Faculty of Natural and Applied Sciences Journal of Mathematical and Statistical Computing
Print ISSN: 3026-8605, e-ISSN 3043-632X
www.fnasjournals.com
Volume 3; Issue 1; March 2026; Page No. 1-14.
DOI: https://doi.org/10.63561/jmsc.v3i1.1207

Change point Detection in Multivariate Climate Time Series: Performance Assessment of Hybrid Pelt and Isolation Forest Against Baseline Pelt

***[1,2]Ademuwagun, A.A., [2]Yahaya, H.U., & [2]Adams, S.O.**
[1]Department of Basic Science and General Studies, Federal College of Forestry Mechanization Afaka, Kaduna
[2]Department of Statistics, University of Abuja

***Corresponding author email**: muwax22@gmail.com

Abstract
The research studied the performance of Hybrid Pruned Exact Linear Time and Isolation Forest (PELT + I Forest) against Baseline PELT in accurately detecting change points in Climate time series data using simulation.
The research adopts a Monte Carlo simulation framework to design and evaluate a hybrid change-point detection technique that combines the Pruned Exact Linear Time (PELT) algorithm with machine learning anomaly detection method (I Forest). The hybrid approach (PELT+I Forest) is compared against baseline PELT using simulated multivariate climate datasets. Across small, moderate, and large sample sizes, the same directional patterns persist, temperature and humidity increase while rainfall decreases with more breaks. However, larger samples make the regime shifts more distinct and less noisy. This finding underscores that robust detection methods must perform well not only in large datasets but also in small samples, where noisy signals make breaks harder to capture. Enhanced detection algorithm like PELT+I Forest is therefore vital for early warning in short observational records or regional climate data with limited length. PELT+I Forest delivered consistent performance across moderate and large sample sizes, with marginal but steady improvements in balanced accuracy, making it more robust for longer or multivariate series.This study carried out a Performance Assessment on the developed, implemented and evaluated Hybrid change-point detection framework which integrates the Pruned Exact Linear Time (PELT) algorithm with machine learning anomaly detection method (Isolation Forest) for robust identification of structural breaks in multivariate climate time series against baseline PELT Algorithm. Relative to all other alternatives for Change Point Detection the hybrid PELT+I Forest approach balanced computational efficiency, interpretability, and robustness which retained PELT's exact segmentation while using ML scorers to capture multivariate, nonlinear anomalies.

Keywords: Change Point Detection (CPD), Pruned Exact Linear Time (PELT), Isolation Forest (IF)

Introduction

The quest to determine and identify the best method for detecting change point or structural breaks in climate data sets requires a method that has the ability to capture multivariate nonlinear components of the climate data sets. Change point analysis is an efficient tool used to understand the fundamental information in environmental data such as rainfall and temperature the result of a reasonable change point detection method can be effectively used in the prediction of flood and draught because it provides the key to resolving the non-stationary or inhomogeneous problems of climate change (Ademuwagun et al., 2024). Recording rainfall and temperature for consecutive years can teach us about fluctuations from year to year and can help us plan accordingly as we learn natures patterns.

The remote sensing of climatic and hydrological variables such as rainfall, humidity, temperature, soil moisture and stream flow is becoming increasingly more important as we strive to improve our understanding of the earth's climate. Because these variables change more rapidly than other land cover characteristics, a high frequency of image acquisition is required. This in turn results in a large volume of data that must be compared to detect changes and trends in moisture flux (Gomez, 2011). As it is important to accurately estimate rainfall for effective use of water resources, so also is Climate variability and change which are characterized not only by gradual long-term trends but also by abrupt regime shifts, often referred to as *structural breaks* or *change points*. Such shifts are critical in climate science because they indicate sudden transitions in temperature, rainfall, or humidity that may have far-reaching consequences for ecosystems, agriculture, water resources, and human livelihoods. Detecting these shifts is therefore central to understanding climate dynamics and informing decision-making. Traditionally, statistical change-point detection (CPD) methods such as the Cumulative Sum (CUSUM) test, binary segmentation, and penalized likelihood approaches have been widely applied to climate datasets. Among these, the Pruned Exact Linear Time (PELT) algorithm has emerged as a particularly powerful tool because it provides exact segmentation with linear average computational cost (Killick, et al, 2012). PELT has been applied in climate research to identify changes in temperature records, precipitation series, and other environmental variables. However, standard PELT methods generally assume univariate Gaussian processes with shifts in mean and/or variance. In practice, climate systems are multivariate, noisy, and nonlinear, where shifts often occur simultaneously across variables (e.g., warmer temperatures coinciding with declining rainfall and rising humidity). This makes classical CPD approaches limited in their ability to detect complex, multivariate regime shifts.

On the other hand, machine learning–based anomaly detection method such as Isolation Forest (Liu, Ting, & Zhou, 2008) have demonstrated strong performance in detecting atypical patterns in multivariate, nonlinear data. This algorithm does not rely on strict parametric assumptions and can capture joint anomalies that classical methods may miss. Yet, machine learning anomaly detection methods lack the statistical rigor of segmentation algorithms like PELT, which provide interpretable change-point estimates and established theoretical guarantees. This methodological gap motivates the development of a hybrid approach that combines the statistical rigor of PELT with the multivariate anomaly sensitivity of machine learning. By first computing anomaly scores using machine learning method (Isolation Forest) and then applying PELT to segment these scores, the hybrid method aims to improve the detection of structural breaks in multivariate climate data. Such an approach holds particular promise in contexts with short, noisy datasets; a common reality in many developing regions where climate monitoring infrastructure is sparse. This study aims to develop, implement, and evaluate a hybrid change-point detection framework that integrates the Pruned Exact Linear Time (PELT) algorithm with machine learning anomaly detection method (Isolation Forest) for robust identification of structural breaks in multivariate climate time series. The study compares the performance of the developed hybrid PELT method by integrating machine learning anomaly scores from Isolation Forest with PELT segmentation; implement the baseline PELT algorithm and hybrid PELT-machine learning method for detecting structural breaks (change points) by comparing their performances across different simulation scenarios using precision.

Abrupt changes, regime shifts, and structural breaks in climate systems are increasingly recognized as critical phenomena that influence long-term variability and adaptation planning. The Intergovernmental Panel on Climate Change (IPCC) highlights that abrupt shifts in temperature, precipitation, and hydrological regimes are not only scientifically significant but also carry high policy relevance for risk management and adaptation strategies (IPCC, 2019). For instance, regime shifts in rainfall patterns directly affect water security and agriculture, while abrupt warming trends alter humidity levels and the frequency of extreme events (Scheffer et al., 2009). Empirical work has also catalogued abrupt ecosystem and climate regime shifts, noting their often multivariate and interacting nature (deYoung et al., 2008; Lenton, 2013). This underscores the necessity of robust change-point detection (CPD) tools that can reliably identify structural breaks in climate time series. Statistical methods for CPD have a long history, including likelihood-ratio tests, CUSUM, binary segmentation, and penalized likelihood approaches (Page, 1954; Hawkins, 2001). Dynamic Programing Algorithm for finding the global minimizers of the sum of squared residuals (SSR) for multiple structural changes in a linear regression model which relies on the principle of dynamic programming (Bai & Perron, 2003). Among modern algorithms, the Pruned Exact Linear Time (PELT) method of Killick, et al. (2012) is particularly attractive. PELT minimizes a penalized cost function and guarantees global optimality while employing pruning to achieve average linear complexity, making it one of the most efficient exact segmentation methods available. PELT has been widely used in environmental and climate contexts, including hydrology, oceanography, and meteorology, because of its computational tractability and theoretical rigor (Reeves et al., 2007; Beaulieu et al., 2012). However, PELT in its standard form typically assumes univariate Gaussian data with shifts in mean and/or variance, which limits its applicability to real-world multivariate and nonlinear climate processes. Moreover, penalty selection in PELT can strongly affect detection outcomes: too low a penalty leads to spurious change points, while too high a penalty misses true breaks (Haynes et al., 2017). These limitations motivate exploration of hybrid methods that preserve the optimal segmentation of PELT while enhancing its robustness to multivariate, nonlinear, and noisy signals. Although effective, classical CPD methods including PELT face several limitations in climate applications: Univariate focus; Most implementations detect changes in a single series (e.g., temperature), whereas climate processes are multivariate and interactive. A shift in rainfall often co-occurs with humidity and temperature changes, making univariate detection incomplete (Killick & Eckley, 2014), Parametric assumptions; Classical methods generally assume Gaussian and linear mean/variance shifts. Yet climate dynamics may involve nonlinear transitions, dependence changes, or distributional shifts beyond mean/variance (Truong et al., 2020), Data length and noise; Regional climate records are often short and noisy, particularly in Africa and other data-sparse regions (Zwiers & von Storch, 2004). In such contexts, classical segmentation either over-segments or fails to detect subtle breaks, Penalty sensitivity; The choice of penalty value is nontrivial; performance depends heavily on penalty calibration, which is rarely straightforward in practice (Haynes et al., 2017).

Together, these limitations underscore the need for more flexible approaches that retain the statistical rigor of PELT while being sensitive to multivariate and nonlinear structures. (Parth Sinroza 2020) studied a review on Climate Change Impact Analysis using Statistical tools. The literature reviewed in this study, showed that, not only globally, but also locally, climate change is now a serious problem. In this field, where not only scientists and engineers, but also social scientists, economists, environmentalists and policy makers put their best to deal with the problem, major research efforts are now being concentrated. There is need to use statistical techniques and software help to explain whether the climate has changed in the past and, more importantly, whether the climate will change in the future. Much of the literature available has used trend detection techniques to estimate the degree of change, such as linear regression. It has now become easier to schedule adaptation to the predicted quantum of change by detecting various magnitudes of change as well. Analyses and prediction of change in critical climatic variables like rainfall and temperature are, therefore, extremely important. Based on this study, we are trying to expand the research area by examining these parameters based on statistical analysis. (Abdulkadir et al 2021) studied Analysis of climate Data using Spatial Techniques to estimate Rainfall in the North West of Nigeria using three different spatial models, Ordinary kriging (OK), Inverse distance weighting (IDW) and geographically weighted regression (GWR). The study inspected the spatial variation of Rainfall in different localities in the North West of Nigeria; Rainfall data present the basic metrological role in many fields of geostatistical and practice that is why it is one of the major climate resources that can be used as a measuring tool of climate change.

The aim was to analysed the data of one decade for thirty sample locations from (2010 – 2019) obtained from NIMET using three different spatial models and compare the model's performance in order to obtained the optimal model that can be used for rainfall prediction in the study Area. The assessment of the optimal model is based on the validation methods used in the research that is the method of RMSE and R^2. The supportive auxiliary variables which have been used in estimating neighbouring locations are Humidity, Temperature, Pressure and Wind speed. The predicted Rainfall in the models has proved the theory of ITCZ, and the locations with a higher predicted Rainfall are in the southern part while the locations with a lower predicted Rainfall are in the northern part of the study Area in all the models, regarding the validation methods used in the research, Geographically weighted Regression (GWR) outperform Ordinary Kring (O.K) and Inverse Distance Weighting (IDW) in terms of RMSE and R^2, we used variables such as temperature, pressure, humidity and win-speed as an independent variables for predicting Rainfall. From the findings, the results show that, the Regression model outperformed geostatistical and deterministic model, but all Univariate models have come out with correlated result, however the outcomes of GWR are best in terms of consistency in prediction, followed by OK and lastly IDW. GWR has proved the used of supplementary variables for rainfall estimation in every location would improve the prediction accuracy. The study recommended the use of GWR for Rainfall estimation in the North West region. Generally statistical data changes over time, and spatial data changes due to the event of ENSO. Therefore, considering the role of auxiliary variables in Rainfall estimate should be more precise considering Models that will incorporate environmental factors as independent variables in their estimation process instead of using univariate models that only depends on its variable of study and the weight to estimate the neighbouring Locations. Therefore, comparing spatial Models that incorporate auxiliary variables in rainfall estimation will be suitable for obtaining good result in the study area. (Kouakou et al 2022) Studied Statistical modelling of monthly maximum temperature in Senegal. They provided the first statistical analysis of maximum temperature in Senegal. The data were from twelve stations spread across Senegal. The generalized extreme value distribution was fitted to maximum temperature by the method of maximum likelihood. Probability and quantile plots showed that the generalized extreme value distribution provided an adequate fit for all stations. The vast majority of stations did not exhibit significant trends in temperature. Three of the stations exhibited positive trends in temperature.

Methodology

Hybrid PELT+I Forest Development and Performance Assessment

PELT is a change point detection algorithm designed to segment time series into homogeneous regimes by minimizing a cost function. Its strength lies in its Computational efficiency (linear in time) and suitability for large climate datasets. It's purely statistical and can miss nonlinear patterns that machine learning can capture this limitation of PELT can adequately be addressed by the developed Hybrid PELT + I Forest Model.

Performance Assessment is measured by High Recall (few missed true changes) High Precision (few false alarms), while F1 Score summarizes the trade-off.

Assume for simplicity that within a segment observation are independent Gaussian:

$$\mathbf{X}_j \sim N_p(\mu_j, \Sigma_j), \qquad t = \tau_{j-1}+1, \dots, \tau_j. \tag{1}$$

For univariate derivation (used when applying PELT to a single series Y_t), assume

$$Y_t \sim N(\mu_j, \sigma_j^2) \qquad \text{on segment } j \tag{2}$$

The negative log-likelihood for a candidate segment $s+1, \dots, e$ (ignoring constants) is

$$\mathcal{L}(s+1, e) = \frac{(e-s)}{2} \log \sigma^2_{(s+1,e)} + \frac{1}{\sigma^2_{(s+1,e)}} \sum_{t=s+1}^{e} \left(Y_t - \hat{\mu}_{(s+1,e)}\right)^2, \tag{3}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are MLEs on that segment. Under the homoscedastic assumption (constant σ^2 across segments) the cost simplifies to the sum of squared errors form:

$$\mathcal{C}(s+1, e) = \sum_{t=s+1}^{e} (Y_t - \bar{Y}_{s+1,e})^2 \tag{4}$$

(up to a constant factor).

Penalized cost formulation

Estimation of change points is framed as minimization of a penalized sum of segment costs:

$$\widehat{m}, \{\hat{\tau}_j\} = \arg \min_{m, \{\tau_1, \dots, \tau_m\}} \sum_{j=0}^{m} \mathcal{C}(\tau_j + 1, \tau_{j+1}) + \beta_m \tag{5}$$

where β is a penalty for each change point (controls trade-off between fit and complexity). Common choices: $\beta = \alpha log\ n (BIC)$ or $(AIC)\ \beta = 2p$ or manual fixed penalty as used in my code.

The Dynamic programming and pruning (PELT)

Let $F(t)$ be optimal value of objective up to time t:

$$F(t) = \min_{s<t} \{F(s) + \mathcal{C}(s+1, t) + \beta\}, \qquad F(0) = -\beta \tag{6}$$

PELT uses a pruning condition to cut candidate s that cannot be optimal for any later t. If for candidate τ and s we have

$$F(\tau) + \mathcal{C}(\tau+1, t) \geq F(s) + \mathcal{C}(s+1, t) \qquad \forall t \geq T, \tag{7}$$

then τ is pruned. Killick et al. (2012) proved that under mild regularity this pruning leads to average linear computational complexity in n.

Key property: PELT gives the *global* minimizer of the penalized cost (not greedy) provided cost is additive over segments and penalty is constant per change.

Developing the Hybrid PELT + Machine Learning

Machine Learning Anomaly Scorer

The hybrid pipeline maps the multivariate series $\{\mathbf{X}_t\}$ to a scalar score time series $\{\boldsymbol{S}_t\}$ that summarizes anomalousness at each time. Then PELT is applied to $\{\boldsymbol{S}_t\}$. Isolation Forest is used:

Isolation Forest (I Forest)

Proposition 1: Let Z_n be a sorted vector of real numbers such that $Z_i \in R$ $i=1,\dots,n$, and $Z_i \leq Z_j$, $i \leq j$. Denote by h the number of splits that it takes to isolate Z_1 and by $g(Z_i, Z_{i+1}|Z_n)$ the probability of a random split occurring on the internal $[Z_i, Z_{i+1})$ given Z_n. Assume that

$$g(Z_i, Z_{i+1}|Z_n) = \frac{g(Z_i, Z_{i+1})}{G(Z_1, Z_n)} \tag{8}$$

Ademuwagun, A.A., Yahaya, H. U., & Adams, S.O. (2026). Change point detection in multivariate climate time series: performance assessment of hybrid pelt and isolation forest against baseline pelt. *FNAS Journal of Mathematical and Statistical Computing*, *3*(1), 1-14. https://doi.org/10.63561/jmsc.v3i1.1207

with $G(Z_1, Z_i) = \sum_{j=1}^{i-1} g(Z_j, Z_{j+1})$, then the moment generating function (mgf) of h is given by

$$E[e^{uh}|Z_n] = \prod_{i=1}^{n-1} \frac{e^u g(Z_i, Z_{i+1}) + G(Z_1, Z_i)}{G(Z_1, Z_{i+1})} \quad (9)$$

Proof: Since h is bounded between 1 and n-1, it is clear that for any function f: R → R

$$E[f(h)|Z_n] = \sum_{i=1}^{n-1} f(\mathrm{i}) P[h = i|Z_n], \quad (10)$$

Furthermore, due to the recursive nature of the random splits

$$P[h = i|Z_n] = \sum_{j=1}^{i-1} g(Z_j, Z_{j+1}|Z_n) P[h = i-1|Z_j] \quad (11)$$

Where Z_j contains the first j elements of Z_n

Substituting equation (11) into equation (3.2110) and swapping the other of summations yields

$$E[f(h)|Z_n] = f(1) g(Z_1, Z_2|Z_n) + \sum_{j=2}^{n-1} g(Z_j, Z_{j+1}|Z_n) \sum_{i=2}^{j} f(\mathrm{i}) P[h = i-1|Z_j] \quad (12)$$

Which can be further simplified as

$$E[f(h)|Z_n] = f(1) g(Z_1, Z_2|Z_n) + \sum_{j=2}^{n-1} g(Z_j, Z_{j+1}|Z_n) E[f(h+1)|Z_j] \quad (13)$$

Equation (13) can then be written as a recursion i.e.,

$$E[f(h)|Z_n] = \frac{g(Z_{n-1}, Z_n)}{G(Z_1, Z_n)} E[f(h+1)|Z_{n-1}] + \frac{G(Z_1, Z_{n-1})}{G(Z_1, Z_n)} E[f(h)|Z_{n-1}] \quad (14)$$

Where the first term on the right-hand side of equation(14) corresponds to the last term of the summation appearing in equation (13) , and the normalization factor $\frac{G(Z_1, Z_{n-1})}{G(Z_1, Z_n)}$ is necessary to move from probabilities conditioned on Z_n to probabilities conditioned on Z_{n-1} . Finally, choosing f(h) = e^{uh} gives

$$E[e^{uh}|Z_n] = \frac{e^u g(Z_{n-1}, Z_n) + G(Z_1, Z_{n-1})}{G(Z_1, Z_n)} E[e^{uh}|Z_{n-1}] \quad (15)$$

And by following the recursion one obtains the desired result of equation (9).

Given the mgf in equation (9) ,it is easy to compute other quantities such as the expectation and the variance of h - through its derivatives with respect to h.

Corollary 1: Let Z_n, h and $g(Z_i, Z_{i+1}|Z_n)$ be defined as in *Proposition 1*. Then

$$E[h|Z_n] = 1 + \sum_{i=2}^{n-1} \frac{g(Z_i, Z_{i+1})}{G(Z_1, Z_{i+1})} \quad (16)$$

$$V[h|Z_n] = \sum_{i=2}^{n-1} \frac{g(Z_i, Z_{i+1})}{G(Z_1, Z_{i+1})} \left(1 - \frac{g(Z_i, Z_{i+1})}{G(Z_1, Z_{i+1})}\right) \quad (17)$$

Proof: The result is immediate after a straightforward computation of the derivatives of log $E[e^{uh}|Z_n]$. While *Proposition 1* provides a general framework for isolation in one dimension from now on we set $g(Z_i, Z_{i+1}) = (Z_{i+1}, Z_i)^\alpha$. Applying this choice to equation (16) and (17) yields the two scoring functions.

$$E[h|Z_n] = 1 + \sum_{i=2}^{n-1} \frac{(Z_{i+1} - Z_i)^\alpha}{\sum_{j=1}^{i} (Z_{j+1} - Z_j)^\alpha} \quad (18)$$

$$V[h|Z_n] = \sum_{i=2}^{n-1} \frac{(Z_{i+1} - Z_i)^\alpha}{\sum_{j=1}^{i} (Z_{j+1} - Z_j)^\alpha} \left(1 - \frac{(Z_{i+1} - Z_i)^\alpha}{\sum_{j=1}^{i} (Z_{j+1} - Z_j)^\alpha}\right) \quad (19)$$

With this formulation, the split probabilities used in I Forest are particular case of equations (18) and (19) with the scoring functions can thus be further simplified as

$$E[h|Z_n] = 1 + \sum_{i=2}^{n-1} \frac{Z_{i+1} - Z_i}{Z_{i+1} - Z_1} \quad (20)$$

$$V[h|Z_n] = \sum_{i=2}^{n-1} \frac{Z_{i+1} - Z_i}{Z_{i+1} - Z_1} \left(1 - \frac{Z_{i+1} - Z_i}{Z_{i+1} - Z_1}\right) \quad (21)$$

An outlier can be isolated in just a few splits and therefore, E[h] should be smaller for outliers than for inliers. An inlier should be surrounded by other inliers and therefore it takes on average many splits to isolate it. However, an inlier can also be isolated in just a few splits giving rise to a large range of variation for h. Hence it is assumed that inlier has a higher V[h] than an outlier. In the extreme case where a point is always isolated in a single split, we have V[h] =0 (Arias 2023)

Ademuwagun, A.A., Yahaya, H. U., & Adams, S.O. (2026). Change point detection in multivariate climate time series: performance assessment of hybrid pelt and isolation forest against baseline pelt. *FNAS Journal of Mathematical and Statistical Computing*, *3*(1), 1-14. https://doi.org/10.63561/jmsc.v3i1.1207

Principle: anomalies are easier to isolate by recursive random partitioning. For an observation x the path length $h(x)$ in a random binary isolation tree is the number of splits required to isolate x. The expected path length $E[h(x)]$ is shorter for anomalies.

Anomaly score definition. For a sample of size n, the score is

$$s(x,n) = 2^{-E[h(x)]/c(n)}, \quad (22)$$

where $c(n)$ is the average path length of unsuccessful search in a Binary Search Tree:

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n}, \quad (23)$$

with $H(\cdot)$ harmonic number. $s(x,n) \in (0,1]$, values near 1 indicate anomalies.

Interpretation of Isolation Forest Scores:

i. If $E[h(x)] \rightarrow$ C(n), the score S approaches 0.5 indicating a normal observation (dataset has no distinct anomalies).
ii. If $E[h(x)] \rightarrow$ 0, the score S approaches 1 indicating a high likelihood of an anomaly.
iii. If $E[h(x)] \rightarrow$ n - 1 is significantly larger than C(n) the score S approaches 0 also indicating a normal observation.

Statistical justification: The random partitioning procedure induces a distribution of path lengths; the expectation across many trees consistently orders points by "isolation" and is robust to scaling and multivariate dependencies. Under mild mixing, the score is consistent for separating rare outliers from bulk.

Hybrid Procedure: Statistical Justification and Derivation

Pipeline

i. Given $\{X_t\}_{t=1}^{n}$, compute scalar scores $S_t = s(\mathbf{X}_t)$ by either IForest or RF proximity
ii. Optionally clean scores: we replace NA with mean.
iii. Apply univariate change-point detection (PELT) to $\{S_t\}$ to find change points in anomalousness:

$$\hat{\tau} = \arg min \sum_{j=0}^{m} \mathcal{C}_S(\tau_j + 1, \tau_{j+1}) + \beta m \quad (24)$$

where $\mathcal{C}_S$ is the cost computed on S (e.g. sum of squared deviations)

iv. Interpret $\hat{\tau}$ as multivariate change points for $\mathbf{X}_t$.

Statistical Logic

The scalar sequence S_t is constructed so that structural changes in the joint distribution of $\mathbf{X}_t$ (mean shifts, covariance changes, new interactions) will typically produce local changes in the distribution of S_t. Thus change-point segmentation of S_t can detect multivariate distributional shifts.

Advantages:

i. Reduces multivariate change-point problem to univariate segmentation (computationally simpler).
ii. ML scorers capture complex, non-linear dependencies and joint anomalies not visible from marginal series.

Caveat: mapping to a univariate score can lose information; the score must be sensitive to the class of changes of interest.

- **Cost Choice on the Score Series and Penalty Calibration**

When applying PELT to $\{S_t\}$, you must choose:

Segment Cost Cs

If S_t is treated as Gaussian-like, one may use the sum of squared deviations cost:

$$\boldsymbol{\mathcal{C}_S}(a,b) = \sum_{t=a}^{b} (S_t - \bar{S}_{a:b})^2 \quad (25)$$

Alternatively, use negative log-likelihood under appropriate distributional assumption for S_t (e.g., if $S_t \in (0,1)$ a Beta model could be more appropriate).

Penalty $\boldsymbol{\beta}$

i. Statistical choices link to model selection criteria: Akaike-type $\beta = 2p, BIC - type\beta = Klogn, MDL,$ or manually tuned penalty.
ii. Penalty affects Type I (false positives) vs. Type II (missed breaks). In practice (your results), high recall but low precision suggests penalty is too small (leading to over-segmentation). Increasing β reduces false positives.

Simulation Setup

This study adopts a **Monte Carlo simulation framework** to design and evaluate a hybrid change-point detection technique that combines the **Pruned Exact Linear Time (PELT) algorithm** with machine learning anomaly

Ademuwagun, A.A., Yahaya, H. U., & Adams, S.O. (2026). Change point detection in multivariate climate time series: performance assessment of hybrid pelt and isolation forest against baseline pelt. *FNAS Journal of Mathematical and Statistical Computing*, *3*(1), 1-14. https://doi.org/10.63561/jmsc.v3i1.1207

detection methods. The hybrid approaches (PELT+I Forest) is compared against the baseline PELT using simulated multivariate climate datasets.

Simulation of Climate Data: Parameter Settings

To generate the synthetic multivariate climate time series used in this study we simulate three variables: **temperature (°C)**, **rainfall (mm)**, and **humidity (%)**. The data-generating mechanism and parameter values are:

i. **Sample sizes (n):** 50, 70, 100, 200, 300, 500, 600, 800, 1000. These cover small, moderate and large sample regimes for robustness checks.

ii. **Number of change points (m):** 0, 1, 2, 3, 4. Change points are randomly positioned subject to a minimum distance constraint (see below).

iii. **Change-point placement:** change-point locations are sampled without replacement from the integer set {20, 21, …, n−20} and then sorted. This ensures a minimum segment length of 20 observations and avoids edge effects.

iv. **Regime means (segment-specific):** for segment index j ($j = 1, 2, \ldots$) the variable means are defined as:

Temperature $mean = 20 + 2 \times j$ (°C) in code this is $mean\ =\ 20\ +\ i * 2$.

Rainfall $mean = 100 - 5 \times j(mm)$ in code $mean = 100 - i * 5$.

Humidity $mean = 60 + 3 \times j(\%)$ in code $mean\ =\ 60\ +\ i * 3$.

This specification produces upward stepping temperature and humidity regimes and a downward stepping rainfall regime across successive segments.

v. **Within-segment dispersion (noise):**

Temperature standard deviation $(\sigma\_temp)\ =\ 1.0°C$.

Rainfall standard deviation $(\sigma\_rain) = 10.0\ mm$.

Humidity standard deviation $(\sigma\_humid)\ =\ 5.0\%$.

These are implemented by adding independent Gaussian noise: $rnorm(length(seg), mean =\ldots, sd\ =\ sd\_value)$.

Random seed: set. Seed (123) for reproducibility in all simulation runs.

Performance Metrics

Let true change points be $\mathcal{T} = \{\tau_1, \ldots, \tau_m\}$ and detected ones $\mathcal{D} = \{\delta_1, \ldots, \delta_k\}$. With tolerance window ω (e.g., $\omega = 5$ time units):

i. True Positive (TP): a true τ has at least one δ with $|\delta - \tau| \leq \omega$.

ii. False Positive (FP): a detected δ is not within www of any true τ.

iii. False Negative (FN): a true τ not matched by any δ within ω.

i. $$Precision = \frac{TP}{TP+FP} \quad (26)$$

ii. $$Recall = \frac{TP}{TP+FN} \quad (27)$$

iii. $$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (28)$$

Results

Presentation of Results / Findings of Simulated Data

The descriptive statistics of the simulated climate data with varying numbers of change points (Tables 4.1-4.3) and the corresponding time series plots (Figures 4.1-4.9) illustrate how structural breaks alter the behaviours of temperature, rainfall, and humidity. These findings have direct implications for the development of enhanced change-point detection techniques in climate research.

Table 1: Summary Statistics of Simulated Climate Data with 0-4 Change Points over Small Sample Sizes

Change Point		50			70			100		
		Temp	Rain	Humid	Temp	Rain	Humid	Temp	Rain	Humid
0	Mean	22.034	96.464	61.731	22.074	94.417	63.022	22.09	93.925	63.602
	SD	0.926	9.054	4.947	0.907	9.341	5.051	0.913	9.67	4.749
	Min	20.033	71.908	52.734	20.033	71.908	54.992	19.691	74.468	54.217
	Max	24.169	116.873	73.501	24.169	116.873	79.205	24.187	127.41	74.465
1	Mean	22.643	93.311	63.95	23.293	89.697	65.056	22.345	93.681	62.759
	SD	1.522	10.63	5.319	1.363	11.504	5.064	1.258	9.587	4.668
	Min	19.601	70.49	53.444	20.43	56.622	52.298	18.962	71.417	50.713

Ademuwagun, A.A., Yahaya, H. U., & Adams, S.O. (2026). Change point detection in multivariate climate time series: performance assessment of hybrid pelt and isolation forest against baseline pelt. *FNAS Journal of Mathematical and Statistical Computing*, *3*(1), 1-14. https://doi.org/10.63561/jmsc.v3i1.1207

	Max	26.454	118.755	76.541	26.46	119.544	77.835	25.535	118.67	75.589
2	Mean	23.976	87.878	66.287	23.897	90.351	65.23	23.26	92.093	64.94
	SD	2.084	11.22	5.517	1.917	10.062	5.514	2.089	11.072	6.253
	Min	20.296	61.417	49.601	19.642	67.77	53.085	18.953	59.486	51.636
	Max	27.861	108.855	76.885	27.918	115.571	77.055	27.673	116.465	81.057
3	Mean	24.849	85.46	67.705	24.359	89.744	66.449	24.762	88.007	66.959
	SD	3.072	9.988	6.434	2.578	12.089	6.399	2.343	11.437	6.391
	Min	20.017	65.424	57.375	20.057	54.921	52.608	20.075	59.813	52.444
	Max	29.536	104.527	87.286	28.887	116.917	85.424	29.4	117.996	84.92
4	Mean	25.85	84.743	69.738	25.968	87.714	69.166	25.377	88.264	68.069
	SD	4.145	12.649	8.742	3.682	13.121	6.662	3.471	13.178	7.173
	Min	19.398	55.563	51.755	20.14	58.515	53.374	19.826	56.483	51.431
	Max	32.458	106.394	88.459	32.305	114.507	84.111	31.851	123.904	86.421

For temperature, at baseline (0 change points), temperature levels remain stable around 22°C with low variability. However, as the number of breaks increases, both the mean and variability rise significantly, reaching 25-27°C and standard deviations above 3°C at four breaks. The time series plots show clear stepwise upward shifts across regimes. This behavior demonstrates that structural breaks can mimic sudden warming episodes. Detecting these shifts reliably is crucial because traditional methods assuming stationarity may overlook abrupt climatic changes, leading to underestimation of warming signals.

Table 2: Summary Statistics of Simulated Climate Data with 0-4 Change Points over Moderate Sample Sizes

Change Point		200			300			500		
		Temp	**Rain**	**Humid**	**Temp**	**Rain**	**Humid**	**Temp**	**Rain**	**Humid**
0	Mean	21.991	95.421	63.159	22.034	95.091	63.083	22.035	94.977	63.131
	SD	0.943	9.959	4.824	0.946	9.889	5.167	0.973	10.109	4.964
	Min	19.691	70.341	48.951	19.691	66.902	49.992	19.339	66.902	49.523
	Max	25.241	120.715	75.151	25.241	120.715	76.459	25.241	121.917	79.952
1	Mean	23.432	91.234	65.234	23.018	92.912	64.4	23.202	91.317	63.99
	SD	1.455	10.199	5.23	1.432	10.04	5.303	1.405	10.102	5.325
	Min	19.398	64.507	51.947	19.451	67.513	50.761	19.411	61.234	43.874
	Max	26.685	119.581	81.92	26.284	126.84	76.952	27.852	121.442	78.599
2	Mean	23.135	92.368	64.837	24.251	88.848	66.406	22.934	92.211	64.369
	SD	1.971	10.581	5.282	1.653	11.174	5.392	1.567	10.404	5.142
	Min	19.686	64.064	52.276	20.322	58.621	43.874	19.218	62.875	47.267
	Max	28.577	115.056	80.825	27.73	128.518	78.783	27.793	123.391	76.637
3	Mean	24.772	87.799	67.221	25.058	86.732	67.772	24.733	88.057	67.389
	SD	2.282	11.2	6.172	2.689	12.379	5.774	1.906	10.513	5.765
	Min	20.322	50.734	42.874	19.646	48.663	50.313	19.862	52.379	50.054
	Max	29.954	120.41	81.729	30.78	120.199	83.155	31.084	112.318	90.579
4	Mean	26.891	83.278	70.182	24.4	88.581	66.67	25.904	85.805	68.309
	SD	2.965	11.694	6.431	2.714	11.443	6.404	2.959	12.229	6.762

Min	19.935	45.561	48.617	18.853	55.631	51.084	20.711	50.388	48.379
Max	31.862	113.892	83.645	32.912	114.257	85.637	32.19	119.289	89.281

Rainfall follows a contrasting pattern: mean levels decline progressively with more breaks, from ~96 mm at baseline to 82–88 mm at four breaks, while variability increases from ~9–10 mm to ~12–13 mm. The plots show successive downward adjustments, representing drying regimes. These results highlight that ignoring change-points could mask shifts towards drier climates. Therefore, enhanced detection techniques are essential for identifying both gradual and abrupt precipitation changes that directly affect water resources and agriculture.

Table 4.3: Summary Statistics of Simulated Climate Data with 0-4 Change Points over Large Sample Sizes

Change Point		**600**			**800**			**1000**		
		Temp	**Rain**	**Humid**	**Temp**	**Rain**	**Humid**	**Temp**	**Rain**	**Humid**
0	Mean	22.022	95.219	63.155	22.011	95.301	62.985	22.016	95.425	62.899
	SD	0.967	10.255	4.946	0.981	0.99	4.979	0.992	10.097	4.892
	Min	19.19	68.983	47.761	19.19	68.047	47.761	19.19	64.521	48.757
	Max	25.241	126.84	79.952	25.241	128.904	79.453	25.241	128.904	80.105
1	Mean	22.849	92.949	64.158	23.615	90.722	65.292	23.583	91.227	65.176
	SD	1.431	10.11	4.937	1.29	9.763	4.927	1.24	10.09	5.214
	Min	18.175	62.181	49.81	19.123	58.535	47.323	19.218	52.647	47.331
	Max	26.52	133.518	80.196	26.839	122.511	83.276	26.725	125.455	81.421
2	Mean	24.337	89.107	66.384	25.295	86.921	67.936	22.819	92.59	64.282
	SD	2.115	11.422	5.6	1.69	10.051	5.411	1.825	10.441	5.354
	Min	19.322	53.626	49.232	19.761	53.199	47.616	18.82	54.706	45.561
	Max	29.236	123.949	83.3	29.084	114.769	84.192	28.778	120.651	79.675
3	Mean	25.068	87.681	67.439	25.037	86.85	67.704	25.521	85.945	68.308
	SD	2.425	11.347	6.085	2.171	11.136	5.829	2.462	11.288	5.971
	Min	19.523	53.199	50.071	19.434	54.005	50.853	19.359	53.572	50.186
	Max	30.927	118.49	84.825	31.218	117.901	85.557	30.605	118.387	86.215
4	Mean	27.248	81.949	70.922	25.684	85.598	68.73	25.339	87.009	67.98
	SD	2.663	11.507	6.136	3.14	12.734	6.517	2.587	11.529	6.130
	Min	19.434	44.706	51.561	19.161	49.259	49.161	19.739	57.191	49.003
	Max	32.558	114.635	85.811	32.667	125.352	88.768	32.857	124.195	86.406

Humidity exhibits rising means and greater dispersion with increasing breaks, moving from ~63% at baseline to ~70% at four breaks. Time series plots show upward shifts in atmospheric moisture regimes. Such shifts, if not correctly identified, may obscure the relationship between temperature rise and atmospheric dynamics. Accurate detection techniques allow researchers to link humidity changes to underlying warming processes and assess compound climate risks. Across small, moderate, and large sample sizes, the same directional patterns persist—temperature and humidity increase while rainfall decreases with more breaks. However, larger samples make the regime shifts more distinct and less noisy. This finding underscores that robust detection methods must perform well not only in large datasets but also in small samples, where noisy signals make breaks harder to capture. Enhancing detection algorithms is therefore vital for early warning in short observational records or regional climate data with limited length.

The simulations show that structural breaks fundamentally reshape climate variable distributions. More breaks correspond to warmer, drier, and more humid conditions, along with higher variability; patterns consistent with observed climate change globally. Crucially, the results emphasize that failure to account for change points risks misinterpreting long-term climate trends as stable, when in fact they are regime-dependent. Developing enhanced change-point detection techniques is therefore not only a methodological improvement but a necessity for accurately characterizing climate dynamics, attributing observed changes to underlying processes, and informing adaptation strategies.

Change Points Detection Performance Assessment of Hybrid PELT Machine Learning Approach against the Baseline PELT for the Simulated Multivariate Climate Dataset

The Performance Assessment of baseline PELT and the hybrid PELT machine learning method (PELT+I Forest) over simulated multivariate climate datasets demonstrates the promise of integrating machine learning into classical change-point detection. The results (Tables 4.4-4.6) reveal consistent strengths of the hybrid approaches, particularly in terms of robustness, balanced accuracy, and adaptability to noisy climate data.

Table 4: Performances of Hybrid PELT Machine Learning Approach against PELT over the Simulated Small Sample Sizes Multivariate Climate Dataset

Metrics	Change Point	50		70		100	
		PELT	PELT+IForest	PELT	PELT+IForest	PELT	PELT+IForest
Precision	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.0602	0.0636	0.0412	0.0423	0.0268	0.0264
	2	0.1344	0.1324	0.0899	0.0886	0.0617	0.0596
	3	0.1856	0.1861	0.1451	0.1415	0.0949	0.0951
	4	0.2507	0.2423	0.1919	0.1980	0.1369	0.1387
Recall	0	0	0	0	0	0	0
	1	1	1	1	1	1	1
	2	1	1	1	1	1	1
	3	1	1	1	1	1	1
	4	1	1	1	1	1	1
F1 Score	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.1135	0.1195	0.0792	0.0811	0.0523	0.0515
	2	0.2367	0.2337	0.1648	0.1626	0.1163	0.1124
	3	0.3128	0.3136	0.2531	0.2477	0.1732	0.1737
	4	0.4004	0.3897	0.3217	0.3301	0.2406	0.2433

Source: *Researchers' Compilations from R-Outputs*

In small-sample scenarios, where variability and noise are most pronounced, all two methods achieve perfect recall (1.0), ensuring that no true change points are missed. This is a critical strength for climate change applications because even short observational records can capture major regime shifts, and missing such breaks could distort scientific conclusions.

The hybrid, however, distinguishes itself in terms of precision and F1-score. At higher change-point counts (3-4), PELT+I Forest demonstrated marginal but consistent improvements over baseline PELT. These increments, though modest, are significant in practice because they reflect the hybrids' capacity to reduce false positives in noisy, short records. In real-world climate analysis, this is especially significant because climate researchers often rely on short records (e.g., station-level data, regional series). The hybrid is better equipped to handle the noise inherent in such small samples, reducing false positives while preserving high sensitivity.

Ademuwagun, A.A., Yahaya, H. U., & Adams, S.O. (2026). Change point detection in multivariate climate time series: performance assessment of hybrid pelt and isolation forest against baseline pelt. *FNAS Journal of Mathematical and Statistical Computing*, *3*(1), 1-14. https://doi.org/10.63561/jmsc.v3i1.1207

Table 5: Performances of Hybrid PELT Machine Learning Approach against PELT over the Simulated Moderate Sample Sizes Multivariate Climate Dataset

Metrics	Change Point	200		300		500	
		PELT	PELT+IForest	PELT	PELT+IForest	PELT	PELT+IForest
Precision	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.0127	0.0128	0.0084	0.0083	0.0049	0.0049
	2	0.0267	0.0268	0.0172	0.0171	0.0100	0.0100
	3	0.0411	0.0421	0.0262	0.0261	0.0154	0.0153
	4	0.0564	0.0559	0.0362	0.0364	0.0205	0.0207
Recall	0	0	0	0	0	0	0
	1	1	1	1	1	1	1
	2	1	1	1	1	1	1
	3	1	1	1	1	1	1
	4	1	1	1	1	1	1
F1 Score	0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	1	0.02518	0.02519	0.01672	0.01643	0.00982	0.00975
	2	0.05199	0.05214	0.03380	0.03355	0.01978	0.01976
	3	0.07894	0.08080	0.05106	0.05090	0.03031	0.03015
	4	0.10683	0.10584	0.06987	0.07030	0.04018	0.04053

Source: *Researchers' Compilations from R-Outputs*

For moderate datasets, precision values decline across all methods, with PELT yielding between 0.01 and 0.05 depending on the number of breaks. However, the hybrid consistently matches or slightly surpass these values. At three change points with n=300, PELT+I Forest achieves a higher F1-score (0.05090) compared to PELT (0.05106), reflecting better balance between sensitivity and precision. Although the numerical gains are not dramatic, the hybrid consistency across scenarios highlights its strength. They neither degrade detection quality nor introduce instability. This robustness suggests that the anomaly signals generated by machine learning models, Isolation Forest's anomaly scores structures, provide meaningful supplementary information that stabilizes change-point detection under increasingly complex climate signals.

Table 6: Performances of Hybrid PELT Machine Learning Approach against PELT over the Simulated Large Sample Sizes Multivariate Climate Dataset

Metrics	Change Point	600		800		1000	
		PELT	PELT+IForest	PELT	PELT+IForest	PELT	PELT+IForest
Precision	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.0041	0.0041	0.0030	0.0030	0.0024	0.0024
	2	0.0082	0.0083	0.0061	0.0061	0.0049	0.0049
	3	0.0126	0.0126	0.0094	0.0093	0.0074	0.0074
	4	0.0170	0.0169	0.0127	0.0125	0.0100	0.0100
Recall	0	0	0	0	0	0	0
	1	1	1	1	1	1	1
	2	1	1	1	1	1	1
	3	1	1	1	1	1	1
	4	1	1	1	1	1	1
F1 Score	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.0082	0.0081	0.0061	0.0060	0.0048	0.0048

Ademuwagun, A.A., Yahaya, H. U., & Adams, S.O. (2026). Change point detection in multivariate climate time series: performance assessment of hybrid pelt and isolation forest against baseline pelt. *FNAS Journal of Mathematical and Statistical Computing*, *3*(1), 1-14. https://doi.org/10.63561/jmsc.v3i1.1207

2	0.0163	0.0164	0.0121	0.0122	0.0098	0.0097
3	0.0248	0.0248	0.0187	0.0184	0.0147	0.0147
4	0.0334	0.0332	0.0250	0.0248	0.0198	0.0197

In large datasets, all methods record very low precision (0.002–0.017), reflecting the difficulty of maintaining specificity when data length and random fluctuations increase. Nevertheless, the hybrids again mirror or slightly exceed baseline performance. For instance, at four change points with n=800, PELT+I Forest yields a precision of 0.0125 versus 0.0127 for PELT, with nearly identical F1-scores.

The key strength here is not dramatic improvement, but stability: the hybrid maintains equivalent recall and do not underperform relative to baseline PELT. This outcome is crucial for climate data analysis, as longer time series often contain multiple overlapping signals (e.g., seasonal cycles, decadal oscillations, anthropogenic shifts). By incorporating anomaly detection into the segmentation process, the hybrid remains robust and computationally efficient, providing a scalable framework for large-scale climate studies.

Discussion

The impact of increasing structural breaks at zero change points, showed that all methods correctly record zero detection (precision, recall, and F1 = 0), demonstrating no false alarms in stable systems. As the number of change points increases, precision and F1-scores systematically rise for all methods. This pattern confirms that the model improves as the true signal becomes stronger. The hybrid consistently tracks or slightly outperforms baseline PELT as the number of breaks increases. Notably, in the 3-4 change-point scenarios, where climate regimes are most unstable, the hybrid deliver their strongest relative advantage. This suggests that PELT+I Forest is particularly effective in capturing complex, regime-rich dynamics, which is highly relevant for climate change analysis where multiple interacting shifts (warming, drying, and humidity increases) may occur simultaneously. The study discovered that PELT+I Forest delivers consistent performance across moderate and large sample sizes, with marginal but steady improvements in balanced accuracy, making it more robust for longer or multivariate series.

Enhancing detection algorithms is therefore vital for early warning in short observational records or regional climate data with limited length. The simulations show that structural breaks fundamentally reshape climate variable distributions. More breaks correspond to warmer, drier, and more humid conditions, along with higher variability; patterns consistent with observed climate change globally. Crucially, the results emphasize that failure to account for change points risks misinterpreting long-term climate trends as stable, when in fact they are regime-dependent. Developing enhanced change-point detection techniques is therefore not only a methodological improvement but a necessity for accurately characterizing climate dynamics, attributing observed changes to underlying processes, and informing adaptation strategies.

Conclusion

The development and implementation of a hybrid change-point detection framework that integrates the Pruned Exact Linear Time (PELT) algorithm with machine learning anomaly detection method (Isolation Forest) for robust identification of structural breaks in multivariate climate time series by integrating machine learning anomaly scores from Isolation Forest with PELT segmentation. Implementation of the baseline PELT algorithm and hybrid PELT-machine learning methods for detecting structural breaks (change points) was done by comparing their performances across different simulation scenarios using precision, recall, and F1-score as performance metrics. PELT+I Forest maintained consistency and stability over moderate and large sample length. The findings reinforce the significance of hybrid PELT method for climate change analysis. By retaining perfect sensitivity and delivering improved balanced accuracy in noisy, multi-break scenarios. This reliability is crucial for: detecting abrupt climate transitions that may affect agriculture, water supply, and health; Supporting attribution studies by distinguishing true shifts from random variability; and enhancing early-warning systems in regions with sparse and noisy datasets. Based on the aforementioned, PELT+I Forest delivers consistent performance across moderate and large sample sizes, with marginal but steady improvements in balanced accuracy, making it more robust for longer or multivariate series.

Limitations

The Performance Assessment shows the discovery of an effective change point detection model for multivariate climate time series data. The Hybrid PELT+I Forest delivers consistent performance across moderate and large sample sizes, with marginal but steady improvements in balanced accuracy, making it more robust for longer or multivariate series. Extensive research should be done on more possible hybrid solutions that could adequately detect change points of multivariate climate data, taking note of the peculiarities associated with climate data. The Limitation of this study might be attributed to the type of data implored, this study was done with simulated multivariate climate datasets, further studies should be carried out on Extreme value analysis / theory (extreme data) and also the application of empirical climate datasets for further validation.

References

Abdulkadir, Y., & Sani, A. (2021). Analysis of climate data using spatial techniques to estimate rainfall in the North West of Nigeria. *FUDMA Journal of Sciences, 5*(4), 174–181.

Ademuwagun, A. A., Suleiman, R. T., Olumuyiwa, S. A., Oyedeji, M. B., & Afolabi, R. T. (2024). Application of change point analysis to Nigerian rainfall data. *The Journal of the Mathematical Association of Nigeria Ogun State (JOMANOGS): Mathematical Sciences and Education Series, 6*(2), 94–108.

Arias, L. A., Oosterlee, C. W., & Cirillo, P. (2023). Analytic isolation and distance-based anomaly detection algorithm. *Pattern Recognition, 141*, 109607.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18*, 1–22.

Beaulieu, C., Chen, J., & Sarmiento, J. L. (2012). Change-point analysis as a tool to detect abrupt climate variations. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370*(1962), 1228–1249.

deYoung, B., Barange, M., Beaugrand, G., Harris, R., Perry, R. I., Scheffer, M., & Werner, F. (2008). Regime shifts in marine ecosystems: Detection, prediction and management. *Trends in Ecology & Evolution, 23*(7), 402–409.

Gomez, C., White, J. C., & Wulder, M. A. (2011). Characterizing the state and process of change in a dynamic forest environment using hierarchical spatio-temporal segmentation. *Remote Sensing of Environment, 115*, 1665–1679.

Hawkins, D. M. (2001). Fitting multiple change-point models to data. *Computational Statistics & Data Analysis, 37*(3), 323–341.

Haynes, K., Eckley, I. A., & Fearnhead, P. (2017). Computationally efficient changepoint detection for a range of penalties. *Journal of Computational and Graphical Statistics, 26*(1), 134–143.

Intergovernmental Panel on Climate Change. (2019). *Special report on the ocean and cryosphere in a changing climate*. Cambridge University Press.

Killick, R., & Eckley, I. A. (2014). *Changepoint: An R package for changepoint analysis*. *Journal of Statistical Software, 58*(3), 1–19.

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association, 107*(500), 1590–1598.

Kouakou Innocent N'dri, & Nadarajah, S. (2022). Statistical modeling of monthly maximum temperature in Senegal. *Environmental Research Communications, 4*, 075002.

Lenton, T. M. (2013). Environmental tipping points. *Annual Review of Environment and Resources, 38*, 1–29.

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining* (pp. 413–422). IEEE.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika, 41*(1–2), 100–115.

Parth, S., & Mohdzuned, S. (2020). A review on climate change impact analysis using statistical tools. In *25th International Conference on Hydraulics, Water Resources and River Engineering (Hydro 2020)* (pp. 16–18). NIT Rourkela, India.

Reeves, J., Chen, J., Wang, X. L., Lund, R., & Lu, Q. Q. (2007). A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology, 46*(6), 900–915.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., et al. (2009). Early-warning signals for critical transitions. *Nature, 461*(7260), 53–59.

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing, 167*, 107299.

Zwiers, F. W., & von Storch, H. (2004). On the role of statistics in climate research. *International Journal of Climatology*, *24*(6), 665–680.