




Influence of instrumentation on long temperature time series

Fiorella Acquavota^{1,2} · Simona Fratianni^{1,2}  · Enric Aguilar³ · Guillaume Fortin⁴

Received: 10 December 2018 / Accepted: 27 August 2019 / Published online: 7 September 2019
© Springer Nature B.V. 2019

Abstract

In time series of essential climatological variables, many discontinuities are created not by climate factors but changes in the measuring system, including relocations, changes in instrumentation, exposure or even observation practices. Some of these changes occur due to reorganization, cost-efficiency or innovation. In the last few decades, station movements have often been accompanied by the introduction of an automatic weather station (AWS). Our study identifies the biases in daily maximum and minimum temperatures using parallel records of manual and automated observations. They are selected to minimize the differences in surrounding environment, exposition, distance and difference in elevation. Therefore, the type of instrumentation is the most important biasing factor between both measurements. The pairs of weather stations are located in Piedmont, a region of Italy, and in Gaspé Peninsula, a region of Canada. They have 6 years of overlapping period on average, and 5110 daily values. The approach implemented for the comparison is divided in four main parts: a statistical characterization of the daily temperature series; a comparison between the daily series; a comparison between the types of events, heat wave, cold wave and normal events; and a verification of the homogeneity of the difference series. Our results show a higher frequency of warm (+10%) and extremely warm (+35%) days in the automated system, compared with the parallel manual record. Consequently, the use of a composite record could significantly bias the calculation of extreme events.

✉ Simona Fratianni
simona.fratianni@unito.it

¹ Department of Earth Sciences, University of Torino, Via Valperga Caluso, 35, 10125 Turin, Italy

² Centro Interdipartimentale sui Rischi Naturali in Ambiente Montano e Collinare, NatRisk, via Leonardo da Vinci 44, Grugliasco, 10095 Turin, Italy

³ Center for Climate Change (C3), Universitat Rovira i Virgili, Tarragona, Spain

⁴ Département d'histoire et de géographie, Université de Moncton, Moncton, NB E1A 3E9, Canada

1 Introduction

The scientific community agrees that global warming is unequivocal, as is demonstrated by the latest IPCC report (Stocker 2014). Since 1850, the Earth's temperature has been increasing rapidly (Brohan et al. 2006; Esper et al. 2002; Acquotta et al. 2015). The trend shows a warming of 0.85 °C from 1880 to 2012 but the magnitude of change detected is very sensitive to the quantity and quality of the data used (Acquotta et al. 2009; Nicholls 1995) as well as the time period considered. For example, the rate of warming over a 15-year period, from 1998 to 2012, is equal to 0.05 °C/decade, but increases to 0.12 °C/decade when it is calculated for about 50 years, from 1951 to 2012. Karl et al. (2015) have identified possible biases in the sea surface temperature series that result from a change in instrumentation, with new data showing systematically cooler temperature than older data (Reynolds and Chelton 2010). In addition, the geographical scale could have a significant influence on trend detection. For example, extreme temperatures show greater amplitudes at the local scale than at the global scale (Nigrelli et al. 2018; Luterbacher et al. 2004), which can have significant implications for understanding local and regional impacts on society and natural environments.

The World Meteorological Organization, WMO, is aware of the importance of identifying the breaks that can be created by station or network changes for reasons of cost-efficiency or innovation. They wrote guidelines recognizing the need for National Meteorological Services to improve their climate data and monitoring services. This work can be done to establish transfer functions either through parallel observations or through alternative approaches in order to ensure that data continuity can be defined (WMO 2007).

Many studies (Hausfather et al. 2016; Fiebrich and Crawford 2009; Davey and Pielke 2005) have tried to identify and calculate the artificial biases in temperature series. These studies show that the biases are rarely constant and depend on the interaction between the weather elements and the local topographic or physiographic features at each site (altitude, exposure, and distance to water bodies). Milewska and Vincent (2016) analysed the daily difference between maximum and minimum temperatures recorded by manual and automatic stations. The differences were classified by season and wind speed conditions. Most of the biases vary with the seasons, and few show wind dependency. Hubbard and Lin (2006) examined the effects of the instrument change in the US Historical Climatology Network. Their results indicate that the magnitude of break changes at individual stations ranges from -1.0 to over $+1.0$ °C. Leeper et al. (2015) evaluated how diverse technological and operational choices at the USCRN (US Climate Reference Network) and COOP program (Cooperative Observer Program) impact temperature observations. They showed that COOP sensors generally have warmer ($+0.48$ °C) daily maximum and cooler (-0.36 °C) minimum temperatures than USCRN, with considerable variability among stations. Gallo (2005) examined the differences between the CRN (Climate Reference Network) station pairs in the USA and he identified significant differences in the annual minimum, maximum and mean temperatures. The microclimate near the weather stations seems to greatly influence the measurements. Karl et al. (1995) underline the importance of overlapping measurements between old and new instruments to increase the knowledge on data record and rehabilitation and enhance our ability to monitor climate change. The effects of changes in instruments, location and observing practices on climate measurements must be known prior to implementing such changes. Changnon and Kunkel (2006) showed the importance of understanding the uncertainties in the historical

climate record comparing the weather stations that possess data of high quality. The quality of the data is enhanced by documenting the history of the weather station and the changes in location, in instrumentation and in observing practices.

The purpose of this paper is to calculate “real” biases that might occur when the more ancient station of the pairs is discontinued. The transition from manual to automatic station introduces noise in the climate change signal. Each change has different characteristics, depending on the exact sensors involved and additional modifications (e.g. relocations, changes in exposure) aliasing with the transition itself. This study evaluates specific changes in the Italian and Canadian networks and should be followed by similar studies to ensure that long-time records can be accurately adjusted.

We developed a methodology to identify the biases in daily maximum and minimum temperatures that are due to “real transitions”, though more often than not, weather station networks suffer relocations which are accompanied by simultaneous changes in the measuring system.

We have selected pairs of manual and automatic weather stations, MWS and AWS, respectively, aiming to minimize the differences in surrounding environment, exposition, distance and difference in elevation. Also, we have shown that the differences influence the records, the extreme values in particular. This comparative approach was applied in two environments, one in Italy and the other in Canada to show the independence and applicability of the methodology in different and contrasted climate conditions. In “[Material and methods](#)”, we present the datasets; in “[Results](#)”, the methods; and in “[Conclusion](#)”, the results.

2 Material and methods

2.1 Study areas

In this study, the daily values of paired maximum and minimum temperature series are analysed for two different areas: Piedmont, a region located in north western Italy, and Gaspé Peninsula, a region located in the southeast of the province of Quebec in Canada (Fig. 1).

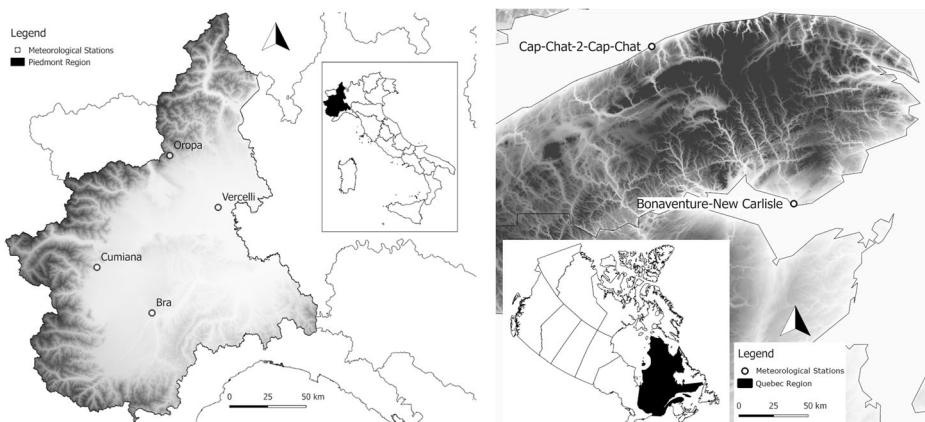


Fig. 1 The locations selected for the study: (left) Piedmont, (right) Gaspé Peninsula

2.1.1 Piedmont (Italy)

Piedmont covers a surface area of 25,402 km². The region is composed of 43.3% mountainous terrain, along with extensive areas of hills (30.3%) and plains (26.4%) (Fig. 1). The hills are composed of Torino, the Langhe and Monferrato. The alpine mountain area is located on the north-western Italian border, while to the east, the region is occupied by the Po Plain (Terzago et al. 2012).

The climate is continental, but the Alpine zones > 1500 m above mean sea level (AMSL) have a typical high mountain climate. Rainfall varies depending on the altitude and the orientation of the mountain slopes: annual precipitation is > 2000 mm in the Alps and in the hills but about 800 mm on the plain (Baronetti et al. 2018). Furthermore, frequent storms make the summer the wettest season in the Alps, whereas autumn and spring are the wettest seasons in the hills and plain. Temperatures are hot in the summer (> 30 °C and even > 35 °C during heat waves), especially on the plain, and the winters are cold (near 0 °C on the plain but considerably colder in the Alps); temperatures are generally fairly mild, although they may be subject to sudden changes when air masses from the north or north west affect the area. Following the Köppen climate classification, Piedmont is classified as continental temperate, Cf, in the plain, and in the Alps, it is classified as cold temperate, Cw (Fратиanni and Acquattrota 2017).

2.1.2 Gaspé Peninsula (Canada)

The Gaspé Peninsula covers a surface area of approximately 30,300 km² and is located in the southeastern part of the province of Quebec in Canada. The relief is relatively flat near the coast and at the bottom of the valleys, while the centre of the peninsula is made up of mountains that are part of the Monts Notre-Dame which represent the Quebec portion of the Appalachian Mountains (Capers et al. 2013). Mean annual temperatures (based on the normal from 1981 to 2010) vary from 8.9 to 7.3 °C for the maximum and from −3.2 to 0.1 °C for the minimum depending on elevation and distance from the sea. Total precipitation received in the study area varies between 933.2 and 1195.1 mm (733–811 mm fall as rain and 269.4 to 387.6 cm as snow) (Environment Canada 2018). According to the Köppen climate classification, the study area can be classified as a humid continental climate (Dfb), which means there is an important seasonal temperature difference. Usually, summers are warm to hot while winters are cold. Precipitation is common all year round.

2.2 Datasets

In the two studied areas, two independent weather networks exist and have measured temperatures independently. For Piedmont, the first is an old network with manual weather stations managed by the Italian Hydrographic Department and Maritime Service (SIMN), founded in 1917 and closed in 2003. The second is a new network with automatic weather stations managed by the Regional Agency for Environmental Protection Piedmont (ARPA), opened in 1986. In 2002, a national law forced the unification of the two networks. ARPA is the new national network and decided to close the SIMN weather stations located very close to the ARPA ones. The situation is different in Gaspé Peninsula where there are two independent

networks (provincial and national) in the study area that were launched at about the same time, but in both cases, most manual stations have been closed or replaced by automatic stations since the late 1980s.

The manual and automatic weather stations were divided into station pairs. The information from parallel measurements is necessary to estimate the contribution of non-climatic changes to the uncertainty budget.

From the two areas, we selected the weather station pairs (Table 1, Fig. 1), adopting the criteria formulated by different authors:

- distance between the station pairs is less than 20 km (Isotta et al. 2013). We chose to use this distance in order to correct the series by difference of latitude. The pairs of stations are aggregated within 1° of latitude (Peterson 2003);
- elevation range is less than 50 m (Biancotti et al. 2005). The temperature lapse rate ranging from -0.54 to -0.58 $^\circ\text{C}$ $(100\text{ m})^{-1}$ for Italy (Rolland 2003) and -0.39 $^\circ\text{C}$ $(100\text{ m})^{-1}$ for Canada (Dodson and Marks 1997). The error associated with the temperature lapse rate is lesser than the maximum error, ± 0.55 $^\circ\text{C}$, associated to the instrumentation utilized in this study;
- overlapping period is more than 5 years (Vincent and Mekis 2009);
- exposition of the instruments and local topography must be similar for both sites. Strong differences in topography and exposition strongly influence the nature of the frequency distribution of temperature (Trewin and Trevitt 1996).

We have decided to follow these criteria to select pairs of stations in order to avoid correcting data for elevation and latitude and reducing the difference due to microclimate. The microclimate influence on temperature measurements at nearby stations is potentially much greater than influences that might be due to latitude or elevation differences between the station pairs (Gallo 2005). Some studies (Trewin 2005, 2010) have identified that the relocation of the stations can introduce great differences in temperature record, and that some relocations introduce changes due to topography, elevation change or proximity to the coast. The differences can range between 3 and 5 $^\circ\text{C}$, for example, when a weather station is relocated from a ridge to a valley. The process used to select station pairs is similar to a “standard procedure” which a climatologist would use to create a long-term network maximizing the use of available series at the cost of introducing artificial biases related to all these factors. In a “standard procedure”, the climatologist would use homogeneity adjustments to minimize these biases. Here we take advantage of the parallel series to describe them and provide valuable a-priori information, both for validation of homogeneity adjustments or for the creation of benchmark datasets (Willett et al. 2014; Williams et al. 2012). As our pairing criteria tries to minimize all the other biases, we highlight the effect of the measuring system change, i.e. from MWS to AWS.

A preliminary quality control (QC) was carried out on the selected weather stations to highlight the incorrect values. The QC shows the values of daily maximum temperature that are lower or equal to minimum daily values, the periods with at least four consecutive days recording the same temperature and the outliers (Alexander and Herold 2015). The outliers were identified from data that exceeded the 99th/1st percentiles calculated on daily series of maximum and minimum temperatures. These values were then rechecked in the original records and compared with neighbouring stations if the analysed series had a reference series

Table 1 Selected locations in Piedmont, Italy, and Gaspé Peninsula, Canada

Station	P	N	MWS [m]	AWS [m]	D [m]	DE [m]	NA (%)
Bra	1998–2003	6	290	285	20	5	2
Cumiana	1990–2003	14	289	327	2800	38	20
Oropa	1995–2002	8	1180	1186	3	6	4
Vercelli	1993–2003	11	135	132	1360	3	8
Cap-Chat-2-Cap-Chat	2008–2014	7	22	46.4	11,200	24.4	2
Bonaventure-New Carlisle	2007–2016	10	22	46	11	24	2

Station	Tx	Tn	Break	Location	Instrument
Bra	AWS 18.6 MWS 20.0	AWS 8.5 MWS 8.1	2002–11 Tx and Tn	MWS and AWS in the garden of the museum Craveri	The MWS (SIMN) manual thermometer at maximum and minimum and AWS (ARPA) thermo-hygrometer TU20AS
Cumiana	AWS 18.0 MWS 16.7	AWS 7.6 MWS 7.1	1993–11 Tx	MWS in the garden of Salesian's school AWS in the garden of municipal school	The MWS (SIMN) manual thermometer at maximum and minimum and AWS (ARPA) thermo-hygrometer TU20AS
Oropa	AWS 12.1 MWS 11.6	AWS 4.7 MWS 5.4		MWS and AWS on the roof of the sanctuary	The MWS (SIMN) thermograph TM26000 and AWS (ARPA) thermo-hygrometer TU20AS
Vercelli	AWS 19.0 MWS 18.0	AWS 7.8 MWS 8.3	1998–10 Tx	MWS and AWS open area, out of the town	The MWS (SIMN) thermograph TM26000 and AWS (ARPA) thermo-hygrometer TU20AS
Cap-Chat-2-Cap-Chat	AWS 7.9 MWS 7.4	AWS 1.1 MWS 0.3	2011–07 Tn	MWS and AWS open space near the coast	The MWS (MSDEFCC) automatic stations used a Vaisala thermistor, 432A HANDAR AWS (ENCA) automatic weather stations with a Vaisala thermistor, Model HMP35C
Bonaventure-New Carlisle	AWS 9.3 MWS 8.8	AWS 0.2 MWS 0.1		MWS and AWS in a garden in the town	MWS (MSDEFCC) manual station Zeal thermometer at maximum with mercury and a Zeal thermometer at minimum with alcohol AWS (ENCA) automatic weather stations with a Vaisala thermistor, model HMP35C

P overlapping period; N, number of years; MWS, manual weather station; SIMN or MSDEFCC, elevation [m] asi; AWS, automatic weather station; ARPA or ENCA, elevation [m] asi; D, distance between the pair of weather stations [m]; DE, difference in elevation [m]; MWS-AWS; Tx, mean values of maximum temperature in the selected period for automatic weather station, AWS, and manual weather station, MWS; Tn, mean values of minimum temperature in the selected period for automatic weather station, AWS, and manual weather station, MWS; NA, missing data in the overlapping period (%); Break, year and month when RHtestV4 identified a break in the monthly difference series for maximum temperature, Tx, and minimum temperature; Tn; Location, brief description of the locations and instruments utilized in the two networks

within a distance of 20 km. Moreover, the missing data were identified directly on the original datasets. The series with more than 20% of daily missing data in the analysis time period were discarded (Giaccone et al. 2015). The QC identified only a small number of errors. For example, in Piedmont, some days were flagged as outliers but the verification process confirmed their correctness. In fact, these suspicious values were affected by foehn episodes that increase the temperatures in some valleys (Fratianni et al. 2009; Fazzini et al. 2004).

Following the QC on the series, a historical research and a homogeneity test were carried out to identify any break or discontinuity in the series. Among the many homogenisation methods found in the literature (e.g. Alexandersson and Moberg 1997; Peterson et al. 1998; Della-Marta and Wanner 2006; Auer et al. 2007; Mestre et al. 2011), we have chosen the RHtestsV4. We have employed this method because it works for periods of less than 20 years and it can carry out with or without the use of a reference series (Fortin et al. 2017; Wang and Feng 2013). RHtestsV4 was used to detect the breaks (shifts or change points) in our candidate series without the presence of the reference series. The test is based on the penalized maximal F tests which allows the time series being tested to have a linear trend throughout the whole period of the series with the annual cycle, linear trend and lag-1 autocorrelation of the base series being estimated in tandem through iterative procedures, while accounting for all the identified mean shift period of data record (Wang et al. 2007; Wang 2008a; Wang 2008b). This test was applied on monthly series.

The series with a discontinuity in the overlapping period were removed from the study. The RHtestsV4 carried out on the series did not show breaks or discontinuities in the overlapping period.

2.2.1 Piedmont dataset

In Piedmont, 4 pairs of stations were selected (Table 1 and Fig. 1) within the two independent networks, SIMN and ARPA.

The SIMN stations used a TM26000 thermograph or a thermometer at maximum with mercury and at minimum with alcohol. In the TM26000, the temperature-sensing element is a bimetallic lamina. The deformations of the lamina are transmitted to a recording system that writes on diagram paper. Every degree of temperature variation is equivalent to a shift of 1.5 mm on the diagram. The measure range is 55 °C with a precision of $\pm 1\%$ (± 0.55 °C) for the entire scale. The thermometer with mercury is a direct reading. The scale is in degrees Celsius, °C, with an error of ± 0.5 °C.

The two instruments are located in a Stevenson Screen at a height of 1.5 m from the ground.

The ARPA stations have a TU20AS CAE thermo-hygrometer. In this instrument, the sensing elements consist of an electrical resistance with a resolution of <0.02 °C and a precision of ± 0.2 °C on the whole scale, from -40 to $+60$ °C. Two of the stations, Piedicavallo and Locana, use a TA20AS thermometer with a resolution of <0.02 °C and a precision of ± 0.3 °C on the whole scale, from -40 to $+60$ °C. The instruments are located a height of 2 m from the ground.

2.2.2 Gaspé Peninsula dataset

Two pairs of stations were selected in the Gaspé Peninsula (Table 1 and Fig. 1). They belong to two networks: ENvironment CANada (ENCA) and the Ministry of

Sustainable Development, Environment and Fight against Climate Change (MSDEFCC 2017). The MSDEFCC and ENCA networks use either manual stations with observers or automatic stations. The MSDEFCC manual station use a Zeal thermometer at maximum with mercury and a Zeal thermometer at minimum with alcohol with a precision of ± 0.05 °C on the whole scale. The instruments are located in a Stevenson screen, 1.5 m from ground (Lepage and Bourgeois 2011; Environment Canada 2018). The MSDEFCC automatic stations use a Vaisala thermistor (432A HANDAR) with an accuracy of ± 0.2 °C on the whole scale, from -40 to $+60$ °C plugged into a data logger. The ENCA automatic weather stations recorded hourly maximum and minimum temperatures with a Vaisala thermistor (Model HMP35C, with an accuracy of ± 0.4 °C on the whole scale from -40 to $+60$ °C) installed in a solar shield and plugged into a data logger to record data.

A metadata file was created for each pair of weather stations. In the files are the location of the weather stations, the latitude, the longitude, the altitude, the type of instrumentation, any changes in location or in instrumentation (metadata), the difference in elevation, the distance between the two stations and the overlapping period. Photographic documentation and maps with different levels of detail were used to complete the information.

In this study, the automatic ARPA and ENCA series were called automatic weather station, AWS, because the two networks are more recent while the manual SIMN and MSDEFCC series were called manual weather station, MWS, because the two networks are older.

All temperature measurements were performed by a proxy and correlated with the reference measurement regardless of whether the measurement was from a maximum thermometer with mercury and a minimum thermometer with alcohol or measurements from a resistance thermometer (TM26000, CAE TU20AS, 432A HANDAR, HMP35C). Calibration, under highly controlled conditions, leads to a comprehensive assessment and definition of component balances to assess uncertainties as well as overall values, depending on the type of sensors used. When used in the field, a thermometer used to measure air temperature actually measures the instantaneous mixture of convective, radiation and conduction heat transfer. All these thermodynamic effects are difficult to correct and the measurement of daily fluctuations in air temperature is generally unstable. Indeed, sensor dynamics can introduce differences due to inertia and delayed response and may not be taken into account (Bertiglia et al. 2014; Grykalowska et al. 2015).

On the other hand, there are also two broad categories of instruments: those where air is artificially aspired, used by AWS, and those where air is not aspirated, used by MWS. For those where air is artificially aspirated, the measurements indicate a considerably lower sensitivity for the most frequent weather conditions, provided that they are adequately protected against direct and indirect radiative effects. They may also tend to read slightly higher temperatures during the day due to imperfect radiation or thermal contact protection and measurements that are slightly lower during the night due to the cooling effects of condensation of the aspired air. The range of measured values, whether aspirated or not aspirated, probably gives different error characteristics according to the meteorological networks. Nevertheless, these errors, associated with the instruments, essentially depend on the climate. We therefore decided not to dissociate them from our analysis (Thorne et al. 2016).

2.3 Methods

The approach implemented for the comparison between the station pairs is divided in four parts: (1) statistical characterization of the daily series of maximum (Tx) and minimum temperature (Tn) and the QC on the difference series; (2) verification of the homogeneity of the difference series, MWS-AWS; (3) comparison between the monthly series of Tx and Tn; and (4) seasonal comparison between the type of events, heat wave, cold wave and normal events.

2.3.1 Statistical characterization

In the first step, for each raw series and for the daily difference series, MWS-AWS, we provide descriptive statistics: mean, median and first and third quartiles, and minimum and maximum values were calculated and missing values were identified.

For each pairs of series, the *t* test, the Kolmogorov-Smirnov test and the Wilcoxon's rank-sum test were carried out (Guenzi et al. 2017). The *t* test and the Wilcoxon's rank-sum test (Lejeune et al. 2015) explore the statistical similarity of the means and the Kolmogorov-Smirnov test shows if both series are drawn from the same distribution (Pauli et al. 2007). For all the tests, we used a $p = 5\%$ significance level.

Furthermore, the root mean squared error, RMSE, and the correlation coefficient by Kendall's method were calculated for each pair of series. The RMSE is interpretable as a typical error between the two recordings, and the correlation coefficient measures the association between the two series (Vincent et al. 2018).

For the daily difference series, Dif, we selected only the values ranging between ± 5 °C. Differences greater or lesser than ± 5 °C are considered excessive between the pairs of series and show an error in the measurements of the temperature in one station, be it manual or automatic (Squintu et al. 2019).

To ensure a fair comparison between the members of each pair, if missing values were found in one series, they were then also removed from its counterpart (Hubbard and Lin 2006; Acquotta et al. 2016). We also removed the days with differences greater or lesser than ± 5 °C to avoid the use of incorrect values.

2.3.2 Breaks in the differences

In the second step, we checked the homogeneity of the monthly differences series, MWS-AWS, by RHtestsV4 without a reference series (Wang and Feng 2013) to show any bias between the pairs of series. If a difference series had a break, the pairs of series were divided into two periods, before and after the break. Step 3 and step 4 were repeated in the new periods and the results, before and after the break, were compared.

2.3.3 Monthly comparison

In the third step, the monthly relative ratio (RR), for each pair of series, was calculated, as well as

$$RR_{i,j} = |MWS_{i,j}|/|AWS_{i,j}|$$

$MWS_{i,j}$ monthly values for manual weather station
 $AWS_{i,j}$ monthly values for automatic weather station
 where i indicates the month and j indicates the year.

Values of $RR > 1$ indicate an overestimation of monthly minimum (maximum) series of the MWS, $0 \leq RR < 1$ highlights an underestimation of monthly minimum (maximum) series of the MWS while with $RR = 1$, the pairs of series recorded the same monthly temperature. If $|AWS_{i,j}|$ is equal to 0.00 °C, we established that RR equals missing data, NA. The trend was also calculated for the monthly RR series. The slope was calculated to establish if the ratio between the pairs of series increases, decreases or is constant. The trends were calculated using Yue Pilon's method (Yue et al. 2002). The slopes were estimated with the TSA, Theil-Sen Approach. The Mann-Kendall test was then applied to assess the significance of the trend. For the pairs of series with a break in the differences, the trend was calculated only for the periods greater than 5 years.

2.3.4 Seasonal class analysis

In the fourth step, the seasonal class analysis, the daily series were divided into 5 classes: extreme cold (EX_C), cold (C), mean (M), warm (W) and extreme warm (EX_W), by percentiles (Table 2). The class analysis was carried out on a seasonal scale. The daily series were divided into 4 seasons: winter (W), January, February and December; spring (Sp), March, April and May; summer (S), June, July and August; and autumn (A), September, October and November. For each season, the percentiles were estimated by combining the daily series pairs (manual and automatic).

For each series and for each season, the frequency of values falling in each class, obtaining the proportion corresponding to each sensor (MWS and AWS), was calculated.

3 Results

We have selected 6 pairs of stations: 4 in Piedmont and 2 in Gaspé. Overlapping period ranges between 6 years for Bra and 14 years for Cumiana. On average, the missing values are 6% of the daily data. The distance, on average, is 2.6 km, and the difference in elevation, on average, is 17 m.

The 6 pairs of stations were divided into two groups: station pair with breaks in the monthly difference between MWS and AWS and station pair without breaks (Table 1).

Table 2 Ranges of the five temperature classes

Class	Range
EX_C	$X \leq 05\text{th } p$
C	$05\text{th } p < X \leq 20\text{th } p$
M	$20\text{th } p < X \leq 80\text{th } p$
W	$80\text{th } p < X \leq 95\text{th } p$
EX_W	$95\text{th } p < X$

The thresholds were estimated for each season from a single distribution created by joining the daily manual and daily automatic series for maximum (Tx) and minimum (Tn) temperatures. With X, we indicate the temperature (Tx or Tn) series.

3.1 Pairs of weather stations with breaks

RHtestsV4 carried out on the monthly difference series showed a break for Tx and Tn in only one location (Bra), a break for Tx in two locations (Cumiana and Vercelli) and a break for Tn in Cap-Chat-2-Cap-Chat (Table 1).

In Bra, the break showed a greater difference in the maximum temperature between the two periods. The mean difference ranges from 1.39 °C in the first period to 1.03 °C in the second period and the RMSE is greater than 1.5 °C in the two periods (Table 3). In the first period, 1998 to 2001, the automatic stations underestimate the events classified as warm, W, and extremely warm, EX_W, for all the seasons. The greater differences are recorded in autumn followed by summer. In autumn, the manual station recorded 27 extremely warm events while the automatic station recorded only 5. In the second period, 2002 to 2003, the automatic series underestimates the events classified as extremely cold, EX_C, for all the seasons, but the differences in the W and EX_W events are reduced, particularly in spring and summer when the AWS overestimates the events (Table 3).

For the Tn, in the two periods, the mean difference ranges between -0.50 and -0.32 °C, and the RMSE is near 1.0 °C. The comparison between the two stations shows little change in the two periods. The AWS underestimates the events classified as EX_C and C. The difference between the two periods increases in the winter and the spring for EX_C events, and reduces in the summer and the autumn.

The locations where there is only a break for Tx (Table 3) show an increase in the difference during the second period. This increase is found in all the variables. For example, the mean ranges are between -0.45 °C for Cumiana and -0.60 °C in Vercelli during the first period, while in the second period, the mean ranges are between -1.13 °C in Vercelli and -1.51 °C in Cumiana. In the class analysis, the greater differences are found in the extremely cold and extremely warm events. For Cumiana, the differences increase in particular in the extremely warm events and in the extremely cold events. For the EX_C in the first period, the automatic station recorded a lesser number of events, on average -21% , while in the second period, it was -54% of events. For the EX_W, the behaviour changes. In the first period, the automatic station recorded $+40\%$ of events, while in the second period, it was $+69\%$ on average (Table 3). Also in Vercelli, the main difference between the two periods is in the EX_C and EX_W. The greatest differences were recorded in winter, during the second period. The AWS recorded 16 EX_C events while the MWS recorded 39 events.

For the Cap-Chat-2-Cap-Chat pair where the break is only for Tn, in the second period, from 2011 to 2014, the difference increases (Table 3). The greater difference is in the mean that ranges between -0.55 and -1.08 °C while the RMSE changes from 1.82 to 2.03 °C. For the class analysis, the greater change is for the events classified as EX_W. In the first period, the AWS recorded fewer events in winter, spring and autumn, while in the second period, it was only in the winter. For the events classified as EX_C, the AWS recorded a lesser number of events in all the seasons except for autumn in the second period, where the two stations recorded an equal number of EX_C events.

3.2 Pairs of weather stations without breaks

The pairs of stations without breaks are 3 for Tn, Cumiana, Oropa and Bonaventure-New Carlisle, and 3 for Tx, Oropa, Bonaventure-New Carlisle and Cap-Chat-2-Cap-Chat. It was possible to select pairs of stations, Cumiana and Cap-Chat-2-Cap-Chat, with breaks in Tx or in

Table 3 Root mean squared error (RMSE), median of the difference series, mean of the difference series, for minimum temperature (Tn), and maximum temperature (Tx), in the two intervals, period

Tx	Period	RMSE	Median	Mean	Season	Ex_C	C	M	W	Ex_W	RR
Bra	1998–2001	1.83	1.40	1.39	W	27	39	0	-61	-83	0.0001
					Sp	15	7	4	-13	-83	
					S	52	34	11	-129	-375	
Bra	2002–2003	1.61	0.80	1.03	A	10	22	8	-27	-440	
					W	-25	17	14	-33	-400	
					Sp	-20	38	-5	7	-117	
Cumiana	1990–1993	1.53	-0.20	-0.45	S	36	49	-11	-57	11	
					A	-75	18	7	-21	-25	
					W	-83	-40	-9	46	53	
Cumiana	1994–2003	2.00	-1.50	-1.51	Sp	24	-6	-2	-2	17	
					S	-20	6	-7	21	33	
					A	-6	-41	2	9	59	
Vercelli	1993–1997	1.28	-0.60	-0.60	W	-112	-38	-9	46	67	
					Sp	-9	-37	-6	25	63	
					S	-50	-67	-9	56	70	
Vercelli	1998–2003	1.50	-1.10	-1.13	A	-44	-25	-12	37	76	
					W	-93	15	-4	11	27	
					Sp	-6	-34	-2	23	36	
TN	1998–2001	1.09	-0.70	-0.50	S	-69	-37	-13	46	79	
					A	-25	0	-1	9	4	
					W	-144	-43	12	4	35	
Bra	2002–2003	0.95	-0.40	-0.32	Sp	-61	-23	0	17	47	
					S	-90	-72	-12	56	80	
					A	-67	-4	1	-7	55	
Bra	1998–2001	1.09	-0.70	-0.50	W	-92	-28	6	3	47	
					Sp	-40	-17	-3	24	43	
					S	-71	-10	0	25	6	
Bra	2002–2003	0.95	-0.40	-0.32	A	-79	20	-3	-6	45	
					W	-200	-22	6	38	-25	
					Sp	-86	0	3	-8	42	

Table 3 (continued)

Period	RMSE	Median	Mean	Season	Ex_C	C	M	W	Ex_W	RR
Cap-Chat-2-Cap-Chat 2008–2010	1.82	-0.40	-0.55	S	0	-57	6	4	63	
				A	-75	7	1	0	27	
				W	-36	-11	7	15	-150	
				Sp	-150	16	5	-5	-15	
Cap-Chat-2-Cap-Chat 2011–2014	2.03	-0.90	-1.08	S	-186	0	4	-8	56	
				A	-15	-2	-6	30	-15	
				W	-50	2	7	-7	-36	
				Sp	-40	-23	-1	27	11	
				S	-46	-25	-17	52	64	
				A	0	-23	-10	33	48	

The percentage of events overestimated (+) or underestimated (-), italicized table entries, indicated in grey, in respect to the automatic series calculated for the five classes: extreme cold (EX_C), cold (C), mean (M), warm (W), extreme warm (EX_W) and in every season: W winter, Sp spring, S summer and A autumn. For periods greater than 5 years, the trend of relative ratio, RR, was estimated. Table entries in bold indicate statistically significant values at 95% level

Tn, because the MWS used two different instruments: a thermometer with mercury for Tx, and a thermometer with alcohol for Tn. Vercelli was deleted because the MWS and AWS used a single instrument for Tx and Tn.

For the series of minimum temperature, Tn, the comparison shows values of RMSE greater than 1.4 °C. The maximum value, 2.24 °C, was calculated for Bonaventure-New Carlisle. As for the mean difference, it shows lesser values. The mean ranges between -0.47 °C for Cumiana and 0.52 °C for Oropa. The different results between RMSE and mean values are due to the events classified as M in the series pairs that correspond to 60% of the daily data.

In the class analysis for EX_C, extreme cold, in Oropa and Bonaventure-New Carlisle, the automatic stations recorded a greater number of events while in Cumiana, they underestimated the events, especially in the winter when the AWS recorded 21 EX_C events while the MWS recorded 90 EX_C events. For EX_W, class, the AWS underestimates the events except for the Bonaventure-New Carlisle station pair (Table 4).

For the Tx, the RMSE has a range between 1.87 and 2.17 °C while the mean varies between -0.75 and -0.49 °C. Only in Cap-Chat-2-Cap-Chat did the statistical tests show the same mean and the same median for the pairs of series (Table 4). The trend of RR exhibits statistically significant slopes in one location, Oropa (Table 4). It identifies gradual overestimations of maximum temperature by the automatic station.

In the class analysis, the greater differences are calculated for EX_W events. In Oropa and Cap-Chat-2-Cap-Chat, the AWS overestimates the number of events, especially in autumn for Oropa, +75%, and in summer for Cap-Chat-2-Cap-Chat, +58% (Table 4). In C and M classes, the AWS underestimates the events (Table 4).

In Table 5, we report the mean season values calculated from the monthly median values for maximum and minimum difference series (MWS-AWS). For Tn and Tx in the major of seasons, the AWS overestimate the temperature except in Oropa for Tn. In Oropa, the AWS underestimates the Tn in all season (Table 5). This behaviour is identified in the season class analysis where in W and Ex_W, the MWS recorded a greater number of events, in average +32% of events, while in Ex_C and C the, MWS underestimates the events, in average -22% of events. For Tn, the greater differences are calculated in Cumiana in winter. The difference is equal -1.30 °C (Table 5). For this location, the class analysis shown an overestimation for AWS in M, W and Ex_W class that corresponding 85% of events. For Tx, the automatic series overestimates the maximum temperature in the major of seasons and locations except in Cap-Chat-2-Cap-Chat in winter (Table 5). In Cap-Chat-2-Cap-Chat, the manual stations recorded greater Tx only in the winter. For this location, the class analysis shows an overestimation of events for MWS in M and W class (Table 4) that correspond 76% of analysed events. For Tx, the greater difference was calculated in Cap-Chat-2-Cap-Chat in summer, -1.60 °C, following by Oropa in winter, -1.32 °C.

In the two locations, Cap-Chat-2-Cap-Chat and Oropa, the AWS recorded a greater number of events classified as W and EX_W and underestimated the events classified Ex_C, C and M (Table 4) underlining that the AWS overestimated the maximum temperature. These differences are also highlighted by the RMSE (Table 4).

To highlight the importance of the difference between the two pairs of stations, we calculated four indices by RCLimDex (Zhang and Yang 2013) for the daily temperature series of Oropa AWS and Oropa MWS. The daily series starts on 1937 January 01 and finishes on 2002 December 31. For Oropa AWS, the series has a break in 1995 when the automatic weather station was added, while for Oropa MWS, the series has no break

Table 4 Root mean squared error (RMSE); median of the difference series; mean of the difference series; trend of the relative ratio, RR; *t* test; K-S test; the Wilcoxon's test, (W); test calculated on the pairs of stations for minimum and maximum temperature

Tx	RMSE	Median	Mean	<i>t</i> test	K-S test	W test	Corr	RR	Season	Ex_C	C	M	W	Ex_W
Oropa	1.99	-0.7	-0.75	0	0.04	0	0.97	-0.0008	W	-3	-19	-23	50	73
									Sp	38	2	-15	16	49
									S	23	-14	-16	31	62
									A	-6	-13	-12	18	75
									W	19	11	-1	-23	13
Cap-Chat-2-Cap-Chat	2.17	-0.5	-0.62	0.07	0	0.05	0.98	0.0005	Sp	-6	-6	-3	8	34
									S	-67	-46	-7	45	58
									A	22	-4	-8	19	18
									W	-7	-2	1	5	-15
									Sp	-7	-21	0	16	15
Bonaventure-New Carlisle	1.87	-0.4	-0.49	0.06	0.01	0.06	0.99	0	S	-94	-17	4	11	29
									A	22	-1	-2	4	-7
									W	17	20	1	-27	-54
									Sp	51	16	-3	-20	-48
									S	27	20	1	-36	-32
Cumiana	1.79	-0.3	-0.47	0	0	0	0.98	0.0003	A	21	9	3	-35	-10
									W	-329	-65	7	45	31
									Sp	-134	-60	22	-11	-33
									S	9	-3	5	-8	-38
									A	-71	3	6	3	-58
Bonaventure-New Carlisle	2.24	0.2	0.12	0.75	0.39	0.53	0.98	0	W	44	3	-3	-12	2
									Sp	23	-13	2	-14	16
									S	30	-15	-7	18	24
									A	-5	23	-7	1	-2
									W	17	20	1	-27	-54

All statistics are offered for minimum temperature (Tn) and maximum temperature (Tx). Table entries in bold indicate statistically significant values at the 95% level. With 0 are indicated the non-statistically significant tests. The percentage of events overestimated (+) or underestimated (-), italicized table entries, in respect to the automatic series calculated for the five classes: extreme cold (EX_C), cold (C), mean (M), warm (W), extreme warm (EX_W) and for every season: W winter, Sp spring, S summer and A autumn

Table 5 The mean season values calculated on median monthly values for minimum temperature, Tn, difference series and for the maximum temperature, Tx, difference series

	Winter	Spring	Summer	Autumn
Tn				
Oropa	<i>0.44</i>	<i>0.64</i>	<i>0.40</i>	<i>0.33</i>
Bonaventure-New Carlisle	<i>0.39</i>	-0.18	-0.25	0.06
Cumiana	-1.30	-0.50	<i>0.28</i>	-0.21
Tx				
Oropa	-1.32	-0.36	-0.58	-0.97
Bonaventure-New Carlisle	-0.18	-0.92	-0.89	-0.18
Cap-Chat-2-Cap-Chat	<i>0.26</i>	-0.85	-1.60	-0.32

The positive mean, italicized entries, equal an overestimation of the manual series, and negative median, in bold, equal an overestimation of automatic series

(Acquaotta et al. 2009). The indices calculated are summer days, SU2, annual number of days where $T_x \geq 25^\circ\text{C}$; frost days, FD = annual number of days where $T_n \leq 0^\circ\text{C}$; cool nights, Tn10p = percentage of days where $T_n > 90$ th percentile calculated on reference period, 1961 to 1990; and warm days, Tx90p = percentage of days where $T_x > 90$ th percentile calculated on the reference period, 1961–1990. In Fig. 2, we plotted the behaviour of the indices and calculated the trends by TSA. For FD and Tn10p, the trends are statistically significant and decrease but with different coefficients. Also, for SU25 and Tx90p for Oropa AWS, the trends are statistically significant and increase, but for Oropa MWS, they are not statistically significant and equal to zero for SU25 and equal to 0.05 for Tx90p (Fig. 2).

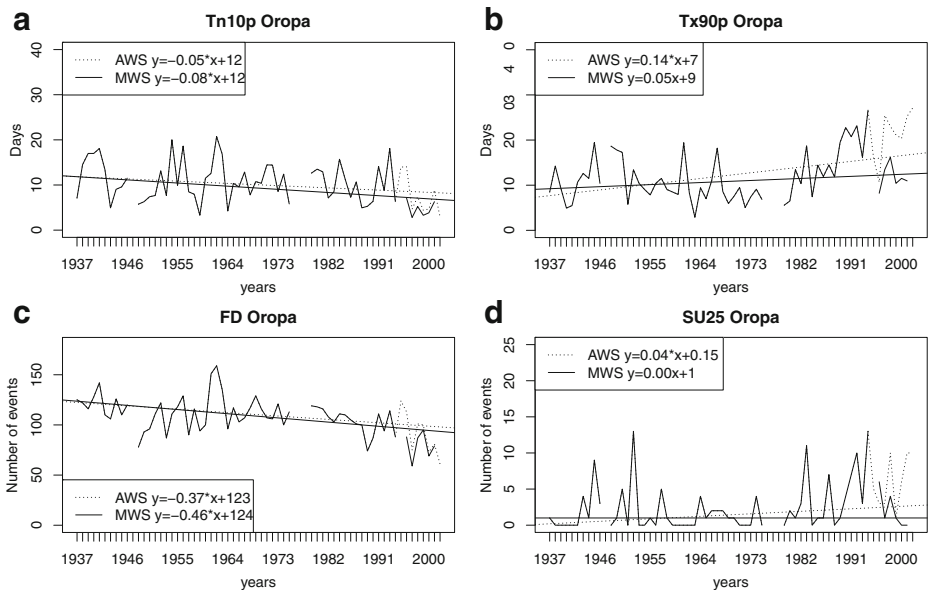


Fig. 2 The behaviour of indices calculated on Oropa AWS, dashed line, and on Oropa MWS, from 1937 to 2002. Top left cool nights, Tn10p; top right warm days, Tx90p; bottom left frost days, FD; and bottom right summer days, SU25. For each series, the linear trends are calculated, dashed line for Oropa AWS and black line for Oropa MWS. In the legend, the coefficient and the intercept are reported. The statistically significant trends are indicated by an asterisk symbol

4 Conclusion

The pairs of weather stations studied in this article were selected due to their similar characteristics in terms of exposition, surrounding environment and station history, as well as absence of breaks or discontinuities and metadata in the overlapping period.

In the first step, we selected 6 pairs of stations: 4 in Piedmont and 2 in Gaspé Peninsula.

In 4 pairs of stations (Bra, Cumiana, Vercelli and Cap-Chat-2-Cap-Chat), the RHtestsV4 identified breaks in the monthly difference series, MWS-AWS, indicating a change in the surrounding conditions or in the operation of the instruments. The historical research had not highlighted any metadata so the breaks can probably be attributed to the instruments. Only in one location, Bra, did we identify a malfunction in the manual thermometer for maximum temperature because it recorded an excessive number of extremely warm events. In the other locations, it was not possible to identify the malfunction of a specific instrument.

The analysis carried out on these locations helps us identify some common features. In all pairs of stations, the greater differences observed concerned the events classified as extremely warm, EX_W, or extremely cold, EX_C. The extreme temperature measurements seem to be more sensitive to changes.

The analysis does not show a clear relationship for the T_n. In some cases, the manual stations recorded higher minimum temperature, and in others, it was the automatic stations. It is not possible to identify a common behaviour for the class analysis either. For the T_x, the study highlighted a common behaviour. The AWS tends to record warmer values of temperature and these weather stations are more sensitive to extreme temperature measurements following the features indicated for the aspired instruments that may tend to read slightly high during the daytime due to imperfect shielding from radiation or thermal contact.

The results of this study indicate how the record of the variables can affect measured values. A change of instrumentation can create non-climatic variations in temperature recording. The greater effects of the change in these two networks (Piedmont and Gaspé Peninsula) are identified in the measurements of extreme temperatures and, consequently, in the analysis of extreme events. The use of data under- or overestimating the values of the temperatures could significantly bias the calculation of climatic indices such as the number of days where very hot or very cold temperatures are recorded, for example, as well as the trends of these indices over the time.

Furthermore, other efforts like this could help the evaluation of daily data homogenisation methods. Knowing real biases helps understand if computed biases are realistic and improve the homogenisation test.

An accurate historical research, quality control and an adequate overlapping period between two instruments can reduce the errors in the series in order to perform better climate analyses.

Acknowledgements This research was developed in the framework of the project Nexdata_Nextsnow (national coordinator, V. Levizzani; unit scientific responsible, S. Fratianni). We greatly thank the Ministry of Sustainable Development, Environment and Fight against Climate Change (MSDEFCC) (Province of Quebec) for providing the weather data for the Gaspé Peninsula area. The authors would like to thank Jeremy Hayhoe and Elisabeth Marchand for proofreading assistance.

Data availability Supplementary data and the program source code associated with this article can be found at <https://github.com/UniToDSTGruppoClima/CoTemp> (Guenzi et al. *in press*).

References

- Acquaotta F, Fratianni S, Cassardo C, Cremonini R (2009) On the continuity and climatic variability of meteorological stations in Torino, Asti, Vercelli and Oropa. *Met Atmos Phys* 103:279–287
- Acquaotta F, Fratianni S, Garzena D (2015) Temperature changes in the North-Western Italian Alps from 1961 to 2010. *Theor Appl Climatol* 122:619–634
- Acquaotta F, Fratianni S, Venema V (2016) Assessment of parallel precipitation measurements networks in Piedmont, Italy. *Int J Climatol* 36(12):3963–3974
- Alexander L, Herold N (2015) ClimPACTv2 indices and software. A document prepared on behalf of the Commission for Climatology (CCI) Expert Team on Sector-Specific Climate Indices (ET-SCI), Sydney
- Alexandersson H, Moberg A (1997) Homogenization of Swedish temperature data. Part I : homogeneity test for linear trends. *Int J Climatol* 14:25–34
- Auer I, Bohm R, Jurkovic A, Lipa W et al (2007) HISTALP: historical instrumental climatological surface time series of the Greater Alpine Region. *Int J Climatol* 27:17–46. <https://doi.org/10.1002/joc.1377>
- Baronetti A, Acquaotta F, Fratianni S (2018) Rainfall variability from a dense rain gauge network in North -West Italy. *Clim Res* 75(3):201–213
- Bertiglia F, Lopardo G, Merlone A, Roggero G, Cat Berro D, Mercalli L, Gilbert A, Brunet M (2014) Traceability of ground-based air-temperature measurements: a case study on the Meteorological Observatory of Moncalieri (Italy). *Int J Thermodyn*. <https://doi.org/10.1007/s10765-014-1806-y>
- Biancotti A, Destefanis E, Fratianni S, Masciocco L (2005) On precipitation and hydrology of Susa Valley (Western Alps). *Geogr Fis Dinam Quat* VII:51–58
- Brohan P, Kennedy JJ, Harris I, Tett SF, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850. *J Geophys Res Atm* 111(D12):1–21
- Capers RS, Kimball KD, McFarland KP, Jones MT, Lloyd AH, Munroe JS, Fortin G, Mattrick C, Goren J, Sperduto DD, Paradis R (2013) Establishing alpine research priorities in northeastern North America. *Northeast Nat* 20(4):559–577
- Changnon SA, Kunkel KE (2006) Changes in instruments and sites affecting historical weather records: a case study. *J Atm Ocean Technol* 23:825–828
- Davey CA, Pielke RA (2005) Microclimate exposures of surface based weather stations. *Bull Am Met Soc* 86: 497–504
- Della-Marta P, Wanner H (2006) A method of homogenizing the extremes and mean of daily temperature measurements. *J Clim* 19:4179–4197. <https://doi.org/10.1175/JCLI3855.1>
- Dodson R, Marks D (1997) Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Clim Res* 8:1–20
- Environment Canada (2018) Canadian climate normals 1981–2010. http://ftp.tor.ec.gc.ca/Pub/Documentation_Canadian_Climate_Normals/1981_2010/Canadian_Climate_Normals_1981_2010_Calculation_Information.pdf
- Esper J, Cook ER, Schweingruber FH (2002) Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* 295(5563):2250–2253
- Fazzini M, Fratianni S, Biancotti A, Billi P (2004) Skiability conditions in several skiing complexes on Piedmontese and Dolomitic Alps. *Meteorol Z* 13(3):253–258
- Fiebrich CA, Crawford KC (2009) Automation: a step toward improving the equality of daily temperature data produced by climate observing networks. *J Atmos Ocean Technol* 26:1246–1260
- Fortin G, Acquaotta F, Fratianni S (2017) The evolution of temperature extremes in the Gaspé Peninsula, Quebec, Canada (1974–2013). *Theor Appl Climatol* 130(1–2):163–172
- Fratianni S, Acquaotta F (2017) The climate of Italy. In: Marchetti M, Soldati M (eds) *Landscapes and landforms of Italy*. Springer Cham, pp 29–38
- Fratianni S, Cassardo C, Cremonini R (2009) Climatic characterization of foehn episodes in Piedmont, Italy. *Geogr Fis Dinam Quat* 22:15–22
- Gallo KP (2005) Evaluation of temperature differences for paired stations of the U. S. Climate Reference Network. *J Clim Notes and Correspondence* 18:1629–1636
- Giaccone E, Colombo N, Acquaotta F, Paro L, Fratianni S (2015) Climate variations in a high altitude Alpine basin and their effects on a glacial environment (Italian Western Alps). *Atmosfera* 28:117–128
- Grykalowska A, Kowal A, Szymrka-Grzebyk A (2015) The basics of calibration procedure and estimation of uncertainty budget for meteorological temperature sensors. *Met Apps* 22:867–872
- Guenzi D, Acquaotta F, Garzena D, Fratianni S (2017) CoRain: a free and open source software for rain series comparison. *Earth Sci Inform* 10(3):405–416
- Guenzi D, Acquaotta F, Garzena D, Baronetti A, Fratianni S (in press) An algorithm for daily temperature comparison: co. temp - comparing series of temperature. *Earth Sci Inform*

- Hausfather Z, Cowtan K, Menne MJ, Claude NW (2016) Evaluating the impact of U.S. Historical Climatology Network homogenization using the U.S. Climate Reference Network. *Geophys Res Lett*:1695–1701
- Hubbard KG, Lin X (2006) Reexamination of instrument change effects in the US Historical Climatology Network. *Geophys Res Lett* 33(15):1–4
- Isotta F, Frei C, Weigluni V, Tadic M, Lassègues P, Rudolf B, Pavan V, Cacciamani C, Antolini G, Ratto S, Munari M, Micheletti S, Bonati V, Lussana C, Ronchi C, Panettieri E, Marigo G, Vertacnik G (2013) The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int J Climatol* 34:1657–1675
- Karl TR, Derr V, Easterling D, Folland C, Hofmann D, Levitus S, Nicholls N, Parker D, Withee G (1995) Critical issue for long term climate monitoring. *Clim Chang* 31:185–221
- Karl TR, Arguez A, Huang B, Lawrimore JH, McMahon JR, Menne MJ, Peterson TC, Zhang HM (2015) Possible artifacts of data biases in the recent global surface warming hiatus. *Science* 348(6242):1469–1472
- Leeper RD, Rennie J, Palecki MA (2015) Observational perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) Network: temperature and precipitation comparison. *J Atm Ocean Technol* 32:703–721
- Lejeune Q, Davin EL, Guillod BP, Seneviratne SI (2015) Influence of Amazonian deforestation on the future evolution of regional surface fluxes, circulation, surface temperature and precipitation. *Clim Dyn* 44(9–10): 2769–2786
- Lepage MP, Bourgeois G (2011) Le réseau québécois de stations météorologiques et l'information générée pour le secteur agricole. Centre de référence en agriculture et agroalimentaire du Québec (CRAAQ), Québec http://www.agrometeo.org/help/le_reseau_quebecois_de_stations_meteo.pdf
- Luterbacher J, Dietrich D, Xoplaki E, Grosjean M, Wanner H (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* 303(5663):1499–1503
- Mestre O, Gruber C, Prieur C, Caussinus H, Jourdain S (2011) SPLIDHOM: a method for homogenization of daily temperature observations. *J Appl Meteorol Climatol* 50:2343–2358
- Milewska EJ, Vincent LA (2016) Preserving continuity of long-term daily maximum and minimum temperature observations with automation of reference climate stations using overlapping data and meteorological conditions. *Atmos-Ocean* 54(1):32–47
- Ministry of Sustainable Development, Environment and Fight against Climate Change (MSDEFCC) (2017) Données du Programme de surveillance du climat, Direction générale du suivi de l'état de l'environnement, Québec
- Nicholls N (1995) Long-term climate monitoring and extreme events. *Clim Chang* 31:231–245
- Nigrelli G, Fratianni S, Zampollo A, Turconi L, Chiarle M (2018) The altitudinal temperature lapse rates applied to high elevation rock falls studies in the Western European Alps. *Theor Appl Climatol* 131(3–4):1479–1491
- Pauli H, Gottfried M, Reiter K, Klettner C, Grabherr G (2007) Signals of range expansions and contractions of vascular plants in the high Alps: observations (1994–2004) at the GLORIA master site Schrankogel, Tyrol, Austria. *Glob Chang Biol* 13(1):147–156
- Peterson TC (2003) Assessment of urban versus rural in situ surface temperatures in the contiguous United States: no difference found. *J Clim* 16:2941–2959
- Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullet D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger JF, Rland E, Alexandersson H, Jones P, Parker D (1998) Homogeneity adjustments of in situ atmospheric climate data : a review. *Int J Climatol* 18:1493–1517
- Reynolds RW, Chelton DB (2010) Comparisons of daily sea surface temperature analyses for 2007–08. *J Clim* 23:3545–3562
- Rolland C (2003) Spatial and seasonal variations of air temperature lapse rates in alpine regions. *Journal of Climate* 16:1032–1046
- Squintu A, van der Schrier G, Brugnara Y, Klein Tank A (2019) Homogenization of daily temperature series in the European Climate Assessment & Dataset. *Int J Climatol* 39:1243–1261
- Stocker T (Ed.), (2014) Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press
- Terzago S, Cremonini R, Cassardo C, Fratianni S (2012) Analysis of snow precipitation during the period 2000–09 and evaluation of a MSG/SEVIRI snow cover algorithm in SW Italian Alps. *Geogr Fis Dinam Quat* 35(1):91–99
- Thorne PW, Menne MJ, Williams CN, Rennie JJ, Lawrimore JH, Vose RS, Peterson TC, Durre I, Davy R, Esau I, Klein-Tank AMG, Merlone A (2016) Reassessing changes in diurnal temperature range: a new data set and characterization of data biases. *J Geophys Res Atmos*. <https://doi.org/10.1002/2015JD024583>
- Trewin BC (2005) A notable frost hollow at Coonabarabran, New South Wales. *Aust Meteorol Mag* 54:15–21
- Trewin B (2010) Exposure, instrumentation, and observing practice effects on land temperature measurements. *Wiley Interdiscip Rev Clim Chang* 1(4):490–506

- Trewin BC, Trevitt ACF (1996) The development of composite temperature records. *Int J Climatol* 16:1227–1242
- Vincent LA, Mekis E (2009) Discontinuities due to joining precipitation station observations in Canada. *J Appl Met Climatol* 48(1):156–166
- Vincent L, Milewska E, Wang X, Hartwell M (2018) Uncertainty in homogenized daily temperatures and derived indices of extremes illustrated using parallel observations in Canada. *Int J Climatol* 38:692–707
- Wang XL (2008a) Penalized maximal F test for detecting undocumented mean shift without trend change. *J Atmos Ocean Technol* 25(3):368–384
- Wang XL (2008b) Accounting for autocorrelation in detecting mean shifts in climate data series using the penalized maximal t or F test. *J Appl Met Climatol* 47(9):2423–2444
- Wang XL and Feng Y (2013) RHtestsV4 User Manual. Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada. 28 pp. [Available online at <http://etccdi.pacificclimate.org/software.shtml>]
- Wang XL, Wen QH, Wu Y (2007) Penalized maximal t test for detecting undocumented mean change in climate data series. *J Appl Met Climatol* 46(6):916–931
- Willett K, Williams C, Jolliffe IT, Lund R, Alexander LV, Bronnimann S, Vincent LA, Easterbrook S, Venema VKC, Berry D, Warren RE, Lopardo G, Auchmann R, Aguilar E, Menne MJ, Gallagher C, Hausfather Z, Thorarindottir T, Thorne PW (2014) A framework for benchmarking homogenisation algorithm performance on the global scale. *Geosci Instrum Method Data Syst* 3:187–200. <https://doi.org/10.5194/gi-3-187-2014>
- Williams CN, Menne MJ, Thorne PW (2012) Benchmarking the performance of pairwise homogenisation of surface temperatures in the United States. *J Geophys Res Atmos* 117:1–16
- WMO (2007) Guidelines for managing changes in climate observation programmes. World Meteorological Organization, World Climate Data and Monitoring Programme series, WCDMP-No. 62, WMO-TD No. 1378, Geneva
- Yue S, Pilon P, Phinney B, Cavadias G (2002) The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol Process* 16:1807–1829
- Zhang X, Yang F (2013) RCLimDex (1.1) User Guide. Climate Research Branch Environment Canada: Downs view, Ontario

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.