Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Calibrating Change-Point Inference in Regional Rainfall Extremes: Evidence from Mizoram, Northeast India

Michelle Lalhruaizeli Hnamte
Mizoram University
Laldinchhuaha Khiangte
khiangtedinchhuaha@gmail.com
Mizoram University
B. P. Mishra
Mizoram University
Vishwambhar Prasad Sati
Mizoram University

Research Article

Keywords: rainfall extremes, change-point detection, selective inference, field significance, wet-spell duration, northeast India

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11253767/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Calibrating Change-Point Inference in Regional Rainfall Extremes: Evidence from Mizoram, Northeast India

Authors

Michelle Lalhruaizeli[1], Laldinchhuaha Khiangte[2]*, B. P. Mishra[1], Vishwambhar Prasad Sati[2]

Affiliations

[1] Department of Environmental Science, Mizoram University, Aizawl 796004, Mizoram, India

[2] Department of Geography and Resource Management, Mizoram University, Aizawl 796004, Mizoram, India

Corresponding Author

Laldinchhuaha Khiangte

Department of Geography and Resource Management
Mizoram University
Aizawl 796004, Mizoram, India

Email: khiangtedinchhuaha@gmail.com
ORCID: 0009-0003-3258-7299

Author Information

Michelle Lalhruaizeli

Department of Environmental Science
Mizoram University, Aizawl 796004, Mizoram, India

Email: michelle_hnamte@hotmail.com
ORCID: 0009-0009-3648-9227

Laldinchhuaha Khiangte

Department of Geography and Resource Management
Mizoram University, Aizawl 796004, Mizoram, India

Email: khiangtedinchhuaha@gmail.com
ORCID: 0009-0003-3258-7299

B. P. Mishra

Department of Environmental Science
Mizoram University, Aizawl 796004, Mizoram, India

Email: mishrabp111@yahoo.com

Vishwambhar Prasad Sati

Department of Geography and Resource Management
Mizoram University, Aizawl 796004, Mizoram, India

Email: sati.vp@gmail.com

Highlights

- Uncalibrated change-point tests inflate false positives 17–265× in Mizoram rainfall.
- Calibrated surrogates confirm consecutive wet days (CWD) rose robustly, by 44.9%.
- Monsoon totals held steady, rainfall shifted to longer, less intense spells.

Abstract

Regional rainfall studies often detect a change point and test the resulting periods as if the break had been specified in advance, a procedure that severely inflates false positives when the break is instead selected from the same data being tested. We quantify this inflation and provide a general calibration framework, applied to ten ETCCDI indices from 0.25° gridded daily rainfall across 29 cells in Mizoram, northeast India (1950–2025). The procedure was calibrated using multivariate AR(1) surrogates that preserve temporal persistence and spatial covariance but contain no change point. In these signal-free series, the uncorrected procedure gave $p < 0.05$ in 83% and $p < 0.001$ in 27%, a 17-fold and 265-fold inflation over nominal levels. Applying the calibration to the observed record, only two of ten epoch contrasts survived correction. Consecutive wet days increased 44.9% across the 2006 break ($p = 0.003$) and was the only spatially coherent signal across all sensitivity settings. SDII declined 24.7%, significant only under the primary null ($p = 0.037$; sensitivity $p = 0.128$). Monsoon totals were unchanged (−2.9%, $p = 0.629$), while wet days increased 17.5% and intensity declined 16.5% (both $p \leq 0.001$), indicating longer, less intense spells rather than a supply change. The 232

mm decline in annual rainfall was not distinguishable from correlated noise ($p = 0.058$). These results show that regional rainfall analyses relying on searched change points must calibrate the full detection-and-testing procedure. Without it, evidence for regional rainfall change is almost entirely an artifact of break selection.

Keywords: rainfall extremes; change-point detection; selective inference; field significance; wet-spell duration; northeast India

1. Introduction

A common approach in regional rainfall analysis is to scan a record for a change point and then test the difference between the periods before and after the selected break. This is problematic because the break is chosen from the same data being tested: the resulting p-value assumes the split was fixed in advance, although it was selected to maximise the observed difference. Conditional methods have been developed to address this selection problem (Jewell et al., 2022), and recent climate studies have highlighted the sensitivity of change-point results to assumptions about trend and serial correlation (Lund et al., 2023), while the number and location of detected breaks is itself conditional on the assumed model and penalty (Shi et al., 2022). Yet regional rainfall studies rarely calibrate the complete detect-then-test procedure they use, and the design remains common even in careful recent work (Amirthanathan et al., 2023). Spatial dependence compounds this problem: serial correlation can produce apparent shifts even without a change in forcing, while spatial correlation can generate clusters of significant cells that appear to represent a coherent regional signal (Cortés et al., 2020). Change-point detection in gridded fields can also produce neighbouring shifts (Aghaei et al., 2025), and counting significant cells without accounting for covariance risks overstating spatial coherence (Huth & Dubrovský, 2021; Farris et al., 2021).

This inference problem is especially consequential in settings where the underlying climate signal is itself contested. A warming atmosphere is generally expected to intensify daily rainfall extremes, but the response is highly uneven across regions. India illustrates this contrast clearly. Extreme rainfall has increased over central India and parts of the coast (Goswami et al., 2006; Roxy et al., 2017), whereas recent assessments indicate declining southwest monsoon rainfall over northeastern India, including a decline of 0.5–1.5 mm day^{-1} per decade during 1951–2024 (Dhara et al., 2025). Other studies report declining annual precipitation over the northeast, linked to moisture-flux divergence (Borah et al., 2026), declining rainfall during both winter and monsoon (Chakraborty et al., 2025), and a regional change point after 1996 (Sharma et al., 2025). Singh et al. (2025) likewise found a decline in extreme rainfall events over the northeast using the same family of gridded data used here.

The evidence is not consistent, however. Das et al. (2024) found no significant change in high-intensity rainfall during 1991–2020, whereas Kolekar et al. (2025) reported increases in high-duration rain spells and maximum spell intensity over 1901–2023. Climate projections also suggest fewer wet days but greater rainfall intensity later this century (Paul & Maity, 2023). These differences partly reflect contrasting periods, datasets and rainfall indices. Given the inference problem outlined above, though, they may also arise from differences in how persistent and autocorrelated rainfall series are analysed, and from uncalibrated inference at searched change points.

We use Mizoram, in the southern part of northeast India, as the setting in which to develop and demonstrate a calibrated detect-then-test framework. The state's rainfall is strongly seasonal, and agriculture and other livelihoods remain highly dependent on the monsoon, making inference errors in either direction, whether over- or under-stating change, practically consequential. Evidence from Mizoram is itself limited and mixed: Lawmchullova et al. (2026) reported increasing rainfall at stations in the Tlawng watershed during 1984–2023, together with declining pre-monsoon rainfall, while earlier work reached similar conclusions for the same basin (Lawmchullova et al., 2024). In neighbouring Meghalaya, the same gridded product used here indicates longer dry periods (Phawa et al., 2022). The direction of rainfall change in Mizoram thus remains uncertain. That uncertainty is itself a symptom of the broader calibration problem this study addresses.

In this study, we address this calibration gap directly rather than abandoning change-point analysis. We analyse ten ETCCDI rainfall indices across 29 grid cells in Mizoram for 1950–2025, identify a change point in rainfall intensity, and evaluate the resulting epoch contrasts using multivariate AR(1) surrogates that preserve temporal persistence and inter-cell covariance but contain no true change point. We ask three questions: how strongly does the uncorrected procedure inflate false positives; which rainfall changes remain after calibration; and what do those changes indicate about the way monsoon rainfall is being delivered? We report both uncorrected and calibrated results because the contrast between them is central to the study.

2. Materials and methods

2.1 Study area and rainfall data

Mizoram occupies the southern part of northeast India (21.9–24.5°N, 92.2–93.4°E), with parallel north–south ridges ranging from below 100 m in the south to over 2000 m along the eastern margin. Rainfall is dominated by the southwest monsoon, with a secondary pre-monsoon maximum linked to Bay of Bengal convection.

Daily rainfall data were obtained from the 0.25° × 0.25° gridded India Meteorological Department product of Pai et al. (2014), which supersedes the earlier 1° analysis of Rajeevan et al. (2006). From the 40 cells covering 92.50–93.25°E and 22.00–24.25°N, 29 fall within Mizoram and form the analysis domain. The record covers 1 January 1950–31 December 2025 (27,759 days; 76 complete years; Figure 1).

All indices were calculated for each cell and then averaged across the 29 cells to obtain the regional series; cell-level results were retained for spatial analyses. A comparison with an earlier rainfall product showed an RMSE of 35.1 mm for the nearest cell relative to a mean Rx1day of 109.5 mm. Because the products differ, they were treated separately and are not combined in the analysis (Supplementary Table S16).

2.2 Quality control and homogeneity

No grid-year had more than 4% missing days, and no negative rainfall values were recorded; therefore, no infilling was applied. Annual rainfall totals were tested for homogeneity using the standard normal homogeneity test (Alexandersson, 1986), Buishand range test (Buishand, 1982), Pettitt test (Pettitt, 1979), and von Neumann ratio. Following Wijngaard et al. (2003), cells were classified as useful (0–1 tests rejecting homogeneity at the 1% level), doubtful (2), or suspect (3–4). The useful cells were retained as a homogeneous subset for sensitivity analysis.

2.3 Extreme indices

Ten ETCCDI indices were calculated: PRCPTOT, Rx1day, Rx5day, R95p, R99p, SDII, R10mm, R20mm, CDD and CWD (Zhang et al., 2011). These definitions follow those used in global assessments (Alexander et al., 2006; Donat et al., 2013). A wet day was defined as ≥1.0 mm, with percentile thresholds referenced to 1961–1990. Seasons followed the Indian convention: pre-monsoon (MAM), monsoon (JJAS), post-monsoon (OND), and winter (JF).

2.4 Trend estimation and serial correlation

Monotonic trends in the rainfall indices were evaluated using the non-parametric Mann–Kendall (MK) test. For a time series (x_1,x_2,\ldots,x_n), the MK statistic is defined as

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}\left(x_j - x_i\right) \quad (1)$$

where,

$$\operatorname{sgn}(x_j - x_i) = \begin{cases} +1, & x_j > x_i \\ 0, & x_j = x_i \\ -1, & x_j < x_i \end{cases} \quad (2)$$

The variance of S was corrected for tied observations according to

$$\operatorname{Var}(S) = \frac{n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p-1)(2t_p+5)}{18} \quad (3)$$

where g denotes the number of tied groups and (t_p) is the number of observations in the $(p) - th$ tied group.

Because annual rainfall indices may exhibit serial correlation, the variance of the MK statistic was adjusted using the Hamed–Rao (1998) approach. The correction was based on the lag-1 autocorrelation of detrended ranks. The adjusted variance can be represented generally as

$$\operatorname{Var}^*(S) = \operatorname{Var}(S) \cdot \mathrm{VIF} \quad (4)$$

where VIF represents the variance-inflation factor associated with the statistically significant autocorrelation structure of the detrended rank series.

The Hamed–Rao correction was used for the primary inference. Two alternative treatments of serial correlation were additionally considered: the effective-sample-size correction of Yue and Wang (2004) and the trend-free pre-whitening method of Yue et al. (2002). These alternative approaches were used to evaluate the sensitivity of the inferred trends to the treatment of temporal persistence.

Trend magnitude was estimated using the Theil–Sen estimator (Sen, 1968),

$$\hat{\beta}_{\mathrm{Sen}} = \operatorname{median}\left\{\frac{x_j - x_i}{j - i} \; : \; 1 \le i < j \le n\right\} \quad (5)$$

where $(\hat{\beta}_{Sen})$ represents the median slope across all possible pairs of observations.

Multiple testing across indices and spatial cells was controlled using the Benjamini–Hochberg false discovery rate (FDR) procedure (Benjamini and Hochberg, 1995), an approach now standard in hydrological trend detection where many sites are tested simultaneously (Douglas et al., 2000; Renard et al., 2008).

2.5 Change-point detection

A change in regional rainfall intensity was identified using the non-parametric Pettitt change-point test (Pettitt, 1979), applied to the regional SDII series. For a candidate change point k, the Pettitt statistic was calculated as

$$U_k = \sum_{i=1}^{k} \sum_{j=k+1}^{n} \operatorname{sgn}\left(x_i - x_j\right) \quad (6)$$

and the estimated change point was identified from the location of the maximum absolute value of U_k,

$$K = \underset{1 \le k < n}{\arg\max} |U_k| \quad (7)$$

The test identified 2006 as the change point in the regional SDII series. Because the estimator assigns the break year to the first segment, the rainfall record was divided into two epochs: 1950–2006 and 2007–2025.

Differences between the two epochs were initially evaluated using the two-sided Wilcoxon rank-sum test. However, because the break was identified through a search over the record, the resulting conventional *p*-value does not account for the selection of the break point. This issue was explicitly addressed through the surrogate calibration procedure described in Section 2.6.

2.6 Calibration of the detect-then-test procedure

Testing the difference between two periods after selecting the break point from the same rainfall record can produce substantially anti-conservative inference. The conventional epoch-comparison *p*-value implicitly assumes that the division between the two periods was specified independently of the data. In the present analysis, the break was instead selected by scanning the record using the Pettitt test. Therefore, the complete detect-then-test procedure was calibrated rather than correcting only the final epoch comparison.

Let the vector of rainfall indices in year (t) be represented by

$$\mathbf{x}_t = \left(x_{1,t}, x_{2,t}, \ldots, x_{m,t}\right)^{\mathsf{T}} \quad (8)$$

where m = 10 represents the ten rainfall indices. Change-point-free surrogate series were generated using a multivariate first-order autoregressive [AR(1)] process,

$$\mathbf{x}_t = \mathbf{A}\,\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t \quad (9)$$

where $(\mathbf{A})$ represents the autoregressive coefficient matrix and the innovation vector satisfies

$$\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon) \quad (10)$$

This formulation preserves the lag-1 temporal persistence and contemporaneous cross-correlation among the rainfall indices while containing neither an imposed trend nor a true change point (Theiler et al., 1992; Schreiber and Schmitz, 2000).

Each surrogate realization was subjected to exactly the same analytical procedure as the observed dataset. Specifically, the Pettitt test was first applied to the surrogate SDII series, the detected break was then used to define the two surrogate epochs, and the Wilcoxon rank-sum test was subsequently applied to each rainfall index. The resulting surrogate distribution was used to obtain a calibrated probability for each observed epoch contrast.

If T_b represents the observed test statistic for rainfall index b, and $\left(T_b^{(r)}\right)$ represents the corresponding statistic from surrogate realization r, the calibrated probability was estimated empirically as

$$p_{\mathrm{corr},b} = \frac{1+\sum_{r=1}^{R} I\left(\left|T_b^{(r)}\right| \geq |T_b|\right)}{R+1} \tag{11}$$

where R is the number of surrogate realizations and (I(\cdot)) is an indicator function.

Two null specifications were considered. Under the primary null, residuals were calculated after removing a step associated with the observed break. This prevents the detected signal itself from artificially increasing the estimated persistence. Under the sensitivity null, the raw series were mean-centred without removing the observed step, producing a more conservative persistence structure. Corrected probabilities were reported under the primary null, with the sensitivity-null results provided as an additional robustness assessment.

2.7 Field significance and spatial dependence

The 29 grid cells exhibit substantial spatial dependence, meaning that they cannot be treated as 29 independent observations (Livezey and Chen, 1983; Wilks, 2006). Consequently, the number of significant cells alone is insufficient to establish regional spatial coherence. The effective number of independent spatial samples was therefore estimated from the mean inter-cell correlation.

For N spatial cells and mean pairwise correlation $(\overline{\rho})$, the effective spatial sample size was calculated as

$$N_{\mathrm{eff}} = \frac{N}{1+(N-1)\bar{\rho}} \tag{12}$$

Field significance was subsequently assessed using the covariance-preserving surrogate framework described in Section 2.6, extended to the 29-cell field. For this spatial analysis, the innovation covariance was estimated from cell-level residuals so that the surrogate realizations preserved both temporal persistence and contemporaneous inter-cell covariance. For every surrogate realization, cell-level trends were calculated and counted using exactly the procedure applied to the observations: a cell was counted when its uncorrected Mann–Kendall trend reached $p \leq 0.05$, with no serial-

correlation correction and no false-discovery-rate control. The local statistic is deliberately left uncorrected because the surrogate null already reproduces both the serial persistence and the inter-cell covariance of the observed field; correcting the statistic as well would apply the same adjustment twice. These counts are therefore not the FDR-controlled counts reported in §3.2 and Supplementary Table S3.

Let (M_{obs})denote the observed number of such cells and (M_r) the corresponding number obtained from surrogate realization r. Field significance was evaluated from the empirical null distribution,

$$p_{\text{field}} = \frac{1+\sum_{r=1}^{R} I(M_r \geq M_{\text{obs}})}{R+1} \tag{13}$$

Because the innovation covariance matrix was estimated from only 75 residual vectors, an unregularized covariance estimate could produce an overly broad or unstable null distribution. Linear covariance shrinkage (Ledoit and Wolf, 2004) was therefore incorporated as an additional sensitivity analysis. The shrinkage covariance estimator was represented as

$$\widehat{\boldsymbol{\Sigma}}_{\lambda} = (1-\lambda)\widehat{\boldsymbol{\Sigma}} + \lambda \mathbf{T} \tag{14}$$

where $\left(\widehat{\boldsymbol{\Sigma}}\right)$ is the empirical covariance matrix, $(\mathbf{T})$ is the shrinkage target, and (λ) is the shrinkage intensity.

Three shrinkage levels, $(\lambda = 0)$, 0.2 and 0.5, were combined with the two null specifications, resulting in six null–shrinkage combinations. An index was classified as spatially coherent only when it achieved field significance under all six combinations. This stringent criterion was adopted to reduce the possibility that spatial coherence would be inferred from the most favourable covariance specification.

2.8 Seasonality and wet-season timing

The change in annual rainfall between the two epochs was decomposed into contributions from the four seasons: pre-monsoon (MAM), monsoon (JJAS), post-monsoon (OND), and winter (JF). For season s, the change in mean rainfall was calculated as

$$\Delta P_s = \bar{P}_{s,\text{post}} - \bar{P}_{s,\text{pre}} \tag{15}$$

where $\bar{P}_{s,\text{pre}}$ and $\bar{P}_{s,\text{post}}$ represent the mean rainfall during the pre-break and post-break epochs, respectively. The contribution of each season to the total annual rainfall change was then expressed as

$$C_s = \frac{\Delta P_s}{\Delta P_{\text{annual}}} \times 100 \tag{16}$$

where (C_s) denotes the percentage contribution of season s to the total annual rainfall change, (ΔP_s) is the change in rainfall for season (s), and (ΔP_{annual}) is the corresponding change in annual rainfall.

Because the seasonal epoch contrasts were based on the same searched 2006 break, their statistical significance was evaluated using the calibrated detect-then-test procedure described in Section 2.6. This approach accounts for the uncertainty introduced by selecting the change point from the observed rainfall record before evaluating the differences between the two epochs.

Wet-season timing was characterized using the calendar day on which 10 %, 50 %, and 90 % of the annual rainfall had accumulated. For year y, the date corresponding to a cumulative rainfall fraction q was defined as

$$D_{q,y} = \min\left\{d: \frac{P_y(d)}{P_y^{\text{annual}}} \geq q\right\}, \qquad q \in \{0.10, 0.50, 0.90\} \tag{17}$$

where $(P_y(d))$ is the cumulative rainfall from the beginning of year (y) through calendar day (d), and $\left(P_y^{annual}\right)$ is the total rainfall accumulated during that year.

The duration of the principal wet-season accumulation period was subsequently calculated as the interval between the 10 % and 90 % accumulation dates:

$$L_y = D_{0.90,y} - D_{0.10,y} \tag{18}$$

Trends in the wet-season timing metrics were assessed using the Mann–Kendall test and Theil–Sen slope estimator described in Sections 2.4, while potential change points were identified using the Pettitt test described in Section 2.5. These timing analyses did not require the additional calibration of Section 2.6 because the Mann–Kendall trend analysis does not involve selection of a change point, whereas the Pettitt test provides its own significance assessment for the change-point search. Differences in timing metrics between the two epochs were therefore reported descriptively.

2.9 Regional frequency analysis

Annual maximum daily rainfall was analysed using the L-moment index-flood framework (Hosking, 1990; Hosking and Wallis, 1997). Sample L-moment ratios, including the L-coefficient of variation (L-CV) and L-skewness, were calculated for the individual grid cells. Regional homogeneity was assessed using the L-moment heterogeneity statistic, with 500 simulations from a kappa distribution used to establish the expected distribution under regional homogeneity.

Candidate probability distributions were ranked using the relevant goodness-of-fit statistic. When the complete 29-cell domain was found to be heterogeneous, the grid cells were grouped into spatially coherent sub-regions and the frequency analysis was repeated within each sub-region.

Regional rainfall quantiles were estimated under the index-flood assumption, whereby site-specific quantiles are represented by the product of an index flood and a common regional growth factor,

$$Q_i(T) = \mu_i \, G(T) \tag{19}$$

where $\left(Q_i(T)\right)$ is the rainfall quantile at site (i)corresponding to return period (T), $(\mu_i)$is the site-specific index flood, and $\left(G(T)\right)$ is the regional growth curve.

At-site annual maximum rainfall was additionally modelled using generalized extreme-value (GEV) distributions (Coles, 2001). The stationary GEV distribution was represented by

$$F(x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{20}$$

subject to

$$1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0 \tag{21}$$

where (μ), $(\sigma > 0)$, and (ξ)denote the location, scale, and shape parameters, respectively.

Stationary GEV models were compared with models incorporating a linear trend in the location parameter, a temporal trend in the logarithm of the scale parameter, and covariates representing Niño 3.4 and the Indian Ocean Dipole. Competing models were evaluated using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

2.10 Coherence with climate modes

Potential coherence between regional annual maximum rainfall and large-scale climate variability was assessed using the Niño 3.4 index and Indian Ocean Dipole (IOD) index. Niño 3.4 was considered for the December–February (DJF) season, whereas the IOD was considered for the September–November (SON) season (Saji et al., 1999).

Cross-wavelet analysis was used to investigate the time–frequency relationship between the regional annual maximum rainfall series and the climate-mode indices (Torrence and Compo, 1998; Grinsted et al., 2004). The cross-wavelet transform was calculated as

$$W_{xy}(s,t) = W_x(s,t)\, W_y^*(s,t) \tag{22}$$

where $\left(W_x(s,t)\right)$and $\left(W_y(s,t)\right)$are the continuous wavelet transforms of the two time series, (s) denotes wavelet scale, (t) represents time, and the superscript $(*)$ denotes complex conjugation.

Wavelet coherence was evaluated using the smoothed cross-wavelet power relative to the individual wavelet powers,

$$R^2(s,t) = \frac{\left|S\left[s^{-1}W_{xy}(s,t)\right]\right|^2}{S\left[s^{-1}|W_x(s,t)|^2\right] S\left[s^{-1}|W_y(s,t)|^2\right]} \quad (23)$$

where (S) denotes the smoothing operator.

Because neighbouring wavelet coefficients are not statistically independent, pointwise significance was not interpreted in isolation. Instead, significance was evaluated using the area of the largest contiguous patch of pointwise-significant coherence. The observed patch area was compared against its corresponding null distribution, following the area-wise approach of Maraun et al. (2007) and Schulte (2016).

2.11 Cartographic representation

Cell-level rainfall results were represented using the original $(0.25°)$ grid cells clipped to the administrative boundary of Mizoram. The results were displayed as discrete grid cells rather than interpolated surfaces.

Further spatial interpolation was avoided because the source rainfall product already represents an interpolation of the IMD gauge network (Pai et al., 2014). Additional interpolation could introduce further smoothing, particularly for rainfall extremes in a relatively sparsely gauged mountainous region, where interpolation is known to damp high percentiles (Hofstra et al., 2010); grid resolution should not be equated with information content (Daly, 2006). Kriging was also not applied because only 29 spatial units were available, far fewer than reliable variogram estimation requires (Oliver and Webster, 2014).

The use of the original grid cells also ensured direct correspondence between the statistical unit used in the analysis and the spatial unit represented in the maps.

For cell-level results, statistical significance was determined after Benjamini–Hochberg FDR correction. Cells satisfying

$$q \leq 0.10 \quad (24)$$

were identified as statistically significant, rather than using unadjusted $p < 0.05$. This approach follows the recommendation of Wilks (2016) for controlling false discoveries in spatial climate analyses. Supplementary figures additionally provide the alternative AR6-style significance representation, in which hatching marks cells that are not robust (Tebaldi et al., 2011; Gutiérrez et al., 2021), while the main figures retain the conventional overlay because denser texture is read as greater confidence irrespective of the legend (Johannsen et al., 2018).

2.12 Reproducibility

All analyses were conducted in R 4.5.1 using trend and modifiedmk for trend and change-point analysis, lmom and lmomRFA for regional frequency analysis, extRemes for GEV modelling, biwavelet for wavelet coherence, and sf for spatial processing. Random seeds were recorded with all outputs. The surrogate procedure in §2.6 was independently re-implemented with a different random-number generator; corrected p-values agreed within Monte Carlo error (maximum absolute difference = 0.038; standard error ≈0.01; Supplementary Table S17). The reconciled results workbook and analysis code are provided as Data S1 and permanently archived on Zenodo (Khiangte, 2026).

3. Results

3.1 Quality control, homogeneity and spatial redundancy

Of the 29 grid cells, 16 were classified as useful, 9 as doubtful and 4 as suspect (Table 1; Figure 1). Nineteen cells showed an FDR-significant break in annual rainfall, but break years varied widely (median 1979; IQR 30 years), indicating no single regional breakpoint (Figure 2; Supplementary Table S2). Breaks were also organised more strongly by longitude than latitude: the western cells had a median break of 2004, whereas the eastern cells had a median of 1976 and contained most of the suspect cells. The eastern margin was therefore treated cautiously, and no substantive conclusion depends on it.

Table 1. Quality control and homogeneity of the 29 in-domain grid cells.

Longitude band	Cells	Useful	Doubtful	Suspect	Median break year PRCPTOT
West (92.50 °E)	6	5	0	1	2004
Central (92.75–93.00 °E)	19	11	8	0	1973
East (93.25 °E)	4	0	1	3	1976
Domain	**29**	**16**	**9**	**4**	**1979**

Classification follows Wijngaard et al. (2003): *useful* = 0–1 of four homogeneity tests reject at the 1 % level, *doubtful* = 2, *suspect* = 3–4. Nineteen of 29 cells carry an FDR-significant break (IQR of break year = 30 yr).

Mean inter-cell correlation was 0.70, reducing the 29 nominal cells to an effective sample of approximately 1.9 for PRCPTOT and 2.9 for Rx1day (Eq. 12; Supplementary Table S1). This strong spatial redundancy was accounted for in the field-significance analysis.

3.2 Monotonic trends, 1950–2025

None of the ten indices showed a significant monotonic trend in the regional series under the primary Hamed–Rao correction (Table 2). The smallest p-value was for SDII ($p = 0.184$), and all FDR-adjusted values were 0.940. SDII was significant under Yue–Wang ($p = 0.023$) and trend-free pre-whitening ($p = 0.003$), with a Sen slope of −0.051 mm day^{-1} yr^{-1} (95% CI: −0.093 to −0.009). Because this result depends on the treatment of serial correlation, we do not regard the regional SDII trend as established.

Table 2. Mann–Kendall trend statistics for the 29-cell area mean, 1950–2025.

Index	*Z*	Sen slope (yr^{-1})	95 % CI	*p*HR	*p*YW	*p*TFPW	*q*FDR
PRCPTOT	-0.076	-0.3344	-6.071 to 6.370	0.940	0.893	1.000	0.940
Rx1day	-0.458	-0.0767	-0.394 to 0.238	0.647	0.458	0.756	0.940
Rx5day	-0.807	-0.2321	-0.986 to 0.503	0.420	0.327	0.784	0.940
R95p	-0.996	-1.4874	-4.178 to 1.150	0.319	0.213	0.297	0.940
R99p	-0.382	-0.2571	-1.677 to 0.797	0.702	0.538	0.763	0.940
SDII	-1.327	-0.0513	-0.093 to -0.009	0.184	0.023	0.003	0.940
R10mm	-0.151	-0.0228	-0.189 to 0.165	0.880	0.759	0.927	0.940
R20mm	-0.253	-0.0228	-0.148 to 0.111	0.800	0.667	0.862	0.940
CDD	-0.076	-0.0057	-0.147 to 0.144	0.939	0.916	0.681	0.940
CWD	0.836	0.0429	-0.034 to 0.133	0.404	0.169	0.280	0.940

HR = Hamed–Rao, YW = Yue–Wang, TFPW = trend-free pre-whitening. *q*FDR applies Benjamini–Hochberg to *p*HR across the ten indices.

At the cell level, CWD showed the clearest spatial pattern, with trends surviving Benjamini–Hochberg control at $q \leq 0.10$ in 20 of 29 cells; R10mm was significant in 2 cells, and PRCPTOT, R20mm and R95p in one each. The remaining indices, SDII among them, had no cell-level trend surviving FDR control (Supplementary Table S3); the larger counts in Table 5 are uncorrected local rejections and are not comparable with these. For R99p, the Theil–Sen slope was zero in all cells because the series contains many zero values; this reflects estimator behaviour rather than evidence of stationarity and is therefore not interpreted as a trend result.

3.3 The 2006 intensity break and its calibration

The Pettitt test identified a change in regional SDII in 2006 ($p = 0.0011$), with mean SDII declining from 19.5 mm day^{-1} during 1950–2006 to 14.7 mm day^{-1} during 2007–2025 (Figure 3). Before calibration, eight of the ten indices showed $p < 0.05$ in the epoch comparison, giving the appearance of widespread change (Table 3).

This apparent significance was strongly inflated by the break-selection procedure. When the same detect-then-test analysis was applied to 2000 AR(1) surrogates containing no change point, SDII produced $p < 0.05$ in 83.1% of realisations under the primary null and 93.1% under the sensitivity null, corresponding to 17- and 19-fold inflation. At $p < 0.001$, rejection rates were 26.5% and 47.8%, respectively, or 265- and 478-fold inflation (Table 4).

Table 4. Empirical rejection rate of the uncorrected detect-then-test procedure applied to change-point-free AR(1) surrogates.

Nominal α	Null	Empirical rejection rate	Inflation
0.05	Signal-adjusted	0.831	17×
0.01	Signal-adjusted	0.552	55×
0.001	Signal-adjusted	0.265	265×
0.05	Raw-fit	0.931	19×
0.01	Raw-fit	0.740	74×
0.001	Raw-fit	0.478	478×

Rates are for SDII, the index used to locate the break; 2000 surrogates.

Conventional p-values evaluated at a searched break therefore overstate significance in this record by a wide margin.

3.4 Epoch differences after correction

After calibration, only two of the ten indices remained significant (Table 3). CWD increased by 44.9%, from 22.3 to 32.3 days, with corrected $p = 0.003$ under the primary null and $p = 0.008$ under the sensitivity null. SDII declined by 24.7%, from 19.5 to 14.7 mm day^{-1}, but was significant only under the primary null ($p = 0.037$; sensitivity $p = 0.128$). We therefore treat the SDII decline as suggestive rather than established. Detailed cell-level epoch contrasts for the corrected split are provided in Supplementary Table S4.

The remaining eight indices showed observed differences but did not survive calibration. R99p declined by 46.5%, R95p by 37.0%, Rx1day by 15.1% and PRCPTOT by 9.5%, but their corrected p-values were ≥0.085. These changes are therefore reported descriptively rather than as evidence of significant shifts.

Table 3. Epoch contrast between 1950–2006 (n = 57) and 2007–2025 (n = 19), before and after correction for the break search.

Index	Pre	Post	Δ (%)	*p*unc	*p*corr primary	*p*corr sensitivity	Inference
CWD	22.32	32.33	+44.9	0.001	0.003	0.008	Significant (both nulls)
SDII	19.48	14.66	-24.7	<0.001	0.037	0.128	Significant (primary only)

Index	Pre	Post	Δ (%)	*p*unc	*p*corr primary	*p*corr sensitivity	Inference
R95p	564.91	355.82	-37.0	0.001	0.085	0.134	Descriptive
CDD	54.08	59.61	+10.2	0.214	0.196	0.200	Descriptive
R20mm	41.24	35.54	-13.8	0.025	0.243	0.270	Descriptive
Rx1day	113.79	96.61	-15.1	0.043	0.264	0.299	Descriptive
R10mm	75.11	71.58	-4.7	0.108	0.269	0.275	Descriptive
R99p	181.73	97.27	-46.5	0.032	0.277	0.319	Descriptive
PRCPTOT	2439.64	2207.24	-9.5	0.036	0.293	0.315	Descriptive
Rx5day	251.19	221.23	-11.9	0.045	0.403	0.419	Descriptive

Units: PRCPTOT, Rx1day, Rx5day, R95p, R99p in mm; SDII in mm day^{-1}; R10mm, R20mm, CDD, CWD in days. *p*unc is the two-sided Wilcoxon rank-sum *p*-value.

Restricting the analysis to the 16 homogeneous cells produced similar results: CWD increased by 38.7% (p = 0.018) and SDII declined by 24.9% (p = 0.038; Supplementary Table S11). The corresponding field-significance results for the homogeneous subset are provided in Supplementary Table S10.

3.5 Spatial coherence and the north–south contrast

CWD was the only index with a spatially coherent signal (Table 5). Its 19 cells with a nominally significant trend exceeded the 95th percentile of the covariance-preserving null under all six null–shrinkage combinations (pfield = 0.000–0.012). SDII had 12 such cells but was not field-significant under any combination (pfield = 0.068–0.146). The eight nominally significant R10mm cells likewise remained within the range expected from correlated noise.

Table 5. Field significance of cell-level trends, evaluated across all six null × shrinkage combinations.

Index	Significant cells (of 29)	*p*field range	Spatially coherent
CWD	19	0.000 – 0.012	**Yes**
SDII	12	0.068 – 0.146	No
R10mm	8	0.096 – 0.202	No
R95p	4	0.414 – 0.534	No
PRCPTOT	3	0.446 – 0.744	No
R20mm	1	0.620 – 0.968	No
CDD	0	1.000 – 1.000	No
R99p	0	1.000 – 1.000	No
Rx1day	0	1.000 – 1.000	No
Rx5day	0	1.000 – 1.000	No

Nulls: signal-adjusted and raw-fit AR(1); covariance shrinkage 0, 0.2, 0.5. Counts are cells whose uncorrected Mann–Kendall trend reaches p ≤ 0.05; the local statistic carries no serial-correlation

correction and no FDR control, because the surrogate null accounts for both (§2.7). They therefore differ from the FDR-controlled counts in §3.2 and Supplementary Table S3.

The CWD pattern showed a clear north–south contrast (Figures 4 and 5; Supplementary Figures S1 and S2). All 11 cells south of 23.00°N showed lengthening wet spells, with a median trend of +8.9% decade^{-1}, whereas only 5 of 18 cells north of 23.25°N were positive, with a median trend of −3.6% decade^{-1}. PRCPTOT showed no comparable spatial gradient, with median trends of −0.5% decade^{-1} in the north and +0.1% decade^{-1} in the south and only one FDR-significant cell. The spatial change therefore reflects a redistribution in rainfall persistence rather than a corresponding change in rainfall amount.

3.6 Seasonal budget and monsoon delivery

Annual rainfall declined by 232 mm (9.5%) between the two epochs (Table 6). The largest contribution came from the pre-monsoon season, which accounted for 62% of the decline (−144 mm; −25.2%), followed by the monsoon (−47 mm; −2.9%), post-monsoon (−22 mm; −9.9%) and winter (−19 mm; −45.5%). However, none of the seasonal contrasts remained significant after correction (pre-monsoon p = 0.058; winter p = 0.069; monsoon p = 0.629; post-monsoon p = 0.637). The annual decline should therefore be treated as descriptive rather than as evidence of a significant reduction in rainfall supply.

Table 6. Seasonal decomposition of the change in annual total between epochs.

Season	Pre (mm)	Post (mm)	Δ (mm)	Δ (%)	Share of change (%)	*p*unc	*p*corr
Pre-monsoon (MAM)	569.5	425.8	-143.7	-25.2	61.8	0.010	0.058
Monsoon (JJAS)	1602.0	1554.8	-47.1	-2.9	20.3	0.230	0.629
Post-monsoon (OND)	225.6	203.4	-22.2	-9.9	9.6	0.606	0.637
Winter (JF)	42.5	23.2	-19.3	-45.5	8.3	0.049	0.069
Annual	**2439.6**	**2207.2**	**-232.4**	**-9.5**	**100.0**	—	—

The seasonal decomposition is shown in Figure 6. The seasonal distribution changed modestly, with the monsoon contribution rising from 65.7% to 70.4% of annual rainfall and the pre-monsoon contribution falling from 23.3% to 19.3%. Cell-level seasonal trend statistics are provided in Supplementary Table S5.

Within the monsoon season, rainfall amount changed little (−2.9%; calibrated p = 0.629). Wet-day frequency increased by 17.5%, CWD by 48.1%, and per-wet-day monsoon intensity declined by 16.5%. These component changes are reported descriptively because corresponding search-calibrated probabilities were not separately estimated. Annual wet days also increased by 11.0% (p = 0.004). Thus, the monsoon is characterised by more frequent and persistent wet periods, with less rainfall per wet day despite little change in seasonal rainfall total. This pattern is consistent with the CWD result

and provides the clearest evidence of a change in rainfall delivery rather than seasonal rainfall amount.

3.7 Wet-season timing

Wet-season onset, defined by the date at which 10% of annual rainfall had accumulated, became later by 2.04 days decade^{-1} ($p = 0.019$; $q = 0.067$; Table 7; Figure 7). The interval between the 10% and 90% accumulation dates shortened by 2.06 days decade^{-1} ($p = 0.034$; $q = 0.067$). Both metrics showed Pettitt breaks in 2005 ($p = 0.0067$ and 0.0118, respectively).

The 50% accumulation date showed a weaker later shift (+0.89 days decade^{-1}; $p = 0.158$), while the withdrawal date showed no trend ($p = 0.737$). The changes therefore indicate a later onset and shorter wet-season accumulation period, with no clear change in the termination date.

Table 7. Wet-season timing.

Metric	Sen slope (d decade^{-1})	*p*trend	*q*FDR	Pettitt break	*p*Pettitt	Epoch shift (d)†
doy_{10} (onset)	+2.04	0.019	0.067	2005	0.0067	+14.9
doy_{50} (centroid)	+0.89	0.158	0.210	2006	0.0446	+10.2
doy_{90} (withdrawal)	-0.18	0.737	0.737	1975	0.3867	+0.5
Span (doy_{10}–doy_{90})	-2.06	0.034	0.067	2005	0.0118	-14.4

† Descriptive; not tested. Trend and Pettitt *p*-values require no search correction.

3.8 Extreme-value behaviour

The full 29-cell domain was heterogeneous for regional frequency analysis (Table 8; Figure 8; Supplementary Table S12). Spatial clustering produced two sub-regions of 16 and 13 cells; the first was acceptably homogeneous and the second possibly heterogeneous, with no discordant sites. Among the candidate distributions, the generalized normal distribution was accepted, whereas the GEV, generalized logistic, Pearson type III and generalized Pareto distributions were rejected. The generalized normal and GEV growth curves differed by less than 3% up to the 100-year return period.

Table 8. Regional frequency analysis of annual maximum daily rainfall.

Quantity	Whole domain	Sub-region 1 ($n = 16$)	Sub-region 2 ($n = 13$)
Heterogeneity H_1	6.03 (definitely heterogeneous)	-0.37 (acceptably homogeneous)	1.81 (possibly heterogeneous)
GoF *Z*: GNO / GEV	1.05 / 2.75	—	—
Max discordancy *D*	2.98 (none discordant)	—	—
50-yr return level, GNO (mm)	—	195 (157–219)	292 (231–387)

GEV is rejected at $|Z| > 1.64$; GNO accepted. Return levels are site medians with the across-site range in parentheses.

The southern and south-western sub-region had the larger index flood and steeper growth curve, indicating greater extreme-rainfall magnitude and a heavier regional upper tail. Sub-region membership and site-level L-moment ratios are reported in Supplementary Table S15. Median 50-year return levels were 195 mm (range 157–219 mm) and 292 mm (231–387 mm) for the two sub-regions, respectively.

For the regional annual maximum series, a stationary GEV provided an adequate fit (K–S D = 0.071, p = 0.816; Figure 9) and gave a 50-year return level of 220 mm (95% CI: 172–313 mm). Adding a linear trend in location did not improve the model (ΔAIC = +1.6 relative to the stationary model), while ocean-covariate models provided only weak AIC support and were not preferred by BIC (Supplementary Table S13). We therefore found no compelling evidence of non-stationarity in annual maximum daily rainfall over the study domain.

3.9 Coherence with climate modes

Neither ENSO nor the Indian Ocean Dipole showed area-wise significant coherence with the regional annual maximum rainfall series (Supplementary Table S14). For ENSO, 8.9% of the wavelet plane was pointwise significant, but the largest contiguous patch (112 units) was below the 95th-percentile null value of 196. For the dipole mode, only 1.0% was pointwise significant and the largest patch (20 units) was also below its null threshold of 191. Thus, neither mode showed evidence of coherent association with annual maximum rainfall.

3.10 Sensitivity analyses

Wet-day threshold. The increase in wet-day frequency was concentrated at light rainfall: days receiving 0–1 mm increased by 0.87 days decade^{-1} (p = 0.007), a rise of 25.9 % between epochs (p < 0.001; Supplementary Table S7). Increasing the wet-day threshold from 1 to 2.5 and 5 mm reduced the epoch change in wet-day count from +11.0% to +7.9% and +3.8%, neither of which was significant (p = 0.164 and p = 0.914; Supplementary Table S6). In contrast, the CWD increase declined from +44.9% to +30.1% and +27.5% but remained significant at all thresholds (p = 0.001, 0.016 and 0.007), while the SDII decline persisted (−24.7%, −21.6% and −18.1%; all p < 0.001).

The concentration of the wet-day signal at low rainfall amounts raises the possibility of changes related to gauge density or interpolation. However, CWD remained evident at higher thresholds and was preserved in the homogeneous-cell subset. We therefore place greater weight on the CWD signal than on the increase in wet-day frequency itself.

Record length. Truncating the record at 2019 removed the nominal significance of the PRCPTOT, R99p, R10mm and R20mm epoch differences, while CWD (+56.7%), SDII (−19.4%) and R95p (−27.0%) remained nominally significant (Supplementary Table S8). Several descriptive epoch differences therefore depend partly on the final six years of the record.

Intensity–persistence relationship. A partial correlation between SDII and CWD, controlling only for annual rainfall, was negative ($r = -0.553$, $p < 0.001$), but this relationship is induced by the definition of SDII as annual rainfall divided by wet-day count. After controlling for both annual rainfall and wet-day count, the correlation between the 95th percentile of wet-day rainfall and CWD was weak and non-significant ($r = -0.125$, 95% CI: −0.341 to 0.103; $p = 0.281$; Supplementary Table S9). Thus, the data do not support an independent physical trade-off between rainfall intensity and wet-spell persistence.

3.11 Supplementary material

Supplementary Information provides detailed cell-level statistics, sensitivity analyses, field-significance assessments, regional frequency-analysis diagnostics, climate-mode coherence results, and independent validation of the surrogate calibration. Supplementary Figures S1 and S2 provide the alternative IPCC AR6-style representations of Figures 4 and 5. Data S1 provides the reconciled results workbook and analysis code.

4. Discussion

4.1 The cost of testing at a searched break

The central methodological result is that conventional inference becomes severely miscalibrated when a change point is selected from the same record and then tested as if it were fixed in advance. In our surrogate experiments, the nominal 5% test rejected in 83–93% of change-point-free series, while the nominal 0.1% test rejected in 27–48%. Thus, the eight apparently significant epoch contrasts obtained from the uncorrected analysis cannot be interpreted as independent findings.

The magnitude of this inflation depends on record length, temporal persistence and the dependence between the index used to locate the break and those subsequently tested. The important point is therefore not the exact inflation factor for Mizoram, but that the composite detect-then-test procedure must be calibrated for the specific record and covariance structure being analysed. This also explains why the two null models matter: SDII is significant under the primary null but not under the more conservative sensitivity null.

Our results support a practical recommendation for regional rainfall studies using searched change points: calibrate the complete detection-and-testing procedure rather than applying a conventional test

only after selecting the break. This provides a direct estimate of how often the analytical pipeline would produce apparently significant results when no change point is present.

4.2 Longer wet spells and where they occur

The most robust physical result is the increase in consecutive wet days. CWD increased by 44.9% across the 2006 break, remained significant under both null models, and was the only index with field significance across all six null–shrinkage combinations. The spatial pattern was equally clear: all 11 cells south of 23.00°N showed lengthening wet spells, whereas only 5 of 18 northern cells were positive. No comparable pattern occurred in rainfall amount.

The result is broadly consistent with recent regional evidence of longer precipitation spells. Sharma et al. (2025) reported increased precipitation duration during July–October across several Indian sub-regions, while Kolekar et al. (2025) found increases in high-duration rain spells over northeast India. These studies do not establish a common mechanism, but their agreement in the direction of change supports the interpretation that the CWD signal warrants further physical investigation.

The seasonal analysis provides an internally consistent picture. Monsoon rainfall totals changed little, while wet-day frequency and CWD increased and per-wet-day intensity declined. The evidence therefore points to a change in the temporal distribution of monsoon rainfall rather than a clear change in seasonal rainfall amount. Because similar patterns appear in both trend and epoch analyses, this interpretation does not depend entirely on the 2006 change point.

4.3 The intensity decline and conflicting evidence

The 24.7% decline in SDII is suggestive but not robust. It is significant under the primary null ($p = 0.037$) but not the sensitivity null ($p = 0.128$), lacks field significance, and is sensitive to the treatment of serial correlation. We therefore do not interpret it as an established regional trend.

This result also differs from recent evidence for increasing rainfall intensity over northeast India. Kolekar et al. (2025) reported increasing maximum spell intensity over 1901–2023, although their longer record and different comparison periods make the results not directly comparable. Lawmchullova et al. (2026) found increasing rainfall totals across stations in the Tlawng watershed during 1984–2023, together with declining pre-monsoon rainfall. That study differs in data source, spatial scale and metric, but its pre-monsoon result is consistent with our seasonal decomposition.

Other recent studies have reported declining extreme rainfall or no significant change across the northeast (Singh et al., 2025; Das et al., 2024), while Dhara et al. (2025) and Borah et al. (2026) identify declines in mean rainfall and changes in moisture transport. The regional evidence therefore

remains mixed, and our SDII result should be viewed as one contribution rather than a resolution of that disagreement.

4.4 What the data can and cannot support

Three limitations constrain the interpretation of the results.

First, the gauge network and gridded interpolation may influence the light-rainfall signal. The increase in wet days is concentrated below 1 mm, and the longitudinal distribution of cell-level break years also raises the possibility of network effects. This is not a remote concern: the IMD gauge network exceeded 3,750 stations during 1991–2016 but fell below 3,000 after 2016, and station density is markedly lower in the northeast than in peninsular India (Peringiyil et al., 2025). However, the CWD increase persists when the wet-day threshold is raised to 2.5 and 5 mm and remains present in the homogeneous-cell subset. We therefore place greater confidence in the CWD result than in the increase in wet-day frequency.

Second, the rainfall product itself is a source of uncertainty. IMD gridded rainfall can underestimate precipitation in mountainous and wet regions, and different gridded products can produce different trends in rainfall extremes (Goteti & Famiglietti, 2024; Paul et al., 2024). The physical interpretation of any single-product signal should therefore remain cautious. Multi-product analysis, particularly with station observations where available, would provide an important test of robustness.

Third, strong spatial correlation limits the effective sample size. With a mean inter-cell correlation of 0.70, the 29 nominal cells correspond to only about 1.9 independent series for PRCPTOT and 2.9 for Rx1day. The absence of a significant trend in annual maximum rainfall should therefore be interpreted as a result about detectability at this spatial scale, not as evidence that extremes are necessarily stationary, and it does not contradict the globally averaged intensification of annual maximum daily precipitation reported by Westra et al. (2013).

4.5 Implications for regional rainfall analysis

The methodological result is more general than the physical findings from Mizoram. Regional rainfall records are often short, serially correlated and spatially redundant, making change-point selection followed by conventional testing particularly vulnerable to false positives.

Three practices follow from our analysis. First, the complete detect-then-test procedure should be calibrated against a null model that preserves the temporal persistence and spatial covariance of the observed field. Second, spatial coherence should be evaluated against a covariance-preserving field null rather than inferred from the number of significant cells alone. Third, studies should report the effective spatial degrees of freedom alongside the nominal number of grid cells.

These steps are relatively simple to implement but materially change the interpretation of regional rainfall signals. In the present case, calibration reduced an apparently widespread set of rainfall changes to one robust change in wet-spell persistence, a less certain decline in intensity, and a redistribution in the temporal delivery of monsoon rainfall. The resulting claim is narrower, but better supported by the data.

4.6 Limitations

The calibration addresses change-point selection and dependence, but not errors in the underlying rainfall product. A change associated with the observing or interpolation system could therefore survive a surrogate test that treats the observed series as climatic. The post-break period is also relatively short (19 years), and some descriptive differences depend on the final six years of the record.

The AR(1) surrogate model may underestimate persistence if the rainfall series has longer memory, which could make corrected p-values somewhat anti-conservative. The raw-fit sensitivity null provides a more conservative check, but higher-order dependence structures were not tested.

Finally, we have not established a physical mechanism for the CWD increase. Moisture-flux divergence provides a plausible avenue for further investigation, but a formal moisture-budget analysis would be required to establish attribution.

5. Conclusions

Using 29 grid cells across Mizoram for 1950–2025, we show that change-point detection followed by conventional epoch testing can substantially overstate evidence for rainfall change. When the complete detect-then-test procedure was calibrated against change-point-free surrogates preserving temporal persistence and spatial covariance, most apparently significant epoch contrasts disappeared.

The most robust rainfall signal is the increase in consecutive wet days. CWD increased by 44.9% across the 2006 break and was the only index with field-significant spatial coherence under all null–shrinkage combinations. The increase was concentrated in the southern part of Mizoram and occurred without a corresponding spatial change in rainfall amount.

The monsoon rainfall total remained statistically unchanged, but wet-day frequency and wet-spell duration increased while rainfall intensity declined. This indicates a shift in the temporal distribution of monsoon rainfall rather than a clear change in seasonal rainfall supply. The SDII decline itself remains uncertain because its significance depends on the treatment of serial correlation and it lacks field significance.

The main implication is methodological: regional rainfall studies that search for change points should calibrate the full detection-and-testing procedure and account explicitly for spatial dependence. For Mizoram, the resulting evidence supports a narrower but more defensible conclusion—longer wet spells and altered monsoon rainfall delivery, with no robust evidence from this dataset for a change in annual maximum rainfall.

Figures

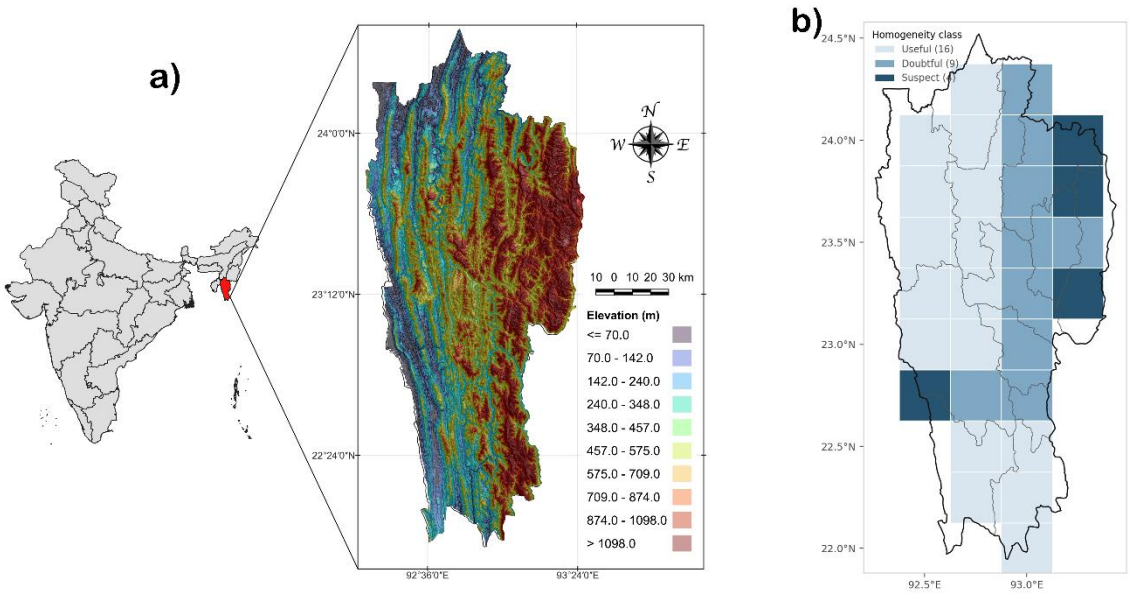
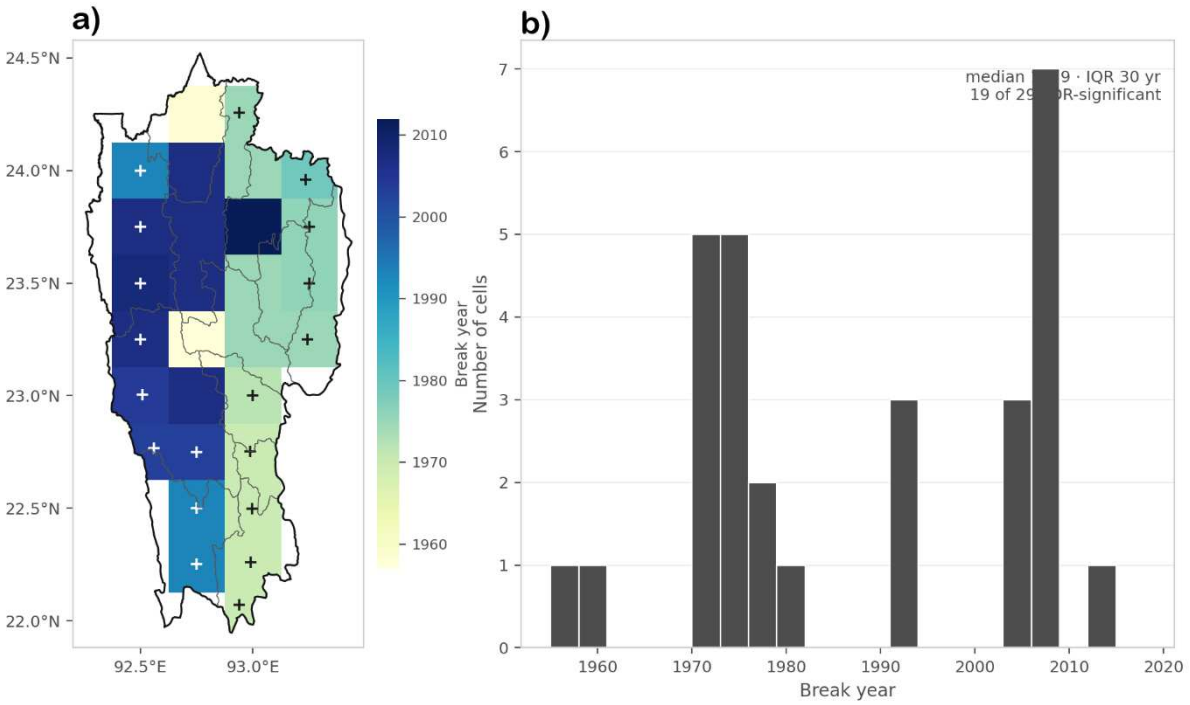
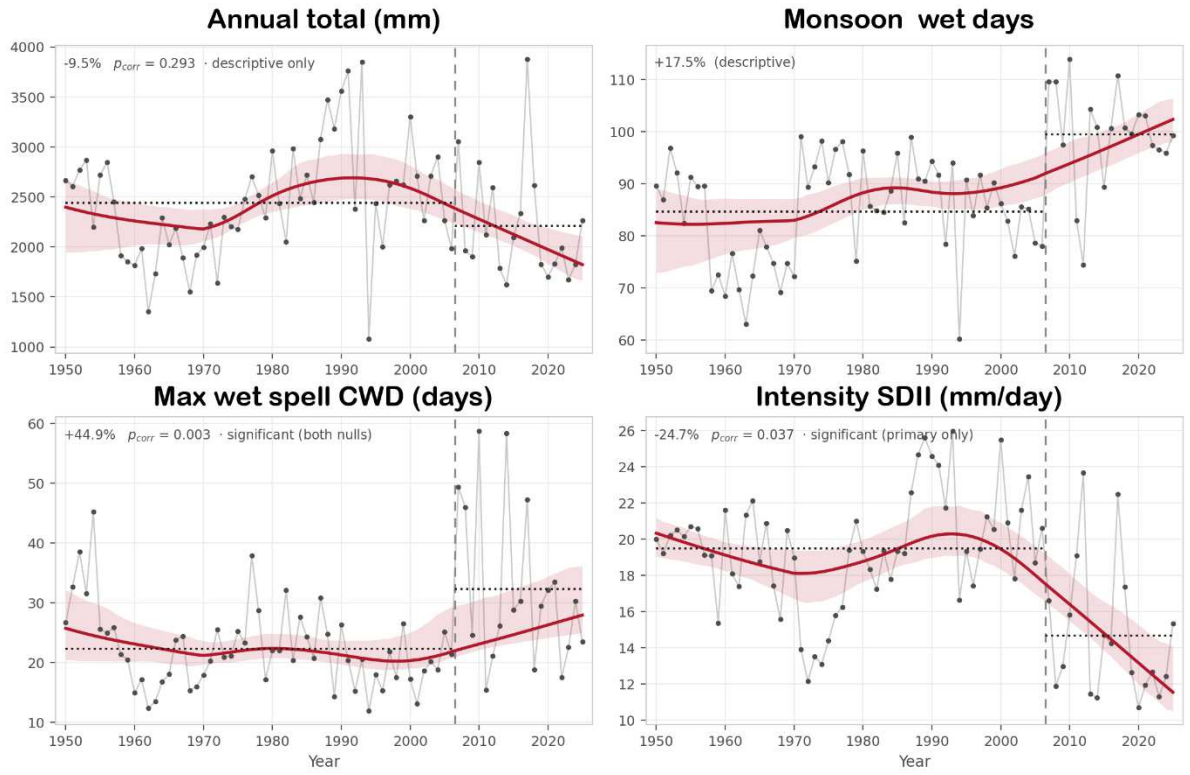


Figure 1. (a) Districts of Mizoram. (b) The 29-cell 0.25° analysis grid, shaded by homogeneity class after Wijngaard et al. (2003). Cells are drawn at full extent to show which cells enter the analysis; cell centres all fall inside the state boundary.



653

654 **Figure 2.** (a) Pettitt break year in annual total, by cell; "+" marks cells whose break survives
655 Benjamini–Hochberg control at $q \leq 0.10$. (b) Distribution of cell-level break years.



656

Figure 3. Regional (29-cell mean) series for annual total, monsoon wet days, CWD and SDII. Red line is a LOESS smoother with bootstrap band; dotted lines are epoch means; dashed vertical line marks the 2006 break, which belongs to the earlier epoch. Corrected *p*-values are annotated.

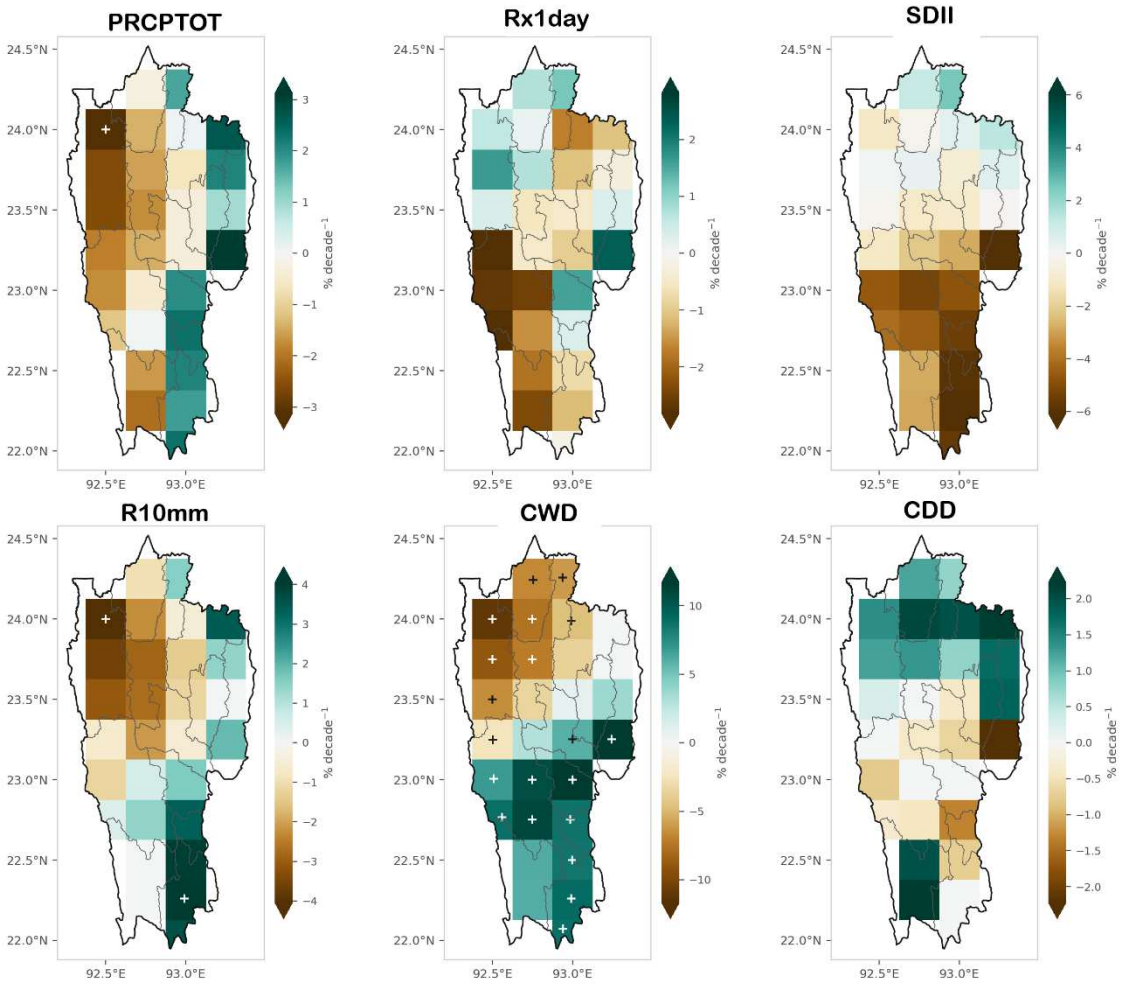
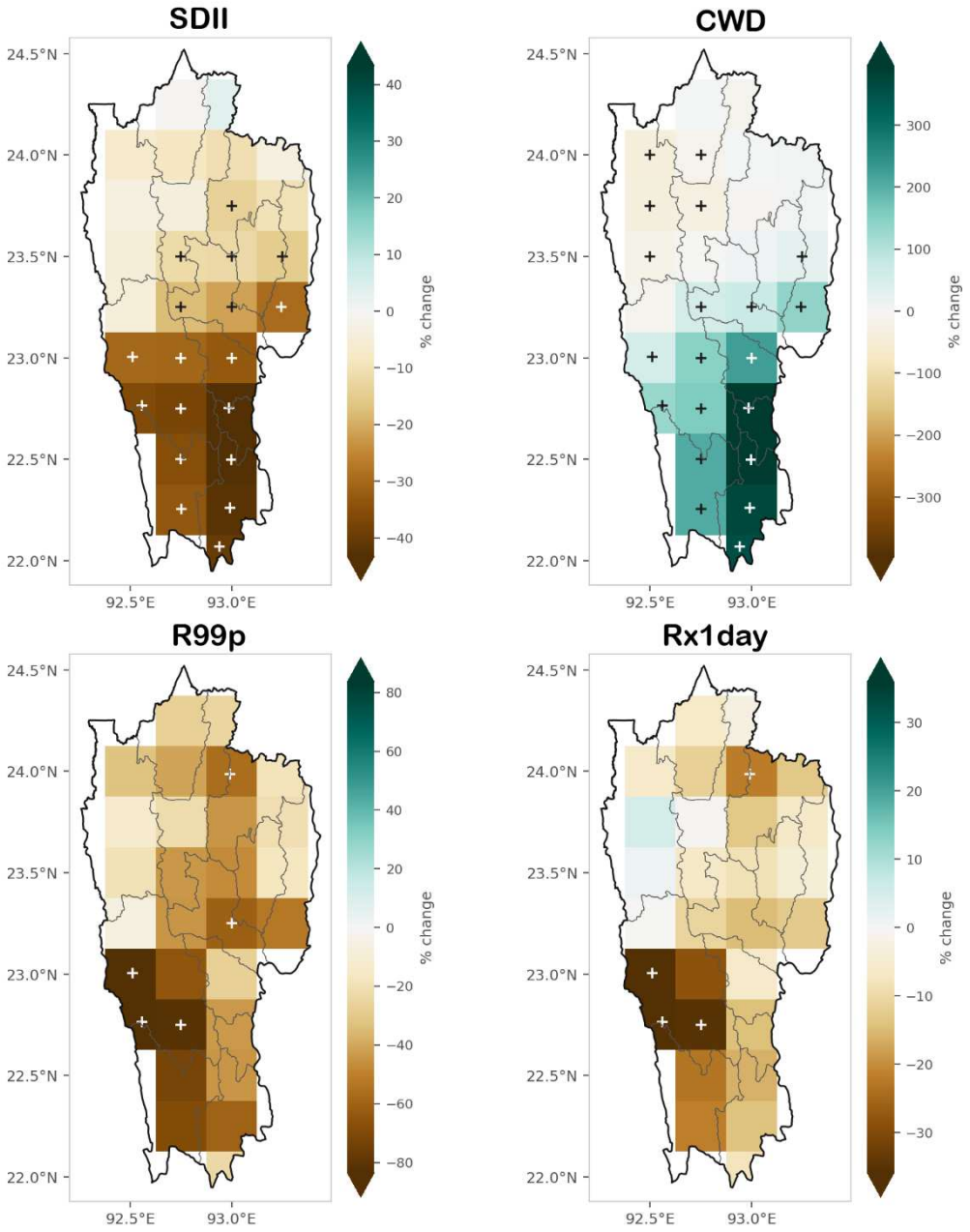
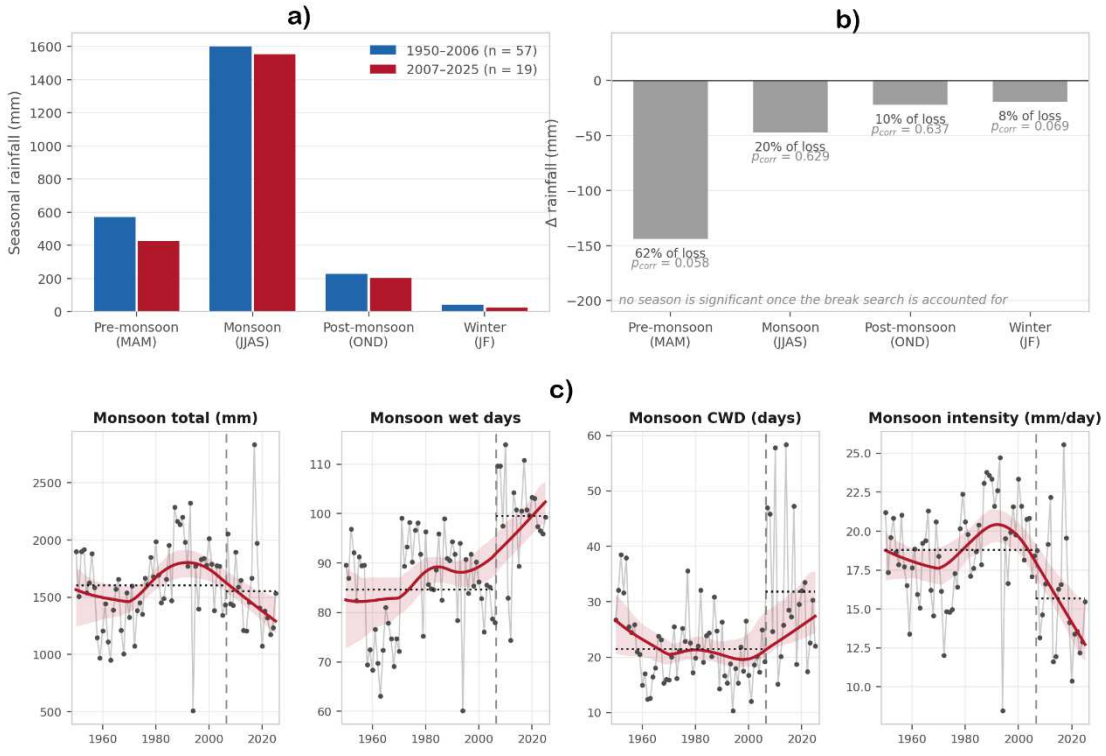


Figure 4. Theil–Sen trend, 1950–2025, by cell, as % decade^{-1}. Cells are clipped to the state boundary; thin lines are district boundaries. "+" marks cells surviving Benjamini–Hochberg control at $q \leq 0.10$, following the recommendation of Wilks (2016) that field maps mark FDR survivors rather than raw per-cell significance. Colour scales are symmetric about zero and capped at the 95th percentile of |value|; arrows on the colour bar denote clipped cells. R99p is omitted because its per-cell Sen slope is exactly zero at all 29 cells (§3.2).



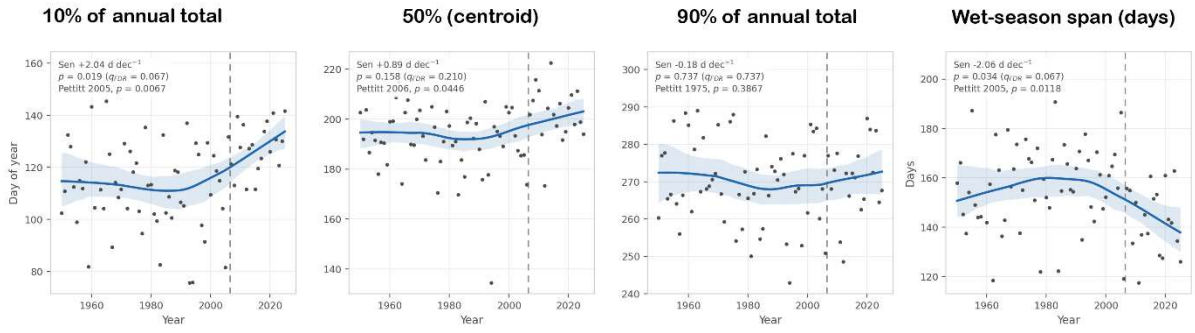
667

668 **Figure 5.** Epoch change, 1950–2006 → 2007–2025, by cell, as % change. Conventions as in Figure 4.



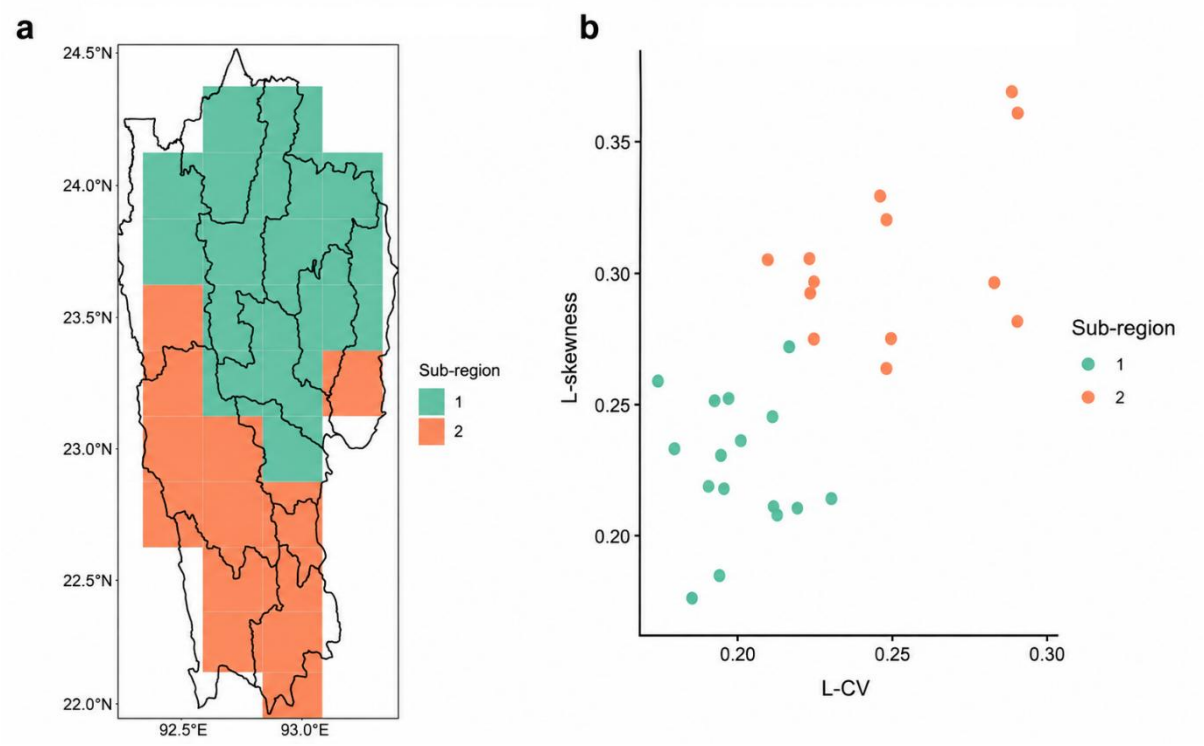
669

670 **Figure 6.** (a) Area-averaged seasonal totals by epoch. (b) Contribution of each season to the −232 mm
671 change, annotated with search-corrected *p*-values; no season reaches $p < 0.05$. (c) Monsoon-season
672 components.



673

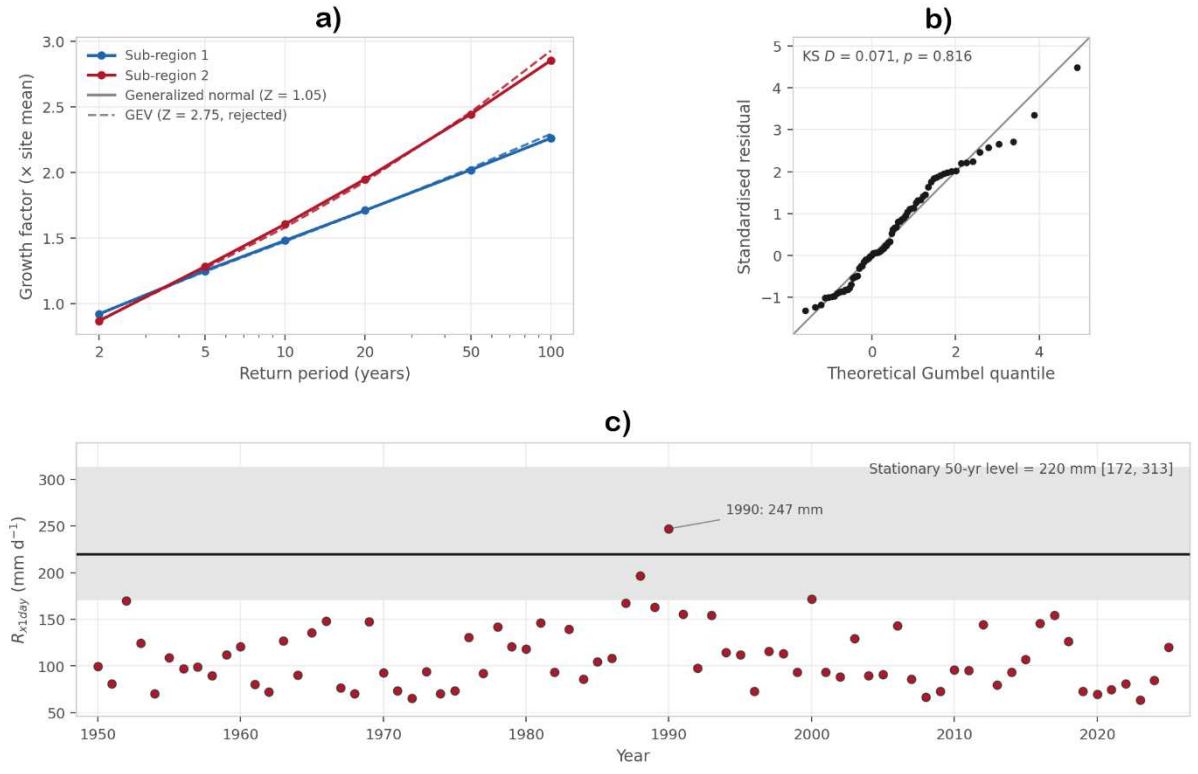
674 **Figure 7.** Wet-season timing metrics with LOESS smoother and bootstrap band.



675

676 **Figure 8.** (a) Extreme-value sub-regions from L-moment clustering. (b) Site L-moment ratios
677 underlying the split.

678



679

Figure 9. (a) Regional growth curves by sub-region, GNO and GEV. (b) Quantile–quantile plot for the stationary GEV fit. (c) 50-year return level with 95 % parametric bootstrap interval.

CRediT Author Statement

Michelle Lalhruaizeli: Conceptualization, Methodology, Data curation, Investigation, Validation, Formal analysis, Resources, Writing – original draft.

Laldinchhuaha Khiangte: Conceptualization, Methodology, Software, Formal analysis, Validation, Resources, Data curation, Visualization

B. P. Mishra: Supervision, Writing – review & editing.

Vishwambhar Prasad Sati: Supervision, Writing – review & editing.

Acknowledgements

The authors gratefully acknowledge the India Meteorological Department (IMD) for providing the long-term gridded daily rainfall dataset used in this study.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. The analysis code is openly available on GitHub and permanently archived on Zenodo (Khiangte, 2026): https://doi.org/10.5281/zenodo.22679527

References

Aghaei, F.A., Phillips, E.T. and Kantz, H. (2025) Change point detection in ERA5 ground temperature time series. *Climatic Change*, 178(11), 193. https://doi.org/10.1007/s10584-025-04021-3

706 Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M.G., et
707 al. (2006) Global observed changes in daily climate extremes of temperature and
708 precipitation. *Journal of Geophysical Research: Atmospheres*, 111, D05109.
709 https://doi.org/10.1029/2005JD006290

710 Alexandersson, H. (1986) A homogeneity test applied to precipitation data. *Journal of Climatology*,
711 6(6), 661–675. https://doi.org/10.1002/joc.3370060607

712 Amirthanathan, G.E., Bari, M.A., Woldemeskel, F.M., Tuteja, N.K. and Feikema, P.M. (2023)
713 Regional significance of historical trends and step changes in Australian streamflow.
714 *Hydrology and Earth System Sciences*, 27(1), 229–254. https://doi.org/10.5194/hess-27-229-
715 2023

716 Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful
717 approach to multiple testing. *Journal of the Royal Statistical Society: Series B*
718 *(Methodological)*, 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

719 Borah, A., Pathak, B., Sahu, P.J. and Bhuyan, P.K. (2026) Exploring moisture budget components
720 and their role on precipitation variability over North-East India and adjoining regions.
721 *Climate Dynamics*, 64(6), 277. https://doi.org/10.1007/s00382-026-08245-6

722 Buishand, T.A. (1982) Some methods for testing the homogeneity of rainfall records. *Journal of*
723 *Hydrology*, 58(1–2), 11–27. https://doi.org/10.1016/0022-1694(82)90066-X

724 Chakraborty, D., Roy, A., Singh, N.U., Saha, S., Das, S.K., Mridha, N., et al. (2025) Assessing
725 climate change impact on rainfall patterns in northeastern India and its consequences on
726 water resources and rainfed agriculture. *Earth*, 6(1), 2. https://doi.org/10.3390/earth6010002

727 Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*. London: Springer.
728 https://doi.org/10.1007/978-1-4471-3675-0

729 Cortés, J., Mahecha, M., Reichstein, M. and Brenning, A. (2020) Accounting for multiple testing in
730 the analysis of spatio-temporal environmental data. *Environmental and Ecological Statistics*,
731 27(2), 293–318. https://doi.org/10.1007/s10651-020-00446-4

732 Daly, C. (2006) Guidelines for assessing the suitability of spatial climate data sets. *International*
733 *Journal of Climatology*, 26(6), 707–721. https://doi.org/10.1002/joc.1322

Das, S., Mohapatra, M., Sahoo, U.K. and Baisya, H. (2024) High-intensity rainfall over northeast India: spatial pattern, short-term fluctuations and associated multiscale oscillations. *International Journal of Climatology*, 44(10), 3315–3327. https://doi.org/10.1002/joc.8525

Dhara, C., Deshpande, A., Roxy, M.K., Dalpadado, P. and Shrestha, M.S. (2025) A post-AR6 update on observed and projected climate change in India. *PLOS Climate*, 4(11), e0000724. https://doi.org/10.1371/journal.pclm.0000724

Donat, M.G., Alexander, L.V., Yang, H., Durre, I., Vose, R., Dunn, R.J.H., et al. (2013) Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: the HadEX2 dataset. *Journal of Geophysical Research: Atmospheres*, 118(5), 2098–2118. https://doi.org/10.1002/jgrd.50150

Douglas, E.M., Vogel, R.M. and Kroll, C.N. (2000) Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of Hydrology*, 240(1–2), 90–105. https://doi.org/10.1016/S0022-1694(00)00336-X

Farris, S., Deidda, R., Viola, F. and Mascaro, G. (2021) On the role of serial correlation and field significance in detecting changes in extreme precipitation frequency. *Water Resources Research*, 57(11), e2021WR030172. https://doi.org/10.1029/2021WR030172

Goswami, B.N., Venugopal, V., Sengupta, D., Madhusoodanan, M.S. and Xavier, P.K. (2006) Increasing trend of extreme rain events over India in a warming environment. *Science*, 314(5804), 1442–1445. https://doi.org/10.1126/science.1132027

Goteti, G. and Famiglietti, J. (2024) Extent of gross underestimation of precipitation in India. *Hydrology and Earth System Sciences*, 28(14), 3435–3455. https://doi.org/10.5194/hess-28-3435-2024

Grinsted, A., Moore, J.C. and Jevrejeva, S. (2004) Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11(5/6), 561–566. https://doi.org/10.5194/npg-11-561-2004

Gutiérrez, J.M., Jones, R.G., Narisma, G.T., Alves, L.M., Amjad, M., Gorodetskaya, I.V., et al. (2021) Atlas. In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press, pp. 1927–2058. https://doi.org/10.1017/9781009157896.021

764 Hamed, K.H. and Rao, A.R. (1998) A modified Mann–Kendall trend test for autocorrelated data.
765 *Journal of Hydrology*, 204(1–4), 182–196. https://doi.org/10.1016/S0022-1694(97)00125-X

766 Hofstra, N., New, M. and McSweeney, C. (2010) The influence of interpolation and station network
767 density on the distributions and trends of climate variables in gridded daily data. *Climate*
768 *Dynamics*, 35(5), 841–858. https://doi.org/10.1007/s00382-009-0698-1

769 Hosking, J.R.M. (1990) L-moments: analysis and estimation of distributions using linear
770 combinations of order statistics. *Journal of the Royal Statistical Society: Series B*
771 *(Methodological)*, 52(1), 105–124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

772 Hosking, J.R.M. and Wallis, J.R. (1997) *Regional Frequency Analysis: An Approach Based on L-*
773 *Moments*. Cambridge: Cambridge University Press.
774 https://doi.org/10.1017/CBO9780511529443

775 Huth, R. and Dubrovský, M. (2021) Testing for trends on a regional scale: beyond local significance.
776 *Journal of Climate*, 34(13). https://doi.org/10.1175/JCLI-D-19-0960.1

777 Jewell, S., Fearnhead, P. and Witten, D. (2022) Testing for a change in mean after changepoint
778 detection. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(4),
779 1082–1104. https://doi.org/10.1111/rssb.12501

780 Johannsen, I.M., Fabrikant, S.I. and Evers, M. (2018) How do texture and color communicate
781 uncertainty in climate change map displays? In: *10th International Conference on*
782 *Geographic Information Science (GIScience 2018)*, LIPIcs 114, 37:1–37:6.
783 https://doi.org/10.4230/LIPIcs.GISCIENCE.2018.37

784 Khiangte, L. (2026) DinosKuri/Calibrating-Change-Point-Inference-in-Mizoram's-Regional-
785 Rainfall-Extremes: Code for the Mizoram rainfall change-point paper (Version v1.0.0.0)
786 [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.22679527

787 Kolekar, V.N., Swarnkar, S. and Joshi, S.K. (2025) Evolving monsoon rain spells and compound
788 extremes across India: a century-long assessment. *Theoretical and Applied Climatology*,
789 156(10), 529. https://doi.org/10.1007/s00704-025-05763-x

790 Lawmchullova, I., Rinkimi, L. and Rao, C.H.U.B. (2024) Trend analysis of rainfall (1984–2023) of
791 Tlawng River basin of Mizoram, India using Mann–Kendall test. *Journal of*
792 *Agrometeorology*, 26(3), 391–394. https://doi.org/10.54386/jam.v26i3.2673

Lawmchullova, I., Lalngaihawma, J., Khumbah, A.L., Rinkimi, L., Sati, V.P., Saitluanga, B.L., et al. (2026) Unraveling spatio-temporal rainfall variability in the Tlawng watershed, Mizoram under a tropical monsoon regime. *Discover Geoscience*, 4(1), 13. https://doi.org/10.1007/s44288-025-00382-5

Ledoit, O. and Wolf, M. (2004) A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2), 365–411. https://doi.org/10.1016/S0047-259X(03)00096-4

Livezey, R.E. and Chen, W.Y. (1983) Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review*, 111(1), 46–59. https://doi.org/10.1175/1520-0493(1983)111<0046:SFSAID>2.0.CO;2

Lund, R.B., Beaulieu, C., Killick, R., Lu, Q. and Shi, X. (2023) Good practices and common pitfalls in climate time series changepoint techniques: a review. *Journal of Climate*, 36(23), 8041–8057. https://doi.org/10.1175/JCLI-D-22-0954.1

Maraun, D., Kurths, J. and Holschneider, M. (2007) Nonstationary Gaussian processes in wavelet domain: synthesis, estimation, and significance testing. *Physical Review E*, 75, 016707. https://doi.org/10.1103/PhysRevE.75.016707

Oliver, M.A. and Webster, R. (2014) A tutorial guide to geostatistics: computing and modelling variograms and kriging. *Catena*, 113, 56–69. https://doi.org/10.1016/j.catena.2013.09.006

Pai, D.S., Sridhar, L., Rajeevan, M., Sreejith, O.P., Satbhai, N.S. and Mukhopadhyay, B. (2014) Development of a new high spatial resolution (0.25° × 0.25°) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *MAUSAM*, 65(1), 1–18. https://doi.org/10.54302/mausam.v65i1.851

Paul, A.R. and Maity, R. (2023) Future projection of climate extremes across contiguous northeast India and Bangladesh. *Scientific Reports*, 13, 15616. https://doi.org/10.1038/s41598-023-42360-2

Paul, S., Sharma, P.J. and Teegavarapu, R.S.V. (2024) Spatial assessment of the reproducibility of Indian summer monsoon rainfall regimes in multiple gridded rainfall products. *Scientific Reports*, 14, 29269. https://doi.org/10.1038/s41598-024-75320-5

Peringiyil, A., Saharia, M., Sreejith, O.P., Wood, A.W. and Mohapatra, M. (2025) A station-based 0.1-degree daily gridded ensemble precipitation dataset for India. *Scientific Data*, 12, 333. https://doi.org/10.1038/s41597-025-04474-2

Pettitt, A.N. (1979) A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(2), 126–135. https://doi.org/10.2307/2346729

Phawa, R., Kusre, B.C. and Gupta, S. (2022) Analysis of a long-term IMD gridded rainfall data for dry period in Meghalaya. *Journal of the Indian Society of Remote Sensing*, 50(10), 1959–1977. https://doi.org/10.1007/s12524-022-01575-y

Rajeevan, M., Bhate, J., Kale, J.D. and Lal, B. (2006) High resolution daily gridded rainfall data for the Indian region: analysis of break and active monsoon spells. *Current Science*, 91(3), 296–306. [No DOI assigned] https://www.currentscience.ac.in/Volumes/91/03/0296.pdf

Renard, B., Lang, M., Bois, P., Dupeyrat, A., Mestre, O., Niel, H., et al. (2008) Regional methods for trend detection: assessing field significance and regional consistency. *Water Resources Research*, 44, W08419. https://doi.org/10.1029/2007WR006268

Roxy, M.K., Ghosh, S., Pathak, A., Athulya, R., Mujumdar, M., Murtugudde, R., Terray, P. and Rajeevan, M. (2017) A threefold rise in widespread extreme rain events over central India. *Nature Communications*, 8, 708. https://doi.org/10.1038/s41467-017-00744-9

Saji, N.H., Goswami, B.N., Vinayachandran, P.N. and Yamagata, T. (1999) A dipole mode in the tropical Indian Ocean. *Nature*, 401(6751), 360–363. https://doi.org/10.1038/43854

Schreiber, T. and Schmitz, A. (2000) Surrogate time series. *Physica D: Nonlinear Phenomena*, 142(3–4), 346–382. https://doi.org/10.1016/S0167-2789(00)00043-9

Schulte, J.A. (2016) Cumulative areawise testing in wavelet analysis and its application to geophysical time series. *Nonlinear Processes in Geophysics*, 23, 45–57. https://doi.org/10.5194/npg-23-45-2016

Sen, P.K. (1968) Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63(324), 1379–1389. https://doi.org/10.1080/01621459.1968.10480934

Sharma, A., Maharana, P. and Dimri, A.P. (2025) Changes in precipitation characteristics across different Indian sub regions. *Meteorological Applications*, 32(6). https://doi.org/10.1002/met.70127

Shi, X., Beaulieu, C., Killick, R. and Lund, R. (2022) Changepoint detection: an analysis of the Central England temperature series. *Journal of Climate*, 35(19), 6329–6342. https://doi.org/10.1175/JCLI-D-21-0489.1

Singh, S., Saini, A. and Kumar, S. (2025) Trends in extreme rainfall events over Northeast India: a novel perspective. *Journal of Earth System Science*, 134(1), 9. https://doi.org/10.1007/s12040-024-02461-6

Tebaldi, C., Arblaster, J.M. and Knutti, R. (2011) Mapping model agreement on future climate projections. *Geophysical Research Letters*, 38(23), L23701. https://doi.org/10.1029/2011GL049863

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J.D. (1992) Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, 58(1–4), 77–94. https://doi.org/10.1016/0167-2789(92)90102-S

Torrence, C. and Compo, G.P. (1998) A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Westra, S., Alexander, L.V. and Zwiers, F.W. (2013) Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, 26(11), 3904–3918. https://doi.org/10.1175/JCLI-D-12-00502.1

Wijngaard, J.B., Klein Tank, A.M.G. and Können, G.P. (2003) Homogeneity of 20th century European daily temperature and precipitation series. *International Journal of Climatology*, 23(6), 679–692. https://doi.org/10.1002/joc.906

Wilks, D.S. (2006) On field significance and the false discovery rate. *Journal of Applied Meteorology and Climatology*, 45(9), 1181–1189. https://doi.org/10.1175/JAM2404.1

Wilks, D.S. (2016) “The stippling shows statistically significant grid points”: how research results are routinely overstated and overinterpreted, and what to do about it. *Bulletin of the American Meteorological Society*, 97(12), 2263–2273. https://doi.org/10.1175/BAMS-D-15-00267.1

879 Yue, S., Pilon, P., Phinney, B. and Cavadias, G. (2002) The influence of autocorrelation on the
880 ability to detect trend in hydrological series. *Hydrological Processes*, 16(9), 1807–1829.
881 https://doi.org/10.1002/hyp.1095

882 Yue, S. and Wang, C.Y. (2004) The Mann–Kendall test modified by effective sample size to detect
883 trend in serially correlated hydrological series. *Water Resources Management*, 18(3), 201–
884 218. https://doi.org/10.1023/B:WARM.0000043140.61082.60

885 Zhang, X., Alexander, L., Hegerl, G.C., Jones, P., Klein Tank, A., Peterson, T.C., Trewin, B. and
886 Zwiers, F.W. (2011) Indices for monitoring changes in extremes based on daily temperature
887 and precipitation data. *WIREs Climate Change*, 2(6), 851–870.
888 https://doi.org/10.1002/wcc.147

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx