water

MDPI

Article

A Multi-Method Approach to Analyzing Precipitation Series and Their Change Points in Semi-Arid Climates: The Case of Dobrogea

Youssef Saliba [1] and Alina Bărbulescu [2,*]

[1] Doctoral School, Technical University of Civil Engineering of Bucharest, 122-124 Bd. Lacul Tei, 020396 Bucharest, Romania; youssefsaliba@gmail.com
[2] Department of Civil Engineering, Transilvania University of Brașov, 5 Turnului Str., 500152 Brașov, Romania
* Correspondence: alina.barbulescu@unitbv.ro

Abstract: The Dobrogea region, located in southeastern Romania, experiences a semi-arid climate. This study provides a deep analysis of monthly precipitation series from 46 meteorological stations spanning 1965–2005, exploring mean and variance characteristics and detecting structural changes in precipitation patterns. The series normality was assessed using the Lilliefors test, and transformation, such as the Yeo–Johnson method, was used to address skewness. Analyses of mean and variance included parametric (*t*-tests, ANOVA) and non-parametric (Mann–Whitney U, Fligner–Killeen) tests to address the homogeneity/inhomogeneity of the data series in mean and variance. Change points were detected using a Minimum Description Length (MDL) framework, modeling the series as piecewise linear regressions with seasonal effects and autocorrelated errors. Pairwise comparisons indicate the low similarity of the series means, and variances, so spatial and temporal variability in precipitation is notable. Validation of the proposed MDL approach on synthetic datasets demonstrated high accuracy, and application to real data identified significant shifts in precipitation regimes. Applied to the monthly series collected at the ten main hydro-meteorological stations, a MDL framework provided at least two change points for each.

Keywords: change point; MDL; homogeneity; statistical analysis

Academic Editors: Chin H. Wu and Alexander Shiklomanov

Received: 30 November 2024
Revised: 28 January 2025
Accepted: 30 January 2025
Published: 31 January 2025

Citation: Saliba, Y.; Bărbulescu, A. A Multi-Method Approach to Analyzing Precipitation Series and Their Change Points in Semi-Arid Climates: The Case of Dobrogea. *Water* **2025**, *17*, 391. https://doi.org/10.3390/w17030391

1. Introduction

Recent research shows that climate change has led to a shift in precipitation evolution worldwide [1–3]. Given the importance of understanding the precipitation variability over time for agricultural planning and water resource management, many studies have analyzed the extreme events at scales from hours to days [4–6] and their immediate economic impact [7–9]. However, significant consequences on human activity and ecosystems also result from longer-duration events, such as seasonal flooding or droughts. They can have profound effects on immediate weather and long-term environmental changes. For instance, research by Wainwright et al. [10], Li et al. [11], and Knapp et al. [12] underscores how shifts in rainfall patterns and prolonged periods of extreme weather can disrupt plant growth, alter biodiversity, and challenge land management practices. The importance of addressing these longer-term events is growing, as their effects can be more widespread and persistent compared to shorter-duration extremes [13,14].

Precipitation varies significantly across different year periods, showing a distinct seasonal pattern. Because the climate system is influenced by predictable cycles and

unpredictable changes, consistent observations over an extended period are essential for a correct analysis that can lead to understanding seasonal events [15,16].

In the Northeast United States (NEUS), the region covering New York city, Philadelphia, and Washinton D. C. shares both climatic (Cfa) and topographic features with Dobrogea, the study by Huang et al. [17] indicates a significant increase in the extreme and total annual rainfall, with change points in 1996 and 2002. In [18], it is shown that 273 extreme events explained 89% of the increase in extreme precipitation events during 1996–2016 compared to 1979–1995. Kunkel et al. [19] showed that since 1991, all regions in the NEUS have experienced the highest number of extreme events compared to the other periods. A statistically significant trend of extreme precipitation from 1957 to 2010 was detected. The increasing tendency of heavy precipitation in the NEUS, documented in the studies of Kunkel et al. [19], Groisman et al. [20], Westra et al. [21], Hoerling et al. [22], and Jong et al. [23], highlights the growing influence of climate change on regional weather patterns. Much of the increase has been observed over the last several decades, corresponding with the broader trends of global warming.

Other studies indicate precipitation variability over Europe [24,25], emphasizing a quasi-period pattern. Studies performed for Romania reveal mutations in the extremes of daily precipitation [26], changes in precipitation recorded in winter [27], modifications in the most probable precipitation [28], and changes in water balance [29].

Since Dobrogea is one of the Romanian regions where agriculture is a widely spread occupation and that was most affected by climate change, different studies have been dedicated to analyzing temperature and precipitation variability and modeling their evolution at a regional scale [30–37]. Most of them emphasize the increased drought periods using drought indices [31], climate water balance [32], trend detection by the Mann–Kendall trend tests, and Sen's slope and other indicators on a daily scale [33–35]. Change points detection of annual meteorological series was also performed [33,37]. Moreover, other scientists analyzed the response of the plants and soil to climate change stress and the necessity of irrigation [38–40].

Despite the extensive amount of literature on Dobrogea's climate, the long-term precipitation series were not investigated using a big database. Therefore, this study will fill a gap in the knowledge about the variability of the precipitation series in this zone using monthly precipitation series collected at 46 meteorological stations from 1965 to 2005. More precisely, we aim to address the similarities/dissimilarities of the data series as an essential stage in a more extensive study for modeling the regional evolution of precipitation in the region. We shall answer the following questions.

(a) Are the series homogenous in mean or variance?

(b) Do the series present breakpoints (change points)?

To answer the first one, we performed parametric and nonparametric statistical tests on the individual and cumulated series, raw or transformed, to achieve normality. To answer the second one, we propose a method for multiple change points detection (implemented in R), which can be applied to a wide range of series.

2. Materials and Methods

2.1. Study Region and Data Series

The Dobrogea region (Figure 1) is located in southeastern Romania, bordered by the Black Sea to the east and the Danube River to the west. It is characterized by a relatively dry climate with notable variations in precipitation patterns.

Figure 1. The map of Dobrogea, Romania, and the hydrological stations.

The climate is semi-arid, especially in the southern parts of Dobrogea. The months of June to August are generally dry and hot. Dobrogea receives relatively low amounts of annual precipitation, averaging around 400–500 mm per year, making it one of the driest areas in Romania. The highest precipitation were recorded in April–May and September–October, when more frequent rainstorms appeared, sometimes leading to localized flooding in areas with poor drainage. The belt situated on the Black Sea benefits from moderate temperatures and a slightly higher precipitation quantity than the inland areas. The central and northern parts of Dobrogea are drier, with more continental influences, so lower precipitation levels.

The precipitation is unevenly distributed across the region, with the zones closer to the Danube River experiencing more precipitation than the Littoral area that benefits from the moderate humidity of the sea [35].

The data series studied in this article is formed by the monthly precipitation series from 46 meteorological stations spanning 1965–2005. This set was used because it is complete, without gaps, and the most recent data series is not publicly available. After obtaining new data series, the study will be extended to cover the entire period from 1965 to 2024. Still, the methodology used here can be used for any number of series and any period.

2.2. Data Transformation

Various modeling techniques rely on series normality. When the original series does not follow a Gaussian distribution, a transformation is necessary to reach normality.

In this study, the data series (individual and global) normality was assessed by the Lilliefors test [41], whose null hypothesis is the series normality, and the alternative is non-normality. The test statistics is given in (1):

$$D = \max|F_n(x) - F_0(x)|, \quad (1)$$

where $F_n(x)$ is the transformed data's empirical cumulative distribution function (ECDF) and $F_0(x)$ is the cumulative distribution function (CDF) of a normal distribution estimated from the sample.

A Yeo–Johnson transformation [42] can be applied to reach normality when the null hypothesis is rejected. This transformation is a variant of the Box–Cox transformation, whose equation is (2):

$$Y(\lambda) = \begin{cases} \frac{\left((Y+1)^{\lambda}-1\right)}{\lambda} \; if \; \lambda \neq 0, \; Y \geq 0 \\ \frac{-\left((-Y+1)^{2\lambda-1}-1\right)}{2-\lambda} \; if \; \lambda \neq 2, \; Y < 0 \end{cases}, \quad (2)$$

where Y represents the original data series, $Y(\lambda)$ is the transformed one, and λ is a parameter whose value is determined using the maximum likelihood estimation for optimization.

The transformed data keeps the original features of the precipitation series, which represent the variability seen in the real world. This property is important in predictive modeling given that the ecological and environmental datasets benefit from keeping data in their original format [43].

2.3. Parametric and Non-Parametric Tests for Assessing the Series Homogeneity in Mean and Variance

To assess the reliability of statistical conclusions, non-parametric tests on the original and transformed datasets can be run simultaneously. The consistency of the results (from the original and transformed datasets) indicates their accuracy, indicating that the transformation probably did not influence the interpretation. Conversely, since transformations frequently improve the validity of parametric analyses (by reducing issues like skewness or heteroscedasticity), significant discrepancies in results suggest that the transformation has affected statistical power or data distribution assumptions [43]. This strategy is helpful because transformations might not be required if non-parametric methods produce reliable results despite the non-normal or heteroscedastic nature of the data [44].

When normality and homoscedasticity hypotheses cannot be rejected, parametric tests are given priority in this analysis. When using parametric methods like *t*-tests and ANOVA, data are assumed to follow a normal distribution with homogeneous variances across groups. Under these circumstances, parametric tests are more useful for analysis because they are sensitive to distributional information and can thus detect smaller differences in variance or mean [45]. On the other hand, the accuracy of parametric tests can be compromised when assumptions are not met, as non-normality and heteroscedasticity introduce biases [46]. In such cases, non-parametric tests, which do not rely on specific distributional assumptions, provide a better alternative. Methods such as the Mann–Whitney U-test for pairwise comparisons [47] and the Fligner–Killeen test [48] for global assessments are applied to guarantee that statistical conclusions hold up under different data properties.

2.3.1. Statistical Tests for Means Equality

Based on the results of normality tests and those of the transformations, stations are classified as "normal" or "non-normal", and the analysis of mean precipitation differences is performed separately for each on both the local (pairwise) and global scales.

For "normal" stations, we utilize pairwise *t*-tests [49] to test the null hypothesis that the means of stations i and j are equal against the inequality of the means. One way to determine the t-statistic is by (3):

$$t = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{\frac{s_i^2}{n_i} - \frac{s_j^2}{n_j}}}, \tag{3}$$

where $\overline{X}_i$ and $\overline{X}_j$ are the sample means, s_i^2 and s_j^2 are the sample variances, and n_i and n_j are the sample sizes at stations i and j, respectively.

A one-way Analysis of Variance (ANOVA) is employed [50] to test the null hypothesis that all station means are equal. The F-statistic is (4):

$$F = \frac{MS_{between}}{MS_{within}}, \tag{4}$$

where $MS_{between}$ is the mean square between groups and MS_{within} is the mean square within groups. Post hoc pairwise comparisons using the Student–Newman–Keuls (SNK) test [51] help distinguish which specific stations differ if the ANOVA indicates significant differences. This goal is achieved using ordered group means and critical values for differences calculated by (5):

$$q = \frac{\overline{X}_{(r)} - \overline{X}_{(s)}}{\sqrt{\frac{MS_{within}}{n}}}, \tag{5}$$

where $\overline{X}_{(r)}$ and $\overline{X}_{(s)}$ are the group means ranked from highest to lowest and n is the sample size per group.

For "non-normal" stations, the Mann–Whitney U-test [52,53] was used to assess pairwise differences in mean precipitation. The null hypothesis is that the precipitation distributions at two stations, i and j, are the same, and the alternative is that they differ. The U statistic is given in (6):

$$U = \min(U_i, U_j), \tag{6}$$

where

$$U_i = n_i n_j + n_i(n_i + 1)/2 - R_i, \tag{7}$$

$$U_j = n_i n_j + n_j(n_j + 1)/2 - R_j, \tag{8}$$

n_i and n_j are the sample sizes of stations i and j, respectively, and R_i and R_j are the ranks of precipitation levels for the two stations combined.

A Welch's ANOVA [54] was performed for global comparisons, accommodating unequal variances. The Welch's F*-statistic for K groups is defined in (9):

$$F^* = \frac{\sum_{k=1}^{K} \frac{\omega_k (\overline{X}_k - \overline{X})^2}{(K-1)}}{1 + \frac{2(K-2)}{K^2-1}\Lambda}, \tag{9}$$

where

- $\omega_k = n_k / s_k^2$ is the weight of group k, based on the group size n_k and variance s_k^2;
- $\overline{X}_k$ is the mean value of group k;
- $\overline{X} = \frac{\sum_{k=1}^{K} \omega_k \overline{X}_k}{\sum_{k=1}^{K} \omega_k}$ is the weighted mean of all groups;

- $\Lambda = \frac{\left(1-\omega_k/\sum_{k=1}^{K}\omega_k\right)^2}{n_k-1}$ is the correction factor used to adjust the unequal variances and sample sizes.

If significant differences are detected, the Dunn test with Bonferroni correction [55] is used for pairwise comparisons. It can be written as in (10):

$$Z = \frac{R_i - R_j}{\sqrt{\left(\frac{N(N+1)}{12}\right)\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}, \tag{10}$$

where R_i and R_j are the rank sums for groups being compared, N is the total number of observations across all groups, n_i and n_j are the sample sizes of the two groups.

The Bonferroni correction adjusts the p-values by multiplying them by the number of comparisons, k, to control the increased likelihood of a Type I error (false positive) across multiple tests. The formula is given in (11):

$$p_{adjusted} = \min(p \times k, 1). \tag{11}$$

2.3.2. Statistical Tests for Homoscedasticity

To assess precipitation variability across weather stations, homoscedasticity, or the homogeneity of variances, must be checked in addition to normality. Homoscedasticity means that the data series from each station have similar variance levels, which is a precondition for using parametric tests.

In the local analysis, pairwise variance comparisons are conducted. For "normal" stations, Levene's test [56] evaluates the equality of variances by analyzing the absolute deviations of observations from the group mean. However, for "non-normal" stations, the Brown–Forsythe test [57], a variation of Levene's test, was applied, in which the median is used instead of the mean as the measure of the central tendency. Both tests are represented in (12):

$$W = \frac{(N-k)}{(k-1)} \cdot \frac{\sum_{j=1}^{k} n_j (M_j - M)^2}{\sum_{j=1}^{k}\sum_{i=1}^{n_j} (Z_{ij} - M_j)^2} \tag{12}$$

where

- $Z_{ij} = |Y_{ij} - Y_j|$, with Y_j being the group mean (for Levene's) or the group median (for the Brow-Forsythe test);
- M_j denotes the mean of Z_{ij} values for group j;
- M is the overall mean value of M_j across all the k groups;
- n_j is the sample size of group j, and N is the number of observations across all groups.

For the global assessment, a one-way ANOVA is applied for "normal" and "homoscedastic" stations, followed by the SNK post hoc test given in (5) if significant differences are detected. As for stations failing to satisfy any of these assumptions (i.e., normality or homoscedasticity), the Fligner–Killeen test [48] is used. The test statistic, FK, is given by:

$$FK = \frac{\sum_{j=1}^{k} n_j \left(\text{rank}(Z_j) - \overline{R}\right)^2}{\sum_{i=1}^{N} \left(\text{rank}(Z_i) - \overline{R}\right)^2 / (N-1)} \tag{13}$$

where

- N is the total number of observations;
- Y is the global series (formed by all the series recorded at all the stations);
- $Z_i = |Y_i - \text{median}(Y)|$;
- n_j is the sample size of group j;

- k is the number of groups;
- $\overline{R}$ is the mean rank.

This method leverages ranking to mitigate the effect of outliers and distributional non-normality [58]. If the Fligner–Killeen test detects significant differences, Dunn's test with Bonferroni correction from (10) is applied to identify specific groups.

2.4. Change Point Detection

2.4.1. The Change Point Problem

In research fields such as climatology, where alterations in instrumentation, station relocations, or external environmental factors often correlate with structural changes in time series, change point analysis is essential to interpreting these phenomena.

The detection of a single change point is a classical problem, often framed as a test for changes in the mean of a time series. Mathematically, the series $\{X_t\}$ for $t = 1, \ldots, N$ can be modeled by (14):

$$X_t = \begin{cases} \mu_1 + \varepsilon_t, & t \leq \tau \\ \mu_2 + \varepsilon_t, & t > \tau \end{cases} \tag{14}$$

where τ is the location of the change point, μ_1 and μ_2 are the mean values of the segments before and after the shift, and ε_t are independent random errors (often assumed to follow a normal distribution with zero mean and constant variance).

The objective is to determine whether $\mu_1 \neq \mu_2$ signifies a change in the mean at location τ. This objective can be formulated as hypothesis testing, where the null hypothesis $H_0: \mu_1 = \mu_2$ assumes no change, and the alternative hypothesis, $H_a: \mu_1 \neq \mu_2$, indicates the opposite.

Classical single change point detection methods include the Pettitt [59], Buishand [60,61], and Lee and Heghinian [62] tests, implemented in Khronostat [63]. These approaches face challenges when extended to time series exhibiting multiple change points [64]. Methodologies such as binary segmentation and penalized likelihood models extend the single change point detection problem to accommodate multiple changes. The first method is an iterative process that detects a single change point, splits the series at an identified point, and then repeats the process within each segment. On the other hand, penalized likelihood models balance the model complexity, and the data fit by applying a penalty that discourages overfitting.

Rissanen [65] rigorously formulated the concept of balancing data fit with model simplicity. He introduced the shortest data description principle, now widely known as the Minimum Description Length (MDL) criterion. His framework defines model complexity in terms of the number of bits required to encode the model and the data it describes. By minimizing the total description length, the MDL principle enables the simultaneous estimation of model coefficients and structural parameters, making it particularly well-suited for handling multiple change points in time series data.

Davis et al. [66] further developed the MDL framework by introducing a methodology to detect structural breaks in nonstationary time series, modeling each segment with autoregressive (AR) processes. Their approach integrated a genetic algorithm to optimize the detection of multiple change points, addressing the computational challenges posed by the large search space of possible model configurations. This method, while general, laid the foundation for subsequent adaptations that are specific to climatic data.

Building on the previous work, Lu et al. [67] adapted the MDL framework for climate time series by accounting for periodicities and autocorrelations that are inherent in such data. Their method added periodic autoregressions (PAR) to capture seasonal dynamics and further refined the MDL objective function to better suit climatic applications. This

approach provided a robust tool for segmenting time series with periodic variability and autocorrelation using a genetic algorithm for optimization.

On the other hand, Killick et al. [68] introduced the Pruned Exact Linear Time (PELT) algorithm to address the computational inefficiencies in change point detection methods. The PELT algorithm achieves a linear computational cost under mild conditions, making it significantly more efficient than quadratic or cubic algorithms like Segment Neighborhood (SN) or Optimal Partitioning (OP). By combining dynamic programming with pruning, the PELT algorithm eliminates candidate change points that cannot improve the cost function. This capacity makes it suitable for large climatological datasets, where multiple change points may need to be identified across extensive records. Importantly, the PELT algorithm supports cost functions such as MDL and penalized likelihood, enabling its integration with autoregressive modeling frameworks.

Metadata, or information about possible change point times derived from station history, was further added to the MDL methodology by Li et al. [69]. This Bayesian MDL (BMDL) framework codifies expert knowledge into a probabilistic framework to improve detection performance in univariate and multivariate contexts. These innovations are essential for climate homogenization, where small changes can throw off long-term trend evaluations. Therefore, a Bayesian penalized likelihood model that considers autocorrelation, seasonal variability, and data-scale intricacies was developed and optimized using genetic algorithms [70]. These additions expand upon previous ground-breaking research, for example, by Caussinus and Mestre [71], who were the first to apply penalized likelihood to the problem of climate series homogenization, and by Menne and Williams [72], who solved the problem of undocumented change points by combining statistical testing with reference series.

Contemporary MDL-based methods have surpassed conventional segmentation techniques by utilizing autoregressive models and hierarchical priors. These methods offer a computationally efficient and statistically robust framework for analyzing structural breaks in intricate datasets. Collectively, these studies illustrate a paradigm shift in change point analysis, improving its efficacy in detecting and reducing the impacts of artificial shifts while preserving the integrity of long-term climate assessments.

2.4.2. The MDL Framework

In this study, we adopt the framework of Lu et al. [67] to formulate the MDL objective function for detecting change points in climatic time series. However, clarifying some foundational statistical concepts that underline this methodology is important, as these may initially seem technical or abstract. The likelihood function is at the framework's core, quantifying how well a statistical model explains the observed data. In simpler terms, the likelihood informs us of how plausible the observed precipitation values (log-transformed in our case) are under an assumed structure of seasonal trends, linear shifts, and change points. It is then transformed into a more manageable log-likelihood form to stabilize numerical computations and simplify optimization. It is also worth mentioning that the log-likelihood for a normal distribution includes terms penalizing variance (to prevent overly broad fits) and measuring the fit quality using residuals.

This approach models time series data as a piecewise linear regression with seasonal effects, accommodating level shifts at change point times and accounting for autocorrelated errors. Specifically, we assume that the data follows model (15):

$$X_{nT+\nu} = \mu_\nu + \alpha(nT+\nu) + \delta_{nT+\nu} + \varepsilon_{nT+\nu} \tag{15}$$

where μ_ν represents the seasonal mean value at period ν, α is the linear trend parameter, $\delta_{nT+\nu}$ captures the mean shift at the change point, and $\varepsilon_{nT+\nu}$ denotes the zero-mean periodic

autoregressive errors with periodic variance $\sigma^2(\nu)$. So, the residuals are assumed to follow a PAR process, which reflects the autocorrelation in climatic time series, with periodic variances that capture the variability in different seasons.

While preserving a good fit to the observed data, the MDL objective function is designed to penalize model complexity. The total description length is given by (16):

$$\begin{aligned} MDL(\mathcal{M}) = \tfrac{1}{2}\sum_{i=1}^{m+1}\ln(\tau_i - \tau_{i-1}) + \frac{pT\ln(2d)}{2} + \ln(m) + \ln(p) \\ + \tfrac{1}{2}\sum_{t=1}^{N}\ln(v_t) + \tfrac{1}{2}\sum_{t=1}^{N}\frac{\left(X_t - \hat{X}_t\right)^2}{v_t} + \sum_{i=2}^{m}\ln(\tau_i) \end{aligned} \tag{16}$$

with

- v_t—the variance of the residuals;
- p—the periodic autoregression order;
- m—the number of change points.

The values of τ_i denote the change point moments, where $\tau_0 = 1$ is the start of the series, and $\tau_{m+1} = N + 1$ is the end of the series. The indices $(\tau_i - \tau_{i-1})$ represent the length of each segment between consecutive change points. $\hat{X}_t$ refers to the next step predicted value of X_t, computed using the model parameters estimated for the segment in question.

The penalty terms aim to prevent overfitting by discouraging overly complex models. The term $\ln(\tau_i - \tau_{i-1})$ penalizes change points that create very short segments, maintaining a meaningful segmentation. Additional penalties, such as $\ln(m) + \ln(p) + \sum_{i=2}^{m}\ln(\tau_i)$, account for the overall number of change points, while $pT/2\ln(2d)$ specifically penalizes the complexity introduced by the PAR error model. These penalties are balanced against the data fidelity terms $1/2\sum_{t=1}^{N}\ln(v_t) + 1/2\sum_{t=1}^{N}\left[\left(X_t - \hat{X}_t\right)^2/v_t\right]$.

For optimization, the model parameters are reformulated into a compact matrix representation. Specifically, the regression equation is written in the form:

$$X = D\beta + \varepsilon, \tag{17}$$

where

- X is the $N \times 1$ vector of observed data;
- D is the design matrix capturing seasonal indicators, trend components, and regimen shifts;
- β is the vector of model parameters (including μ_ν, α, and δ);
- ε is the vector of residuals.

The $m + 1$ regimes are fully represented in the design matrix, which allows efficient parameters' estimation via generalized least squares. These parameters are then optimized using Simulated Annealing (SA).

2.4.3. Simulated Annealing

The iterative optimization technique known as "Simulated Annealing" (SA) is based on the physical phenomenon of "annealing" as seen in materials science, whose goal is to achieve a low-energy crystalline configuration by heating and then slowly cooling a material to minimize defects. For combinatorial optimization problems with complex objective functions, such as finding change points in precipitation time series data, the method initially proposed by Kirkpatrick et al. [73] performs very well.

The SA process begins by initializing the solution to an empty set of change points ($\mathbb{C} = \emptyset$) and computing its corresponding MDL score. The initial temperature (this is

the generic term used by the method, even if the modeling is not about temperature as a meteorological factor), T_0, is set high to allow for a broad exploration of the solution space. The temperature decreases gradually according to a cooling schedule defined by:

$$T_{k+1} = \alpha T_k \tag{18}$$

where α is the cooling rate whose values are between 0 and 1, and k is the iteration's number.

Equation (18) ensures transitions from exploration to exploitation when the algorithm progresses. At each iteration, a new candidate solution is proposed by modifying the current set of change points, $\mathbf{C}$, through one of three actions randomly selected:

- *Addition*: A new change point is sampled randomly from the range $2, \ldots, n-1$ (to avoid boundary points) and added to $\mathbf{C}$, provided that it does not already exist.
- *Removal*: An existing change point is removed from $\mathbf{C}$, provided that the set is not empty.
- *Modification*: An existing change point is randomly replaced with a new position, ensuring no duplicates.

The proposed solution is sorted to maintain chronological order, and duplicates are removed. The MDL score of the proposed solution is then computed. If the new solution improves the score, i.e., $\Delta MDL = MDL_{proposed} - MDL_{current} < 0$, it is immediately accepted. Otherwise, it is accepted probabilistically to allow exploration of suboptimal solutions, with a probability given by:

$$P_{accept} = \exp\left(-\frac{\Delta MDL}{T}\right), \tag{19}$$

allowing the algorithm to escape local minima by occasionally accepting worse solutions early in the process, while gradually forcing better solutions as the temperature decreases. At each iteration, the algorithm keeps track of the best solution encountered so far in terms of its MDL score. This ensures that the final output is the solution with the lowest MDL score, regardless of whether it was accepted during the iterative process. The process continues until a maximum number of iterations is reached or the temperature becomes negligible. The final output consists of the best change point set and its corresponding MDL score.

The pseudocode of the SA algorithm, implemented in R for the current study, is summarized as follows.

1. *Initialization*: Start with $\mathbf{C} = \emptyset$, compute $MDL_{current}$, set $\mathbf{C}_{best} = \mathbf{C}$, and initialize $T = T_0$;
2. *Iteration:*
   - Propose a new solution, $\mathbf{C}'$, by adding, removing, or modifying change points;
   - Compute $MDL_{proposed}$;
   - Accept $\mathbf{C}'$ if $\Delta MDL < 0$ or with the probability given by (19) otherwise;
   - Update $\mathbf{C}_{best}$ if $MDL_{proposed}$ is the lowest seen so far;
   - Reduce T according to (18);
3. Output: Return $\mathbf{C}_{best}$ and its corresponding MDL score.

A summary of the procedure presented in Sections 2.4.2 and 2.4.3 and used to determine the change points is presented in Appendix A.

Figure 2 contains the study flowchart.

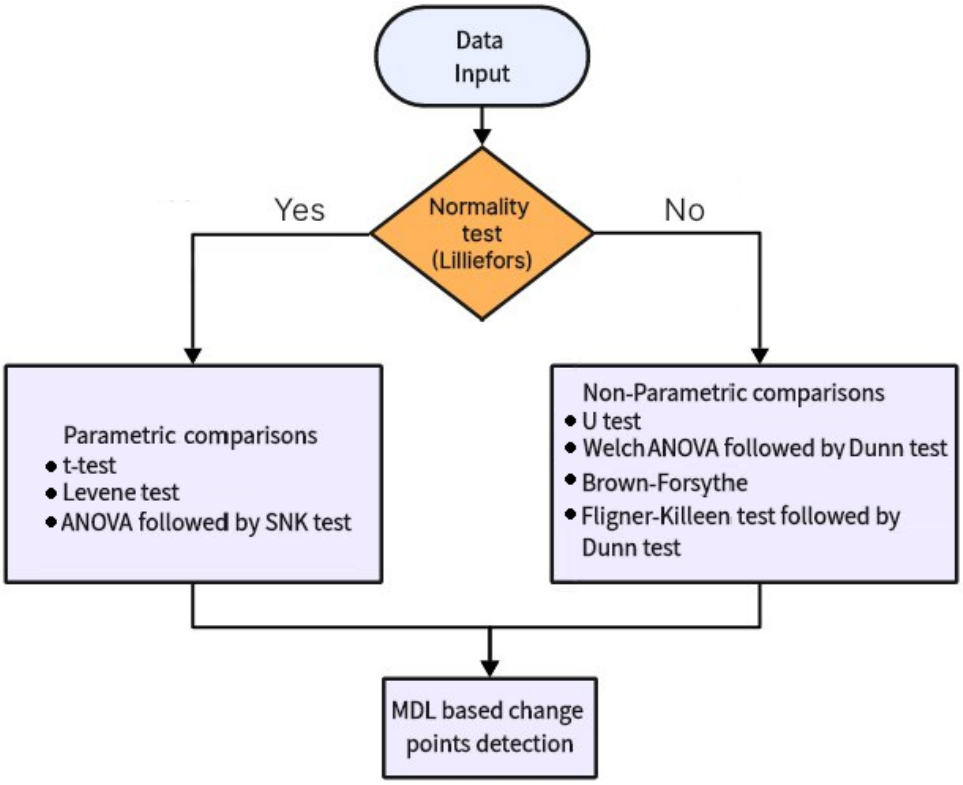

Figure 2. The study flowchart.

3. Results

The basic statistics indicate significant spatial variability in precipitation levels. Total precipitation varies from 382.3 mm to 1444.3 mm. The highest values were recorded at Niculițel, Albești, and Dăeni, indicating a more pronounced variability in precipitation. The average monthly precipitation ranges from 21.8 mm in Sulina to 49.2 mm in Niculițel. Niculițel also has the highest interquartile range (IQR = 44.75), indicating a greater dispersion of monthly precipitation values. The skewness coefficient is positive in all cases, with the highest values found in Mangalia, Pecineaga, and Albești, suggesting more extreme precipitation events. Additionally, most stations exhibit kurtosis values greater than 3, indicating leptokurtic distributions. A global analysis reveals a non-Gaussian distribution with a skewness of 1.852 and a kurtosis of 6.058.

3.1. Normality Tests on the Initial and Transformed Precipitation Series

The Lilliefors test was applied to each series at a significance level of 5% to address the series normality. As shown in Figure 3a, the *p*-values for all series are near zero (so, under the 0.05 significance level), confirming the departure from normality.

The initial Regional series (built by all the recorded series) displayed a right-skewed pattern, as illustrated in Figure 3b. The corresponding Q-Q plot from Figure 3c (blue dots) confirms the departure from normality (red line). The original data showed heavy tails, indicating that parametric statistical tests might yield biased results due to non-normality. Therefore, a Yeo–Johnson transformation was applied to stabilize variance and reduce skewness. The transformed Regional series had a more symmetric and bell-shaped distribution (Figure 4a), suggesting an improvement in normality. This finding is confirmed by the Q-Q plot (Figure 4b), where the samples' quantiles align more closely with the theoretical ones, showing closer adherence to normality.

The Lilliefors test was also applied to each individual transformed series. As illustrated in Figure 4c, *p*-values for most stations were above the 0.05 threshold, confirming that

normality was achieved after transformation. However, ten series (highlighted in red in Figure 4c) still showed significant deviations from normality even after transformation, as their *p*-values remained below the threshold.

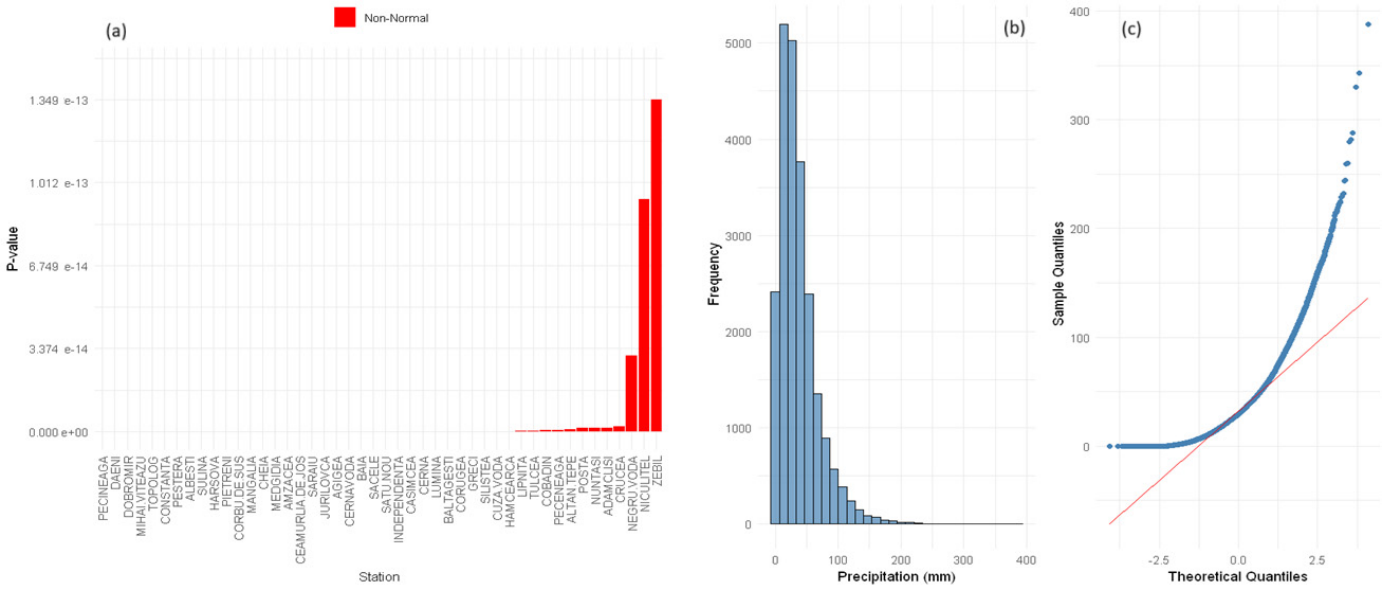

Figure 3. (**a**) Lilliefors *p*-values for the original data, (**b**) the distribution of the original Regional series and (**c**) the Q-Q plot of the original precipitation series, indicating the departure from normality.

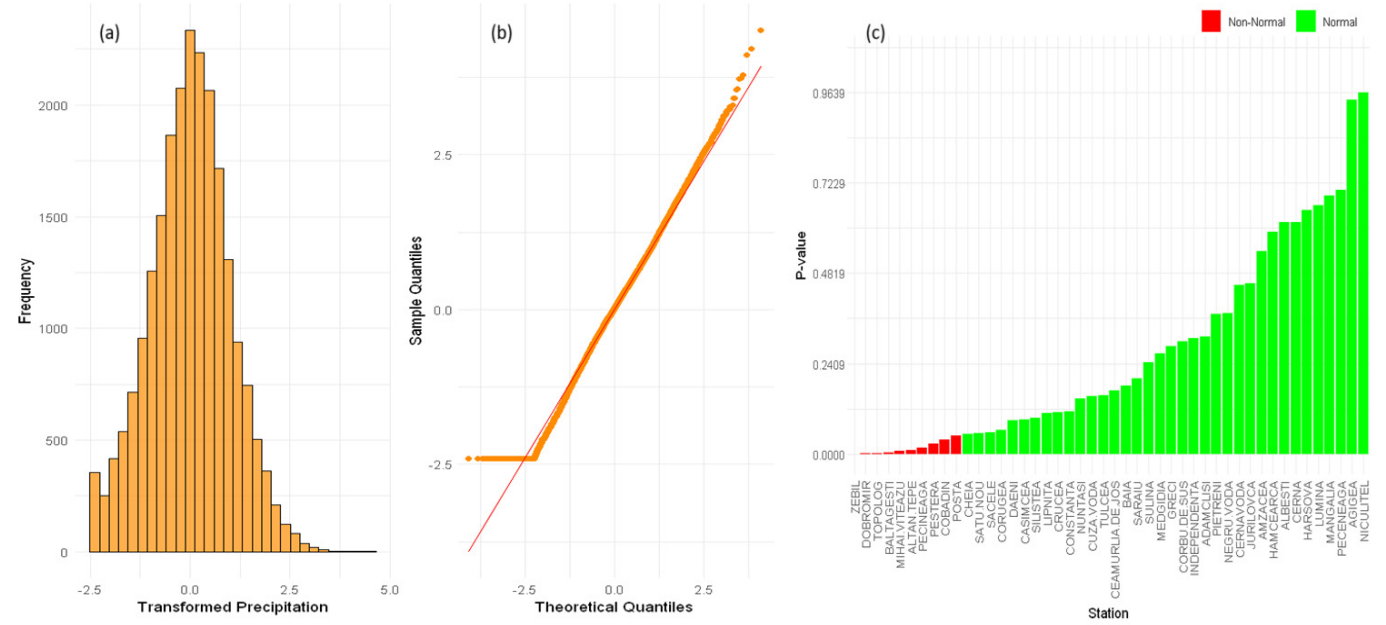

Figure 4. (**a**) The distribution and (**b**) Q-Q plot of the transformed Regional series (λ = 0.31), and (**c**) Lilliefors *p*-values for the transformed individual series. In red is the series that did not reach normality.

3.2. Comparisons of Mean Values

3.2.1. Comparisons for 'Normal' Series

First, a pairwise *t*-test was performed for the Gaussian series. Figure 5a indicates the significance of differences between the pairs of means of normally distributed series. Each cell compares a pair of stations, where green indicates that the equality of the two means cannot be rejected (so the difference between the two means is not significant). The red cells indicate significant differences (i.e., the hypothesis that the two means are equal was

rejected, so $p < 0.05$). This visualization reveals several significant differences between the series averages, reflecting the regional variability of precipitation levels. Notably, some stations exhibit clustering in significance patterns, suggesting possible similarities or regional groupings in mean precipitation.

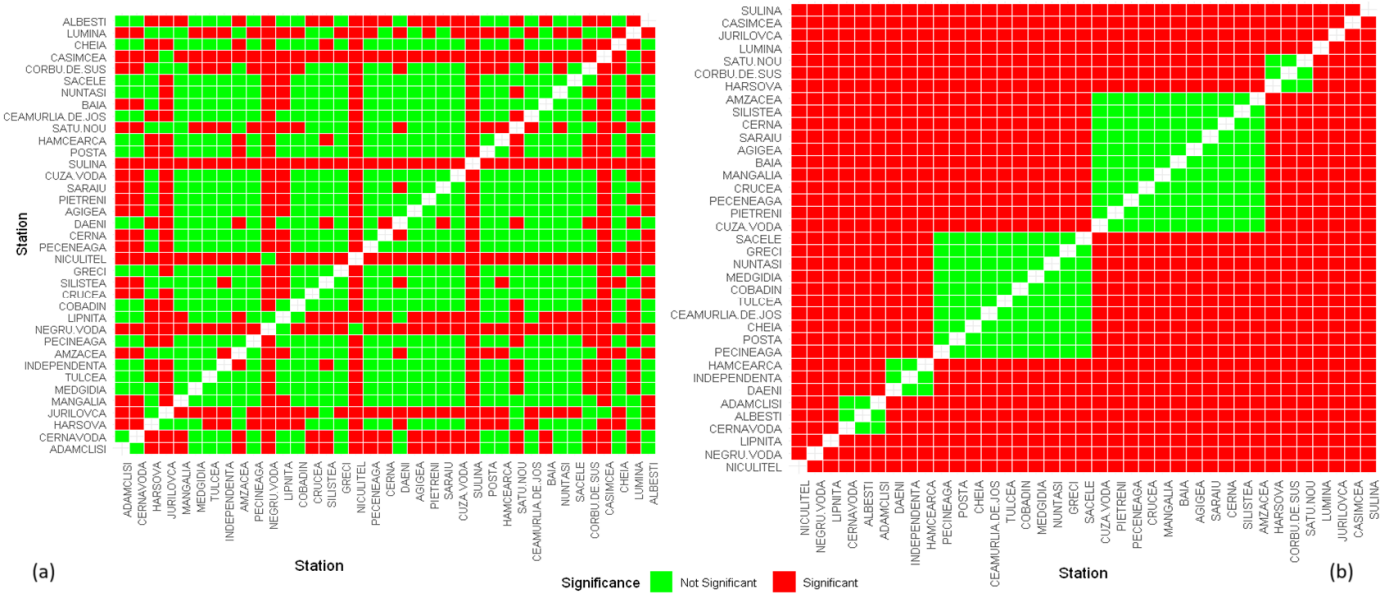

Figure 5. (**a**) Pairwise *t*-tests for normal stations. Green indicates that the equality of the two means of the stations defining the cells in the matrix cannot be rejected; (**b**) SNK test results following one-way ANOVA for normal stations.

The results of the global assessment of the Gaussian series' mean equality (i.e., the SNK post hoc results for the one-way ANOVA) are shown in Figure 5b. Some groups of stations (represented by green cells in Figure 5b) exhibit consistent non-significant differences in means, indicating relatively homogenous average precipitation levels within these clusters. In contrast, other pairs of stations (represented by red cells in Figure 5b) diverge considerably. These patterns may reflect the underlying climatic or geographic factors affecting regional precipitation dynamics.

Following the initial analysis, Figure 6 compares local (*t*-test) and global (SNK) results, highlighting the degree of concordance between these methods. Green cells indicate agreement between the local and global analysis, while orange cells denote disagreement. Stations displaying consistent concordance across both local and global tests represent clusters with stable precipitation characteristics that are unaffected by the scale of analysis. These stations likely belong to regions with uniform climatic influences, making them reference points for understanding broad precipitation patterns in Dobrogea. By contrast, stations with high levels of disagreement may be more sensitive to microclimatic or topographic variations that local tests capture but may not align with the broader regional trends identified by global analysis.

This dual analysis offers two primary insights. First, stations with consistent tests results could serve as stable indicators for studying long-term precipitation trends. Second, stations with discordant results deserve further investigation, as they may highlight localized climatic effects or unique geographic characteristics.

Figure 6. Comparison of the *t*-test and SNK results. A green (orange) cell indicates the concordance (discordance) between the tests' results on a pair of stations.

3.2.2. Comparisons for 'Non-Normal' Series

The Mann–Whitney U-test was employed to compare mean values of non-normal series pairwise. The green cells in Figure 7a indicate that the mean equality hypothesis could not be rejected (so the difference between the means is not significant). The matrix reveals variations in significance, with certain stations forming clusters of non-significance (green cells situated one next to each other) that may point toward local similarities in the precipitation means or microclimatic effects.

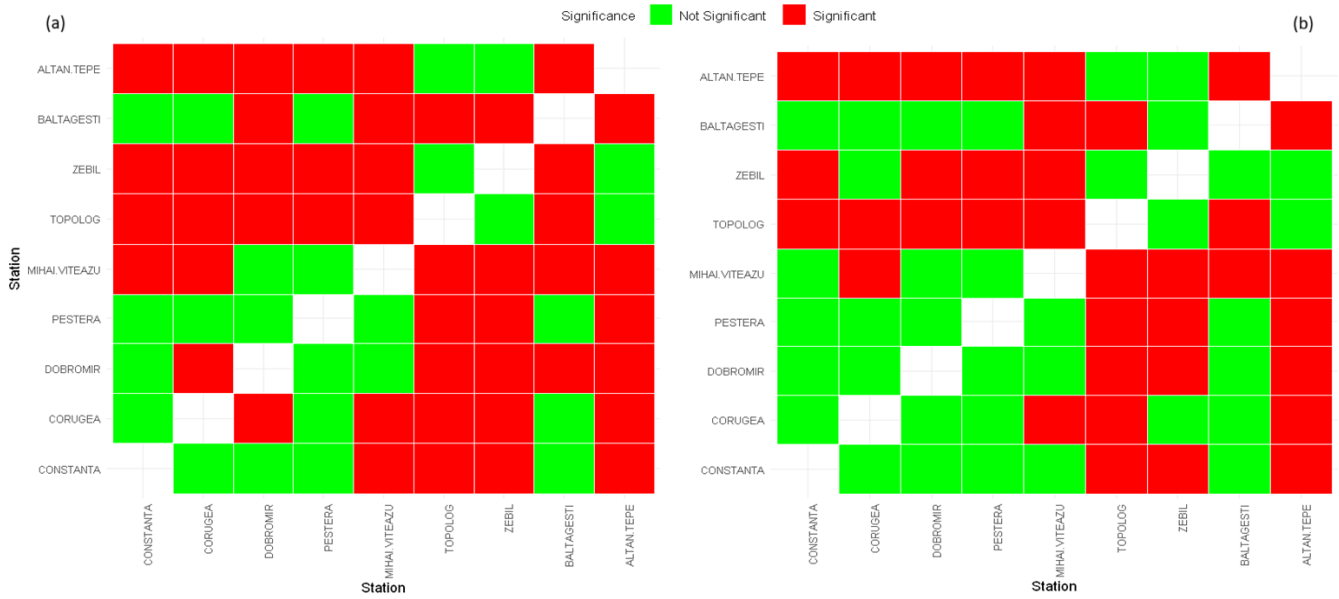

Figure 7. Results of (**a**) U and (**b**) Dunn post hoc tests of Welch's ANOVA. Green (red) cells indicate non-significant (significant) differences in mean.

In addition to the local analysis, the post hoc Dunn test of nonparametric Welch's ANOVA highlights the pairwise significance across non-normal stations, clarifying the station pairs that exhibit significant differences in median precipitation levels (the red cells in Figure 7b) and offering insight into broader regional patterns.

The consistency between local and global analyses is shown in Figure 8, revealing that stations with consistent results (represented by green cells) may reflect stable precipitation characteristics at various scales. In contrast, stations with discrepancies might highlight localized climatic or geographic factors that were not captured by broader assessments.

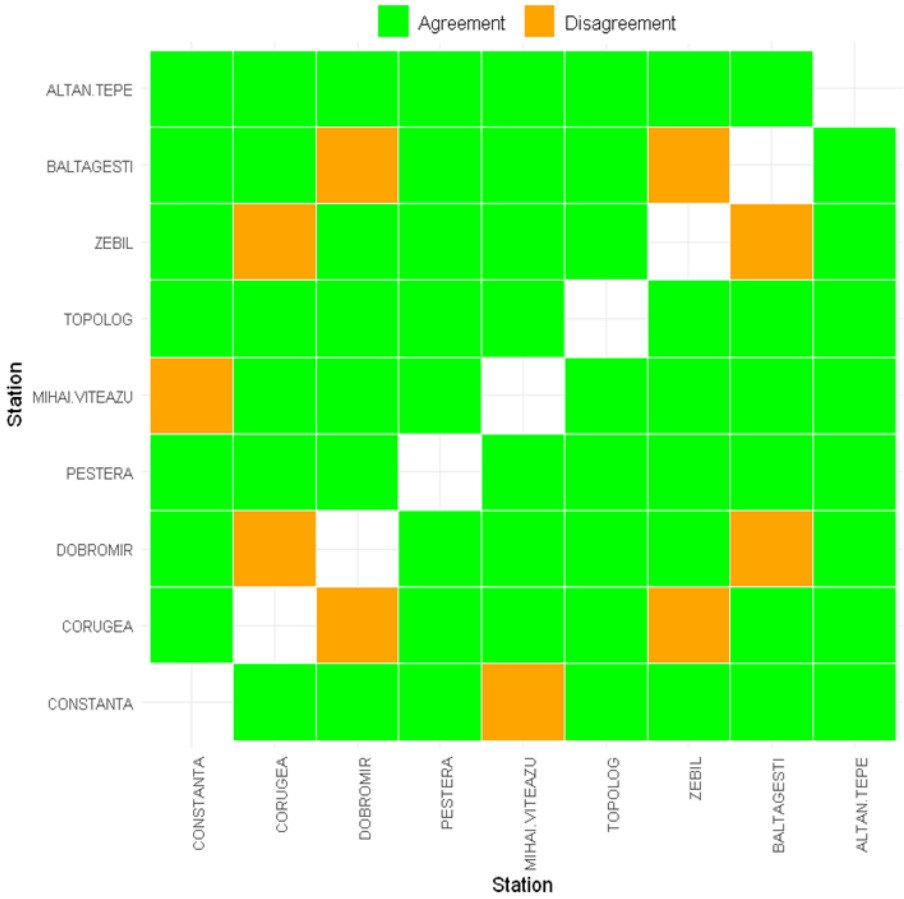

Figure 8. Comparisons of the results of the U-test and Dunn test. Green (orange) cells indicate concordance (discordance) between the results.

3.3. Homoscedasticity Analysis

The variance analysis of each series highlights that most of them fail to satisfy the hypotheses for parametric tests. Figure 9a shows that only Agigea and Jurilovca meet both normality and homoscedasticity hypotheses, as indicated by Levene's test result ($p > 0.05$).

Given that most stations do not satisfy homoscedasticity, the Brown–Forsythe test was applied for variance analysis on the transformed data series (Figure 9b). Most stations are shown in blue, indicating homoscedasticity.

A pairwise Brown–Forsythe test (Figure 9c) was conducted to address the homoscedasticity at the regional level. The results reveal non-significant differences across pairs, suggesting that variances are consistent on a localized basis throughout the dataset. This consistency aligns with the findings from the individual tests, supporting the homogeneity in variance.

To complement the pairwise analysis, the Fligner–Killeen test was used to assess overall differences in variance. A p-value = 0.854 did not reject the hypothesis of equality

of series variances. This result corroborates the findings from the local pairwise tests, affirming that variance differences are not statistically significant across the dataset.

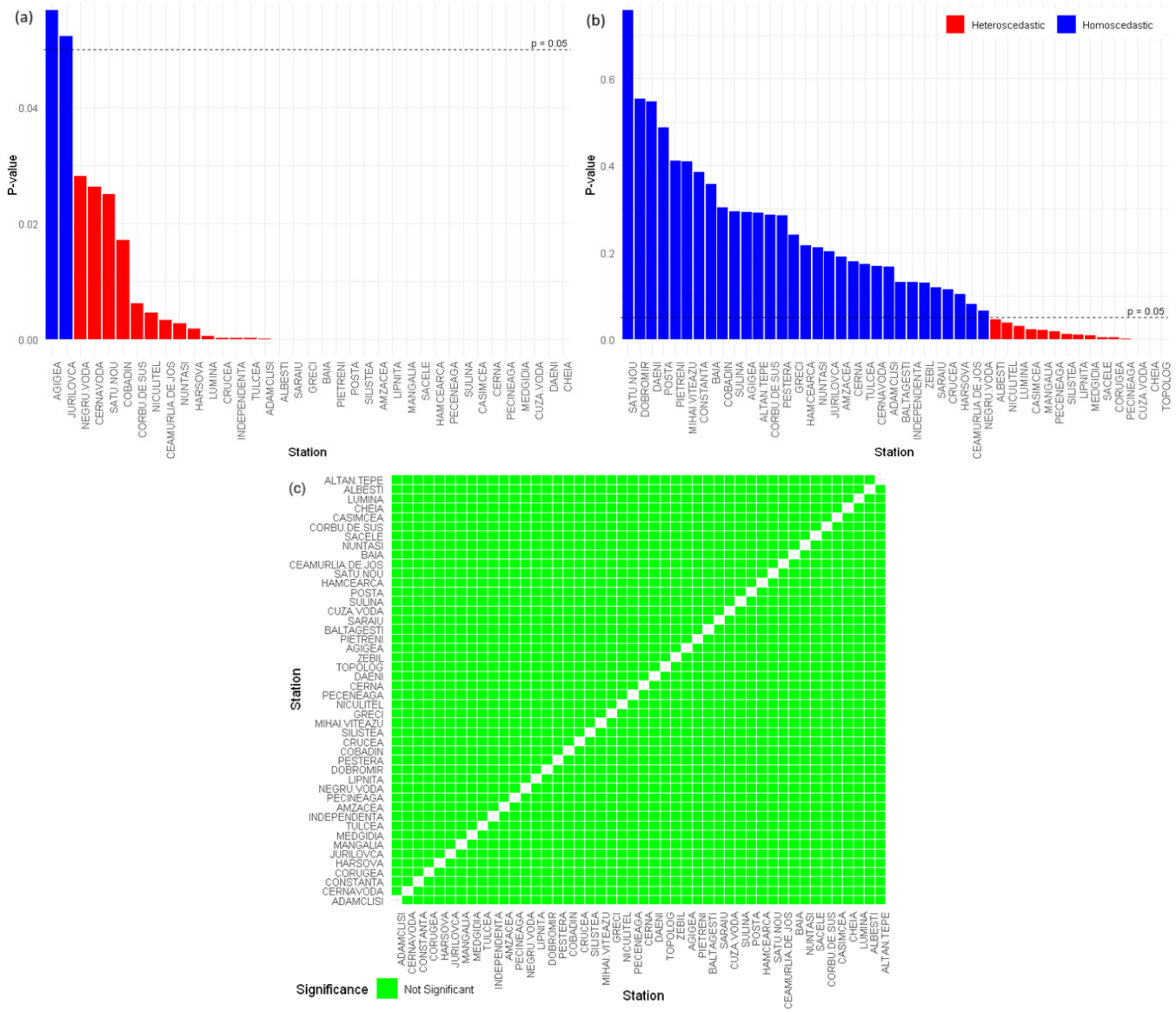

Figure 9. Results of (**a**) the Levene test on normal series and (**b**) the Brown–Forsythe test on transformed series. The null hypothesis was not rejected for the series represented in blue. (**c**) The pairwise Brown–Forsythe test on transformed series indicates homoscedasticity at regional level.

3.4. Comparison of the Analyses' Results on the Original and Transformed Data Series

Figure 10a shows the combined results of the *t*-test (on the Gaussian series) and the U-test (on non-normal series), on the transformed dataset. Figure 10b shows the U-test results for the original series, capturing the non-parametric pairwise comparison of precipitation means across all meteorological stations.

The consistency between the tests of mean equality performed on transformed versus original datasets is represented in Figure 11a, reflecting their agreement. Green (red) cells indicate agreement in non-significance (significance), and yellow cells indicate disagreement between the two analyses. The visual analysis in Figure 11b highlights the outcome of Dunn's test with Bonferroni correction on the original and transformed datasets, revealing the global non-parametric pairwise comparisons. Both datasets maintain nearly identical rank orders across the stations due to the nature of the transformation applied.

Figure 10. Results of (**a**) combined *t*- and U-tests on transformed data, and (**b**) U-test for original data. The red cells indicate significant differences between the means of the pairs of series.

Figure 11. (**a**) Comparison of the results of tests of mean equality on the transformed and original datasets, and (**b**) Dunn's test with Bonferroni correction for the original and the transformed dataset.

In non-parametric tests, like U and Dunn's tests, the ranks are the fundamental basis for statistical comparison. They check whether two or more groups differ in their central tendencies by analyzing the relative positions of observations rather than their absolute values. This approach allows comparisons to remain effective in non-normal distributions, reducing sensitivity to outliers and maintaining the focus on the overall ordering of data points. The results show that transforming the data did not alter its rank order. Thus, the relationships, as represented by ranks, remained unchanged.

We compared the Brown–Forsythe test output on the individual series to determine the transformation impact. Figure 12a provides an overlay map to assess the homoscedasticity consistency between the original and the transformed datasets.

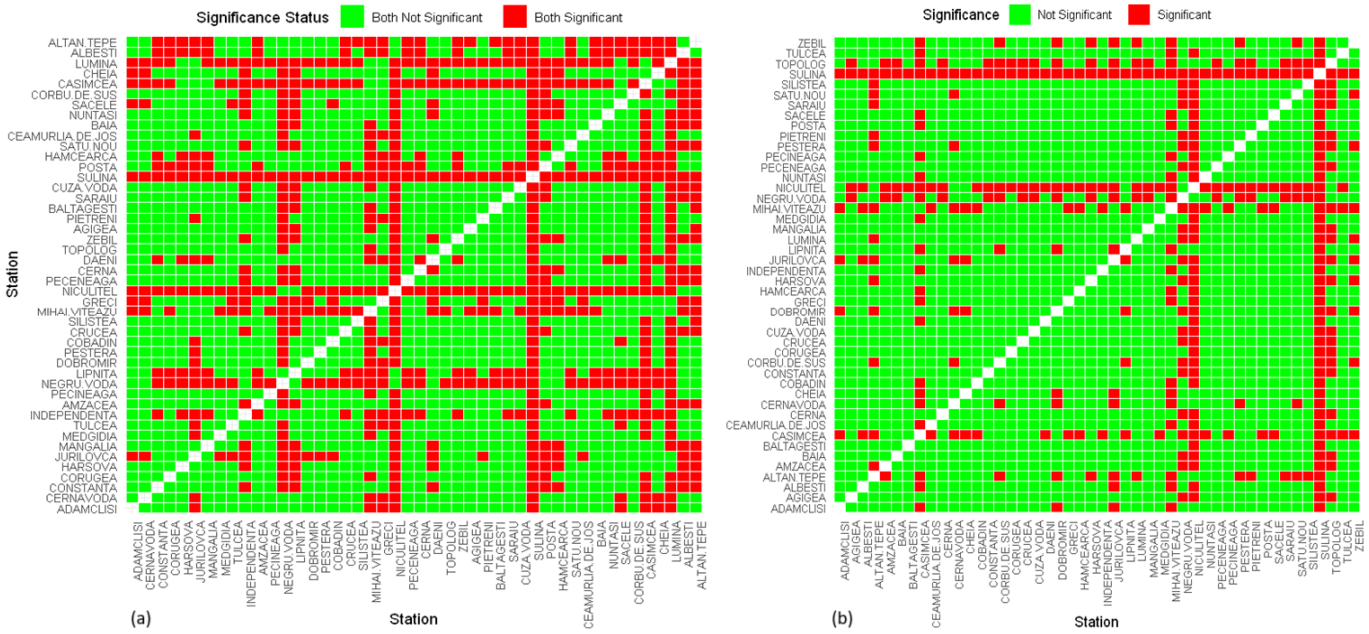

Figure 12. (**a**) Overlay of the Brown–Forsythe test results: local variance comparison for the transformed and original sets. The red (green) cells indicate that the tests on the two datasets are both significant (not significant). (**b**) Results of Dunn's test for the original (and transformed) datasets. The red cells indicate the series pairs for which the hypothesis of variances' equality was rejected.

The results reveal that the transformation either preserved or did not impact homoscedasticity for most series pairs. Specifically, consistent regions of significance or non-significance are highlighted in red and green. The absence of yellow regions indicates no discrepancies between the original and transformed datasets. This outcome is directly related to preserving rank order during the transformation process. By maintaining the rank structure, the transformation ensured that relationships in variance were consistent across both datasets. This visual summary supports the stability and effectiveness of the transformation in preserving fundamental variance relationships.

We extended the homoscedasticity evaluation to the Regional series. We applied the Flinger–Killeen test to both datasets—before and after transformation. Since the homoscedasticity hypothesis was rejected, Dunn's test was used to identify the groups whose variances were unequal. Figure 12b shows the results on the original dataset, noting that the transformed dataset produced the same heatmap. The similarity between the results on both datasets is due to the rank preservation throughout the transformation process.

In conclusion, the above analysis indicates that working on the initial or transformed data series will not affect further analysis, which consists of the change point detection, which assesses the abrupt changes in the series mean. The same is true for the change point in variance, which will be discussed in another article.

3.5. Change Points Detection with MDL and Comparison with Known Results

To validate the proposed MDL framework, we initially applied the methodology to synthetic time series with pre-specified change points. These test datasets, generated under normality assumptions, allowed the evaluation of the algorithm's performance in identifying known change point locations.

Across 1,000 optimization runs, we tested series with no change points ($m = 0$), with one ($m = 1$) and two change points ($m = 2$). The results' accuracy was 99.99%, 99.70%, and 98.80%, respectively, confirming the capability of the proposed framework to reliably detect both the number and location of change points in controlled conditions. Following this

validation, we applied the framework to the real precipitation series from the ten main meteorological stations—Adamclisi, Cernavodă, Constanța, Corugea, Hârșova, Jurilovca, Mangalia, Medgidia, Sulina, and Tulcea, and two secondary stations—Peceneaga and Cheia. The change points of the first ten series are presented in Table 1.

Table 1. Change points determined by MDL.

Station	MDL
Adamclisi	Feb. 1965, Jul. 1994, Dec. 1999
Cernavodă	Feb. 1965, Nov. 1965, Dec. 1994, May 2004
Constanța	Feb. 1965, May 2004
Corugea	Feb. 1967, Jun. 1970, May 2004
Hârșova	Dec. 1965, Feb. 1978, May 2004
Jurilovca	Mar. 1995, Aug. 1998
Mangalia	Feb. 1965, Jun. 1968, Sep. 1984, Jul. 1994
Medgidia	Nov. 1965, May 2004
Sulina	Jul. 1976, May 2004
Tulcea	Mar. 1967, May 2004

Whereas February is situated at the beginning of the series, and the corresponding values might be extreme precipitation values, the other moments are change points. Figure 13 shows the detected change points (red dots) and the corresponding shifts in the mean segments for the stations in Table 1.

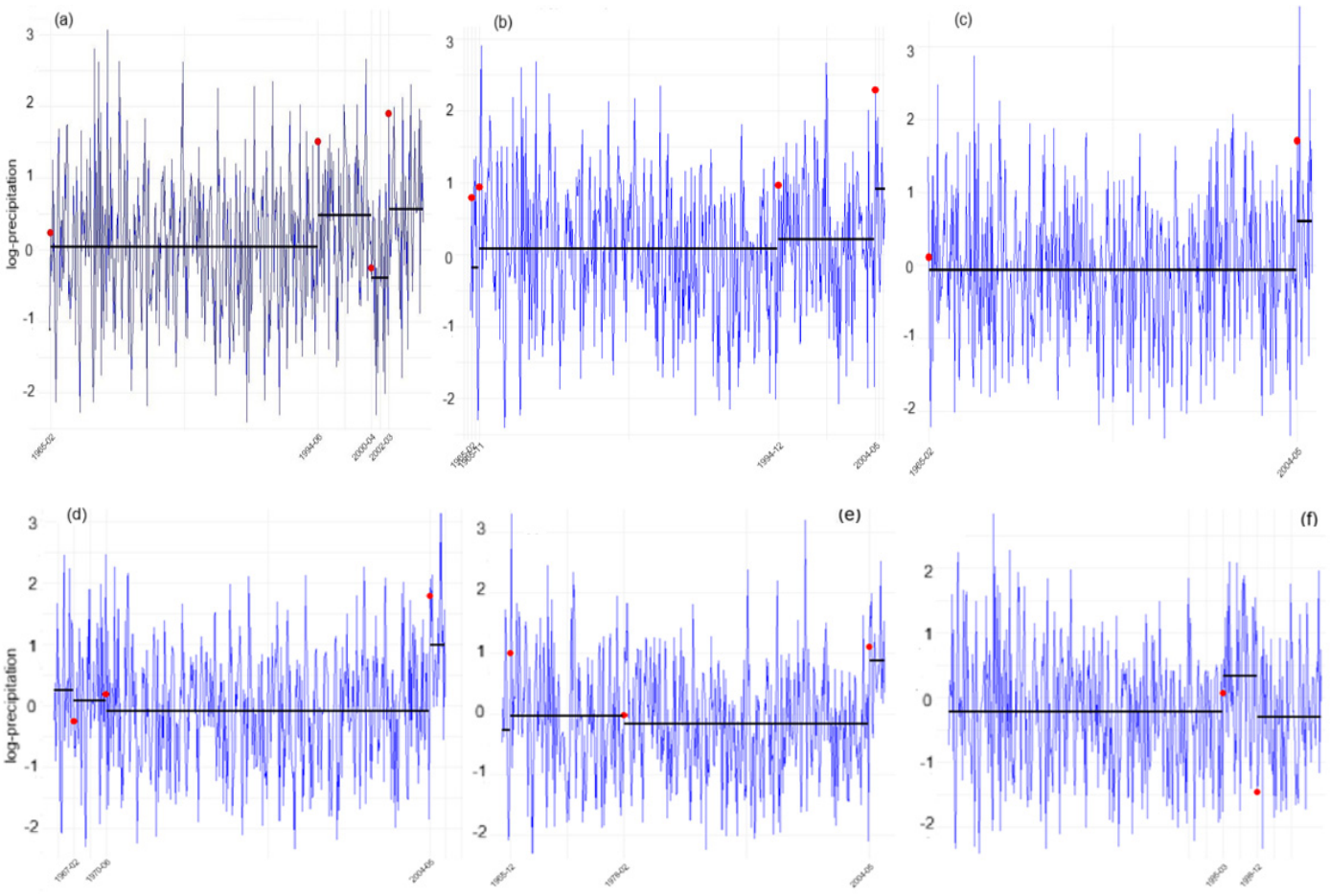

Figure 13. *Cont.*

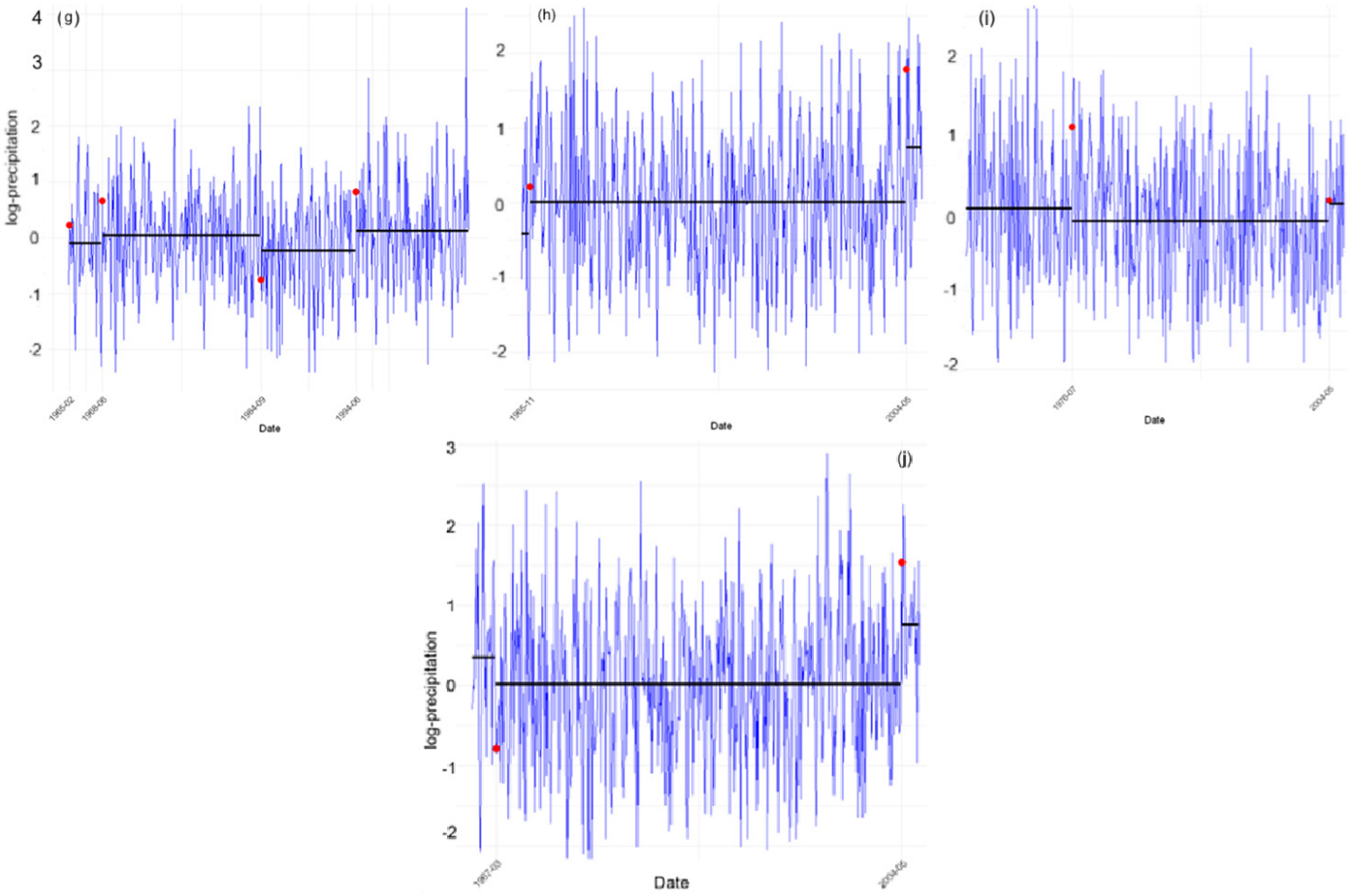

Figure 13. Change points of (**a**) Adamclisi, (**b**) Cernavodă, (**c**) Constanța, (**d**) Corugea, (**e**) Hârșova, (**f**) Jurilovca, (**g**) Mangalia, (**h**) Medgidia, (**i**) Sulina, (**j**) Tulcea series.

Each red dot represents the exact temporal location where the shift was identified, while the horizontal black segments delineate the mean precipitation levels before and after the detected change points.

4. Discussion

The study presented an extended analysis of the statistical features of 46 monthly precipitation total from the Dobrogea region. A key advantage lies in its multifaceted approach, which integrates mean and variance analyses, data transformations, and change point detection. These components can improve statistical inferences while addressing the inherent complexities of climatic data.

Parametric and non-parametric statistical tests were performed to emphasize the mean and variance differences. Specifically, the Lilliefors test addressed the normality of the initial and transformed series. Based on its output, the homoscedasticity of the individual series and that of the Regional one was tested by the Levene and Brown–Forsythe tests. For datasets that satisfy assumptions of normality and homoscedasticity, parametric methods such as ANOVA and the Student–Newman–Keuls test were applied, allowing the detailed exploration of regional precipitation patterns. For datasets where these assumptions did not hold, non-parametric alternatives like the Mann–Whitney U and Fligner–Killeen tests ensured reliable inferences. This dual-framework approach highlights the study's adaptability and underscores the importance of methodological rigor in assessing precipitation variability across heterogeneous datasets.

Figure 14 presents the stations and their similarity in mean and variance (resulting from the tests presented in detail in the previous sections).

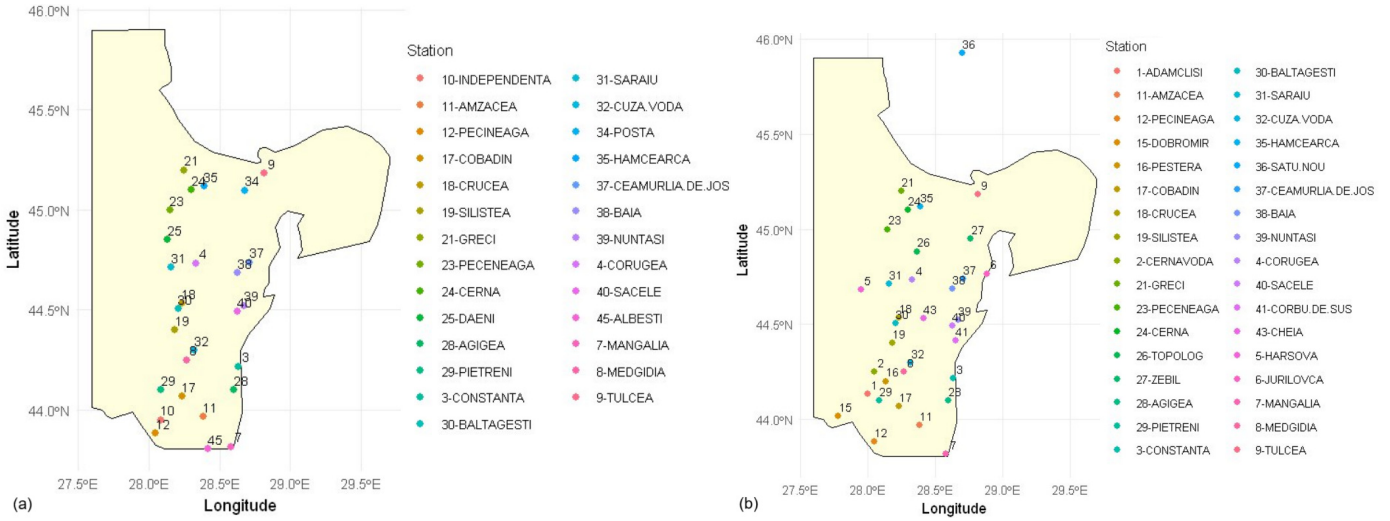

Figure 14. Series with (**a**) similar mean values, and (**b**) similar variance.

The spatial distribution of stations with similar mean values reveals some groupings that could reflect an underlying regional climatic pattern. Stations such as Greci, Peceneaga, Cerna, and Dăeni are concentrated in the inner region toward the northern-left part of the map and exhibit a similar average precipitation. This remark suggests some localized climatic conditions affecting the precipitation patterns, possibly tied to the proximity to the northern inland zone. On the southeast coast, Albești and Mangalia have similar means, showing the coastal influences in this region, where the Black Sea likely moderates precipitation. Analogous Constanța and Agigea, located along the coastline, present similar rainfall averages. In the southern part of Dobrogea, Independența, Amzacea, Pecineaga, and Cobadin also present similar patterns that can be related to continental inland factors.

Figure 13 indicates the presence of detectable climatic shifts, showing that the MDL-based methodology captures change points in climatic datasets. Several validation approaches could strengthen the reliability of these findings. However, the classical Buishand and Lee & Heghinian tests could not be performed for the Constanța series (because the series normality could not be reached by transformations). The Buishand and Pettitt tests (performed at a significance level of 5%) rejected the hypothesis that there is no breakpoint for Adamclisi and Sulina. The Lee &Heghinian test found May 2004 to be a breakpoint for six out of ten main series, confirming our findings. Relying on the absence of the data series, CUSUM cannot always be applied. Therefore, other methods should be explored to validate the MDL results.

In the MDL framework, the penalty scale is the parameter that governs the trade-off between fitting the observed data and penalizing model complexity. A lower penalty scale allows the model to add more change points, prioritizing an exact fit to the data but risking the inclusion of spurious changes driven by noise. Conversely, a higher penalty scale discourages adding unnecessary complexity, focusing on capturing only the most prominent structural shifts in the series. Therefore, striking the right balance is mandatory. The penalty scale operates indirectly, affecting the optimization process without providing real-time feedback on its suitability. Here, the residuals come into play. After fitting, the residuals can confirm if the chosen scale let to an appropriate balance. For example, if the residuals are very small and exhibit no patterns, the model might have been overfitted, suggesting that the scale was too low.

The presence of patterns or trends (e.g., periodic behavior or long-time drifts) could indicate underfitting, the use of a penalty scale too large, that prevented the model from capturing important features. Thus, the penalty scale's adequacy can be tested by utilizing a diagnostic tool like the residuals' properties such as normality, autocorrelation, and homoscedasticity). Residuals that align with theoretical assumptions indicate the model performance. For instance, a failure of residuals to pass normality or homoscedasticity tests will indicate the need for a higher penalty scale to simplify the model. Conversely, significant autocorrelation in residuals might point to an un-modeled structure in the data requiring a lower scale.

While residuals are invaluable for post-optimization validation, their dependence on the model fit limits their role during optimization. They cannot directly guide the choice of penalty scale in real time because they are outcomes of the optimization process. However, residual analysis can inform on retrospective refinement of the penalty scale.

The change point detection methodology, grounded in the Minimum Description Length (MDL) principle, is adapted for precipitation (or climate in general) time series by incorporating seasonal periodicities and autocorrelation. Using a metaheuristic approach (Simulated Annealing) for optimizing the MDL criterion, this study provides a precise segmentation of time series, identifying significant structural shifts (in mean) that are indicative of both climatic variability and potential observational inhomogeneity. These change points are the cornerstone for homogenizing precipitation datasets, as undetected inhomogeneity can distort analyses of long-term climatic trends and variability.

The findings of this study have direct implications for water resource management and climate adaptation strategies. For instance, identifying mean shifts and variance changes in precipitation patterns aids in anticipating shifts in water availability, enabling stakeholders to design adaptive infrastructure to mitigate the impacts of extreme events. Moreover, the segmented precipitation series provides a reliable foundation for regional-scale hydrological and climatological modeling.

Future research directions include incorporating metadata, such as station histories, to validate detected change points and distinguish between artificial and natural shifts in precipitation patterns. Extending the MDL framework to probabilistic models, as suggested by Lu et al. [67] could enhance the process of change point detection by quantifying uncertainties and integrating expert knowledge into the segmentation process. Additionally, coupling break points' detection with spatial clustering methods could be a factor in understanding the geographical distribution of precipitation variability, elucidating links between climatic factors and regional precipitation dynamics.

Another direction for future work is exploring Bayesian extensions to quantify uncertainties and enhance methodological precision. Unlike traditional methods, Bayesian techniques incorporate prior knowledge and provide posterior distributions for parameters and change point locations. They offer a probabilistic view of the segmentation process, allow the detection of change points, and quantify the uncertainty associated with their identification. These models can incorporate hierarchical structures to account for known climatic processes, leading to more context-aware and interpretable results.

5. Conclusions

This study offers a statistical evaluation of precipitation variability in Dobrogea, using data from 46 meteorological stations spanning 40 years (1965–2005). Following the normality assessment by the Lilliefors test, it was found that most series are not Gaussian. Therefore, a Yeo–Johnson transformation was employed to transform the non-normal series into normal ones to apply parametric methods for testing the individual series'

homoscedasticity, equality in means, and the variance of series pairs. Non-parametric tests were employed for the non-normal series.

The *t*-test indicates that some groups of pairs of series exhibit consistent non-significant differences in means. The SNK post hoc test after the one-way ANOVA for normal stations indicates some clusters with relatively homogenous average precipitation levels. The U and Dunn tests on the non-Gaussian series provided similar results. All results reveal high spatial heterogeneity in precipitation means. However, the Brown–Forsythe test output shows that variances are consistent on a localized basis throughout the dataset. Heatmaps used to illustrate pairwise and global test results revealed regional groupings of stations that might be affected by common climatic factors.

The MDL framework, optimized via SA, identified change points in both synthetic and real datasets. The validation with synthetic data demonstrated the framework's precision in detecting pre-specified change points under controlled conditions. Furthermore, applying it to twelve precipitation series revealed significant shifts in mean, such as those in 1994 and 2000 for Adamclisi and 1997 and 2004 for Cheia, reflecting structural changes in precipitation regimes.

On the other hand, to validate detected change points and differentiate between real climatic shifts and observational artifacts, it is important to integrate metadata such as station relocation histories and instrumentation changes. Future work will explore Bayesian extensions to quantify uncertainties further and enhance methodological precision. Complementary change point detection methods could further enhance the robustness of findings. Comparisons with the output of other methods, such as the Hubert segmentation procedure [74], AUG-Segmenter [75], and mDP [76], will also be performed. The findings are necessary for a correct assessment of precipitation variability and the impact of climate change on different regions, particularly in Dobrogea.

The heatmaps' clustering patterns show potential directions for future study. These groupings could be quantitatively confirmed using formal clustering techniques like k-means, hierarchical clustering, or DBSCAN.

One of the main drawbacks of the analysis is the use of data before 2005, limiting the relevance of the results for the current agricultural or water management applications. After obtaining new data, which is a work in progress, we shall update the analysis, and provide comparisons with the new results, emphasizing the climate change extent in the study area.

Expanding the analysis to include spatial modeling would provide a broader context for these results, strengthening their relevance for regional water resource management and climate adaptation strategies. Together, these efforts would enhance the reliability of precipitation studies and support informed decision-making in addressing the challenges posed by increasing climatic variability in semi-arid environments like Dobrogea.

Author Contributions: Conceptualization, Y.S. and A.B.; methodology, Y.S. and A.B.; software, Y.S. and A.B.; validation, Y.S. and A.B.; formal analysis, Y.S. and A.B.; investigation, Y.S. and A.B.; resources, A.B.; data curation, A.B.; writing—original draft preparation, Y.S. and A.B.; writing—review and editing, A.B.; visualization, Y.S. and A.B.; supervision, A.B.; project administration, A.B.; funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data will be available on request from the second author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

We provide an explanation, while avoiding rigorous mathematics, of the Minimum Description Length (MDL) principle used in our work for the detection and selection of change points in time series data. The explanation includes parameter estimation, penalty formulation, and the optimization process.

1. Parameter Estimation

The parameters of the model in Equation (15), including the seasonal means (μ), trend (α), level shifts (δ), and periodic residuals (ε), were estimated through a two-step process, following the methodology from [67]: Ordinary Least Squares (OLS) and Generalized Least Squares (GLS).

- OLS provides an initial estimate of the parameters, ignoring periodicity and autocorrelation. It minimizes the residual sum of squares, offering a straightforward solution.
- GLS refines the OLS estimates, adding periodic autoregressive residuals (PAR(p)) and accounting for autocorrelation within residuals, seasonal effects, and level shifts at change points. It minimizes the residual variance iteratively under the assumption of periodic stationarity. This step requires a few iterations to converge.

Remark. This iterative refinement aligns closely with the approach described by Lu et al. [67], wherein the OLS estimates serve as a starting point for GLS optimization. Importantly, this stage does not involve penalty terms; it optimizes the model fit to the data series.

2. MDL Score Computation

The MDL score evaluates the fit and complexity of proposed change point configurations (so, here we always start from known change points configurations). It is given by:

$$\text{MDL(M)} = -\ell(\theta) + \text{Penalties}$$

where

- $\ell(\theta)$: Negative log-likelihood of the fitted model;
- Penalties: Terms to discourage overly complex models.

The penalty terms include the following:

- Number of change points (m): Penalizes excessive change points;
- Spacing of change points ($\tau_i - \tau_{i-1}$): Ensures sufficient separation;
- Dimensionality of seasonal means, trend, and autoregressive parameters.

3. Penalty Formulation

Penalties in the MDL framework are derived from information theory, accounting for both real and integer parameters:

- Real-valued parameters (e.g., μ, α) require $log_2(N)/2$ bits to encode;
- Integer-valued parameters (e.g., τ_j) require $log_2\tau_j$ bits to encode.

These penalties ensure a balance between model complexity and fidelity.

4. Optimization Process

The MDL framework evaluates multiple change point configurations. The optimization ensures that the best-fit model is identified while avoiding overfitting. It involves the following:

- Proposing combinations of change points (e.g., number and location);
- Fitting the model to each configuration using OLS and GLS to compute $\ell(\theta)$;

- Adding penalty terms to compute the MDL score for each configuration;
- Selecting the configuration with the minimum MDL score.

Remark: Distinction from Tikhonov Regularization

While MDL penalties may appear to be analogous to regularization methods like Tikhonov, their purpose and derivation differ:

- Tikhonov Regularization:
  - Stabilizes ill-posed problems during the parameters estimation;
  - Uses penalties (e.g., $\lambda\|\theta\|^2$) to control parameter magnitudes.
- MDL Penalties:
  - Are applied after model fitting, during model selection;
  - Are derived from coding theory, quantifying the cost of encoding the model and residuals.

Conclusion. The MDL framework balances model fit and complexity through a rigorous process of parameter estimation, penalty formulation, and optimization. By systematically evaluating change point configurations, it identifies parsimonious models that capture the essential structure of the data.

References

1. Douville, H.; Raghavan, K.; Renwick, J.; Allan, R.P.; Arias, P.A.; Barlow, M.; Cerezo-Mota, R.; Cherchi, A.; Gan, T.Y.; Gergis, J.; et al. Water cycle changes. 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*; 2021. Available online: https://www.ipcc.ch/report/ar6/wg1/chapter/chapter-8/ (accessed on 29 November 2024).
2. Marsh, T.; Parry, S.; Kendon, J.; Hannaford, M. The 2010-12 Drought and Subsequent Extensive Flooding: A Remarkable Hydrological Transformation. Available online: https://nora.nerc.ac.uk/id/eprint/503643/1/N503643CR.pdf (accessed on 29 November 2024).
3. Berényi, A.; Bartholy, J.; Pongrácz, R. Analysis of precipitation-related climatic conditions in European plain regions. *Weather Clim. Extrem.* **2023**, *42*, 100610. [CrossRef]
4. Klein Tank, M.G.; Können, G.P. Trends in indices of daily temperature and precipitation extremes in Europe, 1946–1999. *J. Clim.* **2003**, *16*, 3665–3680. [CrossRef]
5. Zolina, O.; Simmer, C.; Belyaev, K.; Kapala, A.; Gulev, S. Improving estimates of heavy and extreme precipitation using daily records from European rain gauges. *J. Hydrometeorol.* **2009**, *10*, 701–716. [CrossRef]
6. Jiang, Y.; He, X.; Li, J.; Zhang, X. On the response of daily precipitation extremes to local mean temperature in the Yangtze River basin. *Atmos. Res.* **2024**, *300*, 107265. [CrossRef]
7. Fowler, H.J.; Lenderink, G.; Prein, A.F.; Westra, S.; Allan, R.P.; Ban, N.; Barber, R.; Berg, P.; Blenkinsop, S.; Do, H.X.; et al. Anthropogenic intensification of short-duration rainfall extremes. *Nat. Rev. Earth Environ.* **2021**, *2*, 107–122. [CrossRef]
8. Szymczak, S.; Backendorf, F.; Bott, F.; Fricke, K.; Junghänel, T.; Walawender, E. Impacts of Heavy and Persistent Precipitation on Railroad Infrastructure in July 2021: A Case Study from the Ahr Valley, Rhineland-Palatinate, Germany. *Atmosphere* **2022**, *13*, 1118. [CrossRef]
9. Rufat, S.; Tate, E.; Burton, C.G.; Maroof, A.S. Social vulnerability to floods: Review of case studies and implications for measurement. *Int. J. Disast. Risk Reduct.* **2015**, *14*, 470–486. [CrossRef]
10. Bărbulescu, A.; Dumitriu, C.S.; Maftei, C. On the Probable Maximum Precipitation Method. *Rom. J. Phys.* **2022**, *67*, 801.
11. Wainwright, C.; Marsham, J.H.; Rowell, D.P.; Finney, D.L.; Black, E. Future Changes in Seasonality in East Africa from Regional Simulations with Explicit and Parameterized Convection. *J. Clim.* **2021**, *34*, 1367–1385. [CrossRef]
12. Li, L.; Zheng, Z.; Biederman, J.A.; Xu, C.; Xu, Z.; Che, R.; Wang, Y.; Cui, X.; Hao, Y. Ecological responses to heavy rainfall depend on seasonal timing and multi-year recurrence. *New Phytol.* **2019**, *223*, 647–660. [CrossRef] [PubMed]
13. Knapp, A.K.; Fay, P.A.; Blair, J.M.; Collins, S.L.; Smith, M.D.; Carlisle, J.D.; Harper, C.W.; Danner, B.T.; Lett, M.S.; McCarron, J.K. Rainfall variability, carbon cycling, and plant species diversity in a mesic grassland. *Science* **2002**, *298*, 2202–2205. [CrossRef] [PubMed]
14. Raymond, C.; Horton, R.M.; Zscheischler, J.; Martius, O.; AghaKouchak, A.; Balch, J.; Bowen, S.G.; Camargo, S.J.; Hess, J.; Kornhuber, K.; et al. Understanding and managing connected extreme events. *Nat. Clim. Change* **2020**, *10*, 611–621. [CrossRef]
15. Kalkuhl, M.; Wenz, L. The impact of climate conditions on economic production. Evidence from a global panel of regions. *J. Environ. Econ. Manag.* **2020**, *103*, 102360. [CrossRef]

16. Zeder, J.; Fischer, E.M. Observed extreme precipitation trends and scaling in Central Europe. *Weather Clim. Extrem.* **2020**, *29*, 100266. [CrossRef]
17. Huang, H.; Winter, J.M.; Osterberg, E.C.; Horton, R.M.; Beckage, B. Total and Extreme Precipitation Changes over the Northeastern United States. *J. Hydrometeor.* **2017**, *18*, 1783–1798. [CrossRef]
18. Huang, H.; Winter, J.M.; Osterberg, E.C. Mechanisms of Abrupt Extreme Precipitation Change over the Northeastern United States. *J. Geophys. Res.-Atmos.* **2018**, *123*, 7179–7192. [CrossRef]
19. Kunkel, K.E.; Karl, T.R.; Brooks, H.; Kossin, J.; Lawrimore, J.H.; Arndt, D.; Bosart, L.; Changnon, D.; Cutter, S.L.; Doesken, N.; et al. Monitoring and Understanding Trends in Extreme Storms: State of Knowledge. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 499–514. [CrossRef]
20. Groisman, P.Y.; Knight, R.W.; Karl, T.R.; Easterling, D.R.; Sun, B.; Lawrimore, J.H. Contemporary changes of the hydrologic cycle over the contiguous United States: Trends derived from in situ observations. *J. Hydrometeor.* **2004**, *5*, 64–85. [CrossRef]
21. Westra, S.; Alexander, L.V.; Zwiers, F.W. 2013: Global increasing trends in annual maximum daily precipitation. *J. Clim.* **2013**, *26*, 3904–3918. [CrossRef]
22. Hoerling, M.; Eischeid, J.; Perlwitz, J.; Quan, X.; Wolter, K.; Cheng, L. Characterizing Recent Trends in U.S. Heavy Precipitation. *J. Clim.* **2016**, *29*, 2313–2332.
23. Jong, B.-T.; Delworth, T.L.; Cooke, W.F.; Tseng, K.-C.; Murakami, H. Increases in extreme precipitation over the Northeast United States using high-resolution climate model simulations. *npj Clim. Atmos. Sci.* **2023**, *6*, 18. [CrossRef]
24. Zveryaev, I.I. Seasonality in precipitation variability over Europe. *J. Geophys. Res. Atmos.* **2004**, *109*, D05103. [CrossRef]
25. Qian, B.; Xu, H.; Corte-Real, J. Spatial-temporal structures of quasi-periodic oscillations in precipitation over Europe. *Int. J. Climatol.* **2000**, *20*, 1583–1598. [CrossRef]
26. Croitoru, A.-E.; Piticar, A.; Burada, D.C. Changes in precipitation extremes in Romania. *Quat. Int.* **2016**, *415*, 325–335. [CrossRef]
27. Busuioc, A.; von Storch, H. Changes in the winter precipitation in Romania and its and its relation to the large-scale circulation. *Tellus* **1996**, *48*, 538–552. [CrossRef]
28. Bărbulescu, A.; Maftei, C.E. Evaluating the Probable Maximum Precipitation. Case study from the Dobrogea region, Romania. *Rom. Rep. Phys.* **2023**, *75*, 704. [CrossRef]
29. Prăvălie, R.; Piticar, A.; Roșca, B.; Sfîcă, L.B.; Tiscovschi, A.; Patriche, C. Spatiotemporal changes of the climatic water balance in Romania as a response to precipitation and reference evapotranspiration trends during 1961–2013. *Catena* **2019**, *172*, 295–312. [CrossRef]
30. Bosneagu, R.; Lupu, C.E.; Torica, E.; Lupu, S.; Vatu, N.; Tanase, V.M.; Vasilache, C.; Daneci-Patrau, D.; Scurtu, I.C. Long-term analysis of air temperatures variability and trends on the Romanian Black Sea Coast. *Acta Geophys.* **2022**, *70*, 2179–2197. [CrossRef]
31. Păltineanu, C.; Mihăilescu, I.; Seceleanu, I.; Dragotă, C.; Vasenciuc, F. Using aridity indices to describe some climate and soil features in Eastern Europe: A Romanian case study. *Theor. Appl. Climatol.* **2007**, *90*, 263–274. [CrossRef]
32. Bandoc, G.; Prăvălie, R. Climatic water balance dynamics over the last five decades in Romania's most arid region, Dobrogea. *J. Geogr. Sci.* **2015**, *25*, 1307–1327. [CrossRef]
33. Bărbulescu, A.; Deguenon, J. Change point detection and models for precipitation evolution. Case study. *Rom. J. Phys.* **2014**, *59*, 590–600.
34. Bărbulescu, A.; Deguenon, J. About the variations of precipitation and temperature evolution in the Romanian Black Sea Littoral. *Rom. Rep. Phys.* **2015**, *67*, 625–637.
35. Maftei, C.; Bărbulescu, A. Statistical analysis of climate evolution in Dobrudja region. In Proceedings of the World Congress on Engineering 2008 (WCE 2008), London, UK, 2–4 July 2008; Volume II, pp. 1082–1087.
36. Bărbulescu, A. A new method for estimation the regional precipitation. *Water Resour. Manag.* **2016**, *30*, 33–42. [CrossRef]
37. Maftei, C.; Bărbulescu, A.; Buta, C.; Șerban, C. Change Points Detection and Variability Analysis of Some Precipitation Series. Available online: https://dl.acm.org/doi/10.5555/2039846.2039887 (accessed on 10 November 2024).
38. Bandoc, G.; Prăvălie, R.; Patriche, C.; Dragomir, E.; Tomescu, M. Response of phenological events to climate warming in the southern and south-eastern regions of Romania. *Stoch. Environ. Res. Risk A* **2018**, *32*, 1113–1129. [CrossRef]
39. Bandoc, G.; Piticar, A.; Patriche, C.; Roșca, B.; Dragomir, E. Climate Warming-Induced Changes in Plant Phenology in the Most Important Agricultural Region of Romania. *Sustainability* **2022**, *14*, 2776. [CrossRef]
40. Dumitriu, C.S.; Bărbulescu, A.; Maftei, C. IrrigTool—A New Tool for Determining the Irrigation Rate Based on Evapotranspiration Estimated by the Thornthwaite Equation. *Water* **2022**, *14*, 2399. [CrossRef]
41. Dallal, G.E.; Wilkinson, L. An analytic approximation to the distribution of Lilliefors's test statistic for normality. *Am. Stat.* **1986**, *40*, 294–296. [CrossRef]
42. Yeo, I.N.K.; Johnson, R.A. A new family of power transformations to improve normality or symmetry. *Biometrika* **2000**, *87*, 954–959. [CrossRef]
43. Helsel, D.R.; Hirsch, R.M.; Ryberg, K.R.; Archfield, S.A.; Gilroy, E.J. *Statistical Methods in Water Resources*; U.S. Geological Survey: Reston, VA, USA, 2020. Available online: https://pubs.usgs.gov/publication/tm4A3 (accessed on 2 December 2024).

44. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 4th ed.; Elsevier: Amsterdam, The Netherlands, 2020.
45. Gatz, D.F.; Smith, L. The standard error of a weighted mean concentration-I. Bootstrapping vs other methods. *Atmos. Environ.* **1995**, *29*, 1185–1193. [CrossRef]
46. Box, G.E.P. Non-normality and tests on variances. *Biometrika* **1953**, *40*, 318–335. [CrossRef]
47. Nachar, N. The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. *Tutor. Quant. Meth. Psychol.* **2008**, *4*, 13–20. [CrossRef]
48. Fligner, M.A. A class of two-sample distribution-free tests for scale. *J. Am. Stat. Assoc.* **1979**, *74*, 889–893. [CrossRef]
49. Student. The Probable Error of a Mean. *Biometrika* **1908**, *6*, 1–25. [CrossRef]
50. Fisher, R.A. Statistical Methods for Research Workers. In *Breakthroughs in Statistics*; Kotz, S., Johnson, N.L., Eds.; Springer: New York, NY, USA, 1992. [CrossRef]
51. Abdi, H.; Williams, L.J. Newman-Keuls Test and Tukey Test. Available online: https://personal.utdallas.edu/~herve/abdi-NewmanKeuls2010-pretty.pdf (accessed on 12 October 2024).
52. Mann, H.B.; Whitney, D.R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [CrossRef]
53. Fay, M.P.; Proschan, M.A. Wilcoxon–Mann–Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Stat. Surv.* **2010**, *4*, 1–39. [CrossRef]
54. Welch, B.L. On the Comparison of Several Mean Values: An Alternative Approach. *Biometrika* **1951**, *38*, 330–336. [CrossRef]
55. Dunn, O.J. Multiple Comparisons Using Rank Sums. *Technometrics* **1964**, *6*, 241–252. [CrossRef]
56. Levene, H. Robust tests for equality of variances. In *Contributions to Probability and Statistics*; Olkin, I., Ed.; Stanford University Press: Palo Alto, CA, USA, 1960; pp. 278–292.
57. Brown, M.B.; Forsythe, A.B. Robust tests for the equality of variances. *J. Am. Stat. Assoc.* **1974**, *69*, 362–367. [CrossRef]
58. Conover, W.J.; Johnson, M.E.; Johnson, M.M. A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* **1981**, *23*, 351–361. [CrossRef]
59. Pettitt, A.N. A non—Parametric approach to the change-point problem. *Appl. Stat.* **1979**, *28*, 126–135. [CrossRef]
60. Buishand, T.A. Some methods for testing the homogeneity of rainfall records. *J. Hydrol.* **1982**, *58*, 11–27. [CrossRef]
61. Buishand, T.A. Tests for detecting a shift in the mean of hydrological time series. *J. Hydrol.* **1984**, *58*, 51–69. [CrossRef]
62. Lee, A.F.S.; Heghinian, M.S. A Shift of the Mean Level in a Sequence of Independent Normal Random Variables—A Bayesian Approach. *Technometrics* **1977**, *19*, 503–506.
63. Boyer, J.F. Logiciel Khronostat D'analyse Statistiquede Séries Chronologique. 2002. Available online: http://www.hydrosciences.org/mytech/khronostat.html (accessed on 21 October 2024).
64. Horváth, L.; Rice, G. *Change Point Analysis for Time Series*; Springer Nature: Cham, Switzerland, 2024.
65. Rissanen, J. Modeling by shortest data description. *Automatica* **1978**, *14*, 465–471. [CrossRef]
66. Davis, R.A.; Lee, T.C.M.; Rodriguez-Yam, G.A. Structural break estimation for nonstationary time series models. *J. Am. Stat. Assoc.* **2006**, *101*, 223–239. [CrossRef]
67. Lu, Q.; Lund, R.; Lee, T.C.M. An MDL approach to the climate segmentation problem. *Ann. Appl. Stat.* **2012**, *6*, 299–319. [CrossRef]
68. Killick, R.; Fearnhead, P.; Eckley, A. Optimal detection of changepoints with a linear computational cost. *J. Am. Stat. Assoc.* **2012**, *107*, 1590–1598. [CrossRef]
69. Li, Y.; Lund, R.; Hewaarachchi, A. Multiple change point detection with partial information on change point times. *Electr. J. Stat.* **2019**, *13*, 2462–2520. [CrossRef]
70. Hewaarachchi, A.P.; Li, Y.; Lund, R.; Rennie, J. Homogenization of daily temperature data. *J. Clim.* **2017**, *30*, 985–999. [CrossRef]
71. Caussinus, H.; Mestre, O. Detection and correction of artificial shifts in climate series. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2004**, *53*, 405–435. [CrossRef]
72. Menne, M.J.; Williams, C.N. Detection of undocumented change points using multiple test statistics and composite reference series. *J. Clim.* **2005**, *18*, 4271–4286. [CrossRef]
73. Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680. [CrossRef] [PubMed]
74. Hubert, P. The segmentation procedure as a tool for discrete modeling of hydrometeorological regimes. *Stoch. Environ. Res. Risk A* **2000**, *14*, 297–304. [CrossRef]
75. Gedikli, A.; Aksoy, H.; Unal, N.E. AUG-Segmenter: A user-friendly tool for segmentation of long time series. *J. Hydroinform.* **2010**, *12*, 318–328. [CrossRef]
76. Gedikli, A.; Aksoy, H.; Erdem Unal, N.; Kehagias, A. Modified dynamic programming approach for offline segmentation of long hydrometeorological time series. *Stoch. Environ. Res. Risk A* **2010**, *24*, 547–557. [CrossRef]