atmosphere

MDPI

Article

Development of Region-Specific Rainfall Design Storms Using Machine Learning in Southwestern Saudi Arabia

Raied Alharbi

Department of Civil Engineering, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia; rsalharbi@ksu.edu.sa

Abstract

The mountainous southwest of Saudi Arabia exhibits complex, highly seasonal precipitation driven by Indian Ocean monsoon inflows and orographic lifting. To characterize storm hyetographs, cluster analysis was applied to 8972 rainfall events recorded at 151 gauges. Two primary clusters emerged: one with early, intense peaks and another with later peak intensities, broadly reflecting windward versus leeward storm behavior. A locally derived hyetograph profile (AI) was constructed from the cluster centroids and benchmarked against standard design-storm distributions (Uniform, SCS Type II, Huff quartiles). Across fit metrics—cumulative RMSE, Kolmogorov–Smirnov distance, and cosine-intensity similarity—the AI distribution provided the best match for ~46% of storms, markedly outperforming canonical profiles (Uniform and SCS Type II each best-fit only ~11–12%). These results indicate that region-specific rainfall distributions more accurately represent precipitation patterns than conventional profiles, and that tailored hyetographs can improve hydrologic modeling and water-resources assessments in this climatically heterogeneous region.

Keywords: precipitation hyetograph; rainfall distribution; cluster analysis

1. Introduction

In recent decades, greater climate variation and fast land use change have been linked to stronger extreme rainfall around the world. In the southwestern region of Saudi Arabia, which has complex terrain, semi-arid climate and fast-growing cities, high risk from hydrologic extremes has been reported, including local flash floods and short-duration high-intensity convective storms. These events are seen to challenge traditional hydrologic design methods, which are usually based on general and idealized storm profiles such as the Chicago storm model and SCS Type II distributions. These common profiles are based on stationarity and on standard intensity duration frequency (IDF) relationships, and they do not capture the real spatial and temporal changes in storm behavior [1–3].

In this region, the local nature of storm events has been observed, often with steep rainfall changes across short distances and with several peaks and irregular timing. Because of this, traditional design methods have been found to be less effective. As a result, stormwater systems have not been well matched to current weather and climate conditions, and both resilience and public safety have been reduced. A need has been recognized for data-driven and adaptive storm profiling methods that reflect regional climate and hydrologic conditions, especially as flood risk continues to rise across southwestern Saudi Arabia [4,5].

Check for updates

Academic Editor: Tin Lukić

Received: 12 October 2025
Revised: 13 April 2026
Accepted: 17 April 2026
Published: 27 April 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Existing efforts to improve storm characterization have largely depended on deterministic or statistical techniques. These include frequency analysis, cluster-based classification (e.g., K-means, hierarchical clustering), and empirical parameterization of design storms using rainfall records. While analytically tractable, these approaches are limited in their ability to detect nonlinear interactions, accommodate high-dimensional inputs, or adapt to non-stationary climatic signals [6–8]. Moreover, these methods assume homogeneity across storm structures and often overlook rare, yet hydrologically significant, storm events. Such limitations are particularly evident in regions like Saudi Arabia, where meteorological conditions are shaped by mesoscale convective systems and orographic influences that produce highly variable rainfall intensities across short distances and durations [9–11].

Conventional, idealized design storms—such as the Chicago hyetograph, the alternating-block method, and the SCS Type II distribution—have long underpinned engineering analyses of peak runoff and infrastructure stress conditions [1,2]. However, intensifying extremes and non-stationary behavior have been widely reported, challenging time-invariant return-period assumptions and the sufficiency of fixed storm templates [3]. In Saudi Arabia, recent climatology also documents marked variability in seasonal rainfall and temperature over the last four decades [4]. These pressures are especially consequential in the mountainous, semi-arid southwest (e.g., Asir–Al-Baha–Jazan), where complex orography and rapid urbanization heighten exposure to short-duration, high-intensity convective storms and localized flash floods [5–7].

In this region, steep rainfall gradients over short distances, multi-peaked intensity structures, and temporally asymmetric accumulation are frequently observed, rendering traditional design hyetographs increasingly misaligned with contemporary hydrometeorological realities. The result is a risk of under- or over-design in stormwater systems, with implications for public safety and asset resilience. A shift toward regional, data-driven storm profiling that reflects local climatic controls, and topographic forcing is therefore warranted.

Past efforts to refine storm characterization have relied on deterministic and statistical techniques—regional frequency analysis, cluster-based classification (K-means, hierarchical), principal component analysis, and empirical parameterization from historical records [2,6,8]. While analytically tractable, these approaches impose rigid structure, struggle with high-dimensional inputs, and typically assume stationarity, limiting their ability to detect nonlinear interactions and rare—but hydrologically significant—events. Such limitations are particularly evident for orographically forced convective systems that dominate the rainfall climatology of southwestern Saudi Arabia [10,11].

In response to these challenges, Artificial Intelligence (AI) has emerged as a powerful tool in the field of hydrological modeling, offering advanced capabilities in pattern recognition, anomaly detection, and high-dimensional classification. AI-based models, particularly those derived from machine learning (ML) and deep learning (DL) paradigms, have demonstrated strong performance across a range of applications including flood forecasting [12], rainfall–runoff modeling [13], spatiotemporal hazard analysis [14], and radar–rainfall estimation or disaggregation [15]. These models, by learning directly from observed data, offer the ability to uncover complex, nonlinear patterns within storm structures that would typically elude deterministic or statistically based methods.

Recent advancements in unsupervised and semi-supervised learning have further enabled the classification of complex rainfall events without the need for predefined templates or training labels. Convolutional neural networks (CNNs), recurrent networks such as long short-term memory (LSTM), and attention-based architectures have been successfully employed to classify storm cells, detect flood-prone zones, and capture temporal dependencies in hydrological time series [16–18]. Within the Arabian Peninsula,

including Saudi Arabia, predictive frameworks leveraging AI have already been explored for drought monitoring [19], rainfall-pattern analysis [20], and urban flood susceptibility mapping [21], reflecting growing regional adoption of AI in hydroclimatic analysis. Furthermore, several recent studies have emphasized the effectiveness of advanced and hybrid machine learning approaches for rainfall prediction, environmental modeling, and intelligent water resource management, demonstrating their robustness across diverse hydroclimatic conditions [22–26].

Despite these advances, the application of AI specifically to the development of local storm profiles, that is, the generation of design-relevant, empirically based rainfall distributions derived from AI-classified storm events, remains largely unaddressed. To the best of current knowledge, no documented effort has yet been made to implement unsupervised AI models for storm-profile development in southwestern Saudi Arabia. The current study addresses this research gap by proposing a framework that utilizes unsupervised and semi-supervised machine learning methods to classify storms and produce physically coherent, localized storm hyetographs. To the best of current knowledge, no previous study has applied AI-based clustering for the development of region-specific rainfall design storm profiles in southwestern Saudi Arabia or comparable arid mountainous environments, thereby underscoring the originality of the present work.

This study builds upon the premise that AI techniques can be employed not merely as forecasting tools, but as analytical instruments capable of redefining storm classification and profile generation processes. Diverging schools of thought persist in the literature: one supports hybrid pipelines in which AI augments or calibrates physics-based hydrological models [13,16], while another emphasizes black-box, data-driven approaches focused on maximizing predictive skill [12,14]. An integrative stance is adopted herein, wherein AI is applied to structure-agnostic clustering and followed by physically interpretable evaluation using post hoc techniques such as SHAP (SHapley Additive exPlanations) and feature importance metrics.

The study focuses on rainfall data from 2017 to 2024 in the southwestern region of Saudi Arabia—an area subject to complex orographic forcing, high convective activity, and hydrological infrastructure sensitivity. The novelty of the approach lies in coupling unsupervised AI-based classification with high-resolution time series data to derive dimensionless storm profiles that capture storm structure features poorly represented by conventional templates—such as early-peaking intensity curves, multi-pulse accumulations, and temporally asymmetric distributions. These profiles are rooted in empirical evidence and region-specific hydrometeorological conditions.

The study is aimed at establishing and evaluating an AI-enhanced pipeline for classifying and profiling local storm events in southwestern Saudi Arabia that are not well represented by standard design storms.

By grounding design profiles in observed rainfall records (2017–2024) and applying AI models to expose latent structural features, more realistic storm hyetographs tailored to southwestern Saudi Arabia are produced. The generated profiles better replicate hydrologically critical features—such as short-duration peaks and temporally variable accumulations—thereby supporting improved flood modeling fidelity and offering enhanced design inputs for climate-resilient stormwater infrastructure [3–5,7,10–15,19–21].

2. Materials and Methods

The study covers the Red Sea coastal plain and Sarawat (Hejaz–Asir) mountain ranges of southwestern Saudi Arabia (Asir, Al-Baha, Makkah, Jazan). This region lies between the Red Sea to the west and the Arabian Desert to the east. Topography rises steeply from near sea level on the Tihamah coast (a hot, humid strip ~25–50 km wide) to the

Sarawat peaks (Jabal Sawda ≈ 3000 m). The geology is Arabian Shield bedrock deeply incised by wadis; farther east the terrain levels onto a broad plateau (≈1600–2750 m) descending gently toward the Empty Quarter. Climatic zones vary with elevation: the lowlands are hot-arid (Köppen BWh/BSh), whereas the higher mountains have a cooler, monsoon-influenced subtropical highland climate (Cwb). Rainfall in the study area is strongly seasonal, with peaks occurring during the spring–summer monsoon period, and exhibits pronounced spatial heterogeneity. Higher rainfall intensities are observed along the windward escarpments, where annual totals may exceed 500–600 mm, while significantly lower amounts are recorded in the leeward and rain-shadow regions. In contrast, the coastal Tihamah and inland plains receive substantially reduced rainfall, averaging approximately 100–120 mm per year. A network of 151 MEWA rain gauges (blue dots in Figure 1) monitors this variability, providing critical data for hydrologic modeling and water resource management in the region.

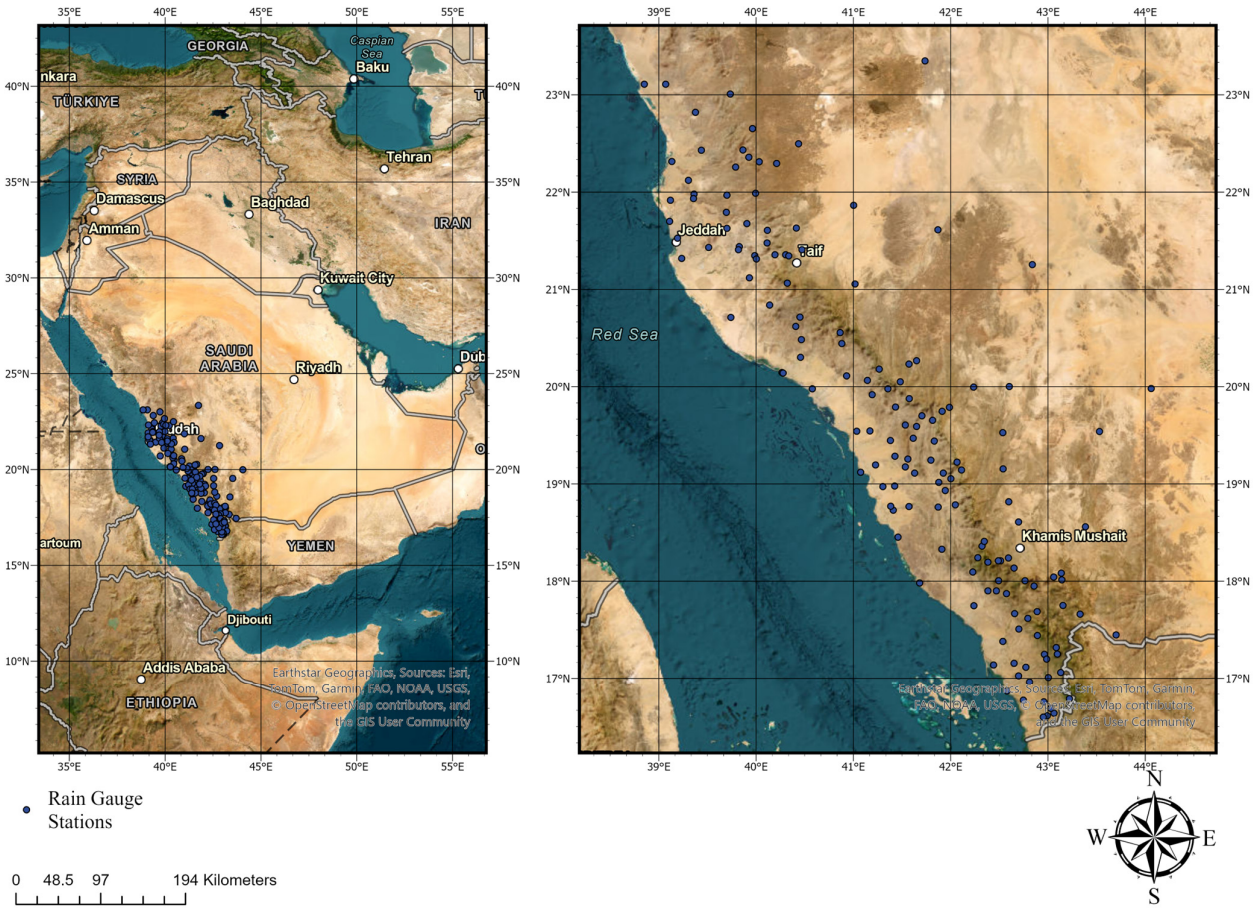

Figure 1. Study Area Location and MEWA Rain Gauge Network in Southwestern Saudi Arabia.

3. Methodology

In this study, a two-part methodology was developed and applied to analyze high-resolution rainfall events in southwestern Saudi Arabia. In the first part, an unsupervised clustering approach based on k-means within a statistical analysis framework is employed to classify observed storm events into distinct temporal patterns. It is noted that k-means is implemented here as a statistical clustering technique, without modification to its underlying mathematical formulation. It is noted that the objective of this study is not to modify the k-means algorithm itself, but to employ it as a clustering tool within a data-driven framework for deriving representative storm profiles. In the second part, the performance

of the resulting AI-derived profiles is systematically evaluated against conventional design storms to identify the most representative synthetic hyetographs for the region.

3.1. Part One: Storm Event Cumulative Profile Clustering

Rainfall events recorded at 5 min resolution from the Ministry of Environment, Water and Agriculture (MEWA) network in southwest Saudi Arabia (2017–2024) were analyzed as shown in Figure 1. The data were pre-segmented into qualified storm events. These events were defined as continuous rainfall sequences with non-zero incremental depths, separated by periods of no rainfall. Each event is represented by a time series of nonnegative incremental depths $\Delta p_k \geq 0$ at ascending timestamps t_k. For each storm event, the total duration T and total depth P_{tot} are defined by Equation (1):

$$T = t_{\text{end}} - t_0 > 0, P_{\text{tot}} = \sum_k \Delta p_k > 0. \quad (1)$$

where t_0 and t_{end} denote the start and end times of the event, respectively. Each event was then transformed into a dimensionless time profile and its cumulative depth fraction. This normalization ensures comparability across events of different durations and magnitudes by removing scale effects and preserving only temporal structure. Specifically, $\tau = (t - t_0)/T \in [0, 1]$ was introduced and the cumulative fraction of rainfall was defined as

$$F(\tau) = \frac{P(t)}{P_\text{tot}} \quad (2)$$

where $P(t) = \sum_{t_k \leq t} \Delta p_k$ is the accumulated depth up to time t. By construction, $F(0) = 0$ and $F(1) = 1$. To enable direct comparison of storm shapes, each empirical cumulative curve $F(\tau)$ was resampled onto a uniform grid of $M = 101$ points in $[0, 1]$ (Equation (3)) using linear interpolation (Equation (4)). Concretely, $F(\tau)$ was evaluated at $\tau_i = i/(M-1)$ for $i = 0, 1, \ldots, M-1$; if $\tau_k \leq \tau_i \leq \tau_{k+1}$ for some index k, then

$$\tau_i = \frac{i}{M-1}, i = 0, 1, \ldots, M-1 (M = 101), \quad (3)$$

using linear interpolation: if $\tau_k \leq \tau_i \leq \tau_{k+1}$,

$$F(\tau_i) = (1-\lambda)F(\tau_k) + \lambda F(\tau_{k+1}), \lambda = \frac{\tau_i - \tau_k}{\tau_{k+1} - \tau_k}. \quad (4)$$

The result of this normalization is that each storm event is represented by a dimensionless cumulative rainfall vector:

$$\mathrm{x} = (F(\tau_0), F(\tau_1), \ldots, F(\tau_{M-1}))^\top \in \left[0, 1\right]^M. \quad (5)$$

Before clustering, the presence of meaningful cluster structure was assessed using the Hopkins statistic (Equation (6)). This statistic contrasts nearest-neighbor distances for uniformly sampled pseudo-points within the data's bounding hyper-rectangle with nearest-neighbor distances for randomly sampled data points. Let u_i denote the distance from the i-th uniformly drawn pseudo-point to its nearest neighbor in the observed dataset, and let w_i denote the distance from the i-th randomly selected point to its nearest neighbor among the remaining observed points (self-matches excluded). For a sample size of m pairs, the Hopkins statistic is defined as

$$H = \frac{\sum_{i=1}^m u_i}{\sum_{i=1}^m u_i + \sum_{i=1}^m w_i}. \quad (6)$$

Values $H \approx 0.5$ indicate spatial randomness, $H > 0.5$ indicate a tendency toward clustered structure, and $H < 0.5$ indicate a tendency toward regular (over-dispersed) structure.

A k-means clustering procedure was applied to partition storms on the basis of their dimensionless cumulative profiles $x \in [0,1]^M$ (constructed as samples of $F(\tau)$ on the uniform grid $\{\tau_i\}_{i=0}^{M-1}$). Similarity between two storm profiles x and y was quantified using the cosine dissimilarity (Equation (7)):

$$d_{\cos}(x,y) = 1 - \frac{x^\top y}{\| x \|_2 \| y \|_2}, \tag{7}$$

which, for non-negative cumulative vectors, ranges from 0 (identical shape) to 1 (maximally dissimilar in direction). This measure emphasizes differences in the temporal shape of rainfall accumulation and is insensitive to total depth (removed by normalization).

For a fixed number of clusters k, storms x_i were assigned to clusters $C_j (j = 1, \ldots, k)$ and cluster prototypes μ_j were estimated by minimizing the total within-cluster distortion, defined as the sum of cosine dissimilarities from each profile to its cluster prototype (Equation (8)):

$$J_k = \sum_{j=1}^{k} \sum_{x_i \in C_j} d_{\cos}(x_i, \mu_j). \tag{8}$$

Multiple random centroid initializations were employed to reduce sensitivity to local minimum. This formulation yields groups of storms with similar cumulative rainfall profiles and corresponding representative prototypes.

To determine an appropriate number of clusters, solutions with $k = 2, 3, \ldots, 10$ were evaluated using three complementary criteria.

(a) Elbow (distortion) method. For each k, the total within-cluster distortion was computed as

$$D_k = \sum_{j=1}^{k} \sum_{x_i \in C_j} d_{\cos}(x_i, \mu_j), \tag{9}$$

where $d_{\cos}$ denotes the cosine dissimilarity and μ_j is the prototype of cluster C_j. Plotting D_k versus k (the "elbow curve") reveals the diminishing returns in distortion reduction as clusters are added; a noticeable inflection was inspected to suggest a parsimonious k (see the elbow–silhouette–gap panel in the Results).

(b) Mean silhouette coefficient. Clustering quality was assessed using the silhouette statistic with $\delta(x, x') = d_{\cos}(x, x')$ as dissimilarity. For storm i assigned to cluster $c(i)$, the intra-cluster dissimilarity a_i and the nearest-neighbor cluster dissimilarity b_i were defined by

$$a_i = \frac{1}{| C_{c(i)} | - 1} \sum_{\substack{x_j \in C_{c(i)} \\ j \neq i}} \delta(x_i, x_j), b_i = \min_{\updownarrow \neq c(i)} \frac{1}{| C_{\updownarrow} |} \sum_{x_j \in C_{\updownarrow}} \delta(x_i, x_j), \tag{10}$$

and the silhouette value by

$$s_i = \frac{b_i - a_i}{\max\{a_i, b_i\}}, \bar{s}(k) = \frac{1}{N} \sum_{i=1}^{N} s_i. \tag{11}$$

The value of k maximizing $\bar{s}(k)$ was recorded (see the elbow–silhouette–gap panel in the Results).

(c) Gap statistic (Tibshirani's rule). To compare observed clustering dispersion against a null reference, the within-cluster Euclidean dispersion was computed as

$$W_k = \sum_{j=1}^{k} \sum_{x_i \in C_j} \| x_i - \overline{x}_j \|_2^2, \tag{12}$$

where $\overline{x}_j$ is the mean of cluster C_j. For B Monte Carlo reference sets, monotone "no-structure" cumulative curves were generated by sampling IID non-negative increments $r_q \sim \mathrm{Exp}(1)$ and forming

$$R(\tau_i) = \frac{\sum_{q=0}^{i} r_q}{\sum_{q=0}^{M-1} r_q}, i = 0, \ldots, M-1,$$

after which $W_k^*(b)$ was computed as in (12). The value gap was then

$$\mathrm{Gap}(k) = \frac{1}{B} \sum_{b=1}^{B} \ln W_k^*(b) - \ln W_k. \tag{13}$$

The selection rule chose the smallest k such that $\mathrm{Gap}(k) \geq \mathrm{Gap}(k+1) - s_{k+1}$, where s_{k+1} denotes the standard-error adjustment for $\ln W_{k+1}^*$.

Decision rule. Each criterion yielded a recommended cluster count (e.g., an elbow at a given k; a maximum $\overline{s}(k)$; a gap-statistic choice). The final number of clusters K^* was taken as the majority among the three; in the absence of a majority, the smaller k was selected for parsimony. This procedure balances complementary notions of compactness, separation, and deviation from null structure.

3.2. Cluster Representative Profiles and Timing Metrics

Once K^* had been fixed, each cluster was summarized by a representative centroid cumulative profile and by timing diagnostics derived from it. For cluster C_j, the centroid cumulative curve $\widetilde{F}_j(\tau)$ was defined pointwise as the median of member curves:

$$\widetilde{F}_j(\tau_i) = \mathrm{median}\left\{ F^{(n)}(\tau_i) : x_n \in C_j \right\}, i = 0, \ldots, M-1. \tag{14}$$

Quartile timing metrics were obtained by inverting $\widetilde{F}_j$:

$$T_p^{(j)} = \inf\left\{ \tau \in [0,1] : \widetilde{F}_j(\tau) \geq p \right\}, p \in \{0.25, 0.50, 0.75\}. \tag{15}$$

Finally, the discrete incremental accumulation

$$s_j(\tau_i) = \widetilde{F}_j(\tau_{i+1}) - \widetilde{F}_j(\tau_i)$$

was evaluated, and the peak-intensity location was identified as

$$\tau_{\mathrm{peak}}^{(j)} = \arg\max_i s_j(\tau_i). \tag{16}$$

These centroid curves and timing metrics provide concise descriptors of each cluster's dimensionless hyetograph and are referenced in the Results (e.g., centroid overlays and associated quartile/peak locations).

3.3. Part Two: Synthetic Profile Evaluation and Selection

In Part Two, each observed storm's temporal rainfall distribution was compared against a suite of candidate synthetic hyetographs to identify the best-fitting profile on a storm-by-storm basis. Two profile families were included:

1. Conventional design profiles. Eight widely used patterns were considered to span front-loaded, near-uniform, and back-loaded shapes: Huff quartiles (Q1–Q4), SCS Type II, Alternating Block Method (ABM), Uniform, and Euler/Elder II.
2. Cluster-centroid profiles. One representative profile per cluster from Part One was included, defined as the cluster centroid cumulative curve $\widetilde{F}_j(\tau)$ (cf. Equations (14)–(16) in Part One).

All profiles and all observed events were expressed in the dimensionless cumulative format, i.e., fraction of storm depth versus fraction of storm duration, $F(\tau) \in [0,1]$ for $\tau \in [0,1]$ (see Equations (2)–(5) in Part One). Conventional profiles that were originally defined as intensity patterns were integrated in τ and then normalized so that $F_{\text{prof}}(0) = 0$ and $F_{\text{prof}}(1) = 1$. This representation enabled direct, scale-free comparisons of timing.

3.4. Event–Profile Comparison Set-Up

Let $F_{\text{obs}}(\tau)$ denote the observed event's cumulative curve and $F_{\text{prof}}(\tau)$ the candidate profile's cumulative curve, both sampled on the common uniform grid $\left\{\tau_i\right\}_{i=0}^{M-1}$ with $\tau_0 = 0$, $\tau_{M-1} = 1$. Goodness-of-fit was quantified for every observed–profile pair on this grid, and the resulting scores were used to select a best-fit profile for each storm.

3.5. Goodness-of-Fit Metrics

Three complementary metrics were employed; equation numbering continues from Part One.

(i) Root-Mean-Square Error (RMSE) on cumulative curves.

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{i=0}^{M-1}\left(F_{\text{obs}}(\tau_i) - F_{\text{prof}}(\tau_i)\right)^2}. \tag{17}$$

(ii) Kolmogorov–Smirnov (KS) distance on cumulative curves.

$$D_{\text{KS}} = \max_{0 \le i \le M-1} \mid F_{\text{obs}}(\tau_i) - F_{\text{prof}}(\tau_i) \mid . \tag{18}$$

(iii) Cosine similarity of pseudo-intensity (increment) sequences.

Pseudo-intensities were obtained as first differences of the cumulative curves on the τ-grid:

$$\Delta F_{\text{obs},i} = F_{\text{obs}}(\tau_{i+1}) - F_{\text{obs}}(\tau_i), \Delta F_{\text{prof},i} = F_{\text{prof}}(\tau_{i+1}) - F_{\text{prof}}(\tau_i), i = 0, \ldots, M-2.$$

The (non-negative) vectors $\Delta \text{F}_{\text{obs}}$ and $\Delta \text{F}_{\text{prof}}$ were then compared via

$$\cos\left(\Delta \text{F}_{\text{obs}}, \Delta \text{F}_{\text{prof}}\right) = \frac{\sum_{i=0}^{M-2} \Delta F_{\text{obs},i} \Delta F_{\text{prof},i}}{\left(\sum_{i=0}^{M-2} \Delta F_{\text{obs},i}^2\right)^{1/2} \left(\sum_{i=0}^{M-2} \Delta F_{\text{prof},i}^2\right)^{1/2}}, \in [0,1]. \tag{19}$$

Equations (17)–(18) capture cumulative-pattern errors (overall and worst-case timing), while Equation (19) captures similarity in the fine-scale temporal distribution of rainfall.

3.6. Best-Fit Profile Selection (Per Storm)

A hierarchical, metric-based rule was applied to select a single best profile for each observed event.

Stage 1—RMSE screening. For storm i, all profiles $p = 1, \ldots, P$ were ranked by $\mathrm{RMSE}_{i,p}$. A short-list of the K_{screen} lowest-RMSE profiles was retained (with $K_{\mathrm{screen}} = 5$ in this study):

$$\mathcal{S}_i = \mathrm{TopK}\{\mathrm{RMSE}_{i,p} : p = 1, \ldots, P\}. \tag{20}$$

Stage 2—Cosine refinement. Within $\mathcal{S}_i$, the profile with the largest cosine similarity of pseudo-intensities was selected,

$$p_i^\star = \underset{p \in \mathcal{S}_i}{\mathrm{argmax}} \cos(\Delta \mathrm{F}_{\mathrm{obs}}^{(i)}, \Delta \mathrm{F}_{\mathrm{prof}}^{(p)}). \tag{21}$$

Tie-break (if needed). If multiple profiles attained statistically indistinguishable cosine values, the one with the smaller D_{KS} (Equation (18)) was preferred.

This two-stage rule ensured that the selected profile simultaneously matched the overall cumulative trajectory closely and (b) reproduced the intra-storm distribution of rainfall increments.

3.7. Aggregation of Wins and Optional Depth Weighting

After assigning a best-fit profile $p_i^\star$ to every storm i, profile performance was summarized by "win" shares. For unweighted wins,

$$W_p = \frac{\{i : p_i^\star = p\}}{N_{\mathrm{storms}}} \times 100\%. \tag{22}$$

An optional depth-weighted variant was also computed to emphasize storms with larger totals. Let $w_i \geq 0$ denote the weight for storm i (e.g., proportional to total depth and normalized to $\sum_i w_i = 1$). Then

$$W_p^{(\mathrm{wt})} = \frac{\sum_{i:p_i^\star = p} w_i}{\sum_{i=1}^{N_{\mathrm{storms}}} w_i} \times 100\% = \Big(\sum_{i:p_i^\star = p} w_i\Big) \times 100\%. \tag{23}$$

Cluster-wise win percentages were additionally computed by restricting the numerator and denominator of (22)–(23) to storms belonging to a given Part-One cluster. These aggregated quantities are reported and visualized in the Results (e.g., the winners bar plot and cluster-breakdown tables) to document how each profile performs overall and within storm types.

4. Results

Figure 2 shows the variation of clustering evaluation metrics as a function of the number of clusters (K). The left panel shows the distortion (within-cluster sum of squared distances), where an elbow or inflection point is visible around K = 4. The middle panel displays the average silhouette coefficient; this curve reaches a clear maximum at K = 2 (silhouette $\approx$ 0.606), indicating the best separation at two clusters. The right panel depicts the Gap statistic, which increases almost monotonically with K over the tested range (becoming less negative from K = 2 to K = 10) and does not exhibit a well-defined peak. Based on these results, a two-cluster solution was adopted for subsequent analysis.

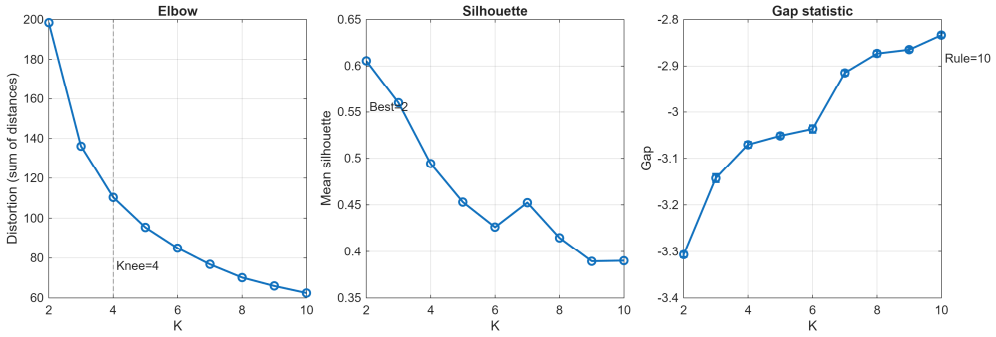

Figure 2. Cluster-number diagnostics (elbow, silhouette, and Gap statistic).

The final partition of the dataset consisted of a larger Cluster 1 and a smaller Cluster 2. Cluster 1 included 6264 storms (~70% of all events), whereas Cluster 2 comprised 2708 storms (~30%). These two clusters exhibited markedly different rainfall temporal patterns, as described below. These complementary diagnostics collectively support the selection of a two-cluster solution, balancing model simplicity and interpretability with statistically meaningful separation between storm types.

Mean normalized cumulative rainfall depth curves (centroid profiles) for each cluster Figure 3. The blue curve corresponds to Cluster 1 and the orange curve to Cluster 2. It can be seen that the Cluster 1 profile rises much more steeply in the early part of the storm, indicating that a large fraction of the total rainfall depth is accumulated in a short portion of the storm's duration. For example, by ~50% of the storm duration, the Cluster 1 average event has already delivered over 80% of its total depth, whereas the Cluster 2 average has delivered only around half of its depth by the same point. Both cluster profiles ultimately reach 100% of depth at the end of the storm (by definition), but the trajectory of accumulation over time differs substantially between the two clusters, with Cluster 1 being much more front-loaded and Cluster 2 more evenly distributed.

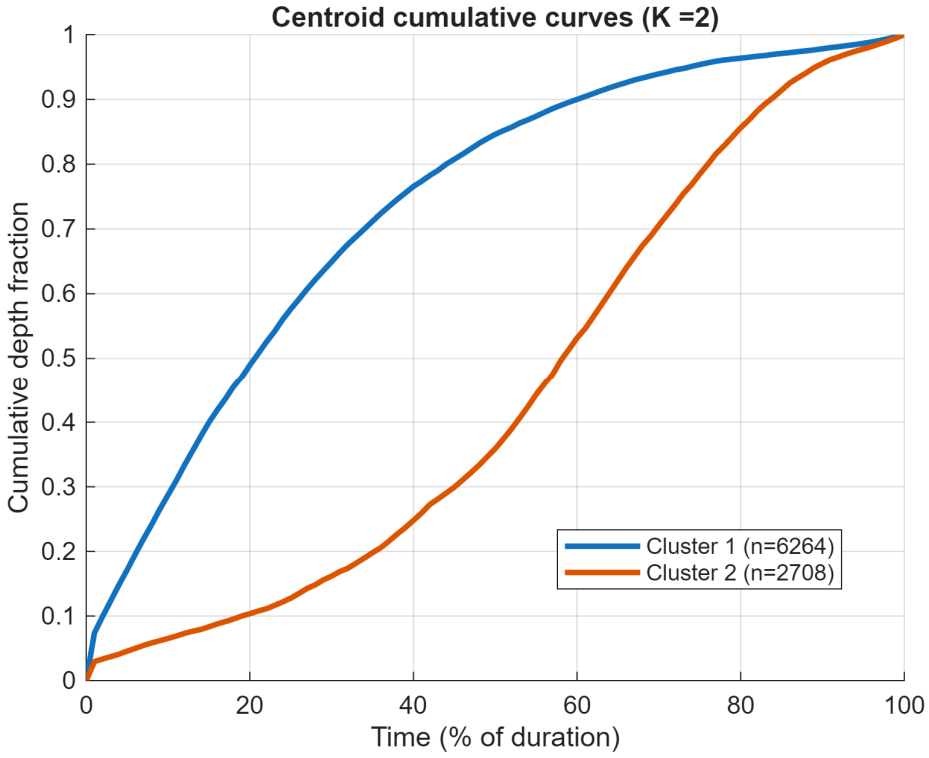

Figure 3. Centroid cumulative rainfall curves for the two clusters (K = 2).

From a hydrological perspective, the Cluster 1 profile represents short-duration, high-intensity storm events characterized by rapid accumulation of rainfall, whereas Cluster 2 reflects longer-duration events with delayed peak intensities. This distinction is critical for flood response, as early peaking storms are more likely to generate rapid runoff and flash flooding.

Key quantitative metrics for these temporal profiles are summarized in Table 1. As shown, Cluster 1 exhibited consistently lower T25%, T50%, and T75% values than Cluster 2, meaning that Cluster 1 storms reach each fraction of total depth much earlier in their duration compared to Cluster 2 storms (Table 1). For instance, the median time to accumulate 25% of total depth (T25%) was only 8.3% of the storm duration in Cluster 1, whereas it was 40.2% in Cluster 2. Similarly, the median T50% (time to 50% depth) was 20.6% for Cluster 1 versus 58.2% for Cluster 2, and the median T75% was 38.4% versus 72.6%, respectively. Cluster 2 also showed greater variability among events, as reflected by much wider interquartile ranges. For example, T25% in Cluster 2 ranged from 29.4% to 59.0% (IQR), compared to 3.2–15.3% in Cluster 1 (a considerably narrower IQR), indicating that the timing of early rainfall accumulation is far more variable in Cluster 2 (Table 1). These statistical summaries quantitatively confirm that Cluster 1 storms tend to be front-loaded (most rainfall occurs early), whereas Cluster 2 storms are delayed in their rainfall accumulation.

Table 1. Summary of normalized time to, 50% of total storm depth in each cluster. Values are given as mean ± standard deviation, with median [interquartile range] in parentheses (all values in % of storm duration). These timing metrics provide a quantitative basis for distinguishing storm types and demonstrate the strong temporal asymmetry between clusters, reinforcing the limitations of conventional symmetric or fixed-shape design storms.

Cluster (n)	T25 Mean ± SD	T50 Mean ± SD	T75 Mean ± SD
1 (6264)	10.0 ± 8.2	22.4 ± 15.3	39.9 ± 21.6
2 (2708)	44.7 ± 20.8	61.3 ± 15.0	72.8 ± 12.6

Cumulative rainfall profiles for all storm events in Cluster 1 (left panel) and Cluster 2 (right panel) as shown in Figure 4. In each panel, the thin gray lines represent individual storms normalized cumulative depth over time (from 0% to 100% on both axes), illustrating the variability of temporal patterns within the cluster. The thick black line in each panel is the cluster's mean profile (identical to the colored curves shown in Figure 4), and the shaded area denotes the 10th–90th percentile range of the cumulative depth at each time point. It is evident that Cluster 1 storms generally climb to high cumulative fractions very quickly, with most gray curves reaching 80–100% depth within the first half of the duration. In contrast, Cluster 2 storms rise more gradually: many gray curves remain in the lower depth fractions (below 50%) until well past the mid-point of the storm's duration. The spread of the gray curves is also larger in Cluster 2, reflecting the greater event-to-event variability in how rainfall is distributed over time (consistent with the wider IQR values noted in Table 1).

Heatmap of cumulative rainfall fraction over time for all 8972 storms, sorted by cluster and by within-cluster temporal order. Each horizontal line represents an individual storm's progression of accumulated depth (from 0% to 100% on the *x*-axis, normalized time) as shown in Figure 5. Color indicates the fraction of total depth achieved at a given time (blue = 0, yellow = 1, with green shades in between). The top portion of the heatmap corresponds to Cluster 1 storms (n = 6264), and the bottom portion corresponds to Cluster 2 storms (n = 2708). A clear contrast is observed: the Cluster 1 section turns green/yellow much earlier along the time axis, whereas the Cluster 2 section remains blue for a longer

duration and only approaches yellow near the end of the storm period. This visualization underscores the distinct temporal structure of the two clusters, visually confirming that Cluster 1 events concentrate their rainfall early in the storm, while Cluster 2 events have a more delayed accumulation of rainfall depth.

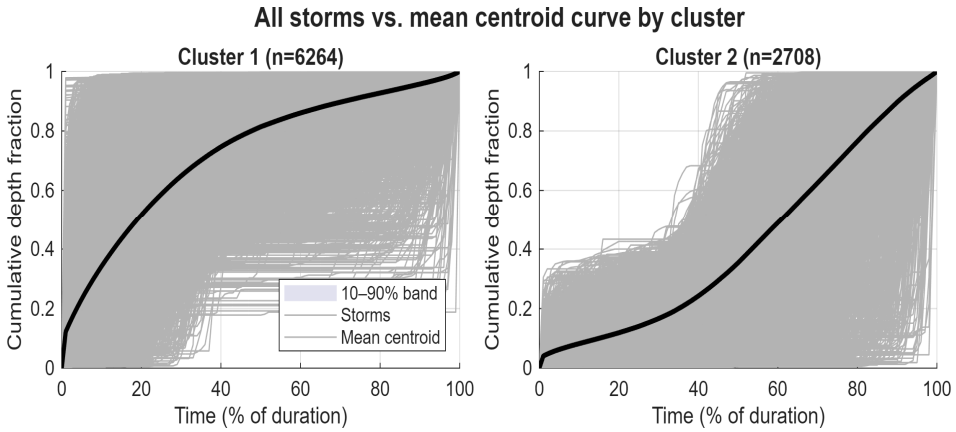

Figure 4. All storms vs. the mean centroid curve by cluster (with 10–90% envelope).

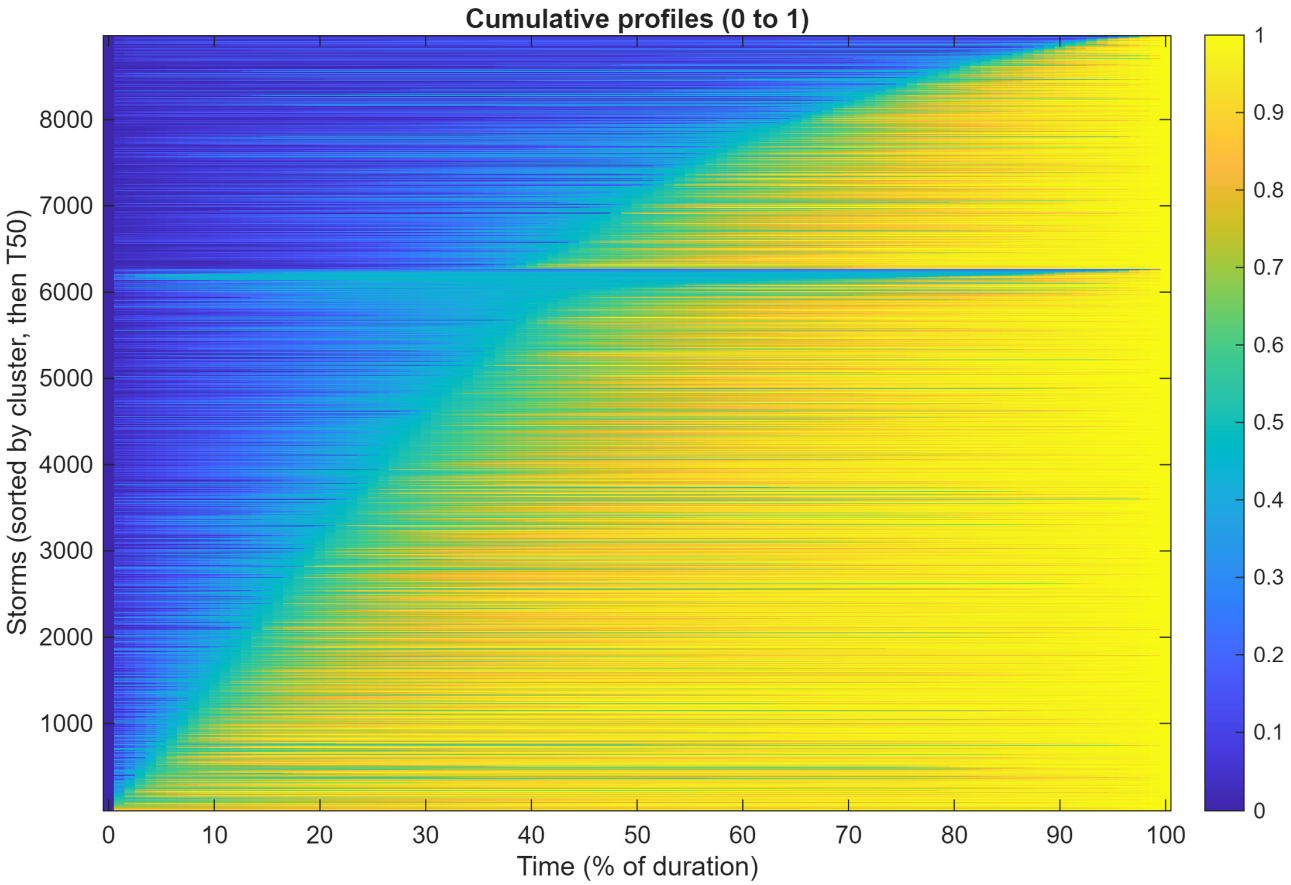

Figure 5. Heatmap of normalized cumulative rainfall profiles, ordered by cluster and median time (T50).

Next, the comparison focuses on evaluating cluster-derived (data-driven) profiles against conventional design storm distributions, rather than comparing alternative clustering algorithms. The performance of each candidate synthetic design storm profile was evaluated against all 8972 events using three metrics: RMSE, KS statistic, and cosine similarity (as defined in Equations (17)–(19) in Part One). Table 2 compiles the error statistics

(mean, median, and standard deviation for RMSE and KS, and likewise for cosine similarity) for each profile across the entire dataset. Lower RMSE and KS values (closer to 0) and higher cosine similarity (closer to 1) indicate a better overall fit to the observed storms. Among the eight conventional design storms, the Euler Type II pattern achieved the lowest overall errors, with a mean RMSE of ~0.113 (median ~ 0.100) and a relatively low mean KS (~0.211). Euler II also had one of the highest mean cosine similarities (~0.74), indicating its temporal shape aligns well with many storm events. The Alternating Block Method (ABM) and Huff Quarter 1 (Q1) profiles also performed relatively well overall, with mean RMSE around 0.14–0.15 and high cosine similarity (~0.72–0.80). These profiles are all front-loaded or evenly distributed patterns, suggesting that many storms in the dataset have rainfall concentrated toward the start or fairly uniform distribution, which these designs can approximate. In contrast, profiles that are heavily back-loaded (e.g., Huff Q3 and Huff Q4) showed much larger errors on average. For instance, Huff Q3 yielded the highest mean RMSE (~0.292) and highest mean KS (~0.471) of all profiles, indicating a poor fit for most events. This is not surprising given that Huff Q3's pattern (rainfall mostly in the middle to late portion of the storm) is quite extreme and did not match the majority of observed storms. Huff Q4 was similarly poor (mean KS ~0.424 and a high RMSE). The widely used SCS Type II profile exhibited intermediate performance overall: its mean RMSE (~0.163) was moderate, and while its median RMSE was fairly low (~0.12, indicating it fits a subset of events well), its error variability was high (standard deviation ~0.069), implying it performed poorly for other events. The Uniform profile (rainfall evenly distributed over time) showed moderate error (mean RMSE ~ 0.148) and a decent average KS (~0.204), reflecting that a purely uniform storm is not a terrible approximation for many events, but often not the best either. Notably, the two AI-derived cluster centroids (AI-C1 for Cluster 1, and AI-C2 for Cluster 2) did not have the lowest errors when evaluated over all storms globally—each had higher mean RMSE and KS than the best conventional patterns. For example, AI-C1's overall mean RMSE was ~0.210 and AI-C2's was ~0.261, both larger than Euler II's 0.113. This is expected, because each cluster-based profile is specialized to its own subset of storms and can deviate significantly from storms belonging to the other cluster. In other words, neither AI-C1 nor AI-C2 is a one-size-fits-all solution across the entire dataset. However, as shown in the next section, within their intended clusters these AI profiles achieve much better fits, outperforming standard designs for those respective groups.

Cumulative rainfall patterns of all candidates' synthetic profiles, illustrating the variety of temporal shapes, are shown in Figure 6. The eight standard design storms (colored solid lines) include front-loaded patterns (e.g., Huff Q1, ABM) that rise very quickly, evenly distributed patterns (e.g., Uniform—nearly a straight diagonal line), and back-loaded patterns (e.g., Huff Q3 and Huff Q4) that rise slowly at first and only steepen later in the storm. The two cluster-derived profiles (AI-C1 for Cluster 1 and AI-C2 for Cluster 2, shown as dashed black lines) capture the representative front-loaded and back-loaded behaviors identified in this study. Notably, AI-C1 is an extremely front-heavy pattern (even more so than Huff Q1, with an earlier rise), whereas AI-C2 peaks later in the storm (somewhat akin to Huff Q3, though less extreme). These shape differences explain the error trends: profiles with a mismatch in timing (e.g., a late-peaking profile applied to an early-peaking storm) yield higher RMSE and KS errors, as observed for Huff Q3 and Q4 in Table 2.

While aggregate performance metrics provide insight into the overall accuracy of synthetic profiles across the full dataset, a more detailed evaluation was conducted to assess which profile best matched each individual storm event. For all 8972 observed rainfall events, the synthetic profile yielding the lowest cumulative error—based on a composite ranking of RMSE, KS distance, and NSE—was identified as the best-fitting

candidate. This event-level analysis, referred to as the "best-fit frequency" or "winner count" approach, quantifies how often each profile provided the closest match to real-world storm behavior. The results are summarized in Table 3 and visually depicted in Figure 7.

Table 2. Performance metrics for each synthetic rainfall profile across all 8972 events. For each profile, the table lists the mean, median, and standard deviation of the RMSE and KS statistics (lower values indicate better fit), as well as the mean, median, and standard deviation of the cosine similarity (higher values indicate a more similar rainfall distribution shape). Profiles are sorted roughly from best to worst overall performance (top to bottom by RMSE_mean). The performance metrics presented in Table 2 quantify the degree to which each candidate profile reproduces observed storm behavior, providing an objective basis for comparing conventional and data-driven approaches.

Profile	RMSE Mean	RMSE Median	RMSE std	KS Mean	KS Median	KS std	Cosine Mean	Cosine Median	Cosine std
Euler Type II	0.113	0.100	0.037	0.211	0.160	0.04	0.740	0.800	0.05
ABM	0.153	0.180	0.046	0.276	0.230	0.05	0.800	0.810	0.07
Huff Q1	0.144	0.120	0.037	0.208	0.200	0.08	0.720	0.730	0.13
Uniform	0.148	0.150	0.080	0.204	0.190	0.09	0.680	0.690	0.10
SCS Type II	0.163	0.120	0.069	0.256	0.180	0.10	0.710	0.720	0.08
Huff Q2	0.236	0.280	0.069	0.417	0.450	0.14	0.660	0.670	0.13
AI-C1 (Cluster 1)	0.210	0.150	0.092	0.390	0.300	0.14	0.694	0.700	0.09
AI-C2 (Cluster 2)	0.261	0.300	0.060	0.450	0.500	0.10	0.539	0.550	0.10
Huff Q4	0.280	0.330	0.073	0.424	0.500	0.11	0.480	0.500	0.14
Huff Q3	0.292	0.340	0.073	0.471	0.500	0.12	0.560	0.580	0.15

Table 3. Overall best-fit profile counts and percentages. For each synthetic profile, the number of events (out of 8972) for which it provided the best fit (i.e., lowest error) is listed, along with the corresponding percentage of all events. Profiles are ordered by decreasing best-fit frequency.

Profile	Best-Fit Count	Best-Fit Percentage (%)
AI-C1	3790	42.24%
AI-C2	1290	14.38%
SCS Type II	920	10.25%
Uniform	682	7.60%
Euler Type II	673	7.50%
ABM	653	7.28%
Huff Q3	319	3.56%
Huff Q2	289	3.22%
Huff Q4	185	2.06%
Huff Q1	171	1.91%

The findings reveal a clear advantage for the cluster-derived synthetic profiles over traditional design storm shapes. AI-C1, representing the front-loaded centroid of Cluster 1, was the most frequently selected best-fit profile, successfully matching 3790 events—equivalent to approximately 42.24% of all storms. AI-C2, corresponding to the back-loaded centroid of Cluster 2, was optimal for 1290 events (14.38%). Together, these two AI-generated profiles accounted for the best representation of over half the dataset (56.6%),

demonstrating their superior ability to capture the temporal variability inherent in observed rainfall patterns.

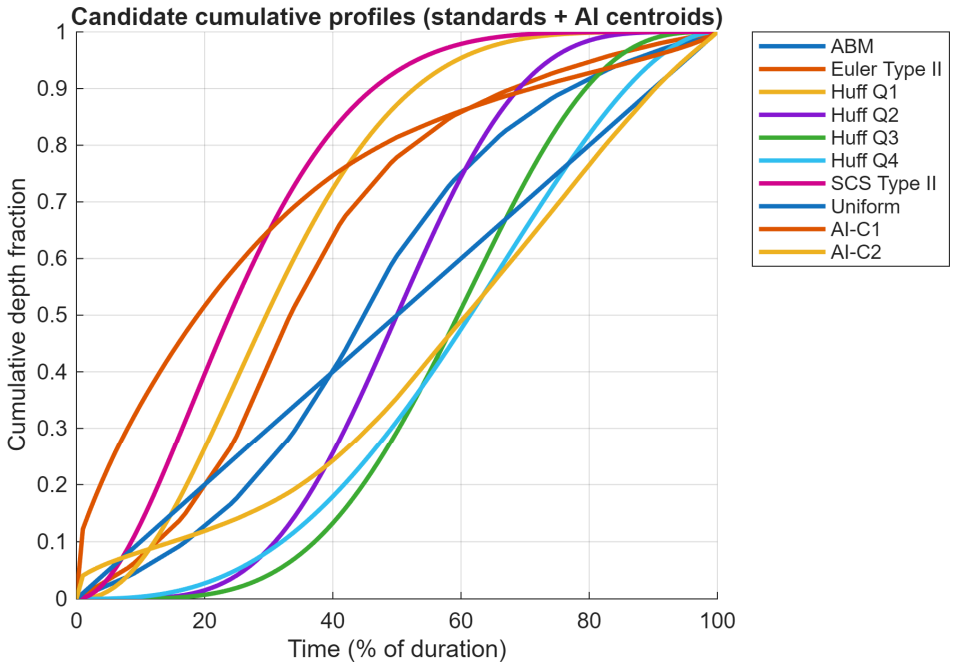

Figure 6. The cumulative rainfall patterns of all candidates synthetic profiles.

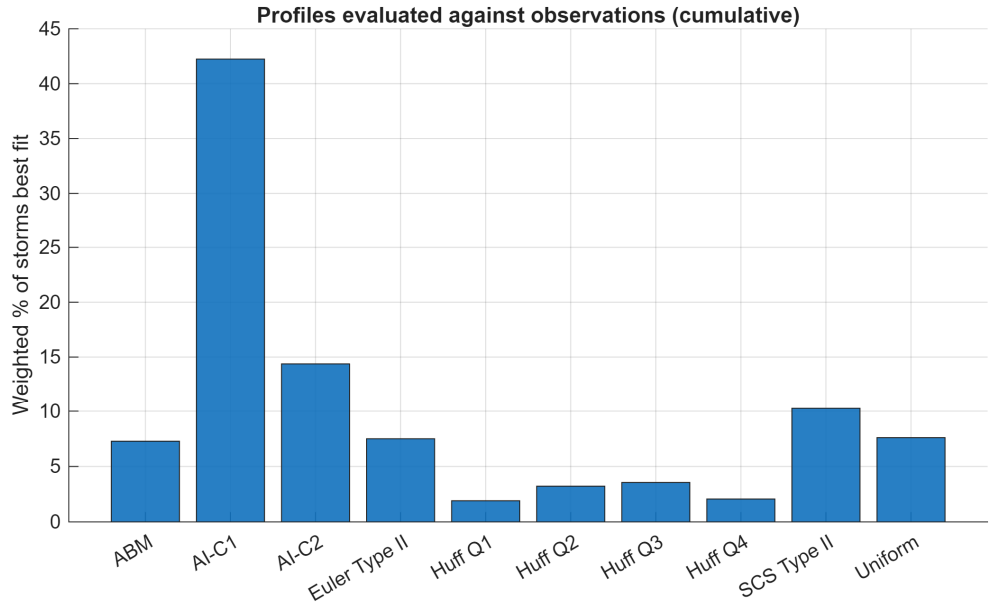

Figure 7. The profile evaluation against each storm.

Among the conventional design storms, SCS Type II was the top performer, providing the best fit for 920 events (10.25%). Other standard profiles showed moderate success: the Uniform pattern matched 682 events (7.60%), Euler Type II matched 673 events (7.50%), and ABM matched 653 events (7.28%). In contrast, the Huff quartile profiles were rarely identified as the best-fitting option. Huff Q3 matched 319 events (3.56%), Huff Q2 matched 289 (3.22%), Huff Q4 matched 185 (2.06%), and Huff Q1 matched only 171 events (1.91%). Each Huff profile accounted for less than 4% of best-fit cases, indicating limited alignment with the temporal structure of observed storms.

This visualization and analysis underscore the substantial portion of events that are more accurately captured by the new AI-derived patterns—specifically, the Cluster 1 and Cluster 2 centroids—compared to any single standard design. The dominance of AI-C1 and AI-C2 in best-fit frequency highlights the effectiveness of data-driven clustering approaches in generating synthetic profiles that better reflect the diversity and dynamics of real storm events. These results support the adoption of cluster-tailored profiles in hydrological modeling and design applications where temporal fidelity is critical.

To further understand the performance dynamics of synthetic storm profiles, the best-fit performance varies greatly between Cluster 1 and Cluster 2 events, as shown in Table 4. Within Cluster 1 (front-loaded storms), the cluster-derived profile AI-C1 was by far the most successful: it provided the best fit for 60.50% of Cluster 1 events. The next closest competitor in Cluster 1 was SCS Type II (best for 14.69% of Cluster 1 storms), followed by Euler Type II (10.04%) and the Uniform profile (6.32%). All other profiles were negligible in Cluster 1—for example, the ABM pattern was the top fit for only ~4.1% of Cluster 1 storms, and each of the Huff profiles was best for <3% (Huff Q1 about 2.7%, and Huff Q2/Q3/Q4 effectively 0% in this cluster). This means the vast majority of front-heavy storms in Cluster 1 were much better captured by the AI-C1 centroid pattern than by any standard design; only a minority of Cluster 1 events aligned better with traditional profiles like SCS Type II or Euler II.

Table 4. Best-fit profile performance within each cluster. Entries indicate the percentage of events in Cluster 1 and in Cluster 2 (respectively) for which a given profile was the best-fitting pattern. For example, AI-C1 was the best match for 60.50% of Cluster 1 storms and 0% of Cluster 2 storms, while ABM was best for ~4.09% of Cluster 1 and 14.66% of Cluster 2 storms.

Profile	Best-Fit Count	Best-Fit Percentage (%)
AI-C1	60.50%	0.00%
AI-C2	1.60%	43.94%
SCS Type II	14.69%	0.00%
Uniform	6.32%	10.56%
Euler Type II	10.04%	1.62%
ABM	4.09%	14.66%
Huff Q1	2.73%	0.00%
Huff Q2	0.03%	10.60%
Huff Q3	0.00%	11.78%
Huff Q4	0.00%	6.83%

In Cluster 2 (back-loaded storms), the cluster centroid AI-C2 also achieved the highest best-fit rate, though its dominance was less absolute than AI-C1's in Cluster 1. AI-C2 was the best match for 43.94% of Cluster 2 events. The remaining Cluster 2 storms were often better described by certain conventional profiles: notably, ABM was the top profile for 14.66% of Cluster 2 storms, and Huff Q3 was best for 11.78%. The Uniform and Huff Q2 profiles each accounted for roughly ~10.6% of Cluster 2 best-fits as well. These results suggest that while the AI-C2 centroid captured a large portion of back-loaded events, a significant subset of Cluster 2 storms were closer in pattern to the traditional ABM or Huff temporal distributions. It is also evident that some profiles which performed well in one cluster performed poorly in the other; for instance, the SCS Type II profile that was frequently successful in Cluster 1 did not best-fit any Cluster 2 storms (0% in Table 4). This underscores that a front-loaded design like SCS II is unsuited to the late-peaking rainfall of

Cluster 2. Overall, the cluster-wise breakdown highlights how each cluster's representative AI profile provided the most accurate fit for that cluster's events, while the performance of standard profiles varied greatly with storm type (front-loaded vs. back-loaded).

In summary, tailoring design storm profiles to data-driven clusters of rainfall patterns offers substantial improvements in fidelity. Each cluster's centroid (AI-C1 for front-loaded storms, AI-C2 for back-loaded storms) outperformed all standard design storms for its respective subset of events. The conventional profiles, by contrast, show clear strengths and weaknesses depending on storm type: a profile like SCS Type II (front-heavy) can mimic Cluster 1 storms reasonably well, but fails for Cluster 2 events, whereas a late-peaking profile like Huff Q3 can capture some Cluster 2 storms but is a poor fit for Cluster 1. The results demonstrate that no single traditional design storm suits all rainfall scenarios; however, by introducing two prototype patterns representative of front- and back-loaded behaviors, it can cover a much broader range of events with higher accuracy. This two-pattern approach (as embodied by AI-C1 and AI-C2) achieves a more robust, site-specific design basis for storm temporal distributions compared to the one-size-fits-all standard profiles, which is a key finding of this study.

To facilitate direct use in hydrologic and hydraulic models, dimensionless cumulative distributions $F(\tau)$ for the two AI-derived centroid profiles were tabulated at decile time steps (Table 5). As shown, the Cluster-1 centroid (AI-C1) is markedly front-loaded: by mid-duration ($\tau = 0.50$), $F(\tau)$ has reached 85% of total depth and exceeds 90% by $\tau = 0.60$. In contrast, the Cluster-2 centroid (AI-C2) exhibits delayed accumulation, reaching only 36% at $\tau = 0.50$ and 53% at $\tau = 0.60$, with rapid gains confined to the latter portion of the storm (e.g., 86% by $\tau = 0.80$ and 96% by $\tau = 0.90$). These distributions provide ready-to-apply, normalized hyetographs that can be linearly interpolated between deciles and differenced to obtain incremental intensities; they therefore offer practical templates for event-based simulation consistent with the observed front-loaded (AI-C1) and back-loaded (AI-C2) regimes identified in the clustering analysis.

Table 5. Dimensionless cumulative hyetographs for AI centroid profiles at decile time steps.

Percentage of the Cumulative Time	AI-C1	AI-C2
0%	0%	0%
10%	29%	7%
20%	49%	10%
30%	65%	16%
40%	77%	25%
50%	85%	36%
60%	90%	53%
70%	94%	71%
80%	96%	86%
90%	98%	96%
100%	100%	100%

5. Discussion

The results identified two distinct clusters of storm hyetograph patterns in the south-western Saudi regions of Jazan, Makkah, Al Bahah, and Aseer. Cluster 1 storms are "front-loaded", meaning a large fraction of the total rainfall occurs in the early part of the event. In fact, the representative Cluster 1 profile reaches about 75% of its total rainfall by only ~38% of the storm duration. This behavior is characteristic of convective cloudbursts common

in arid, mountainous areas—intense rainfall bursts at the storm's onset. Such front-heavy distribution aligns with observations in nearby regions; for example, in northern Oman's Rustaq watershed (an environment climatically comparable to southwest Saudi Arabia), half of all storms delivered ~80–85% of their rain in the first quarter of event time [27]. This tendency for arid-region thunderstorms to be sharply peaked and early-intense has been documented in other dry climates as well, including early comparisons of Texas vs. Illinois storms [28] and in Egypt's Sinai Peninsula, where typical rain events last only 2–3 h and tend to be highly concentrated [29,30]. On the other hand, Cluster 2 storms are "back-loaded", with rainfall intensifying in the later stages of the event. Cluster 2 centroid reaches 75% of total rain by ~73% of the duration—effectively a mirror-image temporal pattern. These back-heavy storms could reflect stratiform or multi-cell storm systems where rainfall persists and peaks towards the end. The existence of two almost opposite profile types highlights the bimodal nature of rainfall temporal distributions in Saudi's arid southwest, likely driven by differing meteorological regimes (e.g., isolated convective cells versus longer-duration orographic storms). Importantly, this broadens the context of the findings: not all storms in arid regions follow a single "design" shape, but rather a spectrum—with some very brief and intense upfront, and others more prolonged with delayed peaks. Recognizing these distinct profiles is essential for designing infrastructure that can handle both rapid flash floods and longer, late-peaking rain events.

When the findings are placed in the context of existing literature, both resonances and advancements are seen. The suitability of imported standard storm profiles for arid climates has long been questioned by prior studies in the Arabian Peninsula. Local storm hyetographs for Saudi Arabia were de-rived by Elfeki et al. (2014), and it was noted that widely used synthetic distributions did not reflect the observed rainfall timing [30]. In Oman, it was explicitly found by Awadallah et al. (2017) that the *Soil Conservation Service (SCS) Type II and UK50 de-sign storm patterns—developed for temperate regions—were "not safe" for use in arid and hyper-arid areas, with peak timing and intensity often misrepresented [31]. These concerns are strongly corroborated by the results. It was observed that standard profiles like the SCS Type II, Huff quartiles, and the Alternating Block Method (ABM) patterns seldom pro-vided the best fit to Saudi events. In fact, the closest match for over 56% of all recorded storms was individually pro-vided by the two AI-derived cluster centroids (AI-C1 and AI-C2), whereas no single traditional design storm exceeded about 10% best-fit frequency. The value of an AI/clustering approach is underscored by this dominance: by objectively learning the patterns inherent in local data, synthetic profiles that outperform legacy designs fine-tuned for other climates were obtained. A similar approach was recently demonstrated in Poland's temperate climate by Wartalska et al. (2020), in which k-means clustering was used to classify heavy rainfall hyetographs [28]. Four typical rain pattern groups were identified, one of which closely resembled the Euler Type II design storm (a front-loaded profile) [28]. Notably, their cluster-based grouping was more effective in characterizing storm shapes than the conventional Huff quartile classification [28]. The data-driven paradigm is extended into an arid-climate context by this work, with one of the first applications of machine learning to develop design storm profiles in Saudi Arabia's dry terrain being pro-vided. It is suggested by the success of the clustering method here that AI techniques can reliably uncover dominant storm structures that might be missed by pre-defined empirical distributions. It is worth mentioning that other forms of AI have also been used in regional hydrology—for instance, machine learning models have been applied to improve rainfall predictions in hyper-arid environments [32] and short-term forecasting in Saudi Arabia [33]. In the study, a straightforward yet powerful way to tailor design hyetographs to local storm dynamics was pro-vided by the unsupervised learning via clustering. A modern step forward in

de-sign storm methodology for arid regions is represented by this integration of AI-derived insights with classical hydro-meteorological analysis.

The identification of front- vs. back-loaded storm profiles is recognized as having tangible implications for engineering design and flood risk management in Saudi Arabia's arid zones. Infrastructure (such as storm sewers, detention basins, and road culverts) is typically designed using design storms—idealized rainfall distributions meant to produce conservative runoff estimates. If the design storm shape does not mimic reality, peak dis-charge calculations can be significantly off-target [31] fayoum.edu.eg. For example, drainage systems can be overwhelmed faster by a front-loaded real storm than would be predicted by a more evenly distributed design storm. It is indicated by the findings that relying on unadjusted imported profiles (like SCS Type II or Huff curves) in this region could cause early peak flows to be underpredicted or the critical intensity period to be mis-timed, potentially leading to flood control structures being undersized. By contrast, a closer match to actual storm temporal patterns is offered by using cluster-derived profiles (AI-C1, AI-C2), thereby allowing the realism of hydrological simulations to be improved. It was demonstrated that a majority of historical storms were best represented by these two centroid patterns rather than any standard profile. In practical terms, the accuracy of flood modeling across southwestern Saudi Arabia could be enhanced by the adoption of the two representative hyetographs for design. The capture of both the flash-flood potential of short, intense storms (via the AI-C1 front-loaded pattern) and the cumulative flooding threat of longer, back-heavy storms (via the AI-C2 pattern) would be better enabled for engineers and planners. This dual-profile approach could be incorporated into regional design guidelines, similar to how the Huff quartile concept provides multiple patterns, but with profiles specifically learned from local data rather than extrapolated from foreign climates.

The broader significance of these results extends beyond Saudi Arabia. Arid and semi-arid regions worldwide—from the Middle East and North Africa to parts of the southwestern United States—face analogous challenges in developing appropriate design storms. The study provides a template for data-driven, region-specific storm profile development. By leveraging long-term local pluviograph records and clustering techniques, water authorities in other dry regions can derive their own synthetic storms that reflect local idiosyncrasies (such as monsoonal burst patterns, orographic effects, etc.). This also dovetails with the increasing emphasis on performance-based and risk-informed design in the era of climate change. As extreme rainfall events grow more frequent and intense in many arid areas [34], having an accurate arsenal of design storm shapes becomes crucial to ensure infrastructure resilience. For instance, recent climate projections for the Arabian Peninsula suggest a potential doubling of extreme precipitation event frequency by late 21st century under high-emission scenarios [34]. Even observationally, the Jazan region has seen unusually heavy rains in recent years (e.g., record-breaking summer 2024 storms) with intensities causing severe floods [35]. These trends imply that historical "worst-case" temporal patterns may be regularly exceeded in the future. Therefore, using locally derived profiles that capture the true rapid-rise behavior of convective storms will be even more critical going forward. The two cluster profiles, grounded in the observed extremes of the past decade, inherently include such aggressive temporal distributions, whereas a generic SCS curve might grossly underestimate the initial burst of a future supercell thunderstorm. In summary, the implications of this work advocate for a shift towards localized design storm libraries in arid regions—an approach that can improve flood estimation accuracy, guide the design of more robust drainage infrastructure, and ultimately enhance public safety against flash floods.

While this study advances the development of storm profiles in Saudi Arabia, it is important to acknowledge its limitations. First, the clustering analysis was performed on dimensionless hyetographs pooled from multiple stations across a broad region. This approach assumes spatial homogeneity in storm shape characteristics, which may overlook micro-climatic differences. For example, coastal lowlands vs. high-altitude mountains within the study area might experience slightly different rainfall distributions (e.g., due to differing storm triggering mechanisms). A possible extension of this work could involve stratifying the data by sub-region or storm season to see if additional profile types emerge. Second, the choice of two clusters was driven by clear differences in the data and supported by cluster validity metrics; however, real storm behaviors form a continuum. Some observed events fell between the front- and back-loaded archetypes. Using only two centroid profiles to represent all events inevitably simplifies that variability. Future research might explore more nuanced classifications (additional clusters or continuous shape parameters) if warranted by larger datasets. Third, data limitations such as rain gauge resolution (hourly aggregation) and record length could affect the analysis. High-resolution (e.g., 5-min) rainfall data would capture storm burst dynamics more precisely, potentially refining the identification of peak timing. This was mitigated by focusing on cumulative patterns and using metrics (RMSE, KS, cosine similarity) that are relatively insensitive to short noise spikes. Nonetheless, any uncertainty in individual storm measurements propagates into the profile fitting and error statistics. It is also noted that the evaluation focused on temporal shape fit and did not explicitly integrate spatial coverage of storms or catchment response. Thus, while a profile might well-represent the timing of rainfall, actual flood impacts also depend on storm movement and spatial extent—factors beyond the scope of the analysis.

Despite these limitations, alignment with theoretical expectations and prior studies strengthens confidence in the results. Tailored, AI-clustered storm profiles substantially improve the representation of rainfall temporal distribution in arid-region contexts. This supports adopting region-specific design storms based on local data whenever possible [31,33]. For future reference, the centroid storm profiles are summarized in Table 5 to facilitate comparison and practical application.

6. Conclusions

In this study, dimensionless storm rainfall profiles representative of the arid, mountainous southwest of Saudi Arabia were developed and analyzed. Through a machine-learning clustering approach, two prototypical storm patterns were identified: a sharply front-loaded profile (AI-C1) that captures short-duration, high-intensity events, and a back-loaded profile (AI-C2) that represents longer storms with delayed peak rainfall. These profiles provided a markedly better fit to ~9000 observed storm events than conventional design-storm shapes drawn from other regions. The AI-derived profiles outperformed standard designs (SCS, Huff, ABM, etc.), confirming that local storms exhibit unique temporal dynamics that generic models fail to mimic. When contextualized with recent literature and regional climate considerations, the findings indicate that locally tailored storm hyetographs can significantly improve hydrological modeling and flood estimation in Saudi Arabia's arid areas. The two proposed profiles effectively bracket the range of storm behaviors in the study region and thus offer a practical, ready-to-use pair of design storms for engineers and planners: one to stress-test systems against sudden cloudburst floods, and another to account for prolonged heavy-rainfall scenarios.

The outcomes of this research carry important implications. On a scientific level, the utility of clustering and data-driven analyses in extracting meaningful patterns from rainfall records is demonstrated—an approach that can be replicated in other data-scarce,

dry environments to derive realistic design storms. On an engineering application level, the findings motivate a re-examination of current design standards in arid and semi-arid regions. Agencies in Saudi Arabia and similar climates are encouraged to update design-storm guidelines to incorporate profiles derived from local data (such as AI-C1 and AI-C2), replacing or supplementing imported curves from humid regions that are currently in use. Implementing these profiles in hydrologic calculations is expected to yield more reliable estimates of peak discharges and flood volumes, ultimately enhancing the safety and resilience of flood-control infrastructure. More broadly, the results reinforce that climate and weather extremes are not one-size-fits-all—regional analyses are essential as the basis for climate adaptation. With extreme precipitation events on the rise across the Arabian Peninsula, the methods and insights presented here provide a timely contribution toward sustainable water management in arid landscapes. It is concluded that, by pairing local observations with modern AI techniques, storm profiles can be developed that faithfully reflect observed behavior, thereby bridging the gap between historical climate data and forward-looking, robust design practices.

These findings are consistent with recent studies that highlight the limitations of conventional design storm approaches under non-stationary climatic conditions, particularly in arid and semi-arid regions. However, existing studies have largely focused on predictive applications of artificial intelligence, such as rainfall forecasting and flood modeling, rather than on the development of design-relevant storm hyetographs. To the best of current knowledge, no prior study has applied AI-based clustering to derive region-specific rainfall design storm profiles in southwestern Saudi Arabia or comparable arid mountainous environments. The present work therefore contributes a novel, data-driven framework that bridges this gap by directly linking observed rainfall structure with engineering design applications.

In addition, further refinement of the proposed framework may be achieved through the implementation of sub-regional clustering approaches, allowing for the delineation of spatially homogeneous zones within the study area. Such an extension would enable a more detailed representation of localized rainfall dynamics influenced by topography and microclimatic variability. Moreover, the incorporation of spatial storm coverage, including areal rainfall distribution and storm footprint characteristics, may further improve the applicability of the derived profiles in hydrological modeling and flood risk assessment. These extensions are expected to enhance the transferability and robustness of the methodology for broader regional applications.

Funding: This work was supported by the Ongoing Research Funding Program (ORF-2026-310), King Saud University, Riyadh, Saudi Arabia.

Data Availability Statement: New data was not created through this work.

Acknowledgments: The author would like to acknowledge King Saud University, Riyadh, Saudi Arabia, for their support of this work. Sincere thanks are extended to colleagues in the Department of Surface Water and Dams at the Ministry of Environment, Water, and Agriculture for providing the essential data that made this research possible. Also, the author is grateful to the peer reviewers for their constructive feedback, which greatly improved the quality of this study.

Conflicts of Interest: The author declares no conflicts of interest. The funding sources have not been involved in the study's design, data collection, analysis, or interpretation, the writing of the manuscript, or the decision to publish the results.

References

1. Almazroui, M.; Nazrul Islam, M.; Athar, H.; Jones, P.D.; Rahman, M.A. Recent climate change in the Arabian Peninsula: Annual rainfall and temperature analysis of Saudi Arabia for 1978–2009. *Int. J. Climatol.* **2012**, *32*, 953–966. [CrossRef]
2. Serinaldi, F.; Kilsby, C.G. Stationarity is undead: Uncertainty dominates the distribution of extremes. *Adv. Water Resour.* **2015**, *77*, 17–36. [CrossRef]
3. Alqadhi, S.; Mallick, J.; Talukdar, S.; Hang, H.T. Assessing short- and long-term drought severity and return periods using bivariate copula models in Saudi Arabia. *Appl Water Sci* **2025**, *15*, 141. [CrossRef]
4. Alharbi, R.S. Development and Evaluation of Dimensionless Design Storm Hyetographs for Southwestern Saudi Arabia in a Hyper-Arid Climate. *Atmosphere* **2025**, *16*, 1237. [CrossRef]
5. El-Haddad, B.A.; Youssef, A.M.; Mahdi, A.M.; Karimi, Z.; Pourghasemi, H.R. National flood susceptibility mapping in Saudi Arabia. *Earth Sci. Inform.* **2025**, *18*, 118. [CrossRef]
6. Mosavi, A.; Ozturk, P.; Chau, K.-w. Flood Prediction Using Machine Learning Models: Literature Review. *Water* **2018**, *10*, 1536. [CrossRef]
7. Zheng, K.; Wei, J.-Z.; Pei, J.-Y.; Cheng, H.; Zhang, X.-L.; Huang, F.-Q.; Li, F.-M.; Ye, J.-S. Impacts of climate change and human activities on grassland vegetation variation in the Chinese Loess Plateau. *Sci. Total Environ.* **2019**, *660*, 236–244. [CrossRef]
8. Kratzert, F.; Klotz, D.; Brenner, C.; Schulz, K.; Herrnegger, M. Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 6005–6022. [CrossRef]
9. Xiang, Z.; Yan, J.; Demir, I. A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water Resour. Res.* **2020**, *56*, e2019WR025326. [CrossRef]
10. Keum, H.J.; Han, K.Y.; Kim, H.I. Real-Time Flood Disaster Prediction System by Applying Machine Learning Technique. *KSCE J. Civ. Eng.* **2020**, *24*, 2835–2848. [CrossRef]
11. Hasanean, H.; Almazroui, M. Rainfall: Features and Variations over Saudi Arabia, A Review. *Climate* **2015**, *3*, 578–626. [CrossRef]
12. Almazroui, M.; Halwani, H.A.; Islam, M.N.; Ghulam, A.B.; Hantoush, A.S. Regional and seasonal variation of climate extremes over Saudi Arabia: Observed evidence for the period 1978–2021. *Arab. J. Geosci.* **2022**, *15*, 1605. [CrossRef]
13. Arsenault, R.; Brissette, F. Multi-model averaging for continuous streamflow prediction in ungauged basins. *Hydrol. Sci. J.* **2016**, *61*, 2443–2454. [CrossRef]
14. Hassan, D.; Isaac, G.A.; Taylor, P.A.; Michelson, D. Optimizing Radar-Based Rainfall Estimation Using Machine Learning Models. *Remote Sens.* **2022**, *14*, 5188. [CrossRef]
15. Corzo Perez, G.A.; Solomatine, D.P. (Eds.) Hydroinformatics and Applications of Artificial Intelligence and Machine Learning in Water-Related Problems. In *Advanced Hydroinformatics: Machine Learning and Optimization for Water Resources*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2024. [CrossRef]
16. Alsumaiei, A.A. Long-term rainfall forecasting in arid climates using artificial intelligence and statistical recurrent models. *J. Eng. Res.* **2025**, *13*, 1594–1602. [CrossRef]
17. Ferrario, D.M.; Sanò, M.; Maraschini, M.; Critto, A.; Torresan, S. *Review Article: Harnessing Machine Learning Methods for Climate Multi-Hazard and Multi-Risk Assessment*; preprint; EGUsphere: Munich, Germany, 2025. [CrossRef]
18. Sit, M.; Demiray, B.Z.; Xiang, Z.; Ewing, G.J.; Sermet, Y.; Demir, I. A comprehensive review of deep learning applications in hydrology and water resources. *Water Sci. Technol.* **2020**, *82*, 2635–2670. [CrossRef]
19. Ebtehaj, I.; Bonakdari, H. CNN vs. LSTM: A Comparative Study of Hourly Precipitation Intensity Prediction as a Key Factor in Flood Forecasting Frameworks. *Atmosphere* **2024**, *15*, 1082. [CrossRef]
20. Khan, T.A.; Alam, M.; Shahid, Z.; Suud, M.M. Prior investigation for flash floods and hurricanes: Concise capsulization of hydrological technologies and instrumentation—A survey. In Proceedings of the 3rd IEEE International Conference on Engineering Technologies and Social Sciences (ICETSS), Bangkok, Thailand, 7–8 August 2017; pp. 1–6.
21. Boero, R. An AI-Enhanced Systematic Review of Climate Adaptation Costs: Approaches and Advancements, 2010–2021. *Climate* **2024**, *12*, 116. [CrossRef]
22. Dotse, S.Q.; Larbi, I.; Limantol, A.M.; De Silva, L.C. A review of the application of hybrid machine learning models to improve rainfall prediction. *Model. Earth Syst. Environ.* **2024**, *10*, 19–44. [CrossRef]
23. Koldasbayeva, D.; Tregubova, P.; Gasanov, M.; Zaytsev, A.; Petrovskaia, A.; Burnaev, E. Challenges in data-driven geospatial modeling for environmental research and practice. *Nat. Commun.* **2024**, *15*, 10700. [CrossRef]
24. Krishnan, S.R.; Nallakaruppan, M.K.; Chengoden, R.; Koppu, S.; Iyapparaja, M.; Sadhasivam, J.; Sethuraman, S. Smart Water Resource Management Using Artificial Intelligence—A Review. *Sustainability* **2022**, *14*, 13384. [CrossRef]
25. Lin, F.; Yuan, X.; Zhang, Y.; Sigdel, P.; Chen, L.; Peng, L.; Tzeng, N.F. Comprehensive Transformer-Based Model Architecture for Real-World Storm Prediction. In *Machine Learning and Knowledge Discovery in Databases: Applied Data Science and Demo Track*; De Francisci Morales, G., Perlich, C., Ruchansky, N., Kourtellis, N., Baralis, E., Bonchi, F., Eds.; ECML PKDD 2023; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2023; Volume 14175. [CrossRef]

26. Marshall, S.R.O.; Tran, T.N.D.; Tapas, M.R.; Nguyen, B.Q. Integrating artificial intelligence and machine learning in hydrological modeling for sustainable resource management. *Int. J. River Basin Manag.* **2025**, 1–17. [CrossRef]
27. Al-Rawas, G.A.; Valeo, C.; Kohl, B.; Kirnbauer, R. Rainstorm characteristics in Oman and Austria. In Proceedings of the 12th Congress INTERPRAEVENT 2012, Grenoble, France, 23–26 April 2012; pp. 15–24.
28. Wartalska, K.; Kaźmierczak, B.; Nowakowska, M.; Kotowski, A. Analysis of Hyetographs for Drainage System Modeling. *Water* **2020**, *12*, 149. [CrossRef]
29. Elkollaly, M.S.; Gado, T.A.; Omara, H.; ElSawwaf, M.A.; Sefelnasr, A.; Ebraheem, A.; Sherif, M. Spatiotemporal Analysis of Event-Based Precipitation in Sinai, Egypt. In Proceedings of the 2nd International Conference on Water Resources Management and Sustainability: Solutions for Arid Regions (WARMS 2024), Dubai, United Arab Emirates, 26–28 February 2024. [CrossRef]
30. Elfeki, A.M.; Ewea, H.A.; Al-Amri, N.S. Development of storm hyetographs for flood forecasting in the Kingdom of Saudi Arabia. *Arab. J. Geosci.* **2014**, *7*, 4387–4398. [CrossRef]
31. Awadallah, A.G.; Elsayed, A.Y.; Abdelbaky, A.M. Development of design storm hyetographs in hyper-arid and arid regions: Case study of Sultanate of Oman. *Arab. J. Geosci.* **2017**, *10*, 456. [CrossRef]
32. Baljon, M.; Sharma, S.K. Rainfall Prediction Rate in Saudi Arabia Using Improved Machine Learning Techniques. *Water* **2023**, *15*, 826. [CrossRef]
33. Soliman, A. Development of Rainfall Design Storm Hyetographs for Al-Quassim Region -Kingdom of Saudi Arabia. *Int. J. Eng. Trends Technol.* **2022**, *70*, 362–372. [CrossRef]
34. Pathak, R.; Dasari, H.P.; Ashok, K.; Hoteit, I. Dynamics of intensification of extreme precipitation events over the Arabian Peninsula derived from CMIP6 simulations. *npj Clim. Atmos. Sci.* **2025**, *8*, 126. [CrossRef]
35. Islam, M.N.; Zamreeq, A.O.; Ismail, M.; Habeebullah, T.M.A.; Ghulam, A.S. Jazan Rainfall's Seasonal Shift in Saudi Arabia: Evidence of a Changing Regional Climate. *Atmosphere* **2025**, *16*, 300. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.