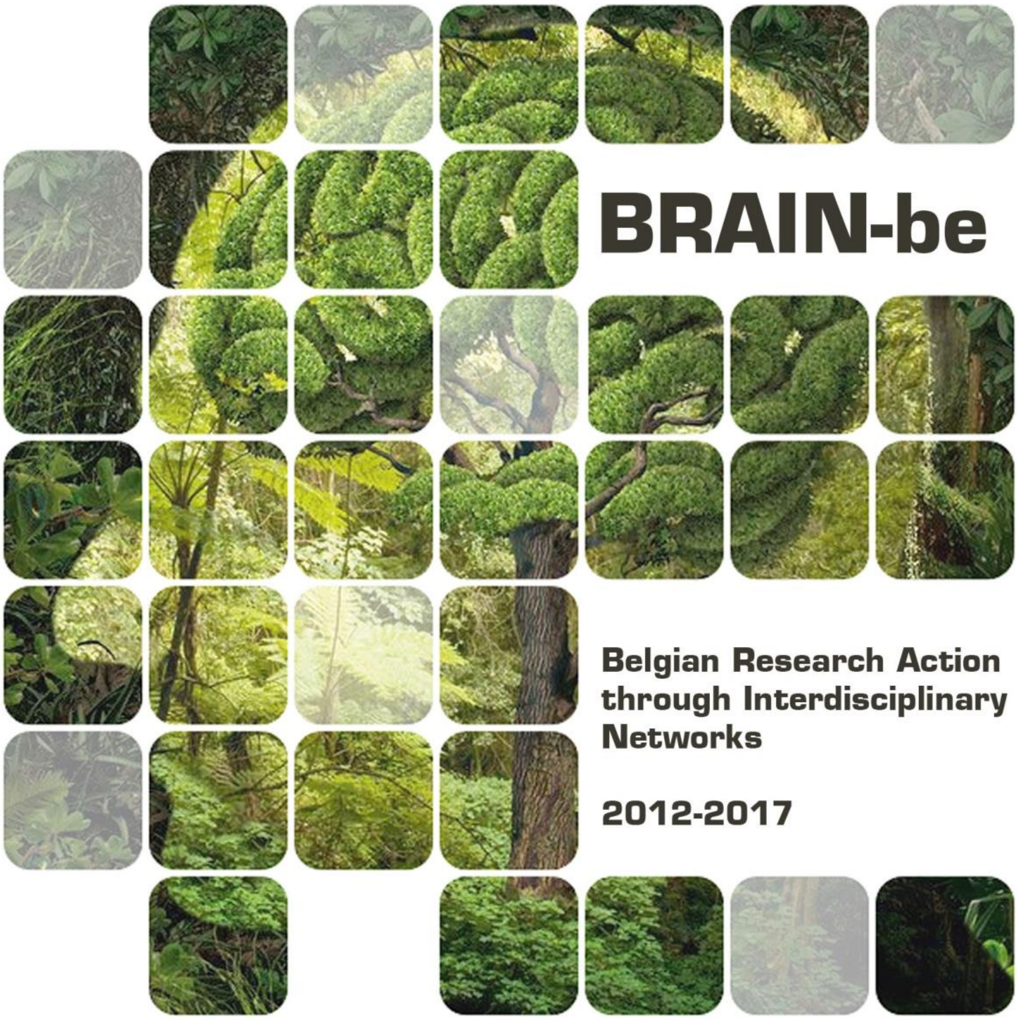


BEL-HORNET

Belgian homogenized long-term reference climate time series

Cedric BERTRAND (IRM/KMI) – Vladimír Vrábeĺ (IRM/KMI) – Charles Delvaux (IRM/KMI) – Michel ALIOSCHA-PEREZ (VUB) – Hichem SAHLI (VUB)

Axis 6: Management of collections

BRAIN-be

Belgian Research Action through Interdisciplinary Networks

2012-2017

NETWORK PROJECT

BEL-HORNET

Belgian homogenized long-term reference climate time series

Contract - BR/154/A6

FINAL REPORT

PROMOTORS: Cedric BERTRAND (IRM/KMI)
Hichem SAHLI (VUB)

AUTHORS: Cedric BERTRAND (IRM/KMI), Vladimír Vrábeĺ (IRM/KMI),
Charles DELVAUX (IRM/KMI) ,

Michel ALIOSCHA-PEREZ (VUB), Hichem SAHLI (VUB)

IRM

Vrije Universiteit Brussel

belspo .be

Published in 2018 by the Belgian Science Policy Office
Avenue Louise 231
Louizalaan 231
B-1050 Brussels
Belgium
Tel: +32 (0)2 238 34 11 - Fax: +32 (0)2 230 59 12
http://www.belspo.be
http://www.belspo.be/brain-be

Contact person: Maaike VANCAUWENBERGHE
Tel: +32 (0)2 238 36 78

Bertrand C., V. Vrábeĺ, C. Delvaux, M. Alioscha-Perez and H. Sahli. ***Belgian homogenized long-term reference climate time series***. Final Report. Brussels : Belgian Science Policy Office 2018 – 62 p. (BRAIN-be - (Belgian Research Action through Interdisciplinary Networks))

TABLE OF CONTENTS

ABSTRACT

Context

Long-term, high-quality and reliable instrumental climate records are indispensable pieces of information required for undertaking robust and consistent studies to better understand, detect, predict and respond to climate variability and change. However, as extensively discussed in literature most instrumental time series contain variations that are not only due to the vagaries of weather or climate. Causes for these variations are manifold, e.g. station relocation, instrumentation changes or changes in observation times. In addition wrong or aberrant observations are common in most observational systems. All these factors reduce the quality of original data and compromise their homogeneity. Because biases in time series frequently have a similar magnitude as the climate signal (i.e. long-term variations, trends or cycles), it is widely accepted that inhomogeneities and aberrant observations in time series have to be detected and if necessary adjusted before performing any kind of climate change analysis.

Objectives

Based on the meteorological records archived at the Royal Meteorological Institute of Belgium, long-term, high-quality and homogeneous temperature and precipitation time series will be produced for Belgium using state-of-the art data quality control procedures and homogenization methods.

Conclusions

A new dataset of quality controlled monthly and daily homogenised temperature and precipitation time series for Belgium has been produced for the period 1880 to 2015, although with a lower station density before the second half of the 20th century. This, in turn, will support century-scale analyses of changes in mean temperatures and precipitations, as well as of extremes temperatures.

Keywords

Time series, data quality control, data mining, artificial neural network, break detection, homogenization, daily extreme temperature, and daily precipitation amount

1. INTRODUCTION

In Belgium, the official climatological network started in 1870s under the auspices of the former Observatory of Brussels located in St-Josse-ten-Noode, Brussels. In 1913, the Royal Meteorological Institute (RMI) was set up at Uccle, Brussels, and the climatological network was put under its responsibility. Mainly relying on volunteers, the main climatic information collected across the time are the daily precipitation amount (RR) and the daily maximum (TX) and minimum (TN) temperature. RR, TX and TN, for the previous 24 hours, are recorded at 8:00 am local time. TN is recorded against the day of the observation while TX and RR against the previous day.

The network currently counts 145 thermometric stations (i.e. 115 manual and 30 automated stations) and 226 pluviometric stations (i.e. 196 manual and 30 automated stations) but the number of climatological stations maintaining daily precipitation amount and temperature records has varied over the time as illustrated in Figure 1. Starting in 1911, an abrupt increase followed a call for collaboration published in the national newspapers. The negative impact of the war periods on the climatological network is also clearly visible in Figure 1. A total of about 1060 stations were identified in the RMI's central database (DB) for the period 1880-2015 with a minimum of 48 operating stations in 1880 and a maximum of 358 operating stations in 1976. (The number of climatological stations maintaining temperature records has varied over time with a minimum of 25 stations in 1925 and a maximum of 177 stations in 2006).

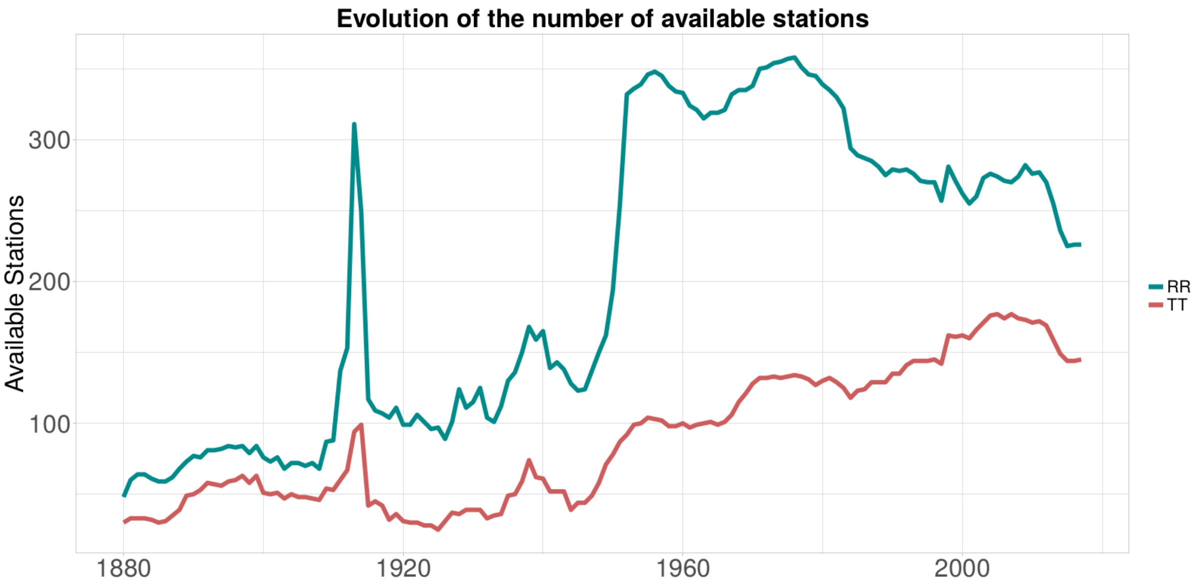


FIGURE 1: Time evolution of the number of available pluviometric (in blue) and thermometric (in red) stations within the central RMI's database.

Huge amounts of climate data have been recorded since the early stage of the climatological network. However, the existing data heritage is largely under-exploited because historical information still remained in hard copy and in fragile media (e.g. data hand-written kept in the original daily weather reports). Easily accessible digital climate data are mostly restricted to the second half of the 20th century. While instrumental data extend back in time at least to the 19th century over most regions of Belgium, only one long-term daily temperature series (1767-1998) representative of central Belgium has been established (Demarée et al., 2002) and 12 long-term daily precipitation time series (e.g. Dupriez and Demarée, 1988 – see table 1-). Thanks to two digitization projects funded by the Belgian Science Policy Office (BELSPO), RMI has undertaken the digitization of the climate observations conducted in the period 1880-1953 by the Belgian climatological network. A total of 623 series of daily precipitation amount and 239 series of daily extreme temperatures data have been digitized (which roughly represents 5 million of daily data encoded). For each of them, the station's name, code, latitude, longitude, altitude, and observer were verified thoroughly. Note that when detailed information about the station's location was missing, the coordinate and altitude of the village's church has been set as default location.

Only some of the stations in the network are suitable for use in climate analysis. Most have not enough data, while others have excessive missing data, poor site characteristics or unreliable observations. Ideally, station used for climate change analysis should meet the following criteria (Trewin, 2013):

1. A long period (preferably 100 years or more) of continuous data with few or no missing observations.
2. No site relocations, changes in observation practices or instruments, or significant changes in local site environment.
3. Located well outside any urban or potential urban growth area.

Because only a few (if any) of such stations exist in practice, it is necessary to construct long time series by merging climate records from neighbouring stations.

SERIES NAME	LATITUDE (°N)	LONGITUDE (°E)	ALTITUDE (m)	TIME PERIOD
Sint-Andries-Brugge	51.159	3.161	11	1880-1983
Ath	50.626	3.778	32	1883-1999
Leopoldsburg	51.107	5.263	48	1880-1999
Uccle	50.798	4.359	100	1880-1999
Gembloux	50.583	4.687	180	1880-1999
Denée-Maredsous	50.287	4.768	222	1882-1998
Rochefort	50.176	5.224	193	1880-1999
Thimister	50.654	5.863	280	1882-1999
Stavelot	50.392	5.923	297	1880-1999
Hives	50.152	5.583	398	1882-1999
Chimay (Forges)	49.982	4.340	318	1880-1999
Chiny	49.739	5.346	374	1882-1998

TABLE 1: Already available centennial daily precipitation time series in the RMI's central DB. Series in bold include records from stations operated by the former Belgian Roads and Bridges Administration in their older parts.

Currently, air temperature is measured in every thermometric station of the Belgian climatological network in a shaded enclosure (i.e. Stevenson screen) at a height of 1.5 m above the ground. Liquid-in-glass manually read thermometers (i.e. mercury thermometers for TX and alcohol thermometers for TN) are still used in about 80 % of the stations; the others being fully automated. The way that thermometers have been exposed to the atmosphere and sheltered from direct or indirect solar radiation has changed over time. The Belgian standard of the Stevenson screen (currently used for both liquid-in-glass manually read thermometers and automated sensors) was introduced in the network in the 1950s. Before this, varied types of exposures and stands were used in the network for protecting thermometers.

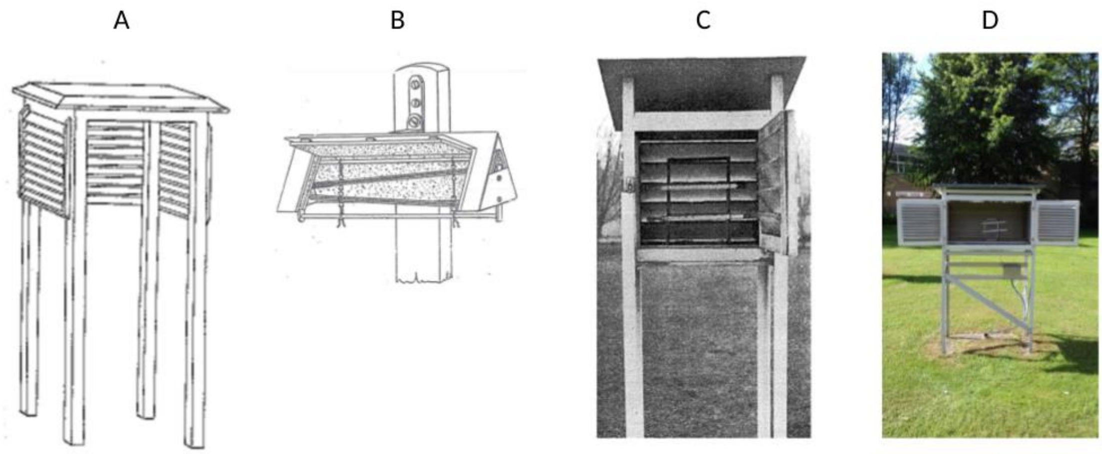


FIGURE 2: The different types of exposures and stands used in the network for protecting thermometers. Type A was used from 1880 to about 1910, type B from 1911 to 1920 and type C from 1921 until the 1950s when type D (the Belgian standard of the Stevenson screen) was introduced in the network.

From 1880 to about 1910, the thermometers were sheltered in a large screen constructed of wood with no northern wall nor floor under the thermometers and single louvered wall on the other sides (see screen of type A in Figure 2). This screen ensured strong natural ventilation, but a poor protection against soil radiation.

From 1911 to 1920 a small prism-shaped double sided screen with no floor constructed of zinc was used (see screen of type B in Figure 2). This screen had no floor and was fixed on the north side of a mast (or a tree). Adopted because inexpensive, it was however particularly defective because of radiation and intense absorption by the zinc walls.

From 1920 to 1950s, a small wooden screen (see screen of type C in Figure 2) with single louvered walls on all sides, a floor built with three boards with the central one slightly shifted to the top and a double roof (the top of the screen was composed of one asbestos board above the wooden roof with an air space between them) was in use. The volume of this screen was twice as small as the volume of type A.

Finally, in 1950s, a large wooden Stevenson screen in a double-louvered design (see screen of type D in Figure 2) was introduced in the network. Note that the asbestos boards are currently replaced by a laminate for health and safety reasons.

The screens were always and are still placed on grass with the opening or door(s) faced north so as to prevent direct sunlight on the thermometers or Pt100 sensors. It is worth pointing out that the introduction of a new type of screen into the network was progressive and could last several years before the whole network has been updated. As an example, archive documents attest that no screen of type B was anymore in use in 1931. Fortunately, two stations did not experience these successive changes in shelter type. In Denée-Maredsous, a screen of type A was used from the late 19th until the beginning of the 1950s when it was replaced by a Stevenson shelter (i.e. the screen of type D). In Uccle, an open shelter slightly different from type A (see Figure 3) is in use since the end of the 19th century till nowadays. This screen differs from type A by the smaller height of the side walls: the single louvered walls do not descend under the level of the thermometers. Since 1983, the Belgian standard of the Stevenson screen (i.e. screen of type D in Figure 2) is used as the reference shelter for climate monitoring in Uccle. Parallel measurements indicate that the open shelter is more sensitive to radiation effects than the Stevenson screen: temperatures recorded in the open shelter tend to have a strong bias in TX records compared to the closed shelter while TN readings have a small cold bias all the years.

FIGURE 3: Open screen in the climatological station in Uccle

As for the meteorological screens different types of rain gauges were used in the network. Atmospheric precipitations, rain, snow, sleet and hail were initially measured in climatological stations using three kinds of rain gauges. The oldest (see panel A in Figure 4), built by Baudin in Paris was a tenfold increasing pluviometer, it consisted of a zinc gauge, fitted on the side with a communicating tube of glass behind which a scale was engraved on the zinc. The gauge had a circular opening of about one-tenth of a square meter equipped with a solid brass ring. After each observation, the device needed to be emptied by turning it over. The second model (see panel B in Figure 4), built by Schubart in Ghent, was very similar to the previous one. It consisted of a gauge and a tank. The gauge was cylindrical and was 20 cm in diameter and 13 cm high. Its bottom was conical. At the tank, which was cylindrical was fitted a glass tube that communicated with it and behind which was the graduated scale. A tap allowed evacuating the collected water. The third model without glass tube but with a tap (see panel C in Figure 4) had an opening of 1 dm^2 so that each cm^3 of water it contained corresponded to a height of one tenth of a millimetre. (Note that it is customary to express the quantity of water fallen in millimetres; one millimetre of water collected representing one litre per m^2). It had a capacity of 2500 cm^3 and its upper part was mobile and used only in summer time to prevent water collected from evaporating and to decrease wetting loss when the rain was falling. In winter it was removed so that the instrument could collect the snow. A sleeve wrapped the zinc cylinder and protected it from the Sun. The three models were fixed on a support or on a stake in a way that their collecting surfaces lied at 1.5 m above the ground surface. This was a double fault. Indeed, it is necessary to bring the rain gauges closer to the ground as much as possible to remove them from the wind which produces eddies and diminishes the amount of water they collect. Moreover, a high rain gauge easily tips, which could artificially increases or decreases the quantities collected, depending on whether the instrument is inclined towards the rain or in the opposite side.

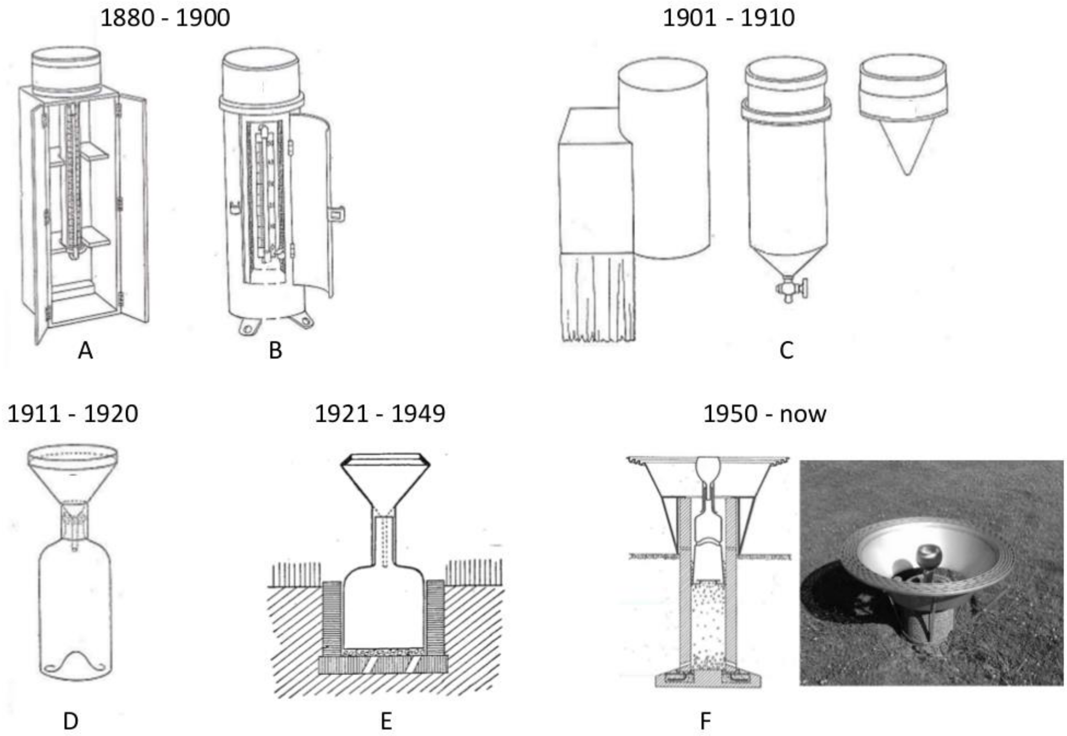


FIGURE 4: The different manual rain gauges used in the Belgian climatological network. Note that the indicated time periods are only approximate since the replacement of the instruments was progressive.

To overcome such deficiencies a new designed rain gauge was introduced in the network at the beginning of the 1910s. The instrument (see panel D in Figure 4) consisted of a bottle with a capacity of about one litre surmounted by a funnel with a diameter of 10 cm. A sleeve welded to the funnel covered the joint between the funnel and the neck of the bottle, so that the rain could not penetrate through this joint. The bottle was placed on the ground (in a cylinder that prevented the bottle from being knocked over) and the receiving surface of the funnel of 78 cm^2 was about 30 cm above the ground. The collected water, measured in cm^3 using a graduated test tube, was divided by 7.85 to convert to mm. In suspicion of snowfall, the funnel was replaced by another whose body was lengthened by a 20 cm cylinder so as to allow the snow to accumulate between two observations.

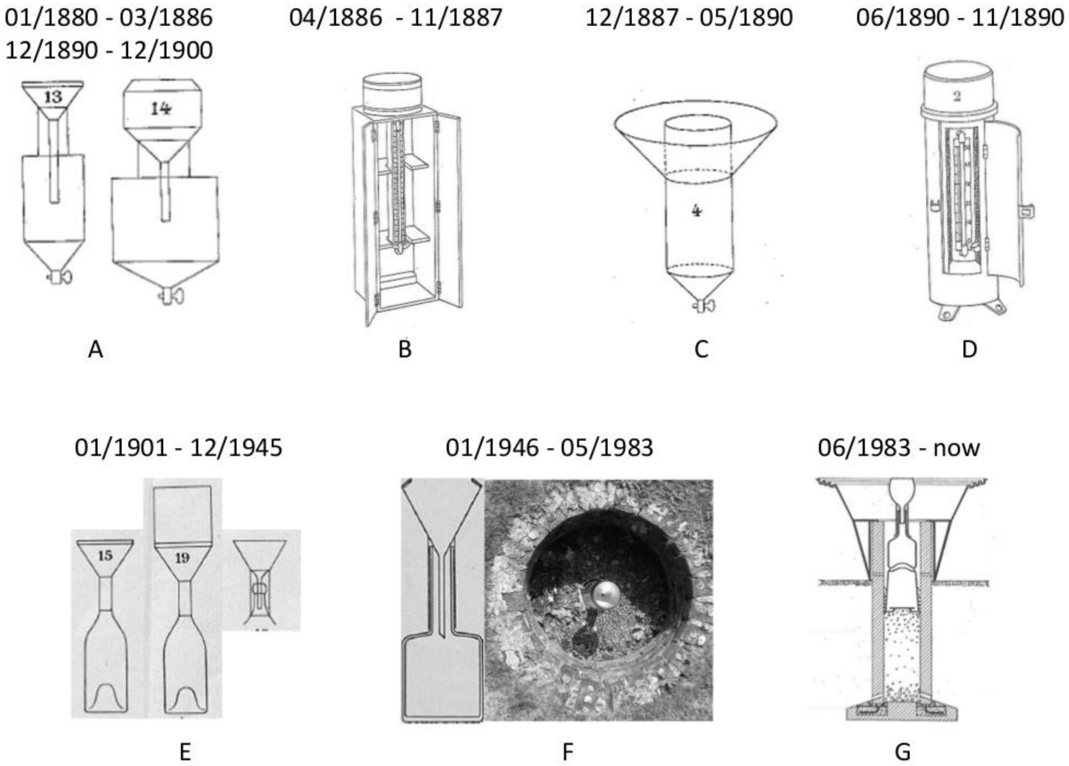
From 1921 to 1950 a funnel with a receiving surface of exactly 1 dm^2 (see panel E in Figure 4) was used so as to convert the collected cm^3 to mm simply by dividing by a factor of ten. The funnel with a rectangular opening surmounted a specially designed 2-liters bottle partly driven into the soil so that receiving surface of this rain gauge was approximately 30 cm above the ground surface. The brass edge of the funnel was sharp-edged to exactly define the receiving surface. The mouth of the funnel was provided with a welded sleeve to cover the junction between the funnel and the bottle in order to prevent water from either entering the bottle laterally or evaporating too easily. This device also ensured a good stability of the funnel on the bottle. A funnel with an elongated 20 cm body was used to collect the snow.

In 1950, as part of a major reorganisation of the Belgian climatological network, a new rain gauge (see panel F in Figure 4) was introduced into the network. This rain gauge is still in use in all manual stations. It consists of a chromed brass funnel whose circular opening has a surface of 1 dm^2, surmounting a special bottle whose capacity is 2 litres. The rain gauge is placed on a tripod support fixed in a stable way in a protective concrete tube. It is surrounded by a protective cone (i.e. Nipher cone) designed to close the wind at the immediate vicinity of the funnel and to ensure the horizontal flow of the air above the rain gauge opening. The opening of the rain gauge is 50 cm above the ground surface in Lower and Middle Belgium and can rise up to 75 cm in Upper Belgium in locations where the snow accumulation can be large in winter time. In suspicion of snowfall, the rain gauge (i.e. the funnel and the bottle) is replaced by a nivometer which consists of a cylindrical gauge 45 cm high with an opening of 1 dm^2.

It is worth pointing out that as for the meteorological screens, the up-mentioned time periods are only indicative because all instrumentation changes have only be progressively implemented in the network. We do not have precise documentation on the date of commissioning of the different models of rain gauges at each station.

Two zinc gauges have been used until the end of November 1890 at the former Observatory of Brussels in St-Josse-ten-Noode (i.e. the No. 13 and No. 14 on panel A in Figure 5). These gauges were rectangular in shape and had an opening of 1 x 2 dm. The gauche No. 13 had the shape of an inverted quadrangular pyramid, opened by the base, but whose walls then extended vertically to form a rim of 2 cm in height. The No. 14 designed to collect the snow differed from No. 13 only in its upper form so as to prevent the snow from being blown away by the wind after its fall. The collected water passed into a lower reservoir/tank through a 1 cm diameter tube. The tank ended with a tap. The rectangular opening of the rain gauges was 1.1 meters above the ground of the terrace of the Old Observatory (which was itself 1.75 m above the garden).

Numerous rain gauges have been in use in the RMI’s climatological park in Uccle, Brussels (e.g. Vincent, 1912; Sneyers, 1964, 1968), but only few of them have served as reference instruments for climate monitoring. The reference rain gauge from March 1889 to October 1947 was placed in the centre of a cylindrical tank 69 cm in diameter dug so that the opening of the rain gauge was flush with the surrounding ground. The rain gauge had a circular collecting surface of 1 dm2 delimited by a cylindrical edge in sheet metal 1 mm thick. From the funnel (with a shape of an inverted cone) the water flowed into a bottle serving as a reservoir (see panel E in Figure 5). To measure the water, the bottle was emptied into a graduated test tube. In case of snowfall suspicion a funnel with an elongated 10 cm cylinder body was used instead of the usual 1 cm height body.



<u>FIGURE 5:</u> The different rain gauges (and associated time periods of use) involved in the Brussels-Uccle centennial time series.

From November 1947 it was preferred another rain gauge whose opening also reached the ground level. This rain gauge consisting of a funnel surmounting a bottle was placed in the centre of a cemented tank with a diameter of 70 cm and a depth of 40 cm. The rain gauge had a circular opening of 1 dm^2 delimited by the bevel stop of a copper ring. The collecting bottle had a 2 litres capacity (see panel F in Figure 5). In suspicion of snowfall, the rain gauche (i.e. the funnel and the bottle) was replaced by a nivometer which consisted of a cylindrical container 40 cm high with an opening of 1 dm^2. It is worth nothing that this rain gauge has been used as the reference instrument until the end of April 1983 while the current reference rain gauge (panel G in Figure 5) was introduced in the network in 1950.

It is worth pointing out that the Brussels-Uccle daily precipitation data retrieved on a digital format from the RMI’s central DB does not exactly correspond to the measurements performed with the reference rain gauges described above. Figure 5 presents the different instruments actually involved in the daily Brussels-Uccle pluviometric time series elaborated in a former digitisation work (e.g. Dupriez and Demarée, 1988) and considered here. From January 1880 to the end of March 1886 measurements from the two zinc gauges (see panel A in Figure 5) located at the former Observatory of Brussels in Saint-Josse-ten-Noode were used. From April 1886 the daily rainfall amounts were taken from rain gauges installed in the Uccle plateau. From April 1886 to the end of November 1890, atmospheric precipitations were measured using rain gauges (see panels B and D in Figure 5) similar to those in use in the regional stations (see panels A and B in Figure 4). Note that from December 1887 to the end of May 1890 measurements were taken from a rain gauge (see panel C in Figure 5) which slightly differs from the one in use in the regional stations during the time period 1901-1910 (see panel C in Figure 4). This rain gauge specially designed to measure the snow consists of a 40 cm long cylinder with a conical bottom and ending with a tap. It is equipped with a truncated cone, inverted, surrounding the container and designed to prevent the air from rising vertically along the gauge during strong wind. This truncated cone makes an angle of 45° with the horizon. The upper base is on the same level as the rain gauge opening. This device was designed by Nipher. The rain gauge was hooked to a support and its circular opening was 20 cm in diameter. From December 1890 to the end of the year 1900 atmospheric precipitations were recorded in zinc gauges similar to those used until the end of November 1890 at the former Observatory of Brussels (see panel A in Figure 5). From January 1901 to the end of the year 1945 measurements were taken from the reference rain gauge at that time (see panel E in Figure 5). From January 1946 to the end of May 1983, the reference rain gauge for the time period November 1947 to April 1983 was considered (see panel F in Figure 5). Finally, the current official manual rain gauche (see panel G in Figure 5) is used since June 1983.

As extensively discussed in literature (e.g. Aguilar et al., 2003; Auer et al., 2007; Brunet et al., 2008, 2011) instrumental time series could be altered by changes in the measurement conditions, such as evolution of the instrumentation, relocation of the measurement site, modification of the surroundings, instrumental inaccuracies, poor installation, and observational and calculation rules. Moreover spurious observations are frequent. As these artificial shifts often have the same magnitude as the climate signal, such as long-term variations, trends or cycles (e.g. Caussinus and Mestre, 2004; Della-Marta et al., 2004) a direct analysis of the raw data series might lead to wrong conclusions about the climate evolution. As an example, Hanssen-Bauer and Forland (1994) reported that, in the case of precipitation, the progressive improvements of instrumentation can introduce artificial systematic increases. Therefore it is widely accepted that inhomogeneities and aberrant observations in time series have to be detected and if necessary adjusted before performing any kind of climate change analysis.

2. STATE OF THE ART AND OBJECTIVES

2.1 Data quality control

The value of any meteorological measurement is dependent on the accuracy and precision with which it represents the physical quantity being measured. No measuring technique is perfect and errors can occur in meteorological observations for a wide variety of reasons, the most common being instrument faults, observer errors, errors in data transmission and clerical error in data processing. A distinction is drawn here between short-term issues which affect observations over a finite period (most commonly a single observation, but sometimes persisting for a period of a number of days or weeks), and longer-term influences on a climate record (inhomogeneities) which are considered separately (see Section 2.2 here below). Without outliers being properly treated, homogenization and analysis may render misleading results.

In the past, little attention was paid to data quality control, believed to be less important than the improvement of numerical weather prediction and data assimilation techniques, and considered as a less “glamorous” topic. Quite early though, it was recognized that the insufficiency of the quality control applied to the observations was an obstacle to the quality of the analysis, also crucial for the skill of numerical forecasts. Since the 1980s more effort has gone in the study and formalization of quality control procedures. Literature of the last two decades suggests an evolution toward complex quality assurance (QA) and quality control (QC) in practice with meteorological data processing. Complex QA is distinguished from more traditional “simple” QA by the use of several different tests or rules against which data are tested and a decision tree to weigh all of the evidence before flagging data (Gandin, 1988; Eskridge et al., 1995). Each procedure will either detect the data as being valid or erroneous. The guiding principle is that no decision to flag a datum should be made until all available approaches have been applied toward the assessment of its validity. False positives (i.e. type I error) increase the burden on the manual QC and false negatives (i.e. type II error) reduce the quality of the data.

In view of the huge amount of values that have to be processed for each selected meteorological parameters (i.e. TX, TN and, RR), automating the QA/QC process is of extreme importance. Fully automated QA procedures required to be developed (with a special care pay to the records digitized by the staff as they require a stricter QC to avoid major mistakes in data digitization) to isolate and flag potentially errant values as well as for ensuring internal consistency and temporal and spatial coherence of the data. Through this objective, we have also investigated how recent data mining techniques can be incorporated in the quality assurance decision making process.

Data mining (e.g Han and Kamber, 2006) is the extraction of hidden predictive information from large databases. It is a powerful technology with great potential to analyse important information in data bases and data warehouses. A variety of data mining tools and techniques are available in the industry, but they have been used in a very limited way for meteorological data (e.g. Abdelaal and Elhadidy, 1995; Lucio et al., 2007; Sciuto et al., 2009). In this project, we use Artificial Neural Network (ANN) to explore the spatiotemporal dependence of meteorological attributes. An ANN is an interconnected group of artificial neurones that uses a mathematical model for information processing based on a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network. In more practical terms neural networks are non-linear statistical data modelling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. The motivation for the development of neural network technology stemmed from the desire to implement an artificial system that could perform intelligent tasks similar to those performed by the human brain.

Supervised deep learning approaches (Lecun et al. 2015) model the underlying distribution of the data by learning from demonstrations, provided that enough examples are available to learn from. Deep learning architectures convert all (raw) data samples under analysis into a stack of intermediate representations that are compact and redundancy-free, in a way that further tasks such as discrimination or detection (among others) can be done in a simpler, more efficient and accurate way, without requiring any specific features engineering. Recurrent neural networks (RNN) models (Lipton et al. 2015), and in particular long-short term memory (LSTM) networks based on auto-encoder representations (Malhotra et al. 2016), have been successfully applied to anomaly detection problems in time series, though mostly for discord rather than contextual anomalies. Essentially, by looking at many examples of raw signals (i.e. including anomalies) and their errors-free (manually corrected) counterparts, it is possible to learn given certain context, the underlying distribution of the residual signal as an indicator of anomaly presence in form of peaks in the residual. Another possibility would be to directly learn to detect anomalous data points from the residual signal and the same context. A third scenario considers combining both detection and prediction in order to better find irregularities on information that follows certain patterns (Feng et al. 2017). Our approach goes in this direction and aims at simultaneously detecting anomalies and predicting the correct values for such data points in weather time series; the benefits of this approach is two folds. Firstly, the detection process can be improved with the prediction model by increasing the detection likelihood whenever the *predicted value is too different with respect to the raw value*, and hence the magnitude of the associated residual will increase (or diminish) the detector signal; this reduces false negatives (errors type II). Secondly, corrections to the raw time series will only be performed if a positive detection happened in the data point in question, which in turns reduces false positives (errors type I). As a result, such a combined detection + prediction model allows to jointly minimizing errors of types I and II.

2.2 Data homogenization

The aim of data homogenization is to adjust observations, if necessary, so that the temporal variations in the adjusted data are caused only by climate processes (Conrad and Pollak, 1950). It is a two-stage process: firstly the detection of inhomogeneities in the data time series and secondly the adjustment of the data to remove these inhomogeneities. Ideally, the date of a change of instruments, location, or observing practices would be recorded as metadata, and parallel measurements made with the original and the new setup for several years, allowing reliable estimation of the inhomogeneity. In practice, metadata are often incomplete and parallel measurements lacking, so statistical homogenisation is required.

Because statistical absolute homogenization (without using neighbouring stations) has the potential to make the data even more inhomogeneous (e.g., Venema et al., 2012) and cannot distinguish small jumps and gradually occurring inhomogeneities from natural variability and climate change, the most commonly used principle over the last decades to detect and remove the artificial changes is relative homogenization. This assumes that nearby stations are exposed to the same climate signal, but that non-climatic changes are station-specific. Differences between nearby stations enable to detect inhomogeneities. By looking at the difference time series, the year to year variability of the climate is removed, as well as regional climatic trends. In such a difference time series, a clear and persistent jump can be therefore easily detected and attributed to changes in the measurement conditions.

In the past, it was a common practice to establish a composite reference time series computed from multiple nearby stations and to assume that the difference between the data at the tested station and the reference series (assumed not perturbed) is fairly constant in time, up to the perturbations to be detected (Alexandersson, 1986). The procedures that were in widespread use among climatologists (e.g., Potter, 1981; Alexandersson, 1986) relied on likelihood ratio tests (Hawkins, 1977; Maronna and Yohai, 1978). In these procedures the null hypothesis was tested against the presence of a single change-point (also referred to as break-point).

These approaches however suffer from two major limitations:

1. The reliability of the so-called reference series cannot be proved. The methods for creating such series (Alexandersson, 1986; Hanssen-Baueur and Forland, 1994; Easterling and Peterson, 1995; Alexandersson and Moberg, 1997) do not guarantee their quality (Caussinus and Mestre, 2004).
2. These methods can usually handle only a single change-point while long time series typically have more than just one jump. Moreover, an unknown number of outliers may spoil the data (Caussinus and Mestre, 2004).

Therefore, more recent homogeneization approaches focus on methods specifically designed to detect and correct multiple change-points and work with inhomogeneous references (e.g. Szentimrey, 1999) or via direct pairwise comparison (e.g. Caussinus and Mestre, 2004).

2.2.1 Monthly homogenization: The HOMER software

The method opted for homogenization of the monthly temperature and precipitation time series in Belgium relies on the use of the HOMER software (Mestre et al., 2013) developed with support of the European Union, through the COST Action ES0601 – Advances in Homogenization Methods of Climate Series: an Integrated Approach (HOME) – for the detection of a multiple number of break-points and the calculation of adjustments. This interactive semi-automatic software includes the best features of some state-of-the-art methods, namely PRODIGE (Caussinus and Mestre, 2004), ACMANT (Domonkos, 2011), and cghseg a joint segmentation method that was developed originally by biostatisticians in the context of DNA segmentation (Picard et al., 2011). These methods identify change-points in a time series by computing an optimal segmentation of this time series in homogeneous subseries. A dynamic programming approach enables to determine the change-points by minimizing the internal variance of the subseries (Caussinus and Lyazrhi, 1997). Combining different break detection algorithms has been demonstrated to be beneficial (Gubler et al., 2017). It results in higher confidence when accepting or rejecting break-points, especially if a break-point cannot be confirmed by metadata (Toreti et al., 2012; Kuglitsch et al., 2012). The correction of the break-points in HOMER is done based on a two-factor analysis of variance (ANOVA) model approach. It allows for the correction of a set of stations simultaneously and automatically (Mestre et al., 2013), and was shown to improve break-point correction over traditional approaches (Domonkos et al., 2011; Domonkos, 2013).

Homogenization with HOMER is an iterative process. Break-point detection procedures are alternated with the correction of the break-points. In case the homogenization does not remove the breaks in a satisfying way, break dates are modified and breaks in the original time series are readjusted. The alternating procedure is stopped once every time series is considered homogeneous by the homogenization operator. Finally, it is worth pointing out that HOMER adjusts the data preceding a detected change-point to make it homogeneous with the data after that change-point. In this way, the most recent data (i.e. following the last change-point) are not modified. This offers the clear advantage for ongoing monitoring that new data can be simply appended to the time series.

2.2.2 Daily homogenization

If methods used to homogenize annual and monthly data are well established (e.g. Peterson et al., 1998; Venema et al., 2012) relatively few methods exist to homogenize daily data. This is not due to limitations in the detection of shifts since this information may be provided by analysis of annual or monthly series (e.g. Kuglitsch et al., 2009; Nemec et al., 2013; Xu et al., 2013; Wang et al., 2014) but is mainly due to an adjustment problem. Indeed, the main challenge of the homogenization of daily compared to monthly data is that, at least for temperature, the magnitude of inhomogeneities may differ with varying weather situations. For temperature correction, multiple regressions including other parameters such as wind-speed and direction, sunshine duration and parallel measurement is probably the best way to proceed (Brandsma, 2004). But such data is extremely rare when considering older data, where usually only precipitation and temperature were observed.

The simplest daily data adjustment method relies on interpolation of monthly adjustment coefficients. However, adjusting monthly or annual mean value is not sufficient to produce homogeneous time series of higher-order statistical properties such as variance, or derived statistics which are a function of those higher-order properties, such as the occurrence of extremes. This issue was initially identified by Trewin and Trevitt (1996), who found that in some cases an inhomogeneity, such as a site move, affected different parts of the frequency distribution of daily temperature in different ways. Different techniques have been proposed to address this problem. These include methods which attempt to homogenize data across the full range of the frequency distribution, by matching percentile points in the frequency distribution (Della-Marta and Wanner, 2006) or by other means (Brandsma and Können, 2006; Toreti et al., 2010; Wang et al., 2010; Mestre et al., 2011), as well as methods which explicitly test the homogeneity of higher-order statistical properties such as mean daily variability (Wijngaard et al., 2003) or exceedances of percentile-based thresholds (Allen and DeGaetano, 2000).

Based on the HOMER detected break-points on the monthly time series, the following daily adjustment methods were applied on the daily extreme temperatures time series: the Vincent's method (Vincent et al., 2002), the Higher-Order-Moments (HOM) method (Della-Marta and Wanner, 2006), the SPLine Daily HOMogenization (SPLIDHOM) method (Mestre et al., 2011) and the Percentile Matching (PM) method (Trewin, 2013) and compared to the Quantile Matching (QM) daily homogenization method proposed by Wang (2009). Because the HOM, SPLDHOM and PM methods are not suited for the computation of daily precipitation adjustments, only the Vincent method using the HOMER adjustments' and the QM method were considered for the homogenization of the daily precipitation time series.

2.2.2.1 The Vincent method

If there is a need for daily data adjustment, the simplest method relies on the interpolation of monthly adjustment coefficient. This method provides adjustments only for the mean of an inhomogeneity and not for its higher-order moments.

2.2.2.2 The HOM method

The Higher-Order Moments (HOM) method uses a nonlinear model capable of handling inhomogeneities in higher moments to estimate the relationship between a candidate series and a highly correlated reference series in the absence of overlapping parallel measurements. The model is built in a homogeneous sub-period before an inhomogeneity and is then used to estimate the observations at the candidate series after the inhomogeneity using observations from the reference series. The differences between the predicted and observed values are binned according to which decile the predicted values fit in the candidate series observed cumulative distribution function defined using homogeneous daily temperature before the inhomogeneity. In this way, adjustments for each decile are produced.

2.2.2.3 The SPLIDHOM method

The SPLIne Daily HOMogenization (SPLIDHOM) method is a variation of the HOM method. Although part of the principle involved is quite similar relying on the definition of homogeneous sub-periods, SPLIDHOM proposes a very different direct non-linear spline regression approach rather than an adjustment based on quantiles. It is a method for sequential adjustment of breaks in time series, relying on the good relationship between the candidate series and the highest correlated reference series. It is based on a nonlinear regression function between the temperature measurements. In a first step, the nonlinear regression between the two series is estimated for both the period before and after the break-point. To circumvent the problem of additional inhomogeneities in the reference series, the regression function is estimated using a classical smoothing spline (i.e. a cubic smoothing spline). The smoothing parameter of the cubic spline is estimated for each regression by means of a standard cross-validation technique, in order to avoid overfitting. Both HOM and SPLIDHOM methods require highly correlated ($r > 0.8$) reference series.

2.2.2.4 The PM method

The percentile-matching (PM) method is similar conceptually to the HOM and SPLIDHOM methods (i.e. differing adjustments are applied to daily data depending on their position in the frequency distribution) although there are some differences, principally in the details of generating transfer functions. By contrast to HOM and SPLIDHOM methods that use a single reference station, the PM method allows considering a combination of multiple reference stations (i.e. up to ten neighbouring stations selected in descending order of correlation with the candidate series, with a lower correlation limit of 0.6). Where overlapping data exist between the candidate and reference series, percentile points (calculated separately by seasons) from the old and new sites are matched to define transfer function. Where no overlap exists, reference series are used in a two-step process which first matches the pre-inhomogeneity series to the neighbouring reference series, then the neighbouring reference series to the post-inhomogeneity series, with the final transfer function taken as the median of the transfer functions derived individually for each neighbouring reference series. Each month is processed individually taking into account the 6 previous and following adjacent months to ensure a smoother passage from one month to another.

2.2.2.5 The QM method

The objective of the Quantile Matching (QM) method is to adjust the candidate series so that the empirical distributions of all segments of the detrended candidate series match each other. The adjustment value depends on the empirical frequency of the datum to be adjusted (i.e. it varies from one datum to another in the same segment, depending on their corresponding empirical frequencies). It is worth pointing out that the QM method does not use reference series.

3. METHODOLOGY

3.1 Series Constitutions and Metadata Completion

3.1.1 Series Constitutions

In view of the limited number of historical daily temperature and precipitation time series available in the RMI's central DB it was necessary to construct long time series by merging records from neighbouring stations. This process was done in two steps. First, stations with more than 360 months (30 years) of continuous records were selected. Second these series were extended as much as possible with measurements from neighbouring stations provided that the following conditions were satisfied: a maximum distance of 15 km and a maximum elevation difference of 50 m between all the stations contributing to the composite series. This allowed records from short-duration sites of similar climatological conditions to be considered in the development of long time series. The location and name of the composite series are those of the longest subseries. Note that when the shift in the location does not involve the change in name and identifier for the station we consider it as a site relocation and otherwise as a station catenation. Site relocation can be as small as a few meters but have the potential to substantially impact on observation (e.g. a short move may place instruments well clear of an obstacle which previously affected the observations or the reverse). On the other hand, station catenations introduce the potential for changes due to meso-scale influences such as elevation changes, site classification changes, local topography, or proximity to the coast (e.g. Trewin, 2010).

The ideal situation in case of composite time series is that a substantial overlap between the two sites exists so that observations in common can be used to determine appropriate adjustments. In most cases however, no sufficient (if any) overlap is available. Records from different sites need then to be merged without overlap, and are treated in the same way as an inhomogeneity identified from metadata within a record from an individual site.

Series Name	Latitude (°N)	Longitude (°E)	Altitude (m)	Time Period
Oostende	51.223	2.906	5	01/1880-12/2015
Jalhay	50.584	5.972	298	01/1880-12/2015
Saint-trond	50.816	5.187	54	01/1880-12/2015
Uccle	50.798	4.359	100	01/1880-12/2015
Antwerpen	51.225	4.400	10	02/1880-12/2015
Thimister	50.654	5.863	280	01/1881-12/2015
Leopoldsburg	51.108	5.271	50	01/1881-12/2015
Gembloux	50.561	4.660	160	05/1881-12/2015
Saint-Josse (Brussels)	50.856	4.366	35	01/1889-12/2015
Stavelot	50.398	5.936	320	01/1890-12/2015
Rochefort	50.176	5.224	193	06/1892-12/2015
Denée Maredsous	50.287	4.768	222	01/1893-11/2013
Chimay (Forges)	49.982	4.340	318	08/1910-12/2015
Ezemaal	50.756	5.083	50	06/1913-12/2015
Baraque-Michel	50.519	6.061	672	01/1928-12/2015
Leuven	50.863	4.685	28	04/1930-12/2015

TABLE 2: Names, coordinates, altitudes and time periods covered by the historical daily extreme temperatures time series. The twelve centennial temperature time series are indicated in italics.

At the end of this process, continuous time series of daily temperature data have been identified for 61 locations in Belgium for the period 1954-2015 (hereafter referred to as long series) including 16 locations with data starting before 1931 (hereafter referred to as historical series, see Table 2). Among them eight locations have data since 1880 (or 1881). Note that the number of stations (i.e. subseries) in a composite temperature series has been limited to three and five for the long and historical series, respectively (see Table 3). The average percentage of missing daily values is in the order of 2.3% for the long and 4.7% for the historical series, respectively. Many historical locations suffer from a lack of data during the World War II. Figure 6 shows the temperature series' location within the Belgian territory, together with additional neighbouring foreign temperature series considered in the homogenization process (see Section 3.3.1).

Number of Stations	**Temperature Times series**		**Precipitation times series**	
	Long Series	Historical Series	Long Series	Centennial Series
1	31	2	72	5
2	25	5	58	6
3	5	5	16	7
4	/	2	3	5
5	/	2	/	/
6	/	/	/	1

TABLE 3: Number of stations (i.e. subseries) involved in the long and historical/centennial temperature and precipitation time series, respectively.

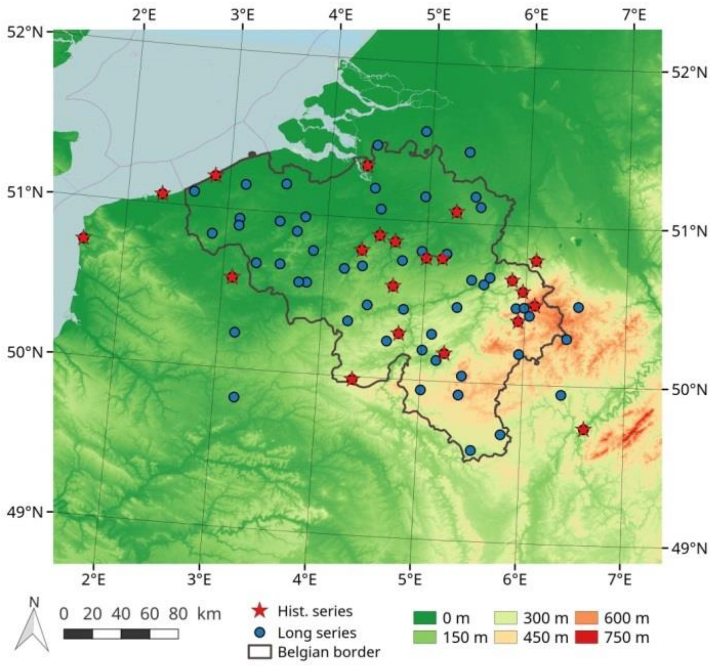


FIGURE 6: Location of the temperature series. Additional neighbouring foreign series considered in the homogenization process are also displayed.

Similarly, time series of daily precipitation data has been established for 149 locations in Belgium for the time period 1951-2015 (hereafter referred to as long series) and for 24 centennial series. It is worth pointing out that because some of the already existing centennial precipitation time series (see Table 1) were not built in compliance with our current criteria of maximum distance and elevation between all the stations contributing to the composite series, they were modified accordingly. More specifically, the starting date of the Denée-Maredsous, Chiny, Hives and Chimay (Forges) time series have been delayed as indicated in Table 4. In addition, in the case of the Chiny series such a temporal reduction combined with the impossibility to extend the series to recent years in a satisfactory way made this series too short to keep referred as a centennial one. Note that it was not possible to extend the Thimister series after the end of March 2000.

Series Name	Latitude (°N)	Longitude (°E)	Altitude (m)	Time Period
Jalhay	50.584	5.972	298	01/1880 – 12/2015
Leuven	50.863	4.685	28	01/1880 – 12/2015
Saint-trond	50.816	5.187	54	09/1899 – 12/2015
Sint-Andries-Brugge	51.159	3.161	11	01/1880 – 12/2015
Stavelot	50.392	5.923	297	01/1880 – 12/2015
Uccle	50.798	4.359	100	01/1880 – 12/2015
Antwerpen	51.225	4.400	10	03/1880 – 12/2015
Gembloux	50.583	4.687	180	03/1880 – 12/2015
Ninove	50.823	4.113	45	03/1880 – 12/2015
Oostende	51.223	2.906	5	03/1880 – 12/2015
Rocheford	50.176	5.224	193	03/1880 – 12/2015
Leopolsburg	51.107	5.263	48	04/1880 – 12/2015
Thimister	50.654	5.863	280	01/1882 – 05/2000
Veurne	51.072	2.669	5	02/1883 – 12/2015
Ath	50.626	3.778	32	04/1883 – 12/2015
Huy	50.534	5.217	71	04/1888 – 12/2015
Schaerbeek	50.847	4.400	72	02/1905 – 12/2015

Etalle	49.674	5.602	336	05/1892 – 12/2015
Denée-Maredsous	50.287	4.768	222	12/1892 – 12/2015
Hives	50.152	5.583	398	09/1909 – 12/2015
Sugny	49.816	4.900	375	09/1909 – 12/2015
Chimay (Forges)	49.982	4.340	318	08/1910 – 12/2015
Landen	50.754	5.079	69	05/1911 – 12/2015
Ezemaal	50.771	5.001	47	05/1913 – 12/2015

TABLE 4: Names, coordinates, altitudes and time periods covered by the established centennial daily precipitation time series. Series in italics are the already existing time series that have been extended to recent years while in bold are already existing time series that have been modified and extended.

Table 3 indicates that the number of stations (i.e. subseries) in the composite precipitation time series has been limited to a maximum of 6 for the centennial series (i.e. 1 series) and 4 for the long series (i.e. 3 series), respectively. Five (resp. 72) centennial (resp. long) series originate from a single station records (including possible relocations) and less than 55% (resp. 13%) of the centennial (resp. long) series contain more than 2 subseries. The location of the precipitation time series within the Belgian territory is displayed in Figure 7 together with the location of additional neighbouring foreign series used to help the homogenization of the national series (see Section 3.3.2).

Finally, it is worth pointing out that the distinction between historical/centennial and long series is due to a reorganization of the Belgian climatological network in the early 1950s. This resulted in the widespread introduction of the current manual rain gauge instrument and the Belgian standard of the Stevenson shelter, the operational start of many new stations and the incorporation of the former Belgian Roads and Bridges Administration observation network within the RMI's climatological network. To this respect, it is important to note that the Belgian Roads and Bridges Administration has used for long time a rain gauge much larger than those of RMI presenting an opening surface of 4 dm^2 and placed on a wooden frame of 2 to 2.5 m height to avoid any spilling in highly frequented locations.

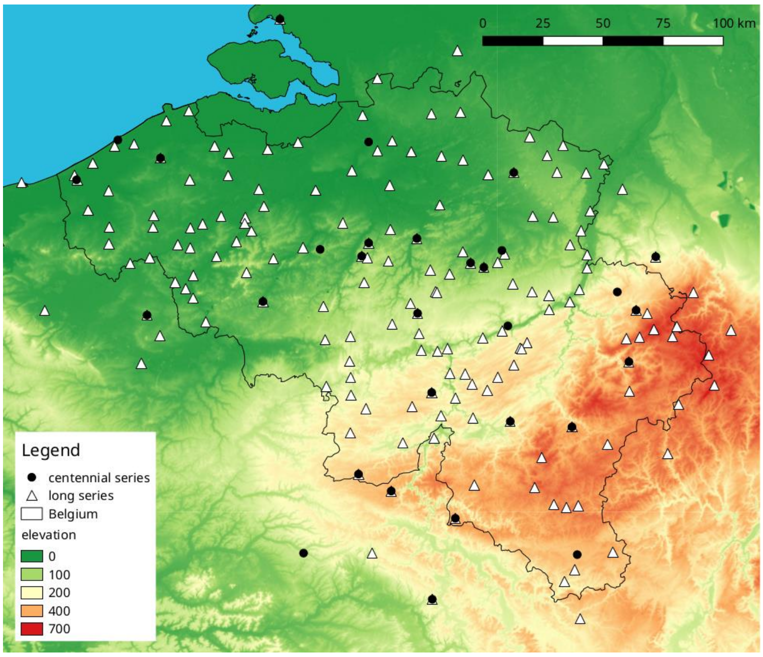


FIGURE 7: Location of the precipitation series. Additional neighbouring foreign series considered in the homogenization process are also displayed.

3.1.2 Metadata completion

Before undertaking any data quality control and homogenization it is necessary to develop complete metadata or station history information. Metadata – that is, information about the way in which the data have been observed – are important in documenting non-climatic influences which can affect measurements. These include issues which are specific to individual sites (such as site relocations, instrument changes or changes in local site conditions), or issues which affect large parts of the observing network (such as changes in observation times). These pieces of information about data are a valuable and essential guide to quality control and homogenization of the raw climate instrumental data.

An inventory of available metadata retrieved from archive documents has been established for each series. However, metadata is harder to recover than observed data: they are often not published, the older documents are sometimes illegible and only a small part of the information provided in these documents is relevant for homogenization purposes. Moreover, the existing information does not always permitted to attribute a precise date of change in a particular station but rather for the entire network in general. As an example, the introduction of a new shelter type or rain gauge instrument in the network was progressive and the date of change in a given station is usually unknown.

3.2 Data Quality Control

Roughly five millions of daily data were encoded at RMI during the digitations phase of the 1880-1949 records. Compared to the supplied readily-available digitized records (i.e. from 1950 to present) additional issues related to the general keying process can affect the quality of recently recovered and digitized data. Data QC was therefore performed in two steps. First the digitized data (i.e. 1880-1949) were checked on a semi-automated basis against typical keying errors. Second, all data (i.e. 1880-2015) were subject to various tests to automatically highlight inconsistencies.

3.2.1 Digitized data

If double-keying is recognized to minimize the effect of human error by having two individuals key the same form and reconciling the differences, historical observation forms were only digitized once at RMI. To overcome such a limitation, a first series of QC tests were applied to the digitized data to ensure that the observations reported on the original documents were accurately recorded in the RMI's DB. A two steps QC approach was implemented:

1. *Visual QC:* Errors in the keyed data are numerous: typo errors (forgetting a comma, doubling a number, adding or forgetting a number, omission of the negative sign, etc.) keying one element as another element, keying 1 day as another day, keying the date of the form and shifting the day of the data up or down to an incorrect day, attributing the form to another station and/or another month, no-data keyed as a zero value, etc.). An example of form leading to keying error is illustrated in Figure 8. In this form, the precipitation quantities measured in millimetres were written in the cubic centimetres column with the decimal in the millimetres column. Unfortunately, in such a case it was frequent that only the latter values were keyed by the digitization staff (e.g. 7 mm instead of 4.7 mm). Various tests (including data visualisation) enabled to highlight periods with suspicious values. In case of non-obvious corrections, the suspicious data were checked against the original hard-written weather reports. Despite the computer assistance, it is worth pointing out that because visual QC is an extensively time-consuming activity; only records from selected stations for the historical/centennial time series constitution were verified.

2. *Automatic QC*: Beside the visual QC automatic procedures were developed to flag and correct (when possible) known systematic errors:

a. 30°C temperature offset. It was common in the past that thermometers in use within the Belgian climatological network added a 30°C offset to the actual temperature record to prevent the omission of the negative sign for temperature values lower than 0°C. Unfortunately, such an offset was not always removed by the digitization staff when keying values reported on the original forms.
b. Unit conversion error. Prior to 1930, it was common to report the daily precipitation amount in cubic centimetres rather than in millimetres. Because various types of rain gauges were in use at that time as discussed in the Introduction, conversion factors of 0.1274, 0.1 or 1.0 needed to be applied by the staff to the precipitation amounts reported on the forms to convert them in millimetres. Unfortunately, it was not always done.
c. One-day time lag in the RR and TX series. Daily precipitation amount (RR) and extreme temperatures (TN and TX) are recorded from 8:00 am of the previous day to 8:00 am of the current day. While most of the observers did shift the RR and TX values from one day directly on the form, some did not.

At the end, it was found that about 16% of the original data were not directly usable. They were corrected or rekeyed accordingly. This error rate is rather high when compared to the recently prepared digital dataset of meteorological observations for Europe and the southern Mediterranean region for the period 1877-2012 (Ashcroft et al., 2018) in which less than 5% of the checked values required correction (in more than 60% of station tested). This highlights the importance of well preparing the digitization work before starting and implements a QC of the keyed data during the digitization phase to minimize the introduction of errors.

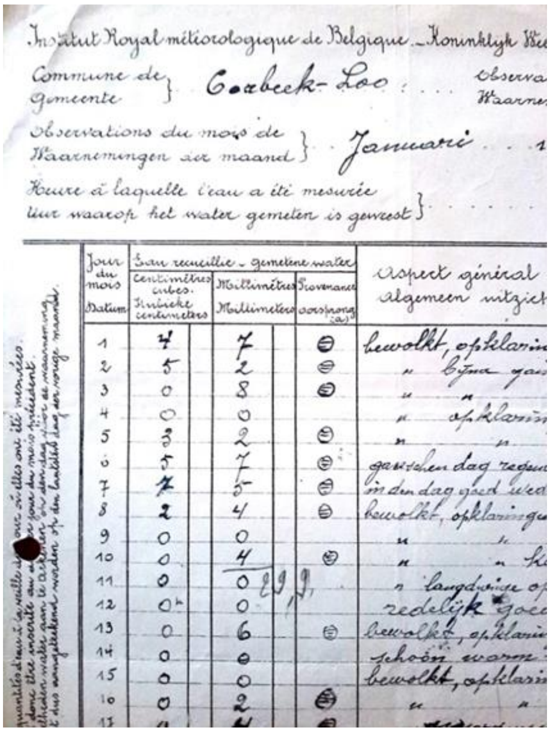


FIGURE 8: Example of form leading to keying error (Corbeek-Loo station)

3.2.2 Daily automated QC procedures

In addition to the data QC especially devoted to the digitized data, fully automated data QA procedures were developed and applied to both the supplied readily-available digitized records (i.e. from 1950 to present) and the data recovered and digitized by the staff (i.e. time period 1880-1949) to isolate and flag potentially errant values as well as for ensuring internal consistency and temporal and spatial coherence for both temperature and precipitation data.

3.2.2.1 "Classical" QC approach

The daily QC process involves a sequence of specific consistency tests. 62 years (i.e. 1954-2015) of in depth manually controlled daily extreme temperature and precipitation values from a subset of well-distributed representative stations within the Belgian territory were used as reference when developing/calibrating the automated tests. As illustrated in Figure 9 for the daily extreme temperature data QC (left panel), record that has successfully passed the first test are checked against the second test and so on.

First, the *physical limits* consistency test verifies whether the values are within acceptable limits depending on the climatic conditions of the measurement site and the season. The *plausible value test* ensures then that the daily records lay within lower and upper bounds determined for each calendar day from the highest and lowest daily values measured in the reference stations. The *internal consistency test* imposes that the maximum temperature of a given day D (corresponding to the observation period D to D + 1 at 08:00 local clock time) cannot be lower than (a) the minimum temperature in the same observation period and (b) the minimum temperature in the preceding and following observation periods, since temperature varies continuously (i.e. not by step). The *temporal* consistency test analyses the daily rate of change in both TN and TX time series in order to detect possible anomalies (e.g. spikes or unusual persistence). Finally, *the spatial consistency test* compares the observations at a given location with temperature values recorded at neighbouring stations. More details about these tests are given in Delvaux et al. (2015) from which the current QC procedures have been derived. At the end of the checks a quality index is attributed to each particular temperature data given the score obtained at the various test:

- "v": validated data
- "c": corrected data
- "sX": suspicious data where "X" is a number ranging from 1 to 5 which indicates the test the data has failed.

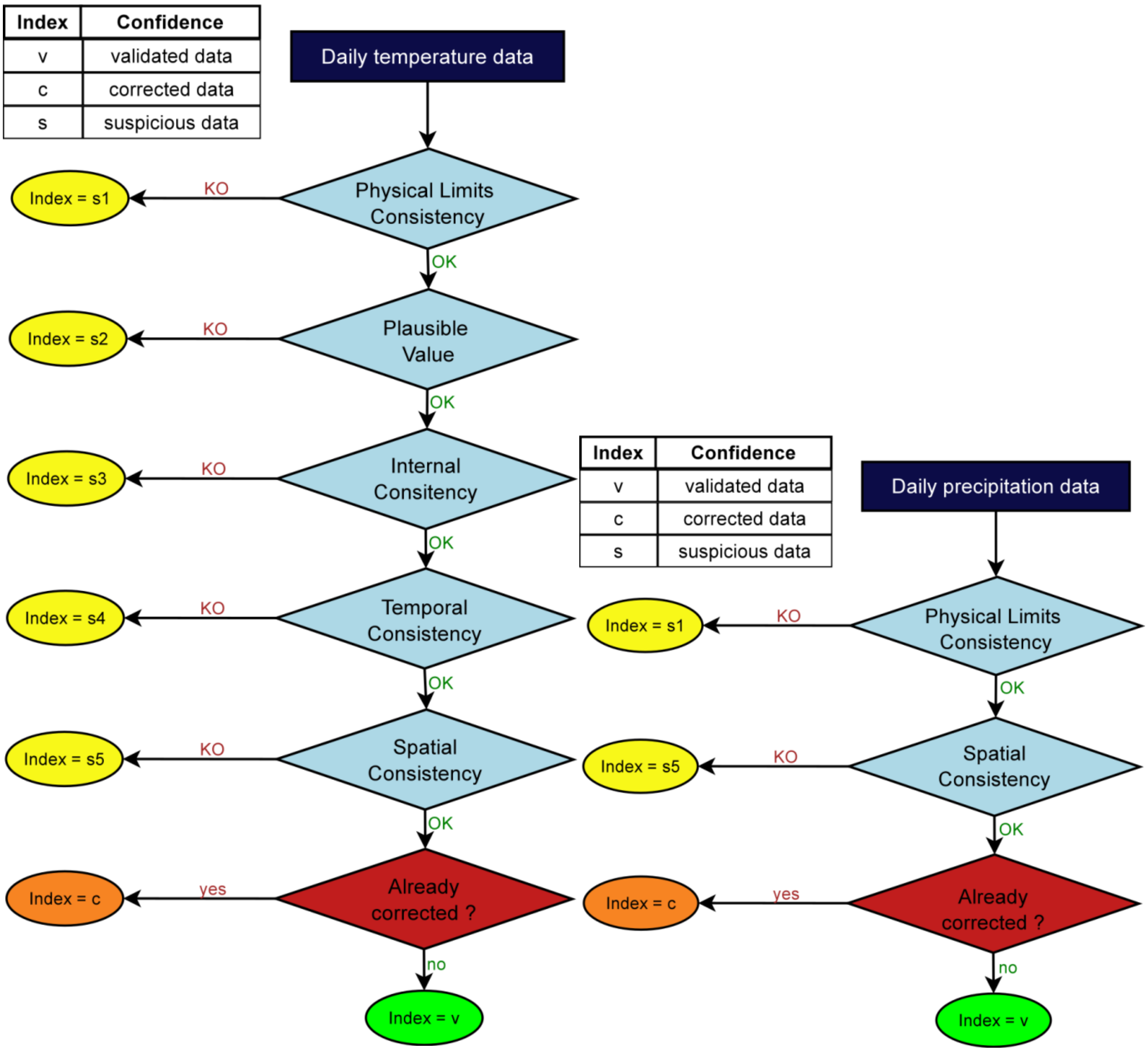


FIGURE 9: Automated QC procedure applied to extremes temperature (left panel) and precipitation (right panel) records together with the associated confidence index.

As for daily temperature records, daily precipitation amounts were faced to a sequence of specific consistency tests to ensure physical limits consistency (i.e. 0 mm ≤ RR < 250 mm) and spatial coherence of the data (see right panel in Figure 9).

3.2.2.2 Deep learning for automated detection and correction of anomalies in time series

The main issue of classical QC approaches in time series anomaly detection is that anomalies presence in the time series is very irregular, and in terms of frequency and magnitude, they are quite similar to unusually high (or low) temperature values that are perfectly natural. As a result, such anomalies can be very difficult distinguish from those extreme values using consistency tests, and require more sophisticated analysis of the local context to properly detect them (Chandola et al., 2009). Contextual anomalies are a kind of single point anomalies, whose values are localized within the range of typical values present in the time series, and thus cannot be considered outliers in the statistical sense; nonetheless, such single points are salient in some local context with respect to the others (Cheboli, 2010) in the time series.

Recall that most of anomalies present in historical temperature records arise from errors related to human and instruments failure, which are largely occasional and unpredictable, and whose nature resembles that of a random error in terms of measurements variability (Brunet et al. 2008). Stochastic Neural Networks (Malhotra et al., 2016) provide a framework to model the appearance and frequency of such anomalies without making any a priori assumption with regards to the errors distribution. The overall idea behind this type models consist in learning how an errors-free time series would looks like given many examples of anomalous and correct temperature values. This allows to ***predict the correct temperature value*** of a specific date in some meteorological station, ***given certain context*** comprising temperature values from previous days of the same station, as well as temperature values of neighbouring meteorological stations. The difference between the predicted and the registered temperature values, known as residual, would then serve as indicator of anomaly likelihood. At the same time, the predicted temperature value can serve to directly correct the anomaly, which paves that way for a fully automated QC solution, or as suggestion for a human operator to revise and correct the detected anomaly in a semi-automated QC solution. Figure 10 illustrates the proposed solution for fully automated and semi-automated QC of daily temperature/precipitation time series.

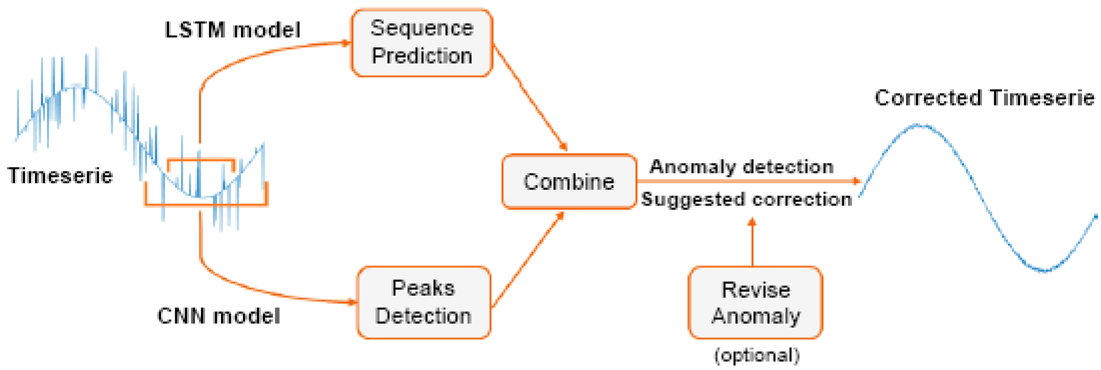


FIGURE 10: Proposed joint peaks detection and sequence to sequence prediction allows to automatically detecting and corrected anomalies in time series.

Due to the low frequency and irregular emergence of anomalies in temperature time series, prediction of temperature values at every date/timestamp could probably lead to too many false positive detections (type I errors) and consequently to unnecessary corrections of the original time series values, which should be avoided. In order to address this issue, corrections are only applied whenever:

- the residual is big enough to suggest the presence of an anomaly
- and a peaks/anomaly detector indicates a positive detection

The peaks detector model's architecture is selected in a way that both type errors I and II are jointly minimized; the NN architecture of both temperature predictor and anomaly detector is provided in Figure 11.

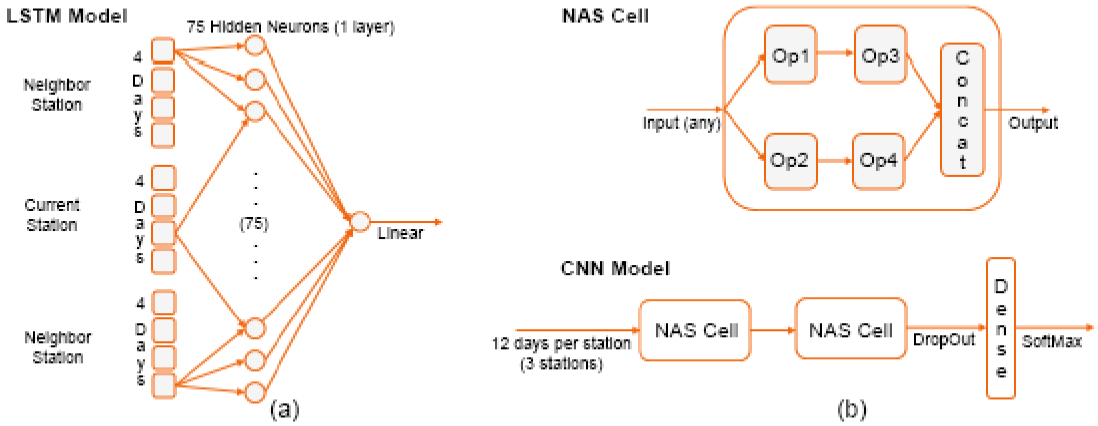


FIGURE 11: Architectures of both models (a) LSTM for temperatures prediction and (b) CNN for anomaly detection; they are trained independently and combined afterwards.

The overall solution was trained in 24 stations for TX and TN temperature values comprising temperature records from 2005 until 2015 (training period). After training, the solution was tested in unseen time series of the same 24 stations involving the period from 1996 until 1989 (testing period). The Figure 12 illustrates the temperature prediction process obtained with the trained model and tested in the period from 1996 until 1997 in one meteorological station.

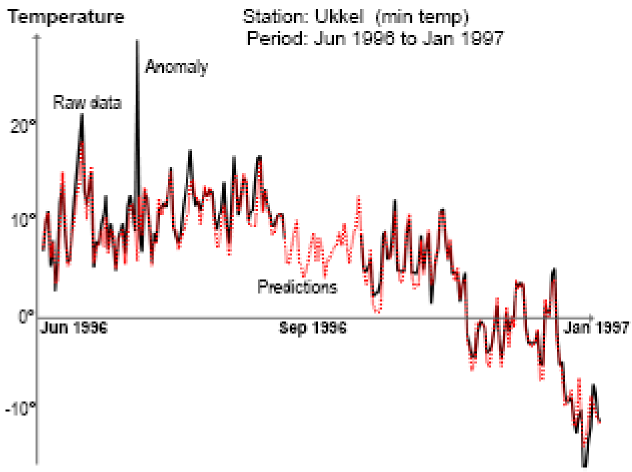


FIGURE 12: Predicted values vs raw time series of temperature records in the Uccle station, tested in the period from Jun 1996 until Jan 1997.

As can be seen, the temperature prediction is quite close to the raw time series and, more importantly, the model is able to predict missing values; the Figure 13 provides a comparison on the predicted values and the errors-free (manually corrected) time series in the same meteorological station. At the same time, the detector was used to predict the likelihood of anomaly presence in each date, and a correction to the time series, using the predicted value, is performed only if both models agree suggest the presence of an anomaly.

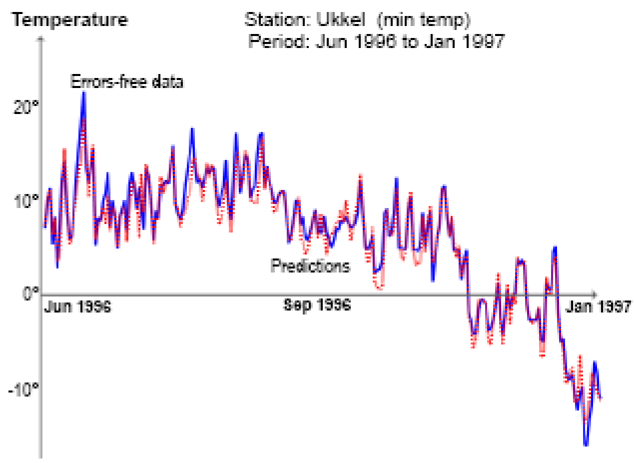


FIGURE 13: Predicted values vs errors-free (manually corrected) time series of temperature records in the Uccle station, tested in the period from Jun 1996 until Jan 1997.

The overall solution was then applied to correct the time series associated to the 24 stations under assessment in the testing period. We can see in Figure 14 that the statistics concerning the errors-free time series (manually corrected) and the time series automatically corrected with the proposed QC (NN) solution are almost the same.

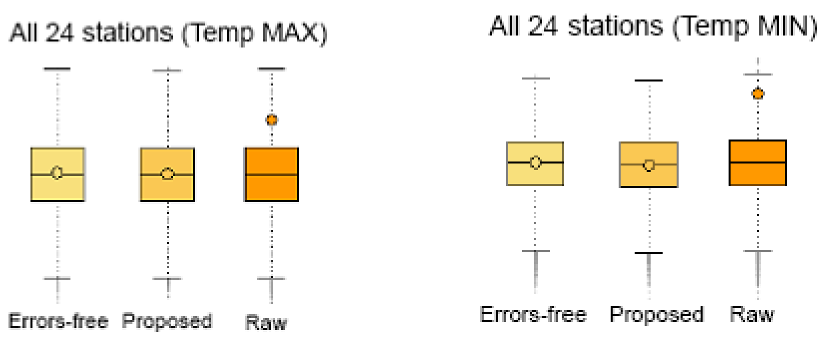


FIGURE 14: Main statistics for TX (MAX) and TN (MIN) temperature values involving the raw time series, the errors-free (manually corrected) time series, and the automatically corrected time series using the proposed QC (NN) solution.

Finally, a set of climate indicators related to temperature values were computed in all the 24 stations to quantify the differences between the raw time series, the errors-free (manually corrected) time series, and the time series automatically corrected with the proposed QC (NN) solution; the results are detailed in Table 5.

Climate Indicator	Raw Data	Errors-free data	Proposed QC (NN) corrected data
Hot days	9	5	5
Summer days	31	27	27
Tropical nights	79	76	76
Frost days	47	48	48
Icy days	7	7	7

TABLE 5: Results in climate indicators frequently used or climate analysis; given values are per-station and per-year covering the testing period from 1989 until 1999 for all the 24 meteorological stations under study.

As can be seen, not only the statistics of errors-free and automatically corrected time series are the same, but also the information extracted from the time series to analyse extreme values are essentially the same. Such a result confirms that both semi-automated and fully automated QC of daily weather time series can be performed via NN with a high accuracy level.

3.3 Data Homogenization

3.3.1 Extreme Temperature

3.3.1.1 Monthly homogenization

For each temperature time series, monthly values were computed when at least 90% of the daily values within a given month were found as valid. The monthly value was otherwise considered as missing. The monthly homogenization was carried out separately for historical series and long series. Twelve additional foreign non-homogenized series (five in France, five in Germany and two in the Netherlands – see Figure 6 for the stations location -) were considered to help the homogenization of the 61 long Belgian series. The same daily QC procedures used to check the Belgian data have been applied to the foreign data. To ensure that nearby locations are exposed to the same climate signal in the homogenisation process, the 73 time series of TX and TN were divided into five clusters of approximately 15 series (see Annexe A.1.1 for the clusters' composition) according to the following criteria:

3. The series of a same cluster should define a coherent climatological area.
4. The series of a same cluster should be correlated and present similarities.
5. A cluster should contain around 15 series

Similarly, five additional foreign series of TX and TN (i.e. two German locations and three French locations – see Figure 6 for the stations location -) were used to help the homogenization of the 16 historical Belgian series of TX and TN. Due to the limited number of locations, a unique cluster was considered for the homogenization of the historical series (see Annexe A.1.2 for the cluster composition).

Because HOMER is an interactive semi-automated software, some decisions may differ from one user to another. Therefore, the full homogenization process has been conducted independently by three different experts to minimize the impact of the cluster definition and the consequences of false break-point attributions and associated adjustments. To ensure the reliability of the homogenized long (resp. historical) temperature series, it was checked that all series satisfy the following conditions:

1. Maximum five (resp. 10) break-points.
2. Maximum one (resp. 2) break-point(s) with a magnitude larger than 0.8°C.
3. No break-point with a magnitude larger than 1°C (excepted if well supported by metadata, for example, Brunet et al., 2006; Mamara et al., 2014, 2016; Yosef et al., 2018).
4. A minimum detectable magnitude (MDM; break-point significance detection threshold provided by HOMER) smaller than 0.25°C.
5. All break-points have a magnitude larger than MDM.

At the end of the homogenisation process, the three experts had to agree on where to place any break-points on the temperature time series. Note that sensitivity tests on the cluster composition in the break-points attribution were also conducted during the homogenization process of long series. Finally, since each historical series is composed of a long series, break-points in the historical series after 1954 are based on the break-points identified when homogenizing the long series.

Table 6 reports the number of break-points detected in the analysed temperature time series. A maximum of 5 and 7 break-points are observed for long and historical series, respectively. Eight long series do not have any break-point while all historical series have at least one break-point. Twenty-nine per cent of break-points are directly supported by metadata. Further 18% break-points are found during the transition period of instrument or shelter type change and could be potentially attributed to this network upgrade. Beside instrumental changes which are not well temporally identified in the stations metadata, station catenations are responsible for the largest number of documented break-points (68%), followed by site relocations (18%) and to a lower extent by observer changes (14%).

Number of breaks	**Historical series (1880 – 2015)**		**Long series (1954 – 2015)**	
	TX	**TN**	**TX**	**TN**
0	0	0	8	8
1	1	1	18	17
2	5	1	18	22
3	2	5	13	7
4	3	3	3	6
5	4	4	1	1
6	1	1	/	/
7	/	1	/	/

TABLE 6: Number of detected break-points per temperature series

Figure 15 presents the frequency distribution of break-point magnitudes for both TX and TN time series. The magnitude range is broader for historical series with 14 break-points with a magnitude higher than 1 °C while the magnitude of the break-points of long series are all between -1 °C and 1 °C. The average distribution is very close to zero for TN and negative for TX, especially for historical series. The station moves from St-Josse-ten-Noode (Brussels downtown) to Zaventem (Brussels airport) in 1954 (distance of 12 km and elevation difference of 23 m between the two stations) presents the largest tolerated break-point (i.e. -2.12 °C for TX). Station catenation often coincides with a change of the observer, of instruments, of altitude and others. In these cases, it is not always possible to determine which of the potential causes has created the break. Multiple causes are common and a break is generally produced by more than one factor. In the present case, it is expected that the dominant cause is the cooling due to the shift from an urban location to an airport site (i.e. the previous site was 2.12 °C warmer than the new site). Note that similar ruptures have been identified in Austrian temperature series in 1940s and attributed to a number of city airport relocations caused by the needs of World War II (Auer et al., 2001).

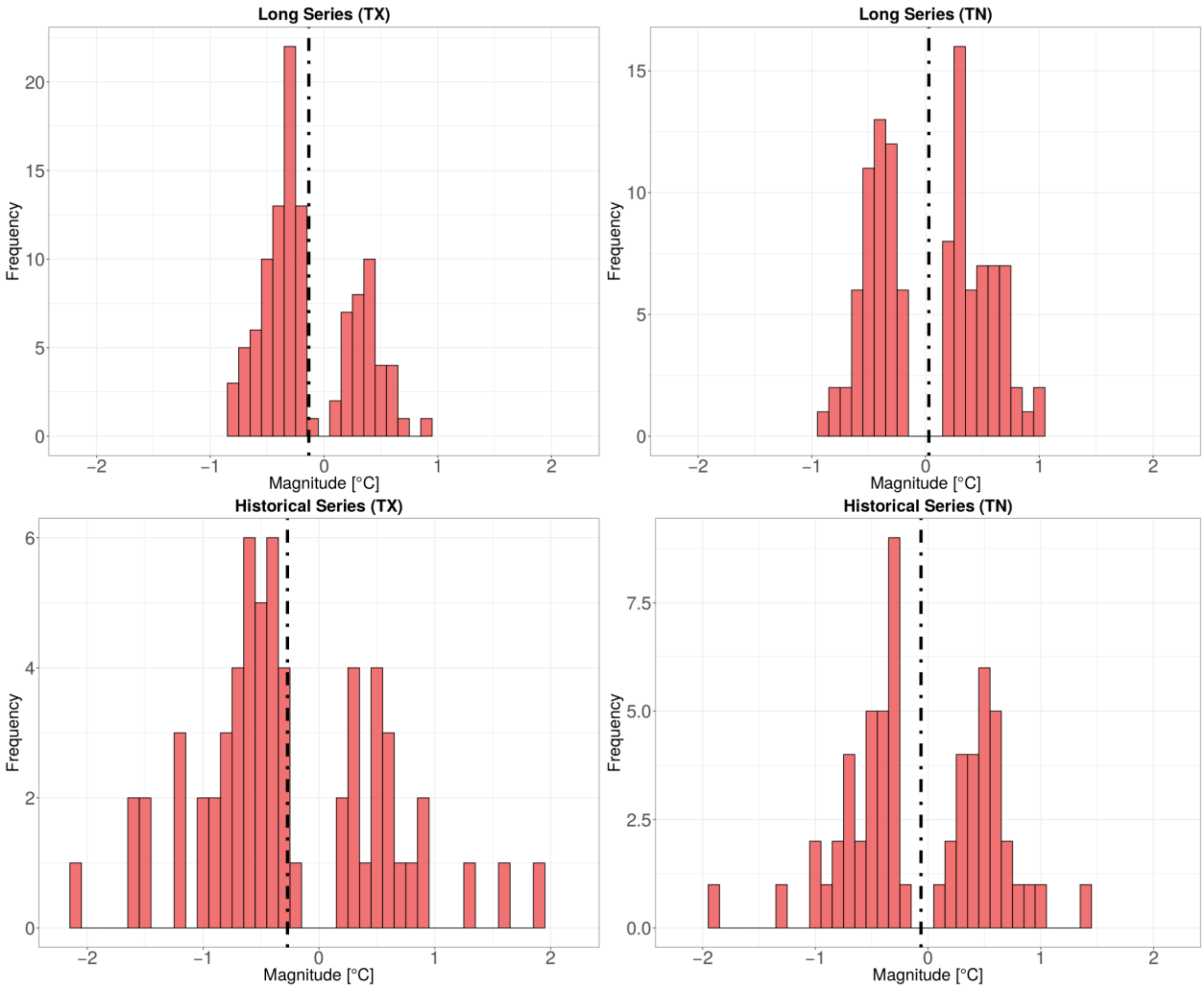


FIGURE 15: Frequency distribution of break-points magnitudes for the maximum (left panels) and minimum (right panels) temperature in historical series (lower panels) and long series (upper panels), respectively. The vertical dotted line represents the average of the distribution.

To assess the impact of the homogenisation on the temperature trend, seasonal and annual trends for TX and TN were computed for historical and long series with both original and homogenized time series. Note that for the historical series only the twelve locations having temperature data since at least 1895 (i.e. centennial series) were considered in the trends analysis to avoid bias. The trends were computed using the Theil-Sen estimator (Fernandes and Leblanc, 2005) (hereafter referred to as TS) which offers the advantage of being robust to outliers and by linear regression (hereafter referred to as LR). Because the TS results do not significantly differ from the LR results, only the LR results are discussed in the following. Table 7 summarizes the temperature trends (in °C/decade) before and after homogenization. The temperature trend of the homogenized long series during the period 1954-2015 is slightly larger for TX (0.30 °C/decade) than TN (0.26 °C/decade). The same observation is true for the period 1880-2015 with a temperature trend of 0.15 °C for TX and 0.13 °C for TN. The respective low p-values (lower than 10^{-6} for the long series and 10^{-12} for the historical series) indicate that these trends are statistically significant.

	TX trend (°C/decade)		TN trend (°C/decade)	
	Mean (min – max)	**p-value**	**Mean (min – max)**	**p-value**
Original (1954-2015)	0.26 (0.01 – 0.46)	0.2	0.29 (0.12 – 0.56)	1e-4
Homogenized (1954-2015)	0.30 (0.23 – 0.35)	5e-6	0.26 (0.21 – 0.33)	2e-6
Original (1880-2015)	0.02 (-0.11 – 0.09)	0.17	0.1 (-0.01 – 0.20)	0.8
Homogenized (1880-2015)	0.15 (0.14 – 0.15)	1e-12	0.13 (0.13 – 0.14)	5e-13

TABLE 7: Mean temperature trends (°C/decade) for long and centennial temperature time series. The minimum and maximum trends observed among all the series are indicated in brackets

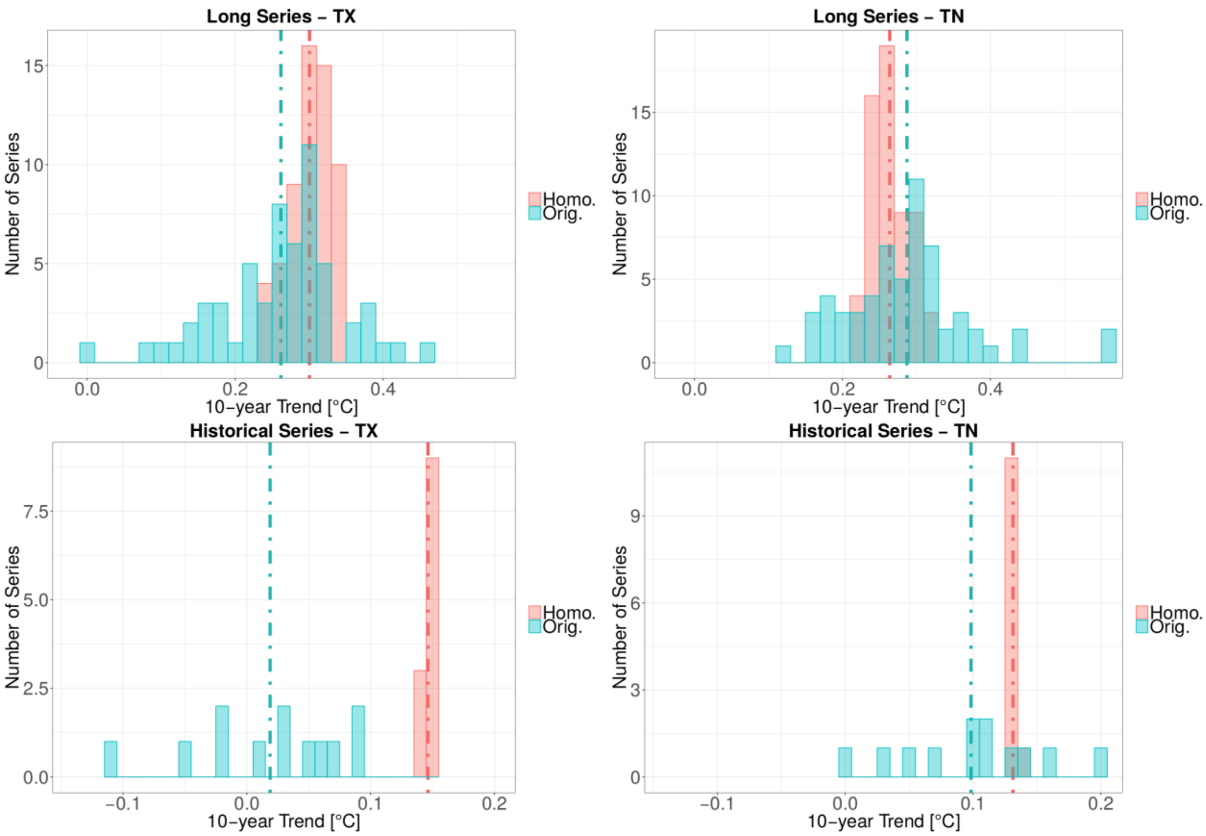


FIGURE 16: Magnitude distribution of the maximum (right panels) and minimum (left panels) temperature trends (°C/decade) for long series (upper panels) and historical (i.e. centennial) series (lower panels), before (i.e. Orig.) and after (i.e. Homo.) homogenization. The dash lines represent the mean of the trends.

Table 7 highlights a strong difference in the 1880-2015 TX time series before and after homogenization (i.e. 0.02 vs. 0.15 °C/decade, respectively). This difference can be explained by the type of stands used in the network for protecting thermometers before the introduction of the closed shelters. Both the open stand and more particularly the semi-opened prism (see types A and B in Figure 2, respectively) used to shelter the thermometers in the late 19th century and early 20th century are known to have biased upward TX (Vincent, 1912; Poncelet and Martin, 1947). Similarly, it is well known that the open shelter in use at Uccle (see Figure 3) induces warm bias on clear days.

Because daily maximum temperatures show original values that are too high during the first part of the period (some original centennial TX series even exhibiting a cooling trend), all homogenized TX time series present a warming trend due to the resulting downward adjustment (see lower left panel in Figure 16).

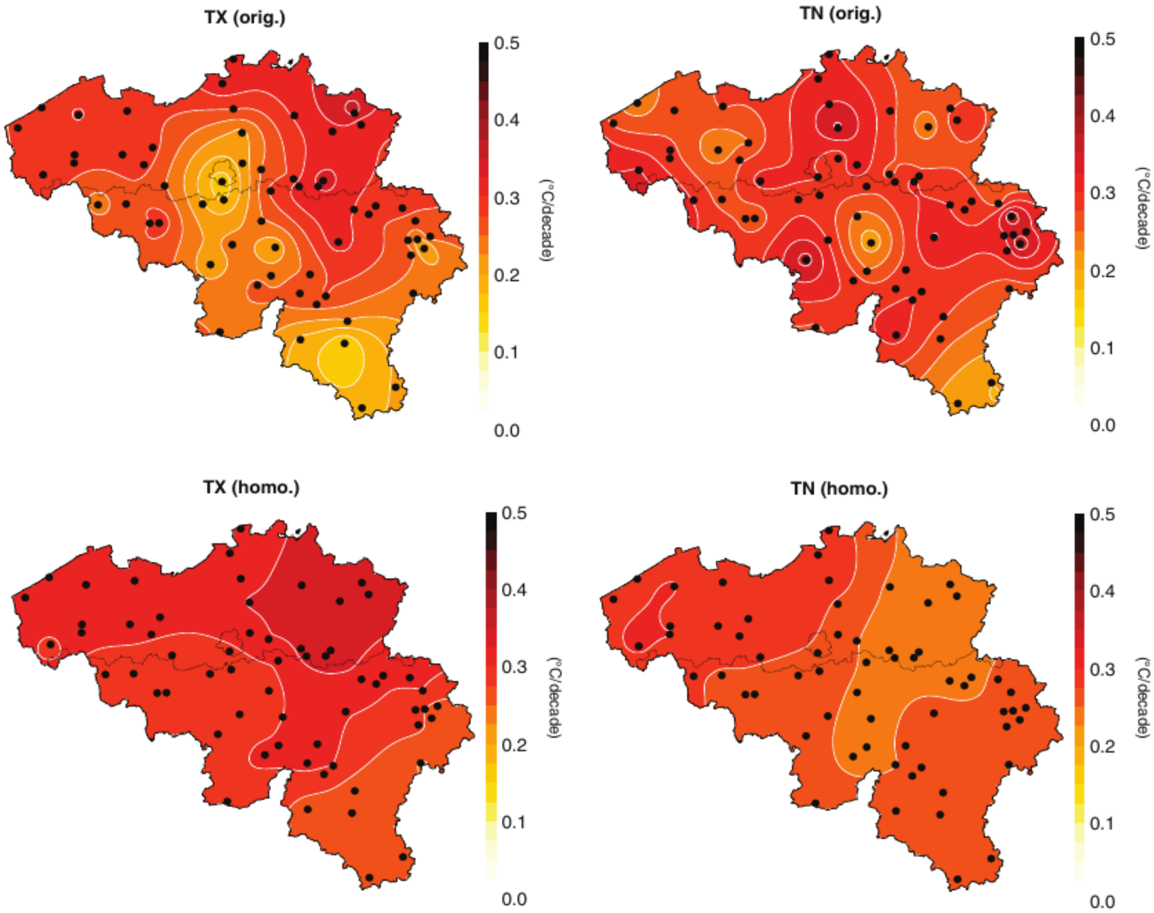


FIGURE 17: Interpolated temperature trends (°C/decade) of original (upper panels) and homogenized (lower panels) data for long series for maximum (left panels) and minimum (right panels) temperature.

Figure 16 presents the distribution of the trend of each time series before and after homogenization. Clearly, the homogenization process tends to reduce the spread of the trends. This is particularly noticeable for the historical TX and TN series (see lower panels in Figure 16). As mentioned previously, the historical series have been treated at once (i.e. in a single cluster) while five clusters were defined for the homogenization of the long series. Because relative homogenization assumes that all series in a given cluster are exposed to the same climate signal, the use of a unique cluster tends to remove the natural variations existing between the different stations.

Figure 17 compares the spatial distribution of the temperature trends for both TX and TN before and after homogenization. This analysis is proposed for the long series due to the limited number of available historical series. The homogenization largely reduces the spatial disparities. For the homogenized data, the spatial differences are moderate over Belgium and not similar for maximum or minimum temperature. In particular, the Campine region (in the northeast of Belgium) shows the most pronounced warming for TX and a relatively small warming for TN.

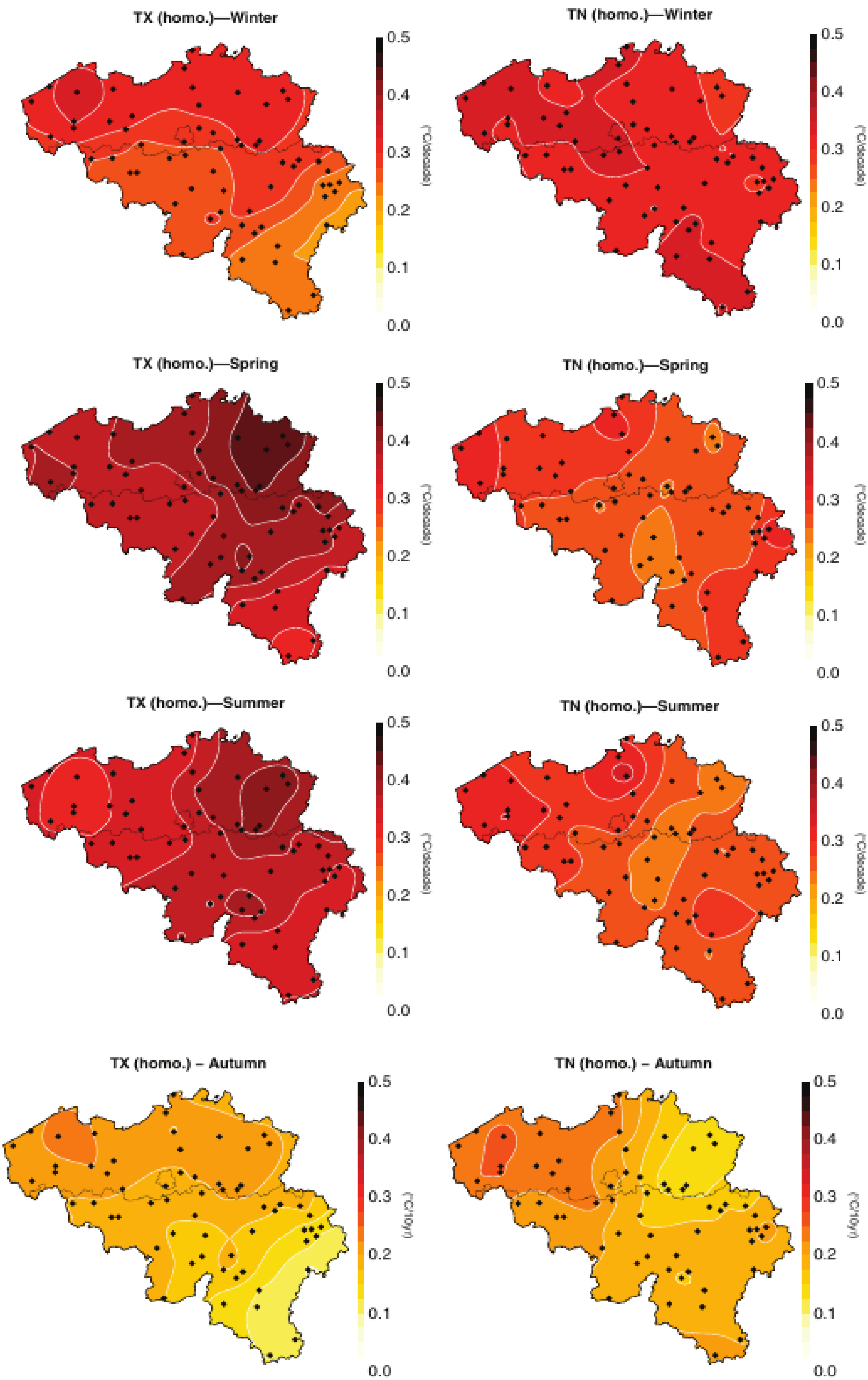


FIGURE 18: Seasonal maximum (left panels) and minimum (right panels) temperature trends (°C/decade) for homogenized long series (from top to bottom: winter, spring, summer, autumn)

Figure 18 presents an overview of the actual seasonal trends for both TX and TN based on the homogenized long series. Even if the temperature trend is always positive, it differs significantly according to the season. Spring and summer show the most pronounced warming concerning maximum temperature with an average trend of 0.38 °C/decade for all long series. For minimum temperature, the winter warming (average trend of 0.32 °C/decade) is slightly larger than the spring and summer warming (average trend of about 0.3 °C/decade). On the other side, the autumn shows the lowest trends with an average of 0.18 °C/decade for both TX and TN. All the seasonal trends are significant with a p-value below 10^{-4}. The Belgian temperature trends are consistent with the dataset of monthly homogeneous series of temperature produced by Météo-France (Gibelin et al., 2014). In the vicinity of the Belgium, the French temperature trends vary between 0.25 and 0.3 °C/decade and are slightly higher for TX than TN. The autumn shows also the smallest temperature increase.

3.3.1.2 Daily homogenisation

Based on the HOMER detected break-points on the monthly time series, the Vincent, HOM, SPLIDHOM and PM daily adjustment methods were first applied to the twelve centennial daily extreme temperature time series (see Table 2) and the resulting homogenization results were then compared to those obtained by using the QM daily homogenization method. In the application of the Vincent method, the HOMER correction values were taken to be the exact correction on the 15th of each month and all the other daily correction values were obtained by linear interpolation between two consecutive HOMER monthly values. For the HOM and SPLIDHOM methods the reference series was taken as the highest correlated series (with a minimum correlation of 0.8 between the candidate and reference detrended daily series). Finally, based on Trewin (2013) the PM95 variation of the PM method (i.e. the fixed points for defining the transfer function are the 5th to 95th percentiles) has been considered using at least 3 and up to ten sufficiently-correlated neighbours selected in descending order of correlation with the candidate series, with a lower correlation limit of 0.6 between the candidate and the neighbouring detrended daily series. Basically, the PM95 variation performs similarly to the PM99 variation (i.e. the fixed points for defining the transfer function are the 1st to 99th percentiles) of the PM method, but is generally much better than PM99 in simulating the highest and lowest values. It is likely that this reflects instability in the transfer functions towards the ends of the distribution when the 1st and 99th percentiles are used. Finally, it was observed that the five-neighbours case performs marginally worse than the ten-neighbours case adopted here.

Figure 19 displays the differences between the five different daily homogenization approaches applied to the centennial temperature time series. Both TX and TN data are additionally divided in three groups: the very low and the very high daily temperatures where the daily observations are respectively less than the 5th percentile and more than the 95th percentile and the intermediate temperatures in between. The percentiles were computed for each series separately from the 1961-1990 Vincent homogenized daily data which were further split by season. It is clearly apparent in Figure 19 that the QM corrections for TX largely differ from the correction made by the four daily adjustment methods using the HOMER detected break-points. The corrections differences between these four methods are relatively small excepted for the low TN observations (i.e. the correction is less than 1°C in 90 % of the cases and less than 2°C in 99.8% of the cases, respectively). Because the HOM and SPLIDHOM methods tend to consider exactly the same reference series for a given break-point, their corrections are only marginally different.

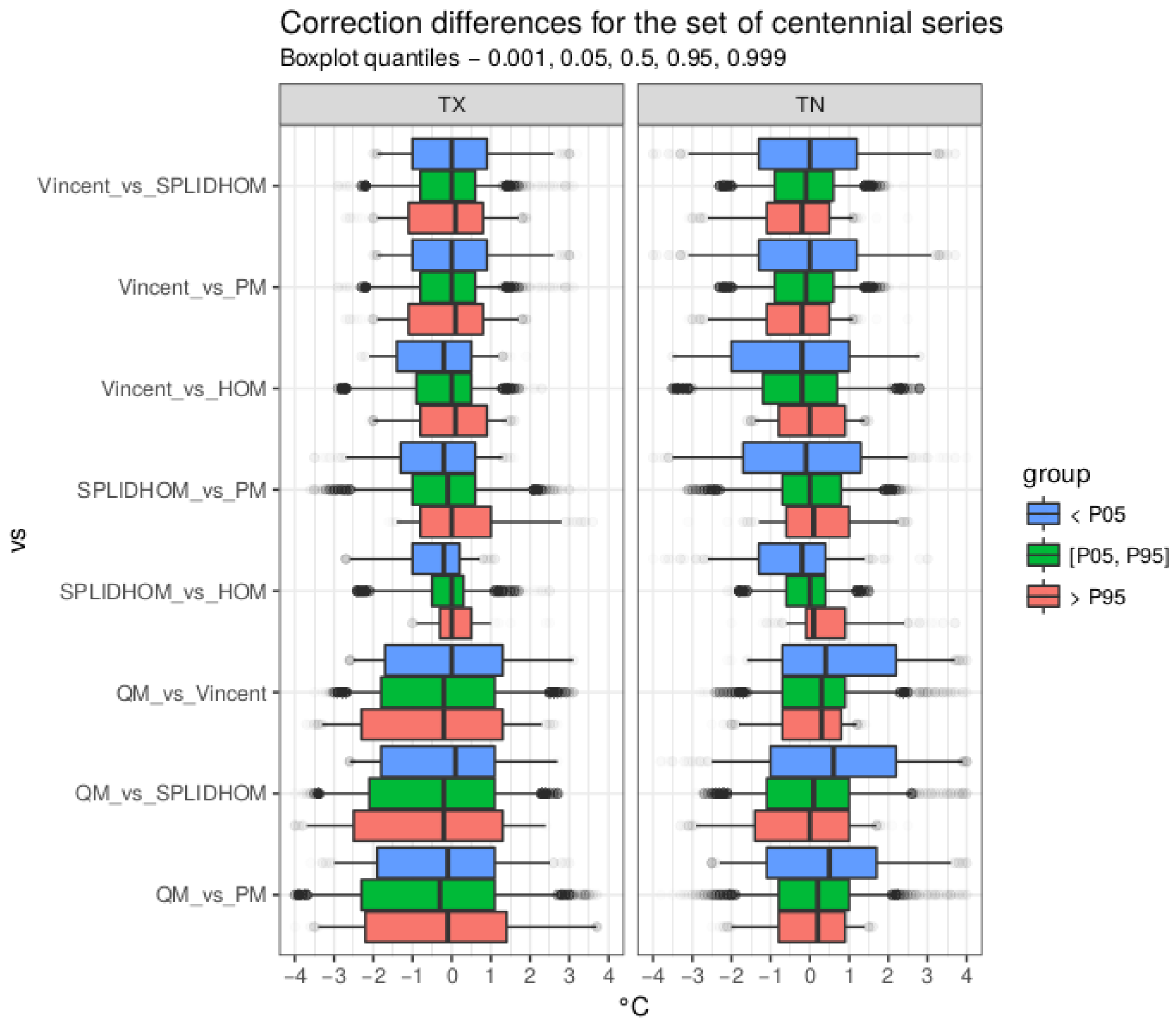


FIGURE 19: Comparison of the daily temperature corrections applied to the centennial temperature time series as provided by the Vincent, HOM, SPLIDHOM, PM and QM methods, respectively. The comparison is provided for both the TX (left panel) and TN (right panel) values and for 3 groups of temperature (i.e. very low, intermediate and very high), respectively.

Figure 20 compares the temperature trends (in °C/decade) between the daily original and homogenised centennial temperature time series at each series location. The decennial trends in original data largely fluctuate from one series to another one with some of them showing negative trends. After homogenization there is less spread in decennial trends with the Vincent methods exhibiting the more uniform trends between stations. There is some variation in the trends resulting from the application of the SPLIDHOM, HOM and PM methods to a given series, but they are globally close to each other and quite similar to those obtained with the Vincent method. In the contrary, the decennial trends obtained after the daily homogenization with the QM method vary just as much as in the original data. In addition the QM method tends to amplify the decennial trends in most of the TX series compared to the four other methods.

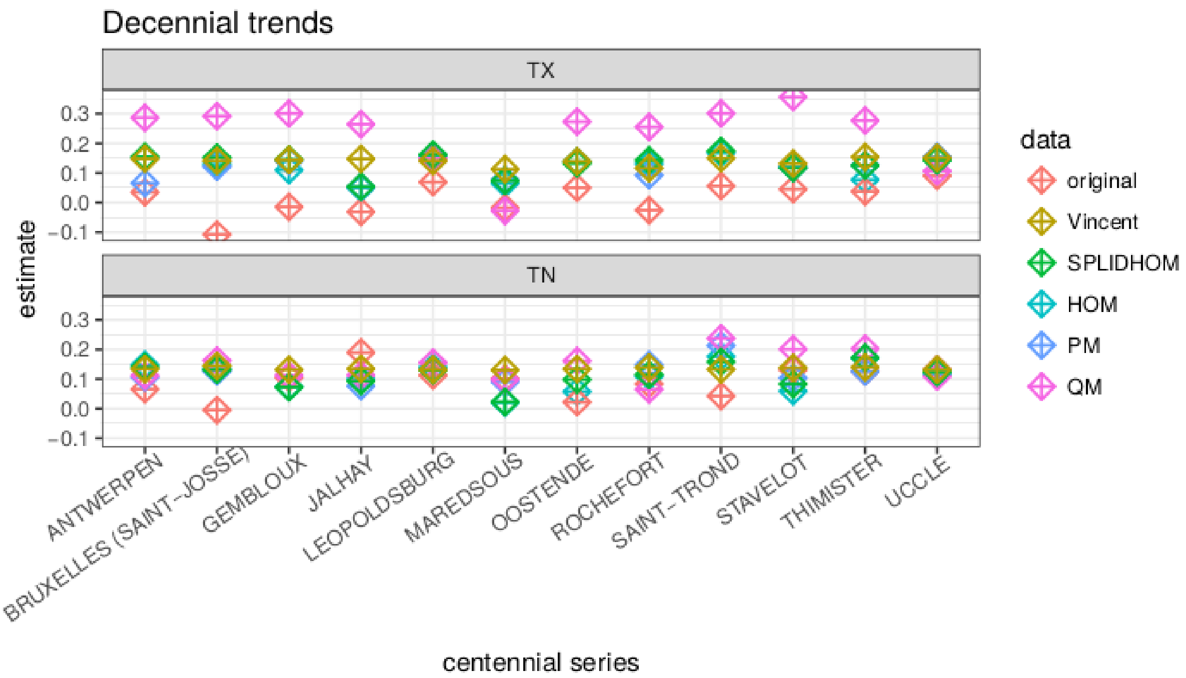


FIGURE 20: Comparison between the decennial temperature trends (in °C/decade) found in the original and homogenised centennial temperature time series. The comparison is provided at each centennial series location for both TX (upper panel) and TN (lower panel) and for the 5 daily adjustment methods, respectively.

Because of the very poor performances of the QM method compared to the four other daily adjustment methods and the fact that the HOM and SPLIDHOM methods produce very similar results, only the Vincent, SPLIDHOM and PM methods were considered to homogenize the TX and TN long series on a daily basis. Figure 21 displays the differences in daily corrections for the 3 adjustment methods applied to the long series of daily extreme temperature. The corrections differences are centred on zero for both TX and TN irrespective of the temperature groups. In almost all the comparison groups, 90 % of the corrections differ by less than about 0.75 °C and for 99.8 % of the corrections the differences are less than about 1.5 °C with the notable exception of the very low TN observations (as already pointed out for the centennial temperature series in Figure 19). There were only around 300 cases (out of more than 5 millions) for which the correction difference was larger than 3°C and most of them concerned TN values. The magnitude of the correction differences between the 3 methods is slightly lower than found for the centennial series (see Figure 19). This is a direct consequence of the larger number of available stations for the long series than for the centennial series which offers the opportunity to use more correlated reference series in the adjustment process.

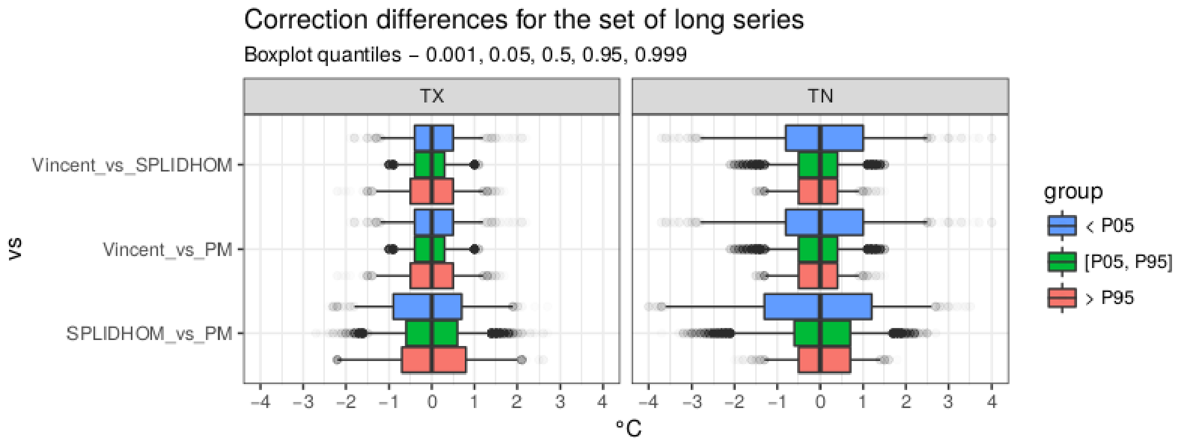


FIGURE 21: Comparison of the daily temperature corrections applied to the long temperature series as provided by the Vincent, SPLIDHOM and PM adjustment methods, respectively. The comparison is provided for both the TX (left panel) and TN (right panel) values and for 3 groups of temperature (i.e. very low, intermediate and very high), respectively.

Figure 22 compares the decennial trends before and after the daily homogenization of the long series of temperature. Clearly, the 3 methods reduce the spread in the decennial trends. The lowest spread in trends is obtained with the Vincent method for both TX and TN. This is not really surprising because the HOMER software, unlike the SPLIDHOM and PM methods which adjust the series one by one, determines the correction coefficients in a given cluster simultaneously for the set of stations.

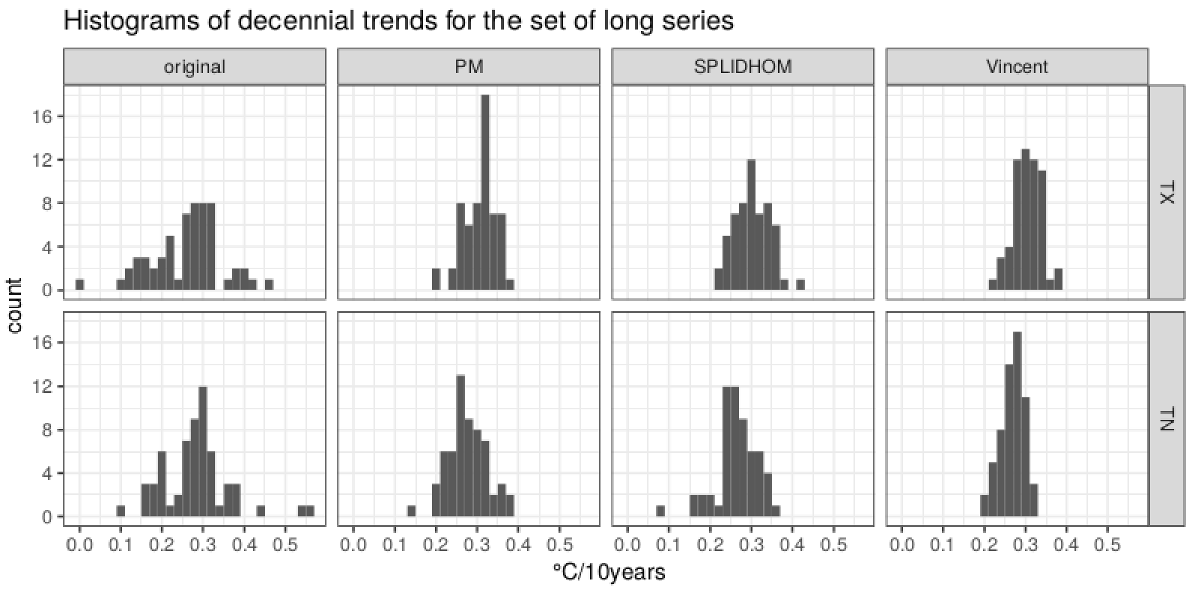


FIGURE 22: Comparison between the decennial temperature trends (in °C/decade) found in the original and homogenised long temperature time series. The comparison is provided for both TX (upper panel) and TN (lower panel) and for the Vincent, PM and SPLIDHOM daily adjustment methods.

3.3.2 Precipitation

3.3.2.1 Monthly homogenization

Similar to temperature, the method opted for homogenization of the monthly precipitation time series relies on the use of the HOMER software. However, in contrast to the temperatures, only the pairwise detection and the joint segmentation components of HOMER were used for the precipitation data. The ACMANT-component of HOMER was not considered because it is not intended for cumulative parameters such as precipitation data. For each precipitation time series, monthly values were computed when at least 90% of the daily values within a given month were found as valid by the QC process and otherwise considered as missing. The monthly homogenization was carried out separately for the centennial series and long series. Twenty tree additional foreign series (i.e. twelve in France, nine in Germany and two in the Netherlands – see Figure 7 for the stations' location) were considered to help the homogenization of the 149 Belgian long series and six of them (i.e. four in France, one in Germany and one in the Netherlands – see Figure 7 for the stations' location) having data since 1880 were also used in the homogenization of the 24 centennial series. To ensure that nearby locations are exposed to the same climate signal in the homogenization process, the monthly precipitation time series were clustered according to their climatological coherence and correlation. The long series were divided into 13 clusters (see Annexe A.2.1 for the clusters' composition) and the centennial series into 3 clusters (see Annexe A.2.2 for the clusters' composition) of about 15 series, respectively.

Following the recommendations by Venema et al. (2012) monthly values were adjusted using a coefficient estimated on annual values. To ensure the reliability of the homogenized long (resp. centennial) precipitation series, it was verified that all homogenized series satisfy the following conditions:

- Maximum 4 (resp. 8) break-points.
- No break-point with a magnitude larger than 30% (for precipitation, the comparison series are computed as the ratio between the candidate and the reference series).
- No accepted break-point with a magnitude lower than the break-point significance detection threshold provided by HOMER.

Number of breaks	**Centennial series (1880-2015)**	**Long series (1951-2015)**
0	6	91
1	5	48
2	6	8
3	5	2
4	1	/
5	1	/

Table 8: Number of detected break-points per precipitation series

Table 8 reports the number of break-points detected in the analysed time series. A maximum of 3 and 5 break-points were identified in the long and centennial series, respectively. The majority of the long series (i.e. 91/149) do not have any break-point and about 93% of the long series have less than 2 break-points. Six centennial series do not present any break-point and the majority of the centennial series (17/24) have less than 3 break-points. While varying from one series to another the HOMER's break-point significance detection threshold is in the order of 8% [6% - 11%] for the centennial series and 6% [3% - 9%] for the long series, respectively. This explains the relatively limited number of break-points detected in the precipitation time series by comparison to the temperature time series.

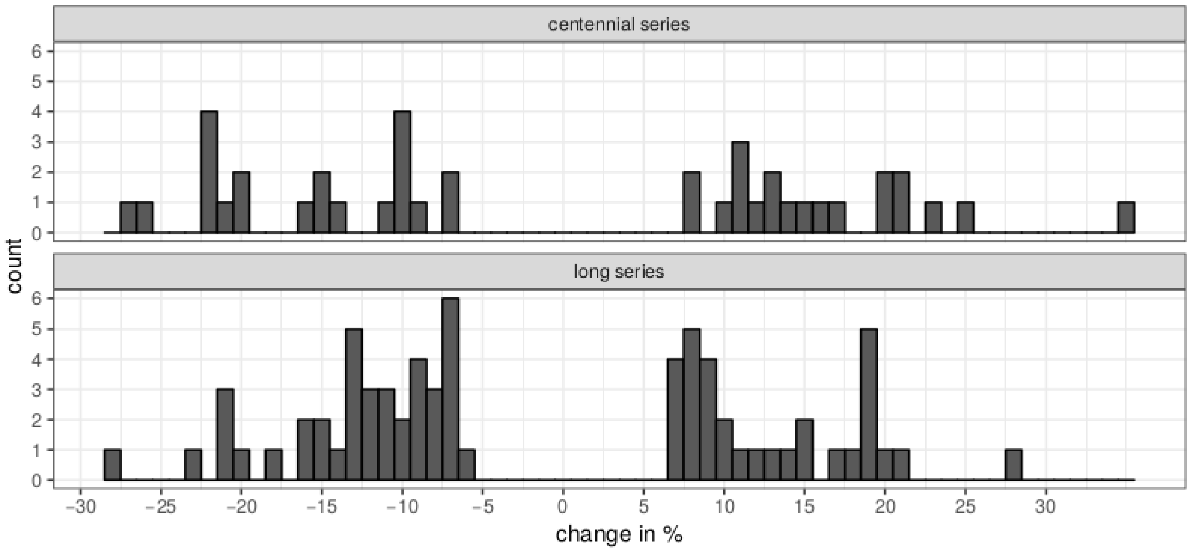


FIGURE 23: Frequency distribution of break-points magnitude (in %) for the centennial (upper panel) and long (lower panel) precipitation time series, respectively.

Figure 23 presents the frequency distribution of break-point magnitudes (in %) for both centennial and long precipitation time series. The largest detected break-point magnitude is +35% and occurs in the Hives series in 1928 (see Figure 24). While the Hives series is one of the 11 long term daily precipitation time series established by Dupriez and Demarée (1988) during a former digitalisation exercise (see Table 1), a break-point magnitude of 35% is larger than the threshold of 30% considered in our study and this series has been discarded. The largest decrease of -27% is found in the Gembloux precipitation time series and results from the station automation at the end of the 1990s. It is worth noting that this decrease is somewhat compensated a few years later by an increase of 25% (see Figure 24) when a new type of automated rain gauche was installed in the station.

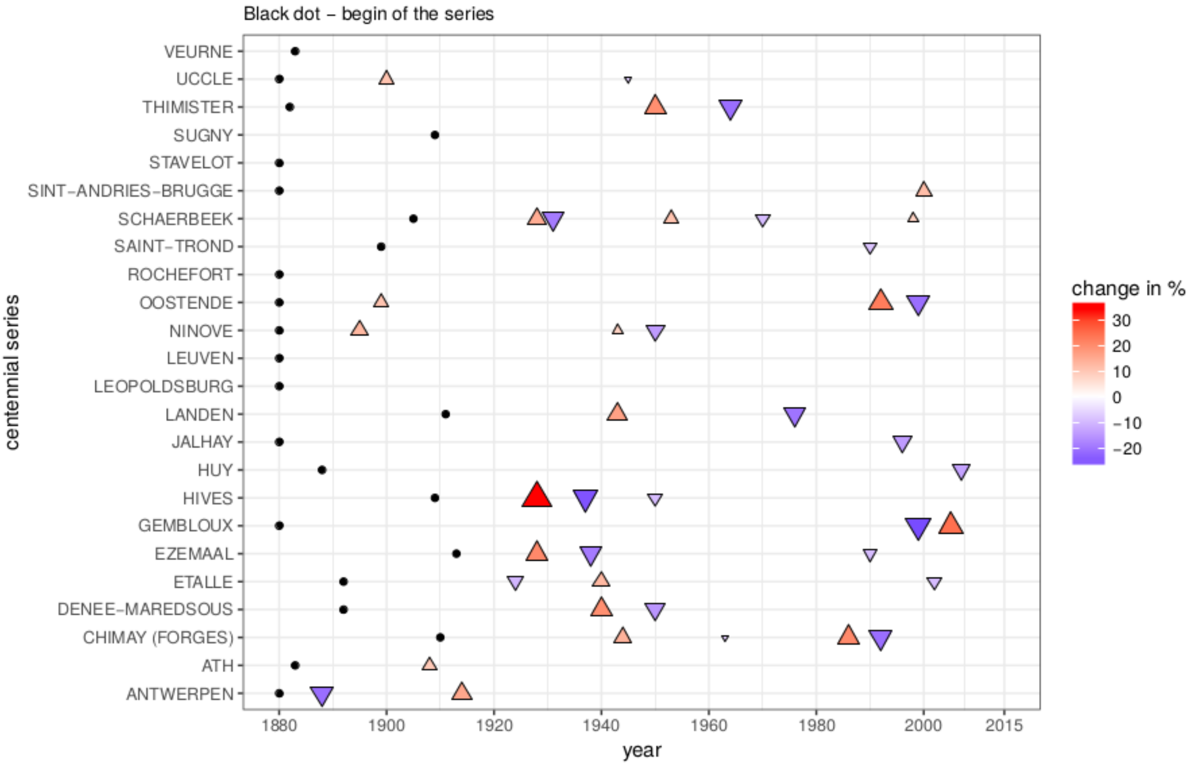


FIGURE 24: Detected breaks position and magnitude (in %) in the centennial precipitation time series.

Roughly 50% of the detected break-points are not directly supported by metadata but many of those are found, at least in the case of the centennial series, during a transition period of instrumental changes. As for the temperature time series, station catenations are responsible for the largest number of documented break-points (about 60%). The rest is attributed to observer changes, station relocations and station automatization.

Series Name	**Original series**		**Homogeneized series**	
	Trend (mm/yr)	***p-value***	***Trend (mm/yr)***	***p-value***
Jalhay	0.882	0.05183	1.370	7e-5
Oostende	1.373	1.9e-4	0.559	0.043
Leopoldsburg	1.330	6e-5	1.143	1.4e-4
Rochefort	1.178	6.3e-4	1.259	1e-5
Stavelot	1.567	0.0013	1.796	1e-5
Uccle	0.760	0.00688	0.898	7.4e-4
Sint-Andries-Brugge	1.340	4e-5	0.908	0.00491
Mean of the 7 longest series	1.260	1e-5	1.333	4e-5
Mean of the 23 centennial series	1.315	1e-5	1.247	3e-5

Table 9: Precipitation trends for the centennial series (mm/year). The trend is provided individually for each of the 7 longest and more complete series, their mean and for the mean of the 23 centennial series, respectively.

Table 9 presents the precipitation trends (in mm/year) before and after homogenization of the centennial series. The trends are given individually for each of the seven longest and more complete precipitation series, their mean and for the mean of the 23 centennial series, respectively. If the trends are positive in both the original and homogenized series the impact of the homogenization differs from one series to another one. As an example the trend is reduced after homogenization in three series and increased in four of them. The largest trend reduction after homogenization is reported for the Oostende series (i.e. 1.373 mm/yr in the original series vs. 0.559 mm/yr in the homogenized series). In the reverse, the largest trend increase after homogenization is for the Jalhay series (i.e. 0.882 mm/yr in the original series vs. 1.370 mm/yr in the homogenized series). When considering all the seven series together the mean precipitation trend is slightly increased after homogenization (i.e. a mean trend of 1.333 mm/yr after vs 1.260 mm/yr before homogenization, respectively) but the reverse situation occurs when all the 23 centennial series are considered (i.e. a mean trend of 1.247 mm/yr after vs. 1.315 mm/yr before homogenization, respectively).

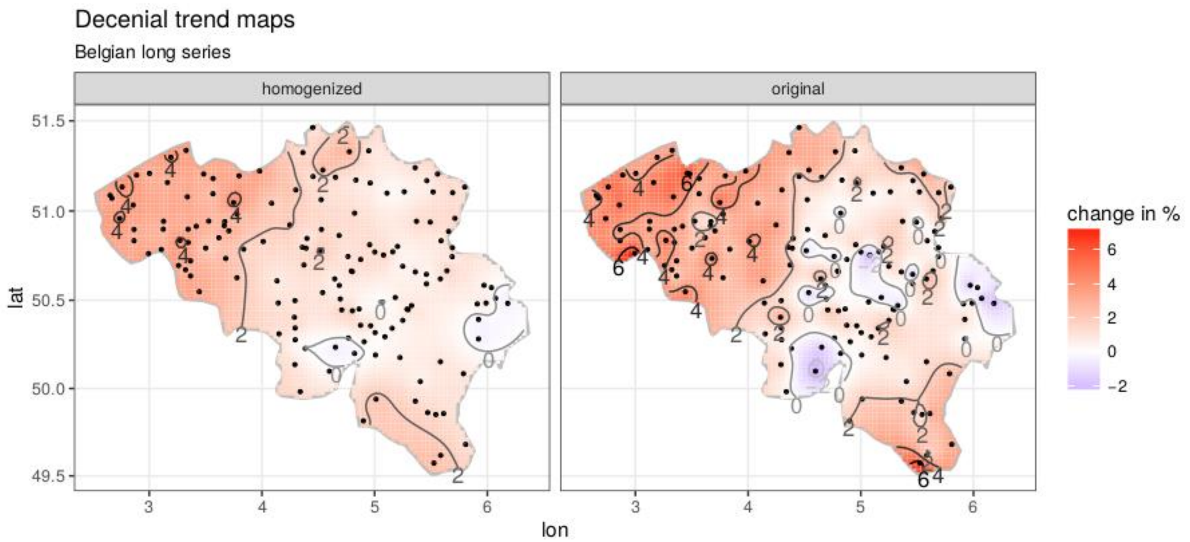


FIGURE 25: Interpolated precipitation trends (decenial change in %) before (right panel) and after (left panel) homogenization of the Belgian long series of precipitation (1951-2015).

Figure 25 compares the spatial distribution of the precipitation trends before and after homogenization. This analysis is proposed for the long series due to the limited number of available centennial series. Note that the trends in Figure 25 are expressed in percentage change per decade rather than in mm change per decade because the annual normals of precipitation in Belgium vary from simple to double (i.e. from 740 mm/year in the northern part of the Hesbaye to more than 1400 mm/year in the Hautes Fagnes). For each long series, the change in % has been computed as following: [(change in mm per decade)/(mean over the years 1961-1990)*100]. The homogenization reduces the spatial disparities within the differents natural regions of Belgium. For the homogenized data, the most apparent changes in the precipitation regime appear in the Polders and Flanders regions and to a lesser extend in Belgian Lorraine and Campine. By contrast, no significant decenial trend is reported in the Hautes Fagnes.

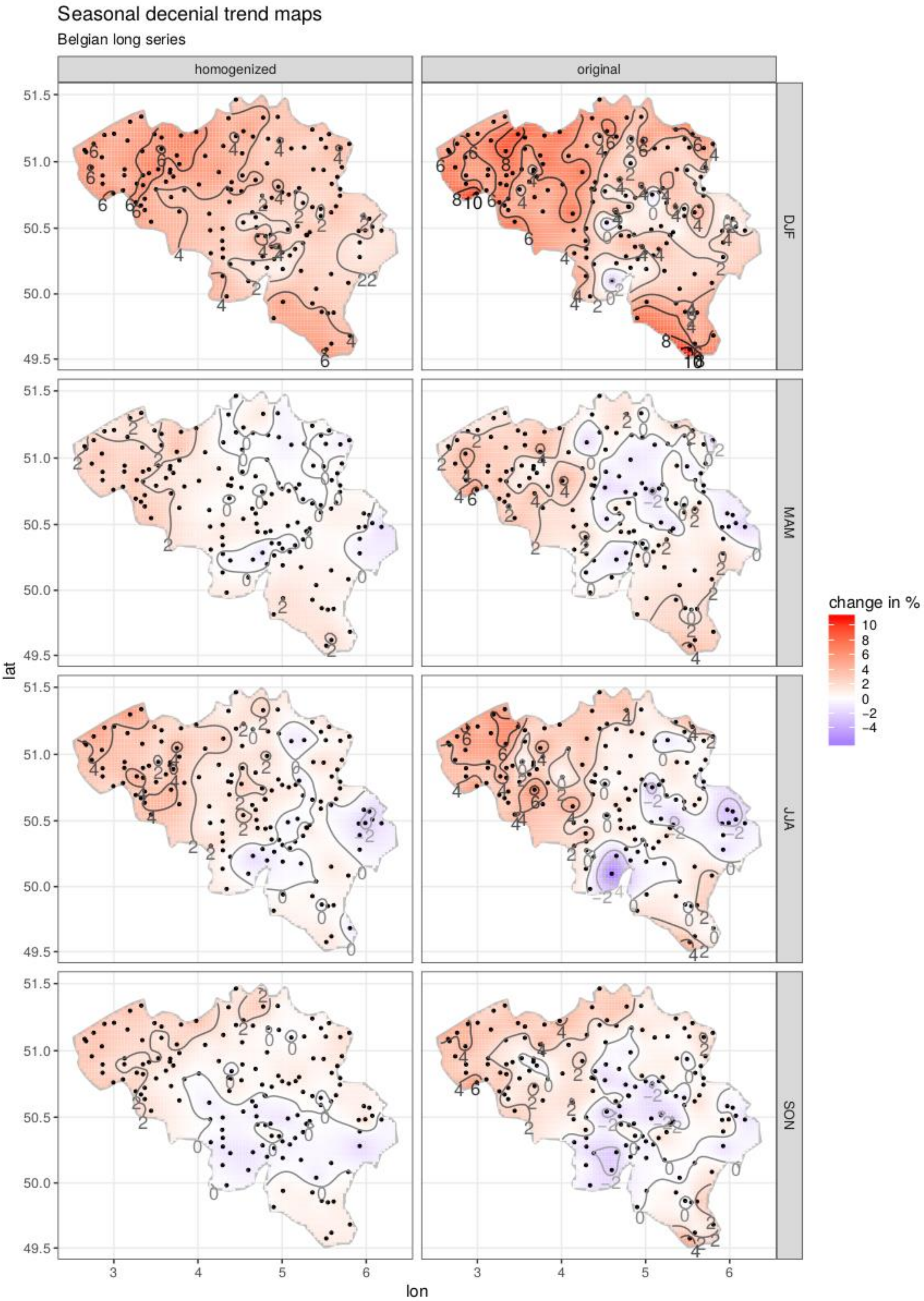


FIGURE 26: Seasonal precipitation trends (decenial change in %) before (right panels) and after (left panels) homogenization of the Belgian long series of precipitation (1951-2015).

Figure 26 presents an overview of the seasonal trends for both the original and homogenized long series. If the precipitation trend is always positive in the Polders and Flanders regions, it differs in the other regions according to the season. The winter shows an increase in precipitation for all long series with the most pronouced increase in the Polders, Flandres and Belgian Lorraine regions. On the other side, the summer shows the largest decreasing trend in the Hautes Fagnes. Spring and autumn show the lowest trends in both precipitation increase and decrease.

3.3.2.2 Daily homogenisation

As mentionned in Section 2.2.2, two methods have been considered for the homogenization of the daily precipitation amount: a direct application of the HOMER's yearly correction factors to the daily data (i.e. the Vincent method) and the QM method. Figure 27 summarizes the correction differences between both methods according to the daily precipitation amount. If both methods produce quite similar results for low precipitation values, the discrepancies between them increase as the precipitations amounts increase. Globally, the QM method tends to produce larger daily precipitation amount than the Vincent method and it is interesting to note that the boxplot median value is very similar for the centennial and the long series within a given precipitation group.

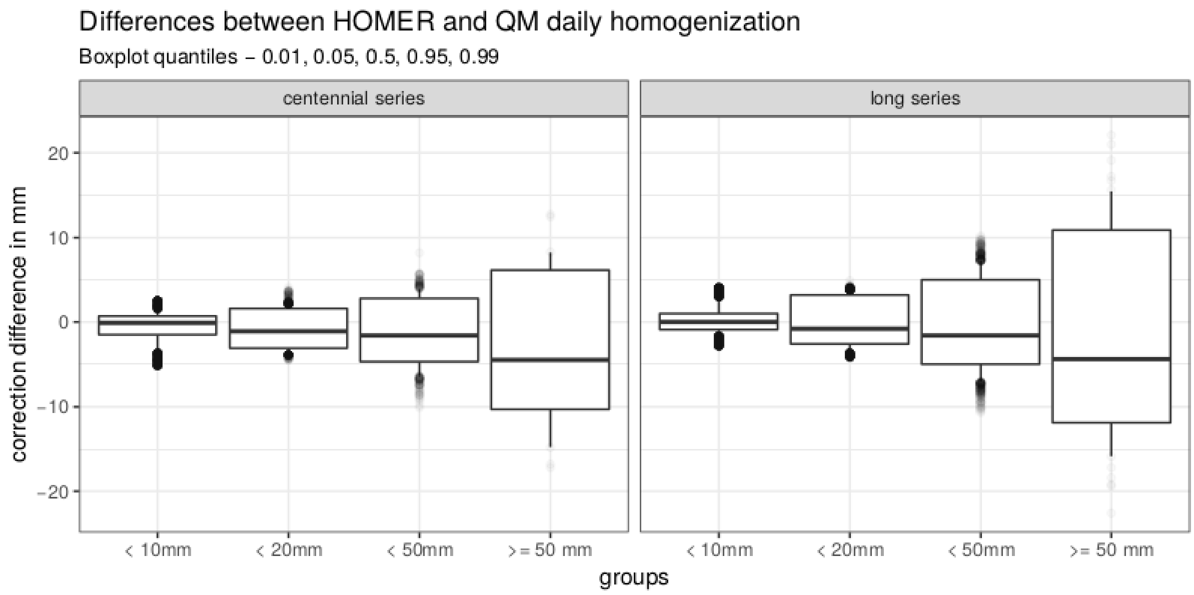


FIGURE 27: Correction differences between a direct application of the HOMER yearly correction factors and the QM method for both centennial series (left panel) and long series (right panel) as a function of the daily precipitation amount.

Figure 28 compares the distribution of the decennial trends before (upper panels) and after the daily homogenization of the long (left panels) and the centennial (right panels) precipitations time series with the QM (middle panels) and Vincent (lower panels) methods, respectively. Clearly, while the daily homogenization performed according to the Vincent method (i.e. a direct application of the HOMER's yearly correction coefficients to the daily values) reduces the spread of the trends by comparison to the original data for both the long and centennial series, the QM method completely fails to produce more uniform decennial trends. In the contrary the spread of the trends is enlarged for both the long and the centennial series meaning that the QM method tends to produce even more inhomogeneous time series than the original ones. This well illustrates the limitations of an absolute homogenization method (i.e. without using neighbouring stations) even when compared to the simplest daily relative homogenization approach based on yearly correction coefficients.

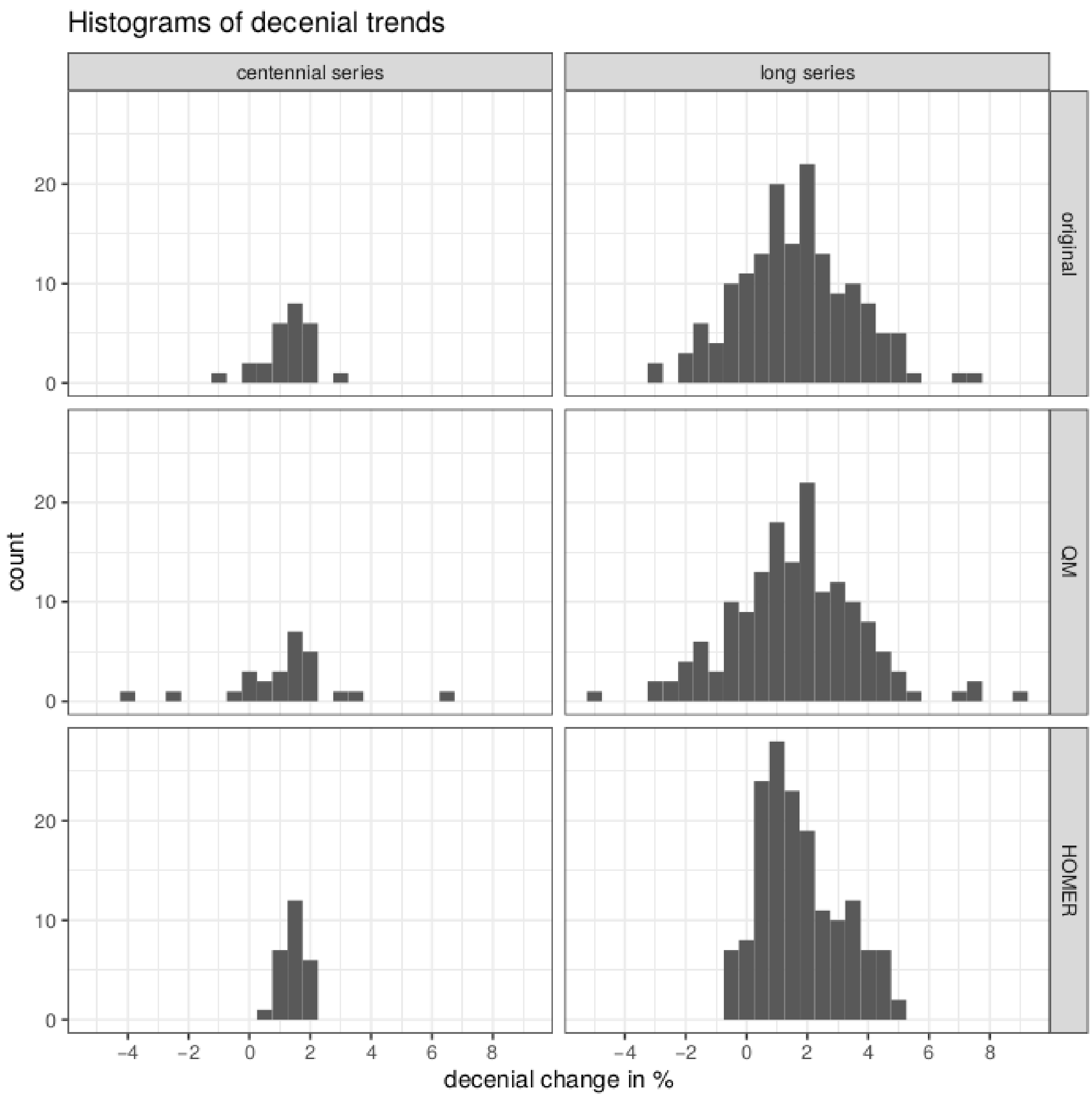


FIGURE 28: Magnitude distribution of the precipitation trends (decennial change in %) for long series (right panels) and centennial series (left panels) before (upper panels) and after daily homogenization using the QM method (middle panels) and the Vincent method (i.e. HOMER's yearly correction coefficients, lower panels), respectively.

4. SCIENTIFIC RESULTS AND RECOMMENDATIONS

A new dataset of quality controlled and homogenized monthly and daily extreme temperature and precipitation amount has been created for Belgium. Homogenization results are provided for 61 long temperature series over the period 1954-2015 and for 16 historical series starting before 1931, including 8 series covering the full time period 1880-2015. Three different versions of daily homogenized temperature time series are available. In a first version, the daily adjustment relies on the interpolation of the monthly adjustment coefficient according to the Vincent method. In a second version the daily adjustments are based on non-parametric regression (by means of cubic smoothing spline) as computed by the SPLIDHOM method. And finally, in a third version, daily adjustments were obtained using the PM method which allows considering a combination of multiple reference stations. Based on simulated examples Mestre et al. (2011) have shown that SPLIDHOM technique improve HOM (especially in terms of RMSE) and Vincent's method for the correction of extreme quantiles if correlation is high enough (above 0.90). When correlation is lower than 0.9 Vincent's method is often superior and thus should not be neglected. Because correlation larger than 0.9 between candidate and reference series was in practice not reached in most of the cases (if any; note that even a correlation of 0.8 was not always found between the candidate and the reference series), we do not recommend the use of the temperature time series homogenized on a daily basis with SPLIDHOM. However, Vincent's method is not designed for adjusting extreme values. In consequence only the daily homogenized temperature time series using the PM algorithm should be considered for analyses of changes in the extremes.

Regarding the precipitation, homogenization results are provided for 149 long series covering the time period 1951-2015 and 23 centennial series. It is worth pointing out that the historical series of Hives established in a former digitalisation exercise (e.g. Dupriez and Demarée, 1988) has been discarded here because of an excessive break-point (magnitude of 35%) identified in this series in 1928. Finally, contrary to the temperature, only one version of the daily homogenized precipitation time series is delivered. Indeed the QM method fails to provide valuable daily homogenised precipitation time series; the resulting series being for some of them even more inhomogeneous than the originals ones. Therefore, only the daily homogenization obtained with the Vincent method is available for the precipitation time series.

5. DISSEMINATION AND VALORISATION

- Alioscha-Perez M. and H. Sahli, 2017: A machine learning perspective towards fully automated quality control in daily weather time series. Paper presented at the 9th seminar for homogenization and quality control in climatological databases and 4th conference on spatial interpolation techniques in climatology and meteorology, 03-07 April 2017, Budapest, Hungary.
- Bertrand C. and M. Journée, 2017: Data QC within the Belgian synoptic and climatological networks: an overview. Paper presented at the 9th seminar for homogenization and quality control in climatological databases and 4th conference on spatial interpolation techniques in climatology and meteorology, 03-07 April 2017, Budapest, Hungary.
- Bertrand C., C. Delvaux, R. Ingels, V. Vrábeĺ and M. Journée, 2017: Belgian homogenized reference climate time series. Poster presented at the 11th EUMETNET Data Management Workshop "Placing climate data to social service: From observations to archives", 18-20 October 2017, Zagreb, Croatia.
- Delvaux C., R. Ingels, V. Vrábeĺ, M. Journée and C. Bertrand, 2017: Quality control and homogenisation of the Belgian historical weather data. Paper presented at the 9th seminar for homogenization and quality control in climatological databases and 4th conference on spatial interpolation techniques in climatology and meteorology, 03-07 April 2017, Budapest, Hungary.
- Delvaux C., R. Ingels, M. Journée, V. Vrábeĺ and C. Bertrand, 2018: Quality Control and Homogenization of Monthly Temperatures in Belgium. Poster presented at the European Conference for Applied Meteorology and Climatology 2017, 04-08 September 2017, Dublin, Ireland.

6. PUBLICATIONS

- Alioscha-Perez M., Oveneke M.C. and Sahli H., 2018: SVRG-MKL: a Fast and Scalable MKL Solution for Features Combination in Classification. Under revision in IEEE Transactions on Neural Networks and Learning Systems.
- Alioscha-Perez M., Journée M., Bertrand C. and Sahli H., 2018: Deep Learning Meets Historical Weather: Automated Detection and Correction of Anomalies in Historical Daily Temperature Records. Draft ready for submission to Geophysical Research Letters.
- Bertrand C., V. Vrábe, C. Delvaux and M. Journée, 2018: Homogenization of the Belgian historical precipitation time series. Article in preparation for submission to International Journal of Climatology.
- Delvaux C., R. Ingels, M. Journée, V. Vrábeĺ and C. Bertrand, 2017: Quality Control and Homogenization of Monthly Temperatures in Belgium. EMS Annual Meeting Abstracts, Vol. 14, EMS2017-285.
- Delvaux C., R. Ingels, V. Vrábeĺ, M. Journée and C. Bertrand, 2018 : Quality control and homogenization of the Belgian historical temperature data. International Journal of Climatology, DOI:10.1002/joc.5792.

7. ACKNOWLEDGEMENTS

The authors acknowledge Anne-Laure Gibelin and Brigitte Dubuisson from Météo-France (Direction de la climatologie) for having provided us with their historical climate data and the precious help for data homogenization with HOMER and SPLIDHOM.

REFERENCES

- Abdelaal, R.E. and Elhadidy, M.A., 1995: Modelling and forecasting the daily maximum temperature using adductive machine learning. Wea. Forecast. 10:310-325.
- Aguilar, E., Auer, I., Brunet, M., Peterson, T.C., Wieringa, J., 2003: Guidelines on Climate Metadata and Homogenization. World Meteorological Organization: WMO-TD No. 1186, WCDMP No. 53, Switzerland, Geneva.
- Allen, R.J., and DeGaetano, A.T., 2000: A method to adjust long-term temperature extreme series for non-climatic inhomogeneities. J. Climate, 13:3680-3695.
- Alexandersson, H., 1986: A homogeneity test applied to precipitation data. Journal of Climatology, 6:661-675.
- Alexandersson, H. and Moberg, A., 1997: Homogenization of Swedish temperature data. Part 1. Homogeneity test for linear trends. Int. J. Climatol., 17:25-34.
- Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Aguillar, E., Sigro, S., Castella, M., Unden, P., Harris, I., Jones, P. and Brunet, M., 2018: A rescuded dataset of sub-daily meteorological observations for Europe and southern Mediterranean region, 1877-2012. Earth System Science Data, Discussion started: 3 May 2018. https://doi.org/10.5194/essd-2018-39, in review.
- Auer, I., Böhm, R. and Schöner, W., 2001: Austrian Long-Term Climate 1767-2000: Multiple Instrumental Climate Time Series from Central Europe. Österreichische Beiträge zu Meteorologie und Geophysik, Vol. 25. Vienna: Zentralanstalt für Meteorologie und Geodynamik (ZAMG), 147 pp.
- Auer, I., et al., 2007: HISTALP: Historical instrumental climatological surface time series of the Greter Alpine Region. Int. J. Climatol., 27:17-46.
- Brandsma, T., 2004: Parallel air temperature measurements at the KNMI-terrain in De Bilt (the Netherlands) May 2003-April 2004 Interim report, Koninklijk Nederlands Meteorologisch Instituut publicatie 207, 29 pp.
- Brandsma, T., and Können, G.P., 2006: Application of nearest-neighbor resampling for homogenizing temperature records on daily to sub-daily level. Int. J. Climatol., 26:75-89.
- Brunet, M., Saladie, O., Jones, P.D., Sigro, J., Aguilar, E., Moberg, A., Lister, D., Walther, A., Lopez, D., Almarza, C., 2006: The development of a new dataset of Spanish daily adjusted temperature series (1850-2005). Int. J. Climatol., 26:1777-1802.
- Brunet, M., Saladié, O., Jones, P.D., Sigro, J., Aguilar, E., Moberg, A., Lister, D., Walther, A. and Almarza, C., 2008: A case-stu dy/guidance on the development of long-term daily adjusted temperature datasets. WCDMP-66/WMO-TD-1425, Geneva, 43 pp.
- Brunet, M., Asin, J., Sigro, J., Banon, M., Garcia, F., Aguilar, E., Palenzuela, J.E., Peterson, T.C., Jones, P.D., 2011: The minimization of the screen bias from ancient Western Mediterranean air temperature records: an exploratory statistical analysis. Int. J. Climatol., 31:1879-1895.

- Caussinus, H. and Lyazrhi, F., 1997: Choosing a linear model with random number of change-points and outliers. Ann. Inst. Statist. Math., 49:761-775.
- Caussinus, H. and Mestre, O., 2004: Detection and correction of artificial shifts in climate series. Appl. Statist., 53:405-425.
- Chandola, V., Banerjee, A. and Kumar, V., 2009: Anomaly detection: A survey. ACM computing surveys (CSUR), 41(3):15
- Cheboli, 2010: Anomaly detection of time series. Master Thesis, 75 pp. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/92985
- Conrad, V. and Pollak, L.W., 1950: Methods in Climatology. Harvard University press, 8 vo., 459 pp.
- Della-Marta, P.M., Collins, D. and Braganza, K., 2004: Updating Australia's high-quality annual temperature dataset. Australian meteorological magazine, 53(2):75-93.
- Della-Marta, P.M. and Wanner, H., 2006: A Method of Homogenizing the Extremes and Mean of Daily Temperature Measurements. J. Climate, 19:4179-4197.
- Delvaux, C., Journée, M. and Bertrand, C., 2015: The FORBIO-Climate data set for climate analyses. Adv. Sci. Res., 12:103-109.
- Demarée, G., Lachaert, P-J., Verhoeve, T. and Thoen, E., 2002: The long-term daily Central Belgium Temperature (CBT) series (1767-1998) and early instrumental meteorological observations in Belgium. Climatic Change, 53:269-293.
- Domonkos, P., 2011: Adapted Caussinus-Mestre Algorithl for Networks of Temperature series (ACMANT). Int. J. Geosci., 2:293-309.
- Domonkos, P., 2013: Measuring performances of homogenization methods. Q. J. Hung. Meteor. Serv., 117 :91-112.
- Domonkos, P., Poza, R. and Efthymiadis, D., 2011: Newest developments of ACMANT. Adv. Sci. Res., 6:7-11.
- Dupriez, G.L. and Demarée, G., 1988: Totaux pluviométriques sur des periods continues de 1 à 30 jours. I. Analyse de 11 series pluviométriques de plus de 80 ans. Institut Royal Météorologique de Belgique, Miscellanea Série A, N° 8, 154 pp., Brussels, Belgium.
- Easterling, D.R., and Peterson, T.C., 1995: A new method for detecting undocumented discontinuities in climatological time series. Int. J. Climatol., 15:369-377.
- Eskridge, R.E., Alduchov, O.A., Chernykh, I.V., Panmao, Z., Polansky, A.C., and Doty, S.R., 1995: A comprehensive aerological reference data set (CARDS): Rough and systematic errors. Bull. Am. Meteorol. Soc., 76:1759-1775.
- Feng, C., Li, T. and Chana, D., 2017: Multi-level anomaly detection in industrial control systems via package signatures and lstm networks. In Dependable systems and networks (dsn), 2017 47th annual IEEE/IFIP International Conference, p. 261-272.
- Fernandes, R. and Leblanc, S., 2005: Parametric (modified least squares) and non-parametric (Theil-Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors. Remote Sensing of Environment, 95:303-316.
- Gandin, L.S., 1988: Complex quality control of meteorological observations. Mon. Weather Rev., 116:1137-1156.

- Gibelin, A.-L., Bubuisson, B., Corre, L., Deaux, N., Jourdain, S., Laval, L., Piquemal, J.-M., Mestre, O., Dennetière, D., Desmidt, S. and Tamburini, A., 2014 : Evolution de la température en France depuis les années 1950 : constitution d'un nouveau jeu de séries homogénéisées de référence. La Météorologie, 87 :45-53.

- Gubler, S., Hunziker, S., Begert, M., Croci-Maspoli, M., Konzelmann, T., Brönnimann, S., Schwierz, C., Oria, C. and Rosa, G., 2017: The influence of station density on climate data homogenization. Int. J. Climatol., 37:4670-4683.

- Han, J. and Kamber, M., 2006: Data Mining: Concepts and Techniques, 2nd ed. The Morgan Kaufmann Publishers, March 2006. ISBN 1-55860-901-6.

- Hanssen-Bauer, I. and Forland, E.J., 1994: Homogenizing long Norwegian Precipitation series. Journal of Climate, 7:1001-1013.

- Hawkins, D.M., 1977: Testing a sequence of observations for a shift in location. J. Amer. Statist. Assoc., 72:180-186.

- Kuglitsch, F.G., Auchmann, R., Bleisch, R., Brönniman, S., Martius, O. and Stewart, M., 2012: Break detection of annual Swiss temperature series. J. Geophys. Res., 117:D13105, doi:10.1029/2012JD017729.

- Kuglitsch, F.G., Toreti, A., Xoplaki, E., Della-Marta, P.M., Luterbacher, J. and Wanner, H., 2009 : Homogenization of daily maximum temperature series in the Mediterranean. Journal of Geophysical Research: Atmosphere, 114, D15108, doi:10.1029/2008JD011606.

- LeCun, Y., Bengio, Y. and Hinton, G., 2015: Deep learning. Nature, 521(7553):436.

- Lipton, Z.C., Kale, D.C., Elkan, C. and Wetzel, R., 2015: Learning to diagnose with lstm recurrent neural networks. arXiv preprint arXiv:1511.03677.

- Lucio, P.S., Conde, F.C., Cavalcanti, I.F.A., Serrano, A.I., Ramos, A.M., and Cardoso, A.O., 2007: Spatiotemporal monthly rainfall reconstruction via artificial neural network – case study: south of Brazil. Adv. Geosci., 10:67-76.

- Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P. And Shroff, G., 2016: Lstm-based encoder-decoder for multi-sensor anomaly detection. arXiv preprint arXiv:1607.00148.

- Mamara, A., Argiriou, A.A. and Anadranistakis, M., 2014: Detection and correction of inhomogeneities in Greek climate temperature series. International Journal of Climatology, 34:3024-3043.

- Mamara, A., Argiriou, A.A. and Anadranistakis, M., 2016: Recent trend analysis of mean air temperature in Greece based on homogenized data. Theoretical and Applied Climatology, 126:543-573.

- Maronna, R. and Yohai, V.J., 1978: A bivariate test for the detection of a systematic change in mean. Journal of the American Statistical Association, 73(363):640-645.

- Mestre, O., Gruber, C., Prieur, C., Caussinus, H., Jourdain, S., 2011: SPLIDHOM: a method for homogenization of daily temperature observations. J. Appl. Meteor. Climatol., 50:2343-2358.

- Mestre, 0., Domonkos, P., Picard, F., Auer, I., Robin, S., Lebarbier, E., Boehm, R., Aguilar, E., Guijarro, J.A., Vertachnik, G., Klancar, M., Dubuisson, V., and Stepanek, P., 2013: HOMER: a homogenization software – methods and applications. Quaterly Journal of the Hungarian Meteorological Service, 117(1):47-67.

- Nemec, J., Gruber, C., Chimani, B. and Auer, I., 2013: Trends in extreme temperature indices in Austria based on a new homogenized dataset. International Journal of Climatology, 33(6):1538-1550.

- Peterson, T.C., Easterling, D.R., Karl, T.R., Groismanm, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullet, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E.J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., and Parker, D., 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. Int. J. Climatol., 18:1493-1517.

- Picard, F., Lebarbier, E., Hoebeke, M., Rigaill, G., Thiam, B., and Robin, S., 2011: Joint segmentation, calling, and normalization of multiple CGH profiles. Biostatistics, 12:413-428.

- Poncelet, L. and Martin, H., 1947: Esquisse Climatologique de la Belgique. Mémoires Volume : XXVII. Institut Royal Météorologique de Belgique, 265 pp, Brussels, Belgium.

- Potter, K.W., 1981: Illustration of a new test for detecting a shift in mean in precipitation series. Mon. Wea. Rev., 109:2040-2045.

- Sciuto, S., Bonaccorso, B., Cancelliere, A, and Rossi, G., 2009: Quality control of daily rainfall data with neural networks. Journal of Hydrology, 364:13-22.

- Sneyers, R., 1964: La statistique des précipitations à Bruxelles-Uccle. Institut Royal Météorologique de Belgique, Contributions, N° 94, 21 pp.

- Sneyers, R., 1968 : La comparaison de pluviomètres de Bruxelles-Uccle. Publication N°78 de l'Association Internationale d'Hydrologie Scientifique, p. 265-273.

- Szentimrey, T, 1999: Multiple Analyses of Series for Homogenization (MASH). In Proceedings of the second seminar for homogenization of surface climatological data, Budapest, Hungary, WMO, WCDMP-No.41:27-46.

- Toreti, A., Kuglitsch, F.G., Xoplaki, E., Luterbacher, Juerg, Wanner, H., 2010: A novel method for the homogenization of daily temperature series and its relevance for climate change analysis. J. Climate, 23:5325-5331.

- Toreti, A., Kuglitsch, F.G., Xoplaki, E. and Luterbacher, J., 2012: A novel approach for the detection of inhomogeneities affecting climate time series. Journal of Applied Meteorology and Climatology, 51:317-326.

- Trewin, B., 2010: Exposure, instrumentation, and observing practice effects on land temperature measurements. Wiley Interdisciplinary Reviews: Climate Change, 1:490-506.

- Trewin, B., 2013: A daily homogenized temperature data set for Australia. International Journal of Climatology, 33:1510-1529.

- Trewin, B.C., and Trevitt, A.C.F., 1996: The development of composite temperature records. Int. J. Climatol., 16:1227-1242.

- Venema, V., Mestre, O., Aguilar, E., Aeur, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Mueller-Westermeier, G., G. Lakatos, M., Williams, C.N., Menne, M., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andersen, L., Acquaotta, F., Fratianni, S., S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Duran, M.P., Likso, T., Esteban, P., and Brandsma, T., 2012: Benchmarking monthly homogenization algorithms. Climate of the Past, 8:89-115.

- Vincent, J., 1912: Annuaire Météorologique pour 1913. Observatoire Royal de Belgique, Brussels, Belgium.

- Vincent, L.A., Zhang, X., Bonsal, B.R., and Hogg, W.D., 2002: Homogenization of daily temperatures over Canada. J. Climate, 15:1322-1334.

- Wang, X.L., 2009: A quantile matching adjustment algorithm for Gaussian data series. Climate Research Division, Science and Technology Branch, Environment Canada, Toronto, Canada, 5 pp.

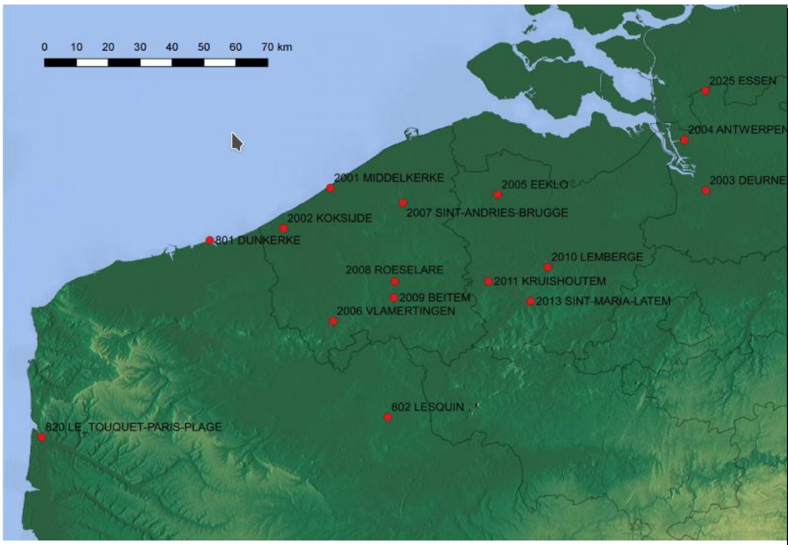
- Wang, X.L., Chen, H., Wu, Y., Feng, Y., Pu, Q., 2010: New techniques for the detection and adjustment of shifts in daily precipitation data series. J. Appl. Meteor. Climatol., 49:2416-2436.
- Wang, X.L., Feng, Y. and Vincent, L.A., 2014: Observed changes in one-in-20 year extremes of Canadian surface air temperatures. Atmopshere-Ocean, 52(3):222-231.
- Wijngaard, J.B., Tank, A., Konnen, G.P., 2003: Homogeneity of 20th century European daily temperature and precipitation series. Int. J. Climatol., 23:679-692.
- Xu, W., LI, Q., Wang, X.L., Yang, S., Cao, L. and Feng, Y., 2013: Homogenization of Chinese daily surface air temperatures and analysis of trends in the extreme temperature indices. Journal of Geophysical Research: Atmosphere, 118(17):9708-9720.
- Yosef, Y., Aguilar, E. and Pinhas, A.: 2018: Detecting and adjusting artificial biases of long term temperature records in Israel. International Journal of Climatology, 38:3273-3289.

ANNEXES

A.1 HOMER Temperature Clusters Constitution

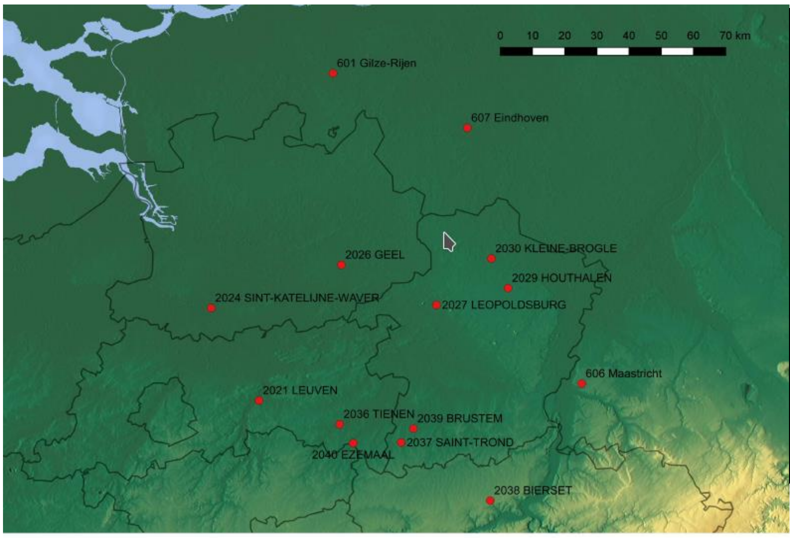
A.1.1 Long Series (LS)

A.1.1.1 Cluster 1



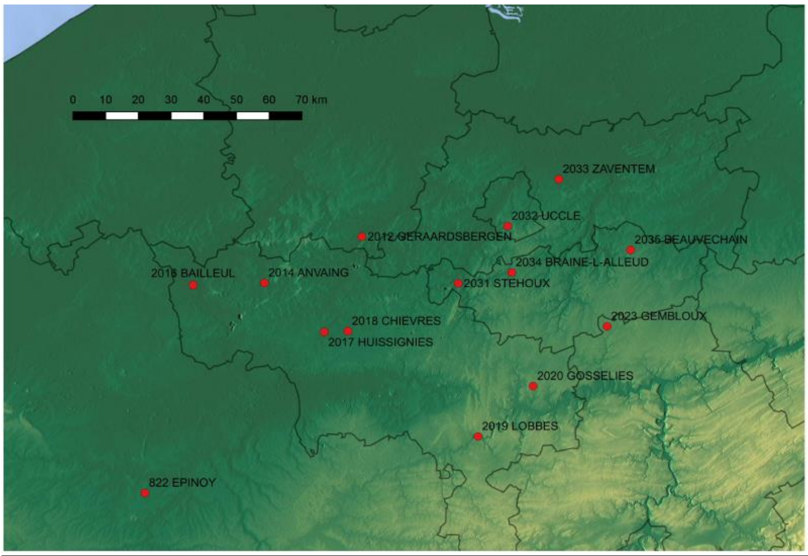
LS CODE	LS NAME
801	Dunkerke
802	Lesquin
820	Le Touquet Paris-Plage
2001	Middelkerke
2002	Koksijde
2003	Deurne
2004	Antwerpen
2005	Eeklo
2006	Vlamertingen
2007	Sint-Andries-Brugge
2008	Roesselare
2009	Beitem
2010	Lemberge
2011	Kruishouten
2013	Sint-Maria-Latem
2025	Essen

A.1.1.2 Cluster 2



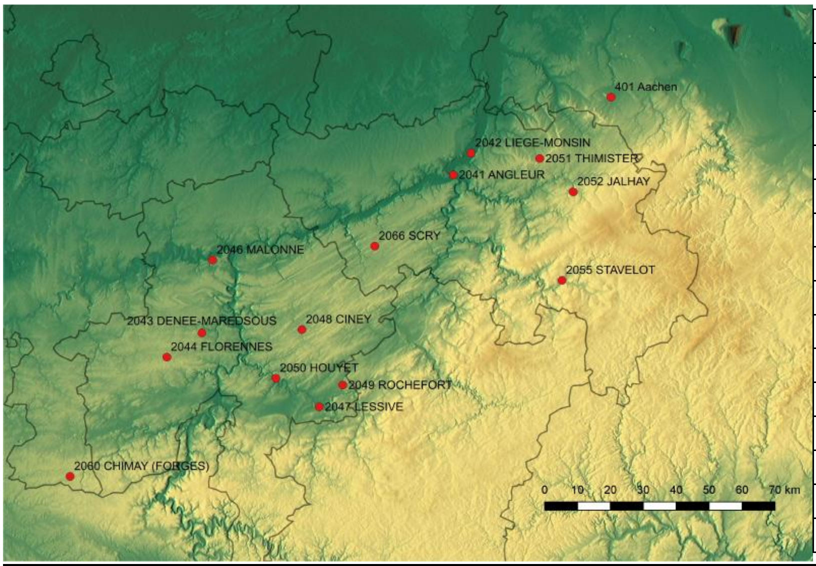
LS CODE	LS NAME
601	Gilze-Rijen
602	Eindhoven
2021	Leuven
2024	Sint-Katelijne-Waver
2026	Geel
2027	Leopoldsburg
2029	Houthalen
2030	Kleine-Brogel
2036	Tienen
2037	Saint-Trond
2038	Bierset
2039	Brustem
2040	Ezemaal

A.1.1.3 Cluster 3



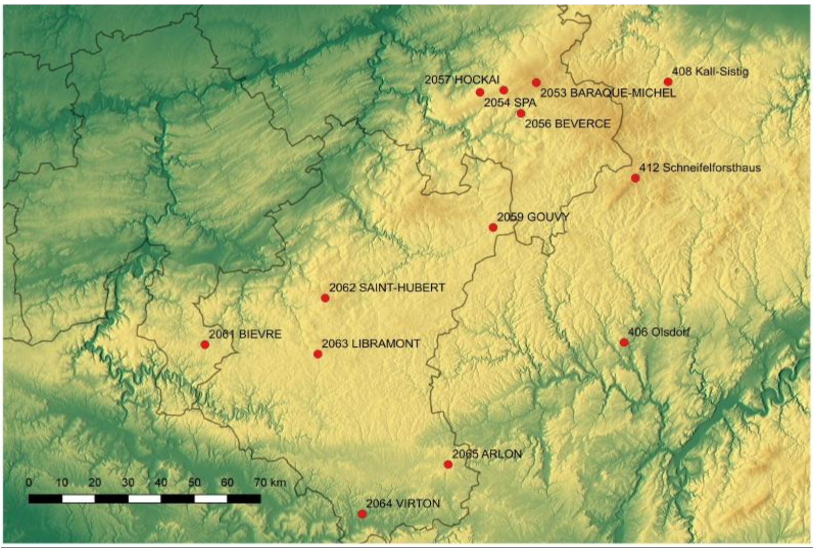
LS CODE	LS NAME
817	Fontaine-les-Clercs
822	Epinoy
2012	Geraardsbergen
2014	Anvaing
2016	Bailleul
2017	Huissignies
2018	Chievres
2019	Lobbes
2020	Gosselies
2023	Gembloux
2031	Stehoux
2032	Uccle
2033	Zaventem
2034	Braine l'Alleud
2035	Beauvechain

A.1.1.4 Cluster 4



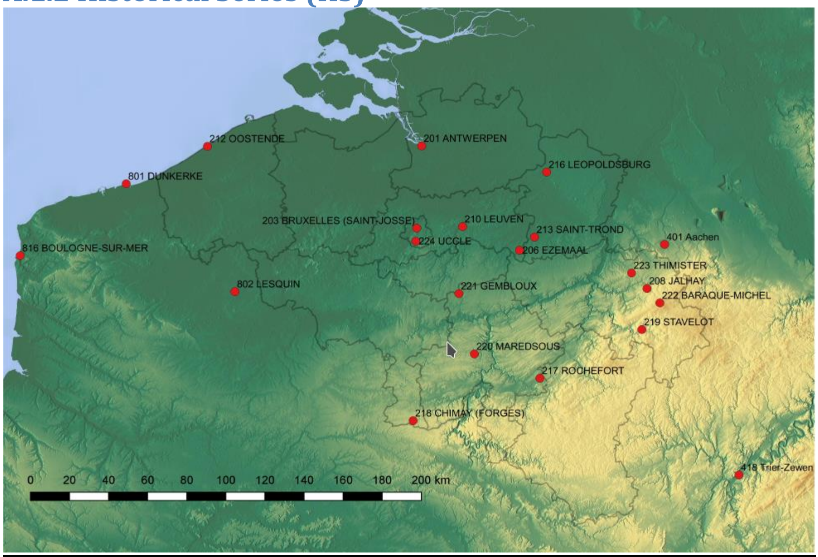
LS CODE	LS NAME
401	Aachen
2041	Angleur
2042	Liège-Monsin
2043	Denée-Maredsous
2044	Florennes
2046	Malonne
2047	Lessive
2048	Ciney
2049	Rochefort
2050	Houyet
2051	Thimister
2052	Jalhay
2055	Stavelot
2060	Chimay-Forges
2066	Scry

A.1.1.5 Cluster 5



LS CODE	LS NAME
406	Olsdorf
408	Kall-Sistig
412	Schneifelforsthaus
2053	Baraque-Michel
2054	Spa
2056	Beverce
2057	Hockai
2059	Gouvy
2061	Bièvre
2062	Saint-Hubert
2063	Libramont
2064	Virton
2065	Arlon

A.1.2 Historical Series (HS)

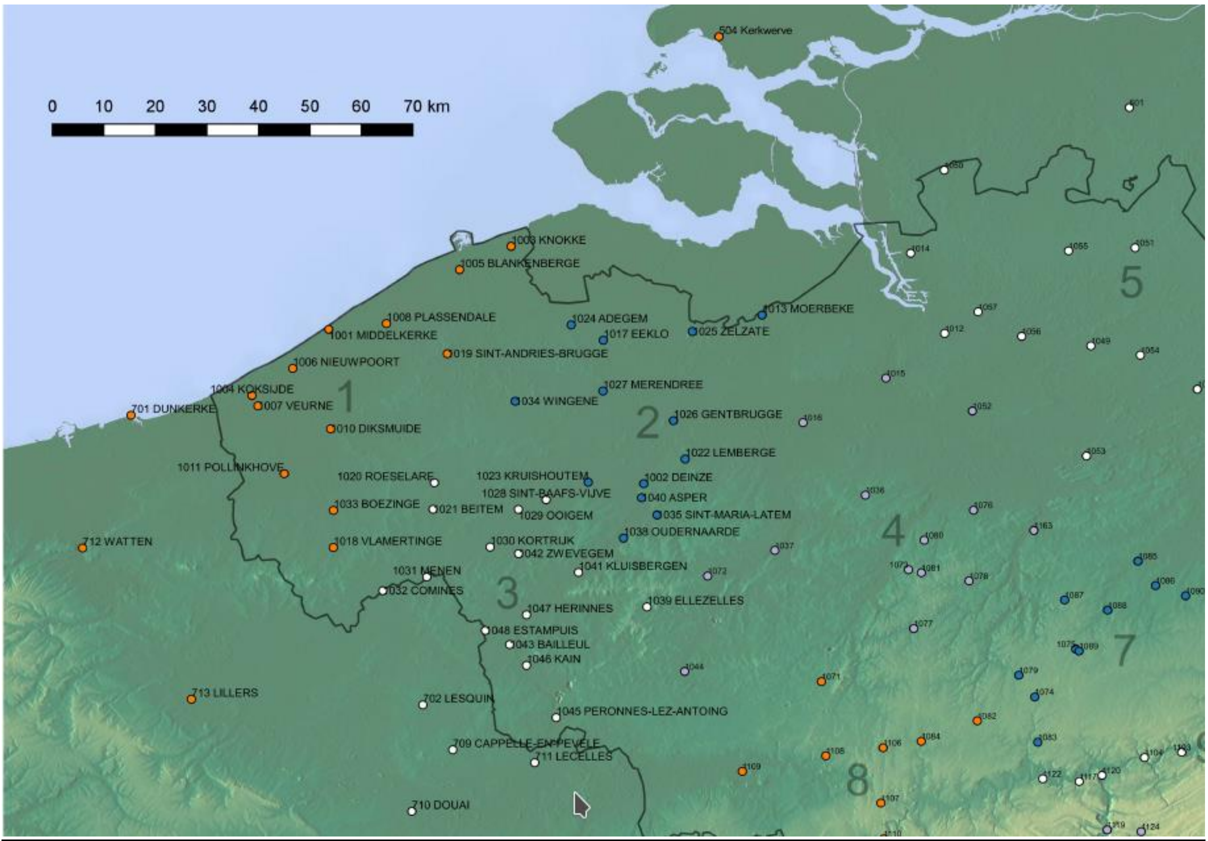


HS CODE	HS NAME
201	ANTWERPEN
203	BRUXELLES (SAINT-JOSSE)
206	EZEMAAL
208	JALHAY
210	LEUVEN
212	OOSTENDE
213	SAINT-TROND
216	LEOPOLSBURG
217	ROCHEFORT
218	CHIMAY-FORGES
219	STAVELOT
220	MAREDSOUS
221	GEMBLOUX
222	BARAQUE-MICHEL
223	THIMISTER
224	UCCLE
401	AACHEN
418	TRIER-ZEWEN
801	DUNKERKE
802	LESQUIN
816	BOULOGNE-SUR-MER

A.2 HOMER Precipitation Clusters Constitution

A.2.1 Long Series

A.2.1.1 Clusters 1, 2 and 3



CLUSTER 1 (orange)		*CLUSTER 2 (blue)*		*CLUSTER 3 (white)*	
LS CODE	**LS NAME**	**LS CODE**	**LS NAME**	**LS CODE**	**LS NAME**
504	Kerkwerve	1002	Deinze	702	Lesquin
701	Dunkerke	1013	Moerbeke	709	Cappelle-en-Pevele
712	Watten	1017	Eeklo	710	Douai
713	Lillers	1022	Lemberge	1020	Roeselare
1001	Middelkerke	1023	Kruishoutem	1021	Beitem
1003	Knokke	1024	Adegem	1028	Sint-Baafs_vijve
1004	Koksijde	1025	Zelzate	1029	Ooigem
1005	Blankenberge	1026	Gentbrugge	1030	Kortrijk
1006	Nieuwpoort	1027	Merendree	1031	Menen
1007	Veurne	1034	Wingene	1032	Comines
1008	Plassendale	1035	Sint-Maria-Latem	1039	Ellezelles
1010	Diksmuide	1038	Oudernaarde	1041	Kluisbergen
1011	Pollinkhove	1040	Asper	1042	Zwevegem
1018	Vlamertinge			1043	Bailleul
1019	Sint-Andries-Brugge			1045	Peronnes-lez-Antoing
1033	Boezinge			1046	Kain

				1047	Herinnes
				1048	Estampuis

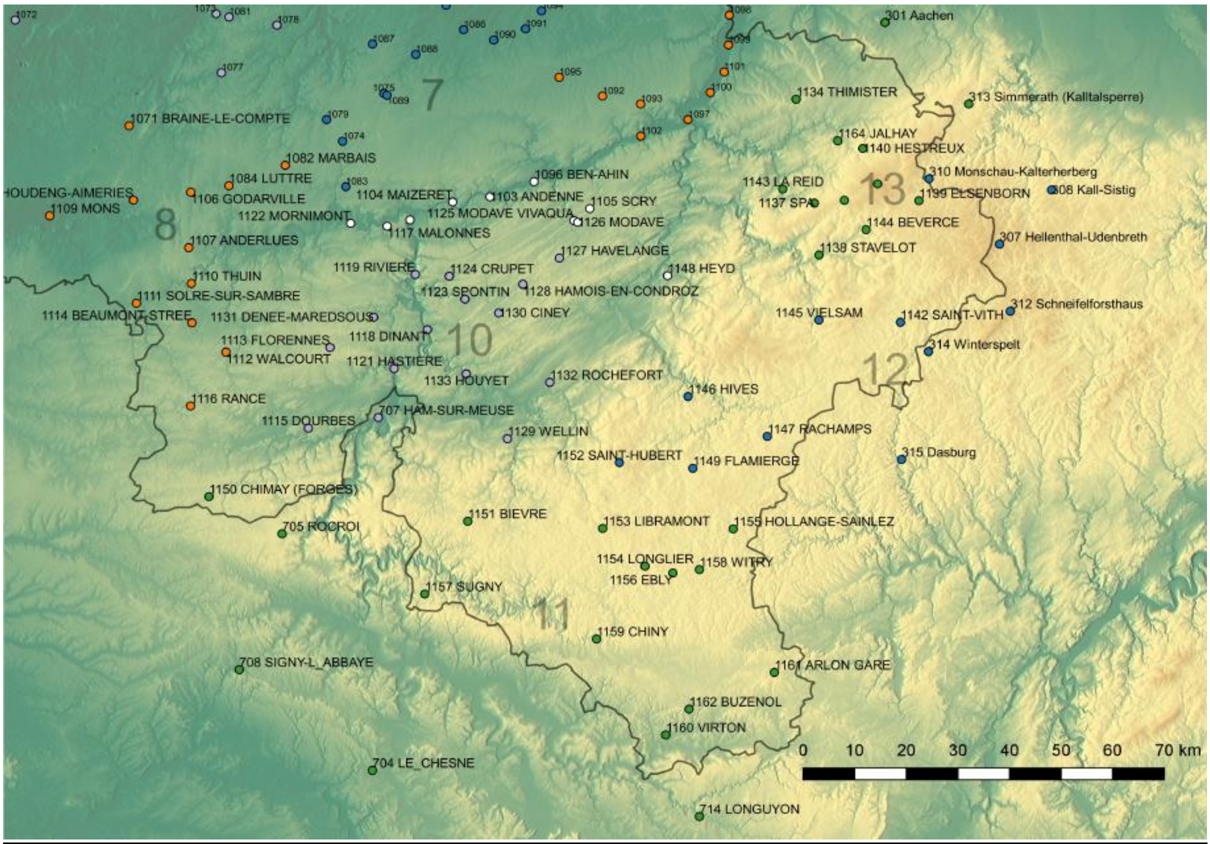
A.2.1.2 Clusters 4, 5, 6, 7, 8 and 9

CLUSTER 4 (grey)		*CLUSTER 5 (white)*		*CLUSTER 6 (red)*	
LS CODE	**LS NAME**	**LS CODE**	**LS NAME**	**LS CODE**	**LS NAME**
1015	Wintam	501	Gilze-Rijen	304	Selkant-Havert
1016	Dendermonde	1012	Deurne	1060	Meeuwen
1036	Brussegem	1014	Stabroek	1061	Riemst
1037	Pollare	1049	Herentals	1062	Lanaken
1044	Ath	1050	Essen	1063	Neeroeteren
1052	Sint-Katelijn-Waver	1053	Retie	1064	Maasmechelen
1072	Deftinge	1054	Aarschot	1065	Kessenich
1073	Uccle	1055	Geel	1066	Hasselt
1076	Zaventem	1056	Rijkevorsel	1069	Genk
1077	Braine l’Alleud	1057	Viersel	1092	Jeneffe
1078	Overijse	1058	Wijnegem	1093	Bierset
1080	Schaerbeek	1059	Leopoldsburg	1095	Waremme
1081	Uccle (Reservoir)	1067	Kleine-Brogel	1097	Angleur
1163	Leuven	1068	Lommel	1098	Lanaye
		1070	Kwaadmechelen	1099	Vise

				1100	Liege-Monsin
				1102	Ivoz-Ramet

CLUSTER 7 (blue)		*CLUSTER 8 (orange)*		*CLUSTER 9 (white)*	
LS CODE	**LS NAME**	**LS CODE**	**LS NAME**	**LS CODE**	**LS NAME**
1074	Gembloux	1071	Braîne-le-Comte	1096	Ben-Ahim
1075	Malèves-Sainte-Marie	1082	Marbais	1103	Andenne
1079	Blamnont	1106	Godarville	1105	Scry
1083	Mazy	1107	Anderlues	1117	Malonnes
1085	Tiennen	1108	Houdeng-Aimeries	1120	La Plante
1086	Ezemaal	1110	Thuin	1122	Mornimont
1087	Beauvechain	1111	Solre-sur-Sambre	1125	MODAVE (vivaqua)
1088	Jodoigne	1112	Walcourt	1126	Modave
1089	Thorembais-les-Beguines	1114	Beaumont-Stree		
1090	Landen	1116	Rance		
1091	Gorsem				
1094	Brustem				

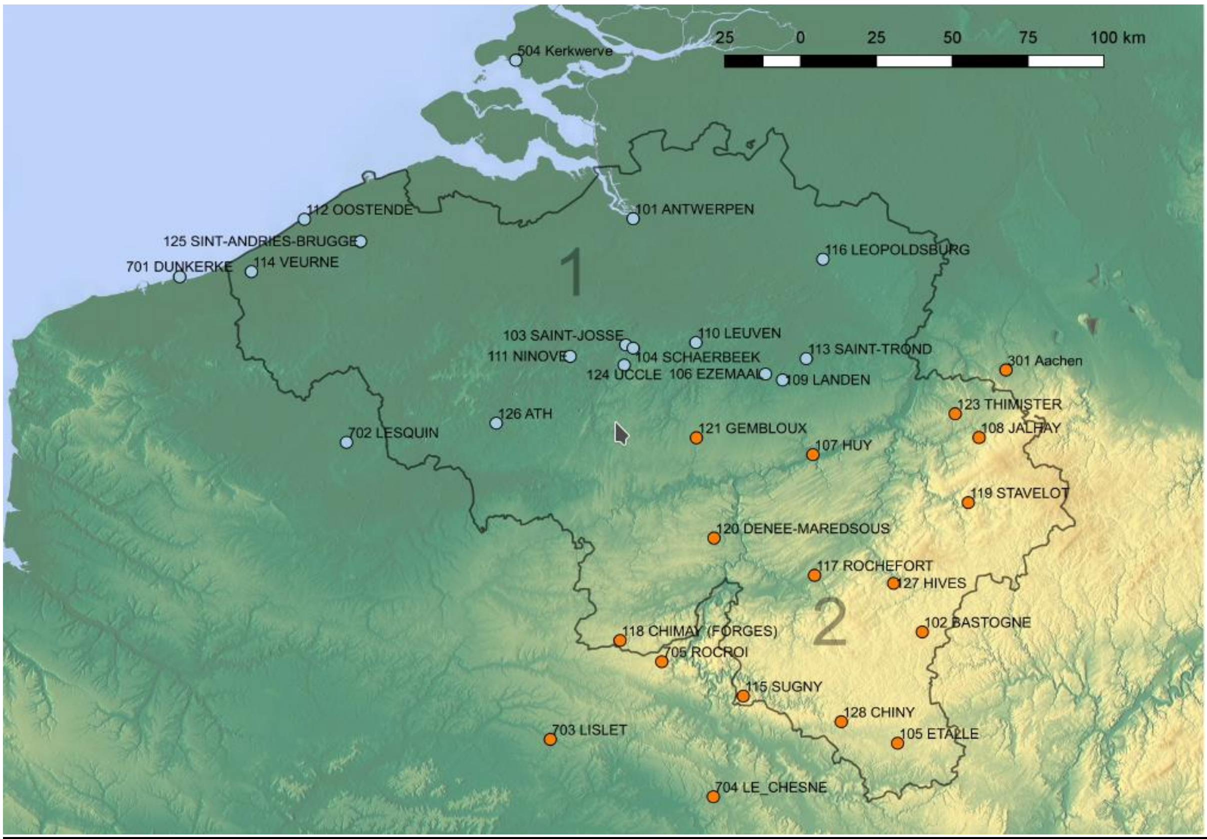
A.2.1.3 Clusters 10, 11, 12, and 13



CLUSTER 10 (grey)		*CLUSTERE 11 (green)*	
LS CODE	**LS NAME**	**LS CODE**	**LS NAME**
707	Ham-sur-Meuse	704	Le Chesne
1113	Florennes	705	Rocroi
1115	Dourbes	708	Signy l'Abbaye
1118	Dinant	714	Longuyon
1119	Rivière	715	Bras-sur-Meuse
1121	Hastière	1150	Chimay-Forges
1123	Spontin	1151	Bièvre
1124	Crupet	1153	Libramont
1127	Havelange	1154	Longlier
1128	Hamois-en-Condroz	1156	Ebly
1130	Ciney	1157	Sugny
1131	Denée-Maredsous	1158	Witry
1132	Rochefort	1160	Virton
1133	Houyet	1161	Arlon (gare)
		1162	Buzenol

CLUSTER 12 (blue)		*CLUSTER 13 (green)*	
LS CODE	**LS NAME**	**LS CODE**	**LS NAME**
307	Hellenthal-udenbreth	301	Aachen
308	Kall-Sistig	313	Simmerath (Kalltalsperre)
310	Monschau-Kalterherberg	1136	Mont-Rigi
312	Schneifelforsthaus	1137	Spa
314	Winterspelt	1138	Stavelot
315	Dasburg	1139	Elsenborn
1145	Vielsam	1140	Hestreux
1146	Hives	1141	Hockai
1147	Rachamps	1164	Jalhay
1152	Saint-Hubert		

A.2.2 Centennial Series



CLUSTER 1 (grey)		***CLUSTER 2 (orange)***	
CS CODE	**CS NAME**	**CS CODE**	**CS NAME**
101	ANTWERPEN	105	ETALLE
104	SCHAERBEEK	108	JALHAY
106	EZEMAAL	115	SUGNY
107	HUY	117	ROCHEFORT
109	LANDEN	118	CHIMAY-FORGES
110	LEUVEN	119	STAVELOT
111	NINOVE	120	DENEE-MAREDSOUS
112	OOSTENDE	121	GEMBLOUX
113	SAINT-TROND	123	THIMISTER
114	VEURNE	127	HIVES
116	LEOPOLDSBURG	301	AACHEN
124	UCCLE	703	LISLET
125	SINT-ANDRIES-BRUGGE	704	LE CHESNE
126	ATH	705	ROCROI
504	KERKWERVE		
702	LESQUIN		