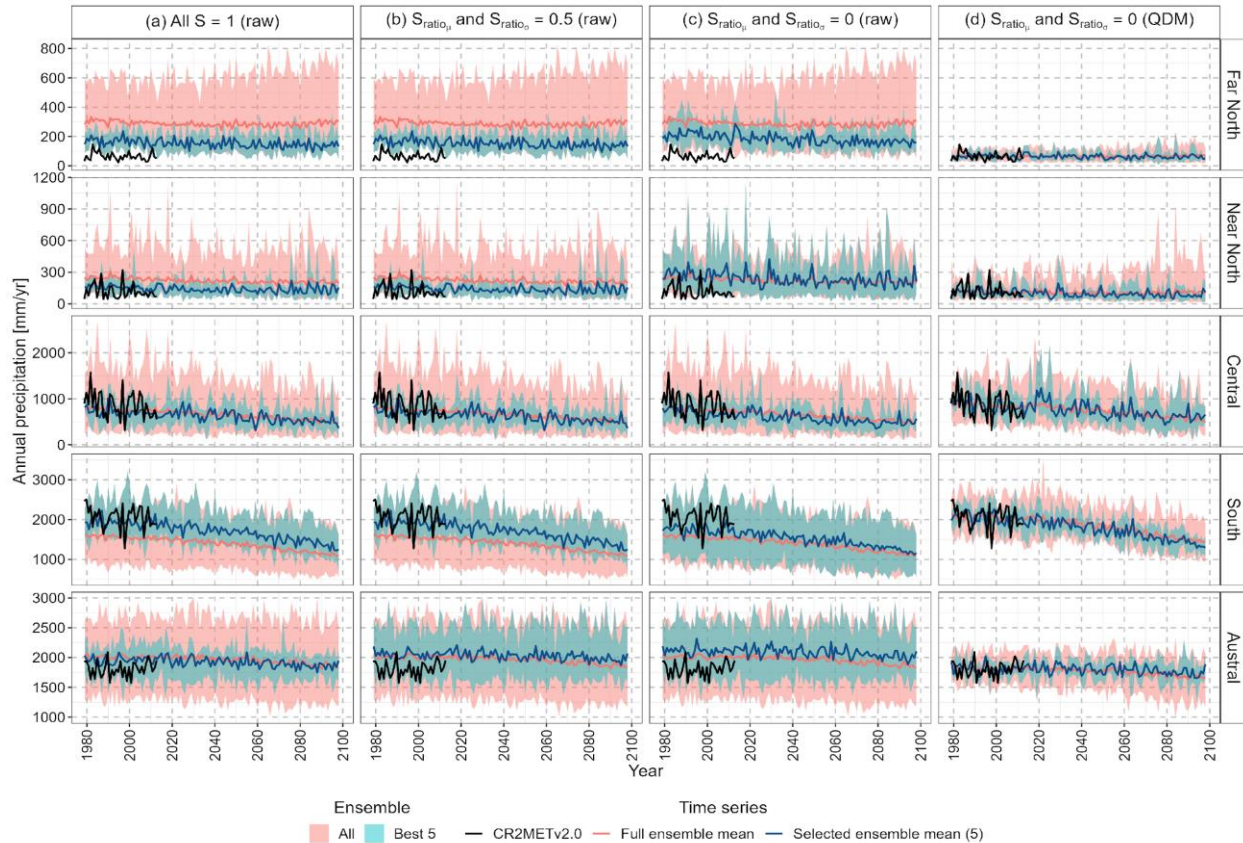


Figura 4: Series temporales CMIP6 de precipitación anual para el periodo 1979-2099 en cada macrozona (filas), donde las observaciones se representan con la línea negra. Las líneas sólidas muestran la media de cada conjunto y la dispersión se representa con el área sombreada: en rosa se presenta el conjunto de los 27 MCGs, mientras que en turquesa se presenta el conjunto conformado por los cinco mejores MCGs según los mejores valores de *PPI* calculados con las combinaciones de  $S_k$  de las figuras anteriores. El panel (d) presenta los mismos conjuntos que el panel (c), pero utilizando los resultados de los modelos corregidos con el método QDM.

La Figura 5 muestra los cambios proyectados en la precipitación media anual y la temperatura. Se puede observar que la estrategia de selección de modelos basada en el *PPI* reduce el rango de cambios proyectados en la precipitación anual. Por ejemplo, la dispersión de los 27 miembros en el Norte Grande varía desde -36% hasta 35%, reduciéndose a -33% y 6% para los cinco MCGs con los valores más altos de *PPI*, de los cuales cuatro (MCGs) proyectan reducciones del 20% en la precipitación anual. De manera similar, el rango del conjunto de 27 miembros en la macrozona Sur varía desde -33% hasta -13%, disminuyendo a -33% hasta -22% después del proceso de selección de MCGs. Se obtienen efectos similares para el resto de las macrozonas, y para la temperatura media anual.

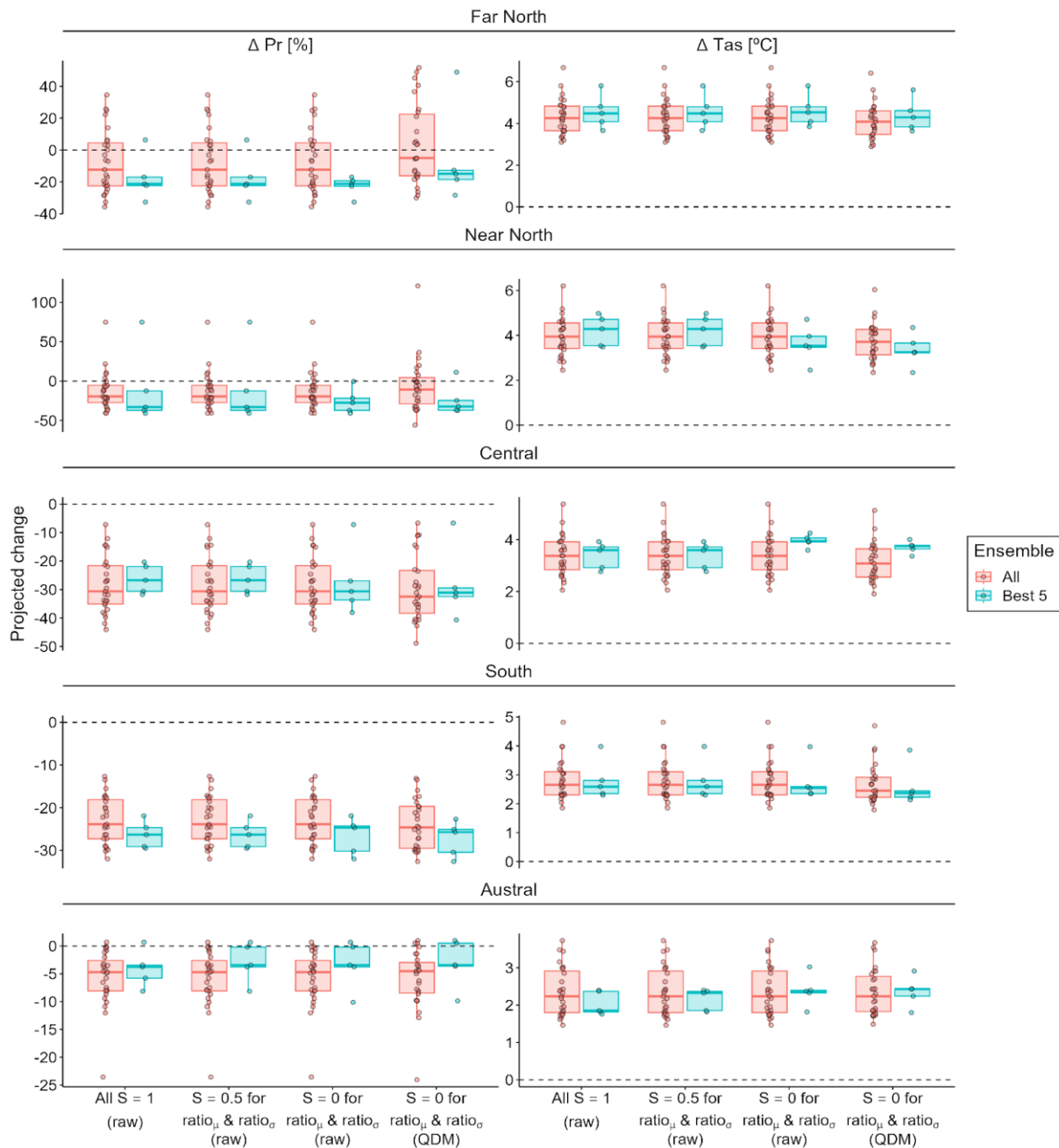


Figura 5: Cambios proyectados en la precipitación media anual (paneles de la izquierda) y la temperatura (paneles de la derecha) para abril/2064-marzo/2098 con respecto al periodo histórico (abril/1979-marzo/2014) en cada macrozona (filas). Los resultados se obtuvieron utilizando todos (es decir, 27) los MCGs (gráficos de caja rosas) y los cinco mejores MCGs (gráficos de caja turquesas) según los valores más altos de  $PPI$  calculados con las mismas combinaciones de  $S_k$  de la Figura 4. Las líneas horizontales discontinuas indican que no hay cambios.

La Figura 6 muestra los cambios proyectados en la precipitación media mensual para todas las macrozonas, utilizando diferentes estrategias de ponderación para calcular el  $PPI$ . Como se encontró para las proyecciones anuales, el enfoque propuesto de evaluación y selección de MCGs reduce el rango en las proyecciones climáticas. Sin embargo, existe un desacuerdo sustancial en el signo de los cambios proyectados para la macrozona Norte Chico, independientemente si se

utilizan los mejores cinco MCGs o el conjunto completo de modelos, aunque la dispersión total se reduce después de seleccionar los modelos basados en el *PPI*. En las demás macrozonas, no se aprecian diferencias en términos de reducción de dispersión entre las estrategias de ponderación para calcular el *PPI*.

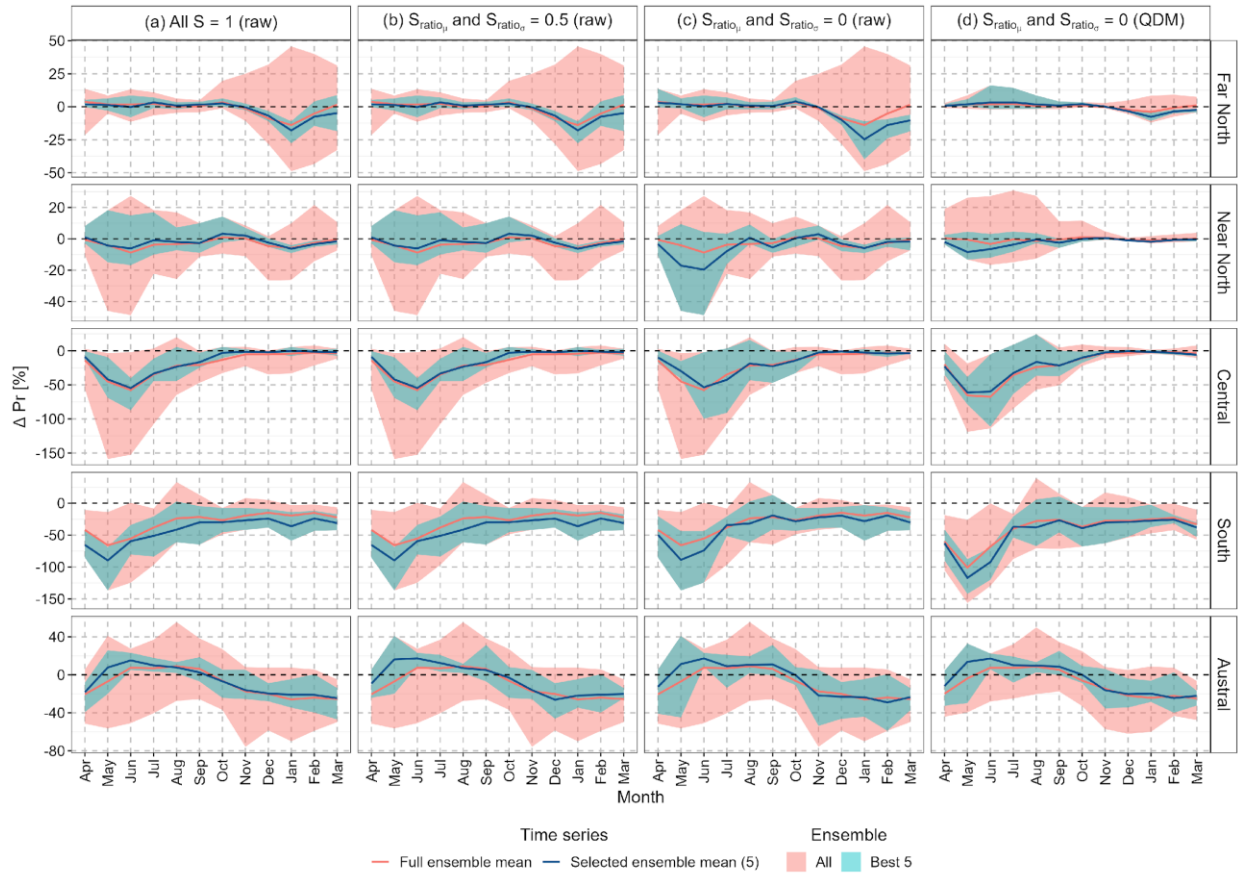


Figura 6: Cambios proyectados en la precipitación media mensual para abril/2064 - marzo/2098 con respecto al periodo histórico (abr/1979 - mar/2014). Las curvas sólidas y los sombreados indican la media del conjunto y la dispersión obtenida con todos los modelos (rosa) y los cinco mejores MCG (turquesa) según los valores más altos de *PPI* calculados con las combinaciones de  $S_k$  de las figuras anteriores.

## 6. CONCLUSIONES

- Se propone un esquema metodológico para diagnosticar el rendimiento histórico de las simulaciones de MCGs del CMIP6 a través de la formulación del *Past Performance Index* (*PPI*), el cual permite incluir variadas métricas de evaluación. Se prueba la efectividad del enfoque para (1) evaluar las simulaciones de precipitación y temperatura mensuales durante 1979-2020 en cinco macrozonas climáticas en Chile, y (2) reducir la dispersión de los cambios proyectados para el período 2064-2098 bajo el escenario SSP5-8.5. También se examina el nivel de acuerdo entre la evaluación según *PPI* y la capacidad de los modelos para replicar teleconexiones históricamente observadas. Nuestras principales conclusiones son las siguientes: el esquema de evaluación y selección propuesto es capaz de discriminar MCGs que no replican la estacionalidad de la precipitación, y también es capaz de reducir

la dispersión en las proyecciones hidroclimáticas anuales y estacionales en todas las macrozonas.

- El conjunto de MCGs con los mejores valores de *PPI* no necesariamente coincide con aquellos que replican teleconexiones, lo que sugiere que ambos criterios podrían utilizarse de manera complementaria.
- Algunas características del modelo que podrían mejorarse mediante la corrección de sesgo (por ejemplo, la media climatológica) podrían excluirse del proceso de evaluación para evitar la eliminación de modelos que replican adecuadamente teleconexiones. Los resultados obtenidos indican que no hay necesariamente una relación entre la magnitud/signo de los cambios proyectados en la temperatura y la precipitación, y la calidad de los MCGs en términos de los valores del *PPI* o la capacidad para replicar teleconexiones observadas.
- Nuestros resultados muestran que existe un conjunto de MCGs que siempre entregan los valores más bajos de *PPI*, independientemente de las ponderaciones de  $ratio_{\mu}$  y  $ratio_{\sigma}$ : CMCC-CM2-SR5 y CMCC-ESM2 en las macrozonas Norte Grande y Norte Chico; INM-CM4-8, INM-CM5-0, MIROC6 y NESM3 en la macrozona Central; e INM-CM4-8, INM-CM5-0 y BCC-CSM2-MR en las macrozonas Sur y Austral. Del mismo modo, encontramos un conjunto de MCGs que siempre presentan un rendimiento satisfactorio para una macrozona determinada: CESM2-WACCM, GFDL-CM4-gr1 y MPI-ESM1-2-LR en el Norte Grande; MPI-ESM1-2-LR y NESM3 en el Norte Chico; CMCC-CM2-SR5 en la macrozona Central; GFDL-CM4-gr1 y MRI-ESM2-0 en el Sur; y GFDL-CM4-gr1 junto con GFDL-CM4-gr2 en la macrozona Austral.

## REFERENCIAS

- Ahmadalipour, A., Rana, A., Moradkhani, H., & Sharma, A. (2017). Multi-criteria evaluation of CMIP5 GCMs for climate change impact analysis. *Theoretical and Applied Climatology*, 128(1–2), 71–87. <https://doi.org/10.1007/s00704-015-1695-4>
- Alvarez-Garretón, C., Mendoza, P. A., Pablo Boisier, J., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., et al. (2018). The CAMELS-CL dataset: Catchment attributes and meteorology for large sample studies-Chile dataset. *Hydrology and Earth System Sciences*, 22(11), 5817–5846. <https://doi.org/10.5194/hess-22-5817-2018>
- Bastola, S., Murphy, C., & Sweeney, J. (2011). The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. *Advances in Water Resources*, 34(5), 562–576. <https://doi.org/10.1016/j.advwatres.2011.01.008>
- Boisier, Juan Pablo, Alvarez-Garretón, C., Cepeda, J., Osses, A., Vásquez, N., & Rondanelli, R. (2018). *CR2MET: A high-resolution precipitation and temperature dataset for hydroclimatic research in Chile*. *Geophysical Research Abstracts*.
- Cannon, A. J., Sobie, S. R., & 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. <https://doi.org/10.1175/JCLI-D-14-00754.1>
- Chegwidden, O. S. S., Nijssen, B., Rupp, D. E. E., Arnold, J. R. R., Clark, M. P. P., Hamman, J. J. J., et al. (2019). How do modeling decisions affect the spread among hydrologic climate change projections? Exploring a large ensemble of simulations across a diversity of hydroclimates. *Earth's Future*, 7(6), 623–637. <https://doi.org/10.1029/2018EF001047>
- Chen, J., Brissette, F. P., Poulin, A., & Leconte, R. (2011). Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resources Research*, 47, W12509. <https://doi.org/10.1029/2011WR010602>

- Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., et al. (2016). Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Current Climate Change Reports*, 2(2), 55–64. <https://doi.org/10.1007/s40641-016-0034-x>
- IPCC. (2021). Assessment Report 6 Climate Change 2021: The Physical Science Basis.
- Karmalkar, A. V., Thibeault, J. M., Bryan, A. M., & Seth, A. (2019). Identifying credible and diverse GCMs for regional climate change studies—case study: Northeastern United States. *Climatic Change*, 154(3–4), 367–386. <https://doi.org/10.1007/s10584-019-02411-y>
- Knutti, R., Furrer, R., Tebaldi, C., Cernak, J., & Meehl, G. A. (2010). Challenges in combining projections from multiple climate models. *Journal of Climate*, 23(10), 2739–2758. <https://doi.org/10.1175/2009JCLI3361.1>
- Lun, Y., Liu, L., Cheng, L., Li, X., Li, H., & Xu, Z. (2021). Assessment of GCMs simulation performance for precipitation and temperature from CMIP5 to CMIP6 over the Tibetan Plateau. *International Journal of Climatology*, 41(7), 3994–4018. <https://doi.org/10.1002/joc.7055>
- Notz, D. (2015). How well must climate models agree with observations? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2052), 20140164. <https://doi.org/10.1098/rsta.2014.0164>
- Perkins, S. E., Pitman, A. J., Holbrook, N. J., & McAneney, J. (2007). Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of Climate*, 20(17), 4356–4376. <https://doi.org/10.1175/JCLI4253.1>
- Rivera, J. A., & Arnould, G. (2020). Evaluation of the ability of CMIP6 models to simulate precipitation over Southwestern South America: Climatic features and long-term trends (1901–2014). *Atmospheric Research*, 241(March), 104953. <https://doi.org/10.1016/j.atmosres.2020.104953>
- Rupp, D. E., Abatzoglou, J. T., Hegewisch, K. C., & Mote, P. W. (2013). Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA. *Journal of Geophysical Research Atmospheres*, 118(19), 10,884–10,906. <https://doi.org/10.1002/jgrd.50843>
- Vidal, J. P., Hingray, B., Magand, C., Sauquet, E., & Ducharne, A. (2016). Hierarchy of climate and hydrological uncertainties in transient low-flow projections. *Hydrology and Earth System Sciences*, 20(9), 3651–3672. <https://doi.org/10.5194/hess-20-3651-2016>
- Di Virgilio, G., Ji, F., Tam, E., Nishant, N., Evans, J. P., Thomas, C., et al. (2022). Selecting CMIP6 GCMs for CORDEX Dynamical Downscaling: Model Performance, Independence, and Climate Change Signals. *Earth's Future*, 10(4), 1–24. <https://doi.org/10.1029/2021EF002625>
- Wilby, R. L., & Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research*, 42(2), W02419. <https://doi.org/10.1029/2005WR004065>
- Zhang, Y., You, Q., Chen, C., Ge, J., & Adnan, M. (2018). Evaluation of downscaled CMIP5 Coupled with VIC model for flash drought simulation in a humid subtropical basin, China. *Journal of Climate*, 31(3), 1075–1090. <https://doi.org/10.1175/JCLI-D-17-0378.1>