In this study, we derive new time-series of monthly-mean surface air temperature for Switzerland that range back to 1864 and represent area-mean conditions over the country and three major sub-regions. The methodology integrates data from a small sample (19 stations) of homogenized long-term series and from a high-resolution (2 km) grid dataset over a short (20 years) period. The statistical combination defines an objective weighting of station data that delivers reliable and time-consistent area-mean estimates, despite coarse and biased coverage with stations in early years. The methodology also quantifies the uncertainty of the estimates. Validation of the method reveals plausible patterns of station weights, and estimation errors of about 0.1 °C, much smaller than inter-annual variations. The new series suggest a warming in Switzerland of almost 1.5 °C from the early-industrial period (1864–1900) till the latest WMO standard period (1981–2010), with a linear trend of 1.29 °C per 100 years between 1864 and 2016. The warming is found to be larger in autumn than in other seasons, larger to the north of the Alps than to the south, and larger below (above) 1000 m asl in winter (summer). In all series, the warming is modulated by inter-decadal variations. Current global temperature datasets exhibit less warming for Switzerland than the present analysis. The pattern of disagreement suggests that a network-wide change in Swiss temperature measurements around 1980 may have been missed in the homogeneity adjustments at global data archives. It is desirable that these archives are better aligned with the latest quality processing of the original data owners.

**KEYWORDS**

area-mean temperature, climate monitoring, global temperature datasets, long-term climate variations, spatial climate analysis, temperature trend

---

**1 | INTRODUCTION**

Surface air temperature is one of the most important and certainly the most widely used metric to measure climate variations and change from regional to global scales. It is a tangible characteristic of the climate system, with comprehensible effects on the environment. This, and the availability of long instrumental series, distinguish surface temperature as a key variable for monitoring and communicating recent changes in climate (e.g., Hartmann et al., 2013). In research, long-term analyses of surface temperature are important for the reconstruction of climate by proxy data (e.g., Jones et al., 2009), for climate change detection and attribution (e.g., Jones et al., 2013), for the evaluation of climate models (Flato et al., 2013).
and to constrain model-based climate change projections (e.g., Knutti and Tomassini, 2008).

Currently available global surface temperature analyses are derived from station measurements collated in global data archives, using statistical spatial analysis (e.g., Hansen et al., 2010; Morice et al., 2012) or model-based data assimilation reanalysis, (e.g., Compo et al., 2011; Dee et al., 2011). The accuracy of such large-scale analyses for a certain region, such as for the territory of Switzerland, is unclear. The station series available for that region in the global data archives may be limited in number, temporal extent, the data may not correspond to the latest quality processing, and the spatial resolution of global analyses may be limited for the complexity of the region. In addition, homogenization of the underlying station series is challenging due to the use of automated methods and the limited access to potential reference stations and metadata. Increasing interest into regional patterns of temperature variations, therefore, calls for dedicated regional surface temperature analyses of high quality. These shall best possibly exploit available long-term instrumental series and integrate knowledge of the regional climate. In this paper, we describe the construction of such a series, representative for monthly-mean surface air temperature averaged over the territory of Switzerland. The series extends back to 1864.

So far, two main procedures have been followed in the construction of high-quality, long-term temperature series representative for a defined region. The first is based on station data directly. The regional temperature is defined either by choosing a particularly representative single station or by averaging data from several stations representing the defined region jointly. The underlying station sample can vary over time. The most prominent example of this category is the Central England Temperature (CET) series. Based on work of Manley (1974) an area-averaged temperature for Central England was published by Parker et al. (1992) and is known as the longest instrumental temperature series available; it dates back to 1659. Depending on the period, the CET is constructed from one to three stations and the variance of the mean was adjusted for the variable sampling. Temperature series for Central Europe (CEUT; Dobrovolný et al., 2010), Central Netherlands (CNT; van der Schrier et al., 2011) and West-Japan (WJT; Zaïki et al., 2006) are constructed in a similar way. Correlation maps with continental or global gridded datasets (e.g., Brohan et al., 2006; Haylock et al., 2008) can be used to define the spatial representativity of the regional series a posteriori (Dobrovolný et al., 2010; van der Schrier et al., 2011). With this approach, regional series are comparatively easy to establish and update, once the laborious quality testing and homogenization of the input station records is accomplished. However, the spatial representativity is somehow undefined. In regions with a spatially highly variable climate, the available stations may also provide a biased representation of the conditions only and, hence, simple averaging may be inaccurate to obtain a region average.

The second approach estimates regional temperatures by averaging grid point values from a prior spatial analysis (interpolation) of available station measurements. The familiar global, hemispheric and land-area temperature series are derived from global gridded analyses (e.g., Hansen et al., 2010; Morice et al., 2012; Harris et al., 2014). van der Schrier et al. (2013) determine a pan-European temperature series dating back to 1950 by averaging over gridpoints of the E-OBS dataset (Haylock et al., 2008). Similarly, long climate series for sectors of the European Alps were derived from the pan-Alpine HISTALP grid dataset (Auer et al., 2007). Examples of national and subnational series are numerous (see e.g., Prior and Perry (2014) for the United Kingdom, Tietäväinen et al. (2010) for Finland, Zhang et al. (2000) for Canada and Kothawale et al. (2010) for India). An advantage of averaging over grid points is the explicit account of the spatial distribution of the available stations in relation to the climatological and geophysical characteristics of the region. However, the reliability of a spatial analysis becomes limited for early periods with only few stations (e.g., Kaspar et al., 2013; Vincent et al., 2015) and the long-term consistency of a grid dataset can be compromised by variations in the station sample and density over time (e.g. Hofstra et al., 2010; Frei, 2014).

For its climate monitoring and communication, Meteo-Swiss has used, over the past 10 years, a series of Swiss temperature anomalies that was defined as an equally weighted mean of measurements at 12 stations, each representing one of the climate regions of the country (Schiüpp and Gensler, 1980). The 12 stations (see Figure 1) are part of the Swiss National Basic Climatological Network (Swiss NBCN; Begert et al., 2007), which is the reference station network for climate monitoring in Switzerland and offers homogenized long-term temperature series back till 1864 (Begert et al., 2005). However, the derived Swiss mean temperature was likely biased, considering that most of the 12 stations are located on the Swiss Plateau and at the bottom of mountain valleys. Elevation-dependent temperature signals may not be adequately represented.

This apprehension led to the re-definition of the Swiss temperature series presented here. The re-definition aims at preserving the high standards in long-term consistency and, hence, is still relying exclusively on data from high-quality homogenized series at a set of stations that remains constant over time. But the new definition also aims at better representing climate conditions that may be under-sampled in the available set of stations. This is accomplished by reference to a high-resolution temperature grid dataset over Switzerland, available for the most recent decades only. To this end, a statistical combination of station and grid data over the common period serves to derive coefficients for a linear combination of station values to best possibly represent “true” area-mean temperatures over the country. In summary, our approach combines ideas and advantages of the two commonly adopted methodological approaches mentioned above.
It is important to clarify what we understand from a country-wide average of surface air temperature. Unlike the naïve understanding of an average that would be obtained from thousands of evenly distributed thermometer measurements, the average is a rather hypothetical measure that could be obtained from those many measurements if WMO standard conditions (grass soil cover, no artificial heat source within 100 m, etc.; WMO, 2008.) would exist everywhere. The available measurements are purposely avoiding conditions over pavement, rock, glacier or lakes and, hence, any projection into an area average will not capture thermal effects in these environments, even though they are prevalent in the domain. This implies that the variations reproduced by the Swiss-mean air temperature represent macro-climatic conditions only. Variations related to, for example, urbanization or land use changes are largely excluded because of the rural and suburban location of Swiss NBCN stations, because of the standardization of the measurements and because of the homogenization of the time series. These considerations are essential for a professional interpretation of the area-mean series derived here.

Section 2 gives an overview of the data used in this study and the study domain. The statistical method is developed in section 3. Section 4 presents the resulting new area-mean temperature for Switzerland and discusses its evolution over time. The uncertainty of the estimates, and the validity of assumptions are assessed and the performance of the method is validated. We also demonstrate how the method can be used to calculate mean values for interesting sub-regions such as the north and south of Switzerland or high and low altitudes. Finally, also in section 4, we compare the temperature evolution for Switzerland as seen by prominent global grid datasets with the new Swiss series and pinpoint to some of their limitations. A summary and conclusions are provided in section 5.

2 | STUDY DOMAIN AND DATA

2.1 | Study domain

The territory of Switzerland covers an area of 41’285 km$^2$ and can be divided into four orographically distinct regions (Figure 1). The mountain ridge of the Alps extends from west to east over the southern portion of the country and covers about half of the terrain. Elevations range from less than 600 m asl at the floor of several broad valleys to more than 4000 m asl in major massifs. The Jura hill range extends along the north-western border of the country and reaches elevations above 1600 m asl. It encompasses several elevated plains and narrow valleys. Enclosed between the two mountain ranges is an extended south-west to north-east oriented basin, the Swiss Plateau, with gentle hills and generally northbound valleys at elevations of 350–600 m asl. The southern part of Switzerland (Ticino) reaches into the southern foothills of the Alps and encompasses several valleys running southward to the Po valley, a major basin below 200 m asl. On Swiss territory, the Alpine main crest runs along the southern border of the Alpine ridge and basically separates the Ticino and several southern valleys of Wallis and Graubünden from the rest of the country (see Figure 1).

Related to the complex topography, anomalies of monthly mean temperature from the long-term climate can exhibit remarkable and convoluted spatial variations in Switzerland. Patterns frequently observed are gradients across the Alpine main crest, contrasts between low and high elevations (seasonally varying mesoscale inversions over the Swiss Plateau) and differences between valleys and high elevations (see e.g. Wanner and Kunz, 1983; Schär et al., 1998; Frei, 2014; Scherrer and Appenzeller, 2014). Accurate sampling of these variations with a coarse station network is a key challenge for estimating a Swiss temperature series that reaches back into the 19th century.
2.2 Data

Monthly long-term temperature series of the Swiss National Basic Climatological Network (Swiss NBCN) and monthly temperature fields of Switzerland on a 2-km grid are the basic input data for this study. The Swiss NBCN (Begert et al., 2007) consists of 29 ground-based measuring stations with long-term measurements of different meteorological variables extending back as far as 1864, when the first nation-wide measurement network in Switzerland was founded. The NBCN stations are a subset of the much larger station network of MeteoSwiss (SwissMetNet; MeteoSwiss, 2010). They are characterized by particularly long series of measurements and exhibit a spatial distribution representative across Switzerland. The NBCN selection covers all the main climate regions of Switzerland. In an objective assessment, using cluster analysis based on correlation, Begert (2008) has shown that the NBCN stations represent the objectively determined climate regions for temperature with a correlation coefficient of at least 0.93, when related to all the SwissMetNet stations in the respective region. The long-term temperature series of the Swiss NBCN were carefully homogenized (Begert et al., 2003; Begert et al., 2005) for the effects of changes in instrumentation, automatization and site relocation. Rarely occurring missing months were interpolated from neighbouring stations using linear regression models. However, only 19 out of the 29 NBCN stations offer the full temperature data record since 1864. Because completeness is a prerequisite for the method proposed here (see section 3) this 19-member NBCN station set (Figure 1 and Table 1) forms the basis for the calculation of the new area-mean temperature of Switzerland.

MeteoSwiss uses measurements from the comprehensive SwissMetNet ground-based station network to derive spatial analyses of temperature across Switzerland on a grid with a spacing of $2 \times 2$ km (Frei, 2014). The statistical method employed addresses the challenges in complex topography by modelling nonlinearities in the vertical temperature profile and by incorporating terrain effects in the spatial representativity of measurements via non-Euclidean distances. In the present application, we use monthly fields from the period 1981–2014, when the coverage with station data is high. Standard errors in the monthly temperature estimates at grid points are between $0.5 \degree C$ (in summer over the Swiss Plateau) and $1.8 \degree C$ (in winter in the Alps). The systematic component of the error is very small and we expect that the mean over all grid points within the national borders reproduces the country-wide average temperature at an accuracy of at least $0.05 \degree C$.

The period with available grid data (1981–2014) is split into a 20-year calibration period (1985–2004) and 14 years for validation (1981–1984 and 2005–2014). The calibration period is used to set up the statistical model whereas the validating period is used to assess the performance of the method independently. The periods were chosen in this way to include the comparatively cold years at the beginning of the 1980’s as well as some warm recent years (see Figure 2) in the validation period.

### Table 1

| Name                  | Abbreviation | Longitude (deg) | Latitude (deg) | Height (m asl) |
|-----------------------|--------------|-----------------|----------------|---------------|
| Altdorf               | ALT          | 8.62            | 46.89          | 438           |
| Andermatt             | ANT          | 8.58            | 46.63          | 1438          |
| Basel/Binningen       | BAS          | 7.58            | 47.54          | 316           |
| Bern/Zollikofen       | BER          | 7.46            | 46.99          | 553           |
| Chauxmont             | CHM          | 6.98            | 47.05          | 1136          |
| Col du Grand St-Bernard| GSB          | 7.17            | 45.87          | 2472          |
| Davos                 | DAV          | 9.84            | 46.81          | 1594          |
| Engelberg             | ENG          | 8.41            | 46.82          | 1036          |
| Genève-Contrin        | GVE          | 6.13            | 46.25          | 411           |
| Grächen               | GRC          | 7.84            | 46.20          | 1605          |
| Lugano                | LUG          | 8.96            | 46.00          | 273           |
| Neuchâtel             | NEU          | 6.95            | 47.00          | 485           |
| Samedan               | SAM          | 9.88            | 46.53          | 1709          |
| Säntis                | SAE          | 9.34            | 47.25          | 2502          |
| S. Bernardino         | SBE          | 9.18            | 46.46          | 1639          |
| Segl-Maria            | SIA          | 9.76            | 46.43          | 1804          |
| Sion                  | SIO          | 7.33            | 46.22          | 482           |
| St. Gallen            | STG          | 9.40            | 47.43          | 776           |
| Zürich/Fluntern       | SMA          | 8.57            | 47.38          | 556           |
3 | METHOD

Our method to calculate a time series of monthly area-mean temperature for Switzerland, ranging back till 1864, builds on a linear combination of observations at the 19 long-term stations. The coefficients of the combination define weights that account for the spatial extent of the climatic regions represented by the stations. They are assumed to be constant over the entire period and show no dependency on the time of the year. Rather than prescribing the coefficients via heuristic considerations, we use the grid dataset, in fact the area average over all grid points, to estimate the coefficients empirically. To this end, a linear model is employed with the grid-average as predictor and station values as predictands. Once calibrated, the linear model is used to predict area-mean temperatures, including pertinent uncertainties, for all months since 1864. The key idea of this approach is to (a) determine the coefficients during a period when reliable estimates of the country-average temperature is available from many stations (the grid dataset), while (b) keeping the input data for the predictions (the long-term stations) constant, in order to avoid inhomogeneities as a result of variations in the station network. An attractive feature of the procedure is that uncertainties can be estimated within the model, which inform about the limited accuracy of area-mean estimates from the small set of stations.

Technically, the proposed method consists of the following steps: (a) The station data and country-mean temperatures from the grid dataset are converted into deviations with respect to the mean of the respective calendar month over 1981–2014. $d(t)$ is the vector of deviations for month $t$, one component for each station. $m(t)$ is the time series of country-mean anomalies. The normalization removes the annual cycle and our forthcoming model focuses on anomalies only. (b) A principal component analysis (PCA, e.g., Wilks, 2006) is conducted with the monthly anomalies $d(t)$ at the 19 long-term stations. This yields a sequence of transformed time series, the PCA scores $s_l(t)$, sorted by the fraction of total variance explained. Vectors $l_l$ denote the pertinent PCA loadings. (c) A linear regression model is then calibrated with the country mean anomaly as predictand and a subset of the PCA scores $s_l(t)$ as predictors:

$$m(t) = \beta_0 + \beta_1 \cdot s_1(t) + \beta_2 \cdot s_2(t) + \cdots + \beta_n \cdot s_n(t).$$  \hspace{1cm} (1)

$\beta_1$, ..., $\beta_n$ denote the linear regression coefficients. Working with PCA scores, instead of station series directly, allows us to constrain the number of predictors to a subset of leading linear modes (i.e. $n < 19$, see below), while still exploiting data from all stations. Also, choosing PCA scores reduces collinearity in the explanatory variables. (d) The linear regression coefficients estimated over the calibration period are finally used to make predictions of $m(t)$ for all months $t$ since 1864, via:

$$\hat{m}(t) = \hat{\beta}_0 + \hat{\beta}_1 \cdot (l_1 \cdot d(t)) + \hat{\beta}_2 \cdot (l_2 \cdot d(t)) + \cdots + \hat{\beta}_n \cdot (l_n \cdot d(t)).$$ \hspace{1cm} (2)

The scalar products $l_l \cdot d(t)$ are estimates of the PCA scores at time $t$. The equation can be rearranged to express the country mean in terms of station values:

$$\hat{m}(t) = w_1 \cdot d_1(t) + w_2 \cdot d_2(t) + \cdots + w_{19} \cdot d_{19}(t),$$ \hspace{1cm} (3)

where the station weights $w_k$ are obtained from the coefficients and PCA loadings via:

$$w_k = l_{k1} \cdot \hat{\beta}_1 + l_{k2} \cdot \hat{\beta}_2 + \cdots + l_{kn} \cdot \hat{\beta}_n.$$ \hspace{1cm} (4)

The number of samples available to calibrate the linear model (1) is limited (20 · 12 = 240) and the predictors (station series) are highly correlated, which points to the risk of non-robust estimates with too many predictors. PCA permits to constrain the number of predictors without undue loss of information. A parsimonious choice of the truncation $n$ was obtained by the following diagnostics: Firstly, the fraction of explained variance in the available station data decreases quickly with higher PCA modes. The first mode already accounts for 89.2% of the variance and with the first nine modes the explained variance reaches 99%. Secondly, experiments with a varying number of predictors to calculate the area-average in an independent validation period showed, that the root mean squared error (RMSE) stopped decreasing (even slightly increased) for $n > 10$. Finally, the number and choice of predictors was assessed based on stepwise regression using the AIC criterion (e.g., Wilks,
Fourteen out of all 19 modes were proposed to be retained, including all modes from 1 to 7. Based on these considerations, admittedly not entirely conclusive diagnostics, we have decided to constrain the linear model to \( n = 9 \), hence, retaining nine leading PCA modes as predictors.

For evaluation purposes the regression model was also calibrated to the monthly data of single seasons individually. It was of interest whether the station weights would differ from one season to another resulting in a better overall performance compared to the standard model based on all calendar months. Results will be discussed later in section 4.2. A suitable truncation for the seasonal models was determined from similar diagnostics like those for the yearly model. The results suggested a truncation of \( n = 9 \) also for the seasonal models.

As a measure of uncertainty of the statistically predicted area-mean temperature we use the prediction interval of the linear regression. The prediction interval describes error statistics for future data (i.e., not used in the calibration) under the assumption that uncertainties in the predictors and the measurements themselves are negligible against those from statistical estimation (e.g., Wilks, 2006). This measure of uncertainty can be used in applications of the predicted area mean series, such as to test the statistical significance of a difference between two independent predictions or to estimate prediction variances for seasonal or annual means, that is, averages of monthly predictions. Relying on the familiar assumptions of linear regression, we calculate approximate variances of a difference (an average) by summing (averaging) the pertinent prediction variances.

All computations are conducted in R (R Core Team, 2012).

4 | RESULTS AND DISCUSSION

In this section, the new temperature series for Switzerland is presented and long-term trends in seasonal and yearly values are analysed (section 4.1). We discuss results from a detailed evaluation (section 4.2) and investigate the risk of non-stationarities affecting the reconstruction (section 4.3). An extension of the method to sub-regions of the country is presented in section 4.4. Finally, in section 4.5, we discuss differences in the long-term temperature variation in Switzerland between the new analysis and that from existing global grid datasets.

4.1 | Area-mean temperatures of Switzerland since 1864

Figures 2 and 3 depict annual and seasonal area-mean temperatures of Switzerland from 1864 to 2016 as derived from the proposed statistical model. The evolutions can be characterized by a superposition of inter-annual and inter-decadal variations plus a long-term trend. The uncertainty (95% confidence interval, indicated in green) in the individual estimates is about ±0.03 °C (±0.06 °C) for annual
(seasonal) estimates, which is quite small compared to the magnitude of the variations.

Annual temperature has increased since the beginning of the instrumental measurements. The average of the latest WMO standard period 1981–2010 lies almost 1.5 °C above the early-industrial level 1864–1900. When expressed by a linear trend, the rate of increase amounts to 1.29 °C per 100 years. The evolution in seasonal area-means reveals similar trends ranging in magnitude from 1.22 to 1.36 °C per 100 years. The largest warming is found in autumn. While autumn and winter temperatures show a more gradual increase over the full period, spring and summer evolutions contain larger decadal fluctuations and more stepwise changes of the temperature level.

A comparison of the present Swiss mean temperature series to the previous climate monitoring series of MeteoSwiss (Begert et al., 2005) reveals minor but noteworthy differences (Table 2). For example, the temperature increase in the new dataset is smaller by 0.13 °C per 100 years in winter, while in summer the trend is larger by 0.1 °C per 100 years. The differences can be explained by the fact, that the old dataset, consisting of a simple average over 12 stations, does overemphasize the lowland areas where stations are predominantly located. As temperature increase in winter is larger in the lowlands than in the mountains (see later section 4.4, Table 4), winter trends for the whole of Switzerland are overestimated in the old area-mean temperature series. The same argument with opposite signs applies for the trends in summer.

Differences also exist with respect to the ranking of single values (Table 2). The most prominent example is the change in ranking between 2011 and 2014 as second warmest years. While 2014 was second on record in the old series, the new series declares 2011 as warmer. The latter year was particularly warm at high elevations. The better representation of these regions in the new method explains the swap. First on record in both datasets is the year 2015. However, when considering the uncertainties delivered by the new method, the area-mean values for 2011 and 2015 are not statistically different at the 5% significance level. In its future communication about climate monitoring, MeteoSwiss will use this more thorough assessment to characterize uncertainties in the relative ranking.

It is interesting to briefly inspect the results of the statistical calibration in terms of the station weights \( w_k \) for the area-mean value (see Equation 4). Results are displayed in Figure 4. Values range from almost zero at LUG to as much as 14% at SAE. The distribution of weights reflects the fact that the rare high-elevation stations in the sample represent larger areas of Switzerland than the more numerous lowland stations. GSB and SAE, the two only stations above 2000 m asl, account together for 24% of the total weight. Adding mid-elevation stations (GRC, DAV, SBE and CHM), the weight sums up to 50% from only 6 out of 19 stations. The remarkable low weights of LUG, SAM and SIO indicate the comparatively unique locations of these stations within the surrounding topography and/or near the border of the country. Both SIO and LUG are located in the very low altitudes of the Valais and the Ticino and, hence, are representative for only a small part of their larger environment. SAM is often affected by a remarkable cold-air pool in winter and may therefore have smaller representativity (valley floor only). Moreover, it is not unexpected that the close-by station SIA, with its broader representativity, draws weight away from SAM. In summary, the station weights, obtained from the calibration, reveal a highly plausible pattern with a comprehensible relation to location and general representativity for the Swiss temperature climate.

### Table 2 Comparison of annual and seasonal values from the present Swiss mean temperature series (new) with those from the previous climate monitoring series of MeteoSwiss (old). With regard to the period 1864–2016 linear trend estimates with a 95% confidence interval and the warmest/coldest years as well as the correlation coefficients are listed

| Season | Old | New | Old | New |
|--------|-----|-----|-----|-----|
| DJF    | 1.35 ± 0.51 | 1.22 ± 0.50 | 1.20 ± 0.35 | 1.24 ± 0.36 |
| MAM    | 1.18 ± 0.33 | 1.28 ± 0.34 | 1.32 ± 0.34 | 1.36 ± 0.35 |
| JJA    | 1.27 ± 0.22 | 1.29 ± 0.22 | 1.27 ± 0.22 | 1.29 ± 0.22 |
| SON    | 1.35 ± 0.51 | 1.22 ± 0.50 | 1.20 ± 0.35 | 1.24 ± 0.36 |
| YYY    | 1.18 ± 0.35 | 1.28 ± 0.34 | 1.32 ± 0.34 | 1.36 ± 0.35 |
| Correlation | 0.991 | 0.998 | 0.996 | 0.993 |

4.2 | **Evaluation of the reconstruction procedure**

In this section, we quantify the accuracy with which the weighted average of station values does reproduce the true country-mean temperature and we verify that the prediction interval of the employed reconstruction model is a reliable measure of the uncertainty. To this end, a comparison is made between area-averages predicted from the 19 long-term stations and area-averages from the temperature grid data set. The comparison is made for all months of the validation period, that is, the 14 years 1981–1984 and
2005–2014, which is fully independent of the calibration process and includes both warm and cold extremes in the period with available grid data.

Figure 5 depicts the distribution of the differences, prediction minus reference from grid dataset, for each calendar month, each season and the year. Results are presented for both, our standard model calibrated over all months (green) and a seasonally stratified model, where weights vary between seasons (yellow). The 95% prediction interval is also indicated (red). To be precise, the plotted intervals are the mean of intervals over all predictions with the standard model in the pertinent time aggregate. In the seasonally stratified case, prediction intervals (not shown) are slightly larger due to the smaller sample size. The root mean square error (RMSE, Table 3) is used as a summary statistic of the prediction errors.

In general, the errors of our reconstruction procedure are rather small. They are almost always within ±0.2°C for monthly, and within ±0.1°C for seasonal and annual predictions. There are seasonal variations in the error characteristics: For winter months, the errors are larger, ranging from −0.2°C to +0.2°C, with a RMSE of around 0.1°C (see Table 3), while for summer months they are smaller, between −0.1°C and +0.1°C and a RMSE around 0.05°C. Larger errors in winter are plausible, considering that spatial temperature distributions are much more convoluted then. For example, radiative cooling at the surface often causes temperature inversions over the Swiss Plateau and in larger mountain valleys, which can manifest in strong temperature gradients along the surface. These are, obviously, more difficult to portray with the limited station sample used in the reconstruction, compared to the more gradual variations in summer.

There is also a sign of seasonal variation in the systematic component of the error with a tendency of under- and overestimation in winter and summer, respectively (Figure 5). This is more readily noticed in the seasonal aggregate. We do not have a clear explanation of this pattern, other than it could originate from peculiarities in the temperature conditions of the calibration period.

Comparing the error distributions (Figure 5) and statistics (Table 3) between the standard model (calibrated for all calendar months) and the seasonally stratified model reveals very similar characteristics. There is no indication of the seasonal model being superior. Possibly, the advantage of seasonally varying station weights is compromised by larger sampling errors, considering that the number of months available for calibrating a seasonal model is only one fourth of that for the standard model. This result may change when many more years become available for calibration. But for now, the comparison suggests that the all-year standard model cannot be measurably improved by stratification.

It is interesting to compare the 95% prediction intervals (Figure 5, red) with the corresponding 95% error spread (box-plot whiskers). Most of the validation errors are contained in the prediction interval. As a result of the seasonal variation of errors, there are more than (less than) the nominal 5% of cases outside the average prediction interval in winter (summer). For the seasonal and annual aggregates, there is a slight but measurable discrepancy between interval and error spread. This may hint to a slight overconfidence in the predictions. But in general, the prediction interval seems to provide a quite reliable measure of the reconstruction uncertainty and it is a valuable complement to the estimates themselves in order to prevent users from over-interpretation.

4.3 Stationarity of station weights

A fundamental assumption of the statistical procedure employed here, is that the statistical characteristics of the temperature distributions in Switzerland are similar for periods when the model is calibrated and when it is applied for prediction. Changes of these characteristics would imply that stationary station weights could lead to biased predictions in some periods. A major
concern in this regard is that the application of our model stretches over a very long period, more than 150 years, but is calibrated in its warmest decades (see Figure 2). Hence, there is a risk that station weights determined in recent decades are not representative for the climate of the 19th and early 20th century. In this section, we test the sensitivity of model predictions to “skewed” calibrations with an experiment where calibration data is chosen deliberately non-representative. The experiment provides insight on the potential effect of, and hence, the likelihood of non-stationarities affecting the results in our standard application.

In our experiment, we split the months of our standard calibration period (1985–2004) into two equal groups, a “warm” and a “cold” group, depending on the anomaly of the country-wide mean temperature (as inferred from the grid dataset). The statistical model of section 3 is then calibrated separately for the warm and the cold group, and the two resulting “skewed” models are employed, each, to predict a temperature series over the entire period. Differences in station weights and predictions between the two models may illustrate the magnitude of error (bias) that may have resulted from the calibration of our standard model in the warmest decades. Indeed, the difference in medians between the warm and cold groups roughly corresponds to the difference between the standard calibration period and the 19th century (Figure 6). Still, our experiment may be viewed as a pessimistic analogue because, unlike in our real application, the two groups are strictly not overlapping.

Results for station weights derived with the three different model calibrations (standard, warm and cold group, Figure 7) reveal mostly minor differences. Only in few cases are the results of the purposely skewed models outside the 95% confidence interval of the standard model. With the “cold” calibration the weights are relatively smaller for SAM and SMA but larger for LUG and BER. During warm months, on the other hand, the weights appear to be larger

| YYY | Single | Season |
|-----|--------|--------|
| DJF | 0.05   | 0.04   |
| MAM | 0.05   | 0.04   |
| JJA | 0.05   | 0.04   |
| SON | 0.05   | 0.04   |

| D  | J  | F  | M  | A  | M  | J  | A  | S  | O  | N  |
|----|----|----|----|----|----|----|----|----|----|----|
| Single | 0.05 | 0.11 | 0.09 | 0.07 | 0.05 | 0.05 | 0.05 | 0.05 | 0.06 | 0.04 | 0.05 | 0.06 |
| Season | 0.05 | 0.10 | 0.08 | 0.07 | 0.05 | 0.05 | 0.05 | 0.05 | 0.06 | 0.04 | 0.06 | 0.07 |

| D  | J  | F  | M  | A  | M  | J  | A  | S  | O  | N  |
|----|----|----|----|----|----|----|----|----|----|----|
| Single | 0.05 | 0.05 | 0.04 | 0.03 |
| Season | 0.05 | 0.04 | 0.04 | 0.03 |
Again, the effect of the skewed calibration on the final time series seems to be rather small: Swiss mean temperatures calculated separately from the cold and the warm model differ by less than ±0.05 °C for annual values (see Figure 8), less than ±0.08 °C for seasonal values and less than ±0.14 °C for monthly values. All “skewed” estimates are well contained in the 95% prediction intervals of the results from the standard model. Note, that these intervals are very small compared to the year-to-year variations and, hence, the skewed results could be hardly distinguished from the standard results in a time series like those in Figures 2 and 3. Nevertheless, Figure 8 reveals a systematic component in the difference of predictions between the warm and cold models: The warm model makes warmer predictions in cold periods (before 1980) and the cold model makes warmer predictions in warm periods (after 1990). This may be a sign for statistical differences between warm and cold months that would violate the stationarity assumption made in our modelling. But its quantitative effects are insignificant. Note that the systematic signal in Figure 8 implies a long-term trend of about 0.04 °C per 100 years, which is very small compared to the actual temperature trend of 1.2 °C per 100 years (see Figures 2 and 3).

Clearly, the test of this section cannot ultimately exclude artefacts from changes in statistical characteristics over time. For example, differences in typical temperature patterns (the part not resolved by the stations) between cold recent and average 19th century months could still lead to biases. Our test is to be seen as an attempt to find violation of stationarity by purposely creating strongly skewed conditions between calibration and prediction periods. The finding that predictions are fairly insensitive to these perturbations suggests that the inevitable prominence of warm conditions in our model calibration may be of lesser concern than expected at first sight.

4.4 | Application to sub-regions

Considering the topographic complexity of Switzerland there is also interest in long-term temperature series for sub-regions. Using the same methodology as for the territory of the country, area-mean temperature series have been derived for the following three sub-regions: the region north of the Alpine main crest at elevations below 1000 m asl (CH-NL; 44% of the land area), the north side above 1000 m asl (CH-NH; 45% of the land area), and the region south of the Alpine main crest (CH-S; 11% of the land area; see Figure 1). The southern region includes, in addition to the Canton of Ticino, the Simplon region and the southern valleys of Grison. It is not stratified by altitude due to the relatively small area and the limited number of stations contained. 75% of the CH-S region is higher than 1000 m asl. The procedure of reconstruction is identical to that for the entire country. Notably, all 19 stations have been used as predictors for the sub-regional mean temperatures.
The application produces plausible patterns of stations weights for the three sub-regions (Figure 9). While the six stations of the Swiss plateau account for 78% of the total weight for low altitudes north of the Alpine main crest, their coefficients become very small and occasionally negative for the high altitudes. For CH-NH the weights of the high elevation stations SAE and GSB together with GRC (Valais), DAV (Grison) and CHM (Jura) dominate (87% of total weight). For the southern region, the two stations south of the Alpine main crest (SBE and LUG) together with the two high elevation stations GSB and SIA close to that border contribute the most (82% of total weight). Due to the fact that CH-S also includes high altitudes, even station SAE, located far north and clearly outside the region, receives a measurable influence. Note, that negative station weights in connection with the sub-regional area-means can occur as a result of anti-correlations (e.g., between the north and south, or high and low elevations) and because of noise in the estimation.

An evaluation of the sub-regional predictions, analogous to that for the whole country in section 4.2, shows that the magnitude of errors for the two sub-regions CH-NL and CH-NH is comparable to that for the whole country but larger in CH-S. Root mean square errors for monthly values range up to 0.1 °C with smaller values in summer than in winter. In CH-S RMSE is around 0.2 °C. The larger errors are likely a result of the smaller station density there. The evaluation pinpoints to the potential of our reconstruction technique, even for smaller than national terrains, provided there is a decent coverage of the region with spatially representative stations.

A trend analysis of the sub-regional temperature series reveals differences that are worthy of a short discussion (Table 4). Firstly, statistically significant warming is observed in all regions and in all seasons. Trend values

![FIGURE 8](image-url) Differences in annual area-mean temperatures of Switzerland 1864–2014 between two experimental models fitted on warm and cold years of the calibration period only (warm minus cold). The mean 95% confidence interval from the full model is shown with a dashed, vertical line

![FIGURE 9](image-url) Station weights (in %) of the 19 stations used to calculate the area-mean temperature for Switzerland (CH) and three sub-regions (CH-NL: north of Alpine main crest below 1000 m asl; CH-NH: north of Alpine main crest above 1000 m asl; CH-S: south of Alpine main crest). Negative station weights are given in green and the extent of the regions the area-mean temperature is calculated for is shown in grey.
range from 0.9 °C per 100 years (CH-S in winter) to 1.43 °C per 100 years (CH-NL in winter). Trends in the south are generally smaller and more uniform between seasons than in the north. The different altitudes north of the Alpine main crest show similar trends in spring and autumn as well as for the year as a whole. Some differences between the two elevation ranges (CH-NL vs. CH-NH) can be observed in winter and summer with a larger warming below (above) 1000 m asl in winter (summer). Most of the differences just mentioned turn out to be statistically significant in that the null hypothesis of a zero trend in the series of differences between regions is rejected (see Table 5).

Evidence for elevation-dependent warming trends in mountain regions has attracted some attention in the literature, not least because of implications for high-mountain cryosphere and ecosystems (e.g., McGuire et al., 2012; Ohmura, 2012; Tao et al., 2014). An earlier study for Switzerland (Ceppi et al., 2012) found, for the period 1959–2008, larger warming trends at low compared to high altitudes in autumn and winter. On the long time-scale considered here (1864–2016) a larger low-level warming is still evident in winter, whereas for autumn the result is different (see Table 4). A sensitivity of elevation-dependent warming upon analysis period has been noted in other studies (e.g., Pepin et al., 2015). The new Swiss regional and elevation-stratified series derived in this study offer a reliable data source for investigating elevation-dependent warming over the whole period of instrumental measurements.

4.5 Comparison of global data sets with the new Swiss temperature series

Global gridded data sets of surface air temperature are the basis for current global and continental temperature monitoring (e.g., Hartmann et al., 2013). Several of such datasets exist, building on archives of historical instrumental records over land and sea, that were compiled by different centres and working groups. The most prominent data sets encompass CRUTEM4 developed and maintained by the UK Met Office and the Climatic Research Unit (Jones et al., 2012), GHCNv3 produced by NOAA’s National Center for Environmental Information (Lawrimore et al., 2011), GISTEMP of NASA’s Goddard Institute for Space Studies (Hansen et al., 2010) and BEST from the University of California, Berkeley (Rohde et al., 2013). The source, volume and processing of the underlying observational series varies between these datasets. CRUTEM4, for example, is the sole dataset to include national and regional homogenized series where available.

A comparison to national high-quality temperature analyses, such as the one developed for Switzerland, is of high interest. Data quality processing at global data archives often has to rely on limited information and meta data, which can affect the accuracy of homogenization and finally impact on long-term variations and trends at the respective grid points. The presented Swiss area-mean temperature series offers a high-quality reference because its data processing builds on extensive information about station and network history, on results from parallel measurements, data from a denser station network and in-house expertise.

Our comparison focuses on the evolution and trend of annual and seasonal mean temperature in the period 1880–2016. Series for Switzerland from the above-mentioned global datasets are obtained by selecting and averaging the grid points encompassing Switzerland (see Table 6).

Table 6 lists estimates of the linear temperature trends in Switzerland over the period 1880–2016 as represented by the four global grid datasets and the high-quality Swiss series (SwissT) developed in this study. Clearly, all datasets report a warming that is statistically significant at the 5% significance level. Interestingly, the global datasets exhibit mostly smaller warming than SwissT, with GHCNv3 and CRUTEM4 being, in general, in closer agreement than GISTEMP and BEST. For annual temperatures, the underestimate is less than 0.1 °C per 100 years for GHCNv3 and CRUTEM4, but almost 0.3 °C

| TABLE 4 | Annual and seasonal temperature trends (in °C per 100 years) from a linear regression trend analysis applied to area-mean temperature series for Switzerland and for three sub-regions over the period 1864–2016. A 95% confidence interval is attributed to each trend value. For the significance of differences in the trend between pairs of regions, see Table 5 |
|------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                  | DJF                 | MAM                 | JJA                 | SON                 | YYY                 |
| CH-NL            | 1.43 ± 0.56         | 1.24 ± 0.35         | 1.19 ± 0.35         | 1.35 ± 0.34         | 1.31 ± 0.23         |
| CH-NH            | 1.10 ± 0.51         | 1.28 ± 0.38         | 1.42 ± 0.34         | 1.43 ± 0.39         | 1.32 ± 0.23         |
| CH-S             | 0.90 ± 0.44         | 1.12 ± 0.36         | 1.14 ± 0.32         | 1.17 ± 0.35         | 1.09 ± 0.21         |
| CH               | 1.22 ± 0.50         | 1.24 ± 0.36         | 1.28 ± 0.34         | 1.36 ± 0.35         | 1.29 ± 0.22         |

| TABLE 5 | Sign and significance of linear trend estimates in difference series between pairs of sub-regions. Results are shown for annual (YYY) and seasonal mean temperature. Significant values on the 5% level are marked ‘Signf.’; not significant values are left out |
|---------|---------------------|---------------------|---------------------|---------------------|---------------------|
| CH-NL – CH-NH | Signf.+               |                      | Signf.–             |                      |                     |
| CH-NL – CH-S   | Signf.+               | Signf.+             |                      |                      |                     |
| CH-NH – CH-S   | Signf.+               | Signf.+             | Signf.+             | Signf.+             | Signf.+             |
The situation is slightly different in DJF where GISTEMP is closest and GHCNv3 and CRUTEM4 overestimate the warming. More insight into these differences is obtained when inspecting the difference series (global minus SwissT). Figure 10 depicts the situation for annual mean temperature. It seems that trend differences in all data sets and seasons are related to a shift-like discontinuity in the difference series around 1980. This coincides with a fundamental change in the meteorological network in Switzerland, that is, when three-fourths of the stations were transformed into automatic operation. Begert et al. (2005) found shifts in the raw measurements at this date, systematically across the network, and requiring homogenization adjustments of around −0.3 °C for yearly values of the manual readings. Homogenization procedures conducted in global data archives are, however, mostly based on relative homogeneity testing, including GHCNv3 (Menne and Williams Jr, 2009), GISTEMP (Hansen et al., 1999) and BEST (Rohde et al., 2013). Network-wide simultaneous inhomogeneities are hardly detectable with such a processing, unless the information is explicitly injected. Our comparison suggests that, at least to a large part, differences found in long-term trends between our high-quality series and global datasets might be the result of insufficient homogenization of the transition from conventional to automated measurements in global data archives. It is interesting to note that CRUTEM4, which builds on station series that were pre-homogenized, exhibits the smallest discrepancy to SwissT (see Figure 10). In the case of Switzerland, the station series in CRUTEM4 originate from the HISTALP project (Auer et al., 2007), a homogenization effort in a denser station network and with more knowledge and metadata from the different data providers.

Apart from an insufficient homogenization of the network-wide inhomogeneity in 1980, there may be

| Data set          | Grid box     | DJF  | MAM  | JJA  | SON  | YYY  |
|-------------------|--------------|------|------|------|------|------|
| SwissT (v 1.0)    | Area-mean    | 1.41 ± 0.57 | 1.47 ± 0.43 | 1.59 ± 0.40 | 1.50 ± 0.42 | 1.48 ± 0.25 |
| GHCNv3 (GHCN-M v 3.3.0) | 45–50°N/5–10°E | 1.61 ± 0.59 | 1.24 ± 0.40 | 1.33 ± 0.38 | 1.54 ± 0.39 | 1.44 ± 0.24 |
| GISTEMP (gistemp250, June 2017) | 46–50°N/6–10°E | 1.34 ± 0.64 | 1.08 ± 0.41 | 1.16 ± 0.38 | 1.25 ± 0.39 | 1.21 ± 0.25 |
| CRUTEM4 (CRUTEM4.5.0.0) | 45–50°N/5–10°E | 1.52 ± 0.62 | 1.30 ± 0.40 | 1.42 ± 0.37 | 1.36 ± 0.39 | 1.41 ± 0.24 |
| BEST (Compl_TAVG_LatLon1) | 46–50°N/6–10°E | 1.22 ± 0.65 | 1.15 ± 0.40 | 1.21 ± 0.37 | 1.21 ± 0.38 | 1.20 ± 0.24 |

---

**TABLE 6** Annual (YYY) and seasonal temperature trends (in °C per 100 years) from a linear regression trend analysis for the period 1880–2016. Results are compared between the country-mean Swiss temperature developed in this study (SwissT) and four global data sets (see section 4.5 for detail). A 95% confidence interval is indicated for each of the estimates.

---

**FIGURE 10** Differences in yearly mean temperature between the high-quality Swiss estimate of this study (SwissT) and values from global temperature data sets (global minus Swiss). A linear trend estimate (thin) and a 20-year Gaussian low-pass filter (bold) are given in black. The year 1980 is marked with a dashed line, indicating the date of the transition from manual to automated measurements in Switzerland.
additional sources for discrepancy. For example, the difference series with GISTEMP and BEST also exhibits a shift around 1920. Further analyses would be required to understand this feature. Moreover, the grid boxes of the global datasets are substantially larger than Switzerland and extend northward into less mountainous areas, which could lead to differences in trend estimates compared to SwissT. Nevertheless, our comparison pinpoints to limitations in global data sets when long-term trends are to be inferred from data at single grid points that merit further study.

5 | CONCLUSIONS

In this study, a new method has been presented for the calculation of monthly temperature series, representative of the area-average over a well-defined region (here Switzerland), extending over long periods (here 153 years), and satisfying high standards in long-term consistency. The proposed method overcomes limitations of more ad-hoc constructions, such as representativity biases in simple averaging over stations or temporal inhomogeneities in conventional spatial interpolations. Regions with complex topography are particularly prone to representativity biases, with high altitudes being notoriously under-represented in long-term records. Our development builds on a linear combination of data from a fully continuous (stationary) station sample (here 19 stations), with coefficients (weights) that are, themselves, data-driven and calibrated to best possibly represent the true domain average, despite biased sampling conditions.

Extensive testing of our application for Switzerland revealed errors (RMSE) in area-mean monthly temperatures ranging from 0.04 to 0.11 °C, with larger errors in winter than in summer. For seasonal and yearly aggregates, errors were found between 0.03 and 0.05 °C. Compared to the variance in the time-series, these errors are small. Our analyses also suggest that the model-implicit prediction error is a good approximation of the uncertainty in the resulting area means, although it is slightly overconfident. The station weights, objectively estimated, show a highly plausible pattern, consistent with expert judgement, and this supports our confidence that the objective approach has successfully disentangled representativity biases. Clearly, the statistical procedure used here relies on a stationarity assumption that may be questioned in the presence of the observed warming. However, tests suggest that our results show a very small sensitivity to purposely biased calibration conditions. The assessments imply that the resulting time series represent the long-term temperature evolution and trends in Switzerland at a high level of accuracy.

The new Swiss temperature series closely follows its older version, based on a simple average over 12 stations, in regards to the year-to-year variations. But there is an interesting change in the magnitude and seasonal pattern of long-term trends. These could be traced down to the under-representation of high-mountain conditions in the old version, a caveat that could now be amended.

The proposed methodology has a high potential for applications in other areas, for other climate variables and over a longer period. The procedure may be particularly beneficial in areas of complex orography where a definition of station weights just from expert knowledge is not obvious. An important pre-requisite is, however, the existence of an explicit an accurate spatial analysis over several decades, in the form of a high-resolution grid dataset. If available, the processing is straightforward and fast; it can be accomplished with statistical procedures that are readily available in most data analysis tools. Obviously, error statistics will depend on the complexity of the climate and the volume and distribution of available observations. Validation experiments, such as that in section 4.2 will have to be conducted for pertinent information. Moreover, precaution should be exercised with the assumption of statistical stationarity, fundamental to the technique. Sensitivity experiments of the sort of section 4.3 are important to assess related risks.

The warming trend since 1880 in Switzerland as inferred from the new regional time series is larger than that found at grid-boxes over Switzerland in the four prominent global datasets GISTEMP, BEST, CRUTEM4 and GHCNv3. Differences of GISTEMP and BEST from the Swiss series are considerable (more than 0.2 °C per 100 years). Our analysis (section 4.5) points to the possibility of a data quality issue in the global datasets. It might be related to an inhomogeneity around 1980 that affected the entire network of Swiss climate stations (automation) but was difficult to detect with relative homogeneity tests commonly applied in global data archives. It seems that time series from individual grid points in global datasets can suffer from residual data quality problems and that the related uncertainties are not necessarily reflected in the spread between datasets. Clearly, it would be desirable if data in global archives could be regularly updated to better reflect the latest homogeneity processing of the data owner or if specific and more metadata information could be included in the homogenization of these archives.

The new Swiss temperature analysis encompasses a country-wide plus three sub-regional series for the south of the Alpine main crest, and the north at elevations below and above 1000 m asl. These data series constitute the basis for operational climate monitoring and public information on temperature changes. Owing to the confidence interval provided with the method for each estimate, the communication can make uncertainties transparent. For a professional interpretation, users should recognize that the series represent macro-climatic variations as registered under standard WMO measurement conditions. Variations of area-mean temperature as a result of, for example, urbanization or changes in land-cover are not reproduced. Also, some of the smaller-scale topo-climatic patterns, such as valley-scale cold-air pools, may be poorly represented in the reconstruction sample.
The regional temperature dataset of Switzerland is referenced with the DOI 10.18751/Climate/Timeseries/CHTM/1.0 and can be downloaded from the MeteoSwiss webpage (www.meteoswiss.ch).

ACKNOWLEDGEMENTS

The authors wish to thank the colleagues from the climate section of MeteoSwiss that contributed with their support and valuable discussions to the present study.

ORCID

Michael Begert http://orcid.org/0000-0002-9569-6842

REFERENCES

Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schönker, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Eyring, V., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westemeier, G., Kveton, V., Bochniak, O., Stauny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z. and Nieplova, E. (2007) HISTALP – historical climatological surface time series of the greater alpine region 1760-2003. International Journal of Climatology, 27, 17–46. https://doi.org/10.1002/joc.1377.

Begert, M. (2008) Die Repräsentativität der Stationen im Swiss National Climatological Network (Swiss NBCN). Arbeitsberichte der MeteoSchweiz. Report number: 217, Zürich.

Begert, M., Schlegel, T. and Kirchhofer, W. (2005) Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. International Journal of Climatology, 25, 65–80. https://doi.org/10.1002/joc.1118.

Begert, M., Seiz, G., Foppa, N., Schlegel, T., Appenzeller, C. and Müller G. (2007) Die Überführung der klimatologischen Referenzstationen der Schweiz in das Swiss National Basical Climatological Network (Swiss NBCN). Arbeitsbericht der MeteoSchweiz, Nr. 215, Zürich.

Begert, M., Seiz G., Schlegel, T., Musa, M., Baudrallaz, G. and Moesch, M. (2003) Homogenisierung von Klimareihen der Schweiz und Bestimmung der Normwerte 1961–1990. Schlussbericht des Projekts NORM90. Veröffentlichung der MeteoSchweiz, Vol. 67, Zürich.

Brohan, P., Kennedy, J.J., Haris, I., Tett, S.F.B. and Jones, P.D. (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. Journal of Geophysical Research, 111, D21106. https://doi.org/10.1029/2005JD006548.

Ceppi, P., Scherrer, S.C., Fischer, A.M. and Appenzeller, C. (2012) Revisiting Swiss temperature trends 1950–2008. International Journal of Climatology, 32, 203–213. https://doi.org/10.1002/joc.2260.

Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsu, N., Allan, R.J., Yun, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouch, R. I., Grant, A.N., Grossman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maurer, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, O., Ross, T.F., Trigo, R., Wang, X.L., Woodruff, S., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Schöner, W., Ungersbü
