Homogenization of Norwegian monthly precipitation series for the period 1961–2018

Elinah Khasandi Kuya, Herdis Motrøen Gjelten, and Ole Einar Tveito
Norwegian Meteorological Institute, Henrik Mohns Plass 1, 0371 Oslo, Norway

Correspondence: Elinah Khasandi Kuya (elinahkk@met.no)

Received: 15 February 2022 – Revised: 27 May 2022 – Accepted: 9 June 2022 – Published: 27 June 2022

Abstract. The primary goal of the analysis was to establish a high-quality precipitation reference dataset, which is both consistent and homogeneous, for calculation of the new standard climate normals (1991–2020). Climatol homogenization method was applied to detect inhomogeneities in 325 Norwegian precipitation series, during the period 1961–2018. Results from homogeneity testing found inhomogeneities in 29% of the 325 series, however, only 25% were classified as inhomogeneous after conferring with metadata and therefore adjusted. Relocation of the precipitation gauge and automation were the main causes of all the inhomogeneities in the Norwegian series, explaining 71% and 12% respectively of all detected breaks. Results further showed benefits of incorporating metadata to the automatically detected inhomogeneities. Linear trend analysis showed increasing trends in the period 1961–2018 except in autumn where a decreasing trend was observed. The homogeneity analysis produced a 58-year long homogenous dataset for 325 monthly precipitation sums with regional temporal variability and spatial coherence that is better than that of non-homogenized series. The dataset is more reliable in explaining the large-scale climate variations and was used to calculate the new climate normals in Norway.

1 Introduction

Climate normals play an important role in weather and climate studies and require high quality data that are both consistent and homogenous. Long-term climatological time series is regarded as homogeneous if the measurements have been consistently done using the same practices, with the same undamaged instruments at the same place and time and in the same environment. In this way the variations in the time series are only a result of variations in weather and climate. But the reality is that many climatic observations have been altered by a variety of external changes such as changes in instruments, observers, observation methods and practices, the station’s geographical location and in the environment surrounding the station. Such changes can introduce sudden shifts (homogeneity breaks) to the time series, others, e.g., environmental changes, lead to gradual biases over time. This can mask the genuine climatic variations and can lead to erroneous interpretations about the evolution of climate (Peterson et al., 1998). Since these shifts and biases are often of the similar magnitude as the climate signal (Menne and Williams, 2009), analyzing and correcting for such external influences to achieve a homogeneous climatic time series is therefore necessary before making a climate assessment.

Several applications have been developed to detect inhomogeneities in climate series (WMO, 2020) including the new emerging techniques HOMER (Mestre et al., 2013), ACMANT (Domonkos, 2015; Domonkos and Coll, 2017), AHOPS (Rustemeier et al., 2017) and Climatol (Guijarro, 2018). The homogeneity approaches often benefit from consulting metadata to validate the breaks and possible outliers.

Homogenization of long-term precipitation series for mainland Norway has previously been done by Hanssen-Bauer and Førland (1994) using the Standard Normal Homogeneity Test (SNHT) (Alexandersson, 1986). The study analyzed 165 Norwegian precipitation series whereby 70% of the series were found to be inhomogeneous. The most recent homogenization analysis on precipitation was done within the MIST-2 project, a collaboration between the Norwegian Meteorological Institute (MET Norway) and Statkraft SF (Lundstad, 2016) to homogenize daily precipitation series in Norway. Five precipitation series representing different energy consumption regions in Norway were successfully
The overarching goal of this study therefore is to perform a homogeneity analysis of Norway’s monthly precipitation series in order to establish a high-quality precipitation reference dataset for calculating the Norwegian Meteorological Institute (MET Norway) new standard climate normal for the period 1991–2020.

2 Data and methods

2.1 Precipitation dataset

370 series of observed monthly precipitation sums were used in the homogenization analysis (Fig. 1). The raw precipitation time series included 325 series from mainland Norway, 44 Swedish series and one Finnish series. The Swedish and Finnish series were included in the analysis only as reference series to support the homogenization of the Norwegian precipitation series. The series were extracted from ClimNorm (a network activity under the Nordic Framework for Climate services) precipitation dataset (Tveito et al., 2020). Since the goal was to have a homogeneous data set for the period of new normals, i.e., 1991–2020, the analyzed period was extended back in time because longer time series gives a more robust result when homogenizing. The homogeneity analysis was performed in 2019. For this reason, the analysis covered a period from 1961–2018. When calculating the new normals, data from 2019 and 2020 were added to the homogenized dataset. While 548 series (including 50 from Sweden and two from Finland) were available during the study period, the 370 series upheld the criteria of 80% data coverage (no more than ten missing years). To compute the climate normals such that it is representative of the 1991–2000 time period and the country, stations with short series were reconstructed. 33% of the 370 series were merged and were reconstructed from up to four original series. The criteria for merging time series were based on a maximum horizontal distance of 10 km, and after the first automatic screening all merged series were manually checked (Tveito et al., 2020).

2.2 Climatol homogenization approach and application

Climatol (Guijarro, 2018, 2019) was used to detect inhomogeneities in Norway’s monthly precipitation series. The method provides automatic quality control (outlier correction), homogenization (break detection and correction) and missing data attribution. In addition, Climatol is excellent for the detection and adjustment of inhomogeneities of large and dense networks like in this study and has the option to incorporate metadata information. Version 3.1.1 of the Climatol R package was used in this analysis. Climatol is highly dependent on difference series between a candidate series and its reference series and is based on the Standard Normal Homogenization Test (SNHT; Alexandersson and Moberg, 1997) for break detection. The difference series is estimated as a weighted average of the nearest and best correlated reference series. Correlation between the series is fundamental, because the higher the correlation, the higher the reliability of the homogenization and in-filling of missing data.

Climatol works with three types of data normalization applying the most appropriate ratios or differences depending on the climatological variable. Here normalizations were achieved by using the normal ratio to the mean method (dividing by the mean climatological values of the series). The process begins by normalizing all the original data and computing a reference series for each candidate series by averaging up to ten data series (if available) at every time step in the detection stages and four data series during the final series reconstruction. While these numbers of reference series can be changed by the user, the default ten series was used in this analysis. In this study, Climatol’s homogenization approach, including the three-step homogenization stages described in Kuya et al. (2021) was adopted.

The results from exploratory analysis of the original data were used to determine the most suitable detection threshold values in this analysis. Diagnostic outputs were also used for quality control and to verify that there were no abnormal features in the data. The default threshold for outliers proved to be too low, as most outliers turned out to be le-
Criteria for inhomogeneities

The criteria for accepting detected breaks were quite strict. For the inhomogeneity to be confirmed, Climatol’s detected breakpoint had to be within five years of the metadata explaining the break. Breaks without metadata were accepted only if significantly high SNHT values were realized. Detected breaks within the first and last five years of the series were rejected (exceptions were however made for inhomogeneities that could be explained with obvious reasons and were four years from the end of the series. Such breaks should be re-evaluated again later when there are more data points on each side of the breaks). Close breaks, especially within five years of each other, were assessed carefully, and weight was placed on the break point with metadata (or the most obvious or dominant reason for a break). In addition, outlier breaks were rejected.

3 Results

Results for homogeneity testing found 25% of the Norwegian precipitation series as inhomogeneous. Climatol detected 121 breaks in the Norwegian precipitation series, but only 90 were accepted and homogenized after conferring with metadata. 16 of the Swedish precipitation series were classified as inhomogeneous and all breaks were accepted since metadata for these series were not available.

The main reason for inhomogeneities was relocation, explaining 71% of the breaks in the Norwegian precipitation series. Figure 2a summarizes the causes for inhomogeneities in the Norwegian series. Figure 2b presents the annual adjustment factors (AFs) for all detected inhomogeneities in the Norwegian precipitation series, sorted by its causes. The distribution of AFs was quite symmetrical, and values ranged between 0.72 and 1.46. The different causes for breaks showed somewhat different patterns.

3.1 Analysis of inhomogeneities

3.1.1 Relocation

Relocation may cause breaks due to differences in either exposure or precipitation conditions between the two sites. The AFs were quite symmetrical around 1 with the mean value 1.008 which showed that there was no systematic tendency in the relocation AFs. Some large relocations of several kilometers led to very minor changes in the precipitation (break with just 1% adjustment, or no break at all) while other small relocations of only a few meters led to major changes in precipitation with adjustments of 10% or more. And vice versa. This illustrates that significant changes can be expected regardless of how far the precipitation gauge is moved, it all depends on the local conditions and environment at the site.

3.1.2 Automation

Station automation caused adjustments of the precipitation series with a mean value of 0.93, hence a net decrease in measured precipitation amount (an average decrease of 7% in annual precipitation). This is in accordance with previous parallel measurements between manual and automatic gauges conducted by MET Norway, where the general tendency is that the measured precipitation amount is lower with automatic gauges on mainland Norway (e.g., Nygård, 2004). But this will of course vary depending on wind exposure, season, and amount of precipitation as snow in addition to whether the manual gauge was equipped with a windscreen. In very wind-exposed areas where a larger fraction of precipitation falls as snow, the automatic gauge may catch more precipitation than the manual gauge. It is important to note that in most cases the automation was coupled with a small or large relocation at the same time, and so the different precipitation conditions and wind exposure would also affect the AFs.
Figure 2. (a) Reasons for inhomogeneities in the Norwegian precipitation series. They include relocation of the rain gauge, automation of the station, environmental changes, change of observer, and other unknown reasons. (b) Annual adjustment factors for all the detected inhomogeneities plotted against time grouped by their causes.

3.1.3 Environmental change
Changes in the local site cause a net 6% increase in the annual precipitation after the change, which implies that most of the changes in the environment increased sheltering of the gauge. Two of the breaks with environmental change as the main reason resulted in a lower gauge catch, while four breaks gave an increased gauge catch.

3.1.4 New observer
Changes in observers should ideally not influence precipitation measurements, but it may in some cases still cause changes in observation practices. The four accepted breaks with change of observer as the main reason all gave adjustments in the same direction, with an average of 11% decrease in annual precipitation after the change.

3.2 Effects of using metadata
In addition to examining all detected breaks against available metadata, a comparison was made between the raw series, Climatol’s automatically generated homogenized series and the homogenized series where metadata were used to verify breaks to assess the benefit of using metadata in the homogenization of Norway’s precipitation series. The analysis was done only on the series with detected breaks. The two homogenized series (with and without metadata) had comparable linear trends: 2.864% and 2.865% per decade, and thus the difference induced by use of metadata was rather insignificant on a regional scale. The non-homogenized series exhibited a 1.8% linear trend per decade. The difference was however large when analyzing single series. Results (not shown) of filtered values of the raw and homogenized series for an individual station showed that, while both homogenized series exhibited an increasing trend with time, in contrast to the raw series which has a decreasing trend, the temporal pattern in the two homogenized series were dissimilar. In this case the automatic adjustment may have masked some of the natural variability in the series when adjusting outlier breaks that were rejected when using metadata. This further proved the benefits of using metadata, as there can be pronounced differences in results when using metadata to accept and reject breaks.

3.3 Impacts of homogenization
Results presented show the impact and efficiency of homogenization on the Norwegian precipitation series. Comparison was made between the non-homogenized and the homogenized series. The series were converted into standardized anomalies using 1961–1990 as the reference period. The average of the annual and seasonal precipitation sums for all 325 Norwegian stations for both the non-homogenized and homogenized data was evaluated.

The results of mean annual total precipitation for the whole Norwegian network did not show too large differences between non-homogenized and homogenized series. This was however not surprising as only about 25% of series analyzed were adjusted. Precipitation is a variable with large spatial and temporal variance, which makes break detection harder than for other variables such as temperature. Nevertheless, some differences were seen where homogenization led to higher deviation from the 1961–1990 average. Values of annual anomalies in the non-homogenized series ranged from $-23\%$ to $27\%$, while those of homogenized series ranged from $-22\%$ to $32\%$.

Both non-homogenized and homogenized time series for the annual and seasonal series showed increasing trends in the period 1961–2018 except in autumn where a decreasing trend is observed (Table 1). The 1961–1990 period was relatively wet in autumn compared to periods before and after. The homogenized series showed a higher trend compared to the non-homogenized series. 56% of all the adjusted inhomogeneities had AFs below 1. As a result, the net effect of the
adjustments led to reduced precipitation values, especially in the earlier parts of the series, causing a more pronounced increasing trend in the homogenized series. The greatest difference in trend size before and after homogenization was observed in the annual series, where linear trends changed by 0.8 % per decade. Most of the increase in the precipitation series has been during the last 30 years.

To further understand the impact of homogenization on the precipitation series, the same calculations highlighted above were applied on some individual series representing different precipitation regions in Norway (Fig. 3). The range of anomalies with respect to the 1961–1990 reference period of the non-homogenized series is clearly wider than those of the homogenized series. Most series also show similar variability with time after homogenization. These results show that after homogenization, the spatial coherence of the temporal evolution of the series is better compared with the non-homogenized series. This proves that homogenization contributes to better temporal and spatial coherence of time series, while preserving the general statistical distribution of the raw time series.

Another comparison between the non-homogenized and homogenized series is presented in Fig. 4 where the difference between the two normal periods 1961–1990 and 1991–2020 is shown. The homogenized series exhibited smoother spatial patterns, meaning that the regional climate signal was well represented, with less local variations than in the maps based on the raw series. This is especially evident in the annual series.

In some regions the ratio between the two normal periods changed from above 100 % to below after homogenization. The greatest variation in the homogenized series was in the south-eastern region of Troms and Finnmark in Northern Norway where the highest decrease in both annual and seasonal precipitation was observed. From the map of annual series for example, the annual precipitation changed from over 140 % of the 1961–1990 mean before homogenization to under 100 % of the 1961–1990 mean after homogenization. This was also observed in the seasonal series. The results here are not surprising as most adjusted inhomogeneities in this region had large AFs above 1.2, hence the net effect of the adjustments led to increased precipitation values in the earlier parts of the series.

In conclusion, results based on the map after homogenization generally showed better regional coherence in comparison to the non-homogenized series. The homogeneous data are therefore more dependable in explaining the large-scale climate variations that should explain the change of “normal” climatologies.
Figure 4. 1991–2020 normal values in percent of the 1961–1990 normal for raw and homogenized annual and seasonal series.
4 Conclusions

The homogeneity analysis produced a 58-year long homogenous dataset for 325 monthly precipitation series, during the period 1961–2018. 370 series (including 44 from Sweden and one from Finland) from the ClimNorm precipitation dataset were used in the homogenization analysis. Application of Climatol homogenization method detected inhomogeneities in 29 % of the Norwegian precipitation series. A strict set of guidelines for inhomogeneity testing was followed to avoid over-adjusting of the dataset while still detecting as many real breaks as possible. Therefore, only 25 % of the series were classified as inhomogeneous after confirming with metadata and therefore adjusted. Precipitation is a variable with large spatial and temporal variance, which makes break detection harder than for other variables such as temperature. Breaks must be of a certain size to be noticeable over the general noise level, and thus many changes may not lead to detectable breaks.

The results of homogeneity testing showed that relocation of the precipitation gauge and automation were the main causes of inhomogeneities in the Norwegian series, explaining 71 % and 12 % respectively of all detected breaks. All but one of the accepted inhomogeneities could be confirmed with metadata. The homogenized series showed increasing trends in the period 1961–2018 except in autumn where a decreasing trend was observed. The highest trend was observed in the winter series while the greatest difference in trend size before and after homogenization was observed in the annual series. The most noticeable changes before and after homogenization were seen in the south-eastern region of Troms and Finnmark in Northern Norway, where the highest decrease in both annual and seasonal precipitation was observed. Nevertheless, results after homogenization showed better regional coherence in comparison to the non-homogenized series.

In general, homogenization greatly improved the quality of precipitation series by reducing the regional variability and improving both the temporal and spatial coherence of the dataset. The dataset was thus more reliable in explaining the large-scale climate variations and was used to calculate the new climate normals in Norway.

Code availability. The homogenization software Climatol is openly accessible at https://cran.r-project.org/web/packages/climatol/index.html (CRAN, 2022). Information about other codes can be obtained from the corresponding author upon request.

Data availability. All unhomogenized precipitation series can be accessed from MET Norway’s archive of historical weather and climate data (https://seklima.met.no/; MET Norway, 2022a). The new normal precipitation series can be accessed through https://frost.met.no/ (MET Norway, 2022b). The homogenized precipitation data set will be accessible in MET Norway’s archive at some point but can for now be accessed from the corresponding author upon request. The unhomogenized precipitation series from the Swedish Meteorological and Hydrological Institute and The Finnish Meteorological Institute can be accessed from the corresponding author upon request.

Author contributions. EKK wrote the manuscript with contributions from all authors. OET was responsible for providing the data series of the analyzed stations while EKK and HMG performed the homogeneity testing and analysis. OET calculated the new precipitation normals.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Special issue statement. This article is part of the special issue “21st EMS Annual Meeting – virtual: European Conference for Applied Meteorology and Climatology 2021”.

Review statement. This paper was edited by Dan Hollis and reviewed by Athanassios Argiriou and one anonymous referee.

References

CRAN: climatol: Climate Tools (Series Homogenization and Derived Products), CRAN [code], https://cran.r-project.org/web/packages/climatol/index.html, last access: 24 June 2022.
Kuya, E. K., Gjelten, H. M., and Tveito, O. E.: Homogenization of Norwegian monthly precipitation series for the period


