



# Homogenized climatological series for temperature and precipitation in Belgium

## Product User Guide

Issued by:

RMIB / Mel Brehon, Romain Ingels

BCC / Thomas Muller

VITO / Hendrik Wouters, Nele Veldeman, Filip Lefebvre

Date: 17/09/2026

Ref: C3S2\_461\_NCP-Action2024-02\_PUG-NOISELESS\_v2

Official reference number service contract: 2025/C3S2\_461-2\_BE\_RMIB-BCC/NCP-  
Action\_2024-2



*This document has been produced in the context of the Copernicus Climate Change Service (C3S). The activities leading to these results have been contracted by the European Centre for Medium-Range Weather Forecasts, operator of C3S on behalf of the European Union (Delegation Agreement signed on 11/11/2014). All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and the European Centre for Medium-Range Weather Forecasts has no liability in respect of this document, which is merely representing the authors view.*



## Contributors

### **RMIB**

Mel Brehon  
Romain Ingels

### **BCC**

Thomas Muller

### **VITO**

Hendrik Wouters  
Nele Veldeman  
Filip Lefebvre



## Table of Contents

|                                                      |           |
|------------------------------------------------------|-----------|
| <b>List of Figures</b>                               | <b>5</b>  |
| <b>1. Introduction</b>                               | <b>6</b>  |
| 1.1. Executive Summary                               | 6         |
| 1.2. Scope of Documentation                          | 6         |
| 1.3. Version History                                 | 6         |
| <b>2. Product Description</b>                        | <b>7</b>  |
| 2.1. Product Target Requirements                     | 7         |
| 2.1.1. Objective                                     | 7         |
| 2.1.2. Spatial and temporal extent                   | 7         |
| 2.1.3. Versioning and update protocol                | 7         |
| 2.1.4. Data access                                   | 8         |
| 2.2. Product Overview                                | 8         |
| 2.2.1. Data Description                              | 8         |
| 2.2.2. Variable Description                          | 9         |
| 2.2.3. Temperature variables                         | 10        |
| 2.2.4. Precipitation variable                        | 10        |
| 2.2.5. Data usage information                        | 10        |
| 2.2.6. Naming convention                             | 10        |
| 2.2.7. Temporal resolution                           | 10        |
| 2.2.8. Interpolation flag                            | 10        |
| 2.2.9. Geographical projection                       | 10        |
| 2.3. Input Data                                      | 10        |
| 2.3.1. Climate observations                          | 10        |
| 2.3.2. Event metadata                                | 11        |
| 2.4. Method                                          | 11        |
| 2.4.1. Clustering of series                          | 11        |
| 2.4.2. Control of outliers                           | 12        |
| 2.4.3. Homogenization                                | 13        |
| 2.4.4. Selection of reference series (neighbourhood) | 13        |
| 2.4.5. Break detection and scoring                   | 14        |
| 2.4.6. Correction                                    | 15        |
| 2.4.7. Iteration strategy                            | 15        |
| 2.4.8. Regionalization                               | 16        |
| 2.4.9. Gridding of anomalies                         | 16        |
| 2.4.10. Computation of regional means                | 17        |
| <b>3. Concluding Remarks</b>                         | <b>17</b> |
| <b>References</b>                                    | <b>18</b> |
| <b>Glossary</b>                                      | <b>19</b> |



|                             |           |
|-----------------------------|-----------|
| <b>General definitions</b>  | <b>19</b> |
| <b>Appendix A</b>           | <b>21</b> |
| <b>Appendix B</b>           | <b>23</b> |
| <b>Dataset Overview</b>     | <b>23</b> |
| <b>Rationale</b>            | <b>23</b> |
| <b>Correction Factors</b>   | <b>23</b> |
| <b>Selection guidelines</b> | <b>23</b> |
| <b>Best practices</b>       | <b>24</b> |

## List of Figures

|                                                                                                            |   |
|------------------------------------------------------------------------------------------------------------|---|
| Figure 1. Homogenized temperature timeseries.....                                                          | 9 |
| Table 1. List of data products.....                                                                        | 8 |
| Table 2. Overview of key characteristics of the homogenized temperature and precipitation dataset<br>..... | 8 |
| Table 3. Overview and description of climate variables.....                                                | 9 |



## 1. Introduction

### 1.1. Executive Summary

Reliable assessment of recent climate trends requires observed climate time series in which variations reflect only climatic changes. Such records are referred to as homogeneous or homogenized climate series. However, weather stations are subject to various non-climatic influences - including instrument changes, modifications in observing practices, relocations or surrounding changes – that can introduce artificial discontinuities known as inhomogeneities or breaks into observational data.

To address this challenge, long-term temperature and precipitation time series in Belgium were processed to produce homogenized datasets suitable for robust climate trend analysis. These corrected data serve as essential reference material for assessing Belgian climate trends. It is important to note that the implemented corrections target abrupt shifts in the climate series but do not address gradual biases.

Regional climate series are also provided as supplementary information on common climate signals over broad geographic domains. Users should note that the reliability of such regional climatological series depends heavily on both the variable analyzed and the region defined. Consequently, we recommend that users without specialized climate expertise avoid drawing broader conclusions from regional series alone. The methodology for producing regional series is documented to demonstrate how individual station-level homogenized series can be combined to extract climate information at larger spatial scales. We encourage users to construct their own regional series based on the boundary definitions most relevant to their research objectives.

### 1.2. Scope of Documentation

This document provides technical guidance on the homogenized temperature and precipitation series for Belgium (NOISELESS) dataset and lists the platforms where the dataset can be accessed. Specifically, it includes:

- A description of the data products, including station level and regional series
- Detailed documentation of the homogenization methodology, including clustering procedures, outlier control, break detection, and correction protocols
- Information on input data aggregation, even metadata classification, temporal and spatial coverage

### 1.3. Version History

This is the first version of the dataset.



## 2. Product Description

### 2.1. Product Target Requirements

#### 2.1.1. Objective

The homogenized climate series support the evaluation and comparison of recent temperature and precipitation trends in Belgium within the context of climate change. Specifically, the dataset addresses two analytical goals:

- **Station-level analysis:** Homogenized series for individual stations capture local climate evolution at specific observation locations.
- **Regional-scale analysis:** Aggregated regional series isolate the common climate signal across broader spatial domains.

#### 2.1.2. Spatial and temporal extent

The dataset aligns with the availability of observational data across Belgium:

- Temporal coverage: 1880-2024
- Temporal resolution: Monthly
- Spatial coverage: National network spanning all elevations and divided into three elevation-based regions
  - o Low-Belgium: Up to 50 meters (m)
  - o Middle-Belgium: Between 50 and 200 m
  - o High-Belgium: Above 200 m

Although Belgium's geographic conventions typically define the Low-Belgium boundary at 100 m, applying this threshold would have yielded an uneven distribution of source stations across the three regions, resulting in inconsistent quality levels among the regional series. To ensure balanced representation, the Low-Belgium threshold was lowered to 50 m.

#### 2.1.3. Versioning and update protocol

Belgian temperature and precipitation series were previously homogenized up to 2015 under an earlier RMIB project, using a distinct homogenization methodology. The present release is **version 1 of a new generation of homogenized data**, and should not be named as version 2 of the earlier dataset. Although this generation builds on the previous homogenization phase as a starting point, the two differ in methodology. In addition, current series extend the record to 2024, incorporate event metadata newly collected for the 2015-2024 period and recovered for earlier periods. This reflects a full re-evaluation of the previous homogenization adjustments.

Within this new generation, the published dataset should be regarded as a provisional release that may be revised in subsequent versions as additional station history information becomes available or methodological refinements are implemented. Updates to the homogenized series are anticipated every 2-5 years (cf Table 2).



2.1.4. Data access

The datasets provided datasets and future updates will be openly accessible through three platforms:

- Copernicus Climate Change Service (C3S) Climate Data Store (CDS)
- Royal Meteorological Institute of Belgium (RMIB) open data platform
- Belgian Climate Centre (BCC) open data platform.

2.2. Product Overview

2.2.1. Data Description

The list of the data products covered by this document is given in Table 2.

Table 1. List of data products

| Product title                                                         | Source data                                                          |
|-----------------------------------------------------------------------|----------------------------------------------------------------------|
| Homogenized monthly temperature time series for individual stations   | 16 historical extreme minimum temperature (TN) series from 1880-2024 |
|                                                                       | 15 historical extreme maximum temperature (TX) series from 1880-2024 |
|                                                                       | 60 long extreme minimum temperature (TN) series from 1954-2024       |
|                                                                       | 60 long extreme maximum temperature (TX) series from 1954-2024       |
| Homogenized monthly precipitation time series for individual stations | 24 historical precipitation series from 1880-2024                    |
|                                                                       | 158 long precipitation series from 1951-2024                         |
| Regional monthly temperature series                                   | 4 extreme minimum temperature (TN) series from 1954-2024             |
|                                                                       | 4 extreme maximum temperature (TX) series from 1954-2024             |
| Regional monthly precipitation series                                 | 4 precipitation series from 1951-2024                                |

Characteristics of the **dataset homogenized temperature and precipitation records for Belgium (NOISELESS)** are given in Table 2. As an example, one indicator is visualized in Figure 1.

Table 2. Overview of key characteristics of the homogenized temperature and precipitation dataset

| Data Description |                                                                           |
|------------------|---------------------------------------------------------------------------|
| Dataset title    | homogenized temperature and precipitation records for Belgium (NOISELESS) |
| Data type        | Homogenized climate observations                                          |
| Topic category   | observations                                                              |
| Sector           | Climate evidence                                                          |
| Keyword          | In situ, observations, homogenization, Belgium                            |



|                       |                                                                                                                                    |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------------|
| Dataset language      | English                                                                                                                            |
| Domain                | Belgium                                                                                                                            |
| Horizontal resolution | -                                                                                                                                  |
| Temporal coverage     | 1880-2024                                                                                                                          |
| Temporal resolution   | monthly                                                                                                                            |
| Vertical coverage     | 0 to 2 metres (variable dependent)                                                                                                 |
| Update frequency      | From 2 to 5 years;                                                                                                                 |
| Version               | V1                                                                                                                                 |
| Provider              | Royal Meterological Institute of Belgium (RMIB), Flemish Institute for Technological Research (VITO), Belgian Climate Centre (BCC) |
| Terms of Use          | No conditions to access and use                                                                                                    |

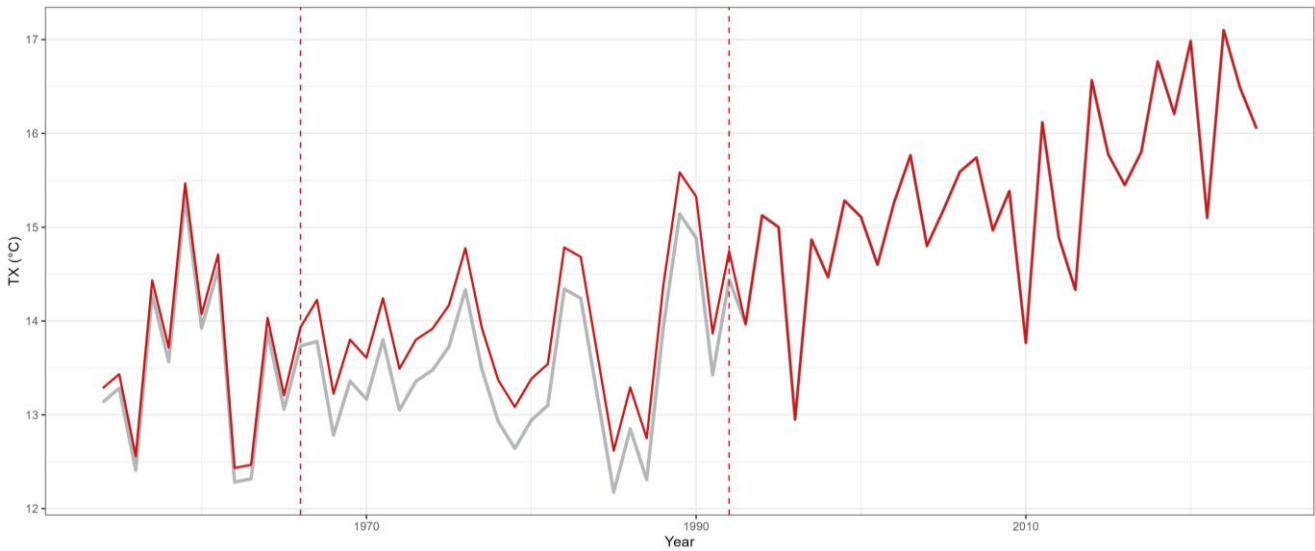


Figure 1. Homogenized temperature timeseries

Comparison between homogenized (red line) with original data (grey line) for the KORTENAKEN (series code 2039) maximal temperature (TX) annual series (code 2039). Dotted lines represent the occurrence dates of the corrected inhomogeneities.

### 2.2.2. Variable Description

The dataset provides homogenized observations of climate variables summarized in Table 3: monthly extreme temperatures and accumulated precipitation. All daily observations were aggregated into monthly statistics (means for temperature, sums for precipitation).

Table 3. Overview and description of climate variables

| Variable                                                    | Units | Definition                                                                     |
|-------------------------------------------------------------|-------|--------------------------------------------------------------------------------|
| Monthly accumulated precipitation                           | mm    | Monthly sum precipitation                                                      |
| Monthly average daily minimum and daily maximum temperature | degC  | Monthly mean of the daily maximum and minimum temperature at 2 m above surface |



### 2.2.3. Temperature variables

Monthly minimum and maximum temperatures represent the monthly means of daily minimum and maximum values measured at 2 meters above ground level under standard shelter conditions. Monthly values are derived by aggregating daily observational records. Daily observations follow a 8:00 local time (LT) cycle:

- Daily Maximum ( $T_x$ ): Highest temperature recorded between 8:00 LT on day  $d$  and 8:00 LT on day  $d + 1$ .
- Daily Minimum ( $T_n$ ): Lowest temperature recorded between 8:00 LT on day  $d - 1$  and 8:00 LT on day  $d$ .

### 2.2.4. Precipitation variable

Monthly accumulated precipitation represents the water-equivalent volume collected at the gauge surface per month. They are obtained by summing the daily observed precipitation amounts. Daily measurements follow the same 8:00 LT observation cycle as temperature: the value logged for day  $d$  corresponds to the sum of precipitation recorded between 08:00 LT on day  $d$  and 08:00 LT on day  $d + 1$ .

### 2.2.5. Data usage information

### 2.2.6. Naming convention

The homogenized series for individual stations are delivered with the names and geographical coordinates of their reference location. The reference location corresponds to the weather station associated with the final segment of the climate series.

### 2.2.7. Temporal resolution

Homogenized series for individual stations and regional series are provided at a monthly resolution.

### 2.2.8. Interpolation flag

Each homogenized series for individual station includes an interpolation flag documenting the interpolation status of monthly values. In a homogenized series, a flag value of 1 indicates that the original observed value was missing and that the corresponding homogenized value was interpolated. In an original (non-homogenized) series, the flag value is set to 0 for every time step by default.

### 2.2.9. Geographical projection

The geographical coordinates of each homogenized series (longitude and latitude in decimal degrees) are provided relative to the WGS84 coordinate reference system.

## 2.3. Input Data

### 2.3.1. Climate observations

#### Aggregation into monthly series



Daily observations were aggregated into monthly statistics: means for temperature variables and sums for precipitation. A monthly value was computed only when the count of days with available data exceeded 90% of the total days in that month. Otherwise, the monthly record was flagged as missing.

### Merging of series

Stations were eligible for series merging based on geographical proximity and correlation. To qualify for merging, the candidate stations required:

- A horizontal distance of a maximum of 20 km and a vertical distance of a maximum of 70 m.
- A Pearson correlation coefficient between first-difference series of minimum 0.96.
- Minimum overlap of one year

Exceptions were permitted for overlaps of 11 months, provided that the horizontal and vertical distances between stations were sufficiently small and that the correlation over this period was sufficiently high.

### 2.3.2. Event metadata

Event metadata documents the date and location of events that impact a weather station and may induce inhomogeneities in observational records. The event metadata was collected and classified into the following categories:

- Station relocation
- Observer change (applicable to manual stations)
- Instrument change
- Change in local environment/surroundings
- Data series merging
- Transmission failure
- Instrument maintenance
- Others

## 2.4. Method

### 2.4.1. Clustering of series

Series were clustered based on similarities in climate patterns (specifically temporal variations rather than overall temperature means) and altitude range.

#### Cluster construction and dissimilarity measures

A dissimilarity metric  $A$  was calculated for all pairs of stations within the Belgian network and selected foreign series. This weighted metric combines correlation-based distance with altitude difference and is defined as:

$$A = \alpha \times (1 - \rho_{xy}) + \beta \times |z_x - z_y|$$

Where  $\rho_{xy}$  is the Pearson correlation coefficient between the first-difference series of stations  $x$  and  $y$ ,  $z_x$  and  $z_y$  are their respective altitudes, and  $\alpha$  and  $\beta$  are weighting coefficients.



A generalized Ward clustering procedure was applied to derive a cluster structure from the dissimilarity matrix. This agglomerative hierarchical method starts with  $n$  clusters, each containing a single station. At each step, the two clusters for which the merging results in the minimum increase in total within-cluster variance (measured by the sum of squared distance, with distance being represented here by dissimilarity  $A$ ) are merged. This process repeats until all the stations are included in a single cluster.

### Clustering constraints

To obtain the optimal cluster structure, the following constraints were applied to adjust the clusters:

- Cluster size: Clusters are limited to a maximum of 15 series whenever possible. These constraints align with recommendations by Domonkos (2017), who suggest that HOMER is best suited for networks of up to 12-15 series.
- Temporal coverage: Stations included in a cluster must have at least 15 years of data available.
- Correlation thresholds:
  - o Standard requirement: Pearson correlation  $\rho \geq 0.9$  on monthly deseasonalized anomalies (derived from first-difference series).
  - o Exceptions: In mountainous and coastal regions, a threshold of respectively 0.85 and 0.89 was accepted.
- Quality control: Within a cluster, monthly boxplots were visually inspected for each station relative to the other stations of the cluster.

### 2.4.2. Control of outliers

Before homogenization, control for outliers is applied at the cluster scale. In this context, "outlier" does not refer to general time series outliers, but to data points exhibiting irregularities relative to the homogenization process. To identify these, two anomaly metrics are used:

- Type 1 anomalies (Deseasonalized): deviations relative to the station's own monthly mean calculated over the entire observation period
- Type 2 anomalies (Anomalies of anomalies): deviations of the Type 1 anomalies relative to their mean across all stations within the cluster.

Both annual and monthly anomalies are computed, alongside their medians. Annual and monthly modified z-score  $Z$  are derived to quantify the deviations:

$$Z = 0.6745 \times \frac{X - M}{MAD}$$

Where  $Z$  is the modified z-score,  $X$  the data value,  $M$  the associated median and  $MAD$  the median absolute deviation. Two thresholds are defined for the modified z-score:  $thresh_1$  for Type 1 anomalies and  $thresh_2$  for Type 2 anomalies, initially set to 3.5 standard deviation units. A data point is flagged as a potential outlier if both its Type 1 and Type 2 z-score exceed their respective thresholds. The flagging and validation procedure follows the steps:

1. Anomaly calculation:  
For every station, Type 1 and Type 2 anomalies are calculated.
2. Initial screening:



Data points are flagged as potential outliers if they simultaneously exceed the defined thresholds ( $thresh_1$  and  $thresh_2$ , default=3.5) for Type 1 and Type 2 anomalies.

3. Expert review

- If neighboring stations show high variance (i.e a wide range of anomaly values) in their Type 1 or Type 2 anomalies at the target date, the outlier flag is removed. This suggests a regional climatic event rather than a measurement error.
- If neighboring stations do not show high variance, the candidate undergo further expert review. The candidate temperature is compared against neighboring station temperatures, taking local sub-climate characteristics into account.

4. Final decision:

If the candidate persists as a divergence from neighbors, then the flag is retained. The data point is retained in the series. The flag serve to guide the expert on how to treat the value in the homogenization process, rather than deleting the data point.

### 2.4.3. Homogenization

Monthly extreme temperature and precipitation series were homogenized using a method based on the HOMER software (Mestre et al., 2013). Developed under the COST Action ES0601—Advances in Homogenization Methods of Climate series: an Integrated Approach (HOME)— HOMER combines three state-of-the-art detection methods: PRODIGE (Caussinus & Mestre, 2004), Segmentation Methods for Array CGH Analysis (CGHSEG; Picard et al., 2011) and ACMANT (Domonkos, 2011). Based on the identified inhomogeneities, a network-wide unified correction model BENOVA (initially named ANOVA; Caussinus & Mestre, 2004) is applied to correct the series.

BART (Joelsson et al., 2021), developed by the Swedish Meteorological and Hydrological Institute (SMHI), is a fully automated version of HOMER. It introduces the integration of event metadata and additional functionalities that significantly improve time efficiency in HOMER methodology. However, automation comes at the cost of losing the expert judgment in validating detected inhomogeneities. To preserve this expertise while still benefiting from BART's improvements, selected BART functions were incorporated into HOMER, resulting in a semi-automated version of the software. This hybrid approach, which can be positioned between the original HOMER and BART, was used to homogenize the Belgian climate series. Its detailed specifications will be described in a forthcoming report comparing homogenization techniques.

### 2.4.4. Selection of reference series (neighbourhood)

The general homogenization workflow iterates through three detection methods to compute break scores based on neighbor comparisons and event metadata. A hybrid procedure developed in BART is used to select the neighboring reference stations for each candidate series. For this, a hierarchical ranking system first categorizes potential references based on correlation thresholds, then sorts them by geographical distance within each category. From this ranking, a number of reference stations are selected, respecting a minimum of 6 references series per candidate. Across all three detection



methods, inhomogeneities are identified in each candidate series relative to its corresponding reference series. Thus, detection is performed in a localized neighborhood within the cluster.

#### **2.4.5. Break detection and scoring**

The detection of breaks combines three different approaches: Pairwise detection from PRODIGE, Joint detection from CGHSEG, and ACMANT detection.

##### **Pairwise detection**

In Pairwise detection derived from PRODIGE, a candidate series is compared against its reference series by computing difference series. These difference series are then tested for discontinuities. While a detected discontinuity may emerge from a change in the candidate or in the neighbor series, the method relies on a consistency principle: if a break signal persists consistently across comparisons between the candidate station and multiple distinct neighbors, then the break signal can be attributed to the candidate station itself. Pairwise detection can be applied to both annual and seasonal series.

When comparing a candidate series to a reference series in pairwise detection, the Minimum Detectable Magnitude (MDM) refers to the estimated standard deviation of the residuals computed for that specific candidate-reference pair. An MDM value is computed for each candidate-reference pair and may differ between the annual and seasonal series. Similarly to Delvaux et al. (2019), we treat the MDM as the significance threshold for pairwise break detection. Consequently, break signals with amplitudes lower than the corresponding MDM are disregarded.

Additionally, maximum noise thresholds were defined by the expert for annual and seasonal MDM values independently. Only breaks signals associated with an MDM value falling below these thresholds were retained for the analysis.

##### **Joint detection**

Derived from CGHSEG package, the joint detection method in HOMER uses a dynamic approach to search for a configuration of breaks on a set of series. First, all optimal solutions for each candidate series are searched separately. Then, the configuration of inhomogeneities is optimized, such that the internal variance of the homogeneous sub-series is minimized (Caussinus & Lyazrhi, 1997).

##### **ACMANT detection**

ACMANT detection in HOMER is derived from ACMANT software (Domonkos, 2011). Seasonal cycles in climate variables can induce seasonal patterns in inhomogeneities. A unified test, applied to both the annual mean series and on the amplitude of the seasonal cycle of the climate variable, enables the identification of inhomogeneities with pronounced seasonal behavior. Since this detection method is not intended for cumulative variables, ACMANT was not applied to the homogenization of Belgian precipitation series.

##### **Break scoring**



Potential breaks are identified through detection rounds using the three detection methods. After a detection round, each method contributes to a weighted score used to filter potential breaks. This scoring system, adapted directly from BART, allows the integration of event metadata. To do so, an additional weight is attributed to break signals occurring near known station history events. A potential break is flagged if its score meets or exceeds the acceptance score of **12 points**. Points are allocated based on the detection method as follows:

- Pairwise detection (Annual): 3 points
- Joint detection: 3 points
- ACMANT detection: 3 points
- Event metadata: 3 points
- Pairwise detection (Seasonal): 1 point

Prior to scoring, specific break signals were excluded from the analysis based on the following criteria:

- **Insufficient amplitude:** Breaks with an amplitude below a threshold set by the expert. This threshold was set to **0.2 °C** for temperature series, and **0.05%** for precipitation series.
- **Excessive noise (high MDM):** Breaks associated with a Minimum Detectable Magnitude (MDM) exceeding the noise threshold set by the expert.
  - o Annual pairwise detection: Thresholds were set to **0.25°C** for temperatures and **0.1%** for precipitation.
  - o Seasonal pairwise detection: Thresholds were set to **0.40°C** for temperatures and **0.2%** for precipitation.
- **Insufficient signal-to-noise ratio:** Breaks with an amplitude lower than its associated MDM

### Expert verification and break validation

Despite algorithmic filtering, a manual review of the breaks detected remains essential. Expert evaluates the list of candidate breaks, along with their calculated score, and select those judged plausible. The validation protocol follows the recommendations established by Domonkos (2017) for the use of HOMER:

- Primary evidence: Break signals from Pairwise Detection serve as the principal evidence of inhomogeneity
- Supported evidence: Break signals from Joint and ACMANT detections are used only to resolve ambiguities arising from Pairwise detection results.

### 2.4.6. Correction

A BENOVA correction model is applied to correct inhomogeneities at the cluster scale. The method, initially called ANOVA, uses the ordinary least squares technique for the joint estimation of the correction factors. The correction method adjusts the data preceding a validated break such that the most recent data of the series is not modified. As the method has few to do with the similarly named ANOVA method for analysis of variance, its new name BENOVA introduced by Domonkos & Joelsson (2024) is used here.

### 2.4.7. Iteration strategy

The homogenization followed an iterative detection-correction cycle of three phases:



- Initial round: focuses on inhomogeneities supported by metadata that exhibit strong detection signals (characterized by a significant amount of break indicators rather than necessarily high break amplitudes).
  - Pairwise detection (annual and seasonal scales)
  - Joint detection (annual scale)
  - Expert validation of candidate breaks
  - Correction
- Intermediary rounds: Iterated until the homogenized series indicate satisfactory quality.
  - Pairwise detection (annual and seasonal scales)
  - Joint detection (annual scale)
  - ACMANT detection (seasonal scale)
  - Expert validation of candidate breaks
  - Correction
- Final round: Refines the correction dates
  - Pairwise detection (annual and seasonal scales)
  - Joint detection (annual scale)
  - ACMANT detection (seasonal scale)
  - Month assessment for correction: adjust the break timing to the most likely month of a break
  - Correction

Temperature series underwent homogenization at both annual and seasonal scales, following the complete workflow described above. Precipitation series, in contrast, were homogenized on an annual scale only. Consequently, ACMANT seasonal detection and monthly date refinement were not applied.

#### 2.4.8. Regionalization

Regional series were developed for Belgium at two spatial scales: one national series and three sub-regional series. The long series previously homogenized were used to construct the regional series for temperature and precipitation. The geographical boundaries of the regions and the reference locations of the homogenized series are presented in Appendix I.

Regional temperature and precipitation series were constructed by first producing gridded data from the homogenized records, then computing spatial averages within each regional boundary for every monthly timestep throughout the observational period.

#### 2.4.9. Gridding of anomalies

A regular longitudinal-latitude grid with  $0.08^\circ$  resolution (approximately 9 km) was established over Belgium. Monthly anomalies relative to monthly reference means were interpolated onto this grid using Inverse Distance Weighting (IDW) with the following parameters:

- Neighborhood radius:  $0.49^\circ$  (approximately 55 km)
- Power exponent: 2



The reference means were derived from the reference period 1961-1990.

#### 2.4.10. Computation of regional means

For each region, a regional anomaly mean series was computed as the area-weighted average of gridded anomalies, where weights correspond to individual grid cell areas. Due to the longitude-latitude regular grid structure, grid cell area decreases towards the pole (here towards higher latitude), requiring area weighting.

Similarly, a regional reference mean was computed from the corresponding gridded reference values over the same spatial domain. The final regional absolute temperature series was then derived by adding the regional reference mean to the regional anomaly series.

### 3. Concluding Remarks

The **dataset the dataset homogenized temperature and precipitation records for Belgium (NOISELESS)** supports the evaluation and comparison of recent temperature and precipitation trends in Belgium. By providing both homogenized individual station series across Belgium, the dataset enables consistent assessment and comparison of local climate evolutions across the country.

Regional series can be built from homogenized individual station series to isolate common climate signals across broader spatial domains. Users are encouraged to develop their own regional series based on boundary definitions aligned with their research objectives, leveraging the document methodology.

As the dataset is considered provisional, future updates may incorporate new station history information or methodological refinements. Subsequent versions are anticipated every 2-5 years. Users are advised to consult version notes and update logs when accessing the data.



## References

- Caussinus, H., & Lyazrhi, F. (1997). Choosing a linear model with a random number of change-points and outliers. *Annals of the Institute of Statistical Mathematics*, 49(4), 761–775. <https://doi.org/10.1023/A:1003230713770/METRICS>
- Caussinus, H., & Mestre, O. (2004). Detection and Correction of Artificial Shifts in Climate Series. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 53(3), 405–425. <https://doi.org/10.1111/J.1467-9876.2004.05155.X>
- Delvaux, C., Ingels, R., Vrábel, V., Journée, M., & Bertrand, C. (2019). Quality control and homogenization of the Belgian historical temperature data. *International Journal of Climatology*, 39(1), 157–171. <https://doi.org/10.1002/joc.5792>
- Domonkos, P. (2011). Adapted Caussinus-Mestre Algorithm for Networks of Temperature series (ACMANT). *International Journal of Geosciences*, 2(3), 293–309. <https://doi.org/10.4236/IJG.2011.23032>
- Domonkos, P. (2017, April). Time series homogenization with optimal segmentation and ANOVA correction: Past, present and future. *Proceedings of the Ninth Seminar for Homogenization and Quality Control in Climatological Databases and Fourth Conference on Spatial Interpolation Techniques in Climatology and Meteorology*.
- Domonkos, P., & Joelsson, L. M. T. (2024). ANOVA (Benova) correction in relative homogenization: Why it is indispensable. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 44(12), 4515–4528. doi:10.1002/joc.8594
- Joelsson, L. M. T., Sturm, C., Södling, J., Engström, E., & Kjellström, E. (2021). Automation and evaluation of the interactive homogenization tool HOMER. *International Journal of Climatology*, 42(5), 2861–2880. <https://doi.org/10.1002/joc.7394>
- Mestre, O., Domonkos, P., Picard, F., Auer, I., Robin, S., Boehm, R., Aguilar, E., Guijarro, J., Vertachnik, G., Klancar, M., Dubuisson, B., & Stepanek, P. (2013). HOMER : a homogenization software—Methods and applications. *IDOJARAS*, 117(1), 47–67.
- Picard, F., Lebarbier, E., Hoebeke, M., Rigai, G., Thiam, B., & Robin, S. (2011). Joint segmentation, calling, and normalization of multiple CGH profiles. *Biostatistics*, 12(3), 413–428. <https://doi.org/10.1093/biostatistics/kxq076>



## Glossary

| Acronym | Description                                 |
|---------|---------------------------------------------|
| C3S     | Copernicus Climate Change Service           |
| CDS     | Climate Data Store                          |
| MDM     | Minimum Detectable Magnitude                |
| CGHSEG  | Segmentation Methods for Array CGH Analysis |

### General definitions

**Inhomogeneity** or **Break**: Artificial shift in a climate series due to non-climatic changes such as changes in measuring instruments, station location, or observation practices. An inhomogeneity can be abrupt or gradual shifts and can lead to inaccurate climate analyses.

**Homogeneous** or **Homogenized climate time series**: Climate time series where variations in the climate variable are caused only by variations in climate.

**Homogenization**: Process by which artificial discontinuities are detected and removed from climatological series.

**Long Belgian climate series** : Long-term Belgian climatological series starting from 1950s to 2024.

**Historical Belgian climate series** : Long-term Belgian climatological series starting from 1880s to 2024.



## Appendix A

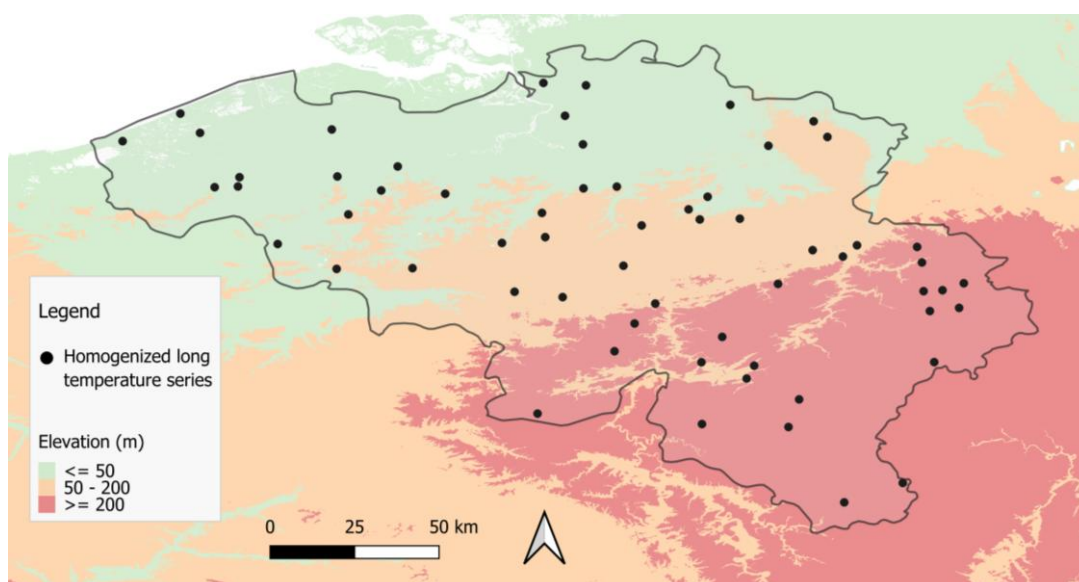


Figure A1. Homogenized long temperature series and contours of elevation-based Belgian sub-regions.

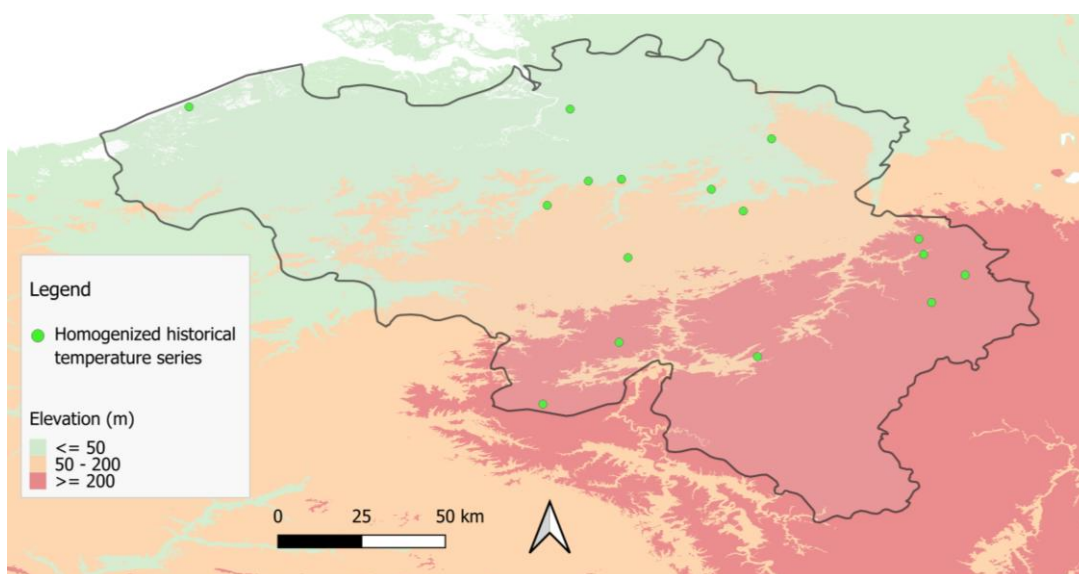


Figure A2. Homogenized historical temperature series.

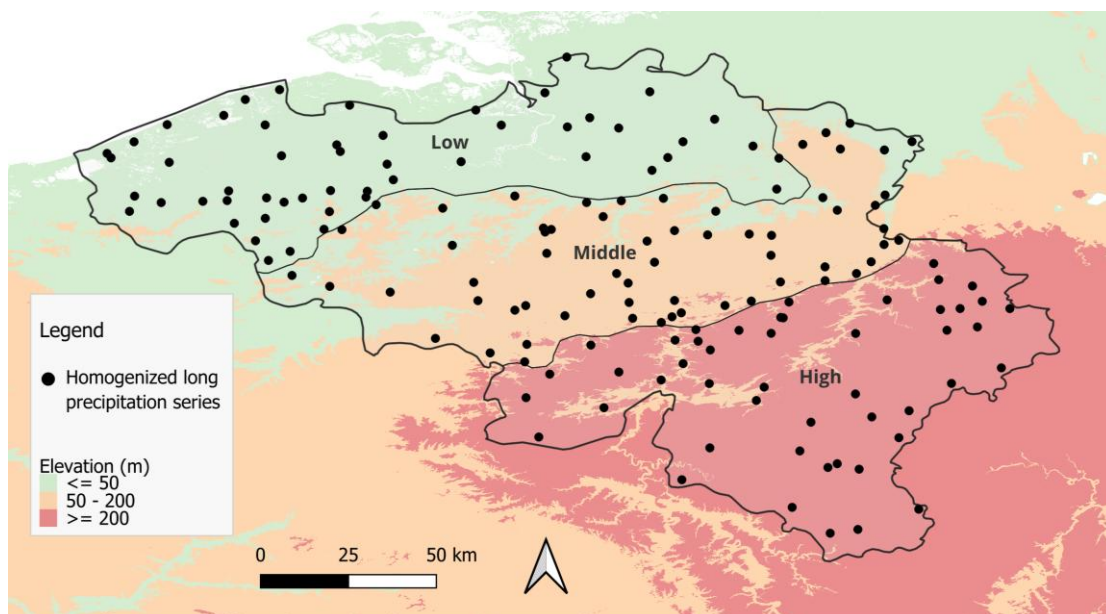


Figure A3. Homogenized long precipitation series and contours of elevation-based Belgian sub-regions.

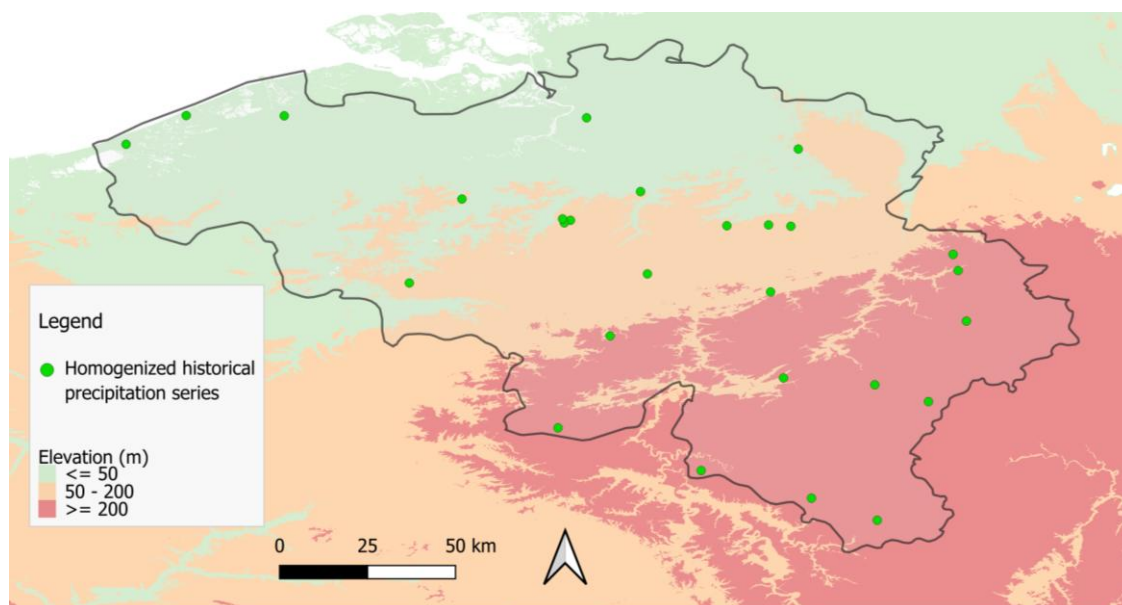


Figure A4. Homogenized historical precipitation series.



## Appendix B

### Homogenized Series Periods

#### Dataset Overview

The NOISELESS dataset consists of homogenized climate time series derived from the same meteorological stations but organized into two distinct versions:

| Version | Coverage                                                   | Recommended use                                                           |
|---------|------------------------------------------------------------|---------------------------------------------------------------------------|
| Recent  | Since 1954 for temperature<br>Since 1951 for precipitation | Recent trend analysis<br>(post-1954 temperature, post-1951 precipitation) |
| Early   | Since 1880                                                 | Early trend analysis (pre-1954 temperature, pre-1951 precipitation)       |

#### Rationale

Homogenization quality scales with **cluster density**: more neighboring series enable better detection and correction of inhomogeneities. The Recent series leverage denser post-1950 station networks, while the Early series are optimized for sparser pre-1950 conditions.

#### Correction Factors

Correction dates are identical between the Recent and Early series for overlapping periods. However, the correction magnitudes themselves may differ slightly. This occurs because:

1. Correction factors are computed based on the other series present in each cluster
2. The composition of each cluster differs between Recent and Early versions (different number of neighboring series)
3. Therefore, even for the same correction event at the same date, the magnitude of adjustment slightly varies by version

The variation in correction magnitude reflects the difference in inputs to the correction algorithm.

#### Selection guidelines

| Your research focus                                               | Recommended dataset | Reason                                                    |
|-------------------------------------------------------------------|---------------------|-----------------------------------------------------------|
| Recent period<br>(post-1954 temperature, post-1951 precipitation) | Recent              | Higher station density<br>→ Higher homogenization quality |



|                                                                |        |                                                                                        |
|----------------------------------------------------------------|--------|----------------------------------------------------------------------------------------|
| Early period<br>(pre-1954 temperature, pre-1951 precipitation) | Early  | Only option for pre-1950 data, optimized for sparser conditions                        |
| Full record (1880-present)                                     | Hybrid | Early for pre-cutoff, Recent for post-cutoff.<br>Acknowledge potential junction shift. |

### Best practices

1. **Match version to period:** Avoid using Early series when Recent is preferable
2. **Document version choice:** Specify selected version and rationale in publications
3. **Test junction transitions:** if combining versions, check for shifts at the cutoff boundary



ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: [info@copernicus-climate.eu](mailto:info@copernicus-climate.eu)