

یک چارچوب چندمعیاره (Entropy-TOPSIS) برای انتخاب ترکیب‌های بهینه مدل-روش CMIP6 جهت تصحیح اریبی بارش و دما در شمال شرق ایران

آرش رازبان^۱، محمد موسوی بایگی^{۲*}، ایمان بابائیان^۳، بهنام کامکار^۲ و نسرين صرافزاده^۴

^۱ دانشجوی دکتری، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران

^۲ استاد، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران

^۳ استاد، پژوهشگاه هواشناسی و علوم جو، پژوهشکده اقلیم‌شناسی، مشهد، ایران

^۴ دانش‌آموخته کارشناسی ارشد، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران

(دریافت: ۱۴۰۵/۰۱/۳۱، پذیرش: ۱۴۰۵/۰۴/۳۱)

چکیده

تغییر اقلیم و افزایش اثرات آن، نیاز به بازتولید دقیق بارش و دما در مقیاس ایستگاهی را تشدید کرده است؛ اما خروجی CMIP6 به‌علت تفکیک مکانی پایین و خطاهای سامانمند، نیازمند تصحیح اریبی آماری و انتخاب مدل-روش است. این پژوهش برای رفع خلأ ارزیابی‌های پراکنده و تک‌متغیره در شرق ایران، چارچوبی ایستگاه‌محور و چندمعیاره برای رتبه‌بندی هم‌زمان ترکیب‌های «مدل-روش» در بازتولید بارش، دمای حداقل و دمای حداکثر سه استان خراسان ارائه می‌کند (۱۹ ایستگاه، ۱۹۸۹-۲۰۱۴). در این چارچوب، خروجی‌های ۸ مدل از CMIP6 (۵ مدل ESM و ۳ مدل GCM) با ۸ روش تصحیح اریبی نرم‌افزار CMHyd (۵ روش بارش، ۴ روش دما) مقایسه شدند. سنجه‌ها: MAE، NRMSE، KGE و r؛ و به‌صورت متغیرمحور PBIAS برای بارش و MBE برای دما می‌باشد. معیارها پس از نرمال‌سازی با آنتروپی شانون (وزن ثابت در منطقه) وزن‌دهی و رتبه‌بندی ایستگاهی با TOPSIS انجام شد؛ سپس نتایج با میانه رتبه‌ها، IQR و تعداد ایستگاه برنده تجمیع و با آزمون‌های فریدمن و نمنی اعتبارسنجی شد. اختلاف عملکرد گزینه‌ها برای هر سه متغیر معنادار بود ($\chi^2 \approx 538$, $p \approx 10^{-94}$; Tmin؛ $\chi^2 \approx 458$, $p \approx 10^{-77}$; Tmax؛ Pr: $\chi^2 \approx 286$, $p \approx 10^{-44}$). MPI|VarianceScaling (میانگین رتبه $\approx 3/1$) برای Tmin و MPI|DistributionMapping (میانگین رتبه $\approx 3/2$) برای Tmax بهترین گزینه‌های منطقه‌ای بودند؛ با این حال، آزمون نمنی نشان داد ۱۲ ترکیب (Tmin) و ۱۱ ترکیب (Tmax) از نظر آماری با گزینه برتر قابل تفکیک نیستند و «گروه ممتاز» با شاخص‌های همگرا و پایدار می‌سازند. برای بارش، HadGEM|DistributionMapping (میانگین رتبه ≈ 1) برتر بود و نسبت به سایر اعضای گروه ممتاز شاخص‌های دقیق‌تری ارائه داد. در مقابل، بسیاری از ترکیب‌های مبتنی بر CMCC، INM و GFDL و نیز برخی روش‌های DeltaChange/LinearScaling و LocalIntensity/PowerTransformation فاصله معناداری با گروه ممتاز داشتند و در رتبه‌های انتهایی قرار گرفتند. بنابراین در خراسان، مدل MPI همراه Distribution Mapping و Variance Scaling و HadGEM همراه Distribution Mapping به‌ترتیب پایدارترین گزینه‌ها برای بازتولید دما و بارش هستند. این چارچوب پیشنهادی مبنایی برای کاهش عدم قطعیت و انتخاب اتکاپذیر مدل-روش در پیش‌نگری اقلیم و مدیریت مخاطره منطقه‌ای است.

کلمه‌های کلیدی: CMIP6، Entropy-TOPSIS، بارش و دما، خراسان، تصحیح اریبی آماری

۱ مقدمه

تغییر اقلیم در دهه‌های اخیر به یکی از چالش‌های بنیادین جهانی تبدیل شده و در قالب افزایش دما، اختلال در الگوی بارش، تشدید خشکسالی و افزایش رخداد‌های حدی، پیامدهای گسترده‌ای بر سامانه‌های طبیعی و انسانی بر جای گذاشته است (آی‌پی‌سی‌سی، ۲۰۲۱). این شرایط ضرورت توسعه ابزارهای پیش‌بینی دقیق و قابل اتکا را برای پشتیبانی از مدیریت مخاطره، سیاست‌های انطباقی و برنامه‌ریزی پایدار در مقیاس‌های منطقه‌ای برجسته کرده است (نگوین و همکاران، ۲۰۲۴).

مدل‌های گردش عمومی و سامانه‌های زمین (GCM/ESM) در چارچوب پروژه بین‌المللی CMIP طی دهه‌های اخیر توسعه یافته‌اند و نسخه ششم آن (CMIP6) پیشرفته‌ترین نسل مدل‌های اقلیمی جهانی محسوب می‌شود (آیرینگ و همکاران، ۲۰۱۶). با وجود پیشرفت‌های ساختاری، محدودیت تفکیک مکانی (در حدود ۱۰۰ تا ۲۵۰ کیلومتر) و وجود خطاهای سامانمند، استفاده مستقیم از خروجی این مدل‌ها در تحلیل‌های محلی را با چالش مواجه می‌سازد. افزون بر این، ارزیابی‌ها نشان می‌دهد عملکرد مدل‌های CMIP6 در بازتولید متغیرهای حساس نظیر بارش و دما به شدت وابسته به شرایط اقلیمی و توپوگرافی منطقه است و حتی در مناطق ناهمگون درون یک کشور نیز تفاوت معناداری دارد (کنوتی و همکاران، ۲۰۱۰). این وابستگی مکانی در کشوری با تنوع اقلیمی بالا مانند ایران—از ارتفاعات زاگرس و البرز تا بیابان‌های خشک شرق—بیش از پیش نمایان است.

مطالعات انجام‌شده در ایران و جهان نشان می‌دهد هیچ مدل واحدی از مجموعه‌های CMIP برای همه مناطق و همه متغیرها بهترین عملکرد را ندارد؛ بلکه دقت مدل‌ها تابعی از موقعیت جغرافیایی، نوع متغیر (بارش یا دما) و دوره زمانی است (حمیدیان‌پور و همکاران، ۲۰۱۶)؛

یزدان‌دوست و همکاران، ۲۰۲۰؛ سراج‌ابراهیمی و همکاران، ۲۰۲۴؛ عبدالعلی‌زاده و همکاران، ۲۰۲۳؛ رشیدی‌قانع و همکاران، ۲۰۲۳). این یافته با شواهد تحلیل‌های مکانی-زمانی و واکنش ناهمگن سامانه‌های کشاورزی به تغییرات اقلیمی نیز هم‌راستا است (عشاقی و کامکار، ۲۰۲۵). در نسل ششم مدل‌ها نیز گرچه بهبودهایی گزارش شده، همچنان تفاوت عملکرد میان مدل‌ها و متغیرها پابرجاست.

مطالعات انجام‌شده در ایران نشان داده‌اند که برخی مدل‌های CMIP6 مانند HadGEM3-GC31-LL، MRI-ESM2-0 و MPI-ESM1-2-LR در شبیه‌سازی دما یا بارش در برخی حوضه‌ها عملکرد مطلوبی داشته‌اند، در حالی که مدل‌هایی نظیر INM-CM5-0 و برخی نسخه‌های BCC-CSM2-MR در برخی مناطق نتایج ضعیف‌تری گزارش کرده‌اند (حمیدیان‌پور و همکاران، ۲۰۱۶؛ یزدان‌دوست و همکاران، ۲۰۲۰؛ سراج‌ابراهیمی و همکاران، ۲۰۲۴؛ عبدالعلی‌زاده و همکاران، ۲۰۲۳؛ انصاری میان‌آبادی و همکاران، ۲۰۲۲؛ زارعیان و همکاران، ۲۰۲۲؛ میان‌آبادی و همکاران، ۲۰۲۳). همچنین، روش‌های مبتنی بر نگاشت توزیع و تصحیح چندکی در اغلب مناطق کشور نسبت به روش‌های خطی عملکرد دقیق‌تری داشته‌اند (آزاد واحمدی، ۲۰۲۴؛ رئیسی و همکاران، ۲۰۲۴). با این حال، این نتایج عمدتاً متغیرمحور و منطقه‌محور بوده و کمتر در قالب یک چارچوب مقایسه‌ای یکپارچه و چندمعیاره برای ارزیابی هم‌زمان ترکیب‌های مدل-روش در مقیاس ایستگاهی و منطقه‌ای ارائه شده‌اند.

به‌منظور استفاده عملی از خروجی مدل‌های جهانی در مقیاس‌های منطقه‌ای، تصحیح اریبی—به‌ویژه از نوع آماری—نقش کلیدی ایفا می‌کند. روش‌های آماری به‌دلیل سهولت اجرا، هزینه محاسباتی کمتر و قابلیت تصحیح خطاهای سامانمند، در بسیاری از مطالعات به‌کار

حداکثر — را به صورت هم‌زمان و در چارچوبی نظام‌مند مقایسه کند، همچنان وجود دارد.

پژوهش حاضر در پاسخ به این خلأ طراحی شده است. در این مطالعه، چارچوبی چندمعیاره برای ارزیابی عملکرد مدل‌های منتخب CMIP6 و روش‌های آماری تصحیح اریبی در بازتولید متغیرهای اقلیمی به کار گرفته شده است. ترکیب رتبه‌بندی ایستگاهی، تجمیع منطقه‌ای و اعتبارسنجی آماری امکان سنجش هم‌زمان کارایی نسبی، پایداری مکانی و قابلیت اعتماد ترکیب‌های مدل-روش را فراهم می‌کند. این رویکرد می‌تواند مبنای علمی قابل اتکایی برای مطالعات پیش‌نگری اقلیم و مدیریت مخاطره در شرق ایران ارائه دهد.

۲ روش پژوهش

۲.۱ منطقه مطالعه و داده‌های مشاهداتی

منطقه مورد مطالعه شامل سه استان خراسان رضوی، شمالی و جنوبی در شرق و شمال‌شرق ایران است که در بازه $30/4^\circ$ تا $37/8^\circ$ درجه عرض شمالی و $55/3^\circ$ تا $61/3^\circ$ درجه طول شرقی قرار دارند (بابائیان و همکاران، ۲۰۲۱). با توجه به ویژگی‌های توپوگرافی و اقلیمی، منطقه مورد مطالعه محیطی پیچیده برای مدل‌سازی اقلیم به شمار می‌رود. به همین دلیل در این تحقیق پس از غربال اولیه داده‌ها، اطلاعات ۱۹ ایستگاه سینوپتیک و اقلیم‌شناسی با دوره آماری ۱۹۸۹-۲۰۱۴ (شکل ۱. محدوده مطالعاتی و موقعیت ایستگاه‌های مورد مطالعه. و شکل ۱) انتخاب شد. ایستگاه‌ها به گونه‌ای برگزیده شدند که نماینده تنوع مکانی و ارتفاعی منطقه باشند.

۲.۲ مدل‌های اقلیمی مورد استفاده

در مجموعه مدل‌های CMIP6 دو نوع مدل اقلیمی GCM و ESM وجود دارد که مدل‌های ESM با در نظر گرفتن بازخوردهای سامانه زمین عملکرد جامع‌تری دارند

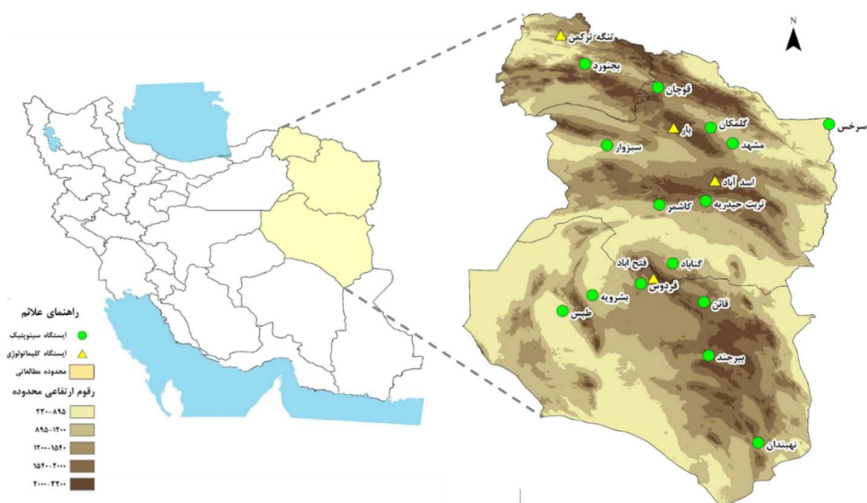
گرفته شده‌اند. نتایج پژوهش‌های داخلی و بین‌المللی نشان می‌دهد روش‌های مبتنی بر نگاشت توزیع، نگاشت چندکی و تبدیل واریانس در بهبود بازتولید بارش و دما عملکرد موثرتری داشته‌اند (ماراون و ویدمن، ۲۰۱۸؛ رئیسی و همکاران، ۲۰۲۴؛ اوزبولدو و ایروم، ۲۰۲۵؛ بروماتی و همکاران، ۲۰۲۴؛ وی و همکاران، ۲۰۲۵؛ لی و لی، ۲۰۲۵). با این حال، کارایی هر روش به اقلیم منطقه، ویژگی‌های ایستگاه‌ها و ساختار خطای مدل پایه وابسته است و هیچ روش واحدی در تمامی شرایط برتر نیست (آزاد و احمدی، ۲۰۲۴).

هم‌زمانی دو چالش «انتخاب مدل اقلیمی مناسب» و «انتخاب روش تصحیح اریبی بهینه» ماهیتی چندمعیاره به مسئله می‌بخشد. در سال‌های اخیر، استفاده از چارچوب‌های تصمیم‌گیری چندمعیاره گسترش یافته است. ترکیب وزن‌دهی عینی مبتنی بر آنالوژی با روش‌هایی نظیر TOPSIS امکان مقایسه کمی گزینه‌ها را تحت معیارهای ناهمگن عملکردی (مانند RMSE، MAE، r و KGE) فراهم کرده و فرآیند انتخاب را از قضاوت ذهنی به چارچوبی داده‌محور و بازتولیدپذیر تبدیل کرده است (هان و همکاران، ۲۰۲۳؛ تاکور و منکار، ۲۰۲۳؛ اگرآوال و موجودمدار، ۲۰۲۵؛ ذبیحی و احمدی، ۲۰۲۴).

منطقه خراسان با دامنه بالای تنوع ارتفاعی، اقلیم خشک تا نیمه‌خشک و نقش کلیدی در امنیت آب و کشاورزی شرق ایران، از مناطق حساس به تغییر اقلیم به‌شمار می‌رود. با وجود مطالعات متعدد در ایران، ارزیابی‌ها غالباً به یک یا دو متغیر محدود بوده، تعداد محدودی مدل یا روش تصحیح اریبی را بررسی کرده‌اند، یا فاقد یک چارچوب یکپارچه چندمعیاره برای رتبه‌بندی هم‌زمان ترکیب‌های مدل-روش در مقیاس منطقه‌ای بوده‌اند. بنابراین، نیاز به ارزیابی ایستگاه‌محور و جامع که بتواند سه متغیر اصلی — بارش، دمای حداقل و دمای

جدول ۱. مشخصات ایستگاه‌های هواشناسی/اقلیم‌شناسی مورد استفاده در مطالعه.

نام ایستگاه	استان	طول جغرافیایی (درجه)	عرض جغرافیایی (درجه)	ارتفاع (متر)	نوع ایستگاه
اسدآباد	خرسان رضوی	۳۵/۶۵	۵۹/۳۵	۱۷۷۰	اقلیم شناسی
بار	خرسان رضوی	۳۶/۴۸	۵۸/۷۰	۱۷۴۶	اقلیم شناسی
بیرجند	خراسان جنوبی	۳۲/۹۰	۵۹/۲۵	۱۴۹۱	سینوپتیک
بجنورد	خراسان شمالی	۳۷/۴۹	۵۷/۳۰	۱۰۶۵	سینوپتیک
بشرویه	خراسان جنوبی	۳۳/۸۵	۵۷/۴۲	۸۷۹	سینوپتیک
فتح آباد- فردوس	خراسان جنوبی	۳۴/۱۲	۵۸/۳۸	۱۸۴۰	اقلیم شناسی
فردوس	خراسان جنوبی	۳۴/۰۳	۵۸/۱۸	۱۲۹۳	سینوپتیک
گلمکان	خرسان رضوی	۳۶/۴۸	۵۹/۲۸	۱۱۷۶	سینوپتیک
گناباد	خرسان رضوی	۳۴/۳۵	۵۸/۶۸	۱۰۵۶	سینوپتیک
کاشمر	خرسان رضوی	۳۵/۲۷	۵۸/۴۷	۱۱۰۹.۷	سینوپتیک
مشهد	خرسان رضوی	۳۶/۲۴	۵۹/۶۳	۹۹۹.۲	سینوپتیک
نهبندان	خراسان جنوبی	۳۱/۵۳	۶۰/۰۳	۱۱۸۸	سینوپتیک
قائن	خراسان جنوبی	۳۳/۷۴	۵۹/۱۸	۱۴۳۲	سینوپتیک
قوچان	خرسان رضوی	۳۷/۱۲	۵۸/۴۵	۱۲۸۷	سینوپتیک
سبزوار	خرسان رضوی	۳۶/۲۱	۵۷/۶۵	۹۶۲	سینوپتیک
سرخس	خرسان رضوی	۳۶/۵۴	۶۱/۱۵	۲۷۸	سینوپتیک
طبس	خراسان جنوبی	۳۳/۶۰	۵۶/۹۵	۷۱۱	سینوپتیک
تنگه ترکمن	خراسان شمالی	۳۷/۹۴	۵۶/۹۲	۹۸۸	اقلیم شناسی
تربت حیدریه	خرسان رضوی	۳۵/۳۳	۵۹/۲۱	۱۴۵۱	سینوپتیک



شکل ۱. محدوده مطالعاتی و موقعیت ایستگاه‌های مورد مطالعه.

(کریشان و همکاران، ۲۰۲۱). داده‌های CMIP6 از دو پایگاه ESGF و Copernicus قابل دریافت‌اند. در این مطالعه، از ۸ مدل اقلیمی (۳ GCM و ۵ ESM) بر اساس دسترسی به داده‌های روزانه بارش، دمای حداقل و حداکثر و عملکرد پیشین در منطقه انتخاب شدند.

جدول ۲. مشخصات مدل‌های اقلیمی مبتنی بر CMIP6 مورد استفاده در مطالعه.

نام مدل	کشور ارائه‌دهنده	سال انتشار	نوع مدل	رزونلوشن افقی	رزونلوشن عمودی
1 ACCESS-ESM1-5	استرالیا	2019	ESM	$1^\circ/25^\circ$	38 لایه تا ارتفاع 40 کیلومتر
2 CMCC-ESM2	ایتالیا	2017	ESM	$1^\circ/25^\circ$	30 لایه تا ۲ هکتوپاسکال
3 CNRM-CM6-1	فرانسه	2017	GCM	$1^\circ/4^\circ$	۹۱ لایه تا ۰/۰۱ هکتوپاسکال
4 GFDL-ESM4	آمریکا	2018	ESM	$1^\circ * 1^\circ$	49 لایه تا 1 هکتوپاسکال
5 HadGEM3-GC31-LL	بریتانیا	2016	GCM	$1^\circ/25^\circ$	۸۵ لایه تا ارتفاع ۸۵ کیلومتر
6 INM-CM5-0	روسیه	2016	GCM	$2^\circ/5^\circ$	۷۳ لایه تا ۵ هکتوپاسکال
7 MPI-ESM1-2-LR	آلمان	2017	ESM	$1^\circ/9^\circ$	۴۷ لایه تا ۰/۰۱ هکتوپاسکال
8 MRI-ESM2-0	ژاپن	2019	ESM	$1^\circ/12^\circ$	80 لایه تا ۰/۰۱ هکتوپاسکال

ماهانه با فرمت ASCII و همچنین داده‌های تاریخی و سناریوهای آینده در قالب فایل‌های NetCDF به صورت برش‌های مکانی متناسب با محدوده جغرافیایی منطقه مورد مطالعه استخراج شده بودند، به عنوان داده‌های ورودی مورد استفاده قرار گرفتند. روش‌های تصحیح

۲.۳ تصحیح اریبی و معیارهای ارزیابی مدل-روش
برای تصحیح سوگیری و انتقال مقیاس به سطح ایستگاه، از نرم‌افزار CMhyd استفاده شد. این نرم‌افزار امکان اجرای روش‌های مختلف تصحیح اریبی آماری را فراهم می‌کند. در این پژوهش، داده‌های دیدبانی در مقیاس

ارایی و مبانی محاسباتی در این نرم افزار به شرح شکل ۳ می باشد. در این مطالعه تمامی روش ها برای هر متغیر اجرا و مقایسه شدند.

همچنین برای مقایسه خروجی مدل های اقلیمی با داده های مشاهداتی در ایستگاه های هواشناسی و انتخاب مدل اقلیمی و روش تصحیح اریبی برتر برای متغیر بارش

از شاخص های آماری جذر میانگین مربعات خطای نرمال شده (NRMSE)، میانگین قدر مطلق خطا (MAE)، ضریب همبستگی پیرسون (r)، شاخص ترکیبی کلینگ-گوپتا (KGE) و (PBIAS) و برای دمای حداقل و دمای حداکثر، از شاخص های آماری MAE، NRMSE، r، KGE و MBE استفاده شد.

جدول ۳. فرآیندها و فرمول های مورد استفاده در روش های تصحیح اریبی نرم افزار CMhyd (بابائیان و همکاران ۲۰۲۳).

نام روش	توضیح روش تصحیح اریبی	منبع	بارش	دما
Linear scaling, multiplicative	نسبت میانگین ماهانه مقادیر دیدبانی به تاریخی مدل در دوره یکسان به سری داده های شبیه سازی شده آینده ضرب می شود. این روش ترجیحاً برای تصحیح اریبی بارش، فشار بخار و تابش استفاده می شود.	مندز و همکاران، ۲۰۲۰	✓	
Linear scaling, additive	تفاوت میانگین ماهانه داده های دیدبانی و تاریخی مدل در دوره یکسان به سری داده های شبیه سازی شده آینده اعمال می شود. این روش ترجیحاً برای تصحیح اریبی دما استفاده می شود			✓
Delta change, additive	تفاوت بین مقادیر متغیر در شبیه سازی تاریخی و آینده به مقادیر دیدبانی افزوده می شود.	دپازنیتو و ویلی، ۲۰۰۵	✓	✓
Delta change, multiplicative	نسبت بین مقادیر متغیر در شبیه سازی تاریخی و آینده در مقادیر دیدبانی ضرب می شود.		✓	
Precipitation local intensity	به طور همزمان شدت و فرکانس بارش تصحیح می شود. ابتدا آستانه شدت بارش هر ماه تعیین می شود، سپس تعداد روزهای تر شبیه سازی فراتر از آستانه یادشده با مقادیر دیدبانی مربوطه تطابق داده می شود.	شمیدلی و همکاران، ۲۰۰۶	✓	
Power transformation	به دلیل ناتوانی روش های انتقال مانند LS و DC در شبیه سازی واریانس، روش غیرخطی به صورت ضریب توانی برای تصحیح واریانس بارش استفاده می شود	لیندر و بوشاند، ۲۰۰۷	✓	
Variance scaling	این روش در سه مرحله میانگین و واریانس دما را تصحیح می کند.	همکاران، ۲۰۱۰	✓	
Distribution mapping	در این روش تابع توزیع داده های شبیه سازی به نحوی تصحیح می شود تا با توزیع دیدبانی مطابق باشد.	پانی و همکاران، ۲۰۱۰	✓	✓

۲.۴ روش ارزیابی عملکرد مدل-روش

در گام نخست مقادیر خام معیارها با استفاده از رابطه (۱) به مقادیر بی واحد در بازه [۰,۱] تبدیل شد؛ به گونه ای که برای تمامی معیارها، مقادیر بزرگ تر نشان دهنده عملکرد بهتر باشند.

ابتدا برای هر پارامتر (بارش و دمای حداقل و حداکثر) ماتریس $X = [x_{ij}]$ تشکیل شد که در آن $i = 1, \dots, m$ تعداد مدل-روش ها برای تمام ایستگاه ها و $j = 1, \dots, n$ تعداد معیارها است. با توجه به ناهمگونی جهت معیارها،

(۱) برای معیارهایی که هر چه بزرگتر باشند بهتر است
برای معیارهایی که هر چه کوچکتر باشند بهتر است

$$x'_{ij} = \begin{cases} (x_{ij} - \min_i x_{ij}) / (\max_i x_{ij} - \min_i x_{ij}) \\ (\max_i x_{ij} - x_{ij}) / (\max_i x_{ij} - \min_i x_{ij}) \end{cases}$$

$$d_j = 1 - E_j \quad (۵)$$

هرچه d_j بزرگتر باشد، معیار j پراکندگی بیشتری دارد و در نتیجه برای تفکیک گزینه‌ها مهم‌تر است. در نهایت، وزن آنتروپی هر معیار با نرمال‌سازی درجه‌های واگرایی از رابطه (۶) به دست آمد:

$$w_j = d_j / (\sum_{j=1}^n d_j) \quad (۶)$$

به این ترتیب، قید $\sum_{j=1}^n w_j = 1$ برقرار و وزن‌ها نشان‌دهنده اهمیت نسبی هر معیار بر اساس اطلاعات درون داده‌ها هستند.

در این پژوهش به منظور اجتناب از تغییرپذیری تصادفی وزن‌ها و حفظ سازگاری مکانی در تحلیل چندایستگاهی، وزن معیارها به صورت مشترک در کل منطقه محاسبه و در تمام ایستگاه‌ها ثابت در نظر گرفته شد. بدین ترتیب، وزن‌های آنتروپی ابتدا به صورت جداگانه برای بارش (Pr)، دمای حداقل (Tmin) و دمای حداکثر (Tmax) با استفاده از کل مجموعه داده‌های مربوط به تمام ایستگاه‌ها و تمام ترکیب‌های مدل-روش تعیین گردیدند. به کارگیری این وزن‌های مشترک در تمام ایستگاه‌ها موجب شد که ساختار تصمیم‌گیری در همه نقاط از منطق یکسانی پیروی کند و مقایسه و تجمیع رتبه‌ها در سطح منطقه‌ای از نظر آماری معتبر باشد. در واقع، در صورتی که وزن‌ها به صورت محلی (ایستگاه‌محور) تعیین می‌شدند، تفاوت در میزان پراکندگی شاخص‌ها در هر ایستگاه می‌توانست موجب تغییر معنی و اهمیت نسبی معیارها بین ایستگاه‌ها گردد و تفسیر فضایی نتایج را مختل کند. بنابراین، با ثابت نگه داشتن وزن‌ها برای هر پارامتر، هم قابلیت مقایسه بین مکانی حفظ شد و هم یکپارچگی در تصمیم‌گیری چندمعیاره تضمین گردید. شایان ذکر است که اگرچه

در گام بعد برای تعیین وزن هر معیار از روش وزن‌دهی آنتروپی شانون استفاده شد. این روش یک روش عینی برای محاسبه وزن معیارهاست که به جای اتکا به قضاوت‌های ذهنی، مستقیماً از پراکندگی داده‌ها در ماتریس تصمیم استفاده می‌کند. ایده‌ی اصلی آن است که اگر مقادیر یک معیار در بین گزینه‌ها تقریباً یکسان باشد، آن معیار اطلاعات چندانی برای تفکیک گزینه‌ها فراهم نمی‌کند (آنتروپی بالا، وزن کم)؛ در مقابل، اگر مقادیر یک معیار در بین گزینه‌ها متنوع و پراکنده باشد، آن معیار در تمایز بین گزینه‌ها نقش مهم‌تری دارد (آنتروپی پایین، وزن بیشتر) (ژو و همکاران، ۲۰۲۰).

برای محاسبه وزن معیارها به روش آنتروپی شانون، از مقادیر نرمال‌شده به عنوان ورودی استفاده شد و برای هر معیار j ، نسبت نسبی هر گزینه i طبق رابطه (۲) محاسبه گردید. این مرحله از نرمال‌سازی مبتنی بر مجموع، صرفاً به منظور برآورده کردن الزامات نظری فرمول آنتروپی و تبدیل مقادیر به توزیع احتمالاتی انجام شده است.

$$p_{ij} = x_{ij} / (\sum_{i=1}^m x_{ij}) \quad (۲)$$

با استفاده از مقادیر نرمال‌شده، آنتروپی اطلاعاتی معیار j به صورت رابطه (۳) تعریف شد:

$$E_j = -k * \sum_{i=1}^m (p_{ij} * \ln(p_{ij})) \quad (۳)$$

که در آن k ثابت نرمال‌سازی است تا مقدار آنتروپی در بازه‌ی $[0,1]$ قرار گیرد و طبق رابطه (۴) محاسبه می‌شود. اگر $p_{ij} = 0$ باشد، جمله‌ی متناظر $p_{ij} * \ln(p_{ij})$ برابر صفر در نظر گرفته می‌شود.

$$k = 1 / \ln(m) \quad (۴)$$

برای تبدیل آنتروپی به معیاری از «اطلاعات مفید» هر معیار، درجه‌ی واگرایی (یا درجه‌ی تنوع) با استفاده از رابطه (۵) محاسبه شد:

این رویکرد یکپارچه، تغییرات فضایی وزن‌ها را نادیده می‌گیرد، اما اثر تفاوت‌های اقلیمی (نظیر تفاوت رفتار مدل‌ها در مناطق خشک جنوب در مقایسه با مناطق مرتفع شمال منطقه) از طریق مقادیر خود سنج‌های عملکردی مستقیماً در ماتریس تصمیم‌گیری در مرحله بعد (رتبه‌بندی تاپسیس) وارد و اعمال شده است.

پس از محاسبه‌ی وزن معیارها، روش تاپسیس برای تجمع وزن‌ها و مقادیر معیارها و به‌دست‌آوردن رتبه‌ی نهایی گزینه‌ها در سطح هر ایستگاه به کار گرفته شد. این روش از پرکاربردترین روش‌های تصمیم‌گیری چندمعیاره محسوب شده که برای نخستین بار در سال ۲۰۱۵ توسط راجو و کومار و در سال‌های اخیر توسط لی و همکاران برای اولویت‌بندی GCM‌ها به کار گرفته شد (هان و همکاران ۲۰۲۳). ایده‌ی اصلی تاپسیس این است که گزینه‌ی مطلوب، کمترین فاصله را از «راه‌حل ایده‌آل مثبت» و بیشترین فاصله را از «راه‌حل ایده‌آل منفی» داشته باشد؛ به عبارت دیگر، گزینه‌ی برتر در عین نزدیکی به بهترین مقادیر ممکن معیارها، از بدترین مقادیر ممکن نیز دور است.

برای رتبه‌بندی مدل-روش‌ها در هر ایستگاه برای هر پارامتر (بارش و دمای حداقل و حداکثر) ماتریس $R = [r_{ij}]$ تشکیل شد که در آن $i = 1, \dots, m$ تعداد مدل-روش‌ها در هر ایستگاه و $j = 1, \dots, n$ تعداد معیارها است. برای هماهنگی جهت معیارها در سطح ایستگاه نیز نرمالسازی با استفاده از رابطه ۱ انجام شد. در گام بعد، برای وارد کردن اهمیت نسبی معیارها در تحلیل، از وزن‌های به‌دست‌آمده از روش آنترپی شانون استفاده شد. ماتریس نرمال وزن‌دار به صورت رابطه (۷) تعریف می‌شود:

$$v_{ij} = w_j * r_{ij} \quad (۷)$$

که در آن w_j وزن معیار j به دست‌آمده از روش آنترپی شانون و $V = [v_{ij}]$ ماتریس نرمال وزن‌دار که عملکرد مدل‌ها را با در نظر گرفتن اهمیت نسبی معیارها نمایش می‌دهد.

با توجه به این که در مرحله‌ی قبل، تمام معیارها به شاخص‌های سودی (هرچه بیشتر، بهتر) تبدیل شده‌اند، راه‌حل ایده‌آل مثبت و منفی به صورت ساده و طبق رابطه (۸) و (۹) تعریف می‌شود:

$$v_j^+ = \max_i(v_{ij}) \quad \text{راه‌حل ایده‌آل مثبت (بهترین وضعیت ممکن برای هر معیار):} \quad (۸)$$

$$v_j^- = \min_i(v_{ij}) \quad \text{راه‌حل ایده‌آل منفی (بدترین وضعیت ممکن برای هر معیار):} \quad (۹)$$

مقدار بزرگ‌تر C_i نشان‌دهنده‌ی عملکرد بهتر مدل است، زیرا در این حالت مدل به راه‌حل ایده‌آل مثبت نزدیک‌تر و از راه‌حل ایده‌آل منفی دورتر است. در نهایت، مدل‌های اقلیمی بر اساس مقادیر C_i به صورت نزولی مرتب شدند و این ترتیب به عنوان رتبه‌بندی نهایی مدل‌ها در هر ایستگاه در نظر گرفته شد.

در انتخاب سنج‌های ارزیابی، اگرچه ضریب همبستگی (r) و کارایی کلینگ-گوپتا (KGE) به دلیل ماهیت ساختاری خود دارای مقداری همپوشانی

برای هر مدل i ، فاصله اقلیدسی آن از راه‌حل ایده‌آل مثبت و منفی با استفاده از رابطه (۱۰) و (۱۱) محاسبه شد.

$$S_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2} \quad (۱۰)$$

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2} \quad (۱۱)$$

در این رابطه‌ها، S_i^+ فاصله مدل i از وضعیت کاملاً مطلوب و S_i^- فاصله مدل i از وضعیت کاملاً نامطلوب است. در گام پایانی، برای هر مدل i ضریب نزدیکی به راه‌حل ایده‌آل مثبت با استفاده از رابطه (۱۲) محاسبه شد:

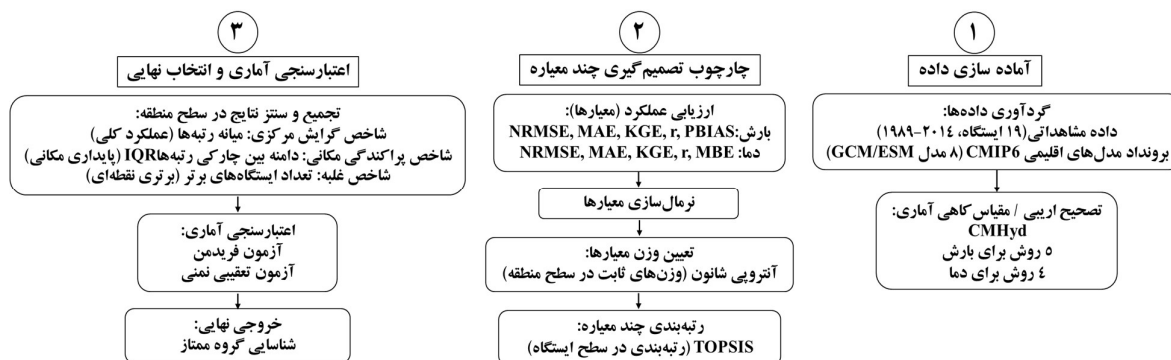
$$C_i = S_i^- / (S_i^+ + S_i^-), \quad 0 \leq C_i \leq 1 \quad (۱۲)$$

اطلاعاتی هستند، اما هر یک ابعاد متفاوتی از رفتار مدل (به ترتیب، همبستگی فاز و رفتار ترکیبی شامل اریبی و تغییرپذیری) را پوشش می‌دهند. انتخاب روش آنروپی شانون برای تعیین وزن سنج‌ها، با تکیه بر قابلیت این روش در ارزیابی محتوای اطلاعاتی شاخص‌ها صورت گرفته است. این رویکرد داده‌محور با تعدیل خودکار وزن سنج‌هایی که دارای اطلاعات موازی یا همبستگی درونی بالایی هستند، از وزن‌دهی مضاعف در فرآیند رتبه‌بندی نهایی به روش تاپسیس جلوگیری می‌کند.

در گام بعدی، به‌منظور دستیابی به یک ارزیابی منطقه‌ای از عملکرد مدل‌ها، در عین حفظ اطلاعات مکانی ایستگاه‌ها، رتبه‌های به‌دست‌آمده برای هر ترکیب مدل-روش در سطح ایستگاهی تجمیع شدند. بدین منظور، سه شاخص مکمل برای توصیف ابعاد مختلف عملکرد منطقه‌ای هر مدل-روش تعریف و محاسبه گردید. در گام نخست، میانه رتبه‌ها به‌عنوان شاخصی از عملکرد مرکزی هر مدل-روش در کل منطقه مورد استفاده قرار گرفت. این شاخص بیانگر جایگاه کلی مدل-روش در میان تمامی ایستگاه‌ها بوده و مقادیر کمتر آن نشان‌دهنده عملکرد بهتر در مقیاس منطقه‌ای است. در گام دوم، فاصله بین چارکی رتبه‌ها به‌منظور ارزیابی پایداری مکانی عملکرد مدل-روش‌ها محاسبه شد. این

شاخص میزان پراکندگی رتبه‌های یک مدل-روش در بین ایستگاه‌های مختلف را نشان می‌دهد؛ به‌گونه‌ای که مقادیر کمتر بیانگر یکنواختی و پایداری مکانی بیشتر و مقادیر بزرگ‌تر نشان‌دهنده ناپایداری و حساسیت مکانی بالاتر عملکرد مدل-روش است. در نهایت، شاخص تعداد ایستگاه‌های برتر به‌عنوان معیاری از برتری موضعی یا قدرت نقطه‌ای مدل-روش‌ها تعریف شد که تعداد دفعات کسب رتبه نخست در ایستگاه‌های مختلف را نشان می‌دهد. در این شاخص، مقادیر بزرگ‌تر بیانگر توانایی بیشتر مدل-روش در ارائه بهترین عملکرد در برخی ایستگاه‌ها است. بدین ترتیب، عملکرد منطقه‌ای مدل-روش‌ها نه تنها از منظر کیفیت کلی، بلکه از نظر پایداری مکانی و میزان برتری موضعی نیز به‌صورت هم‌زمان مورد ارزیابی قرار گرفت.

– آزمون‌های آماری
به‌منظور بررسی معناداری آماری تفاوت بین عملکرد مدل‌ها و روش‌ها، آزمون‌های غیرپارامتری فریدمن برای بررسی فرض برابری میانگین رتبه‌ها بین مدل-روش‌ها در ایستگاه‌های مختلف و آزمون نمنی برای مقایسه‌ی زوجی بین مدل-روش‌ها و شناسایی جفت‌هایی دارای تفاوت معنادار استفاده شدند. خلاصه روش کار در شکل ۲ ارائه شده است.



شکل ۲. چارچوب روش پژوهش.

۳ بحث

در تفسیر نتایج، علاوه بر کارایی نسبی مدل-روش‌ها، محدودیت‌های بالقوه ابزار محاسباتی نیز مدنظر قرار گرفت. گرچه نرم‌افزار CMHyd یکی از ابزارهای متداول در استخراج داده‌ها و اعمال روش‌های تصحیح اریبی در مطالعات اقلیمی و هیدرولوژیکی است، اما همان‌گونه که در برخی مطالعات اخیر (یزدانی و همکاران، ۱۴۰۳) نیز اشاره شده است، امکان بروز خطا در بخشی از محاسبات یا خروجی‌های آن وجود دارد. از این‌رو، در پژوهش حاضر برای کاهش این نگرانی، بخشی از شاخص‌های ارزیابی شامل PBIAS، ضریب همبستگی و NRMSE برای ۳ ایستگاه مشهد، بجنورد و بیرجند به صورت مستقل با محاسبه و با خروجی نرم‌افزار مقایسه شد. نتایج این بررسی نشان داد که اختلاف مشاهده‌شده در حدی نبوده است که بر تفسیر نهایی یا رتبه‌بندی روش‌ها اثر معنادار بگذارد.

- نتایج وزن‌دهی معیارهای خطا به روش آنتروپی نتایج تعیین وزن هر یک از معیارهای خطا به روش آنتروپی، نشان‌دهنده تفاوت ساختاری و معنادار میان سه متغیر اقلیمی بارش، دمای حداقل و دمای حداکثر است. همانطور که شکل ۳ نشان می‌دهد، وزن‌های آنتروپی برای بارش، توزیع تقریباً یکنواختی را در بین معیارهای KGE، NRMSE، MAE و PBIAS نشان می‌دهد ($\sim 0.193-0.206$). این هم‌وزنی بیانگر آن است که عملکرد مدل‌ها در بازتولید بارش وابسته به هیچ معیار منفردی نیست و تمامی جنبه‌های شاخص ترکیبی کلینگ-گوپتا، همبستگی، خطای اندازه و اریبی با اهمیت تقریباً مشابه عمل می‌کنند. چنین الگویی با ماهیت آماری بارش—که دارای نوسانات شدید، رفتار غیرخطی و حساسیت به اریبی و تغییرات نسبی است سازگار بوده و تأکید می‌کند که ارزیابی دقت مدل‌ها در این متغیر باید چندمعیاره و بر پایه هم‌زمان چند شاخص

انجام شود.

در مقابل، دمای حداقل رفتاری کاملاً متفاوت دارد و تأثیر معیارها بسیار نامتوازن است. ضریب همبستگی (r) با وزن 0.337 مهم‌ترین معیار به شمار می‌رود و تقریباً یک سوم وزن کل را تشکیل می‌دهد، در حالی که KGE کمترین اهمیت را دارد (~ 0.103). این ساختار نشان می‌دهد که در بازتولید دمای حداقل، توانایی مدل در بازنمایی الگوی زمانی-مکانی نوسانات اهمیت بیشتری نسبت به بزرگی خطاها دارد. معیارهای NRMSE و MAE نقش مکمل داشته و [MBE] نیز به عنوان شاخص اریبی، وزن متوسطی را داراست. این الگو بیان می‌کند که ترکیب‌های برتر دمای حداقل عمدتاً آن‌هایی هستند که نوسانات و ساختار دینامیکی دما را با دقت بالا (همبستگی قوی) بازتولید کرده‌اند.

برای دمای حداکثر نیز الگوی وزن‌دهی متفاوتی مشاهده می‌شود. در این متغیر، [MBE] با وزن 0.266 مهم‌ترین معیار است و نشان می‌دهد که اریبی سامانمند، مؤلفه اصلی تعیین‌کننده در کیفیت بازتولید دمای حداکثر است. پس از آن معیارهای MAE و r با وزن‌های نسبتاً بالا قرار می‌گیرند و بر اهمیت هم‌زمان دقت خطا و همبستگی تأکید می‌کنند. KGE در دمای حداکثر نیز کمترین وزن را دارد (~ 0.147)، که بیانگر حساسیت پایین‌تر این معیار در تمایز عملکرد مدل‌ها برای این متغیر است. این ساختار منعکس‌کننده ویژگی‌های آماری دمای حداکثر است که معمولاً هموارتر از دمای حداقل بوده؛ اما نسبت به جابه‌جایی میانگین و خطاهای سامانمند حساس‌تر است.

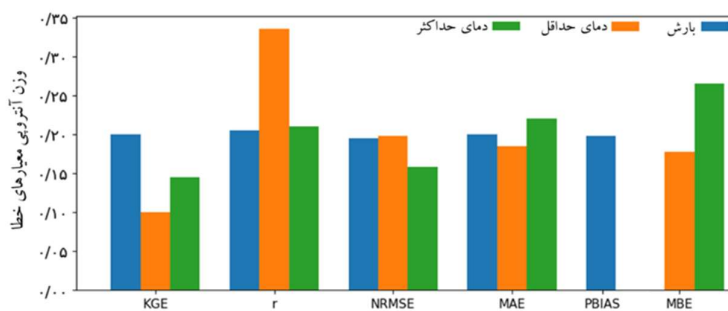
تحلیل سه‌گانه وزن معیارها نشان می‌دهد که هر متغیر اقلیمی، ساختار ارزیابی منحصر به خود را دارد؛ به این ترتیب که در برآورد بارش معیار غالب وجود نداشته و نیازمند ارزیابی جامع است؛ درحالی‌که دمای حداقل نسبت

رتبه‌بندی و درک تفاوت عملکرد مدل-روش‌ها فراهم می‌کند.

– نتایج رتبه‌بندی مدل-روش‌ها در سطح ایستگاه به روش تاپسیس

برای شناسایی بهترین ترکیب مدل-روش تصحیح اریبی در سطح ایستگاهی، رتبه‌بندی همه گزینه‌ها با استفاده از روش تاپسیس انجام شد. در جدول ۴ بهترین ترکیب مدل-روش به ترتیب برای دمای حداکثر، حداقل و بارش ارائه داده شده است. برای دمای حداقل در میان ۱۹ ایستگاه، در ۷ ایستگاه اسدآباد، فتح‌آباد-فردوس،

به ضریب همبستگی پیرسون حساسیت بیشتری داشته و در برآورد دمای حداقل، دقت بازنمایی الگو نسبت به سطح خطا اهمیت بیشتر دارد؛ دمای حداکثر نیز به اریبی حساسیت بیشتری نشان داده است. این تفاوت‌های بنیادین بیان می‌کنند که رتبه‌بندی مدل-روش‌ها نمی‌تواند مبتنی بر یک معیار واحد یا وزن‌دهی ثابت باشد و استفاده از روش‌های تصمیم‌گیری چندمعیاره تنها زمانی معتبر است که وزن معیارها متناسب با رفتار آماری هر متغیر تعیین شود. در نتیجه، تحلیل آنتروپی ضمن شناسایی معیارهای کلیدی هر متغیر، مبنایی آماری برای تفسیر منطقی تر نتایج



شکل ۳. وزن آنتروپی معیارهای ارزیابی عملکرد (تفکیک PBIAS برای بارش و MBE برای دمای حداقل و حداکثر).

بهترین عملکرد را داشت و در ۲ ایستگاه فردوس و گناباد ترکیب HadGEM|LinearScaling بهترین رتبه را به‌دست آورد. در عین حال، روش DistributionMapping در ترکیب با مدل MPI در ۵ ایستگاه بشرویه، فتح‌آباد-فردوس، نهندان، طبس و تنگه ترکمن و ترکیب مدل-روش MPI|VarianceScaling در ۳ ایستگاه بیرجند، کاشمر و قائن رتبه اول را کسب کردند. همچنین در ۲ ایستگاه اسدآباد و مشهد ترکیب DistributionMapping در ۳ ایستگاه طبس توانستند به عنوان روش برتر شناخته شوند.

فردوس، گناباد، مشهد، نهندان و سرخس ترکیب CNRM|DistributionMapping بهترین عملکرد را نشان داد، در حالی که در ۵ ایستگاه بار، بیرجند، کاشمر، قوچان و تربت‌حیدریه ترکیب MPI|VarianceScaling به عنوان گزینه برتر انتخاب شد. همچنین ترکیب‌های مدل ACCESS شامل روش‌های VarianceScaling در ۴ ایستگاه گلمکان، قائن، سبزوار و تنگه ترکمن، LinearScaling در ۲ ایستگاه بجنورد و بشرویه و DistributionMapping در ایستگاه طبس توانستند به عنوان روش برتر شناخته شوند.

بر خلاف دمای حداقل، در متغیر دمای حداکثر، مدل HadGEM نقش برجسته‌تری یافت. در ۷ ایستگاه بار، بجنورد، گلمکان، قوچان، سبزوار، سرخس و تربت‌حیدریه ترکیب HadGEM|VarianceScaling

بارش، رفتاری متفاوت و به مراتب قطبی تر نسبت به دو متغیر دمای حداقل و حداکثر نشان داد. در این متغیر، ترکیب $\text{HadGEM|DistributionMapping}$ به طور مطلق در هر ۱۹ ایستگاه به عنوان بهترین گزینه ظاهر شد و نشان داد که HadGEM در کنار روش $\text{DistributionMapping}$ سازگاری چشمگیری با ساختار بارش منطقه دارد (جدول ۴).

نتایج به روشنی نشان داد که هیچ مدل-روشی در

سطح کل منطقه برای همه متغیرها «بهترین مطلق» نبوده و انتخاب مدل-روش ایده آل به ماهیت متغیر و شرایط محلی ایستگاه وابسته است. همچنین الگوی انتخاب بهترین مدل-روش برای سه متغیر دمای حداقل، دمای حداکثر و بارش کاملاً یکسان نیست و هر متغیر رفتاری متفاوت در مقیاس ایستگاهی ارائه می کند. این تفاوت ها بیانگر نقش برجسته ویژگی های محلی اقلیم و تفاوت مدل های اقلیمی پایه در باز تولید هر متغیر هستند.

جدول ۴. بهترین ترکیب مدل-روش تصحیح اریبی در سطح ایستگاه با استفاده از روش تاسیس.

نام ایستگاه	بهترین مدل-روش	
	دمای حداقل	دمای حداکثر
اسدآباد	$\text{CNRM DistributionMapping}$	$\text{CNRM DistributionMapping}$
بار	$\text{MPI VarianceScaling}$	$\text{HadGEM VarianceScaling}$
بیرجند	$\text{MPI VarianceScaling}$	$\text{MPI VarianceScaling}$
بجنورد	$\text{ACCESS LinearScaling}$	$\text{HadGEM VarianceScaling}$
بشرویه	$\text{ACCESS LinearScaling}$	$\text{MPI DistributionMapping}$
فتح آباد-فردوس	$\text{CNRM DistributionMapping}$	$\text{MPI DistributionMapping}$
فردوس	$\text{CNRM DistributionMapping}$	$\text{HadGEM LinearScaling}$
گلمکان	$\text{ACCESS VarianceScaling}$	$\text{HadGEM VarianceScaling}$
گناباد	$\text{CNRM DistributionMapping}$	$\text{HadGEM LinearScaling}$
کاشمر	$\text{MPI VarianceScaling}$	$\text{MPI VarianceScaling}$
مشهد	$\text{CNRM DistributionMapping}$	$\text{CNRM DistributionMapping}$
نهبندان	$\text{CNRM DistributionMapping}$	$\text{MPI DistributionMapping}$
قائن	$\text{ACCESS VarianceScaling}$	$\text{MPI VarianceScaling}$
قوچان	$\text{MPI VarianceScaling}$	$\text{HadGEM VarianceScaling}$
سبزوار	$\text{ACCESS VarianceScaling}$	$\text{HadGEM VarianceScaling}$
سرخس	$\text{CNRM DistributionMapping}$	$\text{HadGEM VarianceScaling}$
طیس	$\text{ACCESS DistributionMapping}$	$\text{MPI DistributionMapping}$
تنگه ترکمن	$\text{ACCESS VarianceScaling}$	$\text{MPI DistributionMapping}$
تربت حیدریه	$\text{MPI VarianceScaling}$	$\text{HadGEM VarianceScaling}$

- نتایج رتبه بندی مدل-روش ها در سطح منطقه

- دمای حداقل

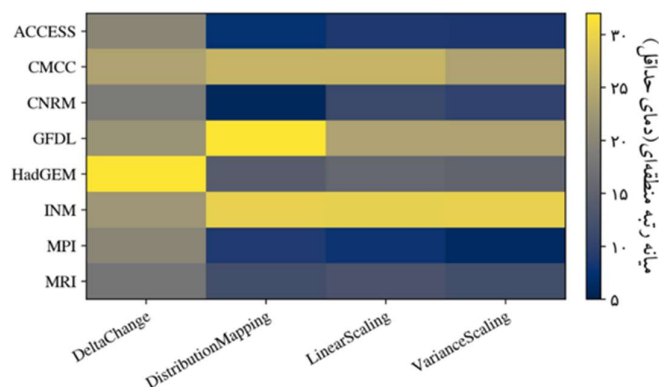
برای ارزیابی عملکرد مدل-روش های تصحیح اریبی دمای حداقل در سطح منطقه، نتایج بر اساس میانه رتبه بندی و با استفاده از نقشه حرارتی ارائه شد (شکل ۴).

همانطور که در شکل ۴، شکل ۵ (نشانگر رتبه بندی نهایی منطقه ای مدل-روش ها برای دمای حداقل) و جدول ۵ مشاهده می شود، ترکیب های مبتنی بر مدل MPI و CNRM با روش های تصحیح اریبی VarianceScaling و $\text{DistributionMapping}$ دارای میانه رتبه پایین تری

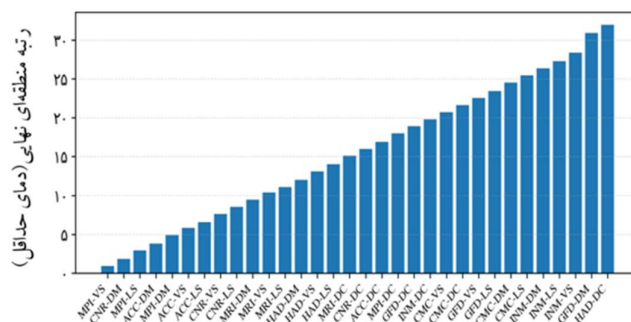
MPI|LinearScaling (Median Rank=5، IQR=3) در طیف عملکرد مناسب قرار می‌گیرند.

در مقابل، مدل‌های CMCC، HadGEM، GFDL و INM در ترکیب با روش DeltaChange در رتبه‌های بالاتری قرار دارند و به‌طور عمومی عملکرد ضعیف‌تری در پیش‌بینی دمای حداقل در منطقه دارند. به‌طور خاص، ترکیب HadGEM|DeltaChange با میانه رتبه ۳۲، IQR برابر ۴ بدترین مدل-روش برای پیش‌بینی دمای حداقل محسوب می‌شوند.

بوده و به‌طور کلی پیش‌بینی‌های بهتری از دمای حداقل داشته‌اند. به‌ویژه، مدل MPI با روش VarianceScaling در رتبه اول قرار دارد، که حاکی از بهترین عملکرد این مدل-روش برای پیش‌بینی دمای حداقل در منطقه است. این مدل با رتبه میانه ۳ و IQR برابر با ۳ و تعداد ایستگاه برنده نشان‌دهنده دقت بالای خود در پیش‌بینی دمای حداقل است. پس از آن ترکیب‌های CNRM|DistributionMapping (Median Rank=2، IQR=5 و Best_per_station =7) و



شکل ۴. ارزیابی عملکرد مدل-روش‌ها برای دمای حداقل در سطح منطقه، با استفاده از نقشه حرارتی میانه رتبه‌بندی.



شکل ۵. رتبه‌بندی نهایی عملکرد مدل-روش‌ها برای دمای حداقل در سطح منطقه.

جدول ۵. رتبه‌بندی مدل-روش‌ها در سطح منطقه برای دمای حداقل.

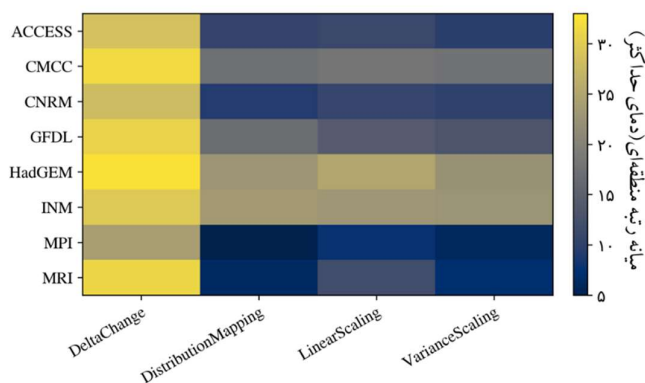
مدل-روش	رتبه منطقه	میان‌رتبه	IQR	تعداد برنده	مدل-روش	رتبه منطقه	میان‌رتبه	IQR	تعداد برنده
MPI VS	۱	۳	۳	۵	CNRM DC	۱۷	۱۸	۶	۰
CNRM DM	۲	۳	۵	۷	ACCESS DC	۱۸	۱۹	۴	۰
MPI LS	۳	۵	۳	۰	MPI DC	۱۹	۱۹	۷	۰
ACCESS DM	۴	۵	۴	۱	GFDL DC	۲۰	۲۰	۵	۰
MPI DM	۵	۶	۳	۰	INM DC	۲۱	۲۱	۲	۰

۰	۴	۲۳	۲۲	CMCC VS	۴	۶	۶	۶	ACCESS VS
۰	۵	۲۳	۲۳	CMCC DC	۲	۷	۶	۷	ACCESS LS
۰	۵	۲۳	۲۳	GFDL VS	۰	۶	۸	۸	CNRM VS
۰	۶	۲۳	۲۵	GFDL LS	۰	۵	۹	۹	CNRM LS
۰	۳	۲۶	۲۶	CMCC DM	۰	۲	۱۰	۱۰	MRI DM
۰	۳	۲۶	۲۶	CMCC LS	۰	۳	۱۰	۱۱	MRI VS
۰	۱	۳۰	۲۸	INM DM	۰	۳	۱۱	۱۲	MRI LS
۰	۲	۳۰	۲۹	INM LS	۰	۰	۱۳	۱۳	HadGEM DM
۰	۲	۳۰	۳۰	INM VS	۰	۲	۱۴	۱۴	HadGEM VS
۰	۲	۳۱	۳۱	GFDL DM	۰	۱	۱۵	۱۵	HadGEM LS
۰	۴	۳۲	۳۲	HadGEM DC	۰	۶	۱۷	۱۶	MRI DC

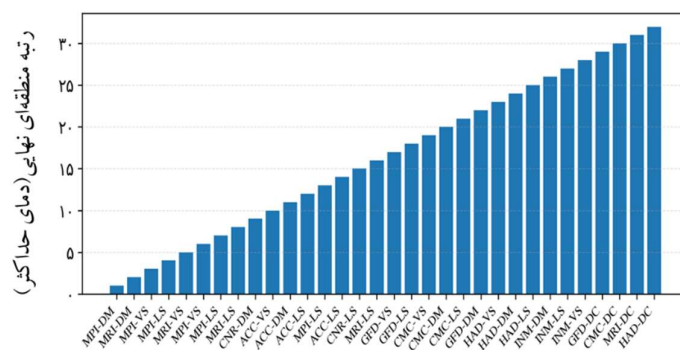
دمای حداکثر

برای دمای حداکثر نیز، ابتدا میانه‌ی رتبه‌ها برای تمام ترکیب‌های مدل-روش محاسبه و در قالب نمودار حرارتی نمایش داده شد. همان‌گونه که در شکل ۴، شکل ۵ (نمایشگر رتبه‌بندی نهایی منطقه‌ای مدل-روش‌ها برای دمای حداکثر) و جدول ۶ مشاهده می‌شود، ترکیب‌های مبتنی بر مدل MPI و MRI، در کنار روش‌های DistributionMapping و VarianceScaling، به‌طور منظم در محدوده‌ی رتبه‌های پایین و رنگ‌های مربوط به عملکرد بهتر قرار گرفته‌اند که نشان‌دهنده‌ی عملکرد نسبتاً پایدار و قابل اتکای این مدل‌ها در روش‌های آماری مختلف تصحیح اریبی دمای حداکثر در منطقه سه استان خراسان هستند. به‌طور مشخص، ترکیب MPI|DistributionMapping با میانه‌ی رتبه‌ی ۳، IQR

برابر ۲/۵ و کسب عنوان بهترین ترکیب در ۵ ایستگاه، مطلوب‌ترین رفتار را در میان تمام ترکیب‌ها نشان داده است. پس از آن، ترکیب‌های (Rank=5 و MPI|VarianceScaling و MRI|DistributionMapping (IQR=3.5, Median Rank=5) در طیف عملکرد مناسب قرار می‌گیرند. در مقابل، ترکیب‌های مبتنی بر مدل‌های INM و HadGEM، در ترکیب با روش DeltaChange عمدتاً در نواحی با رنگ‌های متناظر با رتبه‌های بالا قرار گرفته‌اند؛ برای نمونه، ترکیب HadGEM|DeltaChange دارای میانه‌ی رتبه‌ی ۳۲ و IQR بسیار بزرگ برابر با ۱۱/۵ است که نشان‌دهنده‌ی ضعف و ناپایداری قابل توجه این ترکیب در شبیه‌سازی دمای حداکثر در ایستگاه‌های مختلف می‌باشد.



شکل ۶. ارزیابی عملکرد مدل-روش‌ها برای دمای حداکثر در سطح منطقه، با استفاده از نقشه حرارتی میانه رتبه‌بندی.



شکل ۷. رتبه‌بندی نهایی عملکرد مدل-روش‌ها برای دمای حداکثر در سطح منطقه

جدول ۶. رتبه‌بندی مدل-روش‌ها در سطح منطقه برای دمای حداکثر.

مدل-روش	رتبه منطقه	میان‌رتبه	IQR	تعداد برنده	مدل-روش	رتبه منطقه	میان‌رتبه	IQR	تعداد برنده
MPI DM	۱	۳	۲/۵	۵	CMCC LS	۱۷	۱۸	۳	۰
MRI DM	۲	۵	۳/۵	۰	GFDL DM	۱۷	۱۸	۳	۰
MPI VS	۳	۵	۴	۳	HadGEM VS	۱۹	۲۱	۲۴	۷
MPI LS	۴	۶	۳	۰	INM VS	۲۰	۲۲	۲	۰
MRI VS	۵	۶	۴	۰	HadGEM DM	۲۱	۲۲	۲۳	۰
CNRM DM	۶	۸	۵	۲	INM LS	۲۲	۲۳	۲/۵	۰
ACCESS VS	۷	۹	۲	۰	INM DM	۲۳	۲۳	۵	۰
ACCESS DM	۸	۱۰	۱/۵	۰	MPI DC	۲۴	۲۳	۶/۵	۰
CNRM VS	۹	۱۰	۶	۰	HadGEM LS	۲۵	۲۴	۲۵	۲
ACCESS DC	۱۰	۱۱	۱	۰	CNRM DC	۲۶	۲۷	۵	۰
CNRM LS	۱۱	۱۱	۴/۵	۰	ACCESS DC	۲۷	۲۸	۲/۵	۰
MRI LS	۱۲	۱۲	۳	۰	INM DC	۲۸	۲۹	۳	۰
GFDL VS	۱۳	۱۳	۳	۰	GFDL DC	۲۹	۳۰	۳/۵	۰
GFDL LS	۱۴	۱۵	۲/۵	۰	CMCC DC	۳۰	۳۱	۲	۰
CMCC DM	۱۵	۱۷	۳/۵	۰	MRI DC	۳۱	۳۱	۳	۰
CMCC VS	۱۵	۱۷	۳/۵	۰	HadGEM DC	۳۲	۳۲	۱۱/۵	۰

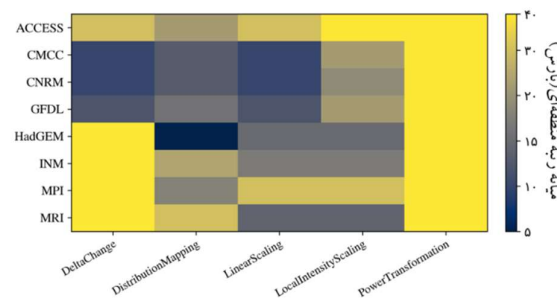
-بارش

برای متغیر بارش نیز، به‌منظور تحلیل جامع رفتار مدل-روش‌ها در مقیاس منطقه‌ای، نقشه حرارتی مبتنی بر میان‌رتبه‌ها برای تمامی ترکیب‌های تصحیح اریبی ترسیم گردید. همان‌گونه که در شکل ۶، شکل ۷ (نمایشگر رتبه‌بندی نهایی منطقه‌ای مدل-روش‌ها برای دمای بارش) و جدول ۷ مشاهده می‌شود، الگوی فضایی رتبه‌ها نشان‌دهنده تفاوت میان ترکیب‌ها است و برخی مدل‌ها در کنار برخی روش‌ها از انسجام و پایداری

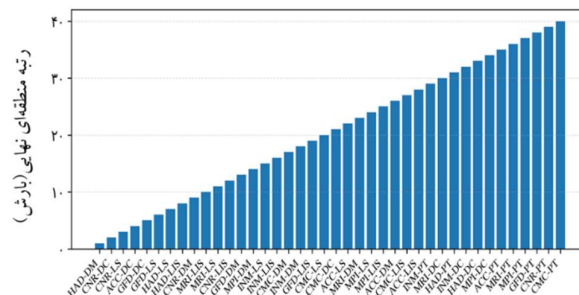
عملکردی بیشتری برخوردارند. در این میان، مدل DisrtibutionMapping در ترکیب با روش جایگاه شاخصی دارد؛ این ترکیب با میان‌رتبه ۱، IQR برابر صفر و کسب عنوان بهترین ترکیب در ۱۹ ایستگاه، نشان‌دهنده ظرفیت بالای این مدل-روش برای بازتولید الگوی بارش منطقه‌ای است و بیانگر هم‌خوانی مناسب‌تر آن با شرایط اقلیمی ایستگاه‌های مختلف نسبت به سایر ترکیب‌ها می‌باشد. پس از آن، ترکیب‌های (IQR=3, Median NRM|DeltaChange

CMCC|PowerTransformation است که با میانه رتبه ۴۰ و IQR برابر ۱، بیانگر ضعف ساختاری و ناپایداری قابل توجه در شبیه‌سازی بارش در ایستگاه‌های مختلف است. این رفتار نشان می‌دهد که مدل‌های مذکور، حتی در ترکیب با روش‌های اصلاح توزیعی، قادر به انطباق مناسب با الگوهای بارش منطقه‌ای نبوده و از این رو برای کاربردهای دقیق اقلیمی توصیه‌پذیر نیستند.

Rank=3) و CNRM|LinearScaling عملکرد قابل‌قبولی را در سطح منطقه ارائه می‌کنند. در مقابل، ترکیب‌های مبتنی بر مدل‌های ACCESS و MPI، خصوصاً در کنار روش‌هایی نظیر PowerTransformation و DeltaChange، عمدتاً در نواحی نقشه حرارتی متناظر با رتبه‌های بالا و نامطلوب دیده می‌شوند. نمونه شاخص این وضعیت، ترکیب



شکل ۸. ارزیابی عملکرد مدل-روش‌ها برای بارش در سطح منطقه، با استفاده از نقشه حرارتی میانه رتبه‌بندی



شکل ۹. رتبه‌بندی نهایی عملکرد مدل-روش‌ها برای بارش در سطح منطقه

جدول ۷. رتبه‌بندی مدل-روش‌ها در سطح منطقه برای دمای بارش

مدل-روش	رتبه منطقه	میان رتبه	I QR	تعداد برنده	مدل-روش	رتبه منطقه	میان رتبه	I QR	تعداد برنده
HadGEM DM	۱	۱	۰	۱۹	CMCC LS	۲۰	۲۱	۶	۰
CNRM DC	۲	۳	۳	۰	ACCESS LS	۲۲	۲۱	۸	۰
CNRM LS	۲	۳	۳	۰	MRI DM	۲۳	۲۲	۵/۵	۰
ACCESS DC	۴	۶	۴	۰	MPI LS	۲۴	۲۲	۷	۰
GFDL DC	۴	۶	۴	۰	MPI LIS	۲۴	۲۲	۷	۰
GFDL LS	۴	۶	۴	۰	ACCESS DM	۲۶	۲۵	۵	۰
HadGEM LS	۷	۶	۱۸	۰	CMCC LIS	۲۷	۲۶	۳	۰

۰	۵/۹	۲۸	۲۸	ACCESS LIS	۰	۱۸	۶	۷	HadGE LIS
۰	۵	۲۹	۲۹	INM PT	۰	۳	۸	۹	CNRM DM
۰	۱	۳۰	۳۰	MRI DC	۰	۵/۵	۱۰	۱۰	MRI LS
۰	۵/۳	۳۰	۳۱	HadGE M PT	۰	۵/۵	۱۰	۱۰	MRI LIS
۰	۱	۳۱	۳۲	INM DC	۰	۶	۱۱	۱۲	CNRM LIS
۰	۳	۳۱	۳۳	HadGE M DC	۰	۵	۱۳	۱۳	GFDL DM
۰	۳	۳۲	۳۴	MPI DC	۰	۴	۱۴	۱۴	MPI DM
۰	۶	۳۳	۳۵	ACCESS PT	۰	۶	۱۵	۱۵	INM LS
۰	۵/۶	۳۳	۳۶	MRI PT	۰	۶	۱۵	۱۵	INM LIS
۰	۵/۱۳	۳۳	۳۷	MPI PT	۰	۵/۶	۱۶	۱۷	CMCC DM
۰	۵/۳۲	۳۴	۳۸	GFDL PT	۰	۵/۶	۱۹	۱۸	INM DM
۰	۵/۳۲	۳۷	۳۹	CNRM PT	۰	۹	۱۹	۱۹	GFDL LIS
۰	۱	۴۰	۴۰	CMCC PT	۰	۶	۲۱	۲۰	CMCC DC

دیگر اختلاف رتبه‌ای کمتر از مقدار اختلاف بحرانی ($CD \approx 10/1$) دارند و از نظر آماری در سطح اطمینان ۹۵٪ از بهترین ترکیب قابل تمایز نیستند که در قالب یک «گروه ممتاز» معرفی می‌شوند (جدول ۸).

بررسی شاخص‌های عملکرد نیز نشان می‌دهد که ترکیب‌های واقع در رأس رتبه‌بندی؛ به‌ویژه MPI|VarianceScaling،

CNRM|DistributionMapping و MPI|LinearScaling علاوه بر برتری آماری، از نظر کیفیت شبیه‌سازی نیز ساختاری هم‌تراز و باثبات ارائه می‌کنند. مقادیر KGE در این گروه در بازه‌ای محدود ($0/67-0/65$) قرار گرفته است و با توجه به پویایی شدید دمای حداقل در مقیاس محلی، بیانگر بازتولید رضایت‌بخش رفتار آماری آن است. همبستگی بسیار بالا در تمام ترکیب‌های ممتاز ($0/96-0/93$) انسجام نتایج

- نتایج ارزیابی عملکرد مدل-روش‌های تصحیح اریبی (آزمون فریدمن و نم‌نی)

-دمای حداقل

برای ارزیابی تفاوت عملکرد مدل-روش‌های تصحیح اریبی در بازتولید دمای حداقل، ابتدا آزمون فریدمن بر روی ۳۲ ترکیب مدل-روش در ۱۹ ایستگاه اجرا شد. مقدار آماری بسیار بزرگ (آزمون $\chi^2 \approx 538$)

($p \approx 10^{-94}$) اختلاف چشمگیر و آماری میان گزینه‌ها را تأیید کرد و نشان داد که اختلاف عملکرد میان گزینه‌ها به‌طور قاطع معنادار است. پس از آن، آزمون نم‌نی برای مقایسه زوجی ترکیب‌ها به‌کار گرفته شد. پایین‌ترین میانگین رتبه مربوط به ترکیب ($3/1$) MPI|VarianceScaling بود که این ترکیب را به‌عنوان بهترین گزینه‌ی منطقه‌ای برای دمای حداقل معرفی می‌کند؛ با این حال، آزمون نم‌نی نشان داد که ۱۲ ترکیب

را تأیید می‌کند و نشان می‌دهد که مدل‌های MPI، CNRM و ACCESS دقیق‌ترین بازتولید را از نوسانات دمای حداقل ارائه کرده‌اند. دامنه تغییرات NRMSE (۰/۴۹-۰/۵۹) و MAE (۳/۴-۳/۸) نیز بسیار محدود است و نشان می‌دهد که اختلاف عملکرد میان اعضای این گروه از نظر خطاهای مقیاس و بزرگی بسیار ناچیز بوده و تمامی آن‌ها رفتار کم‌پراکندگی و قابل اعتماد از خود نشان داده‌اند. همچنین ثابت کامل مقدار $|MBE|$ (۳/۲۵) در سراسر گروه ممتاز، فقدان اریبی سیستماتیک و توانایی این ترکیب‌ها در بازتولید میانگین دمای حداقل بدون جابه‌جایی ساختاری را تأیید می‌کند؛ موضوعی که اهمیت ویژه‌ای در کاربردهای اقلیمی و هیدرولوژیک دارد.

در مقابل، ترکیب‌های مدل GFDL، CMCC، INM و برخی زیر ترکیب‌های HadGEM که مقادیر همبستگی پایین‌تر، خطای بیشتر و میانگین رتبه‌های بالا داشتند، به‌طور معناداری از گروه ممتاز فاصله گرفته و در رتبه‌های انتهایی قرار گرفته‌اند.

-دمای حداکثر

برای ارزیابی تفاوت عملکرد مدل-روش‌های تصحیح اریبی در بازتولید دمای حداکثر نیز، آزمون فریدمن بر

روی ۳۲ ترکیب مدل-روش در ۱۹ ایستگاه اجرا شد. مقدار آماره‌ی فریدمن ($\chi^2 \approx 458$, $p \approx 10^{-77}$) به‌طور قاطع نشان داد که اختلاف عملکرد میان گزینه‌ها در دمای حداکثر نیز معنادار بوده و فرض برابری ترکیب‌ها قابل پذیرش نیست. در ادامه، آزمون نمونی بر پایه میانگین رتبه‌های ایستگاهی اجرا شد و ترکیب MPI|DistributionMapping با میانگین رتبه تقریباً برابر با ۳/۲ بهترین گزینه‌ی منطقه‌ای برای دمای حداکثر شناخته شد. با این حال، نتایج آزمون نمونی نشان داد که ۱۱ ترکیب دیگر اختلاف رتبه‌ای کمتر از مقدار اختلاف بحرانی ($CD \approx 10/1$) داشته و از نظر آماری در سطح اطمینان ۹۵٪ با ترکیب برتر تفاوت معنادار شدیدی ندارند؛ این ترکیب‌ها در کنار ترکیب برتر، یک «گروه ممتاز» را تشکیل می‌دهند (جدول ۹).

بررسی شاخص‌های عملکرد برای گروه ممتاز در دمای حداکثر نشان می‌دهد که ترکیب‌های قرار گرفته در رأس رتبه‌بندی؛ به‌ویژه MPI|DistributionMapping، MRI|DistributionMapping و MPI|VarianceScaling علاوه بر برتری آماری، از نظر کیفیت شبیه‌سازی نیز ساختاری باثبات و هم‌گرا دارند. مقادیر KGE در این گروه در بازه نسبتاً فشرده‌ای قرار

جدول ۷. عملکرد ۱۳ ترکیب (از ۳۲ ترکیب) از مدل-روش‌های تصحیح اریبی در بازتولید دمای حداقل بر روی ایستگاه‌های مورد مطالعه با اختلاف رتبه‌ای کمتر از مقدار بحرانی ($CD \approx 10/1$) و معنادار در سطح اطمینان ۹۵٪ (گروه ممتاز).

مدل-روش	رتبه منطقه	میان رتبه	IQR	تعداد برنده	میان	KGE	r	NRMSE	MAE	MBE
MPI VS	۱	۳	۳	۵	۰/۶۵۳	۰/۹۵۶	۰/۴۹۲	۰/۴۳۵	۳/۲۵۰	۳/۲۵۰
CNRM DM	۲	۳	۵	۷	۰/۶۵۳	۰/۹۵۷	۰/۴۹۲	۳/۴۶۲	۳/۲۵۰	۳/۲۵۰
MPI LS	۳	۵	۳	۰	۰/۶۶۷	۰/۹۵۶	۰/۵۰۹	۳/۴۵۸	۳/۲۵۰	۳/۲۵۰
ACCESS DM	۴	۵	۴	۱	۰/۶۵۳	۰/۹۵۵	۰/۵۰۶	۳/۵۲۴	۳/۲۵۰	۳/۲۵۰
MPI DM	۵	۶	۲/۵	۰	۰/۶۵۳	۰/۹۵۴	۰/۴۹۶	۳/۴۴۸	۳/۲۵۰	۳/۲۵۰
ACCESS VS	۶	۶	۶	۴	۰/۶۵۳	۰/۹۵۷	۰/۴۹۶	۳/۵۰۱	۳/۲۵۰	۳/۲۵۰
ACCESS LS	۷	۶	۶/۵	۲	۰/۶۶۳	۰/۹۵۷	۰/۵۴۷	۳/۵۰۴	۳/۲۵۰	۳/۲۵۰
CNRM VS	۸	۸	۶	۰	۰/۶۵۲	۰/۹۵۲	۰/۵۱۱	۳/۴۶۱	۳/۲۵۰	۳/۲۵۰
CNRM LS	۹	۹	۵	۰	۰/۶۶۳	۰/۹۵۲	۰/۵۲۷	۳/۴۹۰	۳/۲۵۰	۳/۲۵۰
MRI DM	۱۰	۱۰	۲	۰	۰/۶۵۳	۰/۹۵۰	۰/۵۰۲	۳/۵۲۰	۳/۲۵۰	۳/۲۵۰

۳/۲۵۰	۳/۵۳۷	۰/۵۰۸	۰/۹۴۸	۰/۶۵۲	۰	۳	۱۰	۱۱	MRI VS
۳/۲۵۰	۳/۵۳۷	۰/۵۰۹	۰/۹۴۸	۰/۶۶۲	۰	۳	۱۱	۱۲	MRI LS
۳/۱۴۸	۳/۸۴۷	۰/۵۹۹	۰/۹۳۳	۰/۶۵۴	۰	۰	۱۳	۱۳	HadGEM DM

(۴/۴۷-۳/۸۲) MAE و نیز کاهش KGE (تا حدود ۰/۶۸) عملکرد ضعیف‌تری ارائه می‌دهند. این اختلاف نشان می‌دهد که اگرچه همبستگی در این ترکیب‌ها همچنان بالا است، اما خطاهای پراکندگی و بزرگی موجب کاهش توانایی آن‌ها در بازتولید ساختار کامل دمای حداکثر می‌شود.

به‌طور کلی، نتایج نشان می‌دهد که در گروه ممتاز همانند دمای حداقل، ترکیب‌های مبتنی بر مدل MPI و MRI به‌ویژه در روش DistributionMapping و VarianceScaling پایدارترین و دقیق‌ترین بازتولید را از رفتار دمای حداکثر ارائه کرده‌اند، در حالی که روش LinearScaling برای چندین مدل با افت دقت همراه بوده و کارایی کمتری در بازتولید جزئیات آماری دمای حداکثر دارد.

در مقابل با توجه به نتایج آزمون نم‌نی، ترکیب‌های مبتنی بر مدل‌های ACCESS، GFDL، CMCC، INM به‌ویژه در روش DeltaChange فاصله آماری و عملکردی قابل توجهی با این هسته داشته و در رتبه‌های انتهایی قرار گرفته و به طور معنادار دارای عملکرد ضعیفی هستند.

می‌گیرد (۰/۸۳-۰/۸۴) که نشان‌دهنده توانایی مناسب این ترکیب‌ها در بازتولید رفتار کلی دمای حداکثر است. همبستگی بسیار بالا در اغلب ترکیب‌های ممتاز (۰/۹۶-۰/۹۵) انسجام نتایج و دقت مدل‌ها در بازنمایی نوسانات دمای حداکثر را تأیید می‌کند. تحلیل خطا نیز همین الگو را تقویت می‌کند؛ به‌گونه‌ای که مقادیر NRMSE در بخش عمده ترکیب‌های ممتاز در گستره‌ای محدود (۰/۴۳-۰/۴۵) قرار دارد و بیانگر اختلاف اندک میان عملکرد مدل‌ها در بازتولید بزرگی دمای حداکثر است. رفتار MAE نیز دامنه‌ای مشابه (۳/۴۳-۳/۵۵) نشان می‌دهد و حاکی از آن است که تفاوت عملکردی مشاهده‌شده میان ترکیب‌های برتر از نظر خطای مطلق بسیار ناچیز است. همچنین ثبات کامل مقدار $|MBE|$ (۳/۱۹۷) در کل گروه ممتاز نشان می‌دهد که این ترکیب‌ها تقریباً فاقد اریبی بوده و قادرند ساختار میانگین دمای حداکثر را بدون جابه‌جایی سیستماتیک بازتولید کنند.

در مقابل، ترکیب‌های انتهایی به‌ویژه ACCESS|LinearScaling، CNRM|LinearScaling و MRI|LinearScaling افزایش محسوس در (۰/۵۷-۰/۴۸) NRMSE و

جدول ۸. عملکرد ۱۲ ترکیب (از ۳۲ ترکیب) از مدل-روش‌های تصحیح اریبی در بازتولید دمای حداکثر بر روی ایستگاه‌های مورد مطالعه با اختلاف رتبه‌ای کمتر از مقدار بحرانی ($CD \approx 10/1$) و معنادار در سطح اطمینان ۹۵٪ (گروه ممتاز).

مدل-روش	رتبه منطقه	میان رتبه	IQ R	تعداد برنده	میان				
					KGE	r	NRMSE	MAE	MBE
MPI DM	۱	۳	۲/۵	۵	۰/۸۴۰	۰/۹۵۹	۰/۴۳۲	۳/۴۵۷	۳/۱۹۷
MRI DM	۲	۵	۳/۵	۰	۰/۸۳۸	۰/۹۵۴	۰/۴۴۰	۳/۵۱۲	۳/۱۹۷
MPI VS	۳	۵	۴	۳	۰/۸۴۰	۰/۹۵۶	۰/۴۳۰	۳/۴۳۴	۳/۱۹۷
MPI LS	۴	۶	۳	۰	۰/۸۱۸	۰/۹۵۶	۰/۴۴۰	۳/۵۵۱	۳/۱۹۷
MRI VS	۵	۶	۴	۰	۰/۸۳۹	۰/۹۵۵	۰/۴۳۸	۳/۵۴۸	۳/۱۹۷
CNRM DM	۶	۸	۵	۲	۰/۸۳۸	۰/۹۵۴	۰/۴۴۰	۳/۴۶۰	۳/۱۹۷

۳/۱۹۷	۳/۵۸۵	۰/۴۴۹	۰/۹۴۸	۰/۸۳۵	۰	۲	۹	۷	ACCESS VS
۳/۱۹۷	۳/۶۳۶	۰/۴۵۳	۰/۹۴۸	۰/۸۳۴	۰	۱/۵	۱۰	۸	ACCESS DM
۳/۱۹۷	۳/۵۲۸	۰/۴۴۸	۰/۹۵۰	۰/۸۳۸	۰	۶	۱۰	۹	CNRM VS
۳/۱۹۷	۴/۱۱۳	۰/۵۲۲	۰/۹۴۸	۰/۷۵۸	۰	۱	۱۱	۱۰	ACCESS LS
۳/۱۹۷	۳/۸۲۴	۰/۴۸۴	۰/۹۵۰	۰/۷۹۴	۰	۴/۵	۱۱	۱۱	CNRM LS
۳/۱۹۷	۴/۴۷۴	۰/۵۷۴	۰/۹۵۵	۰/۶۸۳	۰	۳	۱۲	۱۲	MRI LS

-بارش

به منظور ارزیابی تفاوت عملکرد مدل‌ها و روش‌های تصحیح اریبی در بازتولید بارش نیز، آزمون غیرپارامتری فریدمن برای ۴۰ ترکیب مدل-روش در ۱۹ ایستگاه اجرا شد. مقدار آماره فریدمن ($\chi^2 \approx 286$, $p \approx 10^{-44}$) بیانگر آن است که در سطح اطمینان ۹۵٪ به طور قاطع فرض برابری عملکرد گزینه‌ها رد و وجود اختلاف معنادار میان مدل-روش‌ها تأیید می‌شود و ضرورت انجام آزمون‌های پساحقی را نشان می‌دهد.

نتایج آزمون نمنی نیز نشان داد کمترین میانگین رتبه مربوط به ترکیب $\text{HadGEM|DistributionMapping} (\approx 1)$ بود و به عنوان بهترین ترکیب منطقه‌ای برای بارش شناخته می‌شود. با این حال، تحلیل نتایج آزمون نمنی نشان داد که ۷ ترکیب دیگر اختلاف رتبه‌ای کمتر از مقدار اختلاف بحرانی ($CD \approx 9/15$) دارند و از نظر آماری در سطح اطمینان ۹۵٪ از بهترین ترکیب قابل تمایز نیستند که در قالب یک «گروه ممتاز» معرفی می‌شوند (جدول ۱۰). این یافته‌ها به وضوح نشان می‌دهد که رفتار بارش نسبت به دما ساختاری متفاوت داشته و وابستگی آن به مدل پایه قوی‌تر و وابستگی آن به ویژگی‌های روش تصحیح اریبی پیچیده‌تر است.

افزون بر نتایج رتبه‌بندی، ارزیابی شاخص‌های عملکردی نشان داد که ترکیب $\text{HadGEM|DistributionMapping}$ پایین‌ترین میانگین رتبه را نیز کسب کرده است؛ از نظر دقت

شبه‌سازی بارش برتری معناداری نسبت به سایر گزینه‌های گروه ممتاز دارد. مقدار KGE این ترکیب ($\approx 0/78$) نشان‌دهنده هماهنگی مناسب میان مقادیر شبه‌سازی و مشاهدات است و مقادیر r ، NRMSE و MAE (به ترتیب ۰/۷۸، ۰/۶۷ و $\approx 7/82$) بیانگر توانایی آن در بازتولید ساختار نوسانات و بزرگای بارش با کمترین میزان خطا هستند. همچنین مقدار بسیار کوچک PBIAS ($\approx 0/32$) بیان می‌کند که این ترکیب تقریباً فاقد اریبی سامانمند بوده و در بازتولید میانگین بارش نیز رفتار پایدار و قابل اتکایی دارد.

در مقابل، سایر ترکیب‌های عضو گروه ممتاز از جمله CNRM|DeltaChange ، $\text{ACCESS|DeltaChange}$ ، $\text{CNRM|LinearScaling}$ و GFDL|DeltaChange —اگرچه از نظر آماری قابل تمایز از ترکیب برتر نیستند، اما در شاخص‌های عملکردی الگوی ضعیف‌تری نشان می‌دهند. کاهش مقادیر KGE و r (عمدتاً در حدود ۰/۴۰–۰/۴۶) و افزایش NRMSE و MAE ($1/1$ – $1/0$ و 13 – 12) حاکی از آن است که این ترکیب‌ها توانایی محدودتری در بازتولید رفتار ساختاری و شدت بارش دارند، هرچند مقادیر بسیار کوچک PBIAS نشان می‌دهد که خطای اریبی در آن‌ها کنترل شده است. به طور خاص، MRI|LinearScaling پایین‌ترین مقادیر ($\approx 0/32$) KGE و ($\approx 0/40$) r را داشته و به عنوان ضعیف‌ترین گزینه در میان ترکیب‌های گزارش شده ظاهر می‌شود.

به صورت کلی، نتایج نشان می‌دهد که اگرچه چندین

مدل-روش از نظر آماری در گروه ممتاز قرار می‌گیرند، فقط HadGEM|DistributionMapping از نظر شاخص‌های خطای ساختاری، پایداری مکانی و نبود اریبی، جایگاه یک ترکیب برتر واقعی را حفظ می‌کند؛ در حالی که سایر ترکیب‌ها گرچه تعادل رتبه‌ای مشابه دارند، اما از منظر دقت بازتولید بارش، رفتار سازگار و قدرتمندی ارائه نمی‌کنند.

در مقابل، با توجه به نتایج آزمون نمونی، ترکیب‌های مدل CMCC، INM و برخی روش‌های

ACCESS و MRI و بسیاری از ترکیب‌های مرتبط با روش‌های PowerTransformation و LocalIntensityScaling که میانگین رتبه آن‌ها اختلافی فراتر از مقدار اختلاف بحرانی با ترکیب HadGEM|DistributionMapping داشته‌اند، در گروه ضعیف‌تر قرار می‌گیرند. اختلاف آماری معنادار این گزینه‌ها با گروه ممتاز نشان‌دهنده ضعف ساختاری آن‌ها در بازتولید بارش در مقیاس منطقه‌ای است.

جدول ۹. عملکرد ۱۰ ترکیب (از ۴۰ ترکیب) از مدل-روش‌های تصحیح اریبی در بازتولید بارش بر روی ایستگاه‌های مورد مطالعه با اختلاف رتبه‌ای کمتر از مقدار بحرانی ($CD=9/15$) و معنادار در سطح اطمینان ۹۵٪ (گروه ممتاز).

مدل-روش	رتبه منطقه	میان رتبه	IQR	تعداد برنده	میان	KGE	r	NRMSE	MAE	PBIAS
HadGEM DM	۱	۱	۰	۱۹	۰/۷۷۵	۰/۷۷۵	۰/۶۷۰	۷/۸۲۴	۰/۳۲۰	
CNRM DC	۲	۳	۳	۰	۰/۴۴۱	۰/۴۵۸	۰/۹۷۹	۱۲/۷۸۴	-۶/۶۰۹۰۲E-۰۷	
CNRM LS	۲	۳	۳	۰	۰/۴۴۰	۰/۴۵۷	۰/۹۹۹	۱۲/۸۸۴	-۶/۹۱E-۰۷	
ACCESS DC	۴	۶	۴	۰	۰/۴۴۸	۰/۴۵۳	۱/۰۲۴	۱۲/۳۰۳	۱/۰۴۴۵۱E-۱۵	
GFDL DC	۴	۶	۴	۰	۰/۴۲۰	۰/۴۳۶	۱/۰۰۶	۱۲/۸۶۷	-۱/۲۲۹۹۷E-۰۶	
GFDL LS	۴	۶	۴	۰	۰/۴۰۰	۰/۴۱۶	۱/۱۰۶	۱۲/۹۶۹	-۱/۳۴E-۰۶	
CNRM DM	۹	۸	۳	۰	۰/۴۵۷	۰/۴۵۷	۱/۰۴۰	۱۳/۴۳۱	-۶/۲۰۵۴۴E-۰۶	
MRI LS	۱۰	۱۰	۵.۵	۰	۰/۳۲۱	۰/۳۹۷	۰/۹۶۴	۱۲/۲۵۰	۱/۱۹۴۶۴E-۱۵	

۴ نتیجه‌گیری

در این پژوهش، عملکرد ۸ مدل اقلیمی از مجموعه مدل‌های مبتنی بر CMIP6 در شبیه‌سازی بارش، دمای حداقل و دمای حداکثر در استان‌های خراسان، با استفاده از روش‌های مختلف تصحیح اریبی آماری مورد ارزیابی قرار گرفت. نتایج به‌دست‌آمده از این مطالعه، به‌ویژه در مقایسه با یافته‌های پیشین، تطابق و تفاوت‌های مهمی را در سطح مدل‌ها و روش‌های تصحیح اریبی نشان می‌دهد که می‌تواند مبنای جدیدی برای انتخاب بهینه مدل‌ها و روش‌ها در مطالعات تغییر اقلیم در ایران فراهم آورد.

• عملکرد مدل‌ها و تفاوت‌های اقلیمی: همانطور که در بسیاری از مطالعات پیشین مشاهده شده است (مانند

حمیدیان‌پور و همکاران، ۲۰۱۶؛ سراج‌ابراهیمی و همکاران، ۲۰۲۴)، عملکرد مدل‌های مبتنی بر CMIP6 به‌ویژه در مناطقی با ناهمگونی اقلیمی بالا، به‌شدت به ویژگی‌های محلی و توپوگرافی منطقه وابسته است. نتایج تحقیق حاضر نیز این الگو را تایید می‌کند، به‌طوری که مدل‌های MPI، MRI و CNRM در ترکیب با روش‌های VarianceScaling و DistributionMapping بهترین نتایج را برای دما و مدل HadGEM در ترکیب با DisrtibutionMapping بهترین نتیجه را برای بارش در استان‌های خراسان به‌دست آوردند. این یافته‌ها با نتایج مشابه در دیگر مطالعات مشابه، مانند سراج‌ابراهیمی و

همکاران (۲۰۲۴)، یزدان‌دوست و همکاران (۲۰۲۱) و آزاد و احمدی (۲۰۲۴) برای دما و انصاری میان‌آبادی و همکاران (۲۰۲۲)، یزدان‌دوست و همکاران (۲۰۲۰) و زارعیان و همکاران (۲۰۲۲) برای بارش همخوانی دارد.

- تصحیح اریبی و روش‌های مناسب: در زمینه تصحیح اریبی، این تحقیق به طور خاص روش‌های VarianceScaling و DistributionMapping را به‌عنوان موثرترین روش‌ها برای اصلاح خروجی مدل‌های مبتنی بر CMIP6 در شبیه‌سازی دما و بارش معرفی می‌کند. این نتایج با شواهد ارائه‌شده در مطالعات آزاد و احمدی (۲۰۲۴)، ماراون و ویدمن (۲۰۱۸) و لی و لی (۲۰۲۵) هم‌راستا است، که نشان می‌دهند روش‌های مبتنی بر توزیع، به‌ویژه در مناطق خشک و نیمه‌خشک، به‌طور قابل توجهی خطای سامانمند مدل‌ها را کاهش می‌دهند.
- در مطالعات گذشته، معمولاً مدل‌های ESM نسبت به مدل‌های GCM از دقت بالاتری در شبیه‌سازی داده‌های تاریخی مدل‌های CMIP6 برخوردار بودند. این تحقیق نیز، به ویژه در مورد مدل‌های MPI-ESM1-2-LR و MRI-ESM2-0 که عملکرد بهتری در شبیه‌سازی دما نسبت به مدل‌های GCM مانند HadGEM و INM نشان داده‌اند، این تمایز را برای شبیه‌سازی دما تایید می‌کند. با این حال، در شبیه‌سازی بارش، نتایج این تحقیق با یافته‌های پیشین در تضاد نیست و مدل HadGEM به‌طور برجسته‌ای بهترین عملکرد را در شبیه‌سازی بارش در تمام ایستگاه‌های مطالعه‌شده به‌دست آورد. در بسیاری از مطالعات گذشته، مدل HadGEM به‌عنوان بهترین مدل برای شبیه‌سازی بارش در مناطق خشک و نیمه‌خشک ایران شناخته شده است، که این نتایج با مطالعه‌ی یزدان‌دوست و همکاران (۲۰۲۰) و دیگر پژوهش‌های مشابه همخوانی دارد.

این پژوهش با ارائه یک ارزیابی منطقه‌ای جامع از عملکرد مدل‌های اقلیمی مبتنی بر CMIP6 در شبیه‌سازی

بارش، دمای حداقل و دمای حداکثر در استان‌های خراسان، نشان می‌دهد که کارایی مدل‌ها اقلیمی و روش‌های تصحیح اریبی به‌طور معناداری به نوع متغیر اقلیمی و ویژگی‌های منطقه‌ای وابسته است. نتایج مطالعه حاکی از آن است که اتکای صرف به یک ترکیب مدل-روش به‌عنوان «بهترین گزینه عمومی»، به‌ویژه در مناطقی با ناهمگونی اقلیمی و توپوگرافی پیچیده، می‌تواند منجر به افزایش عدم قطعیت در مطالعات تغییر اقلیم شود. شناسایی خوشه‌های عملکردی همگن میان ترکیب‌های مدل-روش، رویکردی مؤثر برای کاهش عدم قطعیت ساختاری در پیش‌بینی‌های اقلیمی فراهم می‌کند و امکان گذار از رتبه‌بندی‌های منفرد به چارچوبی مبتنی بر مجموعه‌ای از گزینه‌های هم‌ارز را فراهم می‌سازد. این رویکرد می‌تواند پایه‌ای علمی و قابل اتکا برای کاربردهای عملی، از جمله مدیریت منابع آب، برنامه‌ریزی کشاورزی و ارزیابی مخاطرات اقلیمی در مناطق خشک و نیمه‌خشک باشد. چارچوب ارائه‌شده در این مطالعه می‌تواند به‌عنوان پلی میان پیشرفت‌های مدل‌سازی اقلیم جهانی و نیازهای مطالعات کاربردی در مقیاس منطقه‌ای عمل کرده و مبنایی قابل دفاع برای مطالعات پیش‌نگری تغییر اقلیم در ایران و سایر مناطق با شرایط اقلیمی مشابه فراهم آورد.

با وجود جامعیت روش‌شناسی، این پژوهش محدودیت‌هایی دارد: (۱) اتکا به داده‌های ایستگاهی که ممکن است تحت تأثیر خطاهای اندازه‌گیری و ناهمگنی فضایی، به‌ویژه در مناطق کوهستانی، قرار گیرد؛ استفاده از داده‌های شبکه‌ای با تفکیک مکانی بالا یا محصولات سنجش‌ازدور می‌تواند بازنمایی فضایی را بهبود دهد. (۲) دامنه تصحیح اریبی به روش‌های آماری متداول محدود شده است. اگرچه این روش‌ها از شفافیت و تفسیرپذیری بالایی برخوردارند؛ بهره‌گیری از رویکردهای یادگیری ماشین یا روش‌های ترکیبی می‌تواند در بازتولید

به زون‌های اقلیمی همگن و تعیین وزن‌های اختصاصی برای هر زون مورد ارزیابی قرار گیرد. گسترش چارچوب پیشنهادی به سایر مناطق کشور و سناریوهای انتشار آینده می‌تواند به افزایش اعتبار و کاربردپذیری آن در مطالعات ارزیابی اثرات و سازگاری با تغییر اقلیم منجر شود. پرداختن به این موارد در مطالعات آتی می‌تواند به تکمیل چارچوب انتخاب مدل-روش و افزایش قابلیت اعتماد نتایج پیش‌نگری تغییر اقلیم در ایران کمک کند.

منابع

- فرآیندهای غیرخطی و رخداد‌های حدی مؤثر باشد. (۳)
- نتایج این مطالعه در مقیاس منطقه‌ای استخراج شده و تعمیم‌پذیری الگوهای عملکردی شناسایی شده به مقیاس‌های مکانی کوچک‌تر یا سایر نواحی اقلیمی نیازمند ارزیابی‌های تکمیلی است. (۴) در این پژوهش از وزن‌های منطقه‌ای ثابت برای معیارهای ارزیابی به منظور حفظ سازگاری مکانی و امکان مقایسه آماری استفاده شد؛ با این حال، با توجه به تنوع اقلیمی گسترده در شرق ایران، پیشنهاد می‌شود در پژوهش‌های آتی، عملکرد چارچوب‌های چندمعیاره با تفکیک منطقه‌ای ایستگاه‌ها
- Abdolzadeh, F., Khorshiddoust, A. M., & Jahanbakhsh Asl, S. (2023). Projection of the future outlook of temperature and precipitation in Urmia Lake basin by the CMIP6 models. *Physical Geography Research Quarterly*, 55(1), 95–112. <https://doi.org/10.22059/JPHGR.2023.352727.1007737>
- Agrawal, N., & Mujumdar, S. S. (2025). A multimodel ensemble using the entropy-TOPSIS method for projecting temperature for the Lower Tapi River Basin. *Water Practice & Technology*. <https://doi.org/10.2166/wpt.2025.166>
- Ansari Mahabadi, S., Dehban, H., Zareian, M. J., & Farokhnia, A. (2022). Trend analysis of temperature and precipitation changes in Iran's river basins over a 20-year horizon based on CMIP6 model outputs. *Iran Water Research Journal*, 16(1), 11–24. <https://doi.org/10.22034/iwrj.2022.11204>
- Azad, N., & Ahmadi, A. (2024). Assessment of CMIP6 models and multi-model averaging for temperature and precipitation over Iran. *Scientific Reports*, 14, 24165. <https://doi.org/10.1038/s41598-024-74789-4>
- Babaeian, I., Modirian, R., Khazanedari, L., Karimian, M., Kouzegaran, S., Kouhi, M., Falamarzi, Y., & Malbusi, S. (2023). Projection of Iran's precipitation in 21st century using downscaling of selected CMIP6 models by CMHyd. *Journal of the Earth and Space Physics*, 49(2), 431–449. <https://doi.org/10.22059/jesphys.2023.332410.1007436>
- Babaeian, I., Rahmatinia, A. E., Entezari, A., Baaghdeh, M., Aval, M. B., & Habibi, M. (2021). Future projection of drought vulnerability over northeast provinces of Iran during 2021–2100. *Atmosphere*, 12, 1704. <https://doi.org/10.3390/atmos12121704>
- Brumatti, L. M., Sant'Anna Commar, L. F., de Oliveira Neumann, N., Ferreira Pires, G., & Avila-Diaz, A. (2024). Bias correction in CMIP6 model simulations and projections for Brazil's climate assessment. *Earth Systems and Environment*, 8, 121–134. <https://doi.org/10.1007/s41748-023-00368-8>
- Eshaghi, A., & Kamkar, B. (2025). A remote sensing-based framework for agricultural drought risk monitoring and assessment: Introducing SADFI for disaster risk assessment in Northeastern Iran. *Theoretical and Applied Climatology*, 156, Article 684. <https://doi.org/10.1007/s00704-025-05888-z>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958.
- Ghorbani-Minaei, L., Mosaedi, A., Zakerinia, M., Kalbali, E., & Ghabaei Sough, M. (2024). Study of future climate change on temperature and precipitation trends in Qarasu basin based on the CMIP6 models. *Iranian Journal of Soil and Water Research*, 55(2), 245–268. <https://doi.org/10.22059/ijswr.2024.369146.669613>
- Hamidianpour, M., Baaghdeh, M., & Abbasnia, M. (2016). Assessment of precipitation and temperature changes over southeast Iran

- using downscaling of general circulation model outputs. *Physical Geography Research*, 48(1), 107–123. <https://doi.org/10.22059/jphgr.2016.57030>
- Han, R., Li, Z. L., Han, Y. Y., Huo, P. Y., & Li, Z. J. (2023). A comparative study of TOPSIS-based GCM selection and multi-model ensemble. *International Journal of Climatology*, 43(12), 5348–5368. <https://doi.org/10.1002/joc.8150>
- IPCC. (2021). *Climate Change 2021: The Physical Science Basis: Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY: Cambridge University Press. <https://doi.org/10.1017/9781009157896>
- Krishnan, R., Swapna, P., Dey Choudhury, A., Narayansetti, S., Prajeesh, A. G., Singh, M., Modi, A., Mathew, R., Vellore, R., Jyoti, J., Sabin, T. P., Sanjay, J., & Ingle, S. (2021). The IITM Earth System Model (IITM ESM). *arXiv preprint* [arXiv:2101.03410](https://arxiv.org/abs/2101.03410). <https://arxiv.org/abs/2101.03410>
- Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., & Meehl, G. A. (2010). Challenges in combining projections from multiple climate models. *Journal of Climate*, 23(10), 2739–2758.
- Li, X., & Li, Z. (2025). Assessment of bias correction methods for high-resolution daily precipitation projections with CMIP6 models: A Canadian case study. *Journal of Hydrology: Regional Studies*, 58, 102223. <https://doi.org/10.1016/j.ejrh.2025.102223>
- Maraun, D., & Widmann, M. (2018). *Statistical downscaling and bias correction for climate research*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781107588783>
- Mianabadi, A., Mohammadi, S., & Bateni, M. M. (2023). Projection of changes in precipitation and temperature distributions using bias-corrected simulations of CMIP6 climate models (Case study: Kerman synoptic station). *Climate Change Research*, 4(14), 65–84.
- Nguyen, P. L., Alexander, L. V., Thatcher, M. J., Truong, S. C. H., Isphording, R. N., & McGregor, J. L. (2024). Selecting CMIP6 global climate models for CORDEX downscaling. *Geoscientific Model Development*, 17, 7285–7315.
- Ozbuldu, M., & Irvem, A. (2025). Projecting and downscaling future temperature and precipitation based on CMIP6 models using machine learning in Hatay Province, Türkiye. *Pure and Applied Geophysics*, 182, 1825–1842. <https://doi.org/10.1007/s00024-024-03656-0>
- Raeesi, M., Zolfaghari, A. A., Kaboli, S. H., Rahimi, M., de Vente, J., & Eekhout, J. P. C. (2024). Using quantile mapping and random forest for bias correction of high-resolution reanalysis precipitation data and CMIP6 climate projections over Iran. *International Journal of Climatology*, 44(12), 4495–4514.
- Rashidi Ghane, M., Motavalli, S., Janbaz Ghobadi, G., & Kouhi, M. (2023). Evaluating the capability of three statistical downscaling methods for temperature and precipitation outputs of CMIP6 models in the Kashafrud watershed. *Journal of Climatological Research*, 14(53), 117–132.
- Seraj Ebrahimi, R., Zareian, M. J., & Dehban, H. (2024). Evaluation of the performance of CMIP6 models in estimating temperature and precipitation in the Sefidrood Basin. *Journal of Water and Irrigation Management*, 14(2), 277–289.
- Thakur, R., & Manekar, V. L. (2023). Ranking of CMIP6-based high-resolution global climate models for India using TOPSIS. *ISH Journal of Hydraulic Engineering*, 29(2), 175–188. <https://doi.org/10.1080/09715010.2021.2015462>
- Wei, S., Wang, X., Liu, L., Qie, L., Li, Y., Wang, Q., Wang, T., Wang, J., Gou, X., & Yang, M. (2025). Bias correction of CMIP6 GCMs for historical and future air temperatures across China. *Atmospheric Research*, 323, 108193. <https://doi.org/10.1016/j.atmosres.2025.108193>
- Yazdandoost, F., Moradian, S., Izadi, A., & Aghakouchak, A. (2020). Evaluation of CMIP6 precipitation simulations across different climatic zones: Uncertainty and model intercomparison. *Atmospheric Research*, 250, 105369. <https://doi.org/10.1016/j.atmosres.2020.105369>
- Yazdani, D., Zarrin, A., & Roudbari-Dadashi, A. A. (2024). Statistical downscaling of general circulation models (GCMs): History, principles and methods. *Journal of Water and Sustainable Development*, 11(2), 15–26. <https://doi.org/10.22067/jwsd.v11i2.2401-1305>

- Zabihi, O., & Ahmadi, A. (2024). Multi-criteria evaluation of CMIP6 precipitation and temperature simulations over Iran. *Journal of Hydrology: Regional Studies*, 52, 101707. <https://doi.org/10.1016/j.ejrh.2024.101707>
- Zareian, M. J. (2022). Effects of climate change on temperature and precipitation in Yazd Province based on combined output of CMIP6 models. *Journal of Water and Soil Science*, 26(2), 91–105.
- Zareian, M. J., Dehban, H., & Gohari, A. (2022). Evaluation of the accuracy of CMIP6 models in estimating the temperature and precipitation of Iran based on a network analysis. *Journal of Water and Irrigation Management*, 12(4), 783–797. <https://doi.org/10.22059/jwim.2022.345975>. 1006
- Zhu, Y., Tian, D., & Yan, F. (2020). Effectiveness of entropy weight method in decision-making. *Mathematical Problems in Engineering*, 2020, 3564835. <https://doi.org/10.1155/2020/3564835>

A multi-criteria framework (Entropy–TOPSIS) for selecting optimal CMIP6 model–method combinations for precipitation and temperature bias correction in northeastern Iran

Arash Razban¹, Mohammad Mousavi Baygi,^{2*} Iman Babaeian³, Behnam Kamkar² and Nasrin Sarrafzadeh⁴

¹ Ph.D. Student, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

² Professor, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

³ Professor, Meteorological Research Institute and Atmospheric Science, Mashhad, Iran

⁴ M.Sc., Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

(Received: 20 April 2026, Accepted: 13 July 2026)

Summary

Climate change and the intensification of its impacts have increased the need for accurate station-scale reproduction of precipitation and temperature. However, raw CMIP6 outputs, because of their coarse spatial resolution and systematic biases, require bias correction and a robust procedure for selecting appropriate model–method combinations. To address the gap of fragmented and single-variable evaluations in eastern Iran, this study proposes a station-based, multi-criteria framework for the simultaneous ranking of model–method combinations in reproducing precipitation, minimum temperature, and maximum temperature across the three Khorasan provinces (19 stations, 1989–2014). Within this framework, outputs from eight CMIP6 models (five ESMs and three GCMs) were evaluated using eight statistical bias-correction methods implemented in the CMHyd software (five methods for precipitation and four for temperature). The evaluation metrics included NRMSE, MAE, r , and KGE, together with variable-specific PBIAS for precipitation and MBE for temperature. After normalization, the criteria were weighted using Shannon entropy with regionally fixed weights, and station-level ranking was performed using TOPSIS. The results were then aggregated using median ranks, IQR, and the number of winning stations, and statistically assessed using the Friedman and Nemenyi tests. Performance differences among alternatives were significant for all three variables (Tmin: $\chi^2 \approx 538$, $p \approx 10^{-94}$; Tmax: $\chi^2 \approx 458$, $p \approx 10^{-77}$; Pr: $\chi^2 \approx 286$, $p \approx 10^{-44}$). MPI|VarianceScaling and MPI|DistributionMapping were identified as the best regional options for Tmin and Tmax, respectively. Nevertheless, the Nemenyi test showed that 12 (Tmin) and 11 (Tmax) combinations were not statistically distinguishable from the top-ranked option, forming an excellent group characterized by convergent and stable performance indices. For precipitation, HadGEM|DistributionMapping ranked first and produced more accurate performance indices than other members of the excellent group. In contrast, many INM-, CMCC-, and GFDL-based combinations, together with some DeltaChange/LinearScaling and LocalIntensity/PowerTransformation methods, were significantly separated from the excellent group and occupied the lowest ranks. Accordingly, in the Khorasan region, MPI combined with Variance Scaling and Distribution Mapping, and HadGEM combined with Distribution Mapping, are the most stable options for reproducing temperature and precipitation, respectively. The proposed framework provides a practical basis for reducing structural uncertainty and supporting robust selection of climate model–bias correction combinations for climate change impact assessment and regional hazard management.

Keywords: CMIP6, entropy-TOPSIS, Khorasan, precipitation and temperature, statistical bias correction

* Corresponding author:

mousavib@um.ac.ir