npj | climate and atmospheric science

Article

Published in partnership with CECCR at King Abdulaziz University

https://doi.org/10.1038/s41612-025-01033-9

Developing an ensemble machine learning framework for enhanced climate projections using CMIP6 data in the Middle East

Younes Khosravi[1] ✉, Taha B.M.J. Ouarda[1] & Saeid Homayouni[2]

Climate change in the Middle East has intensified with rising temperatures, shifting rainfall patterns, and more frequent extreme events. This study introduces the Stacking-EML framework, which merges five machine learning models three meta-learners to predict maximum temperature, minimum temperature, and precipitation using CMIP6 data under SSP1-2.6, SSP2-4.5, and SSP5-8.5. The results indicate that Stacking-EML not only significantly improves prediction accuracy compared to individual models and traditional CMIP6 outputs but also enhances climate projections by integrating multiple ML models, offering more reliable, regionally refined forecasts. Findings show R^2 improvements to 0.99 for maximum temperature, 0.98 for minimum temperature, and 0.82 for precipitation. Under SSP5-8.5, summer temperatures in southern regions are expected to exceed 45 °C, exacerbating drought conditions due to reduced rainfall. Spatial analysis reveals that Saudi Arabia, Oman, Yemen, and Iran face the greatest heat and drought impacts, while Turkey and northern Iran may experience increased precipitation and flood risks.

Climate change has become a paramount concern in the 21st century, profoundly impacting ecological systems, economic structures, and societal frameworks[1]. In this context, temperature and precipitation fluctuations significantly impact agricultural productivity[2,3], oceanic acidification[2,3], and occurrences of severe droughts[4–6] or inondations[7,8]. This challenge is especially pronounced in areas like the Middle East, with distinctive geographical and climatic attributes. Recent studies have emphasized that this region's rising temperatures and shifting precipitation patterns may engender considerable challenges, including depleted water resources, heightened health risks, and agricultural instability[9–11]. However, accurately predicting these changes and their impacts remains a formidable challenge due to the complexities of the climate system and the inherent uncertainties of climate modeling[12–14]. In this context, initiatives like the sixth phase of the Coupled Model Intercomparison Project (CMIP6), which explores future climate scenarios, are critically important, yet they still face limitations in delivering precise regional forecasts[15].

The CMIP6 signifies a noteworthy progression in global climate simulation, offering a comprehensive array of simulations that have deepened our understanding of climate behavior and forecasts across diverse socio-economic scenarios. Although previous experiments (i.e., CMIP3 and CMIP5 data) have been utilized in several studies[16–24], recent regional evaluations have demonstrated that CMIP6 outputs have significantly improved compared to the earlier phases[25–27]. These improvements include elevated spatial resolution, more intricate physical representations, and an expanded array of scenarios, notably the Shared Socioeconomic Pathways (SSPs) that encapsulate various future trajectories of greenhouse gas emissions[28]. Additionally, CMIP6 has made significant strides in modeling complex climate feedback, such as the interactions between the atmosphere, oceans, and cryosphere, which are crucial for understanding long-term climate dynamics[29]. These advancements have enabled more precise simulations of temperature and precipitation patterns, vital for elucidating regional climate repercussions[30]. Nonetheless, substantial obstacles persist, particularly in the regional application of these models, where inherent uncertainties in model outputs can result in considerable variability in climate predictions[31,32]. These regional uncertainties are often attributed to structural differences among models and the challenges in accurately representing localized physical processes, leading to divergent climate projections for the same region[33].

Multi-Model Ensemble (MME) has gained significant traction in climate research to address these uncertainties. Ensemble modeling, which

[1]Canada Research Chair in Statistical Hydro-Climatology, Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, INRS-ETE, Québec City, QC, Canada. [2]Centre Eau Terre Environnement, Institut National de la Recherche Scientifique, Quebec City, QC, Canada. ✉e-mail: younes.khosravi@inrs.ca

involves combining outputs from multiple climate models, offers a robust methodology for reducing the biases inherent in individual models and capturing a wider range of possible climate outcomes[34,35]. MMEs have also proven to be particularly effective in evaluating a range of climate scenarios, including the SSPs, by providing probabilistic estimates that help to quantify the uncertainties associated with different future emission trajectories Tebaldi and Knutti[36]. By integrating the diverse outputs of CMIP6 models, ensemble techniques enhance the reliability and robustness of climate projections, particularly by reducing the impact of outlier results and identifying consistent trends across different models[37–39]. This integration is critical for improving the accuracy and reliability of climate predictions, thereby supporting more effective adaptation and mitigation strategies in vulnerable regions. In the classification provided by Wang et al.[40], MME is broadly categorized into two principal types: (1) the Simple Ensemble Mean (SEM), which involves the use of individual models combined through traditional statistical methods such as averaging and median techniques with equal weighting (Sillmann et al.[41] Wang et al.[40]); and (2) the Weighted Ensemble Method (WEM), which employs more sophisticated weighting strategies like the independence weighted mean and multidimensional scaling to ensure spatiotemporal consistency[42]. Each of these methods offers distinct advantages and disadvantages. The SEM, recognized for its straightforwardness, is a widely used approach that generally yields superior performance compared to individual model members Lambert and Boer[43]. However, this approach has certain limitations. A key concern is that many models often share similar parameterizations and components, leading to potential interdependencies among different climate simulations[44]. If this interdependence is not properly addressed, it can result in a misleading consensus among models, diminished accuracy, and inaccurate uncertainty estimation[45]. Conversely, WEM provides the benefit of mitigating systematic biases in the outputs of individual ensemble members, thereby enhancing the ensemble's overall predictive accuracy[46].

In recent years, machine learning-based approaches have been developed to enhance and complement traditional ensemble methods, such as the SEM and WEM in climate modeling[47–50]. Machine learning (ML) has emerged as a vital tool in improving the performance of climate model ensembles by effectively handling the complexities and uncertainties inherent in climate projections[51]. Considering these developments, numerous studies have explored integrating ML algorithms into climate model ensembles, demonstrating significant advancements in prediction accuracy and uncertainty reduction. For instance, Bilbao-Barrenetxea et al.[29] demonstrated that MME techniques, when combined with ML algorithms, can enhance the accuracy of precipitation projections and improve hydrological modeling in complex terrains like the Pyrenees. Similarly, Yılmaz et al.[52] utilized CMIP6 ensembles and ML to project significant temperature increases and precipitation decreases in Türkiye's Altınkaya Dam Basin by 2100, highlighting the critical role of ML in refining these projections. Shao et al.[53] further explored enhancing climate projections using CMIP6 ensembles through Time Variability Correction and Ensemble Dependence Transformation (EDT), significantly improving model and ensemble statistics for more accurate predictions. In another study, Wang et al.[54] combined CMIP6 ensemble models with ML algorithms, such as Random Forest and Gradient Boosting, to project future precipitation changes in the Hanjiang River Basin, outperforming individual methods. The effectiveness of ML-based ensemble predictions was also demonstrated in the Western Ghats of India, where methods like XGBoost and Random Forest (RF) showed superior performance in simulating interseasonal variability and predicting future climate changes[32]. In Australia, Grose et al.[28] focused on developing a CMIP6-based multi-model downscaling ensemble for climate change services, indicating the potential for ML integration to enhance regional climate projections.

Furthermore, Zhang et al.[55] applied ML to CMIP6 ensembles to quantify future climate change under different socio-economic pathways, showing that global warming thresholds could be reached by 2048 under the SSP5-8.5 scenario. Singh et al.[56] used ML models, including Random Forest, to predict increased streamflow in the Sutlej River Basin under various emission scenarios, further demonstrating the utility of ML in refining climate projections. Finally, in Iran, Asadollah et al.[50] employed a Gradient Boosting Regression Tree (GBRT) ensemble model with CMIP6 data to downscale and project climate variables, showing significant improvements in replicating the region's climate.

Previous research has mainly focused on evaluating individual ML methods with a uniform structure, highlighting their unique strengths. However, considering the diverse range of factors that impact climatic elements and the various statistical behaviors they exhibit, a single structured ML model may not be sufficient to fully capture the complex relationships between these elements and their predictors across different climatic regions[57]. In recent years, ensemble learning techniques, which amalgamate multiple ML models, have increasingly demonstrated superior efficacy[58]. To build on this progress, three prominent ensemble methods, Bagging, Boosting, and Stacking have been widely adopted to address the limitations of individual ML models[59]. Among these methods, stacking is particularly favored for its capacity to integrate predictions from several base models via a meta-learner, enhancing overall model performance[60–62].

Building on these ensemble learning advancements, recent studies highlight the substantial benefits of advanced ML-based frameworks for climate projections in arid and hyper-arid contexts. For instance, Aldosary et al.[63] integrated multiple ML algorithms such as Random Forest, LightGBM, and XGBoost to predict specific humidity in Dammam, Saudi Arabia, achieving near-perfect accuracy and enabling effective early-warning insights for heat stress hazards. Similarly, Baig et al.[64] examined monthly rainfall prediction in the hyper-arid United Arab Emirates, revealing that XGBoost, Long Short-Term Memory, and stacked ensembles can substantially outperform conventional models when additional meteorological factors (e.g., wind speed) are incorporated.

Meanwhile, Najafi and Kuchak[65] developed a monthly-to-seasonal precipitation forecasting system for Iran using downscaled global model outputs, reporting notable performance gains across diverse climatic zones. Expanding on these regional applications, Al-Saeedi et al.[66] improved precipitation estimation in Jordan by integrating machine learning and geostatistical techniques, enhancing accuracy for drought assessment and water resource management. Similarly, Asadollah et al.[50] applied ML-based downscaling in Iran, demonstrating how GBRT enhances CMIP6-based temperature and precipitation projections. Their findings indicate significant warming (+8 °C in highlands) and major precipitation shifts, emphasizing the need for localized ML downscaling in arid regions. Beyond these region-specific applications, ensemble learning methods have been widely adopted in broader climate modeling contexts, demonstrating their ability to enhance predictive accuracy across diverse environments. For instance, Tuysuzoglu et al.[67] further reinforced the efficacy of ensemble approaches with their ensemble-based K-stars model for rainfall classification in Australia, underscoring how probability-based aggregating (pagging) significantly enhances predictive reliability. Similarly, Jaiswal et al.[68] developed a stacking ensemble model that combined RF, XGBoost, and Support Vector Machine (SVM) to predict rainfall in India, achieving a notable accuracy of 81.2%. Similarly, Shetty et al.[32] also demonstrated the effectiveness of stacking in improving the reliability of climate predictions. Additionally, Li et al.[69] employed a stacking ensemble technique to integrate SVM, RF, Elastic Net Regression, and XGBoost for mid-term streamflow forecasting. Collectively, these studies underscore that ensemble learning especially stacking can effectively mitigate modeling uncertainties, capture a full range of climate extremes, and offer valuable guidance for adaptation strategies in data-scarce regions.

Despite the potential of ML stacking techniques, research that has applied these methods to enhance the accuracy and robustness of climate model predictions has remained limited. Recognizing that the goal of stacking is to leverage the strengths of base algorithms to build a more reliable framework than individual models[70], this research aimed on developing a stacking ensemble model for predicting precipitation and maximum and minimum temperatures using multiple regressors of ML. The proposed method was structured in two levels: base models and a meta-

Table 1 | Assessment of ML models' performance

Model	Tmax					Tmin					Pre				
	RMSE	R^2	MAE	NSE	MBE	RMSE	R^2	MAE	NSE	MBE	RMSE	R^2	MAE	NSE	MBE
RF	2.446	0.944	1.837	0.945	0.01	2.462	0.926	1.822	0.926	−0.01	0.849	0.645	0.419	0.445	0.001
SMV	2.458	0.943	1.864	0.943	0.008	2.467	0.926	1.839	0.925	−0.01	0.853	0.640	0.420	0.443	0.001
LGBM	2.428	0.945	1.861	0.944	0.007	2.458	0.926	1.835	0.926	−0.01	0.847	0.647	0.419	0.446	0.001
XGB	2.458	0.943	1.872	0.943	0.008	2.470	0.926	1.845	0.926	−0.01	0.850	0.644	0.42	0.444	0.001
CB	2.471	0.943	1.889	0.943	0.008	2.481	0.925	1.857	0.925	−0.01	0.854	0.639	0.423	0.439	0.002

Table 2 | Performance metrics of meta-model regressors under two different scenarios

	Model	Tmax		Tmin		Precipitation	
		RMSE	R^2	RMSE	R^2	RMSE	R^2
SC1	ANN	2.368	0.972	2.7	0.966	0.823	0.705
	MLR	2.364	0.97	2.702	0.964	0.844	0.704
	LASSO	2.364	0.971	2.701	0.965	0.826	0.704
SC2	ANN	1.830	0.99	1.5	0.988	0.689	0.821
	MLR	1.838	0.988	1.506	0.987	0.691	0.817
	LASSO	1.832	0.989	1.503	0.971	0.689	0.819

model. At the first level, five ML models, including RF, XGBoost, LightGBM (LGBM), SVM, and CatBoost, were employed as base learners to model the climate elements of the Middle East. In the second level, the outputs from the base models were used to construct a meta-model by comparing three different models, including Artificial Neural Networks (ANN), Multiple Linear Regression (MLR), and Lasso (Least Absolute Shrinkage and Selection Operator) across two scenarios. After determining the most effective regressor for the meta-model, the Stacking-EML model was proposed under three scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5). The novelty of this study lies in developing a more reliable stacked model for improving the prediction of precipitation and temperature extremes in the Middle East.

Results

Performances of the individual ML models

The performance of five ML models in predicting maximum temperature, minimum temperature, and precipitation was evaluated using RMSE, R^2, MAE, NSE, and MBE metrics (Table 1). The results indicate that the differences between models are statistically significant ($p < 0.05$), highlighting that certain models provide superior predictive accuracy. According to the results, the LGBM and RF models consistently outperform the others. LGBM demonstrates the best overall performance, achieving the lowest RMSE and the highest R^2 values and favorable MAE, NSE, and near-zero MBE across all variables. RF followed closely, particularly excelling in temperature predictions.

In contrast, SVM, XGBoost, and CatBoost displayed comparatively lower accuracy, with CatBoost yielding the least favorable outcomes. Due to their superior performance, LGBM and RF are selected as the top models for integration into the meta-model, designed to enhance predictive capabilities by leveraging their strengths. The other models, including SVM, XGBoost, and CatBoost, showed lower accuracy, with CatBoost performing the worst overall. The decision to select LGBM and RF for meta-modeling is based on their proven ability to minimize errors and effectively capture the variability in climate data, making them suitable for developing robust climate projections for the study area. Consistent with our findings[71], reported that the RF model demonstrates strong potential in precipitation prediction. Our experience here was also similar to the conclusions achieved by ref. 72, who emphasize that RF outperformed traditional ensemble methods, suggesting that ML techniques like RF and LGBM can significantly enhance the accuracy of climate predictions compared to conventional models. The LGBM model is also noted for its efficiency and accuracy in handling large datasets, which is essential for climate modeling[73].

Performances of the Stacking Ensemble Models

The selection of the meta-model plays a crucial role in determining the final fit when stacking models. In this study, two scenarios were considered for the meta-modeling process. As outlined in Table 1, the RF and LGBM models were chosen as the base ML algorithms for the stacking approach. Three regressors, including ANN, MLR, and LASSO, were used to develop the final meta-model under two scenarios. The results indicate that ANN is the most effective regressor for both scenarios, particularly when additional geographic and topographic variables are included, as in SC2 (Table 2). This performance justifies the selection of ANN for constructing the final meta-model, while MLR and LASSO could serve as supplementary models, given their competitive performance. A review of previous studies reveals that, so far, geographical variables have not been utilized to enhance the accuracy of meta-models. However, current findings show that incorporating these variables significantly improves the performance of all models, underscoring the importance of using diverse data sources for robust climate modeling in the study area. Therefore, it is recommended that future research consider including geographical variables to advance the accuracy and reliability of meta-model construction.

To further assess the performance of the models and the influence of geographic variables on prediction accuracy, a Taylor diagram analysis was conducted for three climatic elements: maximum temperature, minimum temperature, and precipitation (Fig. 1). The analysis demonstrated that ML-based stacking models, particularly the stacking model incorporating an ANN in the SC2, outperformed other models, including traditional CMIP6 models such as AWI-CM-1-1-MR, MIROC6, and MRI-ESM2-0. These stacking models exhibited high correlation coefficients (above 0.95 for all variables) and standard deviations closely aligned with the observed data (ERA5 reanalysis dataset), indicating a superior ability to capture complex climatic patterns. Specifically, the stacking model with ANN in SC2 achieved the best performance among all models, with correlation coefficients of 0.99 for maximum temperature, 0.98 for minimum temperature, and 0.82 for precipitation. In contrast, the CMIP6 models, particularly MIROC6, displayed higher standard deviations (15 for maximum temperature, 10 for minimum temperature, and 1.25 for precipitation) and lower correlation coefficients, highlighting their limitations in reproducing climate variability.

Historical simulation assessment and bias analysis of climate variables

Figure 2a–i illustrates the spatial distribution of annual maximum temperature, minimum temperature, and precipitation for both observed and simulated data and the local R^2 maps that evaluate model accuracy across the Middle East. According to Fig. 2a, southern regions, such as the Arabian Peninsula and North Africa, experience higher temperatures, a phenomenon attributable to their proximity to desert areas and the influence of hot, dry winds[74]. These areas are depicted in darker red shades on the map. In contrast, northern regions like Turkey and parts of Iran, characterized by higher elevations or proximity to the Mediterranean and Caspian Seas, exhibit lower temperatures. This variation is primarily due to differences in

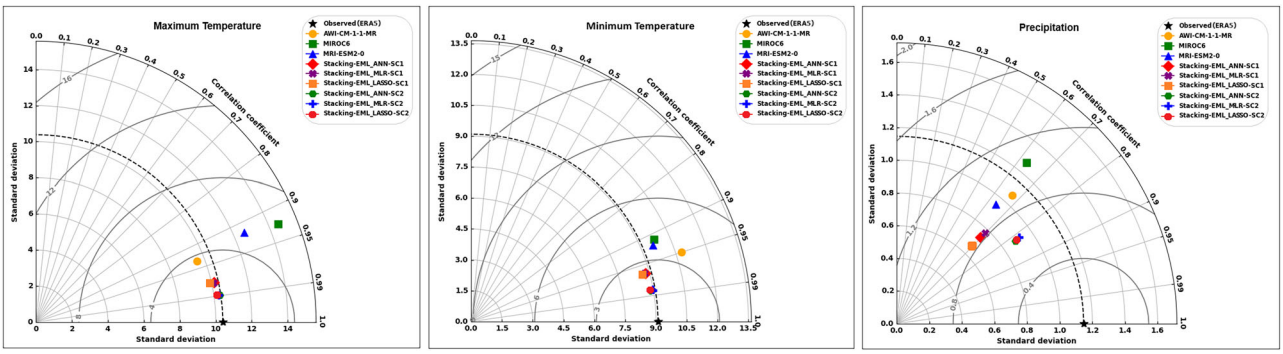

Fig. 1 | Taylor diagrams comparing the predictive performance of CMIP6 models and stacking-based machine learning models for three key climatic variables. Model skill is assessed for **a** maximum temperature, **b** minimum temperature, and **c** precipitation based on correlation coefficient, standard deviation, and centered root mean square error. The diagrams contrast traditional CMIP6 models (AWI-CM-1-1-MR, MIROC6, MRI-ESM2-0) with stacking ensemble models using ANN, MLR, and LASSO meta-learners under two scenarios (SC1 and SC2), with ERA5 reanalysis data serving as the observational benchmark.

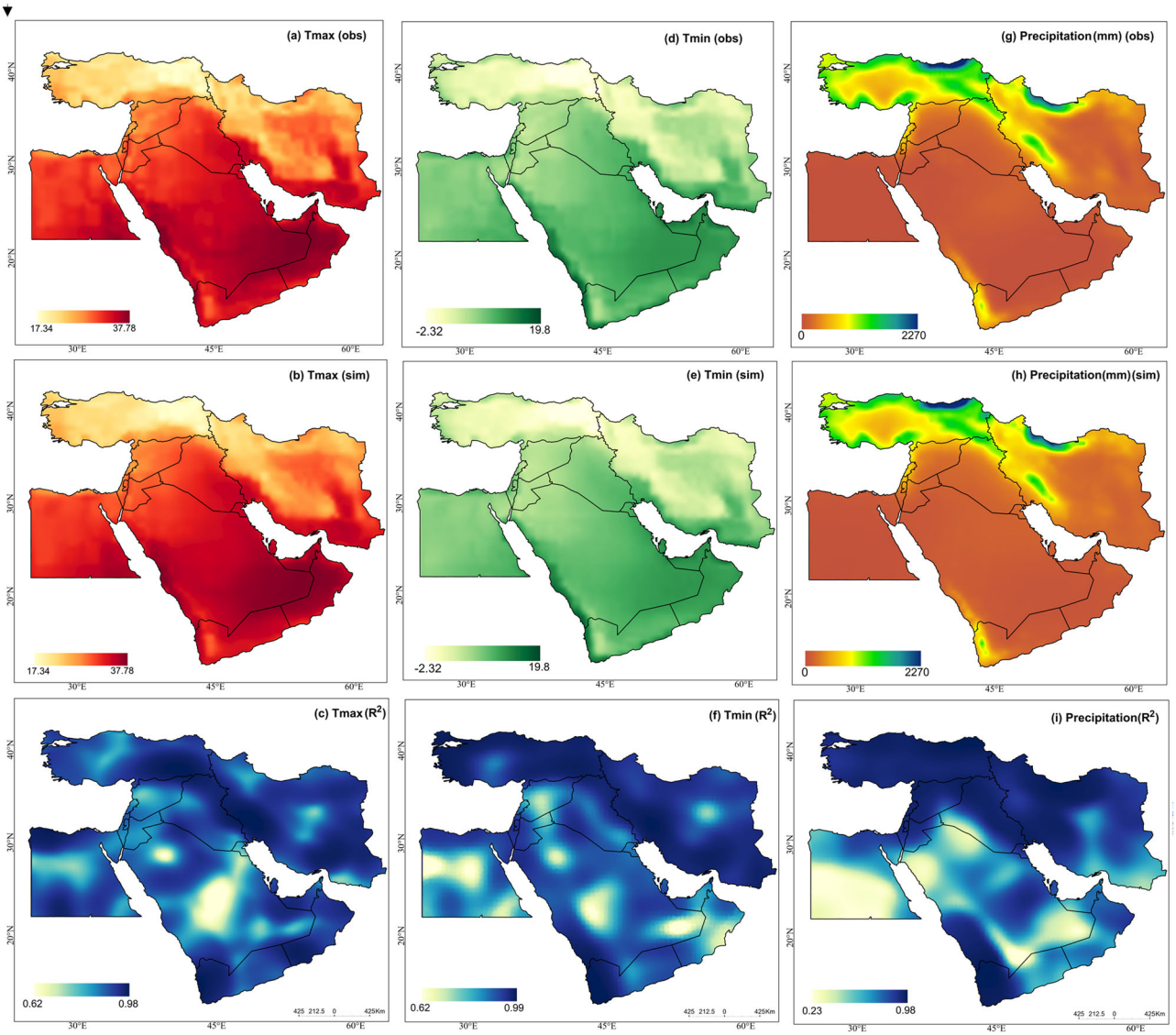

Fig. 2 | Spatial patterns of observed and simulated climatic variables and model performance across the Middle East. a–c Annual average maximum temperature: **a** observed, **b** simulated, **c** local R^2 values. Annual average minimum temperature: **d** observed, **e** simulated, **f** local R^2 values. Annual average precipitation: **g** observed, **h** simulated, **i** local R^2 values.

latitude, elevation, and environmental influences, such as the proximity to water bodies and local climatic conditions, especially the effects of deserts and prevailing wind directions in the region[75]. The climate model employed in this study has successfully reproduced these observed patterns, showing a high correlation between simulated and observed temperatures (Fig. 2b). Figure 2d illustrates the spatial distribution of minimum temperatures across the region. Lower temperatures are noted in the northern areas, such as Turkey and the mountainous regions of Iran, while higher temperatures prevail in the southern areas, including the Arabian Peninsula and North Africa. This spatial distribution reflects the influence of geographical and climatic factors like elevation, latitude, and proximity to deserts. The simulated model closely reproduces the general pattern of observed minimum temperatures, capturing the key features of the region's minimum climate (Fig. 2e). The observed precipitation map (Fig. 2g) provides a clear depiction of rainfall distribution in the Middle East, highlighting higher precipitation levels in the northern and mountainous areas, particularly in northern Turkey, western Iran, and the mountainous regions of Lebanon and Syria. These areas experience elevated rainfall due to specific topographical conditions and the influence of moist air currents. Conversely, the southern and central parts of the Arabian Peninsula and parts of North Africa display the lowest precipitation levels, marked by orange and red shades on the map. The simulated precipitation map (Fig. 2h) aligns closely with the observed distribution, with the discrepancies largely minor and within the acceptable uncertainties of climate models.

The local R^2 maps comprehensively evaluate the model's accuracy in simulating the climatic variables across the Middle East. For maximum temperature, the R^2 values range from 0.62 to 0.98 (Fig. 2c). The highest R^2 values (close to 0.98) are found in the northern parts of the Middle East, including Turkey, western Iran, and parts of northern Iraq. This indicates that the model has a strong performance in capturing the maximum temperature patterns in these areas. In contrast, lower R^2 values (around 0.62) are mainly observed in the central regions of the Arabian Peninsula, parts of Egypt, and central Iran. These spatial differences may result from complex local geographical features, such as deserts, variations in elevation, and diverse climatic conditions that pose challenges to accurate model simulations.

Similarly, the R^2 map for minimum temperature displays a spatial distribution akin to that of maximum temperature, with values ranging from 0.62 to 0.99 (Fig. 2f). However, more regions with lower R^2 values are observed, suggesting that the model encounters greater challenges in simulating minimum temperatures. The model performs better in predicting minimum temperatures in northern areas of the Middle East, including Turkey, northern Iran, and northern Iraq. In contrast, regions in the south, particularly the central parts of the Arabian Peninsula, sections of Egypt, and Yemen, exhibit lower R^2 values. This discrepancy may be due to the complex interactions of geographical and climatic factors in these areas, such as daily and seasonal temperature fluctuations, desert influences, and regional atmospheric dynamics. The R^2 map for precipitation reveals a broader range of values, from 0.23 to 0.98, highlighting the intrinsic complexities of precipitation modeling in the Middle East (Fig. 2i). Higher R^2 values (close to 0.98) are primarily observed in the northern parts of the Middle East, including Turkey and northwestern Iran. Notably, these regions correspond with areas that experience higher precipitation, indicating that the model effectively captures precipitation patterns where rainfall is more abundant. On the other hand, lower R^2 values are evident in the central and southern parts of the Arabian Peninsula, sections of Egypt, and other drier regions. This is likely due to the significant variability in precipitation and the influence of localized geographical and atmospheric conditions, which complicate accurate model simulations.

Spatial analysis of climate hotspots and cold spots in base year (2015)

The following analysis focuses on identifying significant hotspots and cold spots for maximum temperature, minimum temperature, and precipitation in the base year (2015) across the Middle East (Fig. 3). These hotspot maps utilize the Getis-Ord Gi* statistic in ArcGIS Pro V3.3.0 to identify areas with clusters of high or low values, highlighting significant spatial patterns of climate variability. By understanding the spatial distribution of these hotspots, we can better assess baseline climate risks and identify regions likely to experience the most significant impacts of climate change in future scenarios. The spatial distribution of maximum temperatures (Tmax) exhibits clear regional disparities (Fig. 3a). Hotspots characterized by higher maximum temperatures are most prominent in southern Iran, Saudi Arabia, the east of Egypt, and most of Oman. These areas display confidence levels between 90% and 99%, signifying a high degree of statistical confidence that these regions experienced significantly elevated Tmax values in 2015. In contrast, northern Turkey and northwestern Iran present cold spots with confidence levels of 95% and 99%, suggesting significantly lower maximum temperatures in these regions compared to the surrounding areas. These findings underscore the region's thermal heterogeneity, with extreme heat and cooler zones coexisting within relatively close geographic proximity. The minimum temperature (Tmin) map (Fig. 3b) reflects a pattern broadly consistent with the Tmax results.

Hotspots are observed in southwestern Saudi Arabia and along the southern coastline of Iran, where regions exhibit confidence levels from 90% to 99%, indicating areas of significantly higher minimum temperatures. This suggests that these areas experienced relatively higher minimum temperatures in 2015. Conversely, northern Turkey again emerges as a cold spot, showing 95% and 99% confidence levels for lower Tmin values. The consistency of cold spot identification in both Tmax and Tmin data suggests that northern Turkey may have been a region of notable thermal anomaly during this period. The spatial distribution of precipitation in 2015 presents a clear divide between the northern and southern parts of the Middle East (Fig. 3c). Northern Turkey, parts of northern Iraq, and areas along the Zagros Mountains in western Iran appear as prominent hotspots, with confidence levels of 95%. These regions received significantly higher precipitation levels than their surroundings, marking them as key zones of concentrated rainfall. The concentration of hotspots in these northern regions indicates that these areas experienced an unusually wet year in 2015, with precipitation levels significantly above the regional average.

On the other hand, southern Egypt, much of Saudi Arabia, and southwestern Yemen emerge as notable cold spots, reflecting areas of very low precipitation. These regions, particularly in southwestern Saudi Arabia, display confidence levels between 95% and 99%, highlighting a significant reduction in rainfall compared to other parts of the Middle East. The spatial clustering of cold spots in these southern areas emphasizes the persistently dry conditions that characterized these regions in 2015, with extremely low precipitation exacerbating the arid climate. The central regions, including much of Iraq, Syria, and central Iran, are classified as "Not Significant," indicating that precipitation levels did not deviate significantly from the regional norm in 2015. However, the sharp contrast between the wet conditions in the northern hotspots and the dry, cold spots in the south underscores the regional precipitation pattern imbalance.

Projected climate variability and future trends using Stacking-EML

Under various climate scenarios, two analytical methods were employed to assess the projected changes in maximum temperature, minimum temperature, and precipitation in the Middle East from 2015 to 2099. First, we utilized monthly pixel-based data averaged across the entire region to provide a comprehensive overview of temporal patterns in these climatic variables. This approach facilitated the illustration of overall trends over time. Figure 4 presents the variation patterns of maximum temperature, minimum temperature, and precipitation under the SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios. Analysis indicates a significant upward trend in maximum temperatures across all scenarios, particularly during summer. Under the SSP1-2.6 scenario, representing a relatively optimistic pathway, summer maximum temperatures gradually rise to ~35–40 °C, while winter temperatures remain relatively cool between 15 and 20 °C, suggesting a stable seasonal temperature pattern. In the SSP2-4.5 scenario, the increase is more

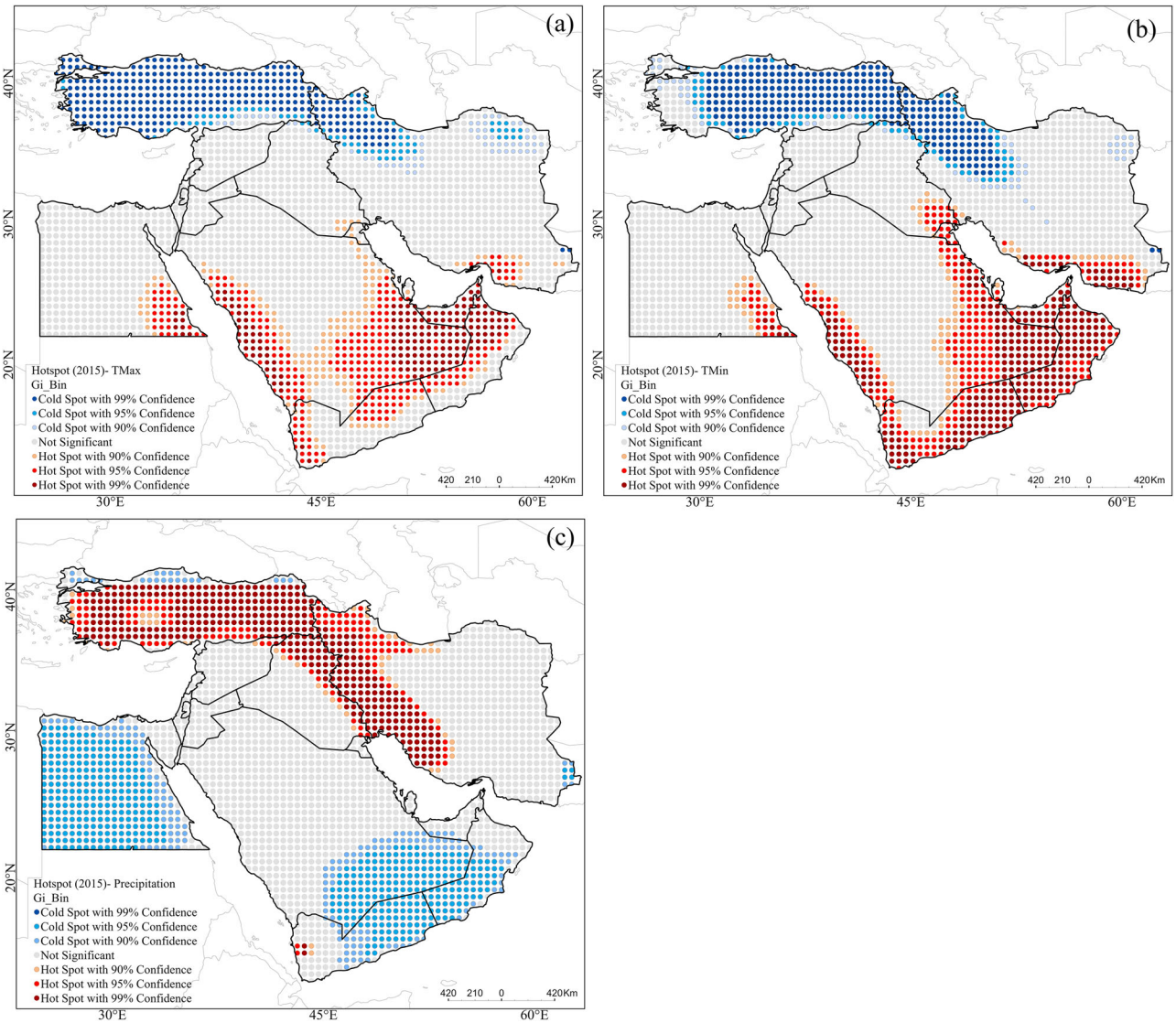

Fig. 3 | Hotspot and cold spot distribution of climate variables across the Middle East in the baseline year 2015. Maps show the spatial clustering of **a** maximum temperature (Tmax), **b** minimum temperature (Tmin), and **c** precipitation based on Getis-Ord Gi* statistics. Hotspots represent statistically significant clusters of high values, and cold spots represent significant clusters of low values at 90%, 95%, and 99% confidence levels. Areas not meeting statistical significance are labeled as Not Significant. These maps highlight spatial extremes and heterogeneity in baseline climate conditions.

pronounced, with summer temperatures frequently exceeding 40 °C and winter temperatures rising to about 20–25 °C. This scenario exhibits greater warming during spring and autumn, indicating more noticeable temperature fluctuations. The SSP5-8.5 scenario, representing high greenhouse gas emissions, projects extreme temperature increases, with summer maximum temperatures surpassing 45°C and winter temperatures exceeding 25 °C. This scenario indicates a substantial loss of seasonal temperature balance, with intense and persistent heat throughout the year. Projections for minimum temperatures indicate a pronounced upward trend, particularly under the more pessimistic climate scenarios. Under SSP1-2.6, minimum temperatures during summer months gradually increase to around 20 °C, while winter months maintain their relative coolness.

In contrast, the SSP2-4.5 scenario shows minimum temperatures exceeding 25 °C during summer, with winter minimum temperatures rising to 10–15 °C. The SSP5-8.5 scenario projects minimum summer temperatures above 30 °C and significant warming during winter to 15–20 °C. Precipitation projections suggest substantial variations and reductions, particularly during the summer months. Under SSP1-2.6, the precipitation pattern remains relatively stable, with marked reductions during summer indicative of seasonal droughts, while winter and spring months experience moderate rainfall. The SSP2-4.5 scenario continues to exhibit dry summer periods but with an increased frequency of heavy rainfall events in certain years and months, implying heightened variability and extreme precipitation events. Winter months in this scenario experience less rainfall compared to SSP1-2.6. The SSP5-8.5 scenario projects the most severe changes, with intensified summer droughts and a near absence of rainfall from June to September. Furthermore, sporadic occurrences of severe and irregular precipitation events during spring and winter reflect a significant reduction in annual precipitation and heightened climatic variability.

Projected spatial changes in climate parameters using Stacking-EML

Analyzing maximum temperature, minimum temperature, and precipitation at a general level lacks the spatial resolution to capture detailed regional changes in the Middle East. To address this, the Hotspot Analysis Comparison tool in ArcGIS Pro V3.3.0 with Fuzzy weights as a similar weighting method was employed to identify areas with significant changes, hot spots, and cold spots across different climate scenarios. This approach examines spatial variations and highlights regions sensitive to climate change. The analysis projects how these climatic elements will evolve under three

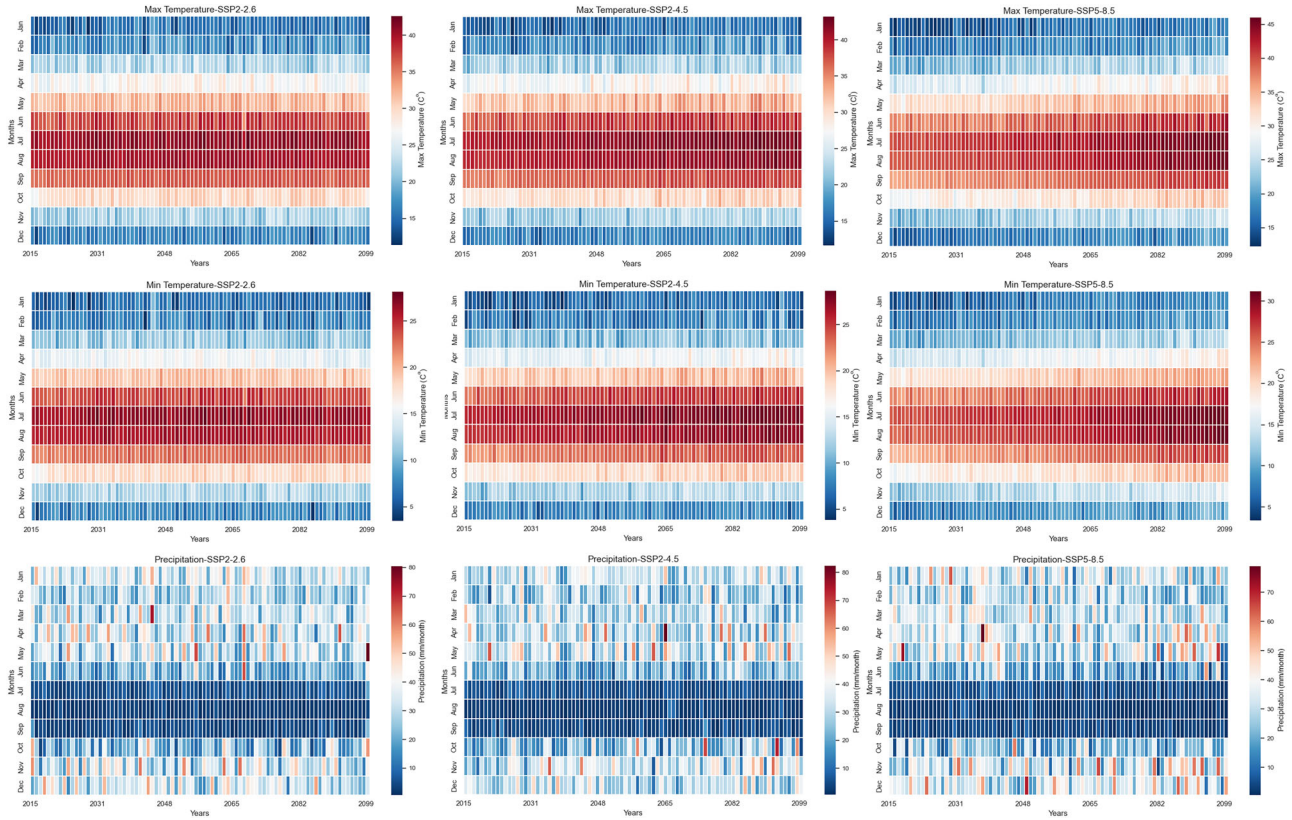

Fig. 4 | Temporal evolution of projected maximum temperature, minimum temperature, and precipitation in the Middle East from 2015 to 2099 under three SSP scenarios. This figure presents heatmaps of monthly climatological variables including maximum temperature (top row), minimum temperature (middle row), and precipitation (bottom row), averaged across the region. Columns correspond to SSP1-2.6 (left), SSP2-4.5 (middle), and SSP5-8.5 (right).

emissions scenarios, SSP1-2.6, SSP2-4.5, and SSP5-8.5, over the periods 2015–2045, 2045–2075, and 2075-2099, offering insights into regional vulnerabilities.

Figure 5 presents the projected hotspots of maximum temperature changes in the Middle East across the SSP scenarios for the specified timeframes. Under the low-emission SSP1-2.6 scenario, the figure reveals limited but noticeable warming, particularly in the southern regions (Fig. 5a). Red areas shift to "Not significant to Hot," mainly in central Saudi Arabia, parts of Yemen, and eastern Egypt. In Turkey and northern Iran, cooler regions maintain their "cold to cold" status. Between 2045 and 2075, the warming continues with less intensity; severe warming is controlled, and northern areas retain their cool status, though slight warming is observable (Fig. 5b). By 2099, significant temperature increases are prevented; while red areas slightly expand in Saudi Arabia and southern Iran, severe heat impact remains limited, and northern regions maintain cooler climates (Fig. 5c). In the moderate-emission SSP2-4.5 scenario, more significant warming occurs. Hot areas expand in central Saudi Arabia, the Red Sea coasts, and southern Iran, highlighting notable increases in maximum temperatures (Fig. 5d). By 2075, these areas will intensify and extend into southern Iraq, Oman, and western Yemen, indicating a shift toward extreme heat conditions (Fig. 5e). In the north, cooler areas remain stable, but initial signs of warming emerge, with some regions shifting to "cold to Not significant." By 2099, a distinct north-south division emerges, with the southern regions engulfed in red zones, signaling persistent extreme heat, while the northern areas show a more widespread warming trend (Fig. 5f).

Under SSP5-8.5, the most severe warming is projected. By 2045, vast sections of the southern Middle East, including southern Saudi Arabia, the Gulf countries, and parts of Iran, will be engulfed in red zones, indicating substantial increases in maximum temperatures (Fig. 5g). Between 2045 and 2075, the southern half of the Middle East remains dominated by these red areas, with prolonged and intense heat affecting Saudi Arabia, Oman, Yemen, and Egypt. In the northern regions, blue areas gradually diminish, transitioning toward neutral conditions (Fig. 5h). As the century draws to a close, severe warming peaks; red zones spread across the southern regions, leaving few areas in the north untouched. This scenario highlights the grave consequences of high emissions, potentially resulting in uninhabitable conditions (Fig. 5i).

For minimum temperatures (Fig. 6), the SSP1-2.6 scenario indicates moderate changes between 2015 and 2045. Blue areas ("cold to cold") are widespread in northern regions, particularly in Turkey and northern Iran, signifying stable cooler temperatures (Fig. 6a). This stability reflects the success of emission reduction efforts in preserving cooler climates in these areas. Meanwhile, southern regions like Yemen, southern Iran, and Saudi Arabia fall within red zones ("Hot to Hot"), pointing to persistently high minimum temperatures. From 2045 to 2075, the red areas in the south expand slightly (Fig. 6b), though emission control efforts effectively limit severe warming. Northern regions remain largely unchanged, except for a slight decrease in blue areas in Iran, suggesting a mild warming trend. By 2099, the overall patterns show little change (Fig. 6c), with the north maintaining its status while the south continues to experience higher temperatures. These mild changes highlight the positive impact of strong climate actions in mitigating severe shifts in minimum temperatures. Under the SSP2-4.5 scenario, more noticeable changes in minimum temperatures occur. In the period from 2015 to 2045, southern regions, particularly Saudi Arabia, southern Iraq, and western Iran, transitioned to red areas ("Not significant to Hot"), indicating an increase in minimum temperatures (Fig. 6d). Northern regions remain cool with blue areas, though this stability appears increasingly threatened. By 2075, red zones in the south expand further, with significant warming observed in Saudi Arabia and southern Iraq (Fig. 6e). This trend suggests growing pressure on water resources, agriculture, and the environment. Although the north maintains cool temperatures, some blue areas gradually shift to neutral zones. By 2099, red

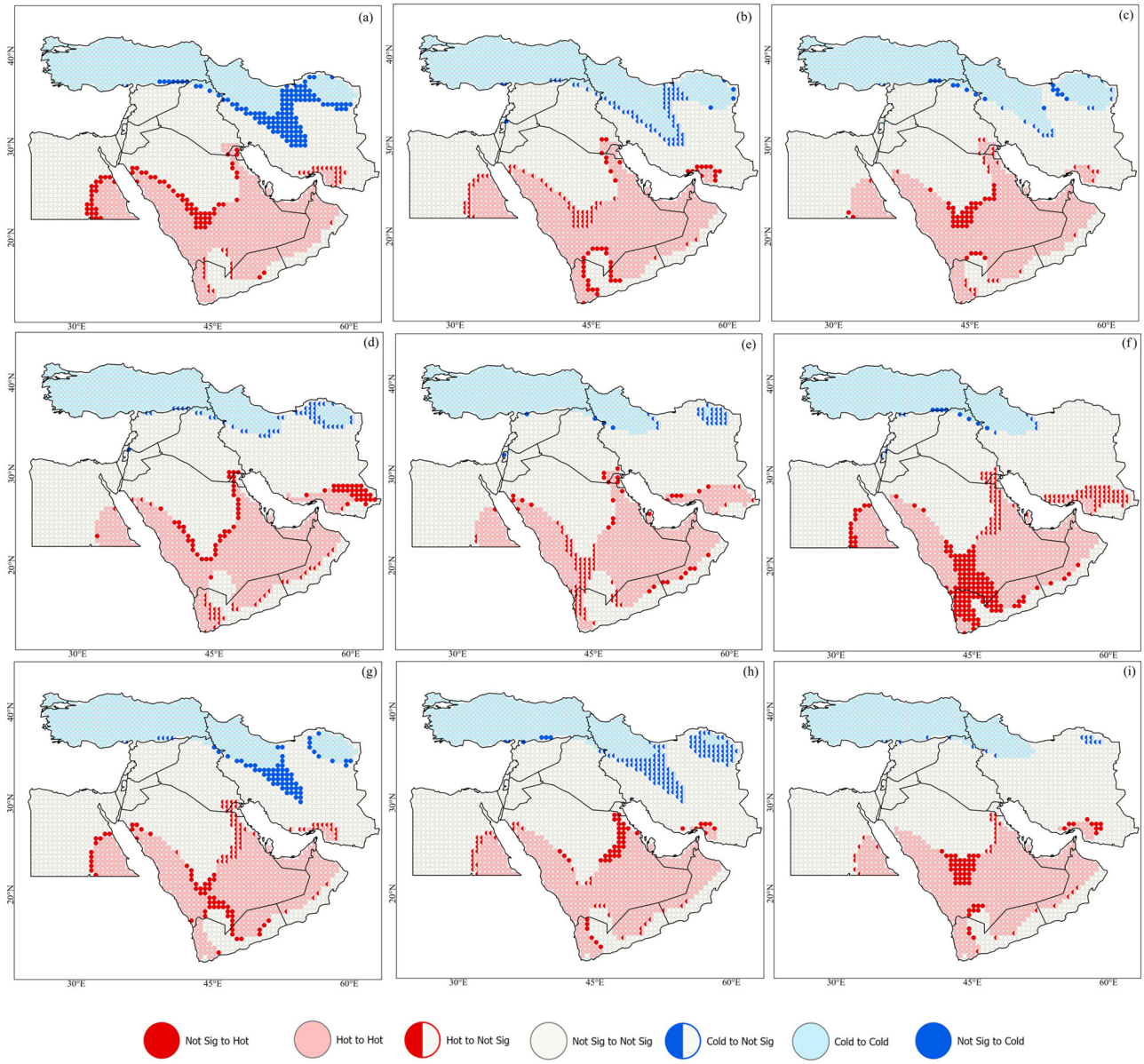

Fig. 5 | Spatial projection of maximum temperature hotspot dynamics in the Middle East under three SSP scenarios (SSP1–2.6, SSP2–4.5, SSP5–8.5) for the future periods. a, **d**, **g** 2015–2045, **b**, **e**, **h** 2045–2075, and **c**, **f**, **i** 2075–2099. The maps display hotspot and cold spot transitions based on the Getis-Ord Gi* statistic using Fuzzy weights. Color coding highlights statistically significant shifts in maximum temperature: deep red (Hot to Hot), pale red (Not Sig to Hot), and light red (Hot to Not Sig) indicate warming trends, while blue shades represent cooling (Cold to Cold, Cold to Not Sig, Not Sig to Cold).

areas dominate the southern Middle East (Fig. 6f), indicating severe warming. While northern regions mostly retain cool temperatures, signs of warming emerge, creating a clear North-South divide. The SSP5-8.5 scenario presents the most drastic changes in minimum temperatures. Between 2015 and 2045, red areas became prominent across the Arabian Peninsula and northern Yemen (Fig. 6g), where regions previously experiencing moderate temperatures now face significant warming. While northern regions initially remain within blue zones, some areas shift to "Not significant to cold." By 2075, warming intensifies, with red areas spreading across most of the southern Middle East (Fig. 6h). Northern regions lose their blue areas, indicating a gradual warming trend. By 2099, severe warming is evident as red zones engulf much of Saudi Arabia, Oman, Yemen, and Iran (Fig. 6i). The northern regions lose most of their blue areas, transitioning to warmer categories, underscoring the grave consequences of continued high emissions. The SSP5-8.5 scenario presents the most drastic changes in minimum temperatures. From 2015 to 2045, red areas became prominent across the Arabian Peninsula and northern Yemen (Fig. 6g), where regions previously experiencing moderate temperatures now face significant warming. Although northern regions initially remain within blue zones, some areas shift to "Not significant to cold." As we move into 2075, warming intensifies, with red areas spreading across most of the southern Middle East (Fig. 6h). In the north, blue areas begin to disappear, indicating a gradual warming trend. By the end of the century, severe warming is evident as red zones engulf much of Saudi Arabia, Oman, Yemen, and Iran (Fig. 6i). Most northern regions lose their blue areas, transitioning to warmer categories, highlighting the grave consequences of continued high emissions.

The analysis of projected precipitation changes under the SSP1-2.6 scenario reveals varying patterns between the northern and southern regions of the Middle East. From 2015 to 2045, northern areas, including Turkey and Iran, experienced a relative increase in precipitation ("Not significant to Hot"), as shown in Fig. 7a, while southern regions like Saudi Arabia faced decreased rainfall. From 2045 to 2075, precipitation patterns in the north remain stable. However, some areas in Iran shift towards unstable

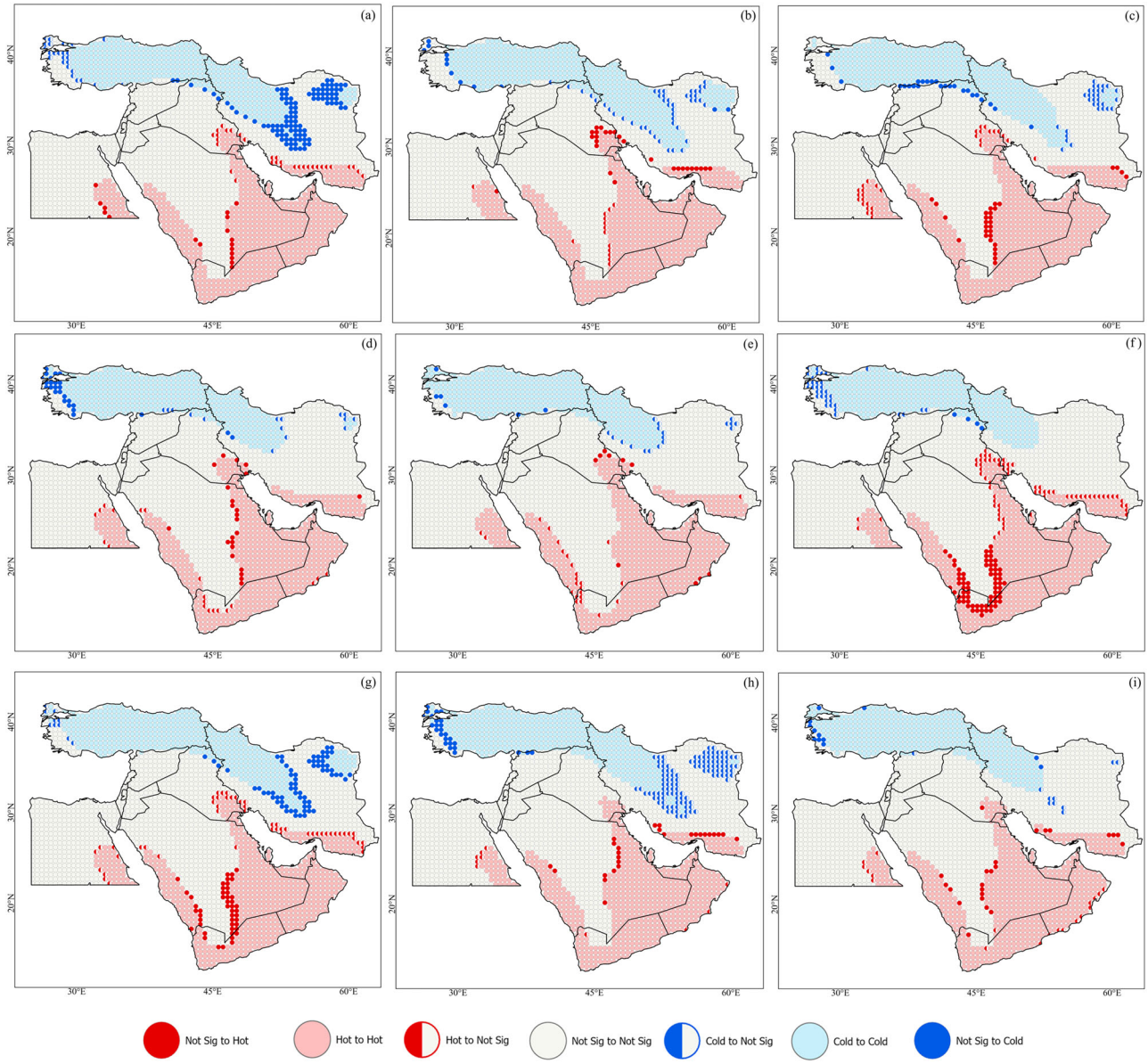

Fig. 6 | Spatial projection of minimum temperature hotspot dynamics in the Middle East under three SSP scenarios (SSP1–2.6, SSP2–4.5, SSP5–8.5) for the future periods. a, **d**, **g** 2015–2045, **b**, **e**, **h** 2045–2075, and **c**, **f**, **i** 2075–2099. This hotspot analysis is based on the Getis-Ord Gi* statistic using fuzzy spatial weights. Red shades indicate significant warming trends, with dark red representing consistent hot spots (Hot to Hot) and pale red showing newly emerging hot areas (Not Sig to Hot). Blue tones reflect areas of relative cooling (Cold to Cold, Not Sig to Cold), while gray zones indicate no significant change.

conditions ("Hot to Not significant"), as depicted in Fig. 7b. Regions with low precipitation extend to higher latitudes, affecting central Middle Eastern areas. By 2099, the SSP1-2.6 scenario prevents more severe changes; northern regions maintain stability, while southern areas continue to suffer from aridity (Fig. 7c). Under the SSP2-4.5 scenario, more significant changes occur. Between 2015 and 2045, northern regions continue to receive increased rainfall (Fig. 7d), whereas southern regions, especially Saudi Arabia and Yemen, encounter significant decreases. In 2075, central and western Iran transformed into "Hot to Hot" areas (Fig. 7e), indicating increased precipitation. In contrast, southern regions move towards severe aridity ("cold to cold"). By the end of the century, red areas dominate the southern Middle East (Fig. 7f), signifying severe aridity, while northern regions show signs of instability. The SSP5-8.5 scenario predicts the most severe precipitation changes. During the period from 2015 to 2045, southern Saudi Arabia, Yemen, and Oman shifted from "cold to Not significant" (Fig. 7g), although vast areas remain "cold to cold." Western Iran experiences decreased precipitation, changing from "Hot to Not significant." Between 2045 and 2075, precipitation increased in southern and western Iran (Fig. 7h), but Saudi Arabia, Oman, and Egypt changed to "cold", signaling intensified aridity. Severe drought conditions dominate towards the end of the century (Fig. 7i), with "cold to cold" areas prevalent in southern regions, while northern areas exhibit widespread instability, even in previously stable rainfall zones.

Discussion

In the Middle East, the climate crisis has become a primary concern for the region's countries[76,77]. Food shortages, exacerbated by high birth and consumption rates and migration from unbearably hot areas to cooler regions, threaten the agricultural sector, which still employs 40% of the Middle Eastern population[78]. These factors illustrate the significant challenges the climate crisis poses for the region. This development not only endangers regional stability but also threatens global security. This study was designed

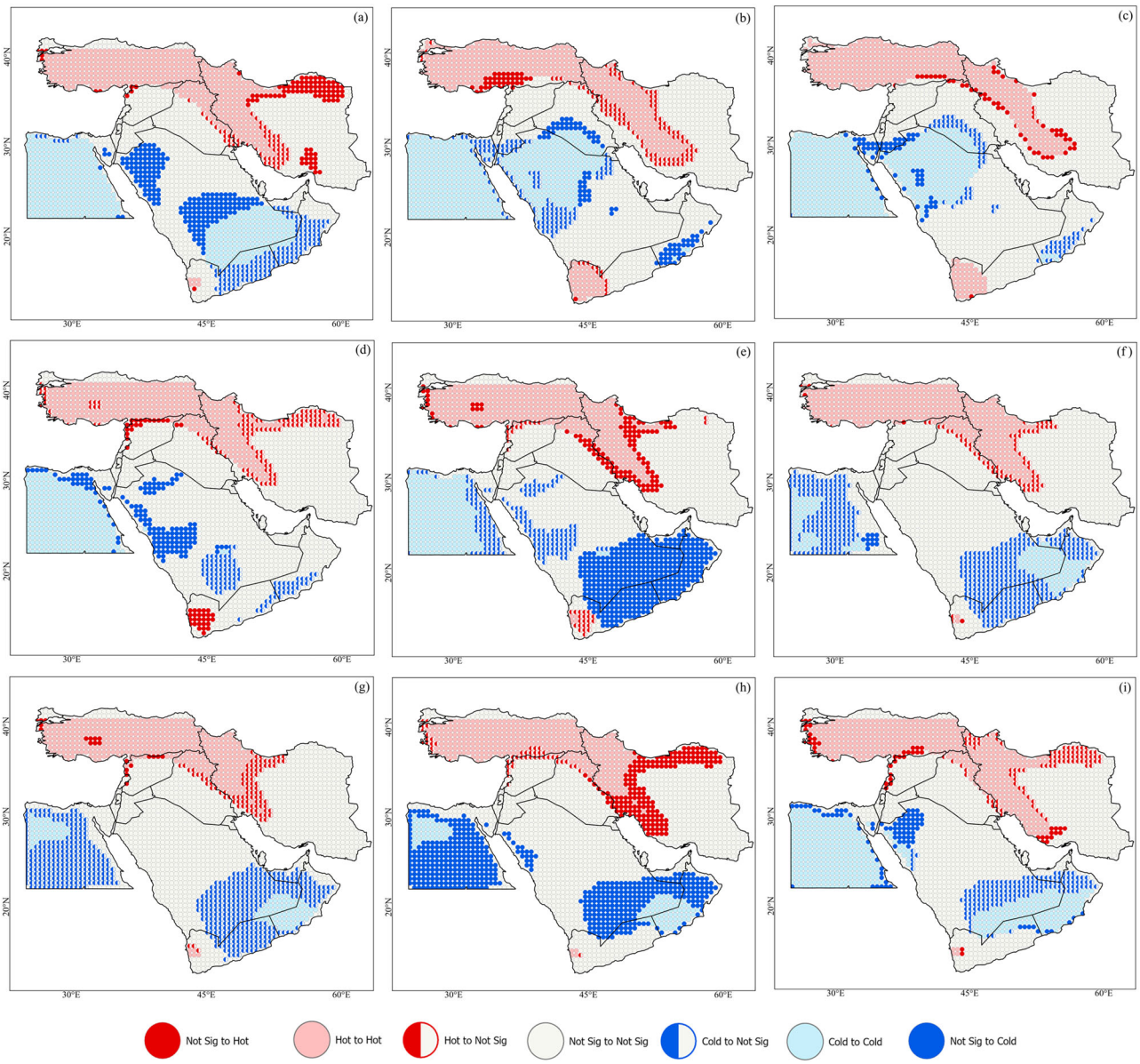

Fig. 7 | Spatial projection of precipitation hotspot dynamics in the Middle East under three SSP scenarios (SSP1–2.6, SSP2–4.5, SSP5–8.5) for the future periods. **a**, **d**, **g** 2015–2045, **b**, **e**, **h** 2045–2075, and **c**, **f**, **i** 2075–2099. This analysis utilizes the Getis-Ord Gi* statistic with fuzzy spatial weighting to identify significant changes in precipitation patterns across the region. Red areas indicate zones of statistically significant increases in precipitation (e.g., "Hot to Hot"), while blue shades show areas of drying (e.g., "Cold to Cold"). The gray category represents areas with no significant change.

to develop a stacking model to improve climate projections' accuracy for key climatic elements in the Middle East. One of the key findings is the superior performance of the "Stacking-EML" framework compared to traditional methods, particularly in the second scenario where geographical factors, including longitude, latitude, elevation, slope, and aspect, were incorporated into the model estimations. These results underscore that ML approaches, especially stacking models based on ANN, can significantly enhance the accuracy of climate predictions. By combining the capabilities of multiple models and leveraging their complementary features, these approaches improve predictive performance for complex climatic variables. This advancement provides a more reliable foundation for policymakers to formulate strategies addressing the climate crisis in the region.

Our study's findings reveal that the expected rise in temperatures, especially under the SSP5-8.5 scenario, could exacerbate these issues, leading to even more profound environmental and societal consequences. The substantial rise in temperatures, especially under the SSP5-8.5 scenario, indicates a severe warming trend, which could lead to profound environmental and societal consequences[79,80]. For instance, it is projected that summer daytime temperatures will exceed 45 °C, which creates hazardous heat conditions, puts immense strain on energy systems required for cooling, and exacerbates water scarcity[81]. Additionally, the findings indicate a projected increase in minimum temperatures in the coming years. Higher winter temperatures will disrupt the seasonal balance, shortening the cooler periods vital for agricultural cycles, natural ecosystems, and public health[25]. The loss of this seasonal equilibrium is expected to trigger significant changes in natural cycles and agricultural patterns across the region[29]. These findings underscore the urgent need for climate adaptation strategies, water resource management, and greenhouse gas emission reduction to mitigate the growing impacts of climate change. Rising minimum temperatures also present serious challenges for climate adaptation in the region. Increasing minimum temperatures, particularly under the SSP5-8.5 scenario, exacerbate climate anomalies by disrupting seasonal cycles and intensifying extreme climate events, such as more frequent warm nights and shorter cooler periods[28]. This trend leads to heightened energy consumption

additional stress on water and energy infrastructures, and poses significant risks to public health through prolonged heatwaves and warm nights, severely impacting human health and ecosystems[82–84]. Long-term climate policies, particularly in response to the SSP5-8.5 scenario, must become a regional priority to address these emerging challenges. Additionally, as highlighted by ref. 85, integrating the prediction of heatwave-related mortality into climate models can improve the forecasting of extreme temperature impacts on public health. This method further supports the need for adaptive management in vulnerable regions like the Middle East, where rising temperatures could substantially increase mortality rates due to heat stress.

The analysis of precipitation patterns under different climate scenarios further emphasizes the critical risks facing the Middle East. A significant reduction in rainfall, especially under the SSP5-8.5 scenario, suggests an intensification of seasonal droughts and greater variability in precipitation, leading to severe challenges such as water resource depletion, reduced agricultural productivity, and increased environmental hazards[79]. Long-term projections, such as those by ref. 86, support these concerns, showing that the region's arid and semi-arid areas will significantly reduce precipitation and soil moisture, further straining water resources. Moreover, the increased variability and frequency of extreme precipitation events under this scenario highlight the need for comprehensive water resource management and climate adaptation planning[87]. Even under the SSP2-4.5 scenario, notable changes in precipitation patterns raise concerns about heightened risks of flash floods and other climate-induced challenges. For example, Kassaye et al.[88] concluded that under the SSP2-4.5 scenario, studies project a general increase in streamflow magnitude, with yearly flow expected to rise by 4.8% during the mid-term (2041–2070). These trends make it clear that urgent climate action and targeted water management strategies are essential to mitigate the long-term impacts of climate change across various scenarios, particularly the pessimistic SSP5-8.5[82].

Given the significant impacts of climate change in the Middle East, identifying the most vulnerable areas is crucial for targeted adaptation and resource management. Initially, a trend analysis using the Mann-Kendall test was conducted to identify these critical areas. However, after calculating the trends, it became clear that under the SSP5-8.5 scenario, the entire Middle East exhibits consistent upward trends in minimum and maximum temperatures and a downward trend in precipitation, making it impossible to distinguish specific vulnerable areas based solely on trend analysis. Therefore, one of the most effective techniques in this context, Hotspot comparison analysis, was applied[89]. This method identifies "hot" and "cold" spots and tracks their shifts over time. In this study, these areas were found to be statistically significant at the 99% confidence level, with at least two intersecting climatic elements, helping policymakers prioritize regions where the impacts of climate change are likely to be most severe. The spatial distribution of sensitive areas in the Middle East under various climate scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5) projected for 2099 offers critical insights into the future climate risks faced by the region. Analysis of these projections reveals five categories of risk areas: (1) regions with high maximum and minimum temperatures alongside low precipitation, (2) regions exposed to high maximum and minimum temperatures, (3) areas vulnerable due to high maximum temperatures combined with reduced precipitation, (4) regions at risk from high minimum temperatures coupled with low precipitation, and (5) regions projected to experience increased precipitation. Under the SSP1-2.6 scenario (Fig. 8a), southern Saudi Arabia, UAE, Yemen, and Oman regions are high-risk, particularly concerning the highest maximum and minimum temperatures. This suggests that, even under a low-emission scenario, these areas will face elevated temperatures, potentially exacerbating heat stress, agricultural challenges, and water scarcity[90]. These findings align with projections by Mora et al.[81] and Lima et al.[91], which indicate an increase in both the frequency and intensity of heatwaves in the Middle East, even under low-emission scenarios. The most critical areas are concentrated in southeastern Saudi Arabia, northern Oman, and eastern Yemen, where all three climatic variables reach their most extreme levels.

Additionally, coastal vulnerability is observed in eastern Egypt, attributable to high maximum and minimum temperatures and a lack of precipitation. Conversely, northern Iran, parts of the Zagros Mountains, and eastern Turkey are projected to experience increased rainfall. In the SSP2-4.5 scenario (Fig. 8b), the expansion of high-risk areas becomes more pronounced, particularly in southern and central Saudi Arabia, Qatar, the UAE, Oman, western Yemen, Iran, and Egypt. The findings indicate that in this scenario, an intensification of extreme heat conditions and a further reduction in precipitation are expected, especially in southern regions. These findings corroborate research by Hamed et al.[92], who reported that CMIP6 models project a decrease in precipitation, and by Malik et al.[93], who noted a significant increase in temperature in arid and semi-arid regions under moderate-emission scenarios. Critical areas concerning maximum and minimum temperatures, previously limited to southern Iran, have expanded considerably in this case.

The most vulnerable regions are in western Yemen, where extreme temperatures coincide with diminished rainfall. As projected in the SSP1-2.6 scenario, most parts of the Middle East are expected to experience a decrease in precipitation. According to the research by Terink et al.[94], most countries in the Middle East are expected to experience a reduction in annual precipitation. In particular, a 15–20% decrease in rainfall is projected during 2040–2050. Regions such as southern Egypt, Saudi Arabia, Syria, and eastern Iran are expected to face the most significant reductions in precipitation.

On the other hand, increased precipitation is expected for northern Iran, areas along the Caspian Sea, the Zagros Mountains, and eastern Turkey. This finding is consistent with previous research indicating a shift toward heavier rainfall events in these areas[95–97]. According to these studies, the western and northern parts of Iran are expected to encounter significant increases in precipitation. Abbaspour, Faramarzi, Ghasemi and Yang[97] suggest that northern and western Iran will experience greater rainfall and a higher likelihood of large and severe floods in these regions under various climate change scenarios.

Similarly[95], found that the number of days with heavy (R10mm) and heavy rainfall (R20mm) will significantly increase, particularly along the southern Caspian Sea coast and the Zagros Mountains. In alignment with these findings, our results also corroborate Sarış[98], who highlighted the increased susceptibility of the Black Sea region to heavy precipitation. Furthermore, Majdi et al.[99] reported a decrease in precipitation, ranging between 5 and 133 mm on average, across most parts of the Middle East and Mediterranean regions, based on an analysis of 23 GCMs. This variability across regions underscores climate change's complex and multifaceted impacts on rainfall patterns in the broader region. Similarly, the SSP2-4.5 scenario also anticipates increased precipitation in some parts of Iran and Turkey. The potential for future increases in precipitation in certain areas of the Middle East may be explained by a theory put forward by Francis and Vavrus[100], which links climate change to extreme weather patterns in mid-latitude regions. This theory suggests that weakened zonal winds, coupled with increased wave amplitude, contribute to the slower movement of Rossby waves, subsequently increasing the likelihood of extreme weather events[101].

The SSP5-8.5 scenario (Fig. 8c) presents the most severe and widespread climate impacts. The southern Middle East, including Saudi Arabia, Oman, Yemen, Iraq, and Iran, is predominantly characterized by extreme heat, with substantial consequences for agriculture, water resources, and human habitability. In this scenario, regions experiencing critical maximum and minimum temperatures and precipitation reduction, which were previously confined to lower latitudes, now extend to Kuwait and southeastern Iraq. These findings align with previous research by Black et al.[102], which projects a decrease in precipitation and an increase in temperatures in the southern Middle East under greenhouse gas emission scenarios. The study also predicts a shift toward polar storm tracks and the weakening of Mediterranean storm systems, potentially leading to reduced winter rainfall in the region. According to research conducted by Almazroui et al.[103], it is projected that by the end of the 21st century, under the RCP8.5 scenario, temperatures in the central and southern regions of the Arabian Peninsula

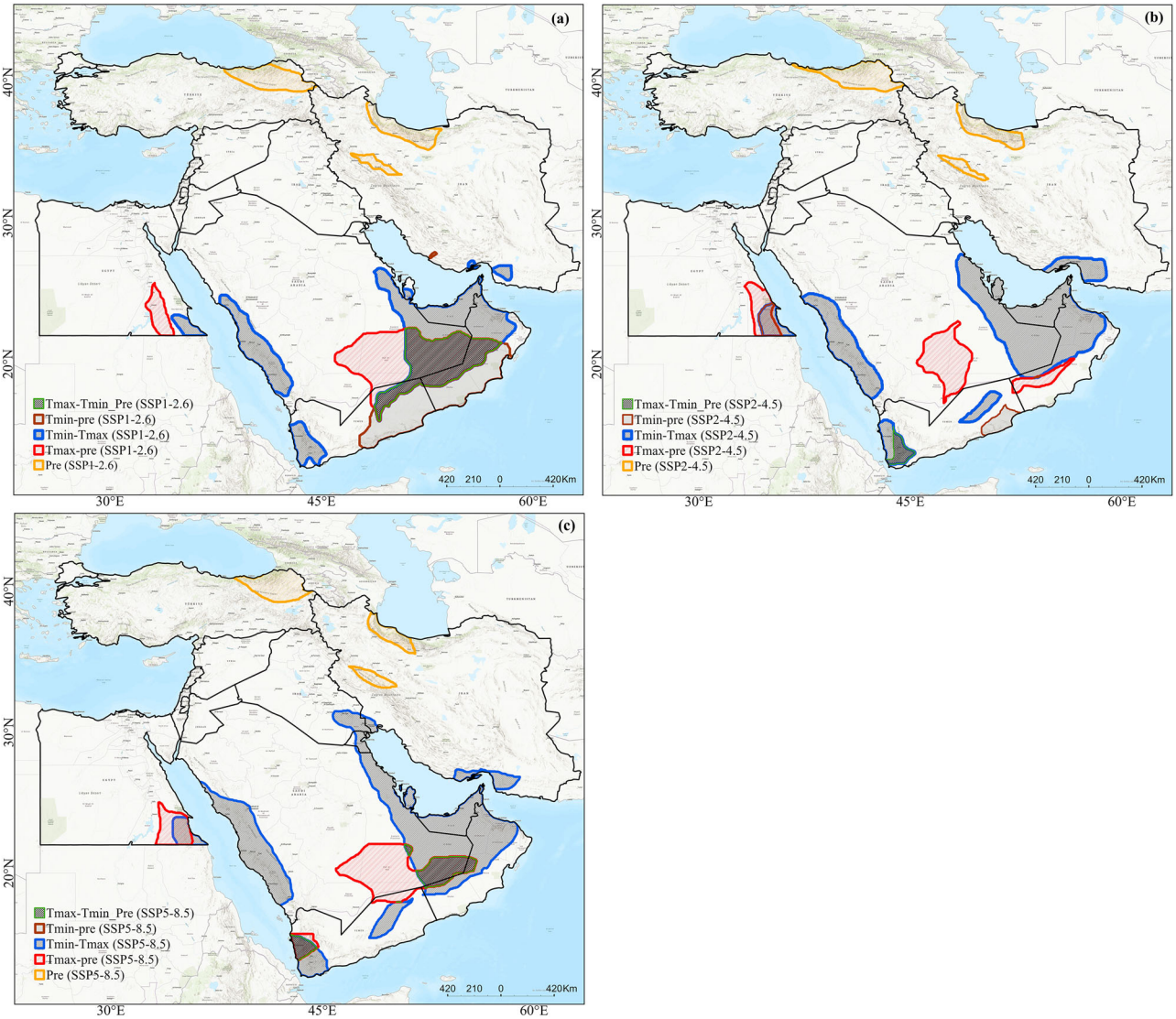

Fig. 8 | Spatial distribution of critical climate-vulnerable areas in the Middle East under three SSP scenarios (SSP1–2.6, SSP2–4.5, and SSP5–8.5) for the year 2099, based on Hotspot Comparison Analysis. Maps (**a**–**c**) illustrate the intersection of significant climatic stressors including maximum temperature, minimum temperature, and precipitation using Getis-Ord Gi* analysis. Colored zones represent areas where at least two extreme climatic variables coincide, identifying the regions at greatest risk from combined heat and moisture stress. Red and orange boundaries indicate compounding temperature extremes, blue.

will increase by up to 6.4 °C. In contrast, the northern regions will experience a temperature rise of 5.2 °C. These projections further corroborate our findings, suggesting a significant rise in temperature in areas such as Kuwait and southeastern Iraq and highlighting the broader trend of increasing temperatures across the Arabian Peninsula. Similarly, Varela et al.[104] concluded that the eastern Arabian Peninsula and North Africa will be among the most affected areas, with extreme temperatures occurring over 80% of days.

In summaty, this study presents a novel ensemble ML framework, Stacking-EML, aimed at enhancing climate forecasting for the Middle East by utilizing CMIP6 datasets. With its diverse climate and geography, the Middle East faces major ecological challenges due to climate change. These challenges include rising temperatures, changes in precipitation trends, and more frequent severe weather events. Accurately predicting and assessing these climate impacts is crucial for effective adaptation and mitigation strategies. Our methodology seeks to address this need by combining various ML algorithms to improve the accuracy and robustness of climate variable forecasts against varying input data and different climatic conditions. The Stacking-EML model was developed through a comprehensive, multi-phase framework. Initially, five distinct ML algorithms, including Random Forest, XGBoost, LGBM, SVM, and CatBoost, were trained using monthly datasets covering precipitation, maximum temperature, and minimum temperature from selected CMIP6 GCMs for 2015–2100. Among these, LGBM and RF showed superior performance, particularly in capturing the variability in climate data across the Middle East. These models were then combined into a meta-model using a stacking ensemble strategy, with an ANN acting as the most effective regressor in the final meta-model. The Stacking-EML framework significantly improved compared to individual models and traditional CMIP6 outputs. Evaluation criteria, including RMSE and R^2, showed enhanced predictive accuracy for precipitation and temperature extremes. Incorporating geographical and topographical factors, such as longitude, latitude, elevation, slope, and aspect, further optimized the model's performance, highlighting the importance of integrating diverse data sources in climate modeling. Spatial analyses using Hotspot Analysis revealed critical insights into expected climate changes under different SSP scenarios. The model forecasts a considerable increase in maximum and minimum temperatures across the Middle East, particularly under the high-emission SSP5-8.5 scenario. This scenario predicts extreme temperature increases, with summer maximum temperatures exceeding 45 °C and a notable disruption of seasonal

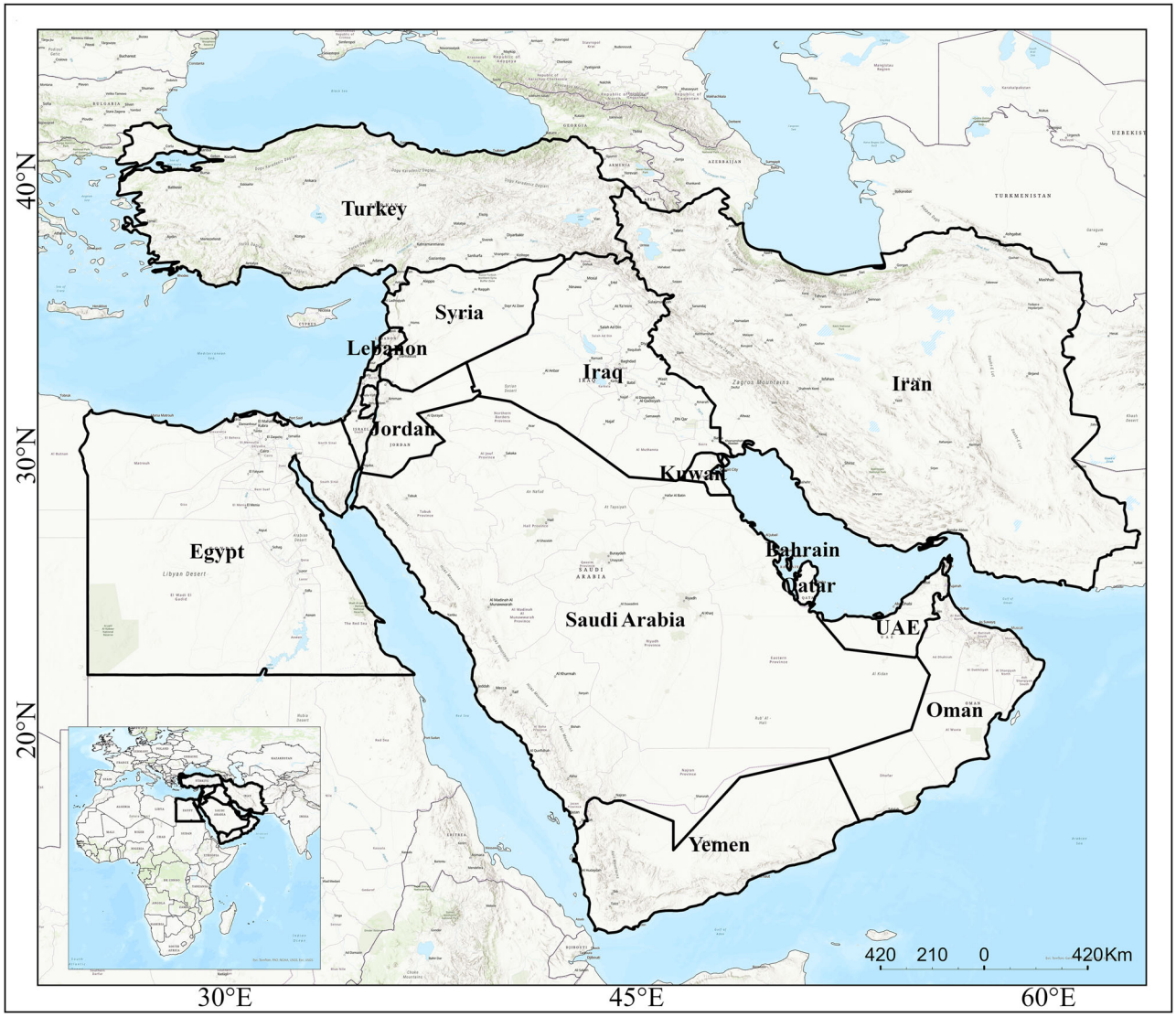

Fig. 9 | Geographic location and extent of the study area.

temperature balance. Precipitation trends are expected to show substantial variability and reductions, especially during summer, leading to intensified drought conditions and greater climate variability.Identifying vulnerable areas through hotspot and cold spot assessments highlights regions likely to experience the most severe impacts of climate change. Southern areas, including Saudi Arabia, Oman, Yemen, and parts of Iran and Egypt, are projected to face extreme heat scenarios and reduced precipitation, potentially exacerbating water scarcity, agricultural productivity, and public health issues. In contrast, northern regions such as Turkey and northern Iran may see increased precipitation, presenting distinct challenges such as flooding and ecosystem disruptions.

Methods

Study area

The Middle East (Fig. 9) is a geographically and climatically diverse region encompassing sixteen countries, including those of the Arabian Peninsula (Oman, the United Arab Emirates, Bahrain, Saudi Arabia, Kuwait, Yemen, and Qatar), as well as Jordan, Syria, Iran, Palestine, Iraq, Turkey, Egypt, and Lebanon[105]. It spans an area of over 69 million km^2 and hosts a wide range of climatic conditions, from the extremely arid deserts of the Arabian Peninsula and parts of Egypt to the semi-arid or more temperate highlands of Turkey and the mountainous regions of Iran[106]. Average annual precipitation typically falls below 250 mm in vast portions of the region; however, areas adjacent to the Mediterranean and Caspian Seas and mountainous corridors such as the Zagros and Taurus ranges can receive substantially higher rainfall[74–76,107]. Temperature extremes are equally striking, with summer highs frequently surpassing 50 °C in the Arabian Peninsula and winter lows dropping below freezing in upland areas[93]. Such spatial heterogeneity and pronounced aridity pose major challenges for climate modeling, particularly given limited in-situ data coverage, complex topographical influences, and rapid demographic expansion in water-scarce zones[77,106]. This study's focus on the Middle East is motivated by its acute vulnerability to climate-related stressors, including heatwaves, rainfall deficits, and water resource depletion, all of which are exacerbated by climate change[76]. The region's significant temperature gradient from intensely hot and hyper-arid conditions along the southern Arabian coasts and deserts to cooler, more humid pockets in northern highlands creates an ideal testbed for evaluating ensemble machine learning approaches that aim to capture the complexity and extremes of regional climatic behaviors. By systematically incorporating both arid lowlands and relatively more temperate uplands, the present framework provides an opportunity to assess the robustness of our Stacking-EML methodology across a broad range of climate profiles.

Methodological framework

The comprehensive framework employed in this study is illustrated in Fig. 10. The methodology is structured into five phases: Phase I involves

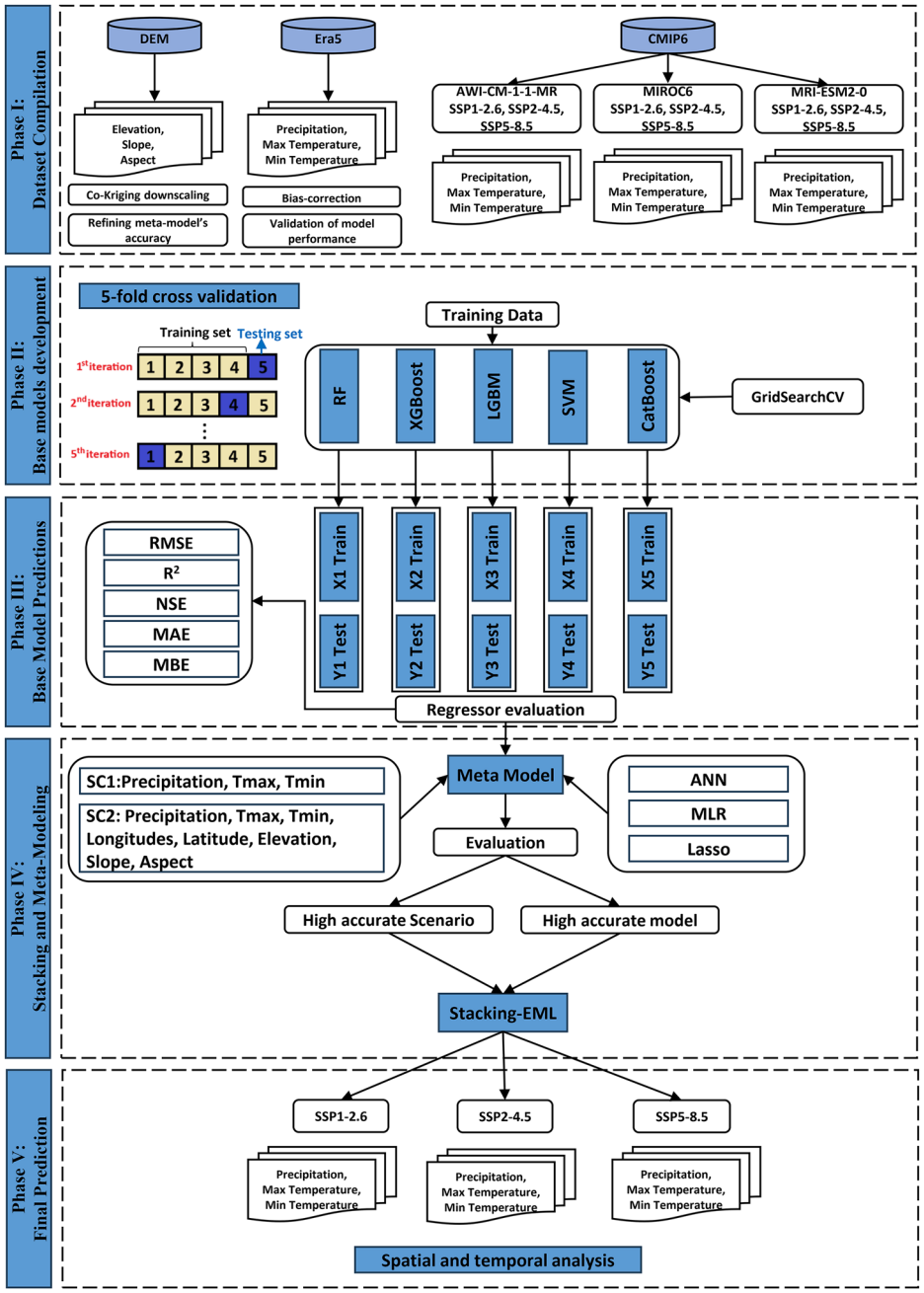

Fig. 10 | Proposed stacking ensemble model workflow for climate projections (Max Temperature, Min Temperature, and Precipitation) in the Middle East.

compiling the Model Dataset, Phase II, Base ML Model Development, entails training five machine learning algorithms, including RF, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Support Vector Machine (SVM), and CatBoost. These models predict the input data in Phase III, Base Model Predictions. Phase IV, Stacking and Meta-Modeling, focuses on creating the Meta Model. Finally, Phase V emphasizes predicting input data using the Stacking-EML approach.

Phase I: Dataset compilation

Phase I involves compiling the Model Dataset, which includes monthly data on precipitation, maximum temperature, and minimum temperature from CMIP6, simulated by selected Global Climate Models (GCMs) for 2015–2100. Additionally, ERA5 reanalysis data (1995–2014) are employed for bias-correcting the historical portion of the GCM outputs and for validating model performance, while the digital elevation model (DEM) provides critical topographical features (elevation, slope, and aspect) that support both the co-Kriging downscaling step and the subsequent refinement of predictions in the meta-model. The datasets employed in this study are outlined as follows:

CMIP6

The monthly maximum temperature, minimum temperature, and precipitation data of CMIP6 simulated by some selected GCMs over 2015–2100 are obtained from the Earth System Grid data distribution portal (https://cds.climate.copernicus.eu/). CMIP6 uses future scenarios to examine how climate change might occur under different greenhouse gas emission scenarios[108]. SSP1-2.6, SSP2-4.5, and SSP5-8.5 climate scenarios represent future global climate change projections until 2100. SSP1-2.6 represents an optimistic scenario where global CO_2 emissions drop to zero by 2050, aligning with the Paris Agreement's goal to limit global warming to 1.5 °C above pre-industrial levels, with a radiative forcing of 2.6 W/m² and

Table 3 | Information on the three cmip6 global climate models employed in this study

Model Name	Country	Institution	Resolution (Lat*Lon)	Reference
AWI-CM-1-1-MR	Germany	Alfred Wegener Institute	1.1° × 1.1°	Semmler, Danilov, Gierz, Goessling, Hegewald, Hinrichs, Koldunov, Khosravi, Mu and Rackow[182]
MIROC6	Japan	Japan Agency for Marine-Earth Science and Technology, Japan	1.4° × 1.4°	Shiogama[183]
MRI-ESM2-0	Japan	Meteorological Research Institute (MRI) of the Japan Meteorological Agency (JMA)	1.125° × 1.125°	Yukimoto, Kawai, Koshiro, Oshima, Yoshida, Urakawa, Tsujino, Deushi, Tanaka and Hosaka[184]

temperatures stabilizing at 1.4 °C by 2100[12]. In contrast, SSP2-4.5 is a moderate scenario with intermediate climate mitigation and adaptation efforts, leading to a radiative forcing of 4.5 W/m² by 2100. It aligns somewhat with the Paris Agreement, predicting a warming of ~2.7 °C by the end of the century[109]. On the other hand, SSP5-8.5 outlines a high-emission scenario with significant climate change mitigation challenges and minimal adaptation issues, leading to a radiative forcing of 8.5 W/m² by 2100. Without additional climate policies, it predicts a 4 °C increase in global temperatures by the end of the century, updating the CMIP5 RCP8.5 scenario with socioeconomic factors[110,111].

This study utilizes monthly outputs from three CMIP6 Global Climate Models (GCMs), namely AWI-CM-1-1-MR, MIROC6, and MRI-ESM2-0 for the period 2015–2100 under three Shared Socioeconomic Pathway (SSP) scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5) (Table 3). These GCMs were selected based on four principal criteria: (i) spatial resolution (nominal 100–250 km), which balances computational efficiency with adequate representation of regional climate processes[112], (ii) scenario availability for all three SSP pathways, essential in assessing a broad range of future emission trajectories, (iii) proven skill in precipitation simulation over arid and semi-arid domains, as demonstrated in prior regional studies[52], and (iv) temporal coverage and consistency with a baseline observational period (1995–2014) to facilitate bias correction against ERA5 reanalysis data[25].

The raw GCM outputs frequently exhibit systematic biases, particularly in topographically complex or data-sparse regions[113]. To address these discrepancies and reconcile the varying spatial resolutions (GCM nominal resolutions of 1.1°–1.4° vs. ERA5 at 0.1° and the high-resolution DEM), a co-Kriging downscaling approach was employed, and the GCM outputs were first spatially resampled to a consistent 0.5° × 0.5° grid. Specifically, a co-Kriging method was used for this spatial interpolation, leveraging topographical features (elevation, slope, and aspect) from the DEM as secondary variables. This approach improves upon ordinary Kriging by incorporating additional geographic information, thereby refining estimates in complex terrains and preserving local variance more effectively[114]. Moreover, Co-Kriging offers a significant advantage in regions with complex topography or sparse observational data by leveraging secondary topographical variables. This method refines spatial patterns in climate simulations by incorporating real-world geographical dependencies, reducing systematic interpolation errors that may arise in simpler spatial resampling methods. Co-Kriging is particularly advantageous for downscaling climate variables in mountainous or data-sparse areas, as it leverages correlations with secondary variables to achieve higher spatial accuracy[115]. Following the co-Kriging downscaling, a two-step bias correction (Linear Scaling and Quantile Mapping) was applied for bias correction[116]. The bias correction process reduces discrepancies between GCM outputs and observed data, facilitating alignment between modeled and observed distributions. While this process cannot eliminate structural uncertainties arising from the incomplete simulation of certain physical processes in climate models, its effectiveness in improving accuracy and consistency with real-world data has been demonstrated in various studies[113]. The combination of Co-Kriging and bias correction in this study provides a comprehensive and efficient approach to enhancing the accuracy of climate data. This method corrects systematic biases in GCM outputs and offers greater flexibility in downscaling data to more accurately represent regional climate variations. Previous research has shown that such an integrated approach significantly improves prediction accuracy and enhances the agreement between simulated and observed values[117]. Therefore, while uncertainties are an inherent part of climate modeling, adopting this combined approach effectively minimizes errors and improves the regional representation of climate data[118]. Prior validation studies confirm that AWI-CM-1-1-MR, MIROC6, and MRI-ESM2-0 exhibit robust skill in simulating temperature trends, precipitation variability, and key climatological extremes over the Middle East[119–121]. Comparative evaluations with broader CMIP6 ensembles suggest that these models effectively capture seasonal precipitation cycles, including winter rainfall peaks in northern Middle Eastern regions, and reproduce long-term warming trends consistent with multi-model assessments[92]. Furthermore, their skill in replicating historical drought frequencies and heatwave intensities has been established in prior regional assessments, reinforcing their credibility for future projections[122]. Despite these advantages, it is important to note that any downscaling procedure may introduce additional uncertainties, particularly in areas where observational data are sparse or topographic gradients are highly variable.

Reanalysis data. ERA5, the fifth generation of reanalysis products developed by the ECMWF, is employed to reanalyze global atmospheric changes using both models and observational data. Data from multiple sources, including satellites, meteorological stations, and aircraft, are combined with advanced physical models to produce high-precision spatiotemporal distributions of various meteorological variables on a global scale[123]. Significant improvements over its predecessor, ERA-Interim, include enhanced spatial resolution (0.25° compared to 0.75° for ERA-Interim), the incorporation of a substantially larger volume of observations for data assimilation, improved representation of radiative forcing (including sulphate aerosols from volcanic eruptions), a better global balance between precipitation and evaporation, and more accurate sea surface temperature and sea ice coverage[124,125]. For this study, ERA5 precipitation and temperature data from 1 January 1995 to 31 December 2014, with a spatial resolution of 0.1° × 0.1° and monthly temporal resolution, were obtained from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/).

Digital elevation model. In this study, the Digital Elevation Model (DEM) was employed to capture the topographical features of the study area, including elevation, slope, and aspect. The high-resolution SRTM15 + DEM, with a spatial resolution of 15 arc seconds (~500 m at the equator), was used to provide detailed bathymetry and topography data. This model integrates over 33.6 million measurements from sources such as the National Geospatial-Intelligence Agency and Scripps Institution of Oceanography. Onshore topography data are primarily derived from SRTM-CGIAR V4.1, ArcticDEM above 60°N, and the Reference Elevation Model of Antarctica below 62°S[126]. These DEM-based variables were instrumental in the co-Kriging downscaling step and were introduced later in Phase IV to refine the meta-model's predictive accuracy.

Phase II: Base model development

Phase II focuses on developing base models using various ML algorithms. This phase aims to create models that can accurately predict climate

Table 4 | The five individual ML models employed in this study

Model	Strengths	Weaknesses	Reference
RF	Handles non-linearity, robust to overfitting, interpretable feature importance	Computationally expensive for large datasets	Breiman, L.[185]
XGBoost	Fast computation, reduces overfitting with regularization, highly efficient.	Sensitive to hyperparameters, requires fine-tuning	Li, P.[142]
LGBM	Faster training, memory-efficient, good for large datasets	Less interpretable than RF, sensitive to outliers	Ke et al.[147]
SVM	Effective in high-dimensional spaces, works well with small datasets	Computationally expensive for large datasets requires careful kernel selection	Vapnik, V.[186]
CatBoost	Handles categorical features well, reduces overfitting, good for imbalanced data	Slower training compared to LGBM and XGBoost	Prokhorenkova et al.[156]

variables, such as maximum temperature, minimum temperature, and precipitation, based on historical data from different scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5) derived from the CMIP6 climate models (AWI-CM-1-1-MR, MIROC6, MRI-ESM2-0). To ensure the reliability of training data, ERA5 reanalysis data was used as a reference for bias correction of the CMIP6 model outputs before being employed in ML model training. The reanalysis dataset provides high-resolution climate variables, reducing systematic biases in raw GCM simulations and ensuring that training data better represents observed climate conditions. ERA5-adjusted climate variables served as input features for training all base ML models, thereby improving model generalization. In this phase, various ML algorithms are employed to develop base models, each designed to effectively capture the complex relationships in climate data. These models were selected based on their proven ability to process nonlinear climate dynamics, handle high-dimensional datasets, and provide reliable predictions in climate modeling applications[127]. We adopted a structured feature selection approach to enhance the models' predictive performance. RF and XGBoost were used to assess the importance of various predictors, given their proven ability to rank features effectively in complex environmental datasets[128].

Based on these analyses, temperature, precipitation, elevation, slope, and aspect emerged as the most influential variables, ensuring spatial consistency and optimizing model outcomes. By contrast, initially considered a potential predictor, wind speed had a negligible impact on overall performance and was thus excluded from the final feature set. Focusing on these high-contribution predictors allowed a clearer understanding of how climate variables interact with geographic factors, ultimately improving forecast accuracy. Subsequently, we partitioned the dataset using an 80-20 split, with 80% of the data allocated for training and 20% reserved for validation. In addition, a stratified 5-fold cross-validation procedure was implemented to reduce predictive variance further and bolster model generalization. This method strikes a pragmatic balance between computational efficiency and dependable performance estimation in datasets of moderate size, where larger k-values often yield diminishing returns while substantially increasing computational overhead. Model hyperparameters were optimized using GridSearchCV, which systematically evaluates different parameter configurations to identify the best fit for each base algorithm. This tuning process was conducted independently for every model, ensuring that each algorithm performed optimally. Once trained, the base models produced predictive outputs covering key variables like temperature and precipitation that were consolidated into a structured dataset, ready for integration into the subsequent meta-modeling phase.

Base ML models. ML offers promising algorithms for analyzing complex environmental phenomena and climate studies[129,130]. Over recent decades, it has been extensively utilized for predicting and forecasting various climate parameters[131–135]. ML models' superior performance and ability to handle large datasets make ML a popular and practical alternative to traditional statistical methods for predicting climatic and complex variables[133,136]. Considering the critical role of ML in enhancing the accuracy of climatic variable predictions, this research applied a range of advanced ML algorithms, including RF, XGBoost, LightGBM, SVM, and CatBoost. These models were chosen for their strong predictive capabilities and suitability in managing complex environmental datasets. Table 4 provides a comparative analysis of these ML models, outlining their strengths and weaknesses. Each algorithm was trained on a portion of the dataset to maximize predictive accuracy.

Random forest

Random forest is a sophisticated ML algorithm that combines tree-based classifiers and is known for its high accuracy in classification, prediction, and regression tasks[137]. This algorithm constructs an ensemble by averaging outputs from multiple trees, and each is created using bootstrap samples from the training data. At each node of the trees, a random selection of predictor variables is assessed, introducing diversity and reducing correlation among the trees. RF's notable resistance to overfitting and its robustness in handling noisy data and irrelevant features make it particularly effective. This has resulted in its outperforming traditional ML models, especially in climate studies and environmental research, where it has found widespread application[138–141].

XGBoost

XGBoost, an optimized version of Gradient Boosting (GB), is widely regarded for its use in optimal classification trees[142]. It effectively addresses overfitting by reducing model complexity, enhancing classification accuracy, and decreasing computation time through high-speed analysis[143,144]. Unlike traditional GB, which relies solely on decision trees, XGBoost incorporates classification trees and linear regression, making it highly effective in processing sparse data through parallel computing[145,146].

LGBM

LGBM is a decision tree-based gradient boosting framework that utilizes boosting techniques[147]. Unlike XGBoost, LGBM employs a histogram-based algorithm, which accelerates training, reduces memory usage, and utilizes a leaf-wise growth strategy with depth constraints[148,149]. The histogram algorithm discretizes continuous floating-point values into bins, constructing a histogram without needing additional storage for pre-sorted data. This approach allows the model to reduce memory consumption significantly, ~1/8 of the original, without compromising accuracy.

SVM

Support Vector Machines (SVMs) extensively utilize supervised learning algorithms that employ linear statistical functions for regression and classification tasks[150]. High levels of accuracy are achieved by SVMs even when dealing with limited data, owing to their maximal-margin classification approach. Input vectors are mapped into an infinite-dimensional feature space, where nonlinear transformations are used to construct an optimal hyperplane that maximizes class separation[151]. The performance of SVMs is significantly influenced by the choice of kernel functions, which include polynomial, sigmoid, radial basis function (RBF), and linear kernels. The RBF kernel is commonly employed among these, particularly in flood vulnerability assessments[152].

CatBoost

CatBoost is a gradient-boosting decision tree model known for its exceptional performance as an individual and a meta-model in ensemble methods[153]. It efficiently handles categorical features, reducing information loss and mitigating overfitting through a random permutation method for selecting tree structures[154,155]. This capability makes CatBoost highly suitable for real-world applications, such as predicting soil water content using multi-sensor data[156].

Phase III: base model predictions

Phase III focuses on generating initial predictions by applying the trained base models from Phase II to the test data for forecasting climate variables. A stratified fivefold cross-validation approach was employed to ensure robust performance, facilitating model generalization and minimizing predictive variance. Hyperparameter optimization was conducted using GridSearchCV, systematically identifying the best configurations for each model. Details on data partitioning and validation methodology are provided in Section 3.2. Once trained, the base models produced predictive outputs for key climate variables, which were subsequently integrated into the meta-modeling phase. This technique reduces the prediction variance and helps prevent overfitting by validating the model across multiple data partitions[157]. After generating predictions, evaluation metrics such as the Root Mean Squared Error (RMSE), the coefficient of determination (R^2), Nash-Sutcliffe Efficiency (NSE), and bias metrics like Mean Bias Error (MBE) are calculated to assess model performance. These results identify the best-performing models, which will be refined and combined in the next phase to create a more robust ensemble model. This approach leverages the strengths of each base model, enhancing overall predictive accuracy and robustness.

Phase IV: Stacking and meta-modeling

Phase IV enhances predictive accuracy by constructing a meta-model that integrates the predictions from the two best-performing base models identified in Phase III. This approach utilizes stacking, a well-known ensemble learning technique originally proposed by Wolpert[158], which leverages the complementary strengths of multiple models to improve overall prediction accuracy. Stacking combines the outputs of several base models as input features for training a meta-model, which produces the final predictions for the target variable. The meta-model is trained to effectively integrate the base models' predictions, enhancing overall performance and generalization capability[159]. In this phase, additional geographical and topographical features, including elevation, slope, aspect, longitude, and latitude, were introduced as supplementary predictors in the meta-model to improve spatial precision and ensure that predictions account for topographic influences on climate variables while preventing multicollinearity issues. The meta-learning model at this stage evaluates whether topographical factors significantly improve predictive performance, allowing for a more adaptive and dynamic modeling approach. The stacking framework consists of multiple base models performing classification or regression tasks on the training dataset, with a meta-learner refining these predictions to surpass the accuracy of any single model.

To mitigate overfitting and ensure robust performance, leave-one-out cross-validation was applied, enabling the meta-learner to effectively leverage the complementary characteristics of the base models[160]. Unlike the base models, which primarily focused on climate variables, the meta-model incorporated topographical features at this stage to refine the final predictions and improve spatial consistency in projections. In this study, two scenarios are considered in the meta-modeling process: Scenario 1 (SC1) utilizes precipitation (P), maximum temperature (Tmax), and minimum temperature (TMin) as input variables, while Scenario 2 (SC2) expands this input set by incorporating additional geographical and topographical features, including longitude, latitude, elevation, slope, and aspect. The meta-learning model typically employs a weighted average approach or a linear regression model[161,162]. However, this study evaluated three regressors, including ANN, MLR, and LASSO, with the best-performing regressor selected to construct the final meta-model. These models were chosen due to their distinct capabilities in handling climate data and improving predictive accuracy: ANN was selected for its ability to model nonlinear relationships and capture complex interactions between variables[163]. MLR served as a benchmark linear model to assess the performance of more advanced approaches[164]. LASSO was incorporated for its ability to enhance model sparsity by selecting the most relevant predictors, thereby reducing overfitting[165].

Multiple linear regression. Multiple Linear Regression is a widely used statistical method for regression analysis that describes the relationship between one dependent variable and two or more independent variables by fitting a linear equation[166]. Climate studies frequently employ this approach for downscaling and impact analysis[167–169]. Mathematically, the general form of MLR can be expressed as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \quad (1)$$

where: y is the dependent variable, *x1, x2,..., xn* are the independent variables, $\beta_0, \beta_1, \ldots, \beta_n$ are the parameters (coefficients) to be estimated, and ϵ represents the error term.

Lasso regression. Traditional regression methods commonly utilize the Ordinary Least Squares (OLS) algorithm to estimate the coefficient β by minimizing the residual sum of squares (RSS) in the model[170]:

$$\left(\hat{\alpha}^{(ols)}, \hat{\beta}^{(ols)}\right) = \arg_{(\alpha,\beta)} \min \sum_{i=1}^{n} \left(y_i - \alpha - \sum_{j=1}^{p} x_{ij}\beta_i\right)^2 \quad (2)$$

While OLS coefficients are generally considered the best-unbiased estimators, this approach is based on the assumption of linearity, which can lead to excessively high variance in the estimates. Moreover, OLS often struggles with accurately estimating and interpreting models, particularly when applied to high-dimensional datasets. Tibshirani[171] introduced the LASSO algorithm to address these limitations, which enhances the OLS method by incorporating an L1 penalty function on β. This addition alters the OLS model, transforming it into an unconstrained quadratic programming problem that improves model performance by managing multicollinearity and facilitating variable selection. The modified objective function for LASSO is given as:

$$\left(\hat{\alpha}^{(lasso)}, \hat{\beta}^{(lasso)}\right) = \arg_{(\alpha,\beta)} \min \sum_{i=1}^{n} \left\{ \left(y_i - \alpha - \sum_{j=1}^{p} x_{ij}\beta_i\right)^2 + \lambda \sum_{j=1}^{p} |\beta_i| \right\} \quad (3)$$

where λ is the penalty coefficient that controls the extent of shrinkage applied to the coefficients. The L1 penalty term encourages sparsity in the model by shrinking some coefficients to zero, allowing the model to select the most relevant predictors, thereby enhancing both predictive accuracy and interpretability. The LASSO algorithm is now commonly employed in climate and environmental studies (Hammami et al.[172]).

Artificial Neural Network (ANN). ANN is a widely used ML algorithm known for its short computational time and capability to model complex nonlinear functions[173]. Inspired by the human brain, ANN is a data-driven model designed to establish relationships between input and output variables[174]. It consists of multiple layers, including input, hidden, and output, each containing neurons that process data. Due to its flexibility and strong performance, ANN has been applied to model various parameters (Ouarda and Shu et al.[175]). The relationship between inputs (x) and output (Y) in an ANN is represented by the equation[176]:

$$Y = f(w_1 x_1 + w_2 x_2 + \ldots + w_n x_n + b) \quad (4)$$

Table 5 | Hyper-parameters of ML models adjusted in this study

Algorithm	Hyper-parameters	Explanation	Grid search values
Random Forest (RF)	n_estimators	Number of trees in a forest	100, 150
	max_depth	Highest depth of the tree	10, 15
Support Vector Machine (SVM)	C	Penalty parameter	0.1, 1, 10
	gamma	Bandwidth parameter	0.01, 0.1, 1
	kernel	Kernel function	RBF
LightGBM (LGBM)	n_estimators	Number of trees in a forest	100, 150
	learning_rate	Learning rate	0.01, 0.1
	num_leaves	Number of leaves in one tree	31, 50
	max_depth	Highest depth of the tree	−1, 10
XGBoost (XGB)	n_estimators	Number of trees in a forest	100, 150
	learning_rate	Learning rate	0.01, 0.1
	max_depth	Highest depth of the tree	3, 5
CatBoost (CB)	iterations	Number of boosting iterations	100, 150
	learning_rate	Learning rate	0.01, 0.1
	depth	Depth of the tree	4, 6

where f is the activation function, b is the bias, and Wn represents the weights of the connections.

Phase V: final prediction

In the final prediction phase, the Stacking-EML framework developed in the previous stage is utilized to generate forecasts for three essential climate variables, including maximum temperature, minimum temperature, and precipitation across three different scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5). These forecasts provide valuable insights into future climatic patterns, facilitating informed decision-making and strategic planning.

Assessment of algorithm performance

In this study, several evaluation criteria, including RMSE, R^2, MAE, MBE, and NSE, were used to assess the performance of the ML algorithms. RMSE measures the accuracy of the model's predictions, quantifying the average magnitude of the prediction errors[177]. The R^2 ranges between 0 and 1, indicating how closely the observed values align with the fitted regression line. It represents the proportion of variance in the observed data explained by the model, providing insight into the model's explanatory power. MAE calculates the mean absolute difference between observed and predicted values, representing the average error magnitude without considering the direction of errors. It provides a straightforward measure of model accuracy, with lower values indicating better performance[178]. MBE evaluates whether a model systematically overestimates or underestimates predictions. A positive MBE indicates overestimation, whereas a negative MBE suggests underestimation[178]. NSE is widely used in hydrological and environmental modeling to assess predictive performance. It compares the model's prediction accuracy against a simple mean-based prediction[179]. An NSE value of 1 represents a perfect model, while values close to 0 or negative indicate poor predictive ability. The definitions and purposes of these evaluation metrics are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{5}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}\right)^2 \tag{6}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{7}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{8}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{9}$$

where n represents the total number of observations. The term x_i refers to the predicted value for the i-th observation, while y_i denotes the actual or observed value for the same observation. The symbol $\bar{y}$ stands for the mean of the observed values and $\bar{x}$ represents the mean of the predicted values.

Optimizing the hyper-parameters of ML models

To enhance the performance of the base learners, a grid search cross-validation (GridSearchCV) method was employed to identify the optimal hyperparameter combinations. This approach involved systematically testing different parameter sets within predefined ranges, as guided by previous research[177,180]. The effectiveness of these combinations was evaluated using a tenfold cross-validation strategy alongside RMSE and the R^2. This strategy enabled the model to be repeatedly validated on various dataset partitions, effectively minimizing variance and reducing the risk of overfitting. This comprehensive search determined the optimal hyperparameters for each base model, and these models were subsequently integrated as base learners in the stacking framework. The specific hyperparameters and their tested values for each base model are presented in Table 5.

Hotspot and cold-spot analyses

Hotspot analyses were commonly used in socio-economic and ecological studies to identify significant spatial clusters, or "hotspots" and "cold spots," of climate parameters across the Middle East region[5]. Hotspots denote areas with elevated values, while cold spots signify areas with lower values relative to the surrounding areas[89]. This investigation focused on the baseline year of 2015, using these analyses to establish a reference point for comparative studies over three future periods: the near-term (2015–2045), medium-term (2045–2075), and long-term (2075–2099), thereby facilitating the observation of spatial shifts in climatic patterns over time. The Stacking-EML approach was utilized to model these climate parameters, and the analyses

were conducted using ArcGIS Pro software. The Hot Spot Analysis (Getis-Ord Gi*) tool was employed to detect statistically significant clusters of high or low values in the dataset, leveraging the Getis-Ord Gi* statistic to calculate z-scores and *p*-values. The z-scores indicate the number of standard deviations a feature's value is from the mean, while the p-values represent the probability of the observed clustering occurring by chance[181]. The Getis-Ord Gi* statistic is computed as follows[89]:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{ij}x_j - \bar{X}\sum_{j=1}^{n} w_{ij}}{\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{ij}^2 - \left(\sum_{j=1}^{n} w_{ij}\right)^2\right]}{n-1}}} \quad (10)$$

where: X_j is the attribute value for feature j, w_{ij} is the spatial weight between features i and j, n is the total number of features, X is the mean of the attribute values, S is the standard deviation of the attribute values. The Getis-Ord Gi* statistic determines whether high or low values cluster spatially by calculating these values for each feature. This analysis identified hotspots and cold spots with confidence levels of 90%, 95%, and 99%, providing a detailed spatial representation of areas with significantly elevated or reduced minimum temperatures, maximum temperatures, and precipitation. Following the initial hotspot detection, a Hotspot Comparison Analysis was performed to compare spatial shifts across the studied periods. This comparison facilitated the identification of trends in the changes in climate parameters, allowing for a comprehensive examination of how hotspots (areas of intense climate parameter values) and cold spots (areas of low values) evolved. The results offer valuable insights into the region's spatial dynamics of climatic changes.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Code availability

The underlying code for this study is not publicly available but may be made available to qualified researchers at a reasonable request from the corresponding author.

Received: 12 November 2024; Accepted: 3 April 2025;
Published online: 08 May 2025

References

1. McCarthy, J. J. Climate Change 2001: Impacts, Adaptation, and Vulnerability: Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Vol. 2 (Cambridge University Press, 2001).
2. Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G. & Lobell, D. B. Anthropogenic climate change has slowed global agricultural productivity growth. *Nat. Clim. Change* **11**, 306–312 (2021).
3. Malhi, G. S., Kaur, M. & Kaushik, P. Impact of climate change on agriculture and its mitigation strategies: a review. *Sustainability* **13**, 1318 (2021).
4. Pokhrel, Y. et al. Global terrestrial water storage and drought severity under climate change. *Nat. Clim. Change* **11**, 226–233 (2021).
5. Khosravi, Y., Homayouni, S. & St-Hilaire, A. An integrated dryness index based on geographically weighted regression and satellite earth observations. *Sci. Total Environ.* **911**, 168807 (2024).
6. Ouarda, T. B. et al. Evolution of the rainfall regime in the United Arab Emirates. *J. Hydrol.* **514**, 258–270 (2014).
7. Salimi, M. & Al-Ghamdi, S. G. Climate change impacts on critical urban infrastructure and urban resiliency strategies for the Middle East. *Sustain. Cities Soc.* **54**, 101948 (2020).
8. Ouarda, T. B. & Charron, C. Changes in the distribution of hydro-climatic extremes in a non-stationary framework. *Sci. Rep.* **9**, 8104 (2019).
9. Lau, W. K. M., Wu, H. T. & Kim, K. M. A canonical response of precipitation characteristics to global warming from CMIP5 models. *Geophys. Res. Lett.* **40**, 3163–3169 (2013).
10. Wasimi, S. A. Climate change in the Middle East and North Africa (MENA) region and implications for water resources project planning and management. *Int. J. Clim. Change Strateg. Manag.* **2**, 297–320 (2010).
11. Gonzalez, R. et al. Water budget analysis in arid regions, application to the United Arab Emirates. *Water* **8**, 415 (2016).
12. Bouramdane, A.-A. Assessment of CMIP6 multi-model projections worldwide: which regions are getting warmer and are going through a drought in Africa and Morocco? What changes from CMIP5 to CMIP6? *Sustainability* **15**, 690 (2022).
13. Seneviratne, S. I. & Hauser, M. Regional climate sensitivity of climate extremes in CMIP6 versus CMIP5 multimodel ensembles. *Earth's Future* **8**, e2019EF001474 (2020).
14. Dibike, Y. B., Gachon, P., St-Hilaire, A., Ouarda, T. B. & Nguyen, V. T.-V. Uncertainty analysis of statistically downscaled temperature and precipitation regimes in Northern Canada. *Theor. Appl. Climatol.* **91**, 149–170 (2008).
15. Kristiansen, T., Miller, J. & Makonnen, S. Regional downscaling of CMIP6 climate projections for the Vancouver coastline. (2024).
16. Capotondi, A., Alexander, M. A., Bond, N. A., Curchitser, E. N. & Scott, J. D. Enhanced upper ocean stratification with climate change in the CMIP3 models. *J. Geophys. Res. Oceans* **117**, 1–23 (2012).
17. Heinrich, G., Gobiet, A. & Mendlik, T. Extended regional climate model projections for Europe until the mid-twenty-first century: combining ENSEMBLES and CMIP3. *Clim. Dyn.* **42**, 521–535 (2014).
18. Kim, H.-J., Wang, B. & Ding, Q. The global monsoon variability simulated by CMIP3 coupled climate models. *J. Clim.* **21**, 5271–5294 (2008).
19. Onyutha, C., Tabari, H., Rutkowska, A., Nyeko-Ogiramoi, P. & Willems, P. Comparison of different statistical downscaling methods for climate change rainfall projections over the Lake Victoria basin considering CMIP3 and CMIP5. *J. Hydro-Environ. Res.* **12**, 31–45 (2016).
20. Perez, J., Menendez, M., Mendez, F. J. & Losada, I. J. Evaluating the performance of CMIP3 and CMIP5 global climate models over the north-east Atlantic region. *Clim. Dyn.* **43**, 2663–2680 (2014).
21. Seth, A., Rojas, M. & Rauscher, S. A. CMIP3 projected changes in the annual cycle of the South American monsoon. *Clim. Change* **98**, 331–357 (2010).
22. Seth, A., Thibeault, J., Garcia, M. & Valdivia, C. Making sense of twenty-first-century climate change in the Altiplano: observed trends and CMIP3 projections. *Ann. Assoc. Am. Geogr.* **100**, 835–847 (2010).
23. Tao, H., Gemmer, M., Jiang, J., Lai, X. & Zhang, Z. Assessment of CMIP3 climate models and projected changes of precipitation and temperature in the Yangtze River Basin, China. *Clim. Change* **111**, 737–751 (2012).
24. Torres, R. R. & Marengo, J. A. Climate change hotspots over South America: from CMIP3 to CMIP5 multi-model datasets. *Theor. Appl. Climatol.* **117**, 579–587 (2014).
25. Chen, H., Sun, J., Lin, W. & Xu, H. Comparison of CMIP6 and CMIP5 models in simulating climate extremes. *Sci. Bull.* **65**, 1415–1418 (2020).
26. Try, S. et al. Comparison of CMIP5 and CMIP6 GCM performance for flood projections in the Mekong River Basin. *J. Hydrol.: Reg. Stud.* **40**, 101035 (2022).
27. Rao, J. & Garfinkel, C. I. CMIP5/6 models project little change in the statistical characteristics of sudden stratospheric warmings in the 21st century. *Environ. Res. Lett.* **16**, 034024 (2021).
28. Grose, M. R. et al. A CMIP6-based multi-model downscaling ensemble to underpin climate change services in Australia. *Clim. Serv.* **30**, 100368 (2023).

29. Bilbao-Barrenetxea, N., Martínez-España, R., Jimeno-Sáez, P., Faria, S. H. & Senent-Aparicio, J. Multi-model ensemble machine learning approaches to project climatic scenarios in a River Basin in the pyrenees. *Earth Syst. Environ.* **8**, 1–19 (2024).
30. Knutti, R. & Sedláček, J. Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Clim. change* **3**, 369–373 (2013).
31. Hawkins, E. & Sutton, R. The potential to narrow uncertainty in regional climate predictions. *Bull. Am. Meteorol. Soc.* **90**, 1095–1108 (2009).
32. Shetty, S., Umesh, P. & Shetty, A. The effectiveness of machine learning-based multi-model ensemble predictions of CMIP6 in Western Ghats of India. *Int. J. Climatol.* **43**, 5029–5054 (2023).
33. Hagos, S. & Zhang, C. Diabatic heating, divergent circulation and moisture transport in the African monsoon system. *Q. J. R. Meteorol. Soc.* **136**, 411–425 (2010).
34. Gleckler, P. J., Taylor, K. E. & Doutriaux, C. Performance metrics for climate models. *J. Geophys. Res.: Atmos.* **113**, 1–20 (2008).
35. Sanderson, B. M., Knutti, R. & Caldwell, P. A representative democracy to reduce interdependency in a multimodel ensemble. *J. Clim.* **28**, 5171–5194 (2015).
36. Tebaldi, C. & Knutti, R. The use of the multi-model ensemble in probabilistic climate projections. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **365**, 2053–2075 (2007).
37. Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
38. Mauritsen, T. & Stevens, B. Missing iris effect as a possible cause of muted hydrological change and high climate sensitivity in models. *Nat. Geosci.* **8**, 346–351 (2015).
39. Duan, K., Wang, X., Liu, B., Zhao, T. & Chen, X. Comparing Bayesian model averaging and reliability ensemble averaging in post-processing runoff projections under climate change. *Water* **13**, 2124 (2021).
40. Wang, B. et al. Using multi-model ensembles of CMIP5 global climate models to reproduce observed monthly rainfall and temperature with machine learning methods in Australia. *Int. J. Climatol.* **38**, 4891–4902 (2018).
41. Sillmann, J., Kharin, V., Zhang, X., Zwiers, F. & Bronaugh, D. Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys Res: atmospheres* **118**, 1716–1733 (2013).
42. Duan, Q., Ajami, N. K., Gao, X. & Sorooshian, S. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Adv. Water Resour.* **30**, 1371–1386 (2007).
43. Lambert, S.J. & Boer, G.J. CMIP1 evaluation and intercomparison of coupled climate models. *Clim. dyn.* **17**, 83–106 (2001).
44. Sanderson, B. M., Knutti, R. & Caldwell, P. Addressing interdependency in a multimodel ensemble by interpolation of model properties. *J. Clim.* **28**, 5150–5170 (2015).
45. Herger, N. et al. Selecting a climate model subset to optimise key ensemble properties. *Earth Syst. Dyn.* **9**, 135–151 (2018).
46. Krishnamurti, T. N. et al. Multimodel ensemble forecasts for weather and seasonal climate. *J. Clim.* **13**, 4196–4216 (2000).
47. Li, T. et al. Machine learning to optimize climate projection over China with multi-model ensemble simulations. *Environ. Res. Lett.* **16**, 094028 (2021).
48. Fu, Y., Zhuang, H., Shen, X. & Li, W. Assessment and prediction of regional climate based on a multimodel ensemble machine learning method. *Clim. Dyn.* **61**, 4139–4158 (2023).
49. Guven, D. Development of multi-model ensembles using tree-based machine learning methods to assess the future renewable energy potential: case of the East Thrace, Turkey. *Environ. Sci. Pollut. Res.* **30**, 87314–87329 (2023).
50. Asadollah, S. B. H. S., Sharafati, A. & Shahid, S. Application of ensemble machine learning model in downscaling and projecting climate variables over different climate regions in Iran. *Environ. Sci. Pollut. Res.* **29,** 1–20 (2022).
51. Bauer, P., Thorpe, A. & Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55 (2015).
52. Yılmaz, B., Aras, E. & Nacar, S. A CMIP6-ensemble-based evaluation of precipitation and temperature projections. *Theor. Appl. Climatol.* **155,** 1–25 (2024).
53. Shao, Y., Bishop, C., Abramowitz, G. & Hobeichi, S. (Copernicus Meetings, 2024).
54. Wang, D. et al. Projection of future precipitation change using CMIP6 multimodel ensemble based on fusion of multiple machine learning algorithms: A case in Hanjiang River Basin, China. *Meteorol. Appl.* **30**, e2144 (2023).
55. Zhang, P., Lu, J. & Chen, X. Machine-learning ensembled CMIP6 projection reveals socio-economic pathways will aggravate global warming and precipitation extreme. *Hydrol. Earth Syst. Sci. Discuss.* **2022**, 1–40 (2022).
56. Singh, D. et al. Machine learning-based streamflow prediction in a hilly catchment for future scenarios 1 using CMIP6 data 2. *Hydrol. Earth Syst. Sci. Discuss.* **2022**, 1–41 (2022).
57. Gu, J., Liu, S., Zhou, Z., Chalov, S. R. & Zhuang, Q. A stacking ensemble learning model for monthly rainfall prediction in the Taihu Basin, China. *Water* **14**, 492 (2022).
58. Tao, S. et al. Retrieving soil moisture from grape growing areas using multi-feature and stacking-based ensemble learning modeling. *Comput. Electron. Agric.* **204**, 107537 (2023).
59. Dou, J. et al. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides* **17**, 641–658 (2020).
60. Rice, J. S. & Emanuel, R. E. How are streamflow responses to the El Nino Southern Oscillation affected by watershed characteristics? *Water Resour. Res.* **53**, 4393–4406 (2017).
61. Zhai, B. & Chen, J. Development of a stacked ensemble model for forecasting and analyzing daily average PM2. 5 concentrations in Beijing, China. *Sci. Total Environ.* **635**, 644–658 (2018).
62. Sun, W. & Trevor, B. A stacking ensemble learning framework for annual river ice breakup dates. *J. Hydrol.* **561**, 636–650 (2018).
63. Aldosary, A. S., Al-Ramadan, B., Kafy, A. A., Altuwaijri, H. A. & Rahaman, Z. A. Forecasting climate risk and heat stress hazards in arid ecosystems: Machine learning and ensemble models for specific humidity prediction in Dammam, Saudi Arabia. *Nat. Hazards*, 1–29 (2025).
64. Baig, F., Ali, L., Faiz, M. A., Chen, H. & Sherif, M. How accurate are the machine learning models in improving monthly rainfall prediction in hyper-arid environment? *J. Hydrol.* **633**, 131040 (2024).
65. Najafi, M. S. & Kuchak, V. S. Ensemble-based monthly to seasonal precipitation forecasting for Iran using a regional weather model. *Int. J. Climatol.* **44**, 4366–4387 (2024).
66. Al-Saeedi, B. A., M. Baez-Villanueva, O., and Ribbe, L.: An optimized representation of precipitation in Jordan: Merging gridded precipitation products and ground-based measurements using machine learning and geostatistical approaches, EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-**11510**, 2024.
67. Tuysuzoglu, G., Birant, K. U. & Birant, D. Rainfall prediction using an ensemble machine learning model based on K-stars. *Sustainability* **15**, 5889 (2023).
68. Jaiswal, P. G. et al. in Proceedings of the 6th International Conference on Information Systems and Computer Networks (ISCON) 1–6 (IEEE, 2023).
69. Li, Y. et al. A multi-model integration method for monthly streamflow prediction: modified stacking ensemble strategy. *J. Hydroinform.* **22**, 310–326 (2020).

70. Morshed-Bozorgdel, A., Kadkhodazadeh, M., Valikhan Anaraki, M. & Farzin, S. A novel framework based on the stacking ensemble machine learning (SEML) method: application in wind speed modeling. *Atmosphere* **13**, 758 (2022).
71. Min, M. et al. Estimating summertime precipitation from Himawari-8 and global forecast system based on machine learning. *IEEE Trans. Geosci. Rremote Sens.* **57**, 2557–2570 (2018).
72. Jose, D. M., Vincent, A. M. & Dwarakish, G. S. Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques. *Sci. Rep.* **12**, 4678 (2022).
73. Zhang, Z. & Li, J. Big Data Mining for Climate Change (2019).
74. Attada, R. et al. Surface Air Temperature Variability over the Arabian Peninsula and Its Links to Circulation Patterns. **39**, 445–464 (2018).
75. Selek, B., Tuncok, I. K. & Selek, Z. Changes in climate zones across Turkey. *J. Water Clim. Change* **9**, 178–195 (2018).
76. Kim, K. & Garcia, T. F. Climate change and violent conflict in the Middle East and North Africa. *Int. Stud. Rev.* **25**, viad053 (2023).
77. Marston, S. A., Knox, P. L. & Liverman, D. M. *World Regions in Global Context: Peoples, Places, and Environments* (2002).
78. East, W. B. M. & Development, N.A.R.S. *Increasing Resilience to Climate Change in the Agricultural Sector of the Middle East: The Cases of Jordan and Lebanon* (World Bank Publications, 2013).
79. Dezfuli, A., Razavi, S. & Zaitchik, B. F. Compound effects of climate change on future transboundary water issues in the Middle East. *Earth's Future* **10**, e2022EF002683 (2022).
80. Naderi, M., Saatsaz, M. & Behrouj Peely, A. Extreme climate events under global warming in Iran. *Hydrol. Sci. J.* **69**, 337–364 (2024).
81. Mora, C. et al. Global risk of deadly heat. *Nat. Clim. change* **7**, 501–506 (2017).
82. Elgendy, M., Hassini, S. & Coulibaly, P. Review of Climate Change Adaptation Strategies in Water Management. *J. Hydrol. Eng.* **29**, 03123001 (2024).
83. Balbus, J. et al. Climate change and human health. The impacts of climate change on human health in the United States: A scientific assessment, 25–42 (2016).
84. Saade, J., Ghanimeh, S., Atieh, M. & Ibrahim, E. in Satellite Monitoring of Water Resources in the Middle East 149–169 (Springer, 2022).
85. Ouarda, T. B. et al. Prediction of heatwave-related mortality magnitude, duration and frequency with climate variability and climate change information. *Stochastic Environ. Res. Risk Assess.* **38**, 1–13 (2024).
86. Basha, G., Ouarda, T. B. & Marpu, P. R. Long-term projections of temperature, precipitation and soil moisture using non-stationary oscillation processes over the UAE region. *Int. J. Climatol.* **35**, 4606–4618 (2015).
87. de Andrade Costa, D. et al. Water availability and extreme events under climate change scenarios in an experimental watershed of the Brazilian Atlantic Forest. *Sci. Total Environ.* **946**, 174417 (2024).
88. Kassaye, S. M., Tadesse, T., Tegegne, G. & Hordofa, A. T. Quantifying the climate change impacts on the magnitude and timing of hydrological extremes in the Baro River Basin, Ethiopia. *Environ. Syst. Res.* **13**, 2 (2024).
89. Getis, A. & Ord, J. K. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **24**, 189–206 (1992).
90. Arias, P. et al. Climate Change 2021: the Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; Technical Summary. **946,** 174417 (2021).
91. Soares, P. M. et al. High-resolution downscaling of CMIP6 Earth system and global climate models using deep learning for Iberia. *Geosci. Model Dev.* **17**, 229–259 (2024).
92. Hamed, M. M., Nashwan, M. S. & Shahid, S. Inter-comparison of historical simulation and future projections of rainfall and temperature by CMIP5 and CMIP6 GCMs over Egypt. *Int. J. Climatol.* **42**, 4316–4332 (2022).
93. Malik, A. et al. Accelerated Historical and Future Warming over the Middle East and North Africa in Response to the Global Temperature Change. *Authorea Preprints* (2023).
94. Terink, W., Immerzeel, W. W. & Droogers, P. Climate change projections of precipitation and reference evapotranspiration for the Middle East and Northern Africa until 2050. *Int. J. Climatol.* **33**, 3055–3072 (2013).
95. Zarrin, A., Dadashi-Roudbari, A. & Hassani, S. Future changes in precipitation extremes over Iran: Insight from a CMIP6 bias-corrected multi-model ensemble. *Pure Appl. Geophys.* **179**, 441–464 (2022).
96. Sharafati, A. & Pezeshki, E. A strategy to assess the uncertainty of a climate change impact on extreme hydrological events in the semi-arid Dehbar catchment in Iran. *Theor. Appl. Climatol.* **139**, 389–402 (2020).
97. Abbaspour, K. C., Faramarzi, M., Ghasemi, S. S. & Yang, H. Assessing the impact of climate change on water resources in Iran. *Water Resour. Res.* **45**, 1–16 (2009).
98. Sarış, F. The spatial pattern of selected extreme precipitation indices for Turkey (1975–2012). *Bull. Geogr. Phys. Geogr. Ser.* **19**, 19–30 (2020).
99. Majdi, F., Hosseini, S. A., Karbalaee, A., Kaseri, M. & Marjanian, S. Future projection of precipitation and temperature changes in the Middle East and North Africa (MENA) region based on CMIP6. *Theor. Appl. Climatol.* **147**, 1–14 (2022).
100. Francis, J. A. & Vavrus, S. J. Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophys. Res. Lett.* **39**, 1–6 (2012).
101. Cohen, J. et al. Recent Arctic amplification and extreme mid-latitude weather. *Nat. Geosci.* **7**, 627–637 (2014).
102. Black, E., Brayshaw, D. J. & Rambeau, C. M. Past, present and future precipitation in the Middle East: insights from models and observations. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **368**, 5173–5184 (2010).
103. Almazroui, M., Khalid, M. S., Islam, M. N. & Saeed, S. Seasonal and regional changes in temperature projections over the Arabian Peninsula based on the CMIP5 multi-model ensemble dataset. *Atmos. Res.* **239**, 104913 (2020).
104. Varela, R., Rodríguez-Díaz, L. & DeCastro, M. Persistent heat waves projected for Middle East and North Africa by the end of the 21st century. *PloS One* **15**, e0242477 (2020).
105. Budhwar, P. & Mellahi, K. Introduction: human resource management in the Middle East. *Int. J. Hum. Resour. Manag.* **18**, 2–10 (2007).
106. Hameed, M., Ahmadalipour, A. & Moradkhani, H. Drought and food security in the Middle East: An analytical framework. *Agric. For. Meteorol.* **281**, 107816 (2020).
107. Ghozat, A., Sharafati, A. & Hosseini, S. A. Satellite-based monitoring of meteorological drought over different regions of Iran: application of the CHIRPS precipitation product. *Environ. Sci. Pollut. Res.* **29**, 36115–36132 (2022).
108. Riahi, K. et al. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Glob. Environ. change* **42**, 153–168 (2017).
109. Fricko, O. et al. The marker quantification of the Shared Socioeconomic Pathway 2: a middle-of-the-road scenario for the 21st century. *Glob. Environ. Change* **42**, 251–267 (2017).
110. Qin, P. et al. Projected impacts of climate change on major dams in the Upper Yangtze River Basin. *Clim. Change* **170**, 8 (2022).
111. Haider, S. et al. Simulation of the potential impacts of projected climate and land use change on runoff under CMIP6 scenarios. *Water* **15**, 3421 (2023).

112. Xin, X., Wu, T., Jie, W. & Zhang, J. Impact of higher resolution on precipitation over China in CMIP6 HighResMIP models. *Atmosphere* **12**, 762 (2021).
113. Gao, W. et al. A deconfounding approach to climate model bias correction. arXiv preprint arXiv:*2408.12063* (2024).
114. Khosravi, Y. & Balyani, S. Spatial modeling of mean annual temperature in Iran: comparing cokriging and geographically weighted regression. *Environ. Model. Assess.* **24**, 341–354 (2019).
115. Goovaerts, P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* **228**, 113–129 (2000).
116. Enayati, M., Bozorg-Haddad, O., Bazrafshan, J., Hejabi, S. & Chu, X. Bias correction capabilities of quantile mapping methods for rainfall and temperature variables. *J. Water Clim. Change* **12**, 401–419 (2021).
117. Nathaniel, J. Bias correction of global climate model using machine learning algorithms to determine meteorological variables in different tropical climates of Indonesia (2020).
118. Wu, Y., et al. Quantifying the uncertainty sources of future climate projections and narrowing uncertainties with bias correction techniques. *Earth's Future* **10**, e2022EF002963 (2022).
119. Shiru, M. S. & Chung, E.-S. Performance evaluation of CMIP6 global climate models for selecting models for climate projection over Nigeria. *Theor. Appl. Climatol.* **146**, 599–615 (2021).
120. Hamed, M. M., Nashwan, M. S. & Shahid, S. A novel selection method of CMIP6 GCMs for robust climate projection. *Int. J. Climatol.* **42**, 4258–4272 (2022).
121. Agyekum, J. et al. Extreme temperature indices over the Volta Basin: CMIP6 model evaluation. *Clim. Dyn.* **61**, 203–228 (2023).
122. Zamani, Y., Hashemi Monfared, S. A., Azhdari Moghaddam, M. & Hamidianpour, M. A comparison of CMIP6 and CMIP5 projections for precipitation to observational data: the case of Northeastern Iran. *Theor. Appl. Climatol.* **142**, 1613–1623 (2020).
123. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
124. Hersbach, H. Global reanalysis: goodbye ERA-Interim, hello ERA5. *ECMWF Newsl.* **159**, 17 (2019).
125. Amjad, M., Yilmaz, M. T., Yucel, I. & Yilmaz, K. K. Performance evaluation of satellite-and model-based precipitation products over varying climate and complex topography. *J. Hydrol.* **584**, 124707 (2020).
126. Tozer, B. et al. Global bathymetry and topography at 15 arc sec: SRTM15+. *Earth Space Sci.* **6**, 1847–1864 (2019).
127. Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
128. Fatima, S., Hussain, A., Amir, S. B., Ahmed, S. H. & Aslam, S. M. H. XGBoost and random forest algorithms: an in-depth analysis. *Pak. J. Sci. Res.* **3**, 26–31 (2023).
129. Deo, R. C., Tiwari, M. K., Adamowski, J. F. & Quilty, J. M. Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stoch. Environ. Res. Risk Assess.* **31**, 1211–1240 (2017).
130. Mohammed, S. et al. A comparative analysis of data mining techniques for agricultural and hydrological drought prediction in the eastern Mediterranean. *Comput. Electron. Agric.* **197**, 106925 (2022).
131. Janizadeh, S. et al. Potential impacts of future climate on the spatio-temporal variability of landslide susceptibility in Iran using machine learning algorithms and CMIP6 climate-change scenarios. *Gondwana Res.* **124**, 1–17 (2023).
132. Tanimu, B. et al. Comparison of conventional and machine learning methods for bias correcting CMIP6 rainfall and temperature in Nigeria. *Theor. Appl. Climatol.* **155**, 1–30 (2024).
133. Yan, Y. et al. Projection of future extreme precipitation in China based on the CMIP6 from a machine learning perspective. *Remote Sens.* **14**, 4033 (2022).
134. Masselot, P. et al. Machine learning approaches to identify thresholds in a heat-health warning system context. *J. R. Stat. Soc. Ser. A Stat. Soc.* **184**, 1326–1346 (2021).
135. Hani, I., St-Hilaire, A. & Ouarda, T. B. Machine-learning modeling of hourly potential thermal refuge area: a case study from the Sainte-Marguerite River (Quebec, Canada). *River Res. Appl.* **39**, 1763–1782 (2023).
136. Souaissi, Z., Ouarda, T. B. & St-Hilaire, A. Non-parametric, semi-parametric, and machine learning models for river temperature frequency analysis at ungauged basins. *Ecol. Inform.* **75**, 102107 (2023).
137. Liu, Y., Wang, Y. & Zhang, J. Information computing and applications. In Proceedings of the Third International Conference, ICICA 2012, Chengde, China, September 14–16, 2012. 246–252 (Springer, 2012).
138. Dey, A., Sahoo, D. P., Kumar, R. & Remesan, R. A multimodel ensemble machine learning approach for CMIP6 climate model projections in an Indian River basin. *Int. J. Climatol.* **42**, 9215–9236 (2022).
139. Akaffou, F. H. et al. Analyzing inflow to Faye reservoir sensitivity to climate change using CMIP6 and random forest algorithm. *Int. J. River Basin Manage.* 1–20 (2024).
140. Limon, G. C. & Jablonowski, C. Probing the skill of random forest emulators for physical parameterizations via a hierarchy of simple CAM6 configurations. *J. Adv. Model. Earth Syst.* **15**, e2022MS003395 (2023).
141. Desai, S. & Ouarda, T. B. Regional hydrological frequency analysis at ungauged sites with random forest regression. *J. Hydrol.* **594**, 125861 (2021).
142. Li, P. Robust logitboost and adaptive base class (abc) logitboost. arXiv preprint arXiv:*1203.3491* (2012).
143. Subramanian, J. & Simon, R. Overfitting in prediction models–is it a problem only in high dimensions? *Contemp. Clin. trials* **36**, 636–641 (2013).
144. Houndekindo, F. & Ouarda, T. B. A non-parametric approach for wind speed distribution mapping. *Energy Convers. Manag.* **296**, 117672 (2023).
145. Chen, T. & Guestrin, C. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 785–794 (2016).
146. Ramraj, S., Uzir, N., Sunil, R. & Banerjee, S. Experimenting XGBoost algorithm for prediction and classification of different datasets. *Int. J. Control Theory Appl.* **9**, 651–662 (2016).
147. Ke, G. et al. Lightgbm: A highly efficient gradient-boosting decision tree. *Advances in. Neural Inf. Process. Syst.* **30**, 1–9 (2017).
148. Fan, J. et al. Light Gradient Boosting Machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agric. Water Manag.* **225**, 105758 (2019).
149. Houndekindo, F. & Ouarda, T. B. Prediction of hourly wind speed time series at unsampled locations using machine learning. *Energy* **299**, 131518 (2024).
150. Mountrakis, G., Im, J. & Ogole, C. Support vector machines in remote sensing: a review. *ISPRS J. Photogramm. Remote Sens.* **66**, 247–259 (2011).
151. Gholami, F., Li, Y., Zhang, J. & Nemati, A. Quantifying the impact of future climate change on flood susceptibility: an integration of CMIP6 models, machine learning, and remote sensing. *J. Water Resour. Plan. Manag.* **150**, 04024031 (2024).
152. Choubin, B. et al. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and

regression trees, and support vector machines. *Sci. Total Environ.* **651**, 2087–2096 (2019).
153. Dorogush, A. V., Ershov, V. & Gulin, A. CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv*:1810.11363* (2018).
154. Ullah, I., Liu, K., Yamamoto, T., Zahid, M. & Jamal, A. Prediction of electric vehicle charging duration time using ensemble machine learning algorithm and Shapley additive explanations. *Int. J. Energy Res.* **46**, 15211–15230 (2022).
155. Ahn, J. M., Kim, J. & Kim, K. Ensemble machine learning of gradient boosting (XGBoost, LightGBM, CatBoost) and attention-based CNN-LSTM for harmful algal blooms forecasting. *Toxins* **15**, 608 (2023).
156. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V. & Gulin, A. CatBoost: unbiased boosting with categorical features. *Adv. Neural Inf. Process. Syst.* **31**, 6639–6649 (2018).
157. Kohavi, R. A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection. (Morgan Kaufman Publishing, 1995).
158. Wolpert, D. H. Stacked generalization. *Neural Netw.* **5**, 241–259 (1992).
159. Breiman, L. Stacked regressions. *Mach. Learn.* **24**, 49–64 (1996).
160. Ribeiro, M. H. D. M. & dos Santos Coelho, L. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Appl. Soft Comput.* **86**, 105837 (2020).
161. Caubet, M., Dobarco, M. R., Arrouays, D., Minasny, B. & Saby, N. P. Merging country, continental and global predictions of soil texture: Lessons from ensemble modelling in France. *Geoderma* **337**, 99–110 (2019).
162. Malone, B. P., Minasny, B., Odgers, N. P. & McBratney, A. B. Using model averaging to combine soil property rasters from legacy soil maps and from point data. *Geoderma* **232**, 34–44 (2014).
163. Ilaboya, I. Performance of multiple linear regression (MLR) and artificial neural network (ANN) for the prediction of monthly maximum rainfall in Benin City, Nigeria. *Int. J. Eng. Sci. Appl.* **3**, 21–37 (2019).
164. Choubin, B., Khalighi-Sigaroodi, S., Malekian, A. & Kişi, Ö Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals. *Hydrol. Sci. J.* **61**, 1001–1009 (2016).
165. Li, J., Pollinger, F. & Paeth, H. Comparing the lasso predictor-selection and regression method with classical approaches of precipitation bias adjustment in decadal climate predictions. *Month. Weather Rev.* **148**, 4339–4351 (2020).
166. Uyanık, G. K. & Güler, N. A study on multiple linear regression analysis. *Procedia-Soc. Behav. Sci.* **106**, 234–240 (2013).
167. Pang, B., Yue, J., Zhao, G. & Xu, Z. Statistical downscaling of temperature with the random forest model. *Adv. Meteorol*. **7265178**, 1–11 (2017).
168. Themeßl, M. J., Gobiet, A. & Leuprecht, A. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int. J. Climatol.* **31**, 1530–1544 (2011).
169. Joshi, D., St-Hilaire, A., Ouarda, T. & Daigle, A. Statistical downscaling of precipitation and temperature using sparse Bayesian learning, multiple linear regression and genetic programming frameworks. *Can. Water Resour. J./Rev. Canadienne Des. Ressour. Hydr.* **40**, 392–408 (2015).
170. Dismuke, C. & Lindrooth, R. Ordinary least squares. *Methods Des. outcomes Res.* **93**, 93–104 (2006).
171. Tibshirani, R. Regression shrinkage and selection via the lasso. *J.R. Stat. Soc. Series B Stat. Methodol.* **58**, 267–288 (1996).
172. Hammami, D., Lee, T.S., Ouarda, T.B. & Lee, J. Predictor selection for downscalingata with LASSO. *J. Geophys. Res.: Atmos.* **117** (2012).
173. Abiodun, O. I. et al. Comprehensive review of artificial neural network applications to pattern recognition. *IEEE Access* **7**, 158820–158846 (2019).
174. Tikhamarine, Y., Malik, A., Souag-Gamane, D. & Kisi, O. Artificial intelligence models versus empirical equations for modeling monthly reference evapotranspiration. *Environ. Sci. Pollut. Res.* **27**, 30001–30019 (2020).
175. Ouarda, T.B. & Shu, C. Regional low-flow frequency analysis using single and ensemble artificial neural networks. *Water Resour. Res.* **45** (2009).
176. Ahmed, K., Shahid, S., Haroon, S. B. & Xiao-Jun, W. Multilayer perceptron neural network for downscaling rainfall in arid region: a case study of Baluchistan, Pakistan. *J. Earth Syst. Sci.* **124**, 1325–1341 (2015).
177. Gholam Ali, M., Younes, K., Esmaeil, A. & Fatemeh, T. Assessment of geostatistical methods for spatial analysis of SPI and EDI drought indices. *World Appl. Sci. J.* **15**, 474–482 (2011).
178. Willmott, C. J. On the validation of models. *Phys. Geogr.* **2**, 184–194 (1981).
179. Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I—a discussion of principles. *J. Hydrol.* **10**, 282–290 (1970).
180. Liashchynskyi, P. Grid search, random search, genetic algorithm: a big comparison for NAS. arXiv preprint arXiv*:1912.06059* (2019).
181. Khosravi, Y. & Abbasi, E. Spatial analysis of environmental data with geostatistics. *Zanjan, Azarkalak* (2016).
182. Semmler, T. et al. Simulations for CMIP6 with the AWI climate model AWI-CM-1-1. *J. Adv. Model. Earth Syst.* **12**, e2019MS002009 (2020).
183. Shiogama, H. MIROC MIROC6 model output prepared for CMIP6 DAMIP hist-CO2. (2019).
184. Yukimoto, S. et al. The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2. 0: Description and basic evaluation of the physical component. *J. Meteorol. Soc. Jpn. Ser. II* **97**, 931–965 (2019).
185. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
186. Vapnik, V. N. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **10**, 988–999 (1999).

Acknowledgements

This study was funded by the National Sciences and Engineering Research Council of Canada (NSERC) (funding number: RGPIN-2024-06736). The funder played no role in the study design, data collection, analysis, and interpretation of data, or the writing of this manuscript.

Author contributions

Y.K.: Methodology, Software, Writing—original draft, Writing—review & editing. T.O.: Supervise, methodology, writing—review & editing, conceptualization. S.H.: Methodology, Investigation, Writing—original draft, Writing—review & editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Younes Khosravi.