https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Convective Environments Over the Arabian Peninsula in Current and Future Climates: A Machine Learning Approach

Ahmed Homoudi[1], Henning W. Rust[2], Klemens Barfus[3], Christian Bernhofer[1], and Matthias Mauder[1]

[1]Institute of Hydrology and Meteorology, TUD Dresden University of Technology, Pienner Strasse 23, 01737 Tharandt, Germany

[2]Institute of Meteorology, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany

[3]Independent Researcher, Dresden, Germany

Correspondence: Ahmed Homoudi (ahmed.homoudi@tu-dresden.de)

Abstract. Climate change is intensifying extreme rainfall and flash floods across the arid Arabian Peninsula (AP). Adaptation requires a large ensemble of high-resolution precipitation projections obtained from dynamical downscaling, which remains computationally prohibitive. Here, we present the first component of a statistical-dynamical downscaling framework designed to reduce this computational burden. Machine learning models were trained to identify convective environments (CEs) using predictors derived from ERA5 reanalysis and a binary predictand from IMERG precipitation data. The best-performing model was applied to the CMIP6 ensemble to assess CE occurrence and components under current and future climates. At the current regional warming level, CEs are less frequent in the CMIP6 ensemble relative to ERA5, accompanied by drier and more stable environments, suppressed updraft range, and stronger wind shear. At an additional +1 °C regional warming, CE occurrence generally increases, excluding spring, accompanied by a diurnal shift towards nocturnal and morning periods. A general decrease is projected at an additional +3 °C, excluding winter. Nevertheless, CE-related moisture and instability exhibit monotonic increases with warming, suggesting more extremes and a shift toward a higher contribution of extremes to total rainfall. The resulting CE probability dataset enables targeted event selection for convection-permitting dynamical downscaling across the AP.

1 Introduction

The Arabian Peninsula (AP) is warming faster than the global mean (Cook et al., 2021; Malik et al., 2024). Combined with regional aridity, high temperatures, and limited freshwater resources, the AP is especially vulnerable to the impacts of climate change. While the region is warming and drying (Almazroui et al., 2020), droughts are increasing in frequency, duration, and spatial extent (Saharwardi et al., 2025), and extreme precipitation events are, in contrast, intensifying (Nelli et al., 2021; Luong et al., 2020). These extremes have been associated with significant casualties and substantial economic losses across the region

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

(e.g., De Vries et al., 2016; Francis et al., 2025). To support climate change adaptation efforts, it is essential to understand the atmospheric conditions associated with precipitation in the region under current and future climates.

Precipitation over the AP is sporadic and primarily convective, occurring with moisture intrusion from adjacent seas coupled with atmospheric instabilities (Luong et al., 2020, 2024). However, conditions for convection varies spatially and seasonally. Generally, in the cold months, convection is limited to the northern AP and linked to mid-latitude cyclones and fronts (Almazroui et al., 2016; Catto and Pfahl, 2013; Luong et al., 2024; Charabi, 2009; Taraphdar et al., 2025). In spring, when Mesoscale Convective Systems (MCSs) dominate, the interaction between cold and dry extratropical airmasses and warm, moist tropical airmasses creates convectively favourable environments (Luong et al., 2024; Nelli et al., 2021). In summer, convection is controlled by sea breeze, orographic lifting, surface heating, and the Indian summer Monsoon (ISM) and is confined to the southern AP (Branch et al., 2020; Nelli et al., 2025; Babu et al., 2016; Attada et al., 2019). This spatiotemporal complexity renders the AP a particularly challenging yet scientifically intriguing region for studying convection.

Earth's mean surface temperature has risen by approximately 1.1 °C in 2011-2020 relative to preindustrial levels (1850-1900), predominantly due to increasing anthropogenic greenhouse gas emissions (IPCC, 2023). In a warming atmosphere, the water vapour holding capacity increases with the Clausius-Clapeyron (CC) scaling rate of 6%/°C - 7%/°C (Trenberth, 2011), suggesting an enhanced thermodynamic instability, thereby increasing convection and thunderstorms likelihood (e.g., Lepore et al., 2021; Allen, 2018; Diffenbaugh et al., 2013; Trapp et al., 2007). Convective environments, hereafter CEs, are synoptic- or meso-scale atmospheric conditions favourable for the initiation and organisation of (deep) convection (e.g., combination of instability, moisture, vertical shear, lifting mechanisms). Their response to warming can be evaluated explicitly or implicitly. Explicit approaches rely on convection-permitting models (CPMs), which resolve convection at kilometre-scale resolution. Implicit approaches instead use convective indices, such as Convective Available Potential Energy (CAPE), derived from coarser climate models operating at O(100 km) resolution, where convection is parameterised. Yet, CPMs are computationally expensive and therefore limited in their spatial and temporal coverage. Consequently, two options emerge: relying solely on implicit approaches or using them as a first step to identify situations worth downscaling (e.g., Beaulant et al., 2011; Meredith et al., 2018; Wang et al., 2026).

Numerous studies have utilised the implicit approach to examine CEs' responses to a warming climate, considering different regions, for example, over the globe (Lepore et al., 2021), the United States, (Williams and Fieweger, 2026; Seeley and Romps, 2015; Diffenbaugh et al., 2013; Trapp et al., 2007), Europe (Púčik et al., 2017; Rädler et al., 2019), and finally over Australia (Allen et al., 2014b, a; Cheung et al., 2025). These studies collectively indicate that the frequency and intensity of CEs increase with rising temperatures, yet the change is spatiotemporally heterogeneous. Critically, such analyses remain scarce for the AP, where convection regimes differ from those that underpin most existing literature (i.e., mid-latitudes).

A few AP-focused studies exist, revealing an emerging picture, yet they leave important gaps. Nelli et al. (2025) found summertime convection is increasing over the southeastern AP in a warming climate. Furthermore, Pathak et al. (2025) reported that extreme precipitation is intensifying between November and April, driven by a southward shift in the subtropical jet and increased moisture intrusion from adjacent seas. Both studies, however, are limited in spatial or seasonal scope and rely on parameterised precipitation from global climate models to define convective or extreme events. Comprehensive evaluations

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

of climate models' ability to represent convective indices and assessment of future change remain absent for the AP, in stark contrast to more extensively studied regions in the Global North.

Machine Learning (ML) has emerged as a powerful tool in weather and climate research due to its ability to handle large datasets efficiently, learn complex, nonlinear relationships, and emulate subgrid-scale processes (de Burgh-Day and Leeuwenburg, 2023). For instance, Hill and Schumacher (2021) applied Random Forest (RF) to forecast excessive rainfall over the United States. Miller et al. (2025) applied RF for understanding the occurrence of deep convection over the Great Plains in the United States. Similarly, Zhang et al. (2025b) employed RF to distinguish triggering conditions between embedded and isolated convection initiation during the Meiyu season in China. Kang and Ebtehaj (2025) applied extreme gradient boosting (XGB) to detect subgrid convective precipitation features over the United States.Vahid Yousefnia et al. (2024) demonstrated the skill of deep learning (DL) in thunderstorm forecasting through post-processing of numerical simulation data. Despite this growing body of work, ML-based CE detection has not been applied to the AP.

To address these limitations, we train supervised ML models to detect CEs using predictors derived from the ERA5 reanalysis (Hersbach et al., 2020), with a binary convective label from the Integrated Multi-satellitE Retrievals for GPM (IMERG V07; Huffman et al., 2023a) as the predictand. The best-performing ML model is applied to a Coupled Model Intercomparison Project Phase 6 (CMIP6) ensemble to detect CEs under current and future climates. The novelty of our approach lies in: (i) independence from parameterised precipitation from CMIP6 models; (ii) comprehensive set of predictors accounting for the diverse nature of convection over the AP and the improbability of an all-encompassing index (e.g., Doswell III and Schultz, 2006); (iii) elimination of the need for heuristic thresholds for convective indices; (iv) a consistent CE definition across reanalysis and model datasets, thereby enabling a more robust assessment of future changes; and (v) reporting changes in CEs at +1 °C and +3 °C above the current regional warming levels circumventing the 'hot models' problem (e.g., Hausfather et al., 2022).

This study is set out with the aim to: (a) assess the ability of the CMIP6 ensemble to reproduce the spatiotemporal distribution and the components of CEs over the AP for the current climate; and (b) evaluate changes in the spatiotemporal distribution and the components of CEs at an additional +1 °C and +3 °C above current regional warming levels. The present approach focuses on inferring the frequency and agglomeration of CEs over the region, while leaving the questions of precipitation intensity and extent to be physically resolved using CPMs. Although recent advances demonstrate that global kilometre-scale climate simulations are now within reach (e.g., Hohenegger et al., 2023), producing large ensembles of climate projections to quantify uncertainty and risk remains computationally prohibitive. This challenge is further compounded in the Global South, where limited access to high-performance computing infrastructure and insufficient technical capacity continue to hinder the development of high resolution regional climate projections (Govett et al., 2024; Hoteit et al., 2025). Thus, this work lays the foundation for a statistical-dynamical downscaling approach, in which CEs identified from the CMIP6 ensemble can be dynamically downscaled at convection-permitting resolution, offering a computationally efficient pathway for regional climate projections over a highly sensitive and vulnerable region to climate-related hazards (Mahmoud et al., 2023; Hoteit et al., 2025).

The remainder of this paper is organised as follows: Sect. 2 describes the datasets, ML models, the definition of CEs, as well as the training and evaluation of ML models. The research findings are presented in Sect. 3. Results are discussed in relation

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

to existing literature in Sect .4. Lastly, Sect. 5 summarises and concludes the study, and suggests potential topics for future research.

2 Data and methodology

2.1 Data

In the study, we utilised the IMERG V07 precipitation dataset developed by NASA's Global Precipitation Measurement mission (Huffman et al., 2023b). IMERG is a high-resolution dataset that combines observations from various satellite precipitation estimates using microwave and infrared sensors, along with precipitation gauge analyses at high temporal and spatial scales, with varying latencies (i.e., early, late, and final runs; Huffman et al., 2023a). The final run incorporates monthly gauge corrections to improve accuracy over land surfaces, making it particularly valuable for hydrological applications and climate studies. We used the research-quality 'precipitation' product from the IMERG V07 final run, enclosing the period from 2001 to 2024 with 0.1°spatial resolution and 30 min temporal resolution.

ERA5 is a state-of-the-art global reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). It combines extensive observations from satellites, weather stations, and other sources with numerical weather models to deliver spatiotemporally consistent data (Hersbach et al., 2020). ERA5 provides global output at a spatial resolution of 0.25°, a temporal resolution of 1 hour, and 137 vertical hybrid levels, covering the period from 1940 to the present. We considered ERA5 data on pressure levels for the period 2001-2024. Pressure level data were used instead of the hybrid levels due to the storage limitations.

The CMIP6 represents the latest generation of coordinated climate model experiments designed to improve understanding of past, present, and future climate change. The CMIP6 ensemble consists of a collection of climate model simulations from different modelling groups contributing to the CMIP6 archive (Eyring et al., 2016). In this study, we utilised the historical and SSP5-8.5 experiments. The historical experiment is an all-forcing simulation of the recent past until 2014. The SSP5-8.5 (2015-2100) represents a future scenario characterised by high greenhouse gas emissions, continued reliance on fossil fuels, and limited climate mitigation efforts (O'Neill et al., 2016).

We analysed 10 models from the CMIP6 archive (Table 1), selected based on the data availability of sub-daily and hybrid vertical levels on the Earth System Grid Federation data nodes. While the sub-daily temporal resolution of 6 h enables representation of the diurnal cycle of CEs, data at hybrid levels provide the complete atmospheric profile for estimating convection indices. The model variables considered in the study are zonal and meridional wind components, air temperature, specific humidity, and pressure at the model's hybrid levels (*table_id: 6hrLev*) and at the near-surface level (*table_id: 3hr*). Since data from vertical and near-surface levels are instantaneous measurements, we extracted timestamps from the *3hr* fields that match the *6hrLev* time coordinates, similar to Lepore et al. (2021).

The average regional warming of the AP for the period 2001 - 2024 relative to the preindustrial level (1850 - 1900) is +1.22 °C, based on the area-weighted average of Berkeley Earth (Rohde and Hausfather, 2020), GISTEMP (GISTEMP Team, 2025; Lenssen et al., 2024), HadCRUT5 (Morice et al., 2021), and NOAA Global Temp (Huang et al., 2024) datasets (see

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Table 1. CMIP6 models utilised in the study, including the ensemble member, horizontal resolution in degrees, number of vertical levels, and abbreviations used throughout the text.

Model	Ensemble Member	Resolution (*lon*×*lat*)	Vertical Levels	Hereafter
ACCESS-CM2	r5i1p1f1	1.875×1.250	85	ACM2
BCC-CSM2-MR	r1i1p1f1	1.125×1.125	46	BCM2
CanESM5	r1i1p2f1	2.812×2.812	49	CAN5
CNRM-CM6-1	r1i1p1f2	1.406×1.406	91	CNR6
EC-Earth3	r1i1p1f1	0.703×0.703	91	ECE3
HadGEM3-GC31-LL	r1i1p1f3	1.875×1.250	85	HG3LL
HadGEM3-GC31-MM	r1i1p1f3	0.833×0.555	85	HG3MM
MPI-ESM1-2-LR	r11i1p1f1	1.875×1.875	47	MPI1
MIROC6	r1i1p1f1	1.406×1.400	81	MRC6
MRI-ESM2-0	r1i1p1f1	1.125×1.125	80	MRI2

supporting information Text S1). We considered comparable periods of regional warming using a 24-year average from CMIP6 historical simulations to evaluate model performance. If the relevant period overlaps with the SSP5-8.5 scenario, only years from the historical simulation up to 2014 were used (Table 2), avoiding mixture of experiments, as noted by Lepore et al. (2021). Additionally, we analysed two future periods with temperature increases of +1 °C and +3 °C additional to the current level of regional warming. This corresponds to temperature increases of +2.22 °C and +4.22 °C above preindustrial levels, respectively. We chose a regional warming approach because regional warming rates are 1.4–2.1 times those in the tropics and 2.3–3.1 times the global mean warming rate (Cook et al., 2021). In addition, extreme precipitation (temperature) over the region is projected to increase at a slower (faster) rate than the global land (Seneviratne et al., 2016). According to Corre et al. (2025), regional warming levels are more suitable for adaptation planning when global warming levels differ from regional ones, as shown in Table S2.

2.2 Convective environments

We compiled a comprehensive set of 102 variables from ERA5 data that are potentially informative for convection (Table S1). The majority are indices describing the main components of convection: (1) moisture; (2) instability; (3) lift; and (4) wind shear. The remainder are spatiotemporal covariates, such as longitude, latitude, and day of the year. The 102 variables represent predictors of a statistical model or features in the ML terminology. Throughout the text, the terms *predictors* and *features* are used synonymously to refer to this full set, whereas the term *indices* is reserved for the subset excluding spatiotemporal covariates.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Table 2. Time periods corresponding to regional warming levels for each model. Current (+1.22 °C) includes historical simulation only, while future warming levels consider only the SSP5-85 scenario.

Model	Current (+1.22 °C)	*FutI* (+2.22 °C)	*FutIII* (+4.22 °C)
ACM2	1994 - 2014	2015 - 2036	2039 - 2062
BCM2	1994 - 2014	2022 - 2045	2053 - 2076
CAN5	1981 - 2004	2015 - 2024	2025 - 2048
CNR6	2001 - 2014	2022 - 2045	2048 - 2071
ECE3	1991 - 2014	2015 - 2034	2041 - 2064
HG3LL	1997 - 2014	2015 - 2034	2036 - 2059
HG3MM	1998 - 2014	2015 - 2038	2042 - 2065
MPI1	1991 - 2014	2026 - 2049	2059 - 2082
MRC6	1998 - 2014	2025 - 2048	2058 - 2081
MRI2	2004 - 2014	2022 - 2045	2058 - 2081

Constrained by CMIP6 temporal (6 h) and spatial ($\sim$1°×$\sim$1°) resolutions, we estimated the 6 h accumulated precipitation from IMERG data centred at timestamps matching the CMIP6 data (i.e., 00:00, 06:00, 12:00, and 18:00 UTC). Then we averaged the accumulated precipitation over a $\sim$1°×$\sim$1°grid cell. A grid cell was classified as convective if the accumulated average precipitation exceeded the local climatological median, which was estimated from wet timesteps only (precipitation $\geq$1 mm 6 h^{-1}). Figure 1 illustrates the median precipitation and the fraction of time rainfall exceeds 1 mm 6 h^{-1}. The median threshold is considered given the convective nature of precipitation in the region and the commonality of virga (Wang et al., 2018), it allows us to focus on above-median events in which precipitation likely reaches the surface.

Convection indices are first derived from ERA5 at its original horizontal resolution, using instantaneous atmospheric profiles at 00:00, 06:00, 12:00, and 18:00 UTC. These indices are subsequently averaged onto a 1°×1°grid. To enhance the signal of convection, we computed anomalies relative to the monthly-diurnal climatology for each grid point (see supporting information Text S1). The resulting indices serve as features ($\xi \in \mathbb{R}^n$) for the ML models, paired with binary convective labels ($y \in \{0, 1\}$) derived from IMERG, reducing the problem to a binary classification task or equivalently CE probability estimation.

For CMIP6 data on hybrid levels, we first interpolated data to pressure levels at 10 hPa increments. Then, we calculated and regridded convection indices onto a 1°×1°grid. Generally, the vertical extent of the atmospheric profile is restricted to between 1000 hPa and 100 hPa for both ERA5 and CMIP6, unless explicitly stated otherwise (Table S1). Our `R`-package `MeteoMate` estimates convection indices efficiently by performing its core calculations in Fortran with OpenMP parallelism (Homoudi and Barfus, 2025). Because climate models often exhibit systematic bias, we applied Multivariate Bias Correction (MBCn; Cannon, 2018) to CMIP6 convection indices. The MBCn accounts for dependence among variables without strong stationarity assumptions, using an N-dimensional probability density function transform. Data were grouped by location, calendar month,

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

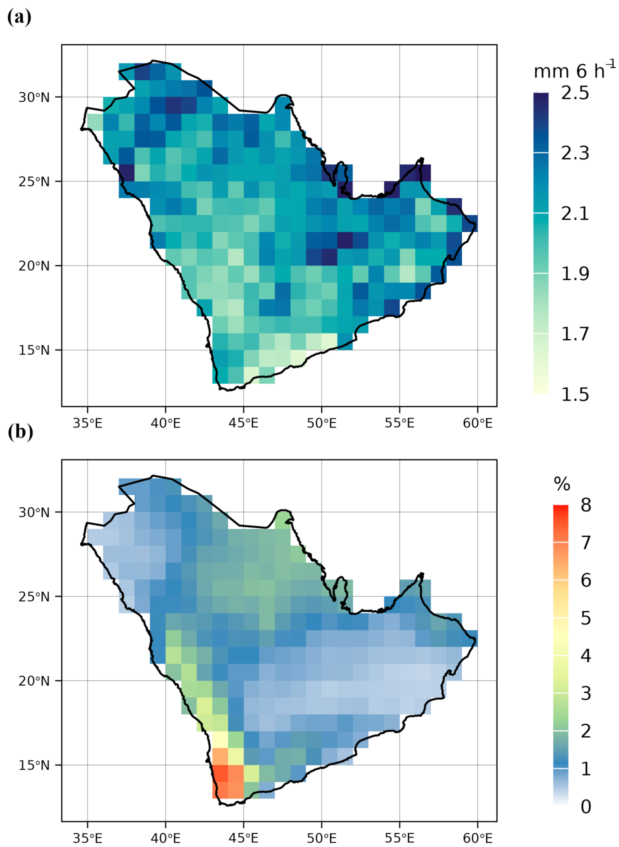

Figure 1. Spatial distribution of (a) 6 h precipitation climatological median over the AP (for precipitation $\geq 1\,\mathrm{mm\,6\,h^{-1}}$) and (b) rain occurrence frequency, calculated as the fraction of 6 h timesteps with precipitation exceeding 1 mm relative to all timesteps (N=35,064).

and hour of the day, and bias correction was applied to each group separately. Probabilities of CE in CMIP6 data were derived from bias-corrected data after removing ERA5 monthly-diurnal climatology.

2.3 Machine learning models

In the study, we considered three ML models: RF, XGB, and DL. Each model learns a non-linear function $f : R^n \to [0,1]$ that estimates the conditional probability of a CE given the feature vector ξ: $f(\xi) = P(y = 1 \mid \xi)$.

RF is an ensemble machine learning method that constructs multiple decision trees (Breiman, 2001). Each tree is trained independently on a bootstrap sample of the training data, and at each node split, it considers only a random subset of features, aiming to minimise the class inhomogeneity in the subsequent nodes. Probabilities are estimated as the fraction of decision trees predicting convection out of the total number of trees. This approach produces more reliable and stable probability

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Table 3. Contingency Table for binary classification.

		Observed	
		Yes	No
Predicted	Yes	hit (h)	false alarm (f)
	No	miss (m)	quiescent (q)

estimates than a single decision tree. For our analyses, we employed the efficient RF implementation of the `R`-package `ranger`, developed by Wright and Ziegler (2017).

XGB is another ensemble technique based on gradient boosting that builds weak decision trees and sequentially reduces their residual errors. It implements techniques such as regularisation and tree pruning to improve generalisation. In XGB, the probabilities of belonging to a certain class are the logarithmic odds transformed by the logistic sigmoid function. We used the `R`-package `XGBoost` in our study, described in Chen and Guestrin (2016).

Herein, DL is a deep feedforward neural network architecture of an input layer, three densely connected hidden layers, and an output layer. Each hidden layer contains multiple neurons, each with associated weights, biases, activation functions, and regularisation mechanisms. However, the output layer consists of a single neuron that applies a sigmoid function to map its output to a probability $\in [0,1]$. We utilised the *Keras* deep learning API (Chollet et al., 2015) for the development and training of the DL model.

2.4 Performance evaluation metrics

We evaluated how well the ML models identify CEs (i.e., classification) and estimate their probabilities. Assessing binary classification performance requires a dichotomous predictand indicating whether or not the atmosphere is convective. This is obtained by converting predicted probabilities to a binary classifier using a threshold probability. The contingency table compares these binary predictions with actual labels from IMERG data (Table 3). Numerous performance metrics exist to further summarise the information in the contingency table; however, we selected metrics based on their suitability for significantly imbalanced data, i.e., Probability of Detection (POD), False Alarm Ratio (FAR), Heidke Skill Score (HSS; Heidke, 1926), and Area under the Precision-Recall Curve (AUCPR). POD is the fraction of observed CEs that were correctly predicted ($h/(h+m)$). FAR is the proportion of predicted CEs that were actually false alarms ($f/(h+f)$). HSS measures the fractional improvement over the reference forecast (i.e., achieved by random chance). HSS is given by

$$HSS = \frac{2(hq - fm)}{(h+m)(m+q)+(h+f)(f+q)}. \tag{1}$$

Different nomenclature exists for these metrics across disciplines. In ML literature, POD is referred to as recall, precision is defined as 1 - FAR, and HSS corresponds to Cohen's kappa. A contingency table can be constructed for various threshold probabilities $\in [0,1]$, allowing performance metrics to be estimated at each threshold. Consequently, plotting POD against 1 - FAR yields Precision-Recall (PR) curves; the area under the PR curve is the AUCPR. It is a threshold-independent metric

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

particularly suited for imbalanced datasets (e.g., Davis and Goadrich, 2006). A ML model without skill would achieve an AUCPR equal to the fraction of CEs in the dataset. This AUCPR is called the baseline AUCPR. Furthermore, optimising ML models for AUCPR implicitly optimises both POD and 1 - FAR. In addition, the threshold probability is determined by maximising the HSS.

For probabilistic assessment of ML models, we used reliability diagrams, which assess calibration by grouping predicted probabilities into bins and comparing them with observed frequencies. A well-calibrated model produces probabilities equal to the observed frequencies. By grouping probabilities into bins, resolution (RES) and reliability (REL) can be estimated to further assess model performance. RES quantifies the model's ability to distinguish between CEs and non-CEs, whereas REL measures calibration accuracy as the mean-squared deviation of the calibration curve from a perfect model (Wilks, 2019). Brier Skill Score (BSS) is the area between the RES and REL curves (Murphy, 1973). A forecast is perfect when BSS equals 1. A BSS of zero indicates no improvement relative to the climatology; conversely, a negative BSS indicates forecast quality is poorer than the climatology. Further details regarding probabilistic assessment are included in supporting information Text S2.

2.5 Training and calibration

In the IMERG-ERA5 dataset, the prevalence of CEs is 0.66%, indicating a significant imbalance. We split the dataset into three parts: a) 60% for training (5.2 million samples), b) 20% for validation and hyperparameter tuning (1.7 million), and c) the remaining 20% for final evaluation (1.7 million). The splitting is stratified, preserving the original distribution of CEs. The rarity of these environments complicates predicting their occurrence, as ML models struggle to learn from imbalanced datasets (e.g., Sun et al., 2009; Vahid Yousefnia et al., 2024). Furthermore, the size of the training dataset exceeds available memory; therefore, we sampled 50% of the non-CEs and oversampled CEs with replacement, increasing the ratio to 13% for RF and XGB. For DL, the complete train dataset was used.

Defining the labels or what constitutes a CE is imperative because it determines the ML models' learning objective and performance. We performed sensitivity analysis to identify the precipitation percentile (i.e., 50^{th}) that yields the best performance. Further information is included in supporting information (Text S3 and Figure S1). In addition, we conducted a separate sensitivity experiment to obtain the optimal fraction of CEs (i.e., 13%) to include in the training data. For further details, see supporting information (Text S4 and Figure S2).

Hyperparameters are the values that control the model's learning process. Optimising these parameters is necessary since the default values are rarely universal and can lead to subpar performance. We have tuned the hyperparameters and model structures for RF, XGB, and DL using Bayesian optimisation (Jones et al., 1998). The optimal hyperparameters for RF and XGB are shown in Tables S3 and S4, respectively, and the DL model architecture and hyperparameters are shown in Table S5. Since we altered the fraction of CEs in the training datasets for RF and XGB, the posterior probabilities from ML models do not align with the prior probabilities of the CEs. Nonetheless, this mismatch can be corrected through an analytical adjustment of the probabilities, as detailed by Pozzolo et al. (2015) and Vahid Yousefnia et al. (2024). The adjusted probability estimate is

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

given as

$$f(\xi,g) = \frac{f(\xi,\hat{g})}{f(\xi,\hat{g}) + \frac{1-g}{g}\frac{\hat{g}}{1-\hat{g}}(1-f(\xi,\hat{g}))}, \quad (2)$$

where g is the climatological probability of CEs (0.66%), $\hat{g}$ is the fraction of CE in the training dataset (13%), and $f(\xi,\hat{g})$ is the model output probability. While this adjustment accounts for the altered class distribution, it does not guarantee predictions match observed distributions in the data. Therefore, we applied beta calibration (Kull et al., 2017) as a post-processing step to further refine the adjusted probabilities

$$f(\xi,g)_{calibrated} = \frac{1}{1+1/(e^{c}\frac{f(\xi,g)^{a}}{(1-f(\xi,g))^{b}})}, \quad (3)$$

where a and b are the shape parameters, and c is the location parameter. Probabilities obtained from DL were directly beta-calibrated since class distribution was not altered. Generally, beta calibration outperforms other calibration approaches because it can model non-linear transformations. In addition, if the input probabilities are already well calibrated, beta calibration learns a function approximating identity, thereby preserving well-calibrated predictions Kull et al. (2017). The calibration model was trained on the validation dataset and applied to the test dataset. Calibrated probabilities are evaluated visually with reliability diagrams and quantified using the Expected Calibration Error (ECE). The ECE is the weighted average difference between predicted probabilities and observed frequencies across probability bins (Naeini et al., 2015, supporting information Text S2). Finally, we used calibrated probabilities from test data only to evaluate the ML models' performance. Noteworthy, probabilities from CMIP6 are also adjusted and calibrated.

3 Results

3.1 ML models performance

The performance of ML models was evaluated on calibrated probabilities from the holdout ERA5 test dataset using 300 bootstrap samples. Fig. 2a shows the reliability diagrams for DL, RF, and XGB. Analytically adjusting the probabilities renders RF underconfident across the full probability range. For XGB, the adjustment results in underconfidence for predicted probabilities below 0.6 and reduces overconfidence for those above 0.6. However, beta calibration recalibrates the probabilities for the three ML models. The parameters of beta calibration are included in Table S6. Furthermore, the ECEs for DL, RF, and XGB are 0.01%, 0.01%, and 0.15%, respectively. Although XGB shows the highest miscalibration, it exhibits high RES, especially for probabilities greater than 0.4 (Fig. 2b). Additionally, AUCPR is larger in the case of XGB (0.60) compared to RF (0.53) and DL (0.54; Fig. 2c,d). Compared to the baseline AUCPR (0.0066), this indicates that XGB, RF, and DL are performing approximately 90, 80, and 82 times better than a classifier with no skill, respectively. The threshold probabilities resulting in the maximum HSS are 0.25, 0.24, and 0.38 for DL, RF, and XGB, respectively. With respect to other metrics, XGB achieves the highest POD, HSS, and BSS and the lowest FAR among ML models (Fig. 2d).

Seasonally, ML models achieve their highest performance in winter (DJF), followed by a gradual decline toward summer (JJA), after which performance recovers (Fig. S3). Overall, XGB outperforms other ML models across all evaluation metrics,

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

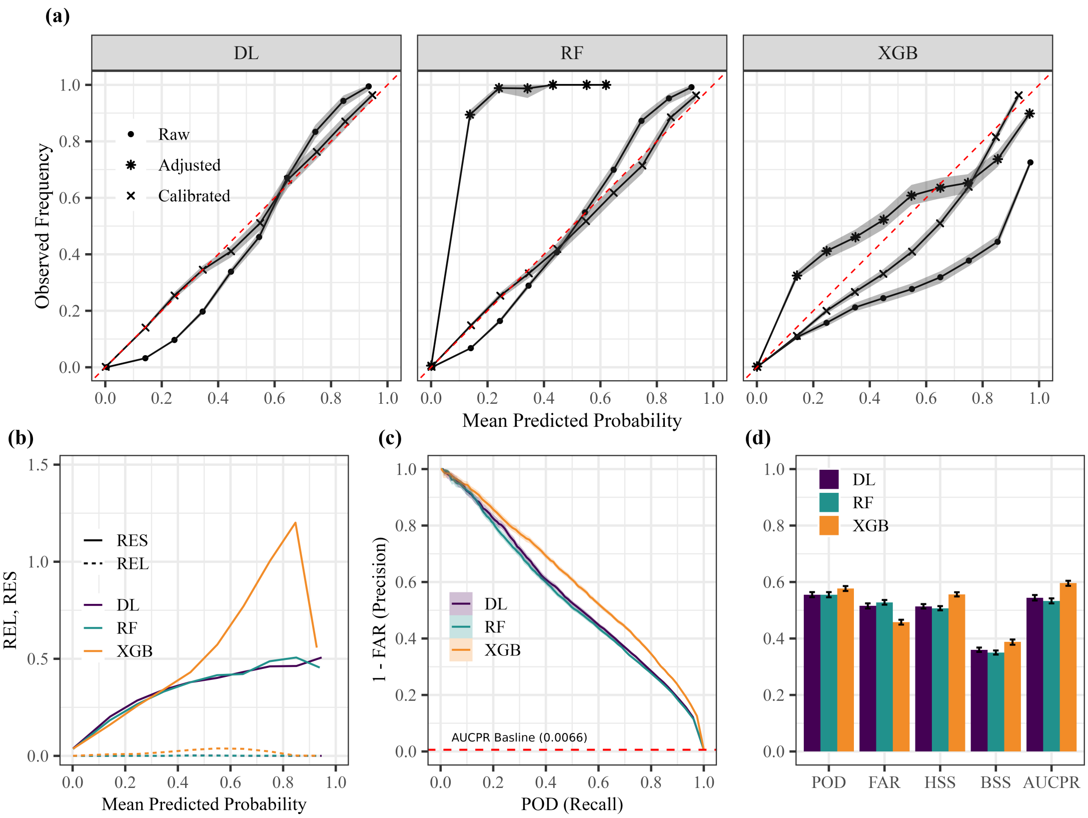

Figure 2. (a) Reliability diagrams for DL, RF, and XGB evaluated on the test data, illustrating the raw, analytically adjusted, and calibrated probabilities. The dashed diagonal line indicates perfect reliability. (b) Bin-wise resolution and reliability, (c) the precision-recall curves, and (d) performance evaluation metrics. The shaded area and error bars denote the symmetric 95% confidence interval obtained from 300 bootstrap resamples.

except in summer, when RF achieves the lowest FAR. Turning to the diurnal cycle, ML models' performance peaks in the early morning (03:00 UTC+3) and weakens toward the afternoon (15:00 UTC+3), then recovers (Fig. S4). Nonetheless, XGB achieves the highest POD, HSS, BSS, and AUCPR across all timesteps; however, it achieves the lowest FAR among ML models only at 09:00 UTC+3. Considering spatial performance, ML models achieve high POD, HSS, BSS, and AUCPR, and low FAR over the northern AP, whereas DL and RF struggle particularly over the southern AP and high terrain. However, XGB performs well in identifying CEs over the southern AP and high terrain (Fig. S5). Overall, these results indicate that XGB outperforms other ML models across evaluation metrics at both temporal and spatial scales; therefore, it was used to identify CEs in CMIP6 data. The rest of the manuscript uses only calibrated probabilities from XGB.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

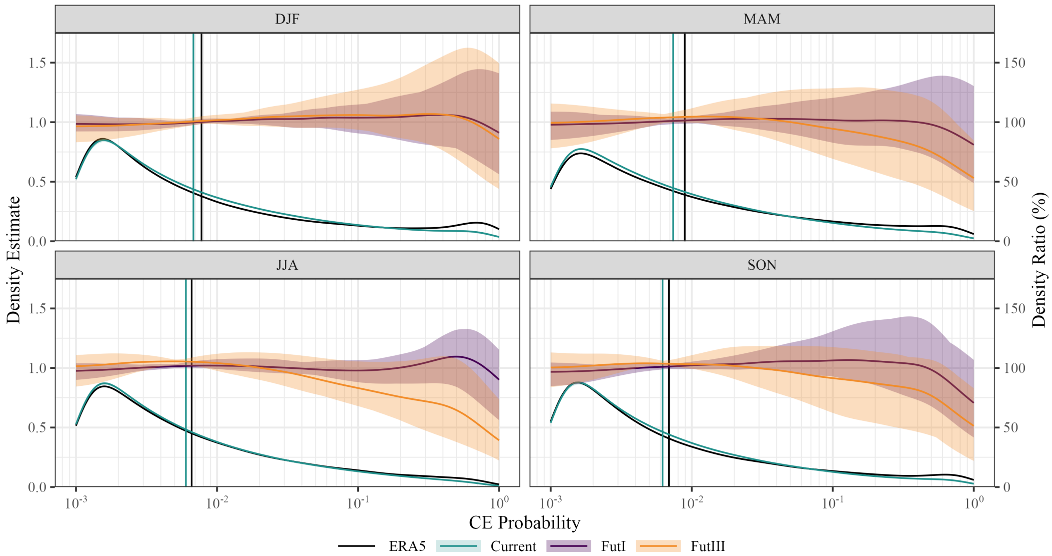

Figure 3. Seasonal kernel density estimate for CE (≥0.001) obtained from XGB for ERA5 test data (black line) and the MM ensemble mean at the current regional warming level (+1.22 °C, cyan) (left y-axis). The vertical lines indicate the seasonal mean CE probability. The right y-axis shows the density ratio of the MM ensemble (FutI: +2.22 °C, orange; FutIII: +4.22 °C, violet) relative to the current warming level. The shaded areas indicate indicate the 2.5th and 97.5th percentile range of the CMIP6 ensemble. The x-axis is $\log_{10}$ transformed.

3.2 Probability and occurrence of convective environments

We evaluated CE probabilities from CMIP6 against the probabilities from ERA5 test data only, because XGB and beta calibration models were trained and calibrated on the training and validation data. Moreover, the CMIP6 CE probabilities presented here are derived from bias-corrected convection indices.

Figure 3 shows the kernel density estimate of the CE seasonal probability for both ERA5 and CMIP6. Generally, the Multi-Model (MM) mean density is lower for high probabilities and higher for low probabilities than ERA5 across all seasons. The mean probability in ERA5 is greatest during MAM, followed by DJF and SON, and lowest in JJA. The MM mean at the current warming level exhibits the same seasonal trend. At *FutI*, the MM density ratio of the CE density estimate relative to the MM current one shows a slight increase up to CE probabilities of 0.7, 0.4, 0.7, and 0.5 for DJF, MAM, JJA, and SON, respectively. Afterwards, it decreases; the smallest decrease is observed in winter (9%) and the largest in autumn (29%). At *FutIII*, the MM density ratio decrease begins at much lower CE probabilities (0.05, 0.02, and 0.028 for MAM, JJA, and SON, respectively), except in DJF, where it begins at 0.65. The largest reduction occurs in JJA (60%) and the smallest in DJF (14%).

The diurnal kernel density estimates for CE probability are shown in Fig. S6. The MM mean density is lower (higher) for high (low) probabilities than for ERA5 across all timesteps. The mean CE probability in ERA5 is greatest at 15:00 and 21:00 UTC+3, followed by 03:00 UTC+3, and lowest at 09:00 UTC+3. The MM exhibits a similar diurnal cycle; however, it is systemically underestimated by 0.001. At the *FutI*, the MM density ratio of the CE density estimate relative to the MM current

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

one remains close to 100% before decreasing, starting at approximately 0.5 in the early morning, 0.7 in the evening, and 0.6 at both the morning and afternoon timesteps. Regarding *FutIII*, the MM density ratio decreases much earlier, starting roughly at 0.04. The decrease is most pronounced in *FutIII* (75%) compared to *FutI* (55%) at both morning and afternoon periods. In addition, the evening (21:00 UTC+3) shows the least reduction in both future climates (41% and 67%).

The seasonal maps of CE probabilities are shown in Fig. 4a. In SON and DJF, elevated probabilities are primarily concentrated over the northeastern AP. During MAM, probabilities are distributed across the region, whereas in JJA, they are limited to the southern AP and high orography. The MM mean reproduces these seasonal patterns, but generally yields lower probabilities with the largest bias in DJF and the smallest in JJA. At *FutI* and *FutIII* (Fig. 4b), most of the AP is projected to experience increases in CE probabilities during DJF. In contrast, MAM is characterised by decreases in CE probabilities in both future climates. In summer, CE probabilities are expected to increase over the AP, except for the Empty Quarter at *FutI*, while the increase is confined to the northern AP at *FutIII*. SON shows an increased CE probability in the northern and central AP during the *FutI* period, with a robust decrease all over the region at *FutIII*.

Diurnal maps of CE probabilities show elevated values over the northeastern AP at all time steps (Fig. S7a), with additional regions appearing in the afternoon and evening along the Red Sea coast and in the southwestern AP. The MM ensemble captures the diurnal pattern, but underestimates across all timesteps, with the largest bias at 03:00 UTC+3 and the smallest at 15:00 UTC+3. In general, the CE probabilities are expected to increase along the Red Sea and regions with high orography at *FutI* (Fig. S7b). An expected increase is also evident in the early morning and morning hours over the central AP. At *FutIII*, a general decrease in mean CE probabilities over the AP across all timesteps is projected.

Converting the CE probabilities into binary or logical variable using the threshold probability enables a comparison of the CE frequency from MM with that of the entire record as defined by IMERG data. Fig. 5a illustrates the seasonal climatology of CEs from IMERG, ERA5 test dataset, and MM. A comparison of CE from IMERG and ERA5 reveals that ERA5 captures the spatial patterns across seasons, with high pattern correlations; however, the CE are underestimated, which is expected given that the ERA5 test data represent only 20% of the total period (2001–2024). MM ensemble reproduces the spatial distribution of CE with the highest pattern correlation in DJF and the lowest in SON. Nevertheless, CEs are underestimated compared to the IMERG climatology, except in SON. The largest underestimation occurs in MAM and the smallest in DJF.

Spatial changes in seasonal mean probabilities (Fig. 4b) are translated into changes in the seasonal mean CE's frequency in future climates (Fig. 5b), with few exceptions. In DJF, the CE probabilities increases primarily over the central to southern AP at *FutIII*, the CE frequencies show a clear increase over the central eastern AP. During JJA, CE probabilities shows a robust increase over the northern AP; however, CEs will be more frequent only over high topography and the southern AP at *FutI*. The southwestern AP is expected to experience a large decrease in CEs at *FutIII*. Finally, MM shows a robust decrease (increase) over the southern (northern) AP and a robust increase over the northern in SON at both future climates. From the diurnal perspective (Fig. S8a), MM captures the spatial pattern, with high correlations. However, the CE frequency in MM is underestimated, with the greatest bias in the evening and the smallest in the morning. An increase in CE frequency is anticipated throughout the day over high topography along the Red Sea at *FutI* (Fig. S8b). Additionally, some internal northern regions

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

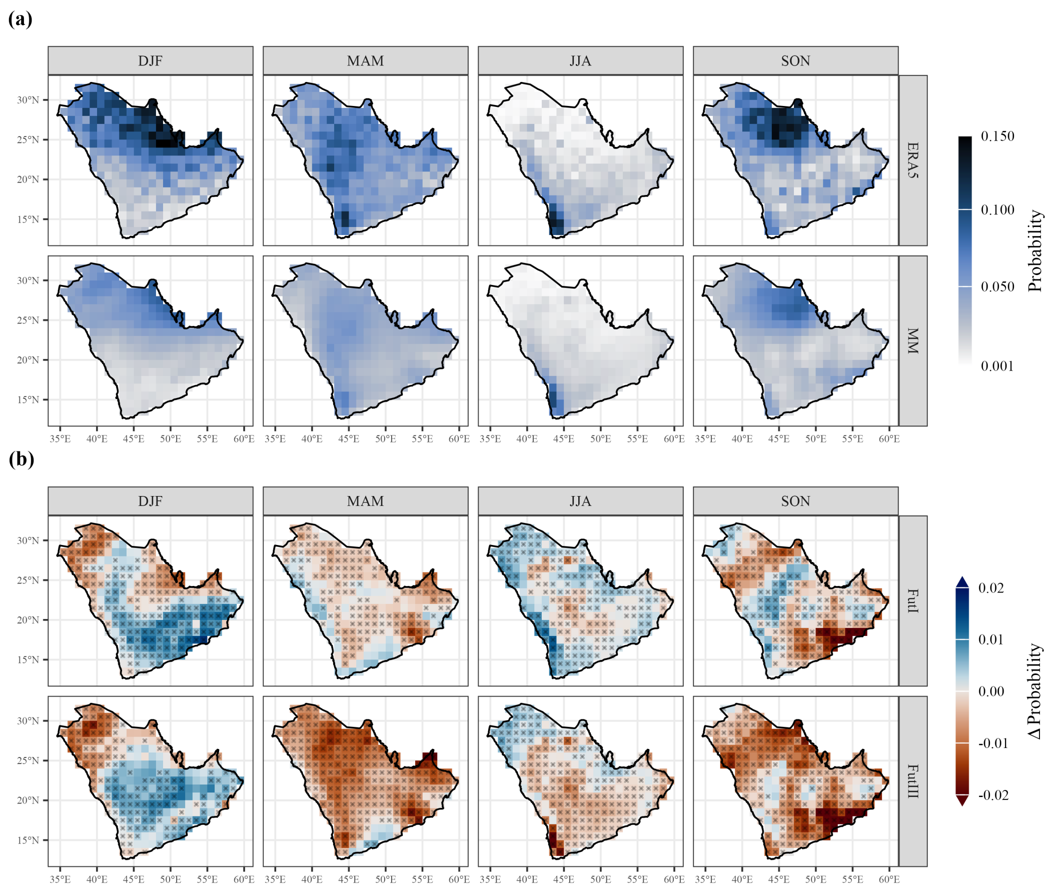

Figure 4. (a) Seasonal CE probabilities (≥0.001) obtained from XGB for ERA5 (test dataset) and the MM ensemble mean at the current (+1.22 °C) regional warming levels, and (b) the seasonal change at *FutI* (+2.22 °C), and *FutIII* (+4.22 °C) warming levels. The stippling indicates at least 70% models' agreement on the sign of change. Colour scale in (a) is $\log_{10}$ transformed.

are expected to experience an increase in CEs at all timesteps, except in the afternoon. At *FutIII*, a robust decrease in CEs is expected across the AP, except over the central AP in the morning, where MM show a robust increase.

The annual cycle of CE frequency across three latitude bands is examined to assess future changes in CE occurrence throughout the year. From IMERG, the CE occurs over the northern AP (NAP; 25.5° N-31.5°N) in the cold months peaking in November (Fig. 6a). The central AP (CAP; 19.5° N-25.0° N) receives rainfall most of the year, with a clear peak in April. The southern AP (SAP: 13.5°N-19.0°N) shows two peaks, one in April and the other in August. The MM ensemble mean reproduces this pattern, with underestimation across the AP and throughout the year. Over the NAP and CAP, the number of CEs slightly increases in SON and MAM, while decreasing in DJF at *FutI*. Nevertheless, the behaviour reverses, and the number of CEs decreases in SON and MAM while increasing in DJF at *FutIII* (Fig. 6b). Over the SAP, CEs show an increase at *FutI*, while a marginal decrease is observed at *FutIII*.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

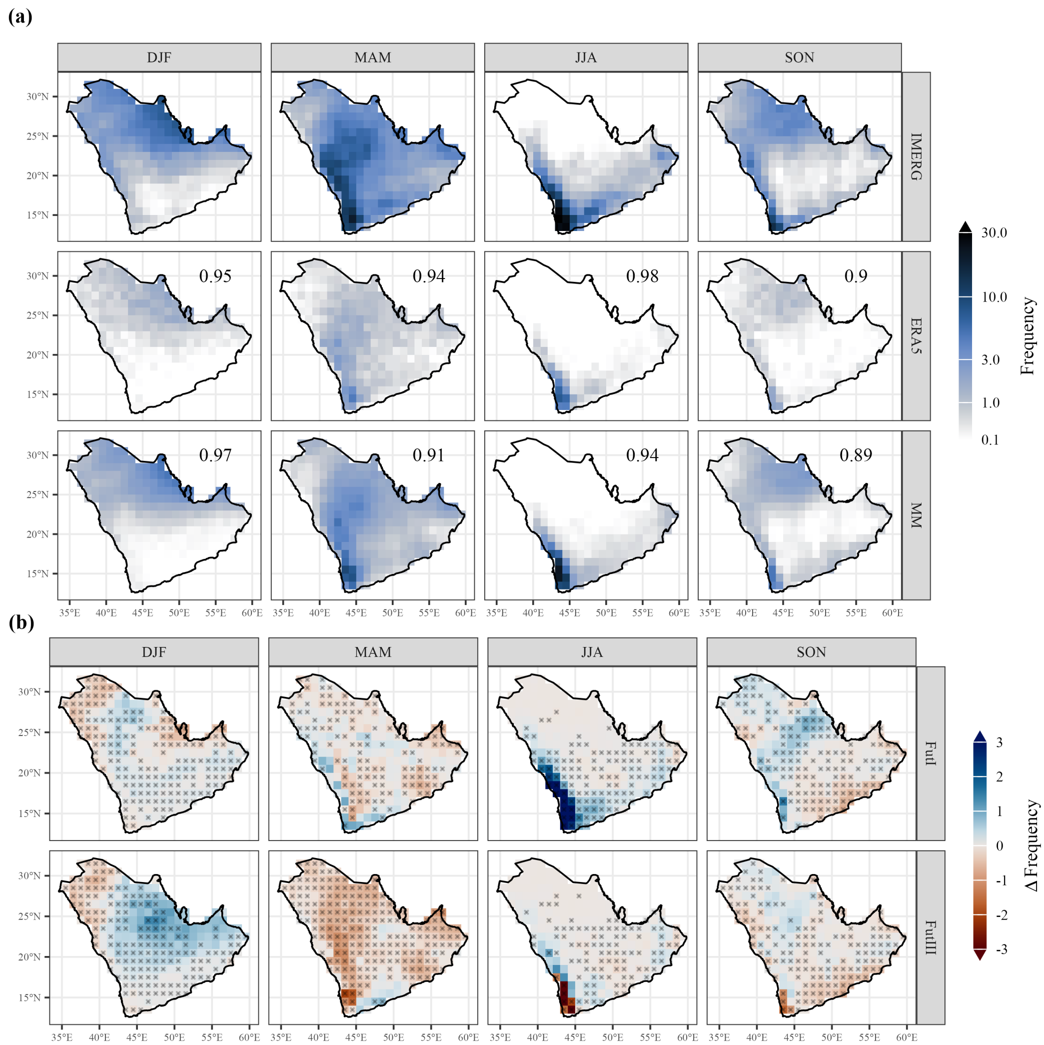

Figure 5. (a) CE frequency per season obtained from the full IMERG record, XGB-detected CEs from ERA5 (test dataset), and the CMIP6 MM ensemble mean at the current (+1.22 °C) regional warming levels, and (b) the seasonal change at *FutI* (+2.22 °C), and *FutIII* (+4.22 °C) warming levels. The stippling indicates at least 70% model agreement on the sign of change. Colour scale in (a) is $\log_{10}$ transformed. The numbers in facets indicate the pattern correlation with CE from IMERG.

3.3 Bias in convective indices

Here, we assess the conditional fidelity of CMIP6 in reproducing CE components at 1°×1°in the current climate, restricted to atmospheric profiles classified as convective by IMERG and XGB for ERA5 and CMIP6, respectively. This restrictions allows a targeted evaluation of CMIP6's ability to simulate the specific thermodynamic and dynamic conditions associated with convection over the region. The analysis here is based on raw (not bias adjusted) convective indices.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

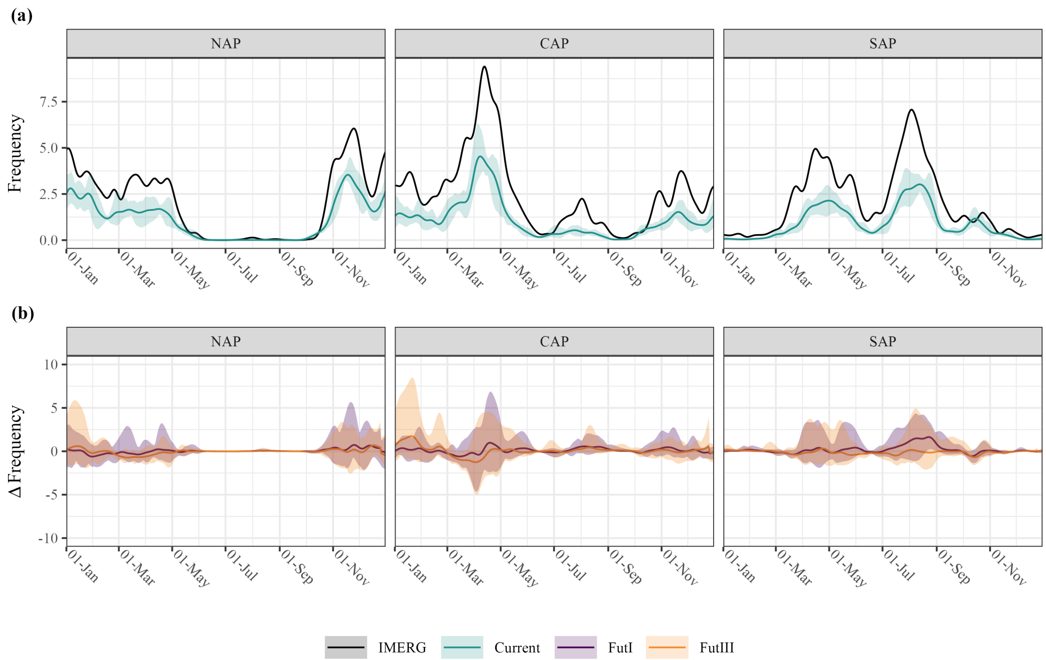

Figure 6. (a) Annual cycle of CE frequency over the northern (NAP), central (CAP), and the southern (SAP) AP from IMERG and the MM at the current warming level. (b) The change in the CE frequency from the MM at the *FutI* and *FutIII* warming levels. The shaded areas indicate indicate the 2.5th and 97.5th percentile range of the CMIP6 ensemble. Data were convoluted using a Gaussian filter with a width of 29 days.

In terms of precipitable water (PW; Fig. 7a) and 2m dew-point temperature (Td_{2m}; Fig. 7b), the dominant signal is dry bias across most models, particularly strong in DJF. However, JJA stands out as BCM2, MRC6, and MPI1 exhibit moist bias, especially at low percentiles. During MAM, MRC6 also shows a strong positive bias. The bias in 500 hPa vertical velocity distribution (ω_{500}) is shown in Fig. 7c, with percentiles sign-reversed for ease of interpretation. Excluding JJA, most CMIP6 models underestimate extremely high (90th-99th) and low (10th-20th) updrafts, while in between, an overestimation signal predominates. In JJA, BCM2 and ECE3 produce stronger extreme updrafts.

Figure 7d and 7e show the bias in percentiles of most unstable CAPE (MUCAPE) and most unstable convection inhibition (MUCIN) for a parcel with the highest equivalent potential temperature (θ_e) in the lowest 300 hPa (Haklander and Van Delden, 2003), considering virtual temperature correction (Doswell and Rasmussen, 1994). Most models show a mostly negative bias in MUCAPE across all seasons, with the largest underestimation at low percentiles. MRC6 is a notable exception, with a rather positively biased MUCAPE all year round, except in SON, where negative bias dominates up to the 70th percentile. CNR6 also shows positive bias during SON and DJF. Regarding MUCIN, the bias is variable across models and seasons, with no consistent direction. Four models persistently overestimate MUCIN, including ACM2, BCM2, HG3LL, and HG3MM, whereas BCM2 shows an underestimation signal in summer. CNR6, ECE3, and MPI1 exhibit a positive bias in DJF, while the rest of the year

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

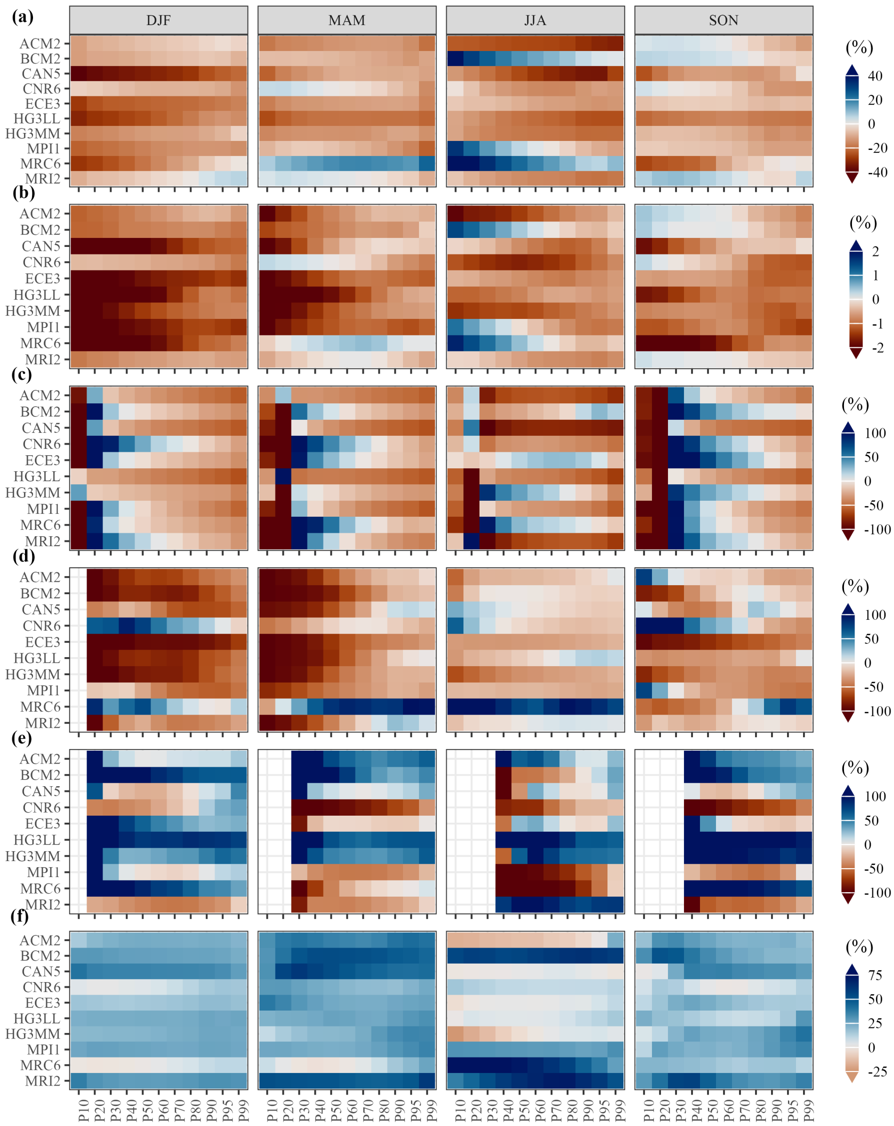

Figure 7. Seasonal bias at different percentiles for (a) precipitable water (PW), (b) 2m dew-point temperature Td_{2m}, (c) vertical velocity at 500 hPa (ω_{500}), (d) most unstable CAPE (MUCAPE), (e) most unstable CIN (MUCIN), and (f) wind shear 0-6 km ($WSh_{0\text{-}6}$), for CEs from CMIP6 at the current regional warming level.

is negatively biased. The only consistent pattern in CAN5 is a positive bias in extreme MUCIN. The missing data in Fig. 7 are due to MUCAPE and MUCIN percentiles from ERA5 equalling zero. Wind shear ($WSh_{0\text{-}6}$; Fig. 7f) is generally stronger across CMIP6 models, except for a few underestimation signals in CNR6 and MRC6 in DJF, MRC6 in MAM, ACM2, CAN5, ECE3, and HG3MM in JJA, and CNR6 in SON. Seasonal patterns of PW, Td_{2m}, and MUCAPE underestimation, MUCIN and $WSh_{0\text{-}6}$ overestimation, and suppressed ω_{500} distribution are generally apparent on the diurnal scale (Fig. S9).

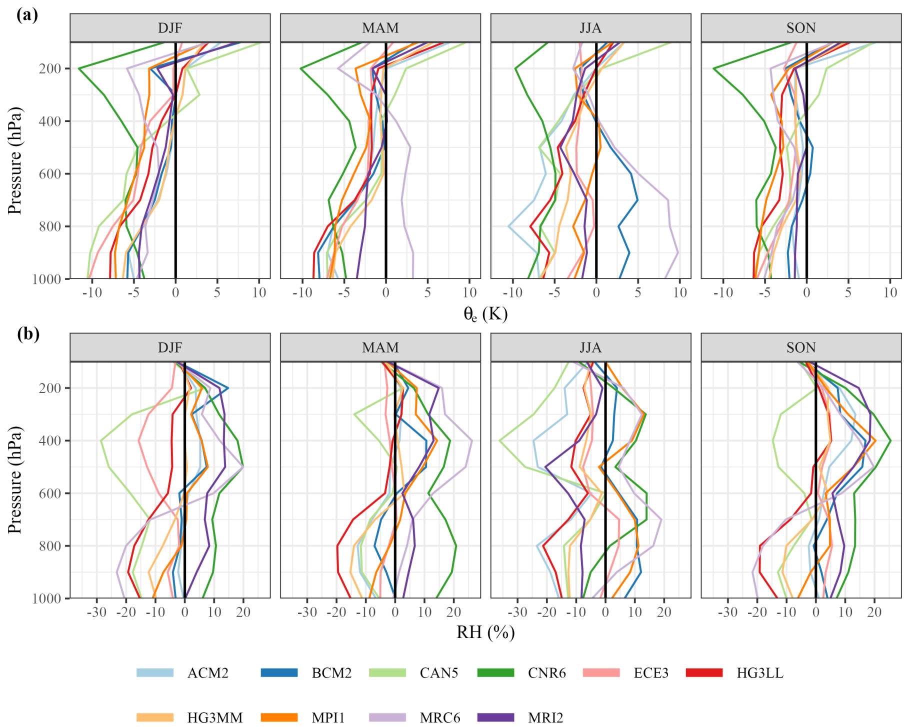

Figure 8. Bias in the mean (a) θ_e and (b) RH profiles for CEs aggregated seasonally for ERA5 and CMIP6 models.

The mean seasonal vertical profiles of θ_e are shown in Figure 8a. Most CMIP6 models reproduce CEs with colder and drier profiles compared to ERA5. Exceptions include MRC6, which shows warmer and moister air masses during MAM and JJA, and BCM2 in JJA. Notably, upper levels above 200 hPa are associated with warmer and moister air masses than ERA5 in most models. On the diurnal scale, CEs in CMIP6 are associated with a colder, drier atmosphere compared to ERA5, except in MRC6 in the afternoon (Fig. S10a).

Turning now to relative humidity (RH, Fig. 8b), a distinct pattern emerges: most models show a negative bias in the lower troposphere. In contrast, the upper troposphere is largely characterised by a positive bias, except in JJA. In summer, the positive bias around 400 hPa is either absent or weak, and a negative bias persists throughout most of the profile in most models. Furthermore, CAN5 and HG3LL show negative bias across most levels in all seasons, while CNR6 and MRI2 are positively biased, with negative bias observed in 1-3 of 10 levels. The diurnal RH bias is illustrated in Fig. S10b, showing a consistent pattern across all timesteps: CAN5, ECE3, HG3LL, and HG3MM exhibit a dry bias across most levels, while CNR6 and MRI2 remain predominantly positively biased.

The difference between MM ensemble and ERA5 in the joint MUCAPE-MUCIN density is illustrated in Fig. 9a. Two patterns are evident. First, CMIP6 underestimates MUCAPE, with the underestimation regimes centred around 1000 $\mathrm{J\,kg^{-1}}$ MUCAPE and ∼70 $\mathrm{J\,kg^{-1}}$ MUCIN in DJF, and shifting to ∼200 $\mathrm{J\,kg^{-1}}$ MUCIN in MAM. In JJA, the MM fail to capture

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

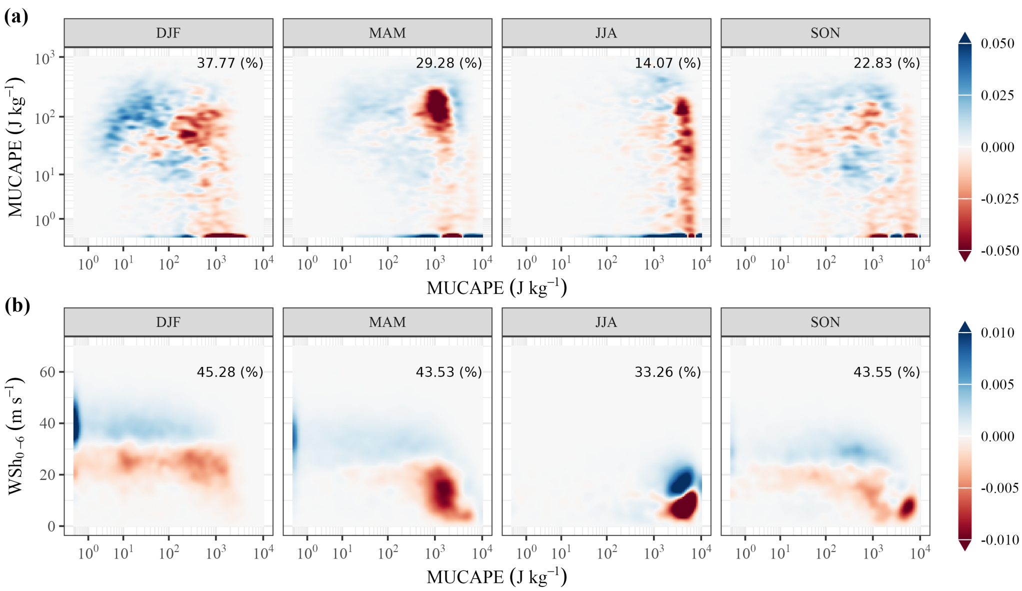

Figure 9. Seasonal difference in the bivariate probability density of CEs between the MM mean and ERA5 at the current warming level: (a) MUCAPE vs MUCIN, and (b) MUCAPE vs $WSh_{0\text{-}6}$. Note the $\log_{1p}$ transformation for MUCAPE and MUCIN axes. Percentages at the top right of each subfigure report the cumulative probability for areas with an increase in occurrence with at least 70% model agreement

a specific MUCAPE regime centred around 7000 J kg^{-1}, including a wide range of MUCIN. In SON, the underestimation is more distributed with no clear centre. Second, in near-zero MUCIN environments, CMIP6 underestimates MUCAPE in DJF and shows mixed over- and underestimation in other seasons.

Figure 9b shows the MM difference to ERA5 in the joint MUCAPE-$WSh_{0\text{-}6}$ density. Generally, CEs occur with stronger $WSh_{0\text{-}6}$ in the MM represented by the horizontal separation. MUCAPE appears to be underestimated across the full range in DJF, while in JJA, the underestimation and overestimation blobs are vertically aligned at similar MUCAPE values, suggesting that MUCAPE is approximately correct while shear is overestimated. MAM and SON represent transition seasons in which the blobs and the broad range of biases coexist.

The diurnal cycle of biases in the joint MUCAPE-MUCIN and MUCAPE-$WSh_{0\text{-}6}$ densities broadly echoes the seasonal patterns, with 03:00 UTC+3, 09:00 UTC+3, 15:00 UTC+3, and 21:00 UTC+3 resembling DJF, SON, JJA, and MAM, respectively (Fig. S11a). A similar correspondence is encountered for MUCAPE-$WSh_{0\text{-}6}$, where morning periods resemble DJF, while afternoon and evening periods reflect a mixture of MAM and SON characteristics (Fig. S11b).

3.4 Projected change in convection indices

In the results' final part, we evaluate changes in convection indices associated with CEs in future climates. PW and Td_{2m} increase monotonically across both future climates, both seasonally (Fig. 10) and diurnally (Fig. S12), with high model agree-

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

ment. Furthermore, PW (Td_{2m}) shows the largest increase in JJA (MAM) in both future climates (Table S7). On the diurnal scale, the increases are consistent with marginal differences across periods (Table S8). Seasonally, ω_{500} associated with CEs generally increases with rising temperature (Fig. 10c); however, it shows a large model disagreement during MAM in both futures (Table S7). Furthermore, the increase in JJA largely deviates from those in other seasons. At *FutI*, 03:00 UTC+3 and 09:00 UTC+3 show an increase in ω_{500}, while a decrease is projected for 15:00 UTC+3 and 21:00 UTC +3 (Table S8 and Fig. S12). *FutIII* shows further increase in ω_{500}; however, with low model agreement at 03:00 UTC+3 and 21:00 UTC+3.

The largest MUCAPE is observed in JJA, and the smallest in DJF, in both the current and future climates (Fig. 10 and Table S7). However, the largest increase occurs in MAM (DJF) at the *FutI* (*FutIII*) warming level by 20.89% (69.29%). SON depicts an increase, with model agreement below 70% in both future climates. An increase at both warming levels characterises diurnal MUCAPE changes (Fig. S12 and Table S8). MUCIN in the current climate exhibits the largest values in MAM (Fig. 10) and at 03:00 UTC+3 (Fig. S12). In a warming Earth atmosphere, MUCIN shows a persistent decrease across seasonal and diurnal scales (Table S7 and S8). Regarding $WSh_{0\text{-}6}$, the highest (lowest) values are observed in DJF (JJA; Fig. 10). A consistent decrease in $WSh_{0\text{-}6}$ is projected in both futures during DJF and MAM, while JJA exhibits an increase signal. The trends in SON depict below 70% model agreement (Table S7). Subdaily, all periods show a decrease in $WSh_{0\text{-}6}$ at *FutI* and mixed-nonrobust signals at *FutIII* (Fig. S12 and Table S8).

The seasonal (diurnal) changes in the mean vertical profiles of θ_e and RH are shown in Fig. 11 (Fig. S13). In both future climates, θ_e associated with CEs increases across seasons and periods, with the enhancement signal strengthening consistently from 100 hPa toward 1000 hPa; yet the largest uncertainties are evident in the lower troposphere. Uncertainties are larger during SON and at 21:00 UTC+3. In terms of RH, large uncertainties predominate in the upper troposphere. In DJF and MAM, the increase in RH is almost identical across future climates at 400 hPa-900 hPa; this behaviour also holds for SON at around 700 hPa. In JJA and at 15:00 UTC+3, a clear increase is observable from *FutI* to *FutIII*.

The change in the seasonal bivariate probability density of MUCAPE-MUCIN is mainly MUCAPE-driven, with CEs being associated with larger MUCAPE values in both future climates (Fig. 12a). These patterns are also evident on the subdaily scale (Fig. S14a). The joint MUCAPE-$WSh_{0\text{-}6}$ distribution shows a shift towards higher MUCAPE and lower $WSh_{0\text{-}6}$ across seasons except in JJA, when rather higher shear is projected (Fig. 12b). Similarly, diurnal MUCAPE-$WSh_{0\text{-}6}$ are shifting towards higher MUCAPE and low $WSh_{0\text{-}6}$ regimes (Fig. S14b).

4 Discussion

Our findings indicate that XGB is the optimal model to identify CEs in the region. Additionally, the sensitivity analysis shows that XGB performance remains relatively stable compared to other ML models as the threshold for precipitating convection increases, suggesting that the XGB can potentially be used to examine extreme events (Fig. S1). Performance metrics demonstrate that XGB spatiotemporally outperforms other ML models (Fig. 2 and S3-S5); however, the ECE is larger for XGB (Sect. 3.1). This is a known issue in supervised learning, in which a classifier can draw perfect decision boundaries separating convective from non-convective situations (i.e., discrimination), but produces miscalibrated probabilities (Jiang et al., 2012).

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

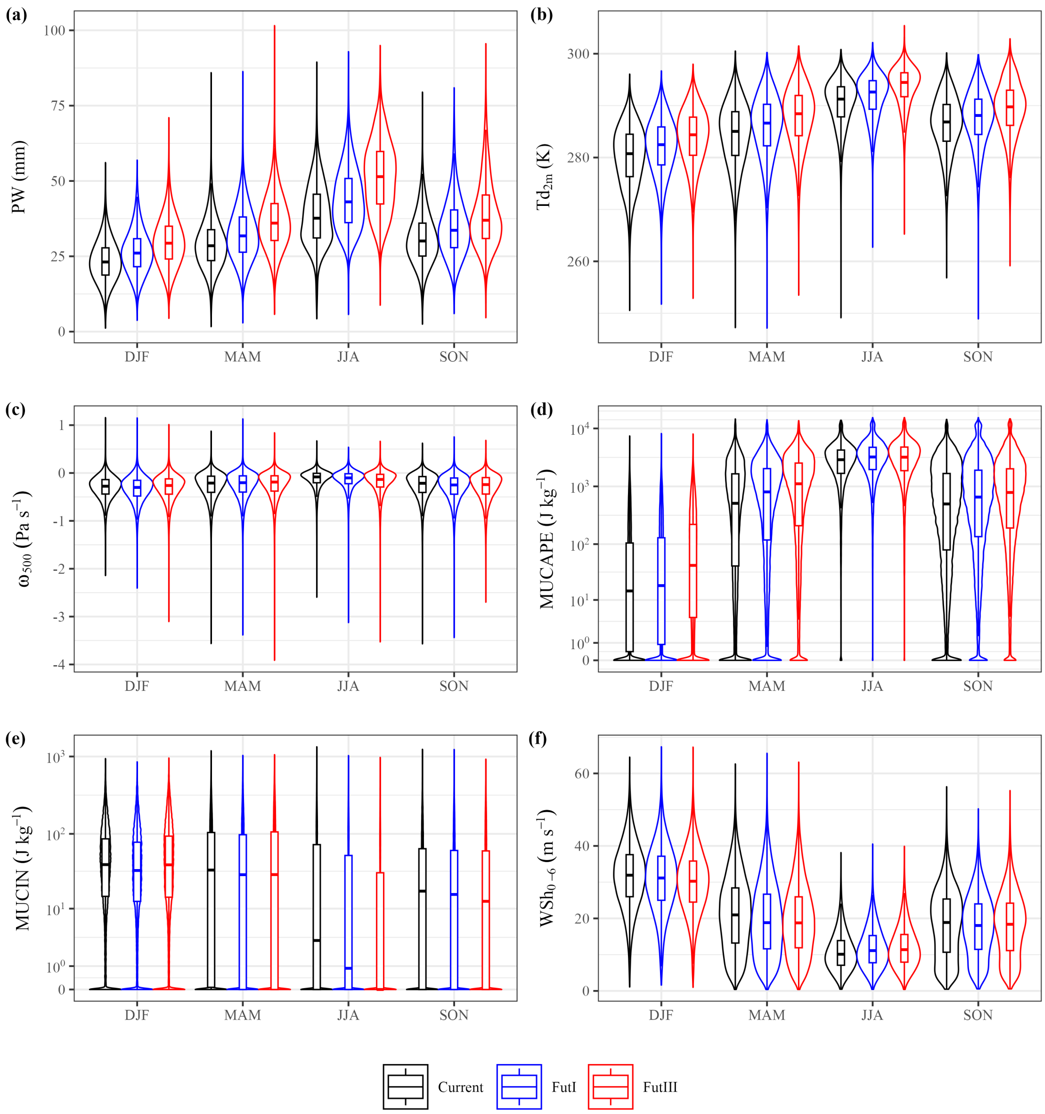

Figure 10. Seasonal box- and violin plots of the pooled MM ensemble distribution for CE components at current and future regional warming levels. (a) PW, (b) Td_{2m}, (c) ω_{500}, (d) MUCAPE, (e) MUCIN, and (f) $WSh_{0\text{-}6}$. Note the $\log_{1p}$ x-axis in (d) and (e). At least 70% of models show a statistically significant difference in mean between the current and future periods ($p < 0.05$) in all subfigures.

One unanticipated finding was the deviation in the RES curve for XGB (Fig. 2b), which could indicate a signature of a model with high discriminating skill in the 0.6-0.9 range. To illustrate, it means that XGB in this range can make skilful predictions and successfully separate high-probability cases from the climatological average. It can also be postulated that CEs in this range are associated with conditions that strongly deviate from the typical state, and that XGB can identify them.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

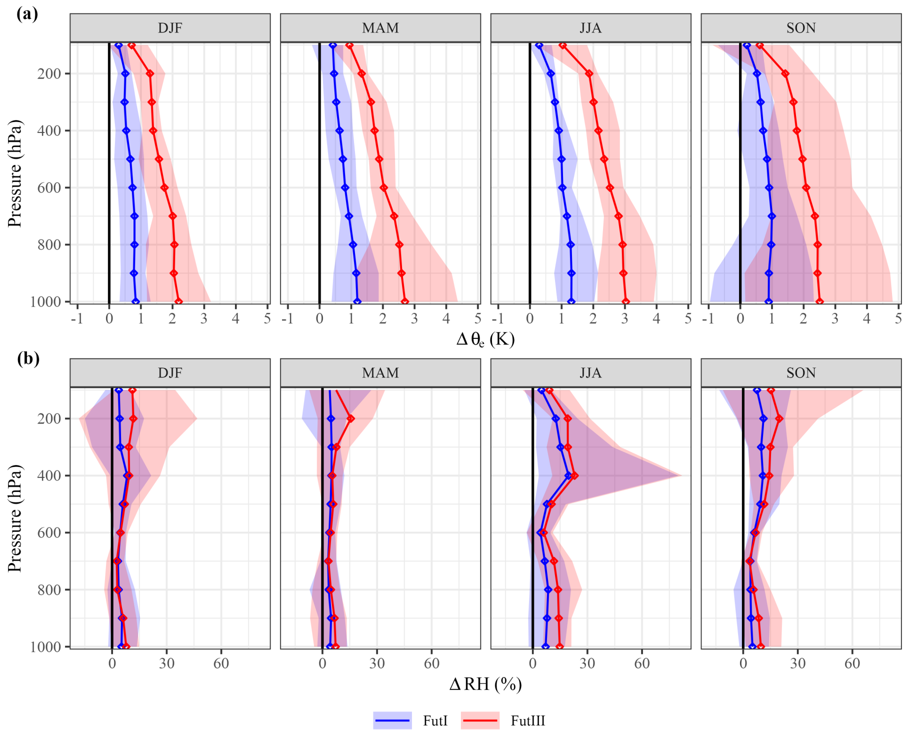

Figure 11. Seasonal change in the MM mean (a) θ_e and (b) RH profiles for CEs. The shaded areas indicate the 2.5th and 97.5th percentile range of the CMIP6 ensemble, and the diamond shapes indicate at least 70% model agreement on the sign of change.

ML model performs the best in the the northern AP and during cold months or cold periods of the day likely because these convective situations are influenced by large-scale environments such as frontal mid-latitude systems (Almazroui et al., 2015, 2016; Luong et al., 2024; Charabi and Al-Hatrushi, 2010), the Red Sea Convergence Zone, the Red Sea Trough (Dasari et al., 2018; de Vries et al., 2013; Khan et al., 2018), and the subtropical jet (Taraphdar et al., 2025). Conversely, performance degrades where convection is localised, such as in the southern AP or during summer and afternoon hours. Similar findings are reported by Gómez-Navarro et al. (2022) over complex terrain in Switzerland. Over the AP, localised convection is influenced by orographic forcing (Abdullah and Al-Mazroui, 1998; Khan et al., 2018; Cook et al., 2021; Branch et al., 2020; Farquharson et al., 1996), interaction between sea-breeze and surface heating or cold pools (Fonseca et al., 2025; Nelli et al., 2022; Eager et al., 2008; Gopalakrishnan et al., 2023), and the remnants of low-pressure systems from the ISM (Charabi, 2009; Babu et al., 2016). These processes are likely poorly represented by convection indices at 1°resolution. Furthermore, an instantaneous profile is labelled convective based on accumulated precipitation; however, this profile may not accurately represent the atmospheric state when localised convection occurs far from the profile timestamp.

CMIP6 models tend to underestimate the probability and occurrence of CEs (Fig. 4-6 and S6-S8), particularly by underestimating (overestimating) the high (low) probability of CEs (Fig. 3 and S6). A plausible explanation lies in the systematic

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

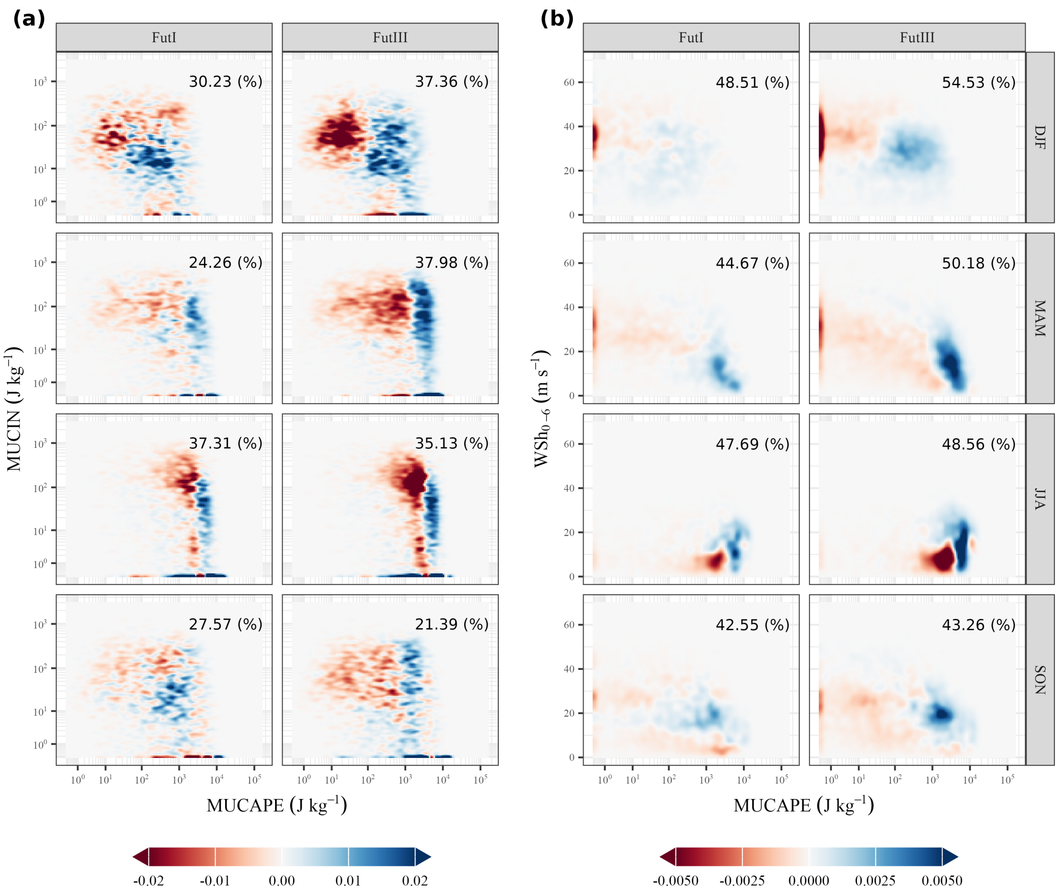

Figure 12. Seasonal change in the bivariate probability density of CEs between the MM mean at current warming and future warming levels: a) MUCAPE vs MUCIN, and b) MUCAPE vs WSh$_{0\text{-}6}$. Note the $\log_{1p}$ transformation for MUCAPE and MUCIN axes. Percentages at the top right of each subfigure report the cumulative probability for areas with an increase in occurrence with at least 70% model agreement.

drizzle bias pervasive in climate models, whereby precipitation occurs too lightly and too frequently (Chen et al., 2021; Hagos et al., 2021; Martinez-Villalobos et al., 2022). Furthermore, Hagos et al. (2021) reported that CMIP6 models require anomalously high moisture accumulation to transition from drizzle to intense convection (Kuo et al., 2020; Hagos et al., 2021). This is further supported by the negatively biased PW, Td$_{2m}$, and θ_e profiles associated with CEs (Fig. 7a,b and 8), reflecting a systematically drier and less buoyant CEs in CMIP6 models. Additionally, convection in cold months over the AP exhibits a secondary nocturnal peak (Homoudi et al., 2026). Most CMIP6 models struggle to represent nocturnal elevated convection above the boundary layer (Tang et al., 2021), providing an additional plausible explanation for the underestimation of CEs' occurrence in CMIP6. Moreover, CMIP6 models may inherit the CMIP5 tendency to underestimate (overestimate) weather types favouring wet (dry) conditions over the region (e.g., El Kenawy and McCabe, 2017). A detailed weather type analysis is outside the scope of this study.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

MUCAPE is underestimated (Fig. 7d and 9), owing to both the pervasive dry bias and the premature release of instability before sufficient thermodynamic buildup can occur (Chen and Dai, 2019; Chen et al., 2021). The overestimation of MUCIN (Fig. 7e) is consistent with Hagos et al. (2021), who demonstrated that most CMIP6 models exhibit stronger convective inhibition than ERA5. The stronger $WSh_{0\text{-}6}$ associated with CEs can be attributed to the positive bias in the 700-400 unithPa layer-mean wind in CMIP6 models relative to ERA5 over the region (Fig. S15 and S16). In general, CMIP6 models exhibit varying degrees of bias and differing behaviour across the distribution of convective indices. Further research should focus on a dedicated, comprehensive evaluation of convective indices and instability parameters over the AP using multiple reanalysis datasets and radiosonde data (e.g., Barfus and Bernhofer, 2015; Gopalakrishnan et al., 2025), given that ERA5 itself may carry uncertainties in the upper atmosphere due to the sparse upper-air observations (Taszarek et al., 2021).

Findings regarding the underestimation (overestimation) of MUCAPE ($WSh_{0\text{-}6}$) associated with CEs contradict previous studies (e.g., Lepore et al., 2021; Chavas and Li, 2022; Gopalakrishnan et al., 2025). However, direct comparison is precluded by key methodological differences. CE definitions differ across studies: all atmospheric profiles are considered in Gopalakrishnan et al. (2025) and Lepore et al. (2021), severe thunderstorm environments are defined by a fixed CAPE$\times WSh_{0\text{-}6}$ threshold in Chavas and Li (2022), whereas the present study investigates CEs identified by XGB. Furthermore, Lepore et al. (2021) and Gopalakrishnan et al. (2025) employ mixed-layer CAPE, and Chavas and Li (2022) surface-based CAPE, whereas the present study utilises the most-unstable parcel formulation (Haklander and Van Delden, 2003).

The consistent increase in trends during DJF could be due to the southward shift and intensification of the subtropical jet, the rise in transient activity, and moisture influx from adjacent water bodies that are expected under a warming climate (Pathak et al., 2025). The low-level jet of the ISM is expected to shift poleward (Sandeep and Ajayamohan, 2015), and this is reflected in the increase in CE during JJA at *FutI*, as well as the monotonic $WSh_{0\text{-}6}$ increase in summer in both futures. The physical mechanisms driving the persistent decrease in CEs during MAM in both futures remain unclear. However, several plausible hypotheses can be postulated: (a) precipitation regimes are shifting into the colder DJF; (b) the coarse resolution of CMIP6 models might prevent adequate representation of the unique moisture influx associated with tropical–extratropical interactions during MAM (Luong et al., 2024); (c) warming of adjacent seas may be insufficiently rapid to satisfy the increased atmospheric moisture demand (Allan et al., 2020); d) CEs are shifting towards nocturnal periods, when CMIP6 struggle with nocturnal elevated convection (Tang et al., 2021). A detailed process-level investigation of the mechanisms driving MAM CE changes is strongly recommended as a priority for future research.

The widespread increase at 03:00, 09:00, and 21:00 UTC+3 at *FutI* (Fig. S7b), along with the decrease at 15:00 UTC+3, aligns with the findings of Meredith et al. (2019). They reported a projected shift in heavy to extreme precipitation occurrence from late afternoon to overnight and morning periods.

PW, Td_{2m}, and θ_e profiles associated with CEs depict a monotonic increase with warming (Fig. 10a,b and 11a). It could be explained by the CC scaling relationship: a warming Earth atmosphere has a higher water-holding capacity Trenberth (2011). MUCAPE is increasing, while $WSh_{0\text{-}6}$ is decreasing under future warming. The results are in line with previous studies (Lepore et al., 2021; Riemann-Campe et al., 2009; Chen et al., 2020). MUCIN associated with CEs is decreasing under future warming. The AP is one of the few regions showing a decrease in convection inhibition (Chen et al., 2020). Furthermore, more low-level

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

moisture leads to lower lifting condensation levels and thus partly results in smaller CIN (Gopalakrishnan et al., 2025). The RH generally does not increase with warming (Fig. 11b and S10b), as stated by previous studies (Douville et al., 2022; Allan et al., 2020; Sherwood et al., 2010), suggesting that global warming occurs under a constant tropospheric relative humidity.

At *FutIII*, CE intensification generally indicate that precipitation over the AP is likely to become more intense, as established by previous studies (Almazroui and Saeed, 2020; Terry et al., 2023), suggesting that the number of wet days will decrease, and the contribution of extreme precipitation to the total amount will increase. The decrease in CE occurrence is expected and consistent with the slower ocean warming rate, which limits moisture supply to maintain continental relative humidity, leading to a drying influence further amplified by land surface feedbacks (Allan et al., 2020). Moreover, the observed peak-like scaling of extreme precipitation with surface temperature north of 20° N and a decreasing relationship south of 20° N by Utsumi et al. (2011) supports the decrease in CE occurrence at *FutIII* warming.

We acknowledge the limitations of the IMERG dataset over the region. Earlier versions of IMERG exhibit a high false alarm rate and a tendency to overestimate light precipitation (e.g., Hussein et al., 2021). Precipitation values greater than $1\,\mathrm{mm}\,6\,\mathrm{h}^{-1}$) were used to define the local climatological median, excluding sporadic light rain. Furthermore, the ML models were trained on a binary flag indicating whether accumulated precipitation exceeded this median, rather than on precipitation quantities.

Another limitation of the study is the absence of a continuous response of CEs to regional temperature increase, as shown in Fig. 6 by Lepore et al. (2021). Unlike Lepore et al. (2021), who analysed CEs generally, this study restricts CEs to those identified by the ML models, requiring bias-corrected indices as predictors. However, applying bias correction across continuous warming levels with overlapping years is computationally inefficient and contradicts the study's goal of reducing the computational footprint. Bias correction methods come with varying strengths and limitations (Sharma et al., 2026); for example, they can introduce temporal jumps (Rust et al., 2015) or degrade the climate model output. A sensitivity analysis to determine the effect of bias correction on CEs is out the scope of the study, particularly since precipitation intensity analysis will be addressed with CPMs.

An arguable weakness is the inconsistent number of years across CMIP6 models (Table 2) that represent the current regional warming observed in ERA5 (2001-2024), a constraint imposed by the length of the IMERG record and the necessity to avoid mixing different experiments Lepore et al. (2021). IMERG is employed to obtain the convection label in preparation for dynamical downscaling, wherein it serves as the observational benchmark for Lagrangian evaluation of convective precipitation systems.

Future research should focus on investigating the CC scaling from CPMs to determine the origin of super-CC scaling in the region (Tuel et al., 2022) and whether it is a statistical shift from stratiform to convective precipitation (Da Silva and Haerter, 2025) is recommended. CEs are sensitive to model biases (Allen et al., 2014b); therefore, an in-depth analysis of biases in moist static energy, weather types and circulation (e.g., El Kenawy and McCabe, 2017), the diurnal cycle of precipitation (e.g., Tang et al., 2021), and soil moisture over the region is imperative. There is ample room for further progress in improving ML model skill by training the models regionally or seasonally, or by testing a 'leave one out' approach to identify locally significant predictors. Further research should test the skill of climate modes of variability indices in identifying CEs over the region.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

5 Summary and conclusions

The study set out to infer the occurrence of precipitating convection over the AP in the current climate using ML models, and to utilise the best-performing ML model to identify potential CEs in CMIP6 data. Furthermore, the study aimed to evaluate CMIP6 models' ability to reproduce the spatiotemporal distribution and the components of CEs over the AP and to quantify changes under a warming climate. In doing so, the study addresses a critical gap in the literature by focusing on convectively favourable conditions rather than the precipitation itself, which is largely governed by CMIP6 parameterisation schemes and therefore of limited reliability over the region (e.g., Pathak et al., 2025; Zhang et al., 2025a). Although the ML models are trained to detect CE or profiles that deliver rainfall above the local median, it is still possible that a CE does not rain. However, the study's approach helps define an approximate frequency for those most likely to result in wet conditions without relying on parameterised precipitation.

The CEs dataset (Homoudi, 2026) produced here lays the foundation for a statistical-dynamical downscaling framework (e.g., Beaulant et al., 2011; Meredith et al., 2018; Wang et al., 2026), wherein identified CEs can be dynamically downscaled at convection-permitting resolution to produce physically consistent regional projection ensembles over the AP while reducing the computational cost. This approach provides an efficient pathway for further scrutiny of the impact of climate change on convective precipitation systems over the AP. Our future work will focus on dynamically downscaling the CEs from the CMIP6 ensemble to O(1km), opening avenues to investigate the Lagrangian properties of convective systems and how land surface characteristics, topography, and climate change influence them.

Beyond climate applications, the ML framework presented here can be further exploited for operational nowcasting and thunderstorm prediction. For instance, an ML model can be trained on radar and ERA5 data to detect extreme events. Consequently, it can be applied to forecast products (e.g., ECMWF HRES), triggering dynamical downscaling only where needed to produce high-resolution risk maps. Such an approach reduces the need to secure adequate computational resources (Hoteit et al., 2025), especially when convection-permitting models fail to capture MCSs over the region (Francis et al., 2021).

At the current level of regional warming (+1.22 °C), CMIP6 models systematically underestimate CE probability and occurrence. This bias is rooted in a constellation of systematic model deficiencies: drier environments, reduced instability, elevated convection inhibition, and compressed vertical velocity distribution. Together, these reflect the tendency of CMIP6 models to operate too frequently in the drizzle regime, requiring anomalously high moisture to transition into organised deep convection (Hagos et al., 2021). Notably, CMIP6 overestimate the 700-400 hPa layer-mean wind and consequently, vertical wind shear is positively biased. Furthermore, we do not explicitly rank CMIP6 models, as all models considered in the study exhibit varying thermodynamic and kinematic biases; selection is therefore application-dependent. As argued by Gopalakrishnan et al. (2025), an ensemble approach is recommended.

At *FutI* (+2.22 °C), CE occurrence generally increases across seasons, excluding the spring, driven by enhanced dynamic and thermodynamic favourability and increased moisture transport from adjacent seas. Furthermore, a diurnal shift towards nocturnal or morning CE occurrence is expected. At *FutIII* (+4.22 °C), CE occurrence decreases across all seasons, except

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

in winter. However, in both future climates, the CEs' components intensify with increasing moisture and instability, and with reduced inhibition and vertical wind shear.

The projected intensification of CEs under future warming suggests that precipitation over the AP is likely to intensify. While it can benefit water resources in the arid AP, it heightens the risk of flash floods and landslides, threatening vulnerable communities, including those engaged in rainfed terrace agriculture in Yemen's highlands (AL-Falahi et al., 2024a, b). The persistent decrease in CE frequency during spring is particularly alarming, given that this is the predominant season of MCS activity (Nelli et al., 2021) and that Saudi Arabia receives approximately 60% of its annual rainfall from these systems (Luong et al., 2024). The physical mechanisms driving this decrease and their implications for regional water resources warrant a dedicated process-oriented investigation.

Code and data availability. IMERG V07 (Huffman et al., 2023a) is obtained from the NASA Goddard Earth Sciences Data and Information Services Centre (https://doi.org/10.5067/GPM/IMERG/3B-HH/07). ERA5 reanalysis data (Hersbach et al., 2020) is available at the Copernicus Climate Change Service (C3S) Climate Data Store (https://cds-beta.climate.copernicus.eu/). CMIP6 data is available from the Earth System Grid Federation nodes (ESGF; e.g., https://esgf-node.llnl.gov/search/cmip6/). Convection indices from ERA5 and CE labels from IMERG are published at Homoudi (2026, ; https://doi.org/10.5281/zenodo.18514414). This repository includes the trained ML models (i.e., RF, XGB, and DL), the convection indices from CMIP6, and the predicted-calibrated probabilities from CMIP6 for the periods shown in Table 2. The containerised software environment to replicate this study is also hosted at Homoudi (2026). It contains all the software mentioned in Sect. 2.1.

Author contributions. Ahmed Homoudi: conceptualization, formal analysis, investigation, methodology, data curation, software, writing – original draft, writing – review and editing, visualization. Henning W. Rust: methodology, supervision, writing – review and editing. Klemens Barfus: conceptualization, methodology, supervision, writing – review and editing. Christian Bernhofer: resources, supervision, writing – review and editing. Matthias Mauder: resources, supervision, writing – review and editing.

Competing interests. Some authors are members of the editorial board of Natural Hazards and Earth System Sciences. Otherwise, the authors declare that they have no conflict of interest.

Disclaimer. The authors received no specific funding for this work.

Acknowledgements. The authors gratefully acknowledge the computing time made available to them on the high-performance computer at the NHR Center of TU Dresden. This center is jointly supported by the Federal Ministry of Research, Technology and Space of Germany and the state governments participating in the NHR (www.nhr-verein.de/unsere-partner). Ahmed Homoudi is grateful for the warm hospitality and fruitful discussions during his visit to the Institute of Meteorology at FU Berlin, Germany. We thank the anonymous reviewers for thoroughly reviewing our manuscript and providing valuable feedback. The authors utilised Grammarly AI Premium to enhance readability and refine specific sentences in this work. Afterwards, the authors carefully reviewed and refined the content, taking full responsibility for the final publication. Open Access funding enabled and organised by Projekt DEAL.

References

Abdullah, M. and Al-Mazroui, M.: Climatological Study of the Southwestern Region of Saudi Arabia. I. Rainfall Analysis, Climate Research, 9, 213–223, https://doi.org/10.3354/cr009213, 1998.

AL-Falahi, A. H., Saddique, N., Spank, U., and Bernhofer, C.: Building High-Resolution Projections of Temperature Potential Changes Using Statistical Downscaling for the Future Period 2026–2100 in the Highland Region of Yemen – A Supportive Approach for Empowering Environmental Planning and Decision-Making, Environmental and Sustainability Indicators, 24, 100 490, https://doi.org/10.1016/j.indic.2024.100490, 2024a.

AL-Falahi, A. H., Saddique, N., Spank, U., Pluntke, T., Gebrechorkos, S. H., Mauder, M., and Bernhofer, C.: Hydrological Investigation of Climate Change Impact on Water Balance Components in the Agricultural Terraced Watersheds of Yemeni Highland, Theoretical and Applied Climatology, 155, 4703–4720, https://doi.org/10.1007/s00704-024-04913-x, 2024b.

Allan, R. P., Barlow, M., Byrne, M. P., Cherchi, A., Douville, H., Fowler, H. J., Gan, T. Y., Pendergrass, A. G., Rosenfeld, D., Swann, A. L. S., Wilcox, L. J., and Zolina, O.: Advances in Understanding Large-Scale Responses of the Water Cycle to Climate Change, Annals of the New York Academy of Sciences, 1472, 49–75, https://doi.org/10.1111/nyas.14337, 2020.

Allen, J. T.: Climate Change and Severe Thunderstorms, in: Oxford Research Encyclopedia of Climate Science, Oxford University Press, ISBN 978-0-19-022862-0, https://doi.org/10.1093/acrefore/9780190228620.013.62, 2018.

Allen, J. T., Karoly, D. J., and Walsh, K. J.: Future Australian Severe Thunderstorm Environments. Part II: The Influence of a Strongly Warming Climate on Convective Environments, Journal of Climate, 27, 3848–3868, https://doi.org/10.1175/JCLI-D-13-00426.1, 2014a.

Allen, J. T., Karoly, D. J., and Walsh, K. J.: Future Australian Severe Thunderstorm Environments. Part I: A Novel Evaluation and Climatology of Convective Parameters from Two Climate Models for the Late Twentieth Century, Journal of Climate, 27, 3827–3847, https://doi.org/10.1175/JCLI-D-13-00425.1, 2014b.

Almazroui, M. and Saeed, S.: Contribution of Extreme Daily Precipitation to Total Rainfall over the Arabian Peninsula, Atmospheric Research, 231, 104 672, https://doi.org/10.1016/j.atmosres.2019.104672, 2020.

Almazroui, M., Dambul, R., Islam, M. N., and Jones, P.: Atmospheric Circulation Patterns in the Arab Region and Its Relationships with Saudi Arabian Surface Climate: A Preliminary Assessment, Atmospheric Research, 161–162, 36–51, https://doi.org/10.1016/j.atmosres.2015.03.014, 2015.

Almazroui, M., Kamil, S., Ammar, K., Keay, K., and Alamoudi, A. O.: Climatology of the 500-hPa Mediterranean Storms Associated with Saudi Arabia Wet Season Precipitation, Climate Dynamics, 47, 3029–3042, https://doi.org/10.1007/s00382-016-3011-0, 2016.

Almazroui, M., Islam, M. N., Saeed, S., Saeed, F., and Ismail, M.: Future Changes in Climate over the Arabian Peninsula Based on CMIP6 Multimodel Simulations, Earth Systems and Environment, 4, 611–630, https://doi.org/10.1007/s41748-020-00183-5, 2020.

Attada, R., Dasari, H. P., Parekh, A., Chowdary, J. S., Langodan, S., Knio, O., and Hoteit, I.: The Role of the Indian Summer Monsoon Variability on Arabian Peninsula Summer Climate, Climate Dynamics, 52, 3389–3404, https://doi.org/10.1007/s00382-018-4333-x, 2019.

Babu, C. A., Jayakrishnan, P. R., and Varikoden, H.: Characteristics of Precipitation Pattern in the Arabian Peninsula and Its Variability Associated with ENSO, Arabian Journal of Geosciences, 9, 186, https://doi.org/10.1007/s12517-015-2265-x, 2016.

Barfus, K. and Bernhofer, C.: Assessment of GCM Capabilities to Simulate Tropospheric Stability on the Arabian Peninsula: GCM SIMULATIONS OF TROPOSPHERIC STABILITY ON THE ARABIAN PENINSULA, International Journal of Climatology, 35, 1682–1696, https://doi.org/10.1002/joc.4092, 2015.

Beaulant, A.-L., Joly, B., Nuissier, O., Somot, S., Ducrocq, V., Joly, A., Sevault, F., Deque, M., and Ricard, D.: Statistico-Dynamical Downscaling for Mediterranean Heavy Precipitation, Quarterly Journal of the Royal Meteorological Society, 137, 736–748, https://doi.org/10.1002/qj.796, 2011.

Branch, O., Behrendt, A., Gong, Z., Schwitalla, T., and Wulfmeyer, V.: Convection Initiation over the Eastern Arabian Peninsula, Meteorologische Zeitschrift, 29, 67–77, https://doi.org/10.1127/metz/2019/0997, 2020.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Cannon, A. J.: Multivariate Quantile Mapping Bias Correction: An N-dimensional Probability Density Function Transform for Climate Model Simulations of Multiple Variables, Climate Dynamics, 50, 31–49, https://doi.org/10.1007/s00382-017-3580-6, 2018.

Catto, J. L. and Pfahl, S.: The Importance of Fronts for Extreme Precipitation: FRONTS AND PRECIPITATION EXTREMES, Journal of Geophysical Research: Atmospheres, 118, 10,791–10,801, https://doi.org/10.1002/jgrd.50852, 2013.

Charabi, Y.: Arabian Summer Monsoon Variability: Teleconexion to ENSO and IOD, Atmospheric Research, 91, 105–117, https://doi.org/10.1016/j.atmosres.2008.07.006, 2009.

Charabi, Y. and Al-Hatrushi, S.: Synoptic Aspects of Winter Rainfall Variability in Oman, Atmospheric Research, 95, 470–486, https://doi.org/10.1016/j.atmosres.2009.11.009, 2010.

Chavas, D. R. and Li, F.: Biases in CMIP6 Historical U.S. Severe Convective Storm Environments Driven by Biases in Mean-State Near-Surface Moist Static Energy, Geophysical Research Letters, 49, e2022GL098 527, https://doi.org/10.1029/2022GL098527, 2022.

Chen, D. and Dai, A.: Precipitation Characteristics in the Community Atmosphere Model and Their Dependence on Model Physics and Resolution, Journal of Advances in Modeling Earth Systems, 11, 2352–2374, https://doi.org/10.1029/2018MS001536, 2019.

Chen, D., Dai, A., and Hall, A.: The Convective-To-Total Precipitation Ratio and the “Drizzling” Bias in Climate Models, Journal of Geophysical Research: Atmospheres, 126, e2020JD034 198, https://doi.org/10.1029/2020JD034198, 2021.

Chen, J., Dai, A., Zhang, Y., and Rasmussen, K. L.: Changes in Convective Available Potential Energy and Convective Inhibition under Global Warming, Journal of Climate, 33, 2025–2050, https://doi.org/10.1175/JCLI-D-19-0461.1, 2020.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, https://doi.org/10.1145/2939672.2939785, 2016.

Cheung, K. K. W., Ji, F., Evans, J. P., Nishant, N., Herold, N., di Virgilio, G., Beyer, K., and Riley, M. L.: Projected Convective Storm Environment in the Australian Region from Two Downscaling Ensemble Systems Under the SRES-A2/RCP8.5 Scenarios, Climate, 13, 229, https://doi.org/10.3390/cli13110229, 2025.

Chollet, F. et al.: Keras, https://keras.io, 2015.

Cook, K. H., Vizy, E. K., Liu, Y., and Liu, W.: Greenhouse-Gas Induced Warming Amplification over the Arabian Peninsula with Implications for Ethiopian Rainfall, Climate Dynamics, 57, 3113–3133, https://doi.org/10.1007/s00382-021-05858-x, 2021.

Corre, L., Ribes, A., Bernus, S., Drouin, A., Morin, S., and Soubeyroux, J.-M.: Using Regional Warming Levels to Describe Future Climate Change for Services and Adaptation: Application to the French Reference Trajectory for Adaptation, Climate Services, 38, 100 553, https://doi.org/10.1016/j.cliser.2025.100553, 2025.

Da Silva, N. A. and Haerter, J. O.: Super-Clausius–Clapeyron Scaling of Extreme Precipitation Explained by Shift from Stratiform to Convective Rain Type, Nature Geoscience, 18, 382–388, https://doi.org/10.1038/s41561-025-01686-4, 2025.

Dasari, H. P., Langodan, S., Viswanadhapalli, Y., Vadlamudi, B. R., Papadopoulos, V. P., and Hoteit, I.: ENSO Influence on the Interannual Variability of the Red Sea Convergence Zone and Associated Rainfall, International Journal of Climatology, 38, 761–775, https://doi.org/10.1002/joc.5208, 2018.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Davis, J. and Goadrich, M.: The Relationship between Precision-Recall and ROC Curves, in: Proceedings of the 23rd International Conference on Machine Learning - ICML '06, pp. 233–240, ACM Press, Pittsburgh, Pennsylvania, ISBN 978-1-59593-383-6, https://doi.org/10.1145/1143844.1143874, 2006.

de Burgh-Day, C. O. and Leeuwenburg, T.: Machine Learning for Numerical Weather and Climate Modelling: A Review, Geoscientific Model Development, 16, 6433–6477, https://doi.org/10.5194/gmd-16-6433-2023, 2023.

de Vries, A. J., Tyrlis, E., Edry, D., Krichak, S. O., Steil, B., and Lelieveld, J.: Extreme Precipitation Events in the Middle East: Dynamics of the Active Red Sea Trough: ACTIVE RED SEA TROUGH DYNAMICS, Journal of Geophysical Research: Atmospheres, 118, 7087–7108, https://doi.org/10.1002/jgrd.50569, 2013.

De Vries, A. J., Feldstein, S. B., Riemer, M., Tyrlis, E., Sprenger, M., Baumgart, M., Fnais, M., and Lelieveld, J.: Dynamics of Tropical–Extratropical Interactions and Extreme Precipitation Events in Saudi Arabia in Autumn, Winter and Spring, Quarterly Journal of the Royal Meteorological Society, 142, 1862–1880, https://doi.org/10.1002/qj.2781, 2016.

Diffenbaugh, N. S., Scherer, M., and Trapp, R. J.: Robust Increases in Severe Thunderstorm Environments in Response to Greenhouse Forcing, Proceedings of the National Academy of Sciences, 110, 16 361–16 366, https://doi.org/10.1073/pnas.1307758110, 2013.

Doswell, C. A. and Rasmussen, E. N.: The Effect of Neglecting the Virtual Temperature Correction on CAPE Calculations, Weather and Forecasting, 9, 625–629, https://doi.org/10.1175/1520-0434(1994)009<0625:TEONTV>2.0.CO;2, 1994.

Doswell III, C. A. and Schultz, D. M.: On the Use of Indices and Parameters in Forecasting Severe Storms, E-Journal of Severe Storms Meteorology, 1, 1–22, https://doi.org/10.55599/ejssm.v1i3.4, 2006.

Douville, H., Qasmi, S., Ribes, A., and Bock, O.: Global Warming at Near-Constant Tropospheric Relative Humidity Is Supported by Observations, Communications Earth & Environment, 3, 237, https://doi.org/10.1038/s43247-022-00561-z, 2022.

Eager, R. E., Raman, S., Wootten, A., Westphal, D. L., Reid, J. S., and Al Mandoos, A.: A Climatological Study of the Sea and Land Breezes in the Arabian Gulf Region, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009710, 2008.

El Kenawy, A. M. and McCabe, M. F.: Future Projections of Synoptic Weather Types over the Arabian Peninsula during the Twenty-First Century Using an Ensemble of CMIP5 Models, Theoretical and Applied Climatology, 130, 173–189, https://doi.org/10.1007/s00704-016-1874-y, 2017.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Farquharson, F. A. K., Plinston, D. T., and Sutcliffe, J. V.: Rainfall and Runoff in Yemen, Hydrological Sciences Journal, 41, 797–811, https://doi.org/10.1080/02626669609491546, 1996.

Fonseca, R., Francis, D., Nelli, N., Yarragunta, Y., Paparella, F., and Pauluis, O. M.: Summertime Secondary Convection and Interaction with Sea-Breeze Circulations, Quarterly Journal of the Royal Meteorological Society, 151, e4907, https://doi.org/10.1002/qj.4907, 2025.

Francis, D., Temimi, M., Fonseca, R., Nelli, N. R., Abida, R., Weston, M., and Whebe, Y.: On the Analysis of a Summertime Convective Event in a Hyperarid Environment, Quarterly Journal of the Royal Meteorological Society, 147, 501–525, https://doi.org/10.1002/qj.3930, 2021.

Francis, D., Fonseca, R., Nelli, N., Cherif, C., Yarragunta, Y., Zittis, G., and Jan de Vries, A.: From Cause to Consequence: Examining the Historic April 2024 Rainstorm in the United Arab Emirates through the Lens of Climate Change, npj Climate and Atmospheric Science, 8, 183, https://doi.org/10.1038/s41612-025-01073-1, 2025.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

GISTEMP Team: GISS Surface Temperature Analysis (GISTEMP), version 4, https://data.giss.nasa.gov/gistemp/, accessed: 2025-11-13, 2025.

Gómez-Navarro, J. J., Raible, C. C., García-Valero, J. A., Messmer, M., Montávez, J. P., and Martius, O.: Event Selection for Dynamical Downscaling: A Neural Network Approach for Physically-Constrained Precipitation Events, Climate Dynamics, 58, 2863–2879, https://doi.org/10.1007/s00382-019-04818-w, 2022.

Gopalakrishnan, D., Taraphdar, S., Pauluis, O. M., Xue, L., Ajayamohan, R. S., Al Shamsi, N., Chen, S., Lee, J. A., Grabowski, W. W., Liu, C., Tessendorf, S. A., and Rasmussen, R. M.: Anatomy of a Summertime Convective Event over the Arabian Region, Monthly Weather Review, 151, 989–1004, https://doi.org/10.1175/MWR-D-22-0082.1, 2023.

Gopalakrishnan, D., Cuervo-Lopez, C., Allen, J. T., Trapp, R. J., and Robinson, E.: A Comprehensive Evaluation of Biases in Convective Storm Parameters in CMIP6 Models over North America, Journal of Climate, 38, 947–971, https://doi.org/10.1175/JCLI-D-24-0165.1, 2025.

Govett, M., Bah, B., Bauer, P., Berod, D., Bouchet, V., Corti, S., Davis, C., Duan, Y., Graham, T., Honda, Y., Hines, A., Jean, M., Ishida, J., Lawrence, B., Li, J., Luterbacher, J., Muroi, C., Rowe, K., Schultz, M., Visbeck, M., and Williams, K.: Exascale Computing and Data Handling: Challenges and Opportunities for Weather and Climate Prediction, Bulletin of the American Meteorological Society, 105, E2385–E2404, https://doi.org/10.1175/BAMS-D-23-0220.1, 2024.

Hagos, S. M., Leung, L. R., Garuba, O. A., Demott, C., Harrop, B., Lu, J., and Ahn, M.-S.: The Relationship between Precipitation and Precipitable Water in CMIP6 Simulations and Implications for Tropical Climatology and Change, Journal of Climate, 34, 1587–1600, https://doi.org/10.1175/JCLI-D-20-0211.1, 2021.

Haklander, A. J. and Van Delden, A.: Thunderstorm Predictors and Their Forecast Skill for the Netherlands, Atmospheric Research, 67–68, 273–299, https://doi.org/10.1016/S0169-8095(03)00056-5, 2003.

Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W., and Zelinka, M.: Climate Simulations: Recognize the 'Hot Model' Problem, Nature, 605, 26–29, https://doi.org/10.1038/d41586-022-01192-2, 2022.

Heidke, P.: Berechnung Des Erfolges Und Der G{\"u}te Der Windst{\"a}rkevorhersagen Im Sturmwarnungsdienst, Geografiska Annaler, 8, 301–349, 1926.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hill, A. J. and Schumacher, R. S.: Forecasting Excessive Rainfall with Random Forests and a Deterministic Convection-Allowing Model, Weather and Forecasting, 36, 1693–1711, https://doi.org/10.1175/WAF-D-21-0026.1, 2021.

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., Bao, J., Bastin, S., Behravesh, M., Bergemann, M., Biercamp, J., Bockelmann, H., Brokopf, R., Brüggemann, N., Casaroli, L., Chegini, F., Datseris, G., Esch, M., George, G., Giorgetta, M., Gutjahr, O., Haak, H., Hanke, M., Ilyina, T., Jahns, T., Jungclaus, J., Kern, M., Klocke, D., Kluft, L., Kölling, T., Kornblueh, L., Kosukhin, S., Kroll, C., Lee, J., Mauritsen, T., Mehlmann, C., Mieslinger, T., Naumann, A. K., Paccini, L., Peinado, A., Praturi, D. S., Putrasahan, D., Rast, S., Riddick, T., Roeber, N., Schmidt, H., Schulzweida, U., Schütte, F., Segura, H., Shevchenko, R., Singh, V., Specht, M., Stephan, C. C., von Storch, J.-S., Vogel, R., Wengel, C., Winkler, M., Ziemen, F., Marotzke, J., and Stevens, B.: ICON-Sapphire: Simulating the

Components of the Earth System and Their Interactions at Kilometer and Subkilometer Scales, Geoscientific Model Development, 16, 779–811, https://doi.org/10.5194/gmd-16-779-2023, 2023.

Homoudi, A.: Dataset for "Convective Environments Over the Arabian Peninsula in Current and Future Climates: A Machine Learning Approach", https://doi.org/10.5281/zenodo.18514414, accessed: 2026-06-05, 2026.

Homoudi, A. and Barfus, K.: MeteoMate: Calculating Convective Indices for Climate Model Outputs, https://ahomoudi.gitlab.io/ahomoudi/MeteoMate/, r package version 1.0.0, 2025.

Homoudi, A., Barfus, K., Bernhofer, C., and Mauder, M.: Climatology of Convective Precipitation Systems Over the Arabian Peninsula Using Object-Tracking, International Journal of Climatology, 46, e70 158, https://doi.org/10.1002/joc.70158, 2026.

Hoteit, I., Abualnaja, Y., Afzal, S., Aman, C., Antony, C., Ashok, K., Asiri, M. I., Attada, R., Al-Azemi, M. A., Bavera, D., Al-Boqami, S., Castro, C., Chang, Hs.-I., Cornuelle, B., Cui, W., Dasari, H. P., Desamsetti, S., Dougherty, W., Ehsan, A., Feki, S., Gandham, H., Hadri, B., Hammoud, M. A. E.-R., Hassan, W. U., Islam, N., Ghulam, A., Gopinathan, P. A., Guo, D., Karumuri, R. K., Krokos, G., Kunchala, R. K., Luong, T., Masabathini, S., Mazloff, M., Pammi, R., Pathak, R., Risanto, C. B., Robertson, A., Lopez Rubio, E., Saharwardi, M. S., Sun, R., Subramanian, A., Sun, J., Titi, E. S., Vankayalapati, K., Viswanadhapalli, Y., Yates, D., Zampieri, M., Zamreeq, A., and Zhan, P.: New Climate Change Center of Saudi Arabia: Advancing Understanding and Prediction for the Arabian Peninsula Climate, Earth's Future, 13, e2025EF006 296, https://doi.org/10.1029/2025EF006296, 2025.

Huang, B., Yin, X., Menne, M. J., Vose, R. S., and Zhang, H.-M.: NOAA Global Surface Temperature Dataset (NOAAGlobalTemp), Version 6.0, https://doi.org/10.25921/rzxg-p717, accessed: 2025-11-27, 2024.

Huffman, G. J., Bolvin, D. T., Joyce, R., Nelkin, E. J., Tan, J., Braithwaite, D., Hsu, K., Kelley, O. A., Nguyen, P., Sorooshian, S., Watters, D. C., West, B. J., and Xie, P.: Algorithm Theoretical Basis Document (ATBD) for the NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) Version 07, Algorithm theoretical basis document, NASA Goddard Space Flight Center, Greenbelt, MD, https://gpm.nasa.gov/sites/default/files/2023-11/IMERG_ATBD_V07.pdf, 2023a.

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., and Tan, J.: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V07, https://doi.org/10.5067/GPM/IMERG/3B-HH/07, accessed: 25-06-2025, 2023b.

Hussein, K. A., Alsumaiti, T. S., Ghebreyesus, D. T., Sharif, H. O., and Abdalati, W.: High-Resolution Spatiotemporal Trend Analysis of Precipitation Using Satellite-Based Products over the United Arab Emirates, Water, 13, 2376, https://doi.org/10.3390/w13172376, 2021.

IPCC: Summary for Policymakers, in: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Team, C. W., Lee, H., and Romero, J., pp. 1–34, IPCC, Geneva, Switzerland, https://doi.org/10.59327/IPCC/AR6-9789291691647.001, in: Climate Change 2023: Synthesis Report. IPCC, Geneva, Switzerland, 2023.

Jiang, X., Menon, A., Wang, S., Kim, J., and Ohno-Machado, L.: Doubly Optimized Calibrated Support Vector Machine (DOC-SVM): An Algorithm for Joint Optimization of Discrimination and Calibration, PLOS ONE, 7, e48 823, https://doi.org/10.1371/journal.pone.0048823, 2012.

Jones, D. R., Schonlau, M., and Welch, W. J.: Efficient Global Optimization of Expensive Black-Box Functions, Journal of Global Optimization, 13, 455–492, https://doi.org/10.1023/A:1008306431147, 1998.

Kang, H. and Ebtehaj, A.: Machine Learning for Explanation of Subgrid Convective Precipitation: A Case Study over CONUS Using a Convection-Allowing Model and SHAP Analysis, Artificial Intelligence for the Earth Systems, 4, https://doi.org/10.1175/AIES-D-24-0062.1, 2025.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Khan, B., Abualnaja, Y., Al-Subhi, A. M., Nellayaputhenpeedika, M., Nellikkattu Thody, M., and Sturman, A. P.: Climatology of Sea Breezes along the Red Sea Coast of Saudi Arabia, International Journal of Climatology, 38, 3633–3650, https://doi.org/10.1002/joc.5523, 2018.

Kull, M., Silva Filho, T. M., and Flach, P.: Beyond Sigmoids: How to Obtain Well-Calibrated Probabilities from Binary Classifiers with Beta Calibration, Electronic Journal of Statistics, 11, https://doi.org/10.1214/17-EJS1338SI, 2017.

Kuo, Y.-H., Neelin, J. D., Chen, C.-C., Chen, W.-T., Donner, L. J., Gettelman, A., Jiang, X., Kuo, K.-T., Maloney, E., Mechoso, C. R., Ming, Y., Schiro, K. A., Seman, C. J., Wu, C.-M., and Zhao, M.: Convective Transition Statistics over Tropical Oceans for Climate Model Diagnostics: GCM Evaluation, Journal of the Atmospheric Sciences, 77, 379–403, https://doi.org/10.1175/JAS-D-19-0132.1, 2020.

Lenssen, N., Schmidt, G. A., Hendrickson, M., Jacobs, P., Menne, M. J., and Ruedy, R.: A NASA GISTEMPv4 Observational Uncertainty Ensemble, Journal of Geophysical Research: Atmospheres, 129, e2023JD040 179, https://doi.org/10.1029/2023JD040179, 2024.

Lepore, C., Abernathey, R., Henderson, N., Allen, J. T., and Tippett, M. K.: Future Global Convective Environments in CMIP6 Models, Earth's Future, 9, e2021EF002 277, https://doi.org/10.1029/2021EF002277, 2021.

Luong, T. M., Dasari, H. P., and Hoteit, I.: Extreme Precipitation Events Are Becoming Less Frequent but More Intense over Jeddah, Saudi Arabia. Are Shifting Weather Regimes the Cause?, Atmospheric Science Letters, 21, e981, https://doi.org/10.1002/asl.981, 2020.

Luong, T. M., Dasari, H. P., Doan, Q.-V., Alduwais, A. K., and Hoteit, I.: Organized Precipitation and Associated Large-Scale Circulation Patterns over the Kingdom of Saudi Arabia, International Journal of Climatology, 44, 3295–3314, https://doi.org/10.1002/joc.8524, 2024.

Mahmoud, S. H., Gan, T. Y., and Zhu, D. Z.: Impacts of Climate Change and Climate Variability on Water Resources and Drought in an Arid Region and Possible Resiliency and Adaptation Measures against Climate Warming, Climate Dynamics, 61, 4079–4105, https://doi.org/10.1007/s00382-023-06795-7, 2023.

Malik, A., Stenchikov, G., Mostamandi, S., Parajuli, S., Lelieveld, J., Zittis, G., Ahsan, M. S., Atique, L., and Usman, M.: Accelerated Historical and Future Warming in the Middle East and North Africa, Journal of Geophysical Research: Atmospheres, 129, e2024JD041 625, https://doi.org/10.1029/2024JD041625, 2024.

Martinez-Villalobos, C., Neelin, J. D., and Pendergrass, A. G.: Metrics for Evaluating CMIP6 Representation of Daily Precipitation Probability Distributions, Journal of Climate, 35, 5719–5743, https://doi.org/10.1175/JCLI-D-21-0617.1, 2022.

Meredith, E. P., Rust, H. W., and Ulbrich, U.: A Classification Algorithm for Selective Dynamical Downscaling of Precipitation Extremes, Hydrology and Earth System Sciences, 22, 4183–4200, https://doi.org/10.5194/hess-22-4183-2018, 2018.

Meredith, E. P., Ulbrich, U., and Rust, H. W.: The Diurnal Nature of Future Extreme Precipitation Intensification, Geophysical Research Letters, 46, 7680–7689, https://doi.org/10.1029/2019GL082385, 2019.

Miller, G. A., Stier, P., and Christensen, H. M.: Characterizing Uncertainty in Deep Convection Triggering Using Explainable Machine Learning, Journal of the Atmospheric Sciences, 82, 1093–1111, https://doi.org/10.1175/JAS-D-24-0085.1, 2025.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, Journal of Geophysical Research: Atmospheres, 126, e2019JD032 361, https://doi.org/10.1029/2019JD032361, 2021.

Murphy, A. H.: A New Vector Partition of the Probability Score, Journal of Applied Meteorology and Climatology, 12, 595–600, https://doi.org/10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2, 1973.

Naeini, M. P., Cooper, G., and Hauskrecht, M.: Obtaining Well Calibrated Probabilities Using Bayesian Binning, Proceedings of the AAAI Conference on Artificial Intelligence, 29, https://doi.org/10.1609/aaai.v29i1.9602, 2015.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Nelli, N., Francis, D., Fonseca, R., Abida, R., Weston, M., Wehbe, Y., and Al Hosary, T.: The Atmospheric Controls of Extreme Convective Events over the Southern Arabian Peninsula during the Spring Season, Atmospheric Research, 262, 105 788, https://doi.org/10.1016/j.atmosres.2021.105788, 2021.

Nelli, N., Francis, D., Fonseca, R., Bosc, E., Addad, Y., Temimi, M., Abida, R., Weston, M., and Cherif, C.: Characterization of the Atmospheric Circulation near the Empty Quarter Desert during Major Weather Events, Frontiers in Environmental Science, 10, https://doi.org/10.3389/fenvs.2022.972380, 2022.

Nelli, N., Francis, D., Fonseca, R., Delle Monache, L., and Al Mandous, A.: Drivers and Trends of Summertime Convection Over the Southeastern Arabian Peninsula, Geophysical Research Letters, 52, e2025GL118 960, https://doi.org/10.1029/2025GL118960, 2025.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

Pathak, R., Dasari, H. P., Ashok, K., and Hoteit, I.: Dynamics of Intensification of Extreme Precipitation Events over the Arabian Peninsula Derived from CMIP6 Simulations, npj Climate and Atmospheric Science, 8, 126, https://doi.org/10.1038/s41612-025-01016-w, 2025.

Pozzolo, A. D., Caelen, O., Johnson, R. A., and Bontempi, G.: Calibrating Probability with Undersampling for Unbalanced Classification, in: 2015 IEEE Symposium Series on Computational Intelligence, pp. 159–166, IEEE, Cape Town, South Africa, ISBN 978-1-4799-7560-0, https://doi.org/10.1109/SSCI.2015.33, 2015.

Púčik, T., Groenemeijer, P., Rädler, A. T., Tijssen, L., Nikulin, G., Prein, A. F., Van Meijgaard, E., Fealy, R., Jacob, D., and Teichmann, C.: Future Changes in European Severe Convection Environments in a Regional Climate Model Ensemble, Journal of Climate, 30, 6771–6794, https://doi.org/10.1175/JCLI-D-16-0777.1, 2017.

Rädler, A. T., Groenemeijer, P. H., Faust, E., Sausen, R., and Púčik, T.: Frequency of Severe Thunderstorms across Europe Expected to Increase in the 21st Century Due to Rising Instability, npj Climate and Atmospheric Science, 2, 30, https://doi.org/10.1038/s41612-019-0083-7, 2019.

Riemann-Campe, K., Fraedrich, K., and Lunkeit, F.: Global Climatology of Convective Available Potential Energy (CAPE) and Convective Inhibition (CIN) in ERA-40 Reanalysis, Atmospheric Research, 93, 534–545, https://doi.org/10.1016/j.atmosres.2008.09.037, 2009.

Rohde, R. A. and Hausfather, Z.: The Berkeley Earth Land/Ocean Temperature Record, Earth System Science Data, 12, 3469–3479, https://doi.org/10.5194/essd-12-3469-2020, 2020.

Rust, H. W., Kruschke, T., Dobler, A., Fischer, M., and Ulbrich, U.: Discontinuous Daily Temperatures in the WATCH Forcing Datasets, Journal of Hydrometeorology, 16, 465–472, https://doi.org/10.1175/JHM-D-14-0123.1, 2015.

Saharwardi, M. S., Hassan, W. U., Dasari, H. P., Abualnaja, Y., and Hoteit, I.: Enhanced Flash Droughts in Recent Decades over the Arabian Peninsula, Journal of Hydrology: Regional Studies, 61, 102 696, https://doi.org/10.1016/j.ejrh.2025.102696, 2025.

Sandeep, S. and Ajayamohan, R. S.: Poleward Shift in Indian Summer Monsoon Low Level Jetstream under Global Warming, Climate Dynamics, 45, 337–351, https://doi.org/10.1007/s00382-014-2261-y, 2015.

Seeley, J. T. and Romps, D. M.: The Effect of Global Warming on Severe Thunderstorms in the United States, Journal of Climate, 28, 2443–2458, https://doi.org/10.1175/JCLI-D-14-00382.1, 2015.

Seneviratne, S. I., Donat, M. G., Pitman, A. J., Knutti, R., and Wilby, R. L.: Allowable CO2 Emissions Based on Regional and Impact-Related Climate Targets, Nature, 529, 477–483, https://doi.org/10.1038/nature16542, 2016.

Sharma, S., Raghuvanshi, A. S., and Agarwal, A.: Evaluating the Performance of Uni- and Multivariate Bias Correction Techniques: Challenges in Preserving Temporal and Dependence Structures, Water Resources Research, 62, e2025WR041 526, https://doi.org/10.1029/2025WR041526, 2026.

Sherwood, S. C., Ingram, W., Tsushima, Y., Satoh, M., Roberts, M., Vidale, P. L., and O'Gorman, P. A.: Relative Humidity Changes in a Warmer Climate, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2009JD012585, 2010.

Sun, Y., Wong, A. K. C., and Kamel, M. S.: Classification of Imbalanced Data: A Review, International Journal of Pattern Recognition and Artificial Intelligence, 23, 687–719, https://doi.org/10.1142/S0218001409007326, 2009.

Tang, S., Gleckler, P., Xie, S., Lee, J., Ahn, M.-S., Covey, C., and Zhang, C.: Evaluating the Diurnal and Semidiurnal Cycle of Precipitation in CMIP6 Models Using Satellite- and Ground-Based Observations, Journal of Climate, 34, 3189–3210, https://doi.org/10.1175/JCLI-D-20-0639.1, 2021.

Taraphdar, S., Gopalakrishnan, D., Liu, C., Pauluis, O. M., Xue, L., Ajayamohan, R. S., Leung, L. R., Hagos, S., Grabowski, W. W., Chen, S., Rasmussen, R. M., and Tessendorf, S. A.: Subtropical Jet Regulates Arabian Winter Precipitation: A Viable Mechanism, Journal of the Atmospheric Sciences, 82, 713–732, https://doi.org/10.1175/JAS-D-23-0213.1, 2025.

Taszarek, M., Allen, J. T., Marchio, M., and Brooks, H. E.: Global Climatology and Trends in Convective Environments from ERA5 and Rawinsonde Data, npj Climate and Atmospheric Science, 4, 1–11, https://doi.org/10.1038/s41612-021-00190-x, 2021.

Terry, J. P., Al Ruheili, A., Almarzooqi, M. A., Almheiri, R. Y., and Alshehhi, A. K.: The Rain Deluge and Flash Floods of Summer 2022 in the United Arab Emirates: Causes, Analysis and Perspectives on Flood-Risk Reduction, Journal of Arid Environments, 215, 105 013, https://doi.org/10.1016/j.jaridenv.2023.105013, 2023.

Trapp, R. J., Diffenbaugh, N. S., Brooks, H. E., Baldwin, M. E., Robinson, E. D., and Pal, J. S.: Changes in Severe Thunderstorm Environment Frequency during the 21st Century Caused by Anthropogenically Enhanced Global Radiative Forcing, Proceedings of the National Academy of Sciences, 104, 19 719–19 723, https://doi.org/10.1073/pnas.0705494104, 2007.

Trenberth, K.: Changes in Precipitation with Climate Change, Climate Research, 47, 123–138, https://doi.org/10.3354/cr00953, 2011.

Tuel, A., Choi, Y.-W., AlRukaibi, D., and Eltahir, E. A. B.: Extreme Storms in Southwest Asia (Northern Arabian Peninsula) under Current and Future Climates, Climate Dynamics, 58, 1509–1524, https://doi.org/10.1007/s00382-021-05975-7, 2022.

Utsumi, N., Seto, S., Kanae, S., Maeda, E. E., and Oki, T.: Does Higher Surface Temperature Intensify Extreme Precipitation?, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL048426, 2011.

Vahid Yousefnia, K., Bölle, T., Zöbisch, I., and Gerz, T.: A Machine-Learning Approach to Thunderstorm Forecasting through Post-Processing of Simulation Data, Quarterly Journal of the Royal Meteorological Society, 150, 3495–3510, https://doi.org/10.1002/qj.4777, 2024.

Wang, S., Trapp, R. J., Allen, J., Gopalakrishnan, D., and Robinson, E.: Future Climate Projections of Hazardous Convective Weather Using an Ensemble of Environment-Informed, Convection-Permitting Dynamical Downscaling Simulations, Journal of Geophysical Research: Atmospheres, 131, e2025JD045 354, https://doi.org/10.1029/2025JD045354, 2026.

Wang, Y., You, Y., and Kulie, M.: Global Virga Precipitation Distribution Derived From Three Spaceborne Radars and Its Contribution to the False Radiometer Precipitation Detection, Geophysical Research Letters, 45, 4446–4455, https://doi.org/10.1029/2018GL077891, 2018.

Wilks, D. S.: Chapter 9 - Forecast Verification, in: Statistical Methods in the Atmospheric Sciences (Fourth Edition), pp. 369–483, Elsevier, fourth edition edn., ISBN 978-0-12-815823-4, https://doi.org/https://doi.org/10.1016/B978-0-12-815823-4.00009-2, 2019.

Williams, I. N. and Fieweger, G. N.: Mechanisms of Projected Changes in Thunderstorm Downburst Environments Across the United States, Geophysical Research Letters, 53, e2025GL117 925, https://doi.org/10.1029/2025GL117925, 2026.

https://doi.org/10.5194/egusphere-2026-3288
Preprint. Discussion started: 3 August 2026

Wright, M. N. and Ziegler, A.: Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R, Journal of Statistical Software, 77, 1–17, https://doi.org/10.18637/jss.v077.i01, 2017.

Zhang, B., Song, S., Wang, H., Guo, T., and Ding, Y.: Evaluation of the Performance of CMIP6 Models in Simulating Extreme Precipitation and Its Projected Changes in Global Climate Regions, Natural Hazards, 121, 1737–1763, https://doi.org/10.1007/s11069-024-06850-4, 2025a.

Zhang, F., Li, M., Zhang, Q., Liu, M., Xu, J., Guo, J., and Fan, H.: Differences in Triggering Conditions between Embedded and Isolated Convection Initiation during a Meiyu Season in China, Quarterly Journal of the Royal Meteorological Society, p. qj.4927, https://doi.org/10.1002/qj.4927, 2025b.