Statistical Homogenization of Undocumented Monthly Temperature Data in British Columbia for Trend Analysis

by

Yaqiong Wang
BSc, Nanjing University of Information Science and Technology, 2013.

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Geography

Statistical Homogenization of Undocumented Monthly Temperature Data in British Columbia for Trend Analysis

by

Yaqiong Wang
BSc, Nanjing University of Information Science and Technology, 2013.

Supervisory Committee

Dr. Francis Zwiers, Department of Mathematics and Statistics
Co-Supervisor

Dr. David Atkinson, Department of Geography
Co-Supervisor

Dr. Faron Anslow, Department of Geography
Departmental Member

Abstract

Statistical Homogenization of Undocumented Monthly Temperature Data in British Columbia for Trend Analysis

by Yaqiong Wang

Homogenization of monthly temperature data in BC is performed for 310 monthly maximum temperature series and 307 minimum temperature series from three networks: BC Hydro, BC Ministry of Forests Land Natural Resource Operations and Rural Development (Wildfire Management Branch) and the BC Ministry of Transportation and Infrastructure. The homogenization procedure is based on a penalized maximum t-test with mean-adjustment to detect inhomogeneities and make adjustments to the data. Before homogenization, quality control is performed on 797 stations at the daily time step.

Trends at each location, in three sub-regions and across the province are analyzed based on resulting homogenized PCIC monthly temperature products. In order to measure the influence homogenization has on trends and validate the trends results calculated from the PCIC homogenized datasets, climate trends derived from the PCIC homogenized dataset are compared to those calculated from PCIC datasets without the homogenization and those from the homogenized temperature products existing in BC from ECCC[1] respectively.

The brief trend analysis components are introduced as follows. Trends before and after homogenization are compared for the averaged time series within three sub-regions based on PCIC station data. Trends based on homogenized PCIC stations and AHCCD[2] stations are also compared. In addition, spatial patterns of trends over BC are analyzed based on PCIC gridded datasets, and compared with those of CANGRD[3].

Homogenization results show that 92 out of 310 stations (29.6%) for maximum temperature and 75 out of 307 stations (24.4%) for minimum temperature have no detected changepoint, which means they appear to be homogenous. BCH has the highest portion of stations with changepoints, with 73.8% and 60.7% for maximum and minimum temperature, whereas FLNRO_WMB has the lowest portion, with 10.5% for Tmax and 27.3 % for Tmin. 80 and 81 stations have sufficient data for Tmax and Tmin variable have been analyzed for single station trend over 1990-2014. Comparing with the trends before homogenization, trends derived from homogenized PCIC stations have similar sign but smaller magnitude in general. The single station trend results are in good agreement with results of AHCCD. Spatial patterns of trends that are based on the interpolated PCIC stations also agree well with those based on CANGRD products. Warming trends predominate. Most of the seasons have distinctive positive trends across the province with exception of spring and some seasons over Vancouver Island.

---

[1] ECCC: Environment and Climate Change Canada

[2] AHCCD: Adjusted and homogenized Canadian climate data; ECCC products.

[3] CANGRD: Canadian Gridded Temperature and Precipitation Anomalies (the climate variable in this paper is only temperature); AHCCD-based ECCC products.

Table of Contents

List of Tables

List of Figures

Acknowledgements

This project was undertaken with the financial support of the Government of Canada through the federal Department of the Environment. Grant number is GCXE17M002. (Ce projet a été réalisé avec I’appui financier du gouvernement du Canada agissant par I’entremise du ministère fédéral de l’Environnemement, # GCXE17M002)

The source data from PCIC’s data portal utilized in this thesis are provided by BC Hydro, BC Ministry of Transportation and Infrastructure, BC Ministry of Forests Lands Natural Resource Operations and Rural Development Wildfire Management Branch. The data are one of the Provincial Climate Data Set that PCIC has assembled under a mandate that was negotiated as part of the Climate Related Monitoring Program of the BC Monitoring of Environment.

I would like to thank many people during the completion of my thesis. Firstly, I express true appreciation to my co-supervisors, Dr. Francis Zwiers and Dr. David Atkinson. Their profound, patient and kind guidance have shed light for this research. The homogenization work in this paper was performed with Dr. Faron Anslow, a knowledgeable climatologist at PCIC, who has contributed a great amount of work such as analyzing the methods, coding to manipulate the complex data and modifying the RClimDex and RHTests software packages. He also provided many constructive suggestions, programming support and much help that improved this thesis. I would also like to thank Lucie Vincent to be my external examiner.
In addition, I am very grateful to have been able to work with so many experts who have instructed me on how to be a good researcher, how to write academically and how to think critically. Other than the above people, I need to express great thanks to the colleagues at PCIC, classmates, advisors, and faculty of Geography department at UVic, all those who have offered sustained support and help through my graduate studies.

Lastly, I would not be able to complete this degree without my parent’s love and support.

Abbreviations

Adjusted and Homogenized Canadian Climate Dataset	AHCCD
Alberta	AB
British Columbia	BC
BC Ministry of Environment's Climate Related Monitoring Program	CRMP
BC Hydro	BCH
BC Ministry of Forests Land Natural Resource Operations and Rural Development (Wildfire Management Branch)	FLNRO_WMB
BC Ministry of Transportation and Infrastructure (manually/automatically operated stations)	MoTI (MoTIm/MoTIe)
CANGRD	Canadian Gridded Temperature and Precipitation Anomalies
Diurnal Temperature Range	DTR
ECCC products	AHCCD and CANGRD
Environment and Climate Change Canada	ECCC
Expert Team on Climate Change Detection and Indices	ETCCDI
Maximum Temperature	Tmax
Minimum Temperature	Tmin
Multiple Linear Regressions	MLR
Missing Value	NA
Northwest Territories	NT
Pacific Climate Impacts Consortium	PCIC
PCIC homogenized stations data	homog PCICstn
PCIC raw stations data	raw PCICstn
PCIC homogenized grid data	homog PCICgrid
PCIC raw grid data	raw PCICgrid
Yukon	YT

Glossary[4]

Autocorrelation: also referred to as lagged correlations or serial correlation, is the correlation of a variable with its own future and past values with different time lags. In particular, lag-1 autocorrelation of a time series is the correlation between this series and a copy of them shifted by one unit of time (Wilks, pp.57).

Base time series*: the time series from the station of interest for detecting inhomogeneity.

Changepoint*: starting time of an inhomogeneity in a time series.

Climatic anomaly: is the difference from an average/baseline of 30 or more years of the data of a certain climatic variable (*e.g.* temperature).

Climatic outliers: are the most extreme anomalies occurring within a time series for any given climatic variable. The magnitudes of an outlier usually are either greater than +3 standard deviations (SD) or smaller than -3 SD (*Hunt, 2007*).

Correlation coefficient: A measure of the linear relationship/association between two variables.

Covariance: is a measure of the joint variability of two variables.

Data adjustment*: a correction applied to data to improve its homogeneity.

Deseasonalized anomaly: climate data is usually subject to strong seasonal variations in BC. Removing the seasonality (e.g. monthly) was achieved through subtracting the climatology of that month from the temperature time series.

Diurnal temperature range*: difference between daily maximum and daily minimum temperature.

Ecoprovince: is defined as an area with consistent climatic processes, oceanography, relief and regional landforms (*Ministry of Environment, BC*).

Environmental temperature lapse rate (ELR): refers to the rate of decrease of temperature with altitude in the stationary atmosphere at a given time and location.

Gaussian distribution, also called the normal distribution. It is the probability distribution that has an arrangement of values in a bell curve with most values clustered around the probability's mean (in the middle) and the rest tail off symmetrically towards either extreme. A random variable X (*e.g.* temperature) has a Gaussian distribution with mean μ and variance σ if it has the following probability density function (Devore, 2008, *pp.145*).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)}$$

Homogeneous time series*: climate data time series where variability and change responds only to true climate variability and change and not to other biases.

Homogenization*: procedure of making a time series homogeneous by using a technique to remove artificial biases.

[4]The Glossary is compiled from various sources. Citations of concepts are given in the parenthesis; others without special denotations are from Wikipedia; concepts flagged with an asterix (*) are cited from Aguilar (2003).

Identical and independently distributed (IID): A collection of random variables is IID if each variable has the same probability distribution as the others and all are mutually independent (the realization of one random variable does not affect the distribution of the other).

***in situ* observations:** observations made as a result of direct measurement at a site.

Metadata*: the information about observational data or data about the observational data, documents how and where the measurements were made.

Quality control*: procedures used to detect erroneous observations before proceeding to the homogenization.

Reference time series*: the time series used to assess homogeneity of a base series from the target station.

Seasonality: or seasonal variations are cycles that repeat regularly over time.

Standard deviation: a measure to quantify the dispersion of an amount of data.

Sticky sensor issue: is a phenomenon when an instrument breaks and keeps recording the same data within a period of time. For example, a wind instrument might record the same wind direction information when it is stuck.

Chapter 1: Introduction

High quality climate data is the basis for accurate climate analysis. It is also crucial to understand climate variability and extreme weather events in the context of climate change. These analyses are important for societal application, including policy making, climate change assessment, and risk management. In addition, having access to good quality data from more observation networks is becoming increasingly more important as the need to understand climate change at the local scale for the purpose of regional climate change adaptations grows.

Without adjustment, weather and climate data, including that collected from land, upper air, ship/aircraft, and satellite observations, can possess a variety of potential inaccuracies. These include factors such as changes in observational surroundings, the changes in instruments or observing procedures, or station relocation. For example, vegetation changes around a station can alter the recorded temperature, while urbanization often raises temperature. A station movement that leads to a change in elevation would cause a change in temperature correspondingly based on the environmental temperature lapse rate (ELR). These sorts of disruptions and errors are termed *inhomogeneities* because they break up an otherwise consistent record. Data discontinuities caused by non-environment factors can occur at all observational time steps (*i.e.* daily, monthly *etc.*). Evidence shows that few long term climate time series from *in situ* observations are free of inhomogeneities (Peterson *et al*. 1998). Even with short records or data at short time scales, data inhomogeniety is inherent (Carrega, 2010, p16). In Canada, a noted cold bias in the means of daily minimum air temperature was introduced as a result of a nation-wide change in observing time in 1961 (Vincent, 2012). Such systematic artifacts in climate data may obscure real climate variability and change. Given, however, that reliable climate trend analysis requires data consistency, the shifts, discontinuities and errors caused by non-climatic factors need to be addressed in a process called *homogenization*, which ensures that the observed trends reflect change in the climate system rather than the observation system.

In general, there are two main approaches that may be used for homogenization: physics-based and statistical. Physics based approaches take the physical process that caused the inhomogeneity into consideration but are not commonly used for homogenization (Brönnimann, 2015, *pp.27*). For example, Auchmann and Brönnimann (2012) found that when homogenizing temperature series that were affected by a change in instrument housing, from a Wild screen to a Stevenson screen (Figure 1), the adjustments can be made through an empirical approach that accounts for differences in the energy balance of the two sensor screen systems.

Figure 1: Screens are used to shelter instruments from solar radiation. Wild screen is a metallic screen with two sides open (left); the Stevenson screen (right) is closed on all sides, and with much smaller blinds than for the Wild screen. A change of the screen system can cause an inhomogeneity (Auchmann and Brönnimann, 2012).

Numerous statistical methods have been developed for homogenization, and have been applied mostly at the annual and monthly observational frequency. One of the most widely used statistical methods is the standard normal homogeneity test (SNHT) developed by Alexandersson

(1986). SNHT is a likelihood ratio method which assumes that the tested data is normally distributed and with no trend. It locates the time point at which a single changepoint is most likely to exist. Another method for homogenization is through pairwise comparisons of monthly temperature series from the target station with a set of its correlated neighbours based on the SNHT method (Menne & Williams, 2009). This automated method can detect step changes and trend changes in a time series through examining the pairwise difference series between the record of interest and its neighbouring stations. One advantage of this method is that the neighboring stations that serve as references do not necessarily have to be known to be homogeneous. A recursive procedure attributes the cause of shifts from the target minus neighboring series after the date of the shifts are identified in the difference series. Specifically, a count is made for the dates of shift each time a station is implicated as having a discontinuity in the difference series. The station that has the highest count in the difference series is then identified as the "culprit". The date of changepoint corresponding to this highest count is then removed from all its neighbours. The process is repeated until no shift date count is greater than one. Other studies have performed comparisons among techniques to assess their ability to detect inhomogeneities in temperature data (Ducre-Robitaille *et al*., 2003; Reeves *et al*., 2007; Venema *et al*., 2012). These studies show that there is not a single method that performs best in all cases because the performance of each method depends on the time step and temporal coverage of the studied data.

In Canada, there have been several homogenization efforts on climate variables such as surface air temperature, precipitation, surface wind speed and sea level pressure at different time scales (Mekis *et al.,* 2011; Vincent *et al.,* 2012; Wan *et al.,* 2010; Wan *et al.,* 2007; Wang *et al.,* 2013). The most recent version for monthly mean surface air temperature of the Adjusted and Homogenized Canadian Climate Dataset (AHCCD) from Environment and Climate Change

Canada (ECCC) includes adjusted temperature time series at 338 locations (Vincent *et al*., 2012). AHCCD temperature data will be used as a homogenization reference in this thesis. Homogenization of monthly temperature data from ECCC is based on a multiple linear regression (MLR) based test and PMTred is based on the penalized maximal t test (PMT) to detect changepoints using the same reference series, together with a quantile-matching algorithm to make adjustments to the data (Vincent *et al*., 2012).

MLR (Vincent, 1998) is based on the application of four linear regression models to the time series of the base and neighboring stations to investigate whether a tested time series is homogeneous, if a non-climatic trend or a step exists before and/or after a step. In other words, The MLR technique identifies both step change and trend change in the series of the base station. In the study of the second generation of AHCCD, the focus of inhomogeneity is only on the shift in mean, therefore, two of the MLR models that are for the step change detection was applied and introduced here. As shown in the block quote below, Model 1 is to determine whether the tested series is homogenous and Model 3 is applied to identify the position of a step. The dependent variable in Model 1 is the time series from the station of the interest and the independent variables are series from several neighboring stations. In Model 3, an additional independent variable was used to describe the potential step in the base series. The other two models of MLR for the detection of trend change are not utilized for the second generation of AHCCD, thus not included here. Full description of MLR method is given in Vincent (1998).

Model 1: Description of a homogeneous series

$$y_i = a_1 + c_1 x_{1i} + d_1 x_{2i} + f_1 x_{3i} + e_i, i = 1,..., n.$$

y_i is the dependent variable from the base station, x_{1i}… are the independent variables from the reference series.

Model 3: Description of a step

$$y_i = a_3 + b_3 I + c_3 x_{1i} + d_3 x_{2i} + f_3 x_{3i} + e_i, i = 1,..., n.$$

$$\text{Where I} = \begin{cases} 0 \ for\ i = 4, \ldots, p-1, \\ 1 \ for\ i = p, \ldots, n-3 \end{cases}$$

I is an independent variable used to describe the step in the base series, b_3 represents the step magnitude, p is the date of a potential step change (changepoint).

(Vincent, 1998)

The PMTred (Wang *et al*., 2007; Wang, 2008a) is used in tandem with MLR for changepoints detection on the ECCC deseasonalized series of monthly means of daily maximum and of daily minimum temperature using the same reference series. PMTred is adopted as the method used for creating the homogenized PCIC station data in this thesis. Details of the implementation of the PMTred will be given in SECTION 2.2.2. Below are several advantages of PMTred (Vincent *et al*., 2012):

- PMTred accounts for the lag-1 autocorrelation, which reduces the false alarms (false detection of changepoints)
- PMTred has been demonstrated to perform well when the lengths of two segments before and after a shift have unequal lengths
- PMTred is capable of detecting multiple changpoints in an individual time series.

The homogenized AHCCD temperature data have been further used for assessing climate trends in Canada. The AHCCD stations have a long temporal coverage that extends back as far as 1840. A hundred and twenty-five (37%) of the adjusted temperature records in the second generation of the AHCCD datasets (Vincent *et al*., 2012) are from stations that have comparatively long records (the lengths range from 58 years to 128 years). The rest of the

AHCCD data were constructed by combining nearby observing sites (usually no more than 20 km apart) to create longer time series (50 to 136 years after merging).

Of the 338 locations where homogenized monthly temperature data were produced for Canada, 54 are available in BC, situated predominantly in areas of high population density. This means much of BC does not have adequate coverage, especially in higher elevation areas. For example, in the northern Rocky Mountains in BC, where few AHCCD stations located, the landscape includes mountainous regions whose snowpack serves as an important water supply for uses that include hydropower and agriculture. The changing climate will have an impact on these water resources, which requires us to have a clearer picture of how climate changes are unfolding in these areas.

There are additional data resources within BC, however. The Climate Related Monitoring Program (CRMP) operated by the British Columbia Ministry of Environment has assembled historical and current weather observations from twelve networks into the Provincial Climate Data Set (PCDS). Thousands of observing locations in the PCDS have records that go back as far as 1960 or earlier. The data will be summarized in SECTION 2.1.1. Thus, large numbers of non-ECCC stations are available in British Columbia for homogenization.

The objective of this project was to create a homogenized temperature dataset for locations across British Columbia so as to improve the accuracy of single station and spatial trend assessment. To do this requires efforts to detect and adjust non-climatic shifts of monthly temperature records for 797 stations from three main networks: BC Hydro (BCH), Ministry of Transportation and Infrastructure (MoTI) and Ministry of Forests Lands Natural Resource Operations and Rural Development Wildfire Management Branch (FLNRO_WMB), with MoTI including a manually observed MoTIm subnetwork and an automated MoTIe subnetwork (see SECTION 2.1.1). Climate trends will be analyzed from the homogenized BC dataset (station data and gridded data) and will be compared to those calculated from datasets without homogenization and with ECCC data products (AHCCD and CANGRD).

The organization of this thesis is as follows: the background for creating homogenized product based on PCIC stations is given in CHAPTER 1. Data and methodology for homogenization is provided in CHAPTER 2. CHAPTER 3 provides details of the applications of data quality control and homogenization. Based on the homogenized PCIC stations, raw observations and ECCC product, temperature trends in BC are analyzed in CHAPTER 4. Discussion and concluding remarks are given in CHAPTER 5. This thesis involves a series of statistical methods and concepts utilized as the basis for conducting the homogenization process. Therefore, simple intuitive definitions of some concepts are given in the Glossary SECTION as well as in the main text. Appendix A provides some relevant statistical methodologies.

Chapter 2: Data and Methods

2.1 Data

2.1.1 Base Data

Observational data from *in situ* stations from three networks were identified for analysis. For this study, the basic metadata such as station location information is available in the PCIC database[5]. Additional metadata, such as that describing changes in observing procedures that might be used to support the detection of changepoints, was not available at the time of conducting this analysis. It should be noted that, therefore, the homogenization efforts in this paper are based primarily on the recorded observations themselves.

The PCIC database contains data from over 7200 station locations in BC. The focus of the homogenization in this paper is on the monthly average of daily minimum and maximum temperature. The FLNRO_WMB, BCH and MoTI networks described above have the longest temperature time series of the data on the database. As shown in Figure 2, the PCIC database contains 1263 stations identified from the three networks in total. Among these, 797 stations have both Tmin and Tmax records (Table 1). They are processed in this research for quality control on the daily time step and further for homogenization analysis on the monthly scale. Specifically, the 797 stations include 84 stations from BCH, 410 stations from FLNRO_WMB, and 303 stations from MoTI. The latter network includes two sub-networks: 173 manually observed stations (MoTIm) and 130 auto-stations (MoTIe). The other 466 stations within the three networks were not suitable for analysis, therefore not included in this work because they:

- contain only once daily observations recorded at noon;
- do not directly record either Tmax or Tmin; or
- Tmax and Tmin cannot be derived from the recorded data.

[5]PCIC data portal: http://tools.pacificclimate.org/dataportal/pcds/map/

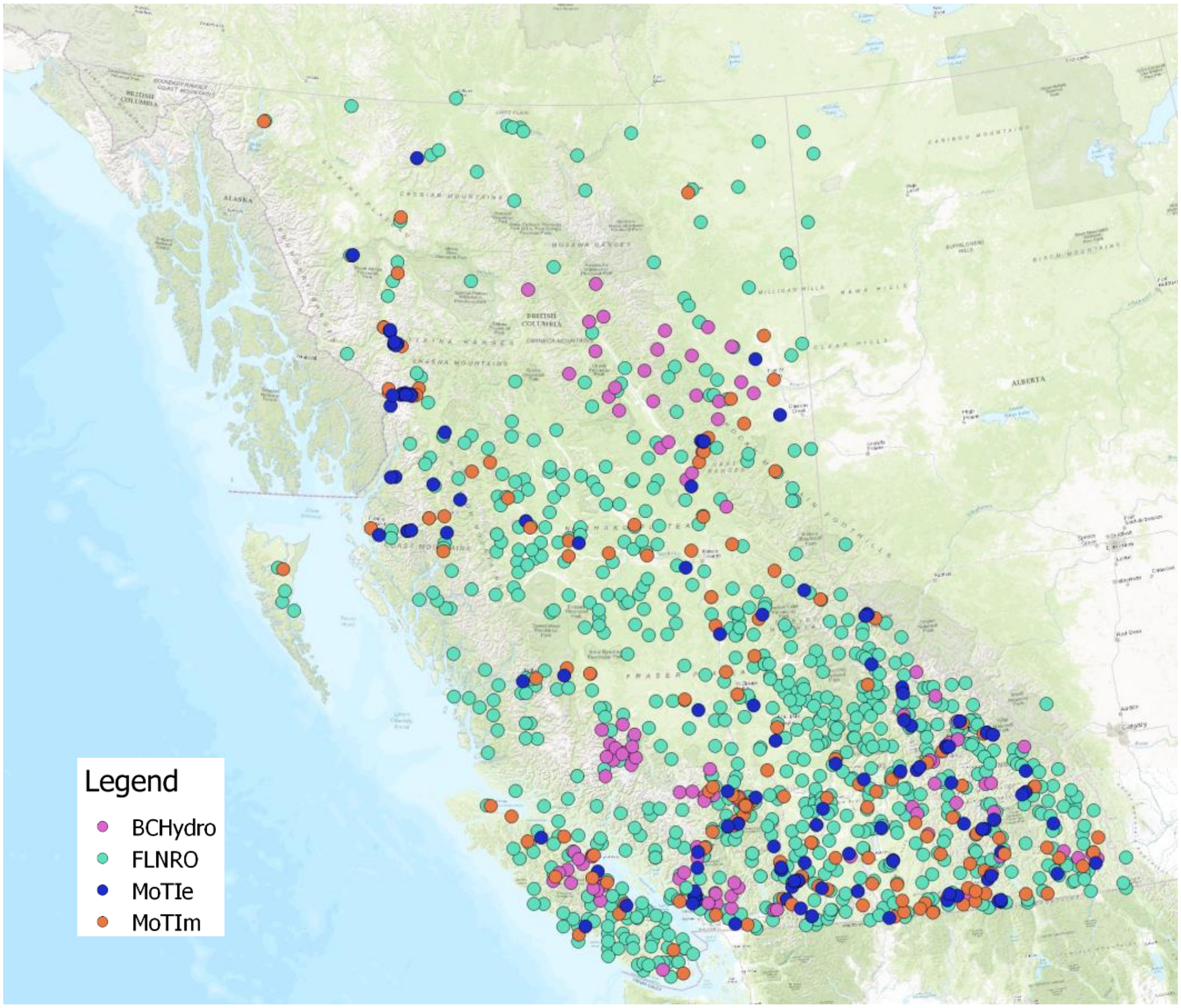


Figure 2: Study area with 1263 observational locations from three main networks across the entire province of British Columbia (Base map courtesy of ESRI)

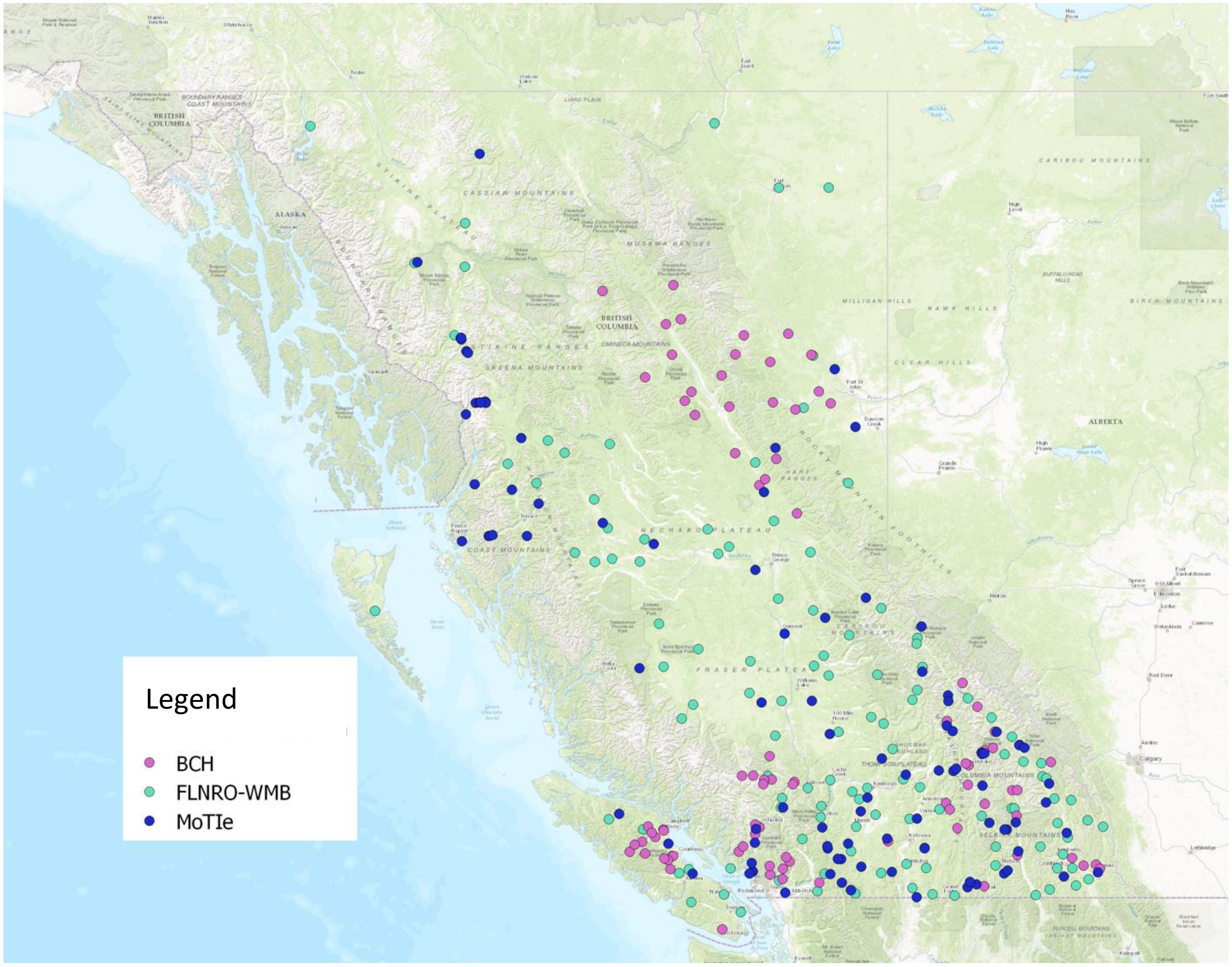


Figure 3: Study area with 307 observational locations that can be homogenized for Tmin (Base map courtesy of ESRI)

The 797 stations are the input stations were used in the quality control (QC) analysis. Homogenization is subsequently performed on these stations that can be quality controlled. The homogenization method (see SECTION 2.2) that we utilized in this project requires that stations have at least 10 years' data; 310 of the 797 stations satisfy the requirement for Tmax and 307 stations do so for Tmin. These are mapped in Figure 3. The reduction in the number of stations mainly results from the large number of short records in the FLNRO_WMB and MoTIm networks as shown in Table 1. None of the stations in MoTIm can be homogenized because they have periods of record that are too short. The efforts haven't been made for stations with short record (less than 10 years). Advancement of procedures for such data would be considered in the future work.

Network	# of studied stations	# of eligible stations	Avg. Start Date	Avg. End Date	Temporal % Cov.
BCH	84	84	1978-10-23	2016-05-27	96%
FLNRO_WMB	410	124 Tmax 121 Tmin	1995-11-11	2010-07-28	47%
MoTIm	173	0	1983-04-14	1997-03-12	39%
MoTIe	130	102	1997-02-11	2014-07-29	75%
Summary	797	310 Tmax 307 Tmin	1988-07-15	2009-06-24	64%

Table 1: Number of stations and data coverage per network for daily maximum temperature (Tmax) and daily minimum temperature (Tmin) in summary. Headers of column 1 is the network name; column 2 is the number of stations that have both Tmax and Tmin records being used in the quality control; column 3 is the number of stations that could be actually homogenized; column 4 and 5 are the average start date and end date of each network; column 6 is the temporal coverage of each network, calculation of the data availability is given in the main text[6].

Each network has its own temporal resolution and other characteristics that reflect the mission of the agency operating the network. BCH Tmin and Tmax are based on hourly instantaneous reading of the air temperature. All BC Hydro data are then converted by this network from such hourly point readings to the daily data and then stored in the PCIC's database as daily maximum and minimum temperatures. This may not reflect the true daily Tmax and Tmin (Tmin derived from hourly data will be a bit warm and Tmax will be a bit cool). Roughly from 2015 (the exact date of the transition is station specific), BCH changed its observing procedure from hourly to 15-minute sampling. This transition could induce a modest changepoint in records. The average end date for BCH data analyzed for homogenization is 2016 (Table 1), which means the potential changepoint due to the observing change would be fall near the end of the time series.

Similar with BCH's hourly collection for Tmax and Tmin, daily Tmax and Tmin from FLNRO_WMB are also based on hourly point data which is the temperature right at the time of observation. The FLNRO_WMB hourly point data were then converted into daily minimum and

[6] Data coverage calculation was performed by Dr. Faron Anslow.

maximum temperature by PCIC staff. The Wildfire Management Branch's stations are primarily operated to support fire hazard analysis. Thus, these stations are maintained for summer observations and many of them do not have all-weather instrumentation for wintertime measurements.

There are two sub-networks within the Ministry of Transportation dataset. The first is a manually observed, principally wintertime network of stations associated with winter road maintenance including avalanche control. This network will be referred to as the MoTIm network. The MoTIm network ceased operation in 2011. MoTIm observations were made twice daily with intervening minimum and maximum temperatures recorded at around 0700 and 1600 hours local time, which is a commonly used operational convention for avalanche forecasting when a uniform12-hour observing interval is challenging, according to the Manual Snow and Weather Observations produced by National Avalanche Center (*pp.*3). To convert these to daily minimum and maximum temperatures, the minimum of the two observation minimums in a given day was used for that day's minimum temperature. For maximum temperature, the greatest value between the maximum from the day's afternoon observation and the maximum from the subsequent morning's observation was chosen. This selection rule takes into consideration that Tmin of the day could happen after 1600 and Tmax of the day could happen prior to 0700 possibly due to the warm and cold fronts. However, the potential drawback of this rule is that it could underestimate Tmin if the observation of Tmin at 0700 was determined as the day's Tmin. Because this recorded value could be the Tmin of the previous day if the day was followed by a warm night for that day. Whereas, it would overestimate Tmax if the observation of Tmax at 0700 in the second day was determined as the day's Tmax when the day's afternoon or evening (between 1600 to 2400 hours) is colder than the subsequent day's night and early morning (between 2400 to 0700 hours).

The second sub-network in the Ministry of Transportation and Infrastructure's dataset consist of auto-stations (MoTIe). MoTIe hourly data is a 2-minute average taken at the hour. The Tmax and Tmin temperatures are the latest Tmax and Tmin during the 12-hour period, which resets at 0600 hours and 1800 hours. Maximum and minimum temperatures for each day are converted from hourly data for further analysis. The MoTIe network is used to monitor road weather conditions, and stations are better maintained in winter due to the importance of monitoring road snow and ice conditions.

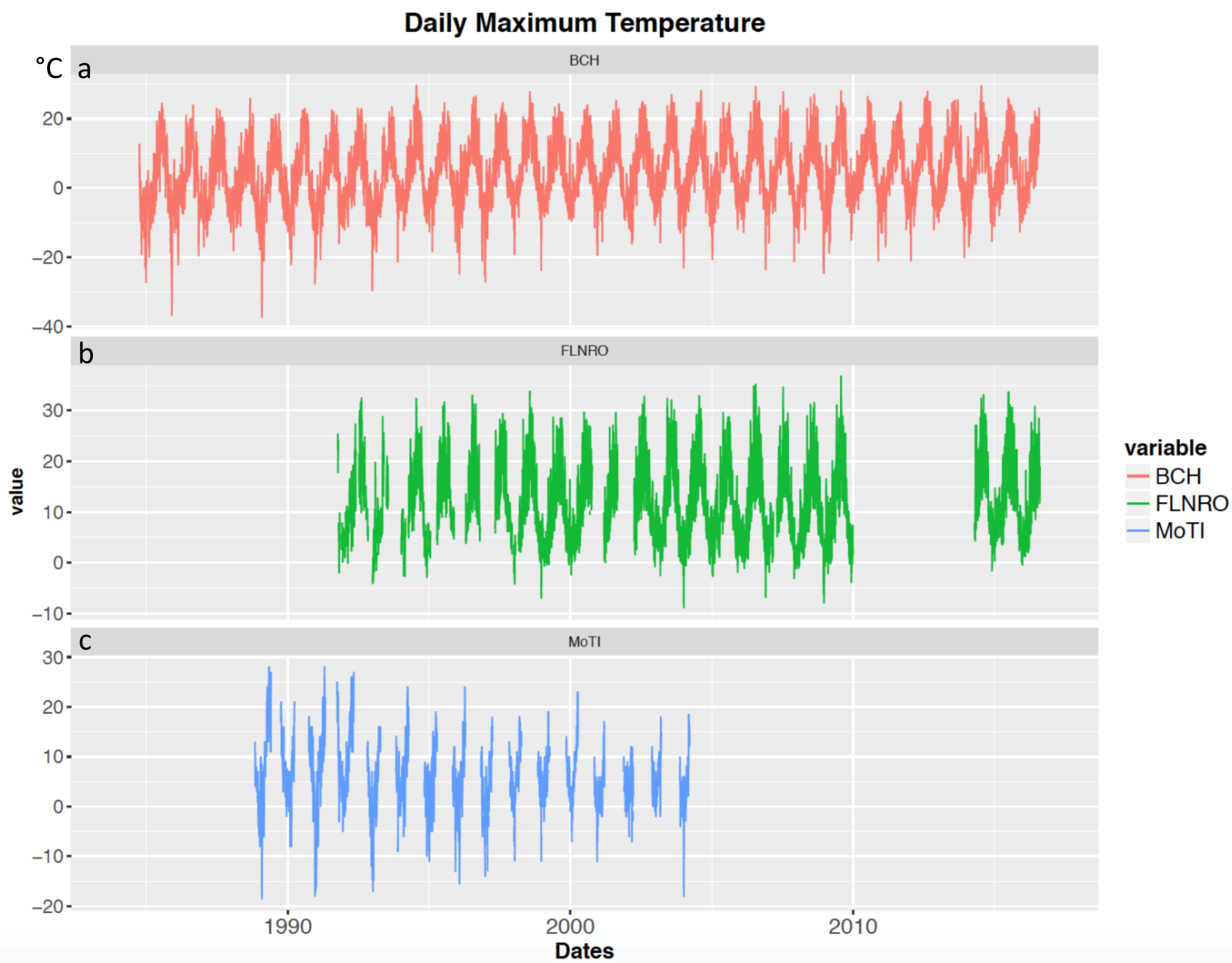


Figure 4: Sample Daily Maximum Temperature at (a) North Tyaughton Ck (ID# NTY) from BC Hydro, (b) SUMMIT (ID# 11) from FLNRO WMB, (c) Keremeos (ID# 24002) from MoTIm

To gain a more straightforward impression of the temporal data coverage for various networks, time series plots from sample stations are displayed in Figure 4, and the number of stations reporting per month is shown in Figure 5. In Figure 4a, the BCH sample station shows an overall long period of record for the station relative to other stations and that the data are continuous. Figure 4b shows a FLNRO time series of that lacks wintertime observations for part of the record and has a data gap between 2010 and 2014. Figure 4c from MoTIm shows a predominance of wintertime observations as well as the relatively short period of record for the MoTIm station that is displayed. These three examples are typical of the three networks.

The number of stations reporting per year has increased over time in general, with most of the stations beginning near from 1990 to recent (Figure 5). Early records (1960-1980) are primarily BCH data. The seasonal fluctuation in station numbers from 1975 to present reflects the service focus of FLNRO and MoTI as explained earlier. The sudden drop in the stations numbers between years 2010 to 2014 arises from the missing FLNRO data described above. Data filling this gap was added to the PCDS in April, 2017 after the homogenization had been completed and the work had proceeded to the trend analysis based on the homogenized products. The updated FLNRO data will be utilized in future homogenization work. Figures are also produced for the number of stations per year per season for all networks combined and are given in Appendix B.

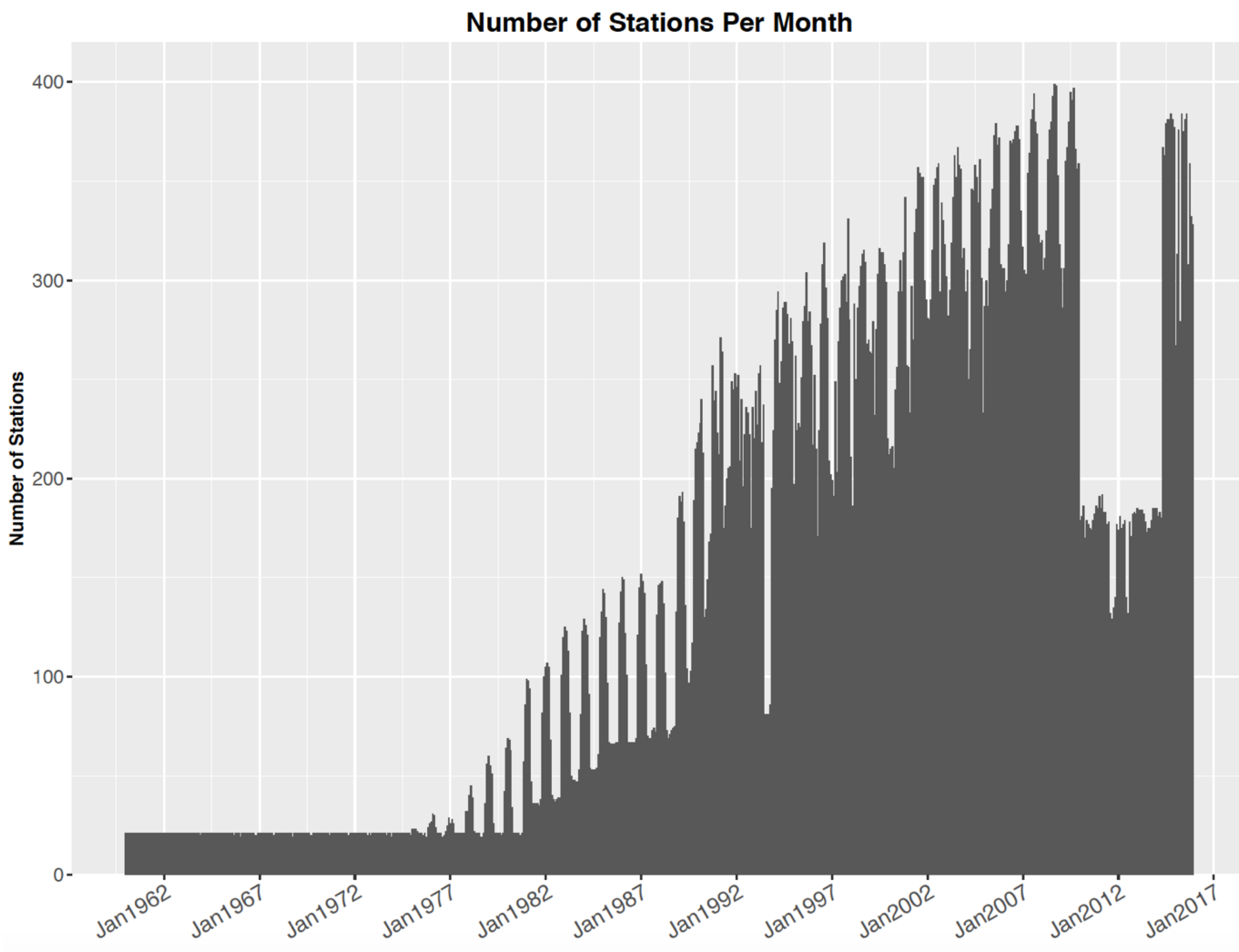


Figure 5: Number of stations reporting per month

A quantitative temporal data coverage summary is shown in Table 1, which demonstrates the average data coverage and start and end dates for each of the network.

Calculations of data coverage are briefly introduced below:

- Based on daily data, the start date and end date of each station for each variable is retrieved.
- The total number of days within this period (denote as A) and the number of days which have non-missing values for that day are counted (B).
- B divided by A is the data coverage for each station of a certain variable (C).

- The network data coverage is calculated from averaging the data coverages calculated for each station (C) and grouped by network.
- The summary of data coverage for stations from all networks is calculated through averaging C of each network.

The statistic of temperature data coverage per network shows that BCH has the longest and most complete records amongst the networks, with an average start year of 1978 and end year in 2016. The percentage of temporal data coverage for both daily maximum and minimum temperature is 96% indicating that only 4% of data are missing. MoTIe has relatively good data coverage as well. The average start year for this network is 1995 and end year is 2010. Temporal coverage is 75% for daily maximum and minimum temperature data. The summary statistics show an overall average start of 1988 and end year of 2009 for stations from all four networks and an average 36% of daily maximum and minimum temperature data are missing.

2.1.2 Reference Data

Reference data in the form of existing homogenized time series or other high quality data are usually essential to the process of homogenizing station data. For this purpose, the carefully homogenized monthly temperature time series from the AHCCD from ECCC (see CHAPTER 1) are well suited as the reference data. For each target station, AHCCD stations can serve as the high quality neighboring stations for homogenization. Their locations in western Canada are displayed in Figure 6. The map shows the density of AHCCD stations decreases with increasing latitude. More stations are located in southern British Columbia and Alberta than in the Yukon and Northwest Territories.

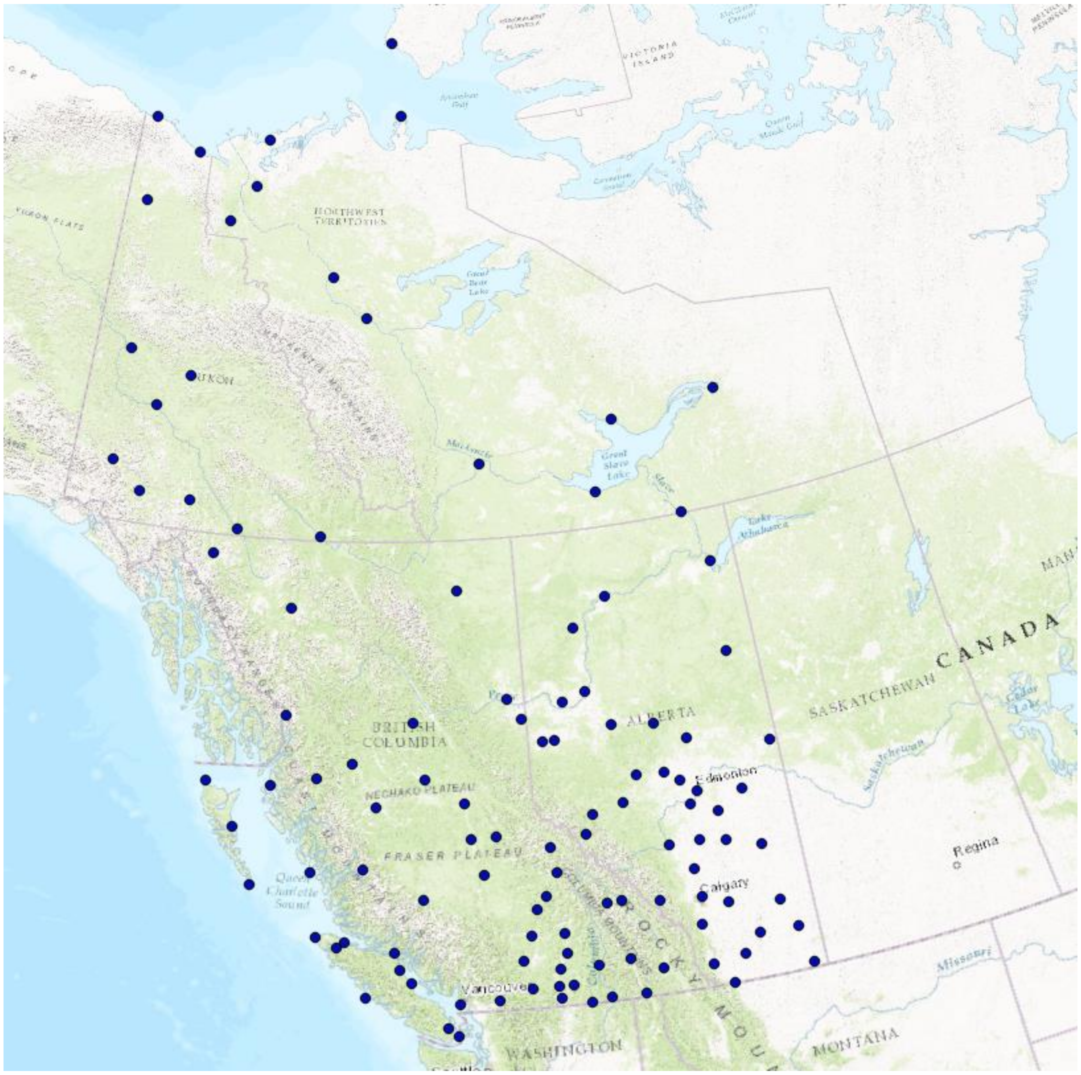


Figure 6: Map of the AHCCD locations (116) in British Columbia (BC), Yukon (YT), NT (Northwest Territories) and Alberta (AB) that could be served as potential reference stations for the homogenization process of PCIC stations in BC (base map courtesy of ESRI).

Because the AHCCD data have been carefully homogenized, any changepoints revealed in the difference between a target station and an AHCCD station likely arises due to a changepoint in the target station. But the homogeneity of the adjusted references series cannot be taken for granted since small amplitude changepoints may not have been detected in the reference series (Menne, 2008) or shifts have been detected and adjusted do not actually exist in the reference series. This may have produced small artifacts in the reference series that could be transferred to the stations that are under analysis for homogenization. The application of multiple reference stations helps to eliminate these issues. In this thesis, three references were utilized for each of the PCIC stations.

For this study, in addition to the AHCCD stations in the province of interest BC, stations in Yukon (YT), Northwest Territories (NT) and Alberta (AB) were also put into the selection pool of possible reference stations. Therefore, in total 116 AHCCD stations from four provinces (Table 2) were available for use as potential reference station. For each of the monthly Tmax or Tmin time series from the PCIC station that is of interest, three monthly AHCCD time series from the potential 116 stations were selected as the reference series. The steps that were used to select AHCCD stations for each PCIC base station is introduced as follows:

1) Based on the monthly averages of 797 daily Tmax and Tmin time series from the PCIC quality-controlled stations, monthly anomalies were calculated for each station relative to its own climatology. The whole period of record was used for calculating the climatology.
2) Monthly anomalies based on 116 AHCCD monthly temperature time series were calculated for both Tmax and Tmin in the same way as stated in step 1.
3) For Tmax and Tmin, based on its deseasonalized time series from each of the 797 PCIC station and the deseasonalized time series from each of the 116 AHCCD, correlation efficiencies were calculated.
4) Three reference AHCCD series were then selected to be those with the largest correlations and they were significant in most cases($p < 0.05$).

The application of using the difference time series is in SECTION 3.2.2.

Correlation coefficient in this thesis refers to the “Pearson product-moment coefficient of linear correlation” between two variables x and y. It ($\boldsymbol{r_{xy}}$) can be obtained as the ratio of the sample covariance of the two variables to the product of the two standard deviations (Wilks, *pp. 50*).

$$r_{xy} = \frac{Cov(x,y)}{s_x s_y} = \frac{\frac{\sum_{i=1}^{n}[(x_i - \bar{x})(y_i - \bar{y})]}{n-1}}{\sqrt[2]{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}\sqrt[2]{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}}}$$

where the bar denote means of the respective datasets. The test of the statistical significance of each correlation coefficient value utilized is the t-test at the 5% significance level: r_{xy} divided by its standard error producing a t-statistic with n-2 degrees of freedom (n is measured in months, it is the length of x, n also equals to the length of y).

To describe the correlation strength, use of absolute value of r_{xy} is suggested by Evans (1996) (Table 3). For most target locations in this research, the correlation coefficient is larger than 0.60, which means the base time series is strongly correlated with the reference time series.

0.00-0.19	very weak
0.20-0.39	weak
0.40-0.59	moderate
0.60-0.79	strong
0.80-1.00	very strong

Table 2: Strength of correlation based on its absolute value, from Evans (1996). Column 1 is the range of absolute values of the correlations; column 2 shows the descriptive strength of the correlation corresponding to its values.

When calculating correlation based on the monthly anomalies, monthly means for all parts of the year are used. Since autocorrelation frequently exists in temperature data, especially at the monthly scale, it would be desirable to take its effects into consideration. This is a complex issue that requires further research, and thus the assessment of significance in this thesis may be slightly over estimated the strength of correlation. In general, temporal autocorrelation is dealt with by calculating an equivalent sample length (Thiebaux and Zwiers, 1984), but the definition of the equivalent sample length for correlation is different from that typically used when conducting tests on the means of time series.

Within the 116 potential reference stations, 99 of the AHCCD stations for Tmax and 101 of the AHCCD stations for Tmin that were significantly corrlated were actually used in the homogenization analysis[7]. Specifically, all the 54 AHCCD stations in BC were actually used as reference for both Tmax and Tmin. Almost all stations except one station in each of YT and NT were utilized as reference stations for Tmax and Tmin. Roughly 2/3 of the AHCCD stations in AB are highly cited as reference stations for the PCIC base stations.

	Tmax		Tmin	
	# potential	# utilized	# potential	# utilized
BC	54	54	54	54
YT	11	10	11	10
NT	13	12	13	12
AB	38	23	38	25
In total	116	99	116	101

Table 3: Number of AHCCD stations in four provinces: British Columbia (BC), Yukon (YT), NT (Northwest Territories) and Alberta (AB). Column name "#potential" means the number of stations that could be served as reference stations; column name "#utilized" means number of AHCCD stations that are actually utilized as reference stations for Tmax and Tmin in each province in the homogenization analysis.

The reason of using stations from YT, NT and AB as reference stations in addition to those from BC is because this larger "pool" of data allows consideration of reference stations outside BC for locations near the border, where the geophysical characteristics and climatic conditions are similar between BC and adjacent provinces and territories.

2.2 Methodology

This section introduces the methodologies used in the quality control and homogenization process. Detailed applications and summary results of both procedures are given in SECTION 3.

2.2.1 Quality Control in Daily Temperature Records

Before the homogenization process, applying basic quality control to the climate data is necessary to avoid erroneous detection of non-climatic changepoints. Climatic outliers are the

[7] The counts of AHCCD stations that were actually utilized was performed by Dr. Faron Anslow.

most extreme anomalies that occur within a time series for a climatic variable (Hunt, 2006). In this thesis, the climatic outliers refer to outliers in the daily minimum and maximum temperatures and diurnal temperature range. During this research, quality control refers to checking for outliers and for inconsistent data, such as the maximum temperature on a given day being less than or equal to the minimum temperature on that day. The RClimDex package[8] developed by the Expert Team on Climate Change Detection and Indices (ETCCDI) was used to perform quality control. It is designed to calculate extremes indices but also contains data quality control algorithms.

To circumvent the problem of comparing extremes against standard deviations calculated from potentially error-prone data, an improved method that is resistant to outliers, called the bi-weight method (Hoaglin et al., 1983), is used for estimating variance. The bi-weight method is also used to obtain a resistant estimate of the median, rather than the non-resistant sample mean to detect outliers. The details of this method are provided in Appendix A1.

An outlier is determined by measuring its difference from the median in units of bi-weight standard deviation. A value is flagged as outlier if it falls more than four standard deviations away from the median. The outlier detection procedure based on the resistant bi-weighted estimates of median and standard deviation has been incorporated into the outlier detection module of the RClimDex software by Faron Anslow at PCIC.

The detection of outliers relies on the estimation of the variance and median of data on a given day of the year for all years in the record. Before modifications, the RClimDex package estimated this using the standard deviation from all of the years available for the calendar day of interest. First, even if a given record has numerous years of data, the number of observations

[8] RClimDex was downloaded on May 2015 from http://etccdi.pacificclimate.org/software.shtml.

corresponding to the day of year of interest will be less than or equal to the number of years of record (depending on whether there are gaps in the data). Even with complete records, this will be a small sample with no more than 60 observations in the data analyzed here and frequently many fewer. Thus, a better approach for computing variance and median appropriate to the time of year is to use a window of days surrounding the day of interest so that the sample size of climatologically similar data increases. A 15-day window (+/-7 days) is introduced centred on the day of interest for estimating the variance for the given day of the year. This allows a better estimation of the true variance and median of the data. The selection of the window involves a variance-bias trade off. For a large window size (e.g. 91-day), more data will be included in the estimation of temperature mean and variance, leading to an estimate with high precision (or small uncertainty), but at the expense of losing accuracy since the estimate indeed represent the averaging feature of the days within the window rather than the day of interest. The window size of 15 is chosen as it allows a reasonably large sample in a 60-year datasets (about 60*15=900) for estimating the variance while does not substantially affect the representativeness of the estimated variance for the day of interest. The second issue is that the standard deviation is strongly influenced by the presence of the outliers themselves. So in cases where large outliers exist in the data, the computed variance will be large and an outlier is less likely to be flagged. To take care of the falsely flagged outlier issue, the robust bi-weighted method came into play.

2.2.2 Detection and Adjustments of Mean Shifts in Monthly Time Series

Since comprehensive metadata describing changes in station recording environment in BC are not available at the time of the analysis, a robust statistical technique for detecting undocumented shifts is therefore needed. Here the term of *undocumented shifts* are referred to as shifts that are statistically significant even without metadata support (Wang *et al*., 2007). The algorithm of penalized maximal *t*-test accounting for the effect of autocorrelation (PMTred) is used to detect changepoints and the mean-adjustment method to adjust shifts at the detected changepoints (Wang 2008a; Wang 2008b). The method of PMTred has its origin from the penalized maximal *t*-test (PMT), which does not take the effect of autocorrelation into account (Wang *et al*. 2007). The PMTred algorithm is chosen as it has been shown to have good performance in detecting multiple undocumented mean-shifts in Gaussian distributed climate data series (Wang, 2008b). Temperature typically follows a distribution that is approximately Gaussian. This thesis focuses only on shifts in mean, although it should be realized that non-climatic factors may also cause changes in variance or in both. Moreover, the PMTred algorithm is formulated for time series without trend. It is therefore the base-minus-reference series that are actually involved in the test. The reference series should be chosen with great care such that it is essentially homogeneous, highly correlated with and ideally have the same trend as the base series. The RHTestV4[9] software written in the R programming language is used for the implementation of the PMTred algorithm. The RHTestV4 software was developed by researchers[10] from ETCCDI supported by ECCC.

Given the base time series $X=(x_1, x_2, \ldots x_N)^T$ and a reference time series $Y=(y_1, y_2, \ldots y_N)^T$, the base-minus-reference times series $Z=X-Y=(z_1, z_2, \ldots z_N)^T$ is obtained, with which the RHTestV4 changepoint detection procedure proceeds as follows:

1) Identify the first most probable changepoint and determine whether or not it is statistically significant at a significance level of 5%. If it is, then this probable changepoint is taken as the fist changepoint, add it to the list of identified changepoints,

[9] RHTestV4 was downloaded on May 2015 from http://etccdi.pacificclimate.org/software.shtml.
[10] Researchers who created RHTestsV4 are Xiaolan Wang and Yang Feng from ECCC.

and proceed to step 2); otherwise, stop searching and claim that the series being tested is homogeneous.

2) Divide the whole time series into two segments by the first changepoint thus detected; in each segment identify the most probable changepoint and determine if it is a statistically significant one. If it is, add it to the list of identified changepoints and go to step 3). If no changepoint is detected in neither of the two segments, stop searching.

3) Suppose now that there are M changepoints that have been detected, these M changepoints divide the whole time series into M+1 segments. Similarly, in each segment search the most probable changepoint and determine if it is a significant one. If so, add it to the list of identified changepoints, and proceed to step 4).

4) Repeat step 3) until no more changepoint can be detected in none of the M+1 segments.

After implementing the above procedure described in 1)-4), the detected changepoints are reassessed. Those are found to be insignificant are removed from the list of identified changepoints. Once all changepoints have been detected, the base-minus-reference will be adjusted following the mean-adjustment method.

Before introducing the adjustment method, the PMTred algorithm that is used in each of the above 4 steps and the reassessment step for changepoint detection is introduced first. Without loss of generality, suppose that there are M ($1 \leq M \ll N$) changepoints that have been detected, and that the goal is to search if there is another changepoint between the i-th and (i+1)-th changepoints, that is, within the i-th segment. It is noted that the changepoint searching procedure does not scan all data points within the segment, but instead exclude the first and last 5 data points excluded from the scan. This ensures at least 5 data values available for a full-model-fit, as will be evident later. For the ease of presentation, denote the *i*-th segment with the first and last 5 data excluded as Z_t^i (t = 1, 2,…, L), where L represents the length of this reduced segment; and keep referring to this reduced segment as the *i*-th segment.

The *first step* of the PMTred test is to locate the most probable changepoint within the *i*-th segment using data in Z_t^i (t = 1, 2,…, L). The most probable changepoint is the one which has the largest penalized *t*-statistic among the L-10 data points. For a given point i_k within this *i*-th segment, the penalized *t*-statistic is calculated as follows:

$$\mathrm{PT}(i_k) = \mathrm{P}(i_k)\mathrm{T}(i_k)$$

where

$$\mathrm{T}(i_k) = \frac{1}{\widehat{\sigma_{i_k}}}\left[\frac{i_k(L-i_k)}{L}\right]^{1/2}\left|\overline{Z_{1st}^i} - \overline{Z_{2nd}^i}\right|$$

$$\widehat{\sigma_{i_k}} = \frac{1}{L-2}\left[\sum_{1\le t\le i_k}\left(Z_t^i - \overline{Z_{1st}^i}\right)^2 + \sum_{i_k+1\le t\le L}\left(Z_t^i - \overline{Z_{2nd}^i}\right)^2\right]$$

with $\overline{Z_{1st}^i}$ being the sample mean of the first part of the time series Z_t^i (t = 1, 2,…, L) up to the i_k-th point being evaluated, and $\overline{Z_{2nd}^i}$ being the sample mean of the remaining part of that time series. The penalty term $\mathrm{P}(i_k)$ is designed to resolve the overly large false alarm rate of the conventional *t*-test near the ends of the time series. Wang *et al.* (2007) suggested the following penalty function,

$$P_0(i_k) = \frac{(11C^{9/8})}{200}F^{\upsilon}$$

where if $L \leq 100$, $\mathrm{F} = 1 - A^{\left(\frac{7B-2BC}{10}\right)}$ and $\upsilon = \frac{15C^{\frac{1}{2}}-11}{100}$; otherwise if $L > 100$, $\mathrm{F} = 1 - A^{\left(\frac{11BC}{50}\right)}$ and $\upsilon = \frac{2C^2+2C-1}{100}$. Quantities A, B, C, and D are given respectively by $\mathrm{A} = \left|1 - \frac{2i_k}{L}\right|$, $B = \log(L)$, $\mathrm{C} = \log(B)$, and $D = \log[\log(L + 150)]$. Wang et al. (2007) realized that this penalty function tends to cause over-penalty for data points near the ends of the time series, and suggested a modified version as follows

$$\mathrm{P}(i_k) = \begin{cases} P_0(i_C) - \Theta(i_C - i_k), & i_k = 1,2,\dots,i_C \\ P_0(i_C), & i_k = i_C + 1, i_C + 2, \dots, L - i_c - 1 \\ P_0(L - i_C) - \Theta(i_k - L + i_C), & i_k = L - i_C, L - i_C + 1, \dots, L - 1 \end{cases}$$

where

$$\Theta = \begin{cases} D^{1/2}[P_0(i_C + 1) - P_0(i_C)], & L \leq 10 \\ D^{1/3}[P_0(i_C + 1) - P_0(i_C)] + \dfrac{3}{10L^{4/3}}, & 10 < L \leq 100 \end{cases}$$

if the series length $L \leq 100$, and

$$\Theta = \begin{cases} \dfrac{[P_0(i_C) - P_0(1)]}{2i_C - 4} A^{C^3}, & i_k = 1,2,\dots,i_C \\ \dfrac{[P_0(\mathrm{N} - \mathrm{L}) - P_0(N - 1)]}{2L - 4} A^{C^3}, & i_k = L - i_C, L - i_C + 1, \dots, L - 1 \end{cases}$$

if the series length $L > 100$. In the above, the threshold location index i_C depends only on the length of the time series, and is given by $i_C = \lfloor i_\alpha/2 \rfloor + 3$ for series of length $10 < L < 50$ and $i_C = \lfloor i_\alpha/2 \rfloor + 2$ otherwise, where $\lfloor\cdot\rfloor$ is the floor function and i_α is taken as the location index at which the original penalty function curve $P_0(i_k) \sim i_k$ first crosses the horizontal line passing through 1. Following the above, a penalty *t*-statistic can be calculated for each data point in the *i*-th segment, resulting in a series of PT(t), $t \coloneqq [1, 2, \dots, L]$. As was previously mentioned, the most probable changepoint is the one with the largest penalized *t*-statistic, denoting the corresponding statistic as PT_{max}.

The M changepoints that are already in the list of identified changepoints plus the most probable changepoint identified in the *first step* divides the whole base-minus-reference time series into M+2 segments. The *second step* of the PMTred algorithm first fits a horizontal line to each of the M+2 segments, that is, the so-called "full-model-fit" in Wang (2008b), and remove the full-model-fit from the base-minus-reference, leading to a residual series R_t (t = 1, 2,…, N). The algorithm then proceeds to remove the lag-1 autocorrelation from the residual series R_t (t = 1, 2,…, N) as follows:

$W_1 = R_1$ and $W_t = R_t - \widehat{\emptyset} R_{t-1}$ for $t = 2, \ldots, N$.

where $\widehat{\emptyset}$ is the estimate of the lag-1 autocorrelation of residual series. The resulted time series W_t (t = 1, 2,…, N) is often called as the prewhitened series. Now pick out the part of the prewhitened series that correspond to the i-th segment referred in the *first step*, and repeat the computation procedures in the *first step* on this segment of the prewhitened series to compute a series of $\mathrm{PT}_W(\mathrm{t})$, $t \coloneqq [1, 2, \ldots, L]$, where the subscript indicates that the values are calculated from the prewhitened series. From $\mathrm{PT}_W(\mathrm{t})$, $t \coloneqq [1, 2, \ldots, L]$, the 1-$\alpha$-percentile is computed and denoted as $\mathrm{PT}_{max,\alpha}(\widehat{\emptyset}, L)$. Here the significance level α is chosen as 5%. In order to account for the uncertainty in the autocorrelation estimate, the RHTestV4 software computes the 95% confidence interval of $\widehat{\emptyset}$, denoted as [$\widehat{\emptyset}^L$, $\widehat{\emptyset}^U$], with which the corresponding uncertainty range of the $\mathrm{PT}_{max,\alpha}$ is computed, denoted as [$\mathrm{PT}_{max,\alpha}(\widehat{\emptyset}^L, L)$, $\mathrm{PT}_{max,\alpha}(\widehat{\emptyset}^U, L)$]. To determine whether the identified most probable changepoint in the *first step* is a significant changepoint with the influence of autocorrelation accounted for, the PT_{max} obtained in the first step is compared with [$\mathrm{PT}_{max,\alpha}(\widehat{\emptyset}^L, L)$, $\mathrm{PT}_{max,\alpha}(\widehat{\emptyset}^U, L)$]. Only if $\mathrm{PT}_{max} > \mathrm{PT}_{max,\alpha}(\widehat{\emptyset}^U, L)$, the identified most probable changepoint can be declared statistically significant at the nominal level, that is, 5%.

Now the mean-adjustment method implemented in RHTestV4 is introduced. This method first fits common trend regression lines (Wang, 2008a), one for each segment formed by the detected changes points, to the deseasonalized base series (rather than the base-minus-reference series), and then adjusts the base series in a backward way, that is, keep the segment of the most recent period as it is, adjust the segment of the second most recent period such that there is no abrupt shift between them, and then adjust the third most recent segment in the same way, and so on.

Chapter 3: Application and Results

3.1 Data Quality Control

3.1.1 Quality Control Procedures

Quality control in this work relies on two main categories for daily temperature data:

- Unreasonable data checks: require Tmax larger than Tmin;
- Outlier checks for: Tmax, Tmin and diurnal temperature range.

All original data from different networks were first reformatted to match the format necessary for RClimDex (Appendix B, Table B1). Input data preparation (to manipulate the raw data to the desired format) was therefore performed. The PCIC database unique internal station identifiers were utilized so that all data are stored in a uniform way. Variables needed in this work are maximum and minimum temperature, but precipitation observations were also accessed for the next step of the homogenization project for the ECCC grant.

The steps for converting the original data to the required daily standard format are briefly described below. The first step is developing the main dataset. In the dataset, hourly recorded data from the FLNRO_WMB and the MoTIe networks are converted into daily maximum and minimum temperature; precipitation is also stored. Twice daily, irregular interval records from

MoTIm are converted into daily Tmax and Tmin (this is described in SECTION 2.1.1). Observations from BCH were provided to PCIC at the daily time scale, so these were pulled directly from the PCIC database.

The reformatted data for each station, including daily maximum temperature (Tmax), minimum temperature (Tmin), served as the input data to RClimDex for the data quality checks. To facilitate the automation of the quality control process, RClimDex was modified to make the interactive graphical user interface optional to allow processing of large amounts of data.

The requirement that daily minimum temperature be lower than the maximum temperature is strictly enforced. This means that values that violate this rule were set to missing for both minimum and maximum temperature. For the potential outlier checks, Tmin, Tmax and DTR that are more than four standard deviations from the mean are defined as outliers. The outliers are then set to missing. The mean and standard deviation were calculated using the outlier resistant bi-weight method (CHAPTER 2).

3.1.2 Quality Control Results

The most frequently occurring issue was the occurrence of "sticky sensor" results for stations in the BCH network and multiple stations in the FLNRO_WMB network. An example is the station Horn Creek (ID# HRN) from BCH network in the Williston Basin that is shown in Figure 7. More than a month of the data have this issue where maximum temperature and minimum temperature do not vary from 8.3°C.

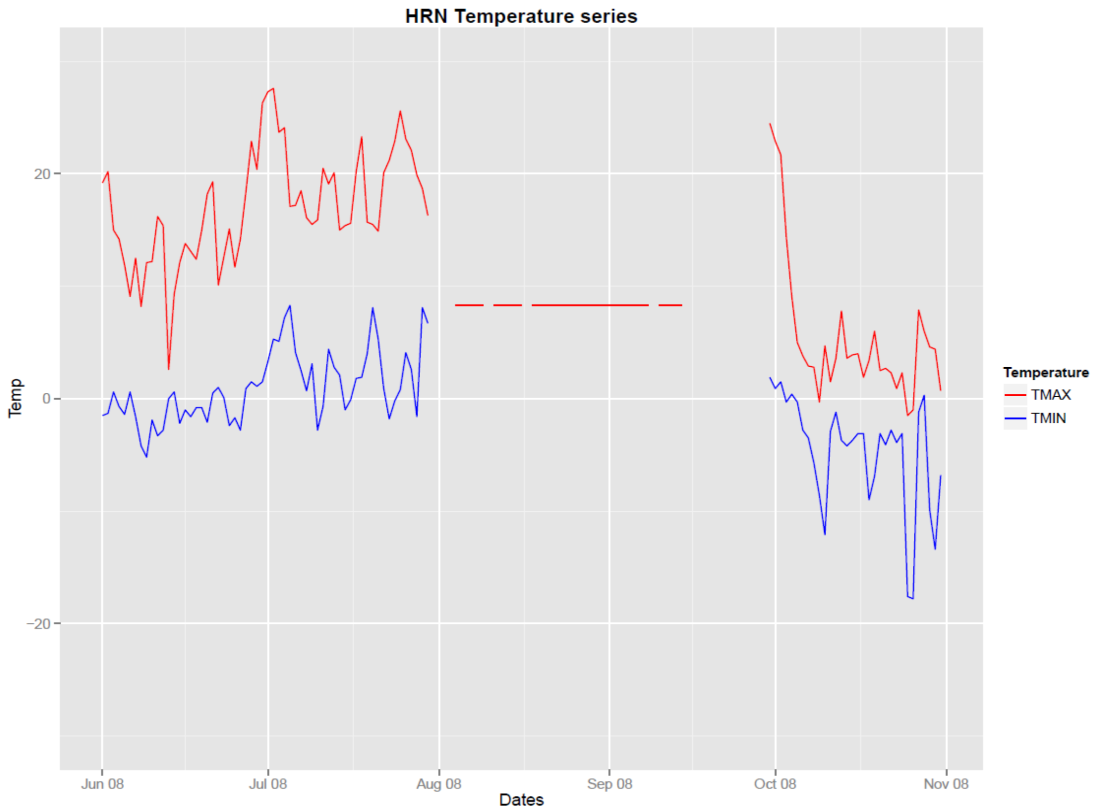


Figure 7: An example of station data with "sticky sensor" from the Williston Basin at Horn Creek (ID# HRN). The data show a more than month long period where Tmin = Tmax = 8.3 °C.

Outliers in the raw observational data were found as well. Table 4 gives an example of flagged observations from RClimDex and shows the QC results for one FLNRO_WMB station (Beaver Creek, ID# 37). The table is described below. The time series from this station, which is displayed in Figure 8, shows the outliers that are marked with red dots. The dates of the potential outliers are listed in the first columns of Table 4. The Tmax (column 3), Tmin (column 6) and DTR (column 9) represent the checked variables maximum temperature, minimum temperature and the diurnal temperature range respectively. The lower and upper bounds of the three variables are calculated from the definition of the outlier, mean values plus or minus four standard deviations. Therefore, temperatures outside of the range would be flagged and regarded as outliers. For this sample station, Table 4 indicate that a set of -20 °C observations clearly fall below the allowable 4 standard deviation range for minimum temperature (Tmin Low).

This -20 °C issue is not uncommon for the FLNRO_WMB network before early 2000. Due to data logger's intermittent malfunction within the FLNRO_WMB network, the instrument record any temperature value as -20 °C when temperature falls below that level or when sensors were poorly connected or malfunctioned. For this research, these values are set to missing once identified. To obtain a quantitative knowledge of the occurrence of this value, the number of identified -20 °C observations is divided by the total number of observations based on the FLNRO_WMB daily data before and after 2008. Before 2008, fraction of the -20 °C observations accounts for the 1.41% and 0.68% for Tmin and Tmax respectively. After 2008, using the complete records, the fraction of the -20 °C turns out to be less than 0.1% and 0 for Tmin and Tmax.[11] This suggests that the majority of -20 °C before 2008 was from data logger.

Date	Tmax Low	Tmax	Tmax Up	Tmin Low	Tmin	Tmin Up	DTR Low	DTR	DTR Up
1992-03-22	-12.23	19.4	34.04	-18.65	**-20**	20.09	-18.75	39.4	38.52
1994-05-29	-12.14	13.6	54.07	-15.15	**-20**	29.41	-24.73	33.6	52.54
1996-06-17	-8.71	16.2	50.72	-9.7	**-20**	26.4	-21.7	36.2	46.95
1996-09-01	-7.92	23.8	55.74	-10.23	**-20**	28.37	-22.44	43.8	52.22
1997-08-22	-6.26	23	56.55	-8.58	**-20**	28.58	-20.75	43	51.13
1998-08-10	-5.91	22.5	58.63	-7.16	**-20**	28.2	-19.75	42.5	51.92
2003-04-22	-13.26	14.2	52.53	-16.88	**-20**	29.58	-24.5	34.2	51.3
2004-07-09	-7.64	15.9	56.59	-7.9	**-20**	28.11	-22.75	35.9	51.66
2007-04-17	-12.81	12.8	40.88	-18.07	**-20**	23.23	-22.47	32.8	45.78

Table 4: Sample data from the FLNRO_WMB Beaver Creek station (ID# 37) showing the occurrence of a Tmin outlier and daily temperature range (DTR) outlier. Column 1 is date of the identified outliers; column 3, 6, 9 are daily maximum temperature (Tmax) daily minimum temperature (Tmin) and daily temperature range (DTR); column 2 and 4 are the mean of total daily maximum temperature minus and plus 4 standard deviation; column 5 and 7 are the mean of total daily minimum temperature minus and plus 4 standard deviation; column 8 and 10 are the mean of total daily temperature range minus and plus 4 standard deviation.

[11] The fraction calculation of -20°C was performed by Dr. Faron Anslow.

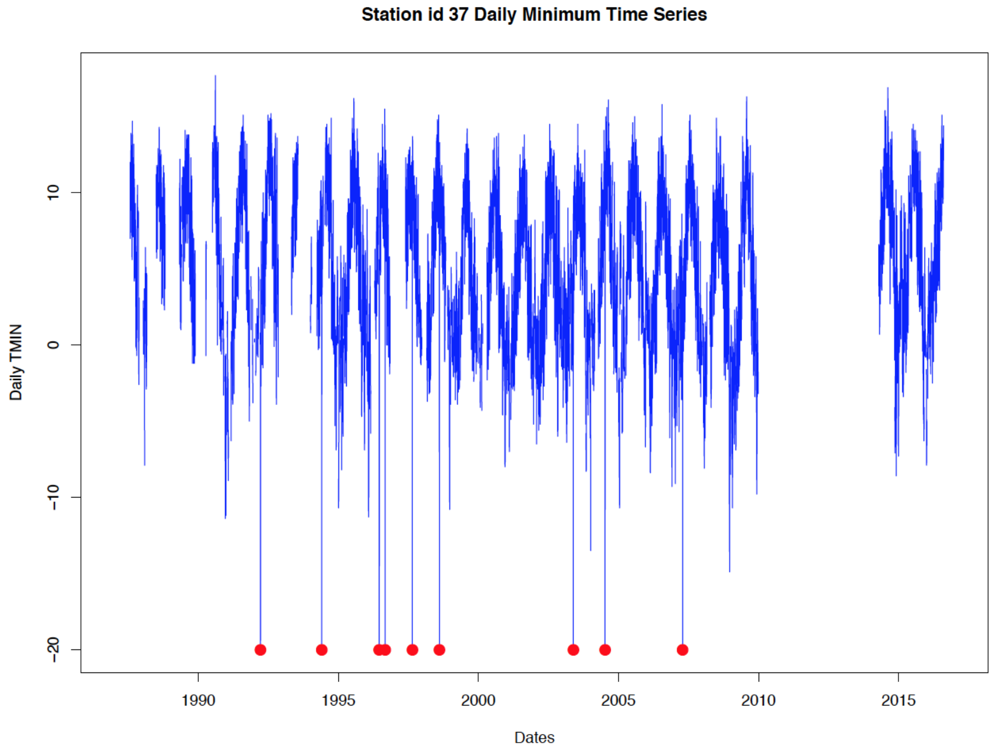


Figure 8: Sample daily minimum temperature record from the FLNRO_WMB Beaver Creek station (ID# 37) indicating the occurrence of outliers.

Another example is provided in Figure 10 and Table 5, depicting a time series of daily minimum temperature for the Gold River station (ID# 64003) from MoTI network in the Campbell River region with a prominent negative outlier in January of 2000. Table 5 shows the QC results for this observation and indicates that the allowable range is for minimum temperatures that are equal to or above -23.73 °C. The value -35 °C clearly violates this range. It is unlikely this value is realistic. In Figure 9, the Tmin values for the nearby stations Gold R. nr Ucona R. (BCH, ID# GLD) and Woss (MoTIm, ID#64006) are found to be 1.5 °C and 0.5 °C for the same day Jan 26th , 2000, which therefore did not record a similar extreme cold value as the one observed in Gold River. Given the climate of the region and the proximity to the Pacific Ocean, the GLD and Woss values are much more plausible. In this case, the outlier is regarded as an error and marked as missing value. This is also supported by the time series plot for Gold River, which is displayed in Figure 10. Note that this type of outliers check with careful comparison with nearby stations and consideration of climatic conditions was performed only for 3% (25 stations) of the total stations from the BCH network. The majority of the outlier check is based solely on the station itself. More sophisticated automation methods for outliers check including multiple stations will be developed in the future work.

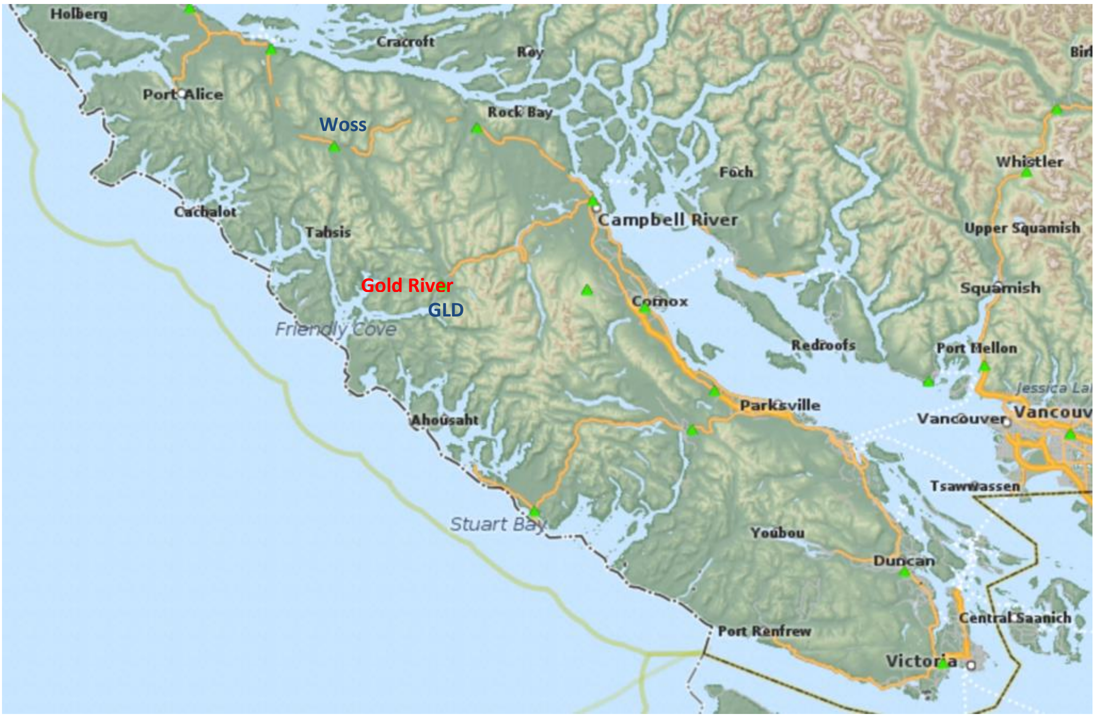


Figure 9: location of sample station Gold River (ID# 64003) from MoTI, and its neighboring stations Gold R. nr Ucona R. (ID# GLD) and Woss (ID# 64006) (Map is from the PCIC data portal)

Date	Tmax Low	Tmax	Tmax Up	Tmin Low	Tmin	Tmin up	DTR Low	DTR	DTR up
2000-01-26	-12.54	1.5	19.23	-23.73	**-35**	22.59	-14.8	**36.5**	23.2
2001-12-29	-11.83	-1.5	17.12	-22.54	-20	20.97	-9.24	**18.5**	15.9

Table 5: Sample data from the MoTI Gold River station (ID# 64003) showing the occurrence of a Tmin outlier and two daily temperature range (DTR) outliers. Column 1 is date of the identified outliers; column 3, 6, 9 are daily maximum temperature (Tmax) daily minimum temperature (Tmin) and daily temperature range (DTR); column 2 and 4 are the mean of total daily maximum temperature minus and plus 4 standard deviation; column 5 and 7 are the mean of total daily minimum temperature minus and plus 4 standard deviation; column 8 and 10 are the mean of total daily temperature range minus and plus 4 standard deviation.

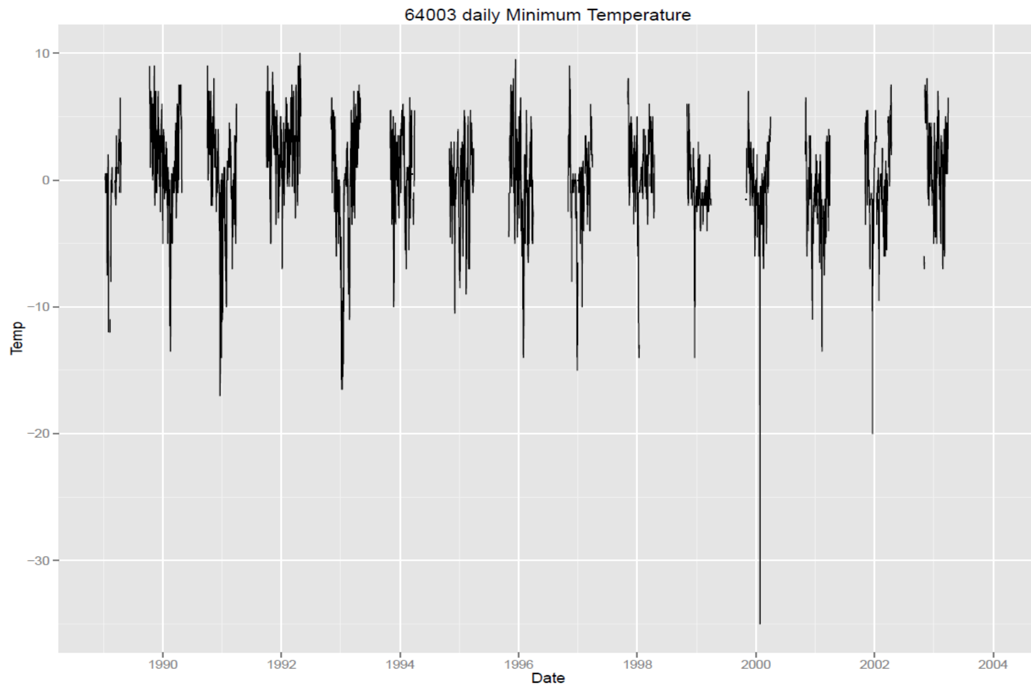


Figure 10: Sample daily minimum temperature record from the Campbell River region at Gold River (ID# 64003) indicating the occurrence of an extreme low outlier early in year 2000.

3.1.3 Data Quality Control Summary

Networks	Total # of stations	# of stations with outliers	% of stations with outliers
BCH	84	47	56%
FLNRO_WMB	410	158	39%
MoTI	303	121	40%
Networks	**Average data length (years)**	**# of daily outliers**	**% of daily outliers**
BCH	39 (1978 to 2016)	134	0.011%
FLNRO_WMB	16 (1995 to 2010)	394	0.035%
MoTI	32 (1983 to 2014)	288	0.014%

Table 6: Summary table of the outliers identified for each network (MoTI include sub-networks MoTIm and MoTIe)

To sum up, 797 stations with daily Tmax and Tmin observation were examined for quality. Table 6 displays a summary the data quality findings for each network, specifically showing the outlier results. Over half of the BCH stations have one or more outliers in their records. In total, 134 outliers were found in the BCH data. FLNRO_WMB has a similar percentage of stations with outliers as observed in the MoTI network with value of 39% and 40% respectively. 394 total outliers and 288 total outliers are identified for FLNRO_WMB and MoTI respectively.

The outlier rate for each network is further calculated from using the flagged daily outliers divided by the estimated daily observations ($\frac{\#daily\ outliers}{\sum \#daily\ observations}$). Take BCH network for an example: BCH has an average of 39-year data length, the average annual days are chosen as 365.25, the data availability of BCH is 96%, and the number of stations is 84 (Table 1). Therefore, there are roughly 1,148,697 daily observations with 134 flagged daily outliers, which lead to 0.011% of daily outliers within the BCH network. Same fraction calculations of outliers are performed for the other two networks, which show 0.035% and 0.014% of the daily outliers in FLNRO_WMB and MoTI networks.

Overall, while BCH has the largest fraction of stations with at least one outliers, it also has the lowest fraction with 0.011% of outliers overall. FLNRO_WMB stations have the highest fraction of outliers 0.038% detected and its average data length is shortest. Outliers in the FLRON_WMB network is also largely due to the occurrence of -20 °C issues regarded as outliers that was previously discussed in the SECTION 3.1.2.

3.2 Homogenization

3.2.1 Description of the Homogenization Procedures

It is recommended that the monthly series be tested for homogeneity before testing the corresponding daily series because day-to-day variability is larger than month-to-month variability, which make it is more difficult to detect changepoints in daily time series. Both daily homogenization and monthly homogenization techniques are implemented for temperature in RHTestsV4 (Wang, 2013). This thesis only focuses on monthly scale homogenization. Daily temperature homogenization is an ongoing project using the results obtained from testing the monthly series.

RHTestsV4 specific data format requirements that are different from the RClimDex data requirement and are shown in Appendix B, Table B4. Thus the first step in the homogenization procedure is to reformat the data to match RHTestsV4 requirements. Quality-controlled daily PCIC station data were converted into monthly averages of daily maximum and minimum temperature, which are on a continuous monthly time scale with NA replacing missing values. The rule for obtaining the monthly means is that there need to be at least 25 valid observations in a month. Reformatting to match the format requirement of RHTests was also done for the AHCCD data. The correlation coefficient and significance were calculated using the monthly anomalies of base time series from PCIC stations and reference time series from AHCCD stations. The top three most significantly correlated reference stations are utilized as three reference series for homogenization (refer to SECTION 2.1.2).

Because comprehensive metadata is not available for this research, we are detecting only "so-called" *Type 1* changepoints which are defined as undocumented changepoints that are statistically significant even without metadata support in Wang (2008a). The iterative process for applying PMTred was described in SECTION 2.2. This process was applied to all base stations with the corresponding three reference stations for each base station.

3.2.2 Homogenization Results Overview

Results for an example station at Brenda Mines (ID# BMN) from BCH are shown in Figure 11 of how discontinuities are tested for using three reference stations. The base minus the first reference series (Base – Ref 1) in the Figure 11a shows a detected shift in January, 2003, which is in approximate agreement with the December, 2002 shift identified in both the base minus the second and the third reference series (Figure 11b and c). The second shift in March 2005 detected by Base – Ref 2 series is in the same season as the changepoint of May 2005 identified in Base – Ref 3. In this case, the changepoints which occurred between December 2002 and January 2003 are significant and are confirmed with the three reference series applied. The changepoint between March and May in 2005 also aligns well within the two references series. This improves confidence that the identified changepoints are legitimite. In addition to that, these changpoints are statistically significant, so they are retained as discontinuities even without metadata support.

This labour-intensive procedure was repeated for all locations, separately for both Tmax and Tmin, using three reference series for each location and variable. The detected dates of changepoints from these various difference series are not always consistent among the three reference time series. The general rule for making adjustments is adjusting the detected changepoints within each of the "bases minus reference series" individually. The method for creating the final homogenized product will be explained in more detail in SECTION 3.2.4.

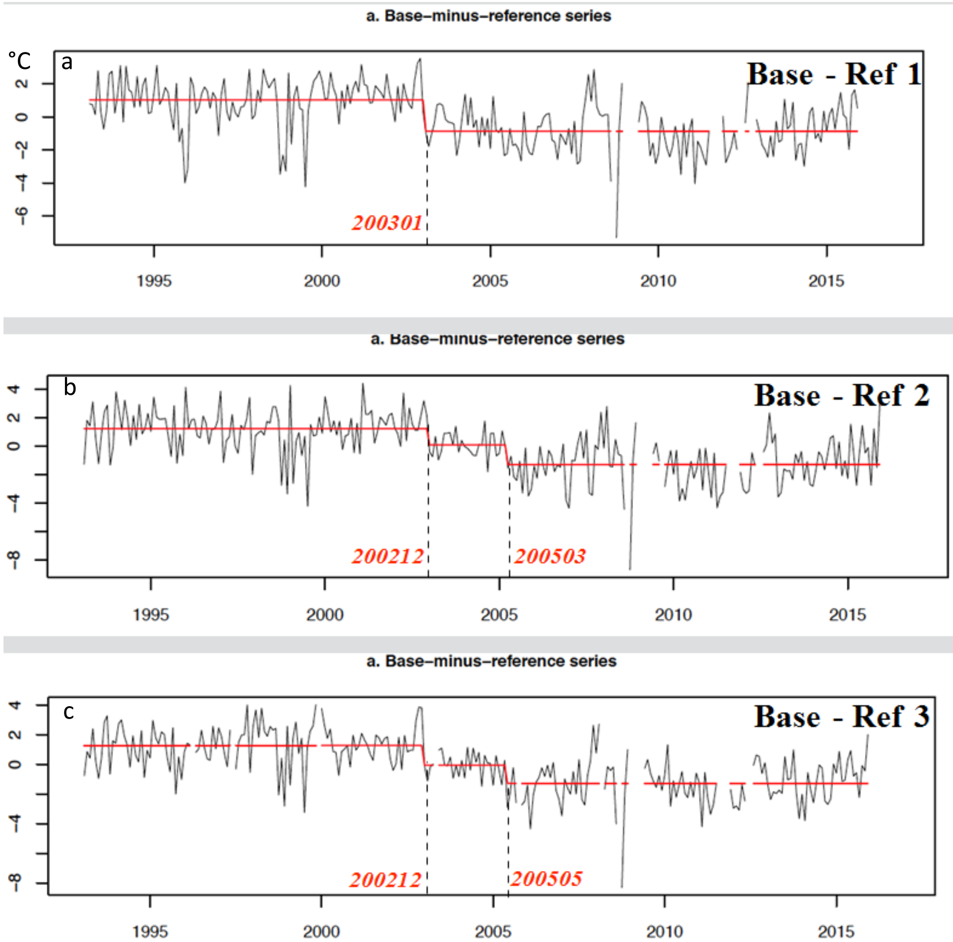


Figure 11: A sample station at Brenda Mines from BCH (ID# BMN) shows that the differencing deseasonalized Tmin time series based on the sample station and its three reference stations (denoted as Base – Ref in the figure) has similar changepoints detected. AHCCD station id for Ref 1, Ref 2 and Ref 3 are 1026639, 1067742 and 1060841 respectively. This figure is one of the sample output results from RHTests that has the homogenization techniques implemented.

Homogenization could alter the trend of the time series. Figure 12 shows the monthly temperature anomalies for a BCH station Stave R. Upper (ID# STV) in the Campbell River basin that exhibits a substantial warming trend. Figure 12a displays the unadjusted data. Figure 12b, obtained using the methods described above, and shows the homogenized record based on one of the reference time series. The latter shows little trend, if any, over the roughly 50 year period. Figure 13c shows the original (black) and adjusted (grey) time series superimposed on each other. Details of temperature trends analysis in BC will be provided in CHAPTER 4.

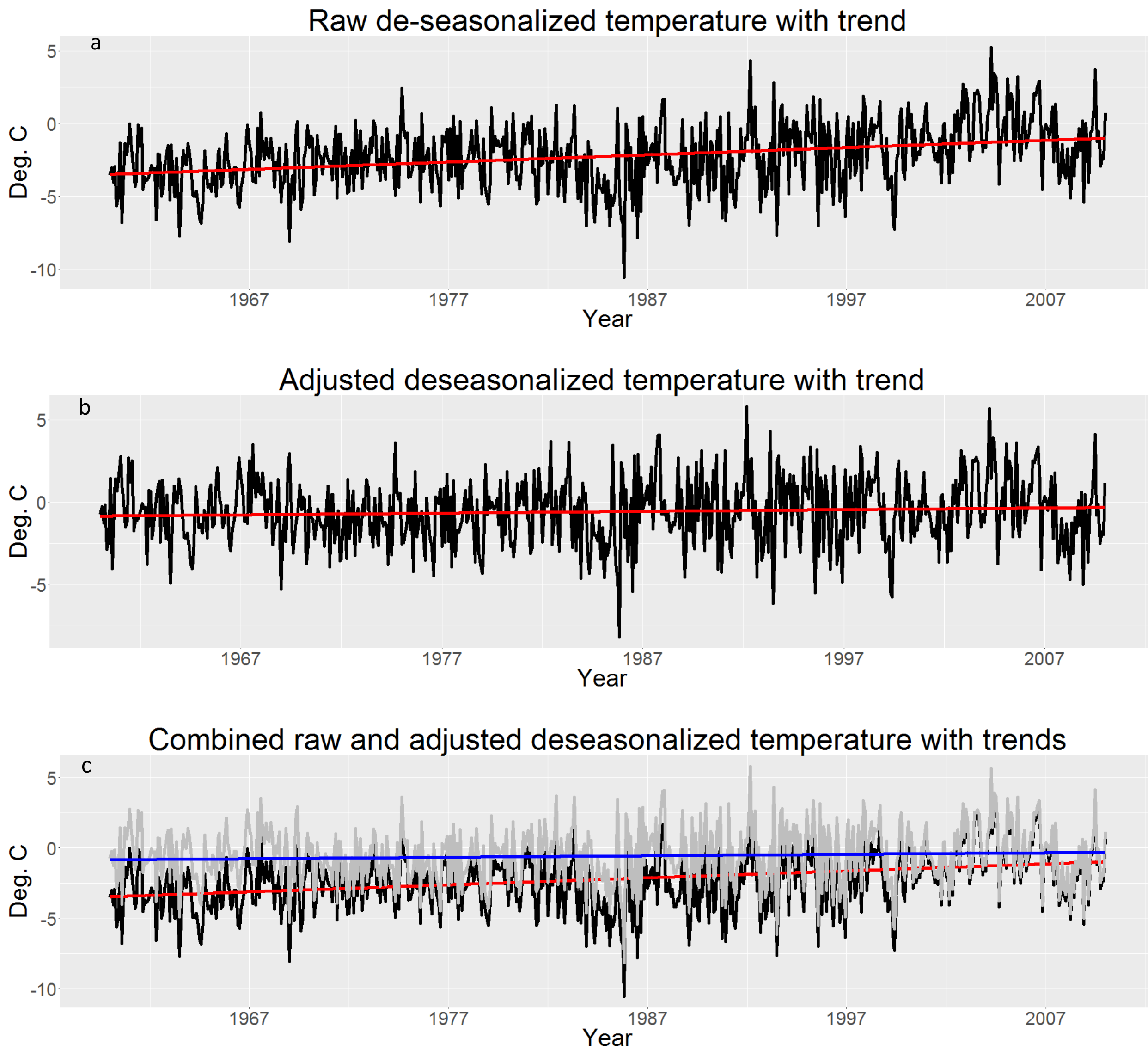


Figure 12: Deseasonalized monthly maximum temperature before (top) and after (middle) adjustments with trend for the sample station Stave R. Upper from BCH network (ID# STV). The lower panel shows the two time series superimposed on each other. The red and blue lines show the temperature trends before and after adjustment, respectively.

3.2.3 Homogenization summary

Table 7 gives a summary of the homogenization results for monthly maximum and minimum temperature as a whole. Although 797 stations for both monthly Tmax and Tmin were tested for inhomogeneities, only 310 (39%) and 307 (38.6%) respectively have sufficient data (time series lengths need to be over 10 years) to be suitable for homogenization. Amongst these stations, 92 Tmax stations and 75 Tmin stations have no detected changepoints, which is confirmed in tandem by their corresponding three references.

	Tmax	**Tmin**
# of Input Stations	797	797
# of Eligible Station	310 (39%)	307 (38.6%)
# of Stations with no detected changepoint	**92 out of 310 (29.6%)**	**75 out of 307 (24.4%)**

Table 7: Summary table of the homogenization results

Tmax	**BCH**	**FLNRO_WMB**	**MoTIe**
# of Stations with CPs	62 out of 84 (73.8%)	13 out of 124 (10.5%)	34 out of 102 (33.3%)
% of CPs	139 out of 35667 (0.4%)	20 out of 17352 (0.1%)	41 out of 10067 (0.4%)
Tmin	**BCH**	**FLNRO_WMB**	**MoTIe**
# of Stations with CPs	51 out of 84 (60.7%)	33 out of 121 (27.3%)	14 out of 102 (13.7%)
% of CPs	137 out of 35188 (0.4%)	38 out of 17200 (0.2%)	18 out of 10194 (0.2%)

Table 8: Summary table of the homogenized results including number of stations that have changepoints and the changpoints rate for Tmax (upper three rows) and Tmin (lower three rows) that break down to each network.

The number of stations with changepoints and the changepoints rates are broken down by network (Table 8). Changepoints rate is calculated from using the number of changepoints divided by the number of monthly temperature data points ($\frac{\sum Number\ of\ CPs}{\sum Number\ of\ the\ monthly\ data\ points}$). It should be noted that the application of three reference time series yields various numbers of changepoints for each of the base time series. Minimum and maximum number of changepoints corresponding to each station is retrieved. In Table 8, the display of the number of changepoints represents only the minimum numbers of changepoints. For example, for the FLNRO_WMB network, the minimum numbers (20) of changepoints is a result of changepoints numbers that were added up across each of the stations for the Tmax variable. The number of the data points (17352) is obtained by counting the number of non-missing values based on the monthly Tmax time series for each of the FLNRO_WMB station and were then combined together. So, the changepoint rate for FLNRO_WMB is 0.1% (20/17352).

In general, BCH has the highest portion of stations with detected changepoints. Specifically, there are 73.8% and 60.7% of the monthly maximum and minimum temperature records in the BCH network that have changepoints. The total number of the changepoints in this network is 139 and 137 for Tmax and Tmin respectively. The changepoint rate of BCH also ranks highest (0.4% for both Tmax and Tmin), with the same percentage as MoTIe for minimum temperature. FLNRO_WMB ranks the lowest within the three networks. It has 10.5% and 27.3 % of the Tmax and Tmin that is detected to have changepoints. In total, there are 20 and 38 changepoints in FLNRO_WMB. The changepoint rates are also smallest for FLNRO_WMB network, with 0.1% and 0.2% for both Tmax and Tmin respectively. MoTIe's changepoint rate is also 0.2% for Tmin. MoTIe is found that the percentage of the stations that have changepoints for Tmax and Tmin are 33.33 and 13.7 respectively. There are 41 and 18 changepoints in total detected for this network.

3.2.4 Selection of Monthly Homogenized Temperature Products

As explained earlier in SECTION 2.2.2, the homogenization approach results in three homogenized time series, one for each reference series. The final step is the determination of the optimal monthly homogenized temperature time series from these three time series as the final homogenized product. Based on the deseasonalized three homogenous time series and the averaged time series of the three, the standard deviation of each of the time series are calculated in order to compare variability differences. An example is given in Table 9. It turns out that the differences among the standard deviations are very small at the sample station Brenda Mines from the BCH network (Figure 13), which illustrates that doing the average will not smooth out the variability of the time series in the relevant way. To a very large extent, similar results are true for other stations as well. In order to take the influence of all reference series into consideration, the averaged homogenized time series is selected as the final product that will be further utilized in the trend analysis presented in CHAPTER 4.

Figure 13: Example BCH station Brenda Mines (ID# BMN) shows the time series of three deseasonalized homogenized monthly mean Tmax datasets obtained using three different reference series (panels a, b, c), and the average time series of the three (panel d)

	standard deviation
Homog_mx1077500	1.90
Homog_mx1085835	1.92
Homog_mx1088015	1.99
Homog_Average	1.93

Table 9: Standard deviation of the four homogenized time series for the Brenda Mines station (ID# BMN)

Chapter 4: Trend Analysis

4.1 Introduction

We have now expanded the availability of homogenized monthly temperature data in BC beyond the ECCC homogenized AHCCD product. We are therefore interested next in analyzing the local trends based on this increased data resource. In this section, trend at each station is examined first. Trends in the sub-regions where the stations are clustered are then compared based on data before and after homogenization. In order to evaluate climate conditions at locations where no direct measurements are collected, interpolation is used to estimate the spatial pattern of trends for BC. In order to validate the trend results based on PCIC homogenized station and gridded datasets, trends are compared based on those calculated from AHCCD (Vincent, 2012) and CANGRD (Gandin, 1965; Bretherton et al., 1976; Alaka and Elvander, 1972).

4.2 Data and Methodological Approach

The station data being utilized in trend analysis include 310 maximum and 307 minimum temperature anomalies of PCIC stations before and after homogenization (denoted as raw PCICstn and homg PCICstn). AHCCD Tmax and Tmin anomalies are available at 54 locations in BC. Gridded products come from anomalies interpolated from PCIC stations (raw PCICgrid and homog PCICgrid) and CANGRD data from ECCC.

Refer to SECTION 4.2.3 for details of the production of PCICgrid.

The CANGRD dataset (CANGRD description document, 2014[12]) has a 50 km spatial resolution and is based on the AHCCD from 1900 for southern Canada (from 1950 for all Canada). The end year of the CANGRD dataset in this research is 2014. CANGRID is produced

[12] Contact: open-ouvert@tbs-sct.gc.ca for details of CANGRD.

using Gandin's Optimal Interpolation method (Gandin, 1965). Monthly, seasonal and annual temperature anomalies from the average over 1961-1990 are interpolated to the 50 km grid boxes. Due to the differences of the projection and grid system, CANGRD is transformed in BC to have the same grid as homog PCIC grid in order to be comparable. The grid will be described in SECTION 4.2.3.

4.2.1 Interpolation

Interpolating point data onto grids is commonly applied in the environmental and biological science, including climate science (Hijmans, 2005). A number of different statistical interpolation methods have been developed to generate gridded climate data, such as Kriging (Hartkamp et al., 1999 for an overview), the thin-plate spline algorithm (Hutchinson, 1995), and Gandin's Statistical Optimal Interpolation (Gandin, 1965). For example, the Thin-Plate spline interpolation method (Hutchinson, 1995), which is the core of the ANUSPLIN package (Hutchinson and Xu, 2013), uses every station as a data point and uses latitude, longitude, and elevation as independent variables.

In this paper, station data are interpolated to evenly spaced grid boxes (0.5°*0.5°) using the Inverse distance weighting (IDW) approach. This method is chosen since the purpose of the interpolation in this study is to estimate trends of spatial temperature anomalies, and trend in anomalies are not strongly affected by topography. In addition, when estimating the temperature beyond the province, such as the borders, IDW based extrapolation usually shows feasible results which tends to be the sample mean of the data points. Another practical reason for using IDW method is that it is relatively simple to implement, and thus was more manageable in the context of the time available to complete this thesis. A potential limitation of IDW is that the interpolated value can be greatly influenced by the local extreme values. This could bring the interpolated value different from points nearby, which leads to an abrupt change in values and is prone to the "bull's-eye effect". Improved interpolation will be employed to handle more complex cases in the future work.

Inverse distance weighted (IDW) interpolation estimates the value (u) at an unknown point (x) using a linearly weighted combination of a set of neighboring sample points (u_i):

$$u(\mathbf{x}) = \sum_{i=1}^{N} w_i'(\mathbf{x}) u_i \quad (4.1)$$

Where $$w_i' = \frac{w_i(\mathbf{x})}{\sum_{i=1}^{N} w_i(\mathbf{x})} \quad (4.2)$$

and

$$w_i(\mathbf{x}) = \frac{1}{d(\mathbf{x}, \mathbf{x}_i)^p} \quad (4.3)$$

w_i':	Normalized/standardized IDW weight. The purpose of using w_i' is to ensure the estimator $u(\mathbf{x})$ is unbiased.
w_i	IDW weighting function defined by Shepard (Shepard, 1968).
$\mathbf{x}$	Location of interpolation point.
$\mathbf{x}_i$	Neighbouring point with known values of u
d	Distance between the point $\mathbf{x}$ and its neighbour $\mathbf{x}_i$.
N	Nearest number of neighbours.
p	Power parameter.

Grandin's method interpolates similarly, except that the weights are chosen to minimize the expected interpolation error. This requires carefully developed estimates of the spatial covariance function for u, which is a task that was beyond the scope of this thesis.

In the case of IDW method, interpolation to unobserved points is optimized through the selection of N and p. Thus cross validation analysis with varying N and p is undertaken in order to minimize the root mean square error of interpolations which is a common measure of interpolation accuracy. RMSE is the square root of Mean Squared Error ($RMSE = \sqrt{MSE}$, Wilks, 3rd, pp. 357). The MSE operates on the gridded forecast and observed fields by spatially averaging the individual squared differences between the two values (y: predicted value, o: observed value) at each of the M grid points in the field.

$$MSE = \frac{1}{M}\sum_{m=1}^{M}(y_m - o_m)^2 \quad (4.4)$$

Cross validation is applied in order to determine the optimal parameters for IDW model. The steps of performing leave-one-out cross validation of IDW interpolation with parameters is as follows.

1) One station at a time is held out and the rest of the stations were used to estimates the value at the holding station point based on the IDW model with certain parameters being tested.
2) Repeat the process in step one for other station locations, which generates a series of predicted values for each station
3) Apply procedures in step one and two for all time instance, being investigated (annual Tmin, Dec 2002, July 1990) so that for each station there is a pair of observed value and estimated value.
4) Compute the RMSE between the observed and estimated values. As a result, spatially one RMSE value for each station is calculated. Median of the RMSE values across all stations is used as the representative RMSE statistics.
5) The optimal parameters combination for IDW interpolation model is chosen as those with the minimum RMSE.

Cross validation was applied for anomalies in the annual averages of daily Tmin interpolation based on the PCIC homogenized station data and displayed in Figure 14. To make the cross validation reliable for producing the gridded products while also being computational efficient, the cross validation is performed on 8 by 5 p (=0, 0.2, 0.5, 1, 1.5, 2, 2.8, 3.5) and N (= 3, 5, 7, 8, 15) parameter combinations. Roughly, p minimized RMSE when N=3, 5 and 7. According to the bar chart, N=3 in general leads to smallest RMSE. However, the small number of stations involved could lead to a biased estimation when one station's data is erroneous. RMSE is larger when N=5 than N=7, therefore N=7 is a better choice. As a result, parameters of N=7 and p=1 (RMSE=0.45) is selected as the optimal combination.

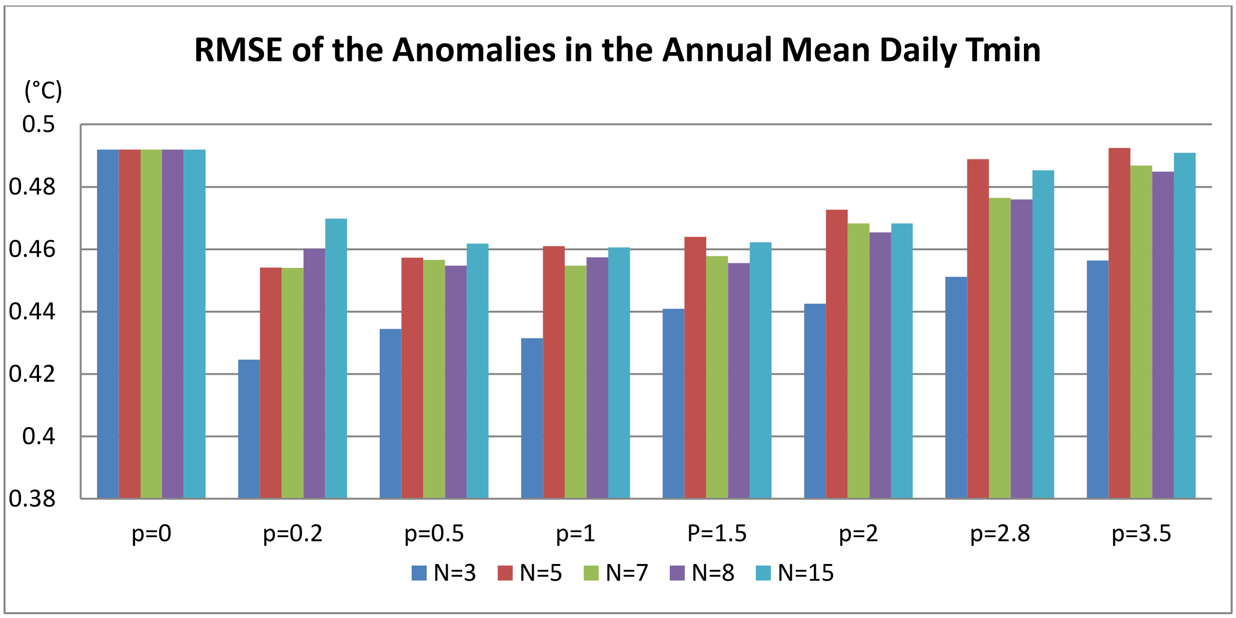


Figure 14: Bar chart shows RMSE of the anomalies in the annual mean daily Tmin interpolation (IDW) based on the PCIC homogenized stations for various N and p

Note that though value N=7 was chosen, fewer than 7 stations may have been used in some cases. Take the interpolation of annual anomalies as an example. The value for a year to be interpolated for a certain grid box is usually determined by the 7 stations that are closest. Due to data limitations however, some stations might not contain any value for that specific year; in this case only the station amongst the 7 with data were used. It is possible that all 7 stations do not have data, in which case the interpolated result is set to missing. Using a fixed set of 7 stations ensures the consistency of the interpolated results, but also has the limitation of increasing uncertainty when most of the stations values are missing.

The cross validation was not carried out for all years and months due to a lack of time. But RMSE calculations for a winter month (Dec, 2002) and a summer month (July, 1990) were processed. The intent was to test the interpolation performance during extreme warm and cold days. The results are shown in the Appendix B (Figure B2 and B3). A complete treatment is therefore a topic for future research.

4.2.2 Calculation of Trends

The Mann Kendall (M-K) trend test (see Appendix A1 for detailed methodology) and Theil-Sen's robust estimator (Sen, 1968) are selected to test the significance of the trend and to estimate the trend slope. The significance level of the M-K test is set to 5% in this paper. Although parametric tests, such as Student-t test, are commonly used to estimate the linear trend and the statistical significance, M-K test is preferable because it tends to be less sensitive to outliers (National Center for Atmospheric Research Staff, 2014).

Auto-correlation (serial correlation) has to be taken into consideration before conducting a test of trend since Monte Carlo simulation experiments have shown that the variance of the estimate of the M-K statistic is altered due to the serial correlation (Yue, 2002). Zhang's iterative pre-whitening method is applied rather than Yue's method since it not only considers the situation that the linear trend and the lag one auto-correlation process might interact with each other when they both exist in a time series (Zhang & Zwiers, 2004), also, the actual significance level that Zhang's method produces is close to the specified significance level, where Yue's tends to inappropriately inflate significance level in the presence of serial correlation. See also Burger (2017).

4.2.3 Steps for Interpolation and trend calculations

Monthly mean maximum and minimum temperature anomalies were first calculated relative to the 1981-2010 climatological mean. A climatological mean was calculated only if there are at least 10 years' data over the 30 years. Annual and seasonal maximum temperature and minimum temperature anomalies were then calculated from these monthly anomalies if there are at least 8 months of data for a year, and at least 2 months' data for a season.

Figure 15 shows the layout of the 0.5° latitude × 0.5° longitude grid system and presents 310 maximum temperature station locations (Table 1) with monthly or seasonal anomalies that are to be interpolated. Anomalies are interpolated onto this grid using the IDW method. The grid system has 62 columns and 28 rows, the coordinates of the center of the grid box in the southwest corner is (47.75° N, 140.25° W). Considering that the area of interest is BC, therefore, the gridded field outside of BC (Pacific Ocean and adjacent provinces) is masked in the display in the trend analysis section.

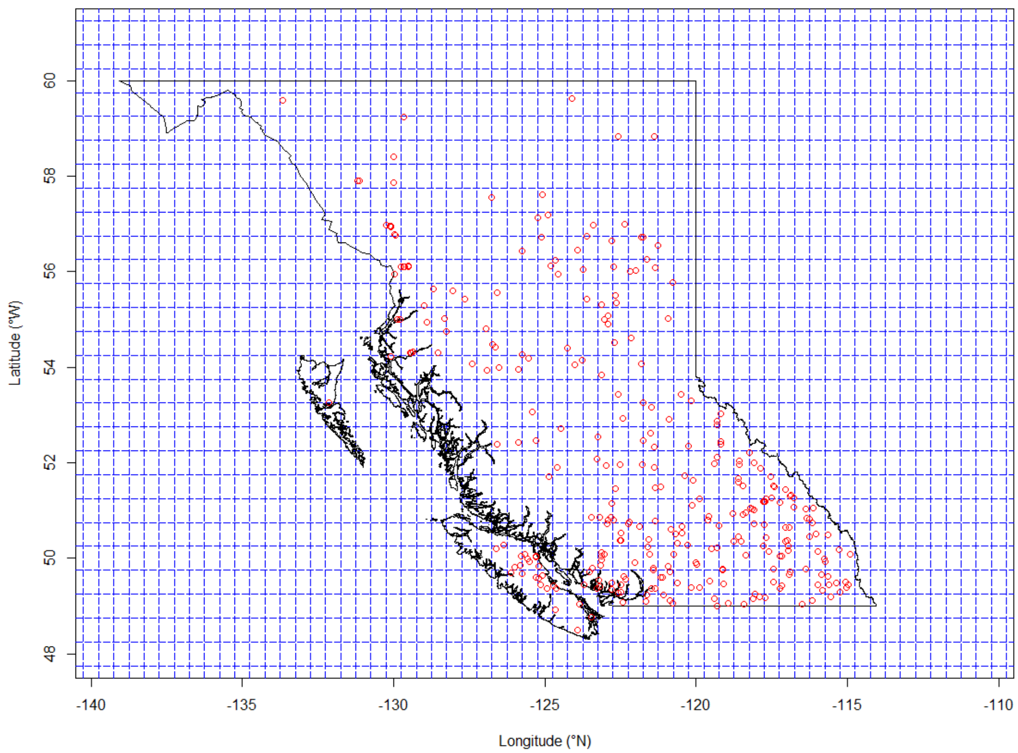


Figure 15: Grid system for interpolation showing an example of 310 homogenized maximum temperature locations

In terms of calculating the trend, the M-K test and Theil-Sen's slope estimator are applied over the 1990-2014 period (25 years). This period was chosen for trend analysis because it overlays well with the periods covered by AHCCD and CANGRD, which end in 2015 and 2014 respectively, and because spatial coverage is much greater for the 25-year period ending 2014 than a longer period that begins earlier, such as 1980-2014. The 1990-2014 encompasses the period when all three networks have relatively large numbers of stations reporting. The data coverage requirement in trend calculation for both station data and each interpolated grid box are that they should have at least 80% data availability, which is more than 20 years. The percentage of stations that satisfy this criteria on average is 81 and 80 for Tmax and Tmin variables respectively (Table 10).

#(%) of stations satisfy the criteria of calculating trend	Tmax (Total #310)	Tmin (Total #307)
Spring	87 (28%)	86 (28%)
Summer	76 (25%)	75 (24%)
Autumn	79 (26%)	78 (25%)
Winter	87 (28%)	87 (28%)
Annual	77 (25%)	76 (25%)
Average	**81 (26%)**	**80 (26%)**

Table 10: Numbers and percentage of PCIC stations (both for raw and homogenized stations) that satisfy the criteria for trend calculation (more than 20 years within the 1990 – 2014 period). Name of column one gives the theme of the table. The second and third columns show the variable name Tmax and Tmin, in the same time, give the number of stations that is used to calculate the percentage for each season.

4.3 Results

In order to accurately locate the regions when describing the trends in BC, the ecoprovinces are used. A map of ecoprovinces (definition is provided in the Glossary) produced by BC Ministry of Environment is displayed in Figure 16.

Figure 16: Map of ecoprovinces in BC (source: Ministry of Environment, BC)

4.3.1 Single Station Trends

Maps of trends at individual stations are produced for all seasons based on PCIC station data before and after homogenization and for AHCCD data for both Tmax and Tmin. In this section, trend results from annual, winter, summer data are presented based on Homog PCICstn and AHCCD stations. The trend maps for spring and autumn are included in Appendix B (Figure B4 and B5). Comparison of trends between homog PCICstn and AHCCD is done to validate the trend results. Differences of trends between PCICstn before and after homogenization will be introduced in SECTION 4.3.2.

Based on the monthly homogenized PCIC Tmax seasonal anomalies time series, trends are examined across the province. As seen in Table 10, only 25% to 28% of the data for Tmax (total 310 stations) satisfy the criteria for calculating trends. AHCCD is filtered with the same criteria; it has only a few stations do not have sufficient data for trends calculation.

The density and spatial locations of homog PCICstn and AHCCD stations are very different. AHCCD are uniformly spread out in BC, whereas PCICstn data with sufficient record length are clustered in three main regions. AHCCD stations have good density in the sub-boreal area where there are no long duration stations available for PCICstn. These two sources of data complement each other.

Homog PCICstn **AHCCD**

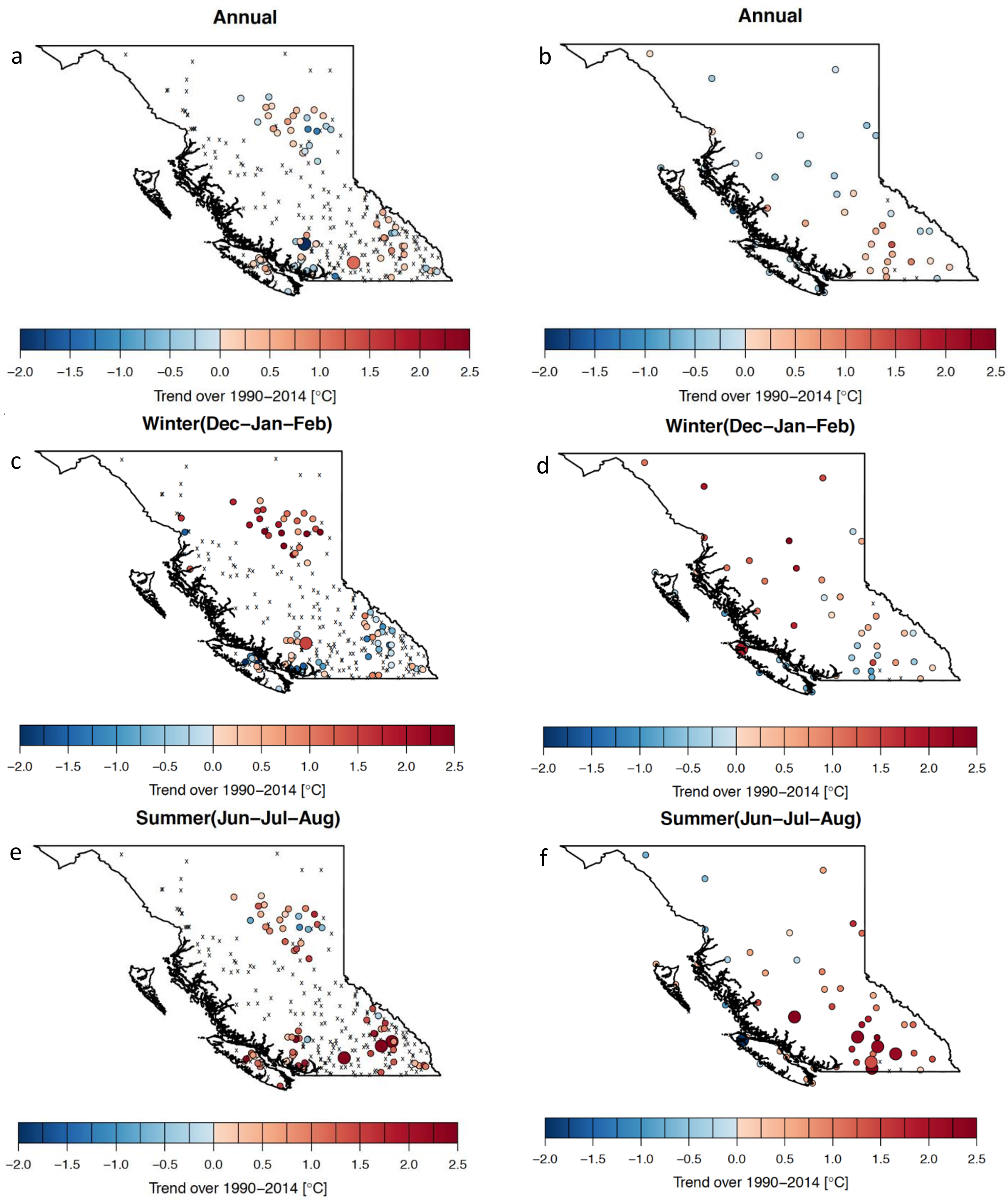


Figure 17: Left panel show the annual, winter and summer Tmax trends over 1990-2014 for each station based on homog PCICstn; trends based on AHCCD station over the same period of time are listed on the right panel. Warm color indicate positive trends, cool color indicate negative trends. Crosses represent stations with inadequate data, which means stations with than 20 years' data for trend analysis. Large dots mean trends are significant at the 5% significant level; small dots mean trends are insignificant at the 5% significant level.

In terms of the trend patterns, starting from the annual trends based on the homog PCICstn, overall mixed trends can been seen across the province, especially in the northern BC, as displayed in Figure 17a. Two stations with statistically significant trends in southern BC are found. Specifically, one has a cooling trend (below -2.0°C/25 years) located in Central Interior and another one has a warming trend (1.1°C/25 years) located in Southern Interior. In Southern Interior Mountains, warming trends (0.5°C~1°C/25 years) are the main pattern for most of the stations, except for four stations with cooling trends which have magnitudes that range from -0.1 to -0.5°C/25 years. For annual trends based on the AHCCD stations as shown in Figure 17b, most stations in Central and Southern Interior also found to have warming trends (0.4 °C~1.5°C/25 years), so do stations in Southern Interior Mountains. AHCCD stations in northern BC in general are found to have negative trends, which is not consistent with trend results based on homog PCIC stations. However, it should be noted that the station locations are very different in the north for these two data sources.

In Figure 17c, the winter trends show that northern BC has a dramatic warming trend over the 25 years (1.0°C~2.5°C). In the southern BC, negative trends (-0.2°C~-1.5°C/25 years) are dominant in Southern Interior Mountains with several warming trends. In the Central and Southern Interior, warming is intertwined with cooling trends, within which there is one significant warming trend with value of 1.3°C. Similar mixed trends are seen in the Coast Mountains and Georgia Depression areas. Vancouver Island has more cooling stations than warming stations, though, in general. Winter trends based on AHCCD data are shown in Figure 17d. Except for several stations with small negative trends in Southern Interior and Southern Interior Mountains, Georgia Depression, Coast Mountains, most AHCCD stations show clear positive trends (0.25°C~2.0°C/25 years) across the province. One statistically significant warm trend with value of over 1.5 °C is identified in the north corner of Vancouver Island. To the extent that one can judge, the overall pattern and magnitude of warming is similar between the two datasets.

For the summer trends, signs are in general opposite with the trends in spring. In Figure 17e and f, we see that most stations have warming trends across the province with few cooling stations. Based on the homog PCICstn, distinct warming trends with values exceeding 1°C are centered in the Southern Interior and Mountains areas, within which three statistically significant trends are observed (close to 2.5°C). Similar with trends based on AHCCD, strong statically significant warming trends with magnitude of 1.8 to around 2.5 °C are found in southern BC. The few cooling stations (below -0.5°C) are mainly located in Northern Boreal Mountain and Sub Boreal Interior for both data sources.

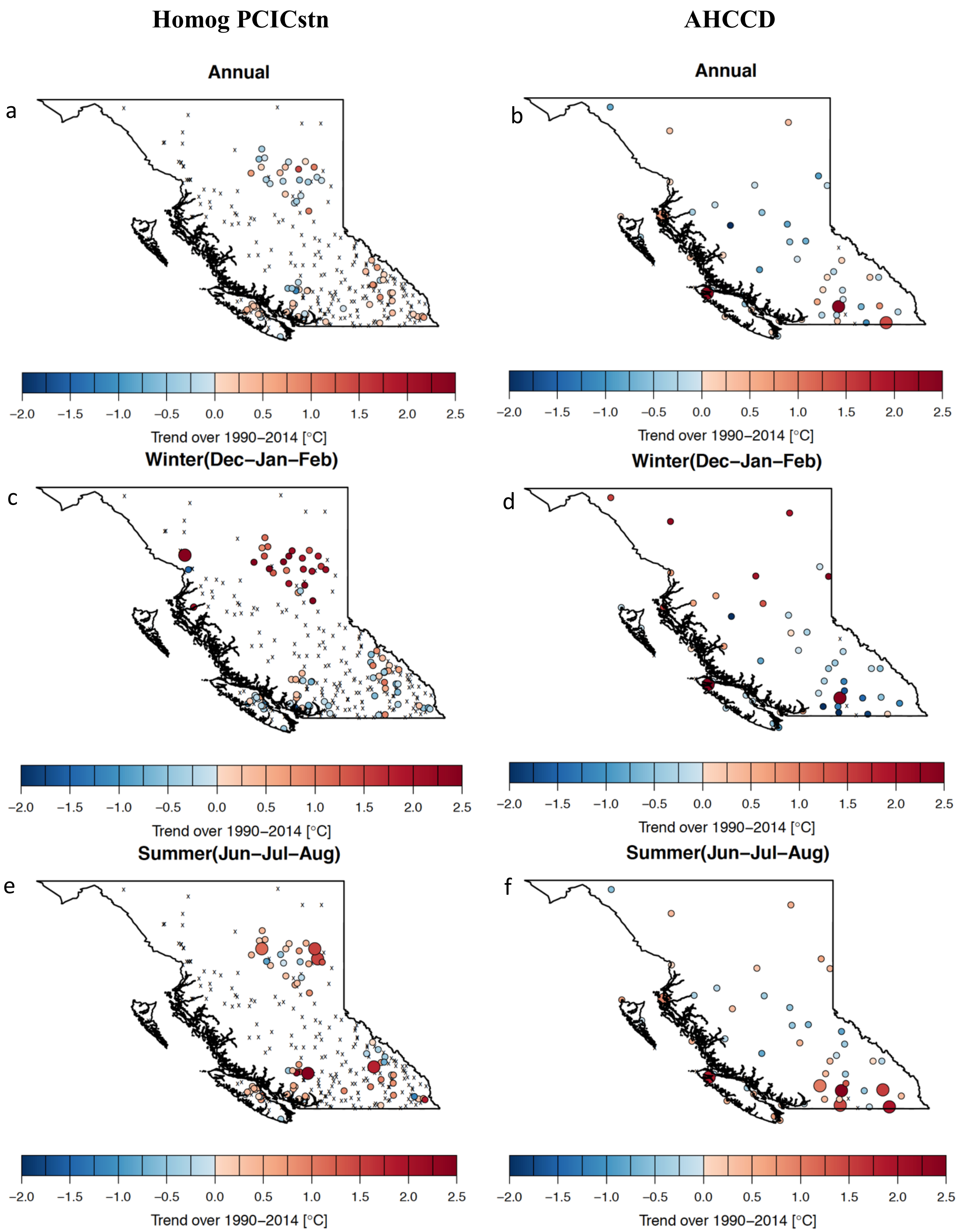


Figure 18: Left panel show the annual, winter and summer Tmin trends over 1990-2014 for each station based on homog PCICstn; trends based on AHCCD station over the same period of time are listed on the right panel. Warm color indicate positive trends, cool color indicate negative trends. Crosses represent stations with inadequate data, which means stations with than 20 years' data for trend analysis. Large dots mean trends are significant at the 5% significant level; small dots mean trends are insignificant at the 5% significant level.

The seasonal trends in Tmin over 1990-2014 (25 years) are first described at the annual average. Trends based on the homog PCICstn in Figure 18a and on AHCCD stations in Figure 18b both show a mixed pattern across the province in general. In Figure 18a, more stations have warm trends in the intersection area of the Northern Boreal Mountains, Taiga Plains, Boreal Plains and Sub-Boreal Interior. Warming trends are the dominant pattern in the Southern Interior Mountains and Vancouver Island. No statistically significant trends are identified based on the homog PCICstn. In the annual trends based on AHCCD (Figure 18b), four stations with statistically significant warming trends are found. Two of them are located close to the south border of BC, one in north corner of Vancouver Island with magnitude of around 2.0°C and the other one is on the Coast and Mountains with the trend value of 0.5°C. Stations in Sub-Boreal Interior and Central Interior based on AHCCD stations show cooling trends with magnitude ranging from 0.25°C to below -2.0°C, over 1990-2014. The differences in trend patterns based on these two types of data mainly due to the location differences.

Winter trends are examined and compared. Trend patterns between homog PCICstn and AHCCD differ somewhat also due to location differences. In Figure 18c, based on the homog PCICstn, Northern BC have distinct warming trends (over 1.75°C/25 years) with one statistically significant trend that has magnitude above 2.5°C. Southern BC trends are mixed with warming and cooling trends without a regular distribution. The trends on Vancouver Island show dominant cooling in Figure 18c, which is similar to trends based on AHCCD on the island (Figure 18d). Nevertheless, one statically significant warming trend with value exceeds 2.5°C is shown in the north corner of the island with AHCCD. Trends based on AHCCD in winter have mixed trends in the Sub Boreal Interior and Central Interior. In the Northern Boreal Mountains and Taiga Plains, strong warming trends (around 2~2.5 °C/25 years) are displayed. In Southern

Interior and Mountains, cooling trends prevail, with exception of one statistically significant warming trend with value of around 2.2°C.

Last, although many warming trends are found in summer Tmin based on both the homog PCICstn and AHCCD, warming trends are found to be slower than that of Tmax. Most homog PCICstn have warming trends in Figure 18e, spread out to the north, the Southern Interior Mountain and Georgia depression. There are five statistically significant warming trends (1.0°C~2.25°C/25 years), with three in the Northern Boreal Mountains and two in the lower interior regions. Among trends at AHCCD stations shown in Figure 18f, about half show positive trends. The pattern of cooling trends is a "belt" cut through northwest corner and southeast of BC. Dramatic statistically significant warming trends with values ranging from 0.3°C to 2.0°C are apparent in the lower Southern Interior and Southern Interior Mountains areas. Two other stations with statically significant warming trends are also found in the Coast and Mountains region (1°C~1.75°C/25 years).

4.3.2 Sub-regions

The potential influence of non-climatic shifts on temperature trends is discussed in SECTION 3.2.2. Here, in order to understand the impacts of the homogenization process on the estimated climate trends, trends calculated using the homogenized data are compared with those calculated using the raw observations. Raw observations refer to the temperature observations after quality control but before homogenization. Due to the identical spatial data coverage for both raw and homogenized PCIC temperature stations, it is found that stations with sufficient data for trend analysis (at least 20 years) are clustered mainly in three areas of BC. Therefore, considering the data density, geophysical terrain types and climatic conditions, the areal average is analyzed in these three sub-regions for regional trend comparison shown in Figure 19. The three sub-regions are: Southern Coastal BC (left corner), Southern Interior (right corner) and boundary intersections among Northern Boreal Mountains, Taiga Plains, Sub Boreal Interior and Boreal Plains (upper box: simply referred to as Northeast BC below).

Annual and seasonal bar charts show trend for each region for Tmax and Tmin over 1990-2014. Figure 19a shows that most Tmax trends in the homogenized data are of the same sign as those in the raw observations with exception of trends for autumn and winter in the Southern Interior region where trend magnitude is small. In general, homogenization reduced trends for most seasons except for spring and winter in the Southern Coastal region. Similarly, for Tmin trends in Figure 19b, homogenization has no effect or a small magnifying effect (increases the trends magnitude) on trends in the Southern Coastal Region. In the other two regions, trend magnitude is smaller after homogenization.

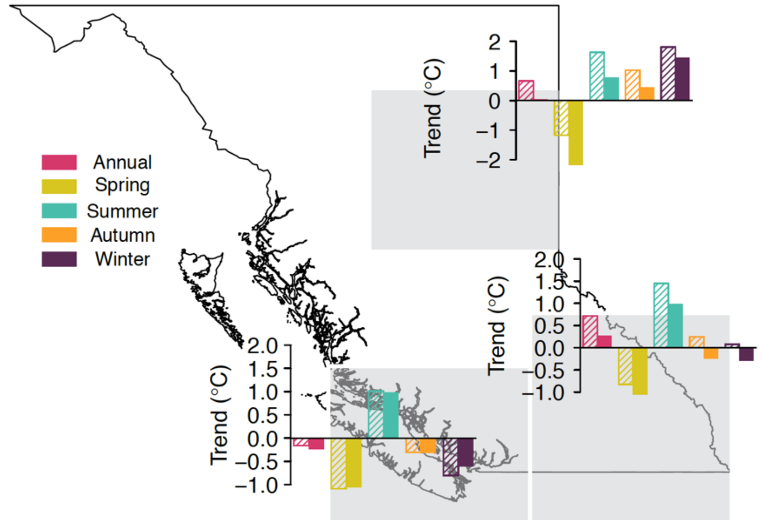
Regions	**Annual** (Raw | Homog)		**Spring** (Raw | Homog)		**Summer** (Raw | Homog)		**Autumn** (Raw | Homog)		**Winter** (Raw | Homog)	
Southern Coastal	-0.16	-0.24	-1.09	-1.06	1.01	0.98	-0.31	-0.32	-0.80	-0.60
Southern Interior	0.71	0.27	-0.83	-1.05	1.45	0.98	0.30	-0.24	0.08	-0.29
Northeast BC	0.63	0.04	-1.17	-2.17	1.63	0.78	1.03	0.44	1.81	1.44

Table 11: Annual and seasonal Tmax trend over 1990-2014 (°C/25 years) for three sub-regions based on raw and homogenized data. Significant trends are denoted with an asterix.

Regions	**Annual** (Raw | Homog)		**Spring** (Raw | Homog)		**Summer** (Raw | Homog)		**Autumn** (Raw | Homog)		**Winter** (Raw | Homog)	
Southern Coastal	-0.27	0.08	-0.66	-0.68	0.16	0.30	0.20	0.34	-0.18	-0.13
Southern Interior	0.45	0.34	-0.30	-0.43	0.50	0.39	0.72	0.48	-0.18	-0.33
Northeast BC	0.30	0.14	-1.24	-1.58	*1.08	0.90	1.2	0.99	1.96	1.43

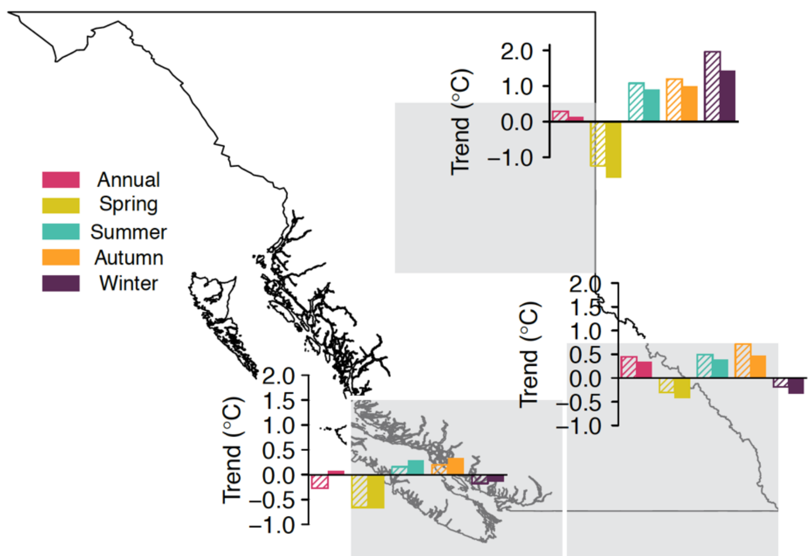
Table 12: Annual and seasonal Tmin trend over 1990-2014 (°C/25 years) for three sub-regions based on raw and homogenized data. Significant trends are denoted with an asterix.

a



Tmax Trend

b



Tmin Trend

Figure 19: 1990-2014 trend (°C over 25-year period) for daily maximum (upper panel) and minimum (lower panel) temperature by season for raw and homogenized data over three sub-regions in BC as indicated by gray shading. Striped bars represent results using raw observations and solid bars represent homogenized observations. Star (*) indicates that the trend for the season is statistically significant at the 5% level.

4.3.3 Spatial Pattern of Trends

Based on the gridded datasets that are interpolated from the homogenized station observations, spatial patterns of trends over 1990-2014 have been calculated. The distribution of stations spatially is quite different between AHCCD and PCIC stations, so in order to compare the trends in the area without data coverage within PCIC stations, trends based on homog PCICstn and CANGRD is considered in this section. Annual, summer and winter trends are discussed in the main text below while the remaining seasons of spring and autumn are displayed in Appendix B (Figure B6 and B7).

As for the trend calculations for single stations, at least 20 years of data were required within the 1990 – 2014 period to calculate trends from the gridded products. Grid boxes with insufficient data are marked as white in Figure 20 below. In order to allow the trends be comparable, a year counted as being available at a grid box only when both CANGRD and PCIC grid have values at that location. It should be noted that CANGRD is complete and only PCIC data have such gaps.

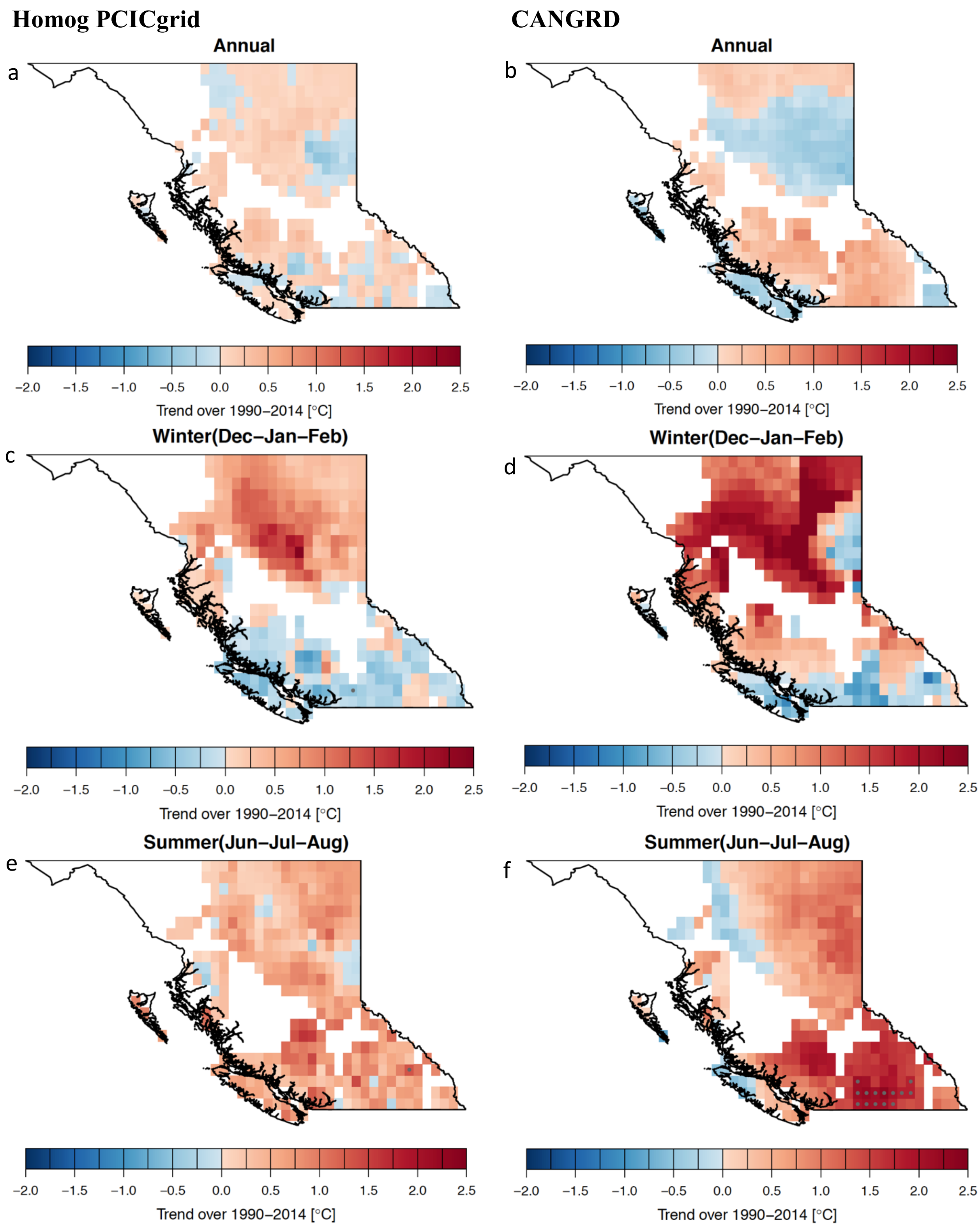


Figure 20: Annual, winter and summer spatial patterns of trends for Tmax over 1990-2014 based on homogPCIC (left) and CANGRD (right). White areas represent grid boxes with insufficient data. Grey dots indicate that the trend in the grid box is statistically significant at the 5% level.

The spatial pattern of annual trends for Tmax based on homog PCICgrid (Figure 20a) shows a generally uniform warming trend provincially with the exception of a cooling trend in the Boreal Plains. The annual trends are similar to those of CANGRD (Figure 20b), except the latter has a larger cooling area in the north that expands to the Northern Boreal Mountains and Taiga Plains. The warming trend is found to be stronger based on CANGRD than homog PCICgrid.

In winter, when CANGRD trends in Figure 20d are compared with those from the homog PCICgrid in Figure 20c, a warming trend pattern that covers larger areas to Central Interior and north of the Southern Interior Mountains. Trend magnitudes are also larger for CANGRD in the Northern Boreal Mountains, Taiga Plains and Sub Boreal Interior with trends exceeding 2.5°C/25 years. In the Boreal Plains, CANGRD data show a cooling pattern not seen in homog PCIC. In southern BC, cooling predominates as shown in Figure 20c, especially in Vancouver Island. Smaller areas of cooling trends are found in Figure 20d from the CANGRD data.

For trends in summer, both data sources show warming across the province (Figure 20e, f). Warming is stronger in CANGRD data, especially in the Boreal Plains and southern BC (Figure 20f). Statistically significant warming trends with values above 1.75°C in Southern Interior are identified.

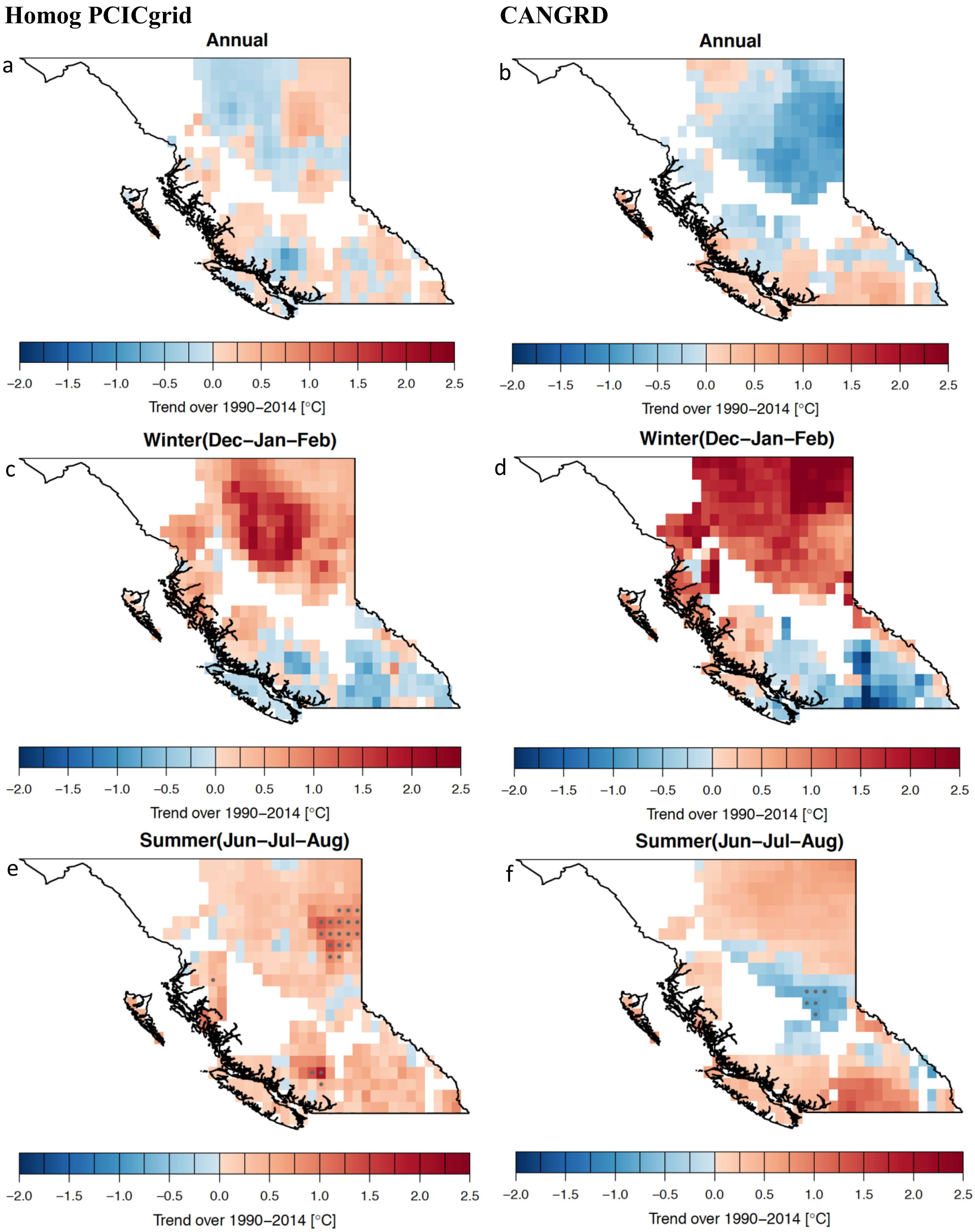


Figure 21: Annual, winter and summer spatial patterns of trends for Tmin over 1990-2014 based on homogPCIC (left) and CANGRD (right). White areas represent grid boxes with insufficient data. Grey dots indicate that the trend in the grid box is statistically significant at the 5% level.

The spatial patterns of trends in annual average Tmin show cooling in the Northern Boreal Mountain and Sub Boreal Interior and warming in the Taiga Plains and part of the Boreal Plains (Figure 21a). In contrast, a strong cooling trend is identified in the CANGRD data for the same areas with a very small area of warming in northern boreal mountains (Figure 21b). Both of the data sources show a mixed pattern of warming and cooling trends in southern BC.

In winter, dramatically large warming of Tmin in northern BC is seen in both datasets (Figure 21c and 21d). For homog PCICgrid, strong warming trends form a circular pattern (above 2.0°C) in the central northern BC where dense network of BCH locations. The magnitude of the warming gets smaller (0.75°C ~1.5°C) on the rim of circular areas in Taiga Plains and Boreal Plains and part of the coastal regions. In comparison, warming is more uniform in the north based on CANGRD. Winter trends in southern BC are comparable between datasets and show a dominant cooling pattern.

Summer trends have consistent warming pattern (0.5°C ~1.5°C/25 years) across BC based on homog PCICgrid in Figure 21e, whereas, cooling trends (below -0.70°C) are found in Sub Boreal Interior and Central Interior based on CANGRD in Figure 21f. Despite these two areas, a warming trend is the primary pattern based on CANGRD. Statistically significant warming trends are identified in the Boreal Plains based on homog PCICgrid. The magnitude of the warming in Southern Interior based on CANGRD is slightly larger (1.0°C ~1.5°C) than that based on homog PCICgrid (0.1°C ~0.8°C).

Chapter 5: Discussion and Conclusion

5.1 Discussion

The patterns of trends based on CANGRD and homog PCICgrid agree well with other, which suggests that IDW produced relatively good results despite occasional problems. Nevertheless, trends based on CANGRD are different in some areas. Possible reasons include data differences and interpolation methods difference, and over-adjustment in the homogenization process, either because of the distance to reference stations or the lack of detailed metadata.

Cooling trends have been found in Vancouver Island in some seasons (Figure 21c,d). This might due to the internal variability since the analyzed period for trend study is not very long (25 years), and coincides with the so called "global warming hiatus" (Fyfe *et al*., 2016). The hiatus coincided with the most recent cold phase of the Pacific Decadal Oscillation (PDO, Whitfield, *et al*., 2010), which covered roughly the last two-thirds of the trend analysis period. Vancouver Island's close proximity to the North Pacific Ocean would suggest strong PDO influences on its climate.

Improvements that can be made in the future are discussed as follows:

1) Overall, corrections and adjustments have been made with statistical confidence, but confidence could be improved further through the use of metadata. Unfortunately, comprehensive metadata documenting the station histories was not available while conducting this research. Once such complementary information is available, it can be used as independent confirmation for the causes of these non-climatic shifts.
2) Efforts have been focused on the monthly temperature homogenization work reported here; daily temperature homogenization work and monthly precipitation homogenization

work to be reported elsewhere. Homogenized data will be available for download in the future.

3) Data quality control has been focused on basic range and consistency checks on individual stations. Future work will advance to multi-location checks with more complex and comprehensive error detection and correction.
4) PMT method requires ten years of data. Except the overall good data continuity in the BCH network, many of the stations data being analyzed have lots of gaps, which make time series that is shorter than 10 years undetectable. For example, summer months within the MoTI network. A homogenization method which takes care of such kind of short records is preferred in the future work.
5) A large data gaps (2010-2014) within the FLNRO_WMB prevented trend analysis for numerous stations. This gap has been filled and will be utilized in future studies, which enable further homogenization and trend analysis.
6) In terms of the spatial trend analysis, spatial interpolation was performed with IDW as a computationally effective method. However, limitations exist for using this method. For example, several grid boxes have odd cooling or warming trends, which might be due to the influence of incorporating one or more station values with large positive or negative value.
7) A better approach for creating a smooth gridded field in BC would be to integrate multiple data sources, such as homogenized PCIC stations and AHCCD and additional stations in the adjacent provinces. The locations of PCIC stations and AHCCD stations complement each other, which produces more uniform station density spatially.

5.2 Conclusion

Detection and adjustment of multiple changepoints based on the statistical methods, with the use of homogenous reference series, have been performed on the monthly observational temperature data across British Columbia. Quality control was undertaken on the daily time scale prior to monthly averaging to reduce data errors before homogenization process. Single station quality control checks were performed using per-station statistics to define thresholds to identify outliers and inconsistent data in this research.

The quality control results show that while BCH has the largest fraction of stations with outliers, it has the lowest fraction of 0.98%. FLNRO_WMB stations have the highest fraction of outliers with 14%. The outlier rate of MoTI is 4%. Homogenization results show that 92 out of 310 stations (29.6%) for maximum temperature and 75 out of 307 stations (24.4%) for minimum temperature have no detectable changepoints, which means they are likely homogenous. BCH has the highest portion of stations with detected changepoints. More specific, there are 73.8% and 60.7% of the maximum and minimum temperature records in the BCH network have changepoints. The changepoint rate of BCH also ranks highest (0.4% for both Tmax and Tmin), with the same percentage as MoTIe for minimum temperature. FLNRO_WMB ranks the lowest within the three networks. It has 10.5% and 27.3 % of the Tmax and Tmin that is detected to have changepoints. The changepoint rates are also smallest for this network, with 0.1% and 0.2% for both Tmax and Tmin respectively. MoTIe's changepoint rate is also 0.2% for Tmin.

In order to evaluate the climate conditions at locations where no direct measurements are collected, interpolation and spatial averages are used in this paper to obtain the spatial pattern of trends for BC and regional average trends for three sub-regions of BC using the homogenized data product we produced and the raw temperature data. In order to validate the trends results, trend based on AHCCD and CANGRD products are compared. Due to the limitations of PCIC station data coverage, trends at 80 and 81 individual stations for Tmax and Tmin variable were analyzed over 1990-2014. The single station trend results have good agreement with results of AHCCD. Spatial patterns of trends that are based on the interpolated PCIC stations also agree well with those based on CANGRD products. Warming trends have been noticed temporally and spatially. Most seasons have been found to have distinctive positive trends across the province, such as summer trends based on homog PCICgrid and CANGRD in Figure 20e and f, with exception of winter and some seasons over Vancouver Island (Figure 20 c, d). Comparing with the trends before homogenization, trends derived from the homogenized PCIC stations generally have the same sign but have smaller magnitude overall (Figure 19).

In general, our results are consistent with the findings from Easterling (1997) that the homogeneity adjustments have little effect on large-area averages, but they can have a noticeable effect on smaller regions, particularly when comparing trends at individual or adjacent grid boxes.

References

Aguilar E., Auer I., Brunet M., Peterson T.C., 2003: Guidelines on Climate Metadata and Homogenization. WMO/TD No. 1186, World Meteorological Organization.

Alaka, M.A., and R.C. Elvander (1972) Optimum interpolation from observations of mixed quality. Mon. Weather Rev. 100(8): 612-624.

Alexandersson, H., 1986: A Homogeneity Test Applied to Precipitation data. *Journal of climatology*, 6, 661-675.

Atkinson, D.E., Gajewski, K., 2002: High-Resolution Estimation of Summer Surface Air Temperature in the Canadian Arctic Archipelago. American Meteorological Society, 15: 3601-3614.

Auchmann R, Brönnimann, S. (2012) A physics based correction model for homogenizing sub-daily temperature series. J Geophys Res 117:D17119

Bennet, K.E., Murdock, T.Q., Rodenhuis. (2009) Update and Errata, Climate Review. Pacific Climate Impacts Consortium, Univeristy of Victoria, Victoria BC: 16pp.

Burger, G. (2017) On trend detection. Hydrological Processes, 31: 4039-4042.

Burrough, P.A., and R.A. McDonnell. (1998) Principles of Geographical Information Systems. New York: Oxford University Press.

Bretherton, F.P., R.E. Davis and B. Fandry. (1976) A technique for objective analysis and design of oceanographic experiments applied to MODE-73. Deep Sea Research, 23: 559-582.

Brönnimann, S., 2015: Climatic Changes Since 1700 (pp.23-29). Switzerland: Springer International Publishing.

Carrega, P., 2010: Geographical Information and Climatology. London & Hoboken, SW & NJ: ISTE & WILEY.

Costa, A.C., and Soares, A., 2008: Homogenization of Climate Data: Review and New Perspectives Using Geostatistics. Math Geosci, 41: 291-305.

Devore, J.L. (2008) Probability and Statistics for Engineering and the Science, Seventh Edition. USA: Thomson Brooks/Cole.

Ducre- Robitaille, J.-F., L. A. Vincent, and G. Boulet, 2003: Comparison of techniques for detection of discontinuities in temperatures, *Int. J, Climatol*., 23, 1087-1101.

Easterling, D.R. and Wehner, M.F. 2009: Is the climate warming or cooling? Geophysical Research Letter. 36: L08706.

Easterling, D.R., Horton, B., Jones, P.D., et al. 1997. Maximum and Minimum Temperature Trends for the Globe. Science Vol.277, 364-366.

Easterling DR, Peterson TC. 1995. A new method for detecting and adjusting for undocumented discontinuities in climatological time series. International Journal of Climatology. 15: 369-377.

Evans, J.D. (1996). Straightforward statistics for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing.

Fannin, R.J., Jaakkola, J., Wilkinson, J.M.T. and Hetherington E.D. (2000) Hydrologic response of soils to precipitation at Carnation Creek, British Columbia, Canada. Water Resources Research, 36(6): 1481-1494.

Gandin, L.S. (1965) Objective Analysis of Meteorological Fields, 242 pp., Israel Program for Sci. Transl., Jerusalem.

Hausfather, Z., Cowtan, K., Menne, M.J., Williams Jr. C. N. (2016) Evaluating the impact of U.S. Historical Climatology Network homogenization using the U.S. Climate Reference Network. Geophysical Research Letters, 43(4): 1695-1701.

Hartkamp A.D., De Beurs K, Stein A, White JW. (1999) Interpolation Techniques for Climate Variables, NRG-GIS Series 99-01. Mexico, D.F.: CLIMMYT.

Hijmans, R.J., Cameron, S.E., Parra, J.J., et al. (2005) Very High Resolution Interpolation Climate Surfaces for Global Land Areas. *Int. J. Climatol.* 25: 1965-1978.

Hunt, B. G. (2007) Climatic outliers. *Int. J. Climatol*. 27: 139-156.

Hutchinson MF., and Xu T. (2013) Anusplin Version 4.4. Fenner School of Environment and Society. The Australian National University: Canberra, Australia.

Hutchinson MF. (1995) Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems. 9: 385-430.

Hoaglin, D.C., F. Mosteller, and J.W. Tukey, 1983: Understanding Robust and Exploratory Data Analysis. *John Wiley and Sons*, New York, 447 p.

Isaaks, E. H and Srivastava, R. M. (1989) An Introduction to applied geostatistics. New York: Oxford University Press.

Fyfe, J.C., Meehl, G.A., et al., (2016) Making sense of the early-2000s warming slowdown. Nature Climate Change. 6: 224-228.

Jiang, T., Su, B., Hartmann, H., 2007. Temporal and spatial trends of precipitation and river flow in the Yangtze River Basin, 1961-2000. Geomorphology 85: 143-154.

Karl, T.R., Arguez, A., etc. (2015). Possible artifacts of data biases in the recent global surface warming hiatus. Science, 348(6242): 1469-1472.

Kendall, M. G., 1955: Rank Correlation Methods, 2nd ed., Charles Griffin, London.

Kosaka, Y. and Xie, S.P. 2013: Recent global-warming hiatus tied to equatorial Pacific surface cooling. Nature, 501: 403-407.

Lanzante, J.R., 1996: Resistant, Robust and Non-Parametric Techniques for the Analysis of Climate Data: Theory and Examples, including Applications to Historical Radiosonde Station Data. *International Journal of Climatol.*, 16, 1197-1226.

Li, Q., L. Zhang, W. Xu, T. Zhou, J. Wang, P. Zhai, and P. Jones (2016) Comparisons of time series of annual mean surface air temperature for China since the 1990s: Observations, model simulations and extended reanalysis. Bull. Amer. Metero. Soc.

MacEachren, A.M., and J.V. Davidson. (1987) Sampling and isometric mapping of continuous geographic surfaces. The American Cartographer. 14:299-320.

McGree, S., & Whan, K., et al. (2014) An updated assessment of trends and variability in total and extreme rainfall in the western Pacific. Int. J. Climatol., 34: 2775-2791.

McKenny, D., W., Hutchinson, F.M., et al. (2011) Customized Spatial Climate Models for North America. American Meteorological Society, BAMS, Dec. 2011: 1611-1622.

Mekis, É. and L.A. Vincent, 2011: An overview of the second generation adjusted daily precipitation dataset for trend analysis in Canada. Atmosphere-Ocean, 49(2): 163-177.

Menne, M.J., Jr., C. N. W., Palecki, M.A., 2010. On the reliability of the U.S. surface temperature record. Journal of Geophysical Research Vol.115, D11108, 1-9.

Menne, M.J., and Williams JR, C.N., 2009: Homogenization of Temperature Series via Pairwise Comparison. *Journal of Climate*, 22, 1700-1717.

Milewska, E., Hogg, W.D. (2001) Spatial representativeness of a long-term climate network in Canada, Atmosphere-Ocean, 39:2, 145-161.

Mishra, N., Khare, D., Shukla, R., Kumar, K., 2014. Trend Analysis of Air Temperature Time Series by Mann Kendall Test- A Case Study of Upper Ganga Canal Command (1901-2002). British Journal of Applied Science & Technology, 4(28): 4066-4082.

National Avalanche Center. "Manual Snow and Weather Observations". Retrieved from https://static1.squarespace.com/static/53fbfe1ce4b04ce86372f366/t/5452f4ece4b04a8f685ec37d/1414722796411/Chapter1.pdf

National Center for Atmospheric Research Staff (Eds). Last modified 05 Sep 2014. "The Climate Data Guide: Trend Analysis." Retrieved from https://climatedataguide.ucar.edu/climate-data-tools-and-analysis/trend-analysis.

Neha Karmeshu, 2012. Trend Detection in Annual Temperature & Precipitation using the Mann Kendall Test- A Case Study to Assess Climate Change on Select States in the Northeastern United States. University of Pennsylvania Scholarly Commons.

Peterson, T.C., Easterling D.R., et al. (1998) Homogeneity Adjustments of in situ Atmospheric Climate Data: A review. International Journal of Climatology, 18: 1493-1517.

Reeves, J., J. Chen, X. L. Wang, R. Lund, and Q. Lu (2007), A review and comparison of changepoints detection techniques for climate data. *J. Appl. Meteorol. Climatol.*, 46, 900-915.

Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63, 1379-1389.

Shepard, D. 1968. A Two-Dimensional Interpolation Function for Irregularly-Spaced Data. Proceedings-ACM National Conference. 517-524.

Thiebaux, H.J. and Zwiers, F.W., 1984: The Interpolation and Estimation of Effective Sample Size. Journal of Climate and Applied Meterology, 23, 800-811.

Wan, H., X. L. Wang, V. R. Swail, 2010: Homogenization and trend analysis of Canadian near-surface wind speeds. Journal of Climate, 23, 1209-1225.

Wan, H., X. L. Wang, V. R. Swail, 2007: A quality assurance system for Canadian hourly pressure data. J. Appl. Meteor. Climatol., 46, 1804-1817.

Wang, X.L., and Y. Feng, published online July 2013: RHtestsV4 User Manual. Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada. 28pp.

Wang, X.L, Y. Feng, L. A. Vincent, 2013. Observed changes in one-in-20 year extremes of Canadian surface air temperatures. Atmosphere-Ocean. Doi:10.1080/07055900.2013.818526.

Wang, X.L., 2008: Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test. *J. Appl. Meteor. Climatol.,* 47, 2423-2444.

Wang, X.L., Q.H. Wen, and Y. Wu, 2007: Penalized maximal t test for detecting undocumented mean change in climate data series. *J. Appl. Meteor. Climatol.,* 46 (No. 6), 916-931. DOI: 10.1175/JAM2504.1

Wang, X.L., 2008a: Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t of F test. *J. Appl. Meteor. Climatol.*, 47, 2423-2444.

Wang, X.L., 2008b: Penalized maximal F-test for detecting undocumented mean-shifts without trend-change. *J. Atmos. Oceanic Tech.*, 25 (No. 3), 368-384. DOI: 10.1175/2007/JTECHA982.1

Wang, X.L., H. Chen, Y. Wu, Y. Feng, and Q. Pu, 2010: New techniques for detection and adjustment of shifts in daily precipitation data series. *J. Appl. Meteor. Climatol.* 49 (No.12), 2416-2436. DOI: 10.1175/2010JAMC2376.1

Whitfield, P.H., Moore, R.D., Flemming, S.W., and Zawadzki, A. (2010) Pacific Decadal Oscillation and the Hydroclimatology of Western Canada – Review and Prospects. Canadian Water Resources Journal, 35(1): 1-28.

Wilks, Statistical Methods in the Atmospheric Sciences, Third Edition. Chapter 5: Frequentist Statistical Inference, 5.3.2 Mann-Kendall Trend Test. P166-168.

Venema, V.K.C., et al., 2012: Benchmarking homogenization algorithms for monthly data. *Clim. Past*, 8, 89-115.

Vincent, L. A., Zhang, X., et al. (2015) Observed Trends in Canada's Climate and Influence of Low- Frequency Variability Modes. Journal of Climate, 28: 4545-4560.

Vincent, L. A., X.L. Wang, E. J. Milewska, H. Wan, Y. Feng, and V. Swail, 2012: A second generation of Homogenized Canadian Monthly Surface Air Temperature for Climate Trend Analysis, *JGR-Atmospheres*, 117, D18110, doi: 10. 1029/2012JD017859. *[CORE PAPER]*

Vincent, L.A., Zhang, X. (2001) Homogenization of Daily Temperature over Canada. Journal of Climate, 15: 1322-1334.

Vincent, L. A., Gullett, D.W. (1999) Canadian Historical and Homogenous Temperature Datasets for Climate Change Analyses. Int. J. Climatol., 19: 1375-1388.

Vincent, L. A. (1997) A Technique for the Identification of Inhomogeneities in Canadian Temperature series. Journal of Climate, 11: 1094-1104.

Whan, K., & Alexander L.V., et al. (2014) Trends and variability of temperature extremes in the tropical Western Pacific. International Journal of Climatology, 34: 2585-2603.

Xu, W. H., Li, Q. X., Yang, S., Xu, Y. (2014) Overview of global monthly surface temperature data in the past century and preliminary integration. Advances in Climate Change Research 5: 111-117.

Yue, S., P. Pilon, B. Phinney and G.Cavadias, 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrological Processes, 16: 1807-1829.

Zhang, X., Vincent, L.A., Hogg, W.D. and Niitsoo, A., 2000. Temperature and Precipitation Trends in Canada during the 20th Century. Atmosphere- Ocean 38(3): 395-429.

Appendix A: Some Statistical methodologies

1. Biweight mean and standard deviation

The biweight is a sophisticated estimator belonging to the category known as redescending estimators (Lanzante, 1996). The more detailed underlying methodology is described in Hoaglin et al. (1983).

This method gives the more extreme observations in a sample reduced or zero weight when calculating the mean and standard deviation. The value beyond a certain critical distance from the centre, which is controlled by the parameter 'c', is given zero weight. A 'c' value of between 6 and 9 is recommended (Hoaglin et al., 1983). In this thesis for the Gaussian case, c equals to 6 is used which censors values more than 4 standard deviations from the mean. The censoring of outliers is performed using resistant estimates of location (median) and scale (median absolute deviation), which are used to determine the weights.

First the median (M) and median absolute deviation (MAD) are estimated. The MAD is the median of the absolute values of the deviations of the sample values from the median. A distance (u_i) corresponding to each of the n observations (X_i) is computed as follows:

$$u_i = (X_i - M)/(c \times MAD) \quad \text{(A1)}$$

For any $|u_i| \geq 1.0$, set $u_i = 1.0$ to accomplish the censoring. The biweight estimate of the mean is then given by:

$$\overline{X}_{bi} = M + \left(\sum_{i=1}^{n} (X_i - M)(1-u_i^2)^2 \Big/ \sum_{i=1}^{n} (1-u_i^2)^2 \right) \text{ (A2)}$$

$$= \sum_{i=1}^{n} x_i (1 - u_i{}^2)^2 / (1 - u_i{}^2)^2 \text{ (A3)}$$

This has the effect downweighting observations that are farther from the median since the weight $(1 - u_i{}^2)^2$ range between 1 for observations equal to the median to 0 for observations more than c MAD's from the median. Similarly, the biweight estimate of the standard deviation is:

$$s_{bi} = \left[n \sum_{i=1}^{n} (X_i - M)^2 (1-u_i^2)^4 \right]^{0.5} \Big/ \left| \sum_{i=1}^{n} (1-u_i^2)(1-5u_i^2) \right| \text{ (A4)}$$

Equations of (A1) to (A4) are from Lanzante (1996).

2. Mann Kendall Trend Test

Mann Kendall test is widely used for the analysis of trends in climatic and hydrologic time series. According to this test, the null hypothesis H_0 assumes that there is no trend for the random data which is independent and distributed. The alternative hypothesis $H_1$is that there is a trend for the sample data.

The Mann Kendall test statistic S is calculated in Equations (A5) and (A6). For a given series with length n that is greater than 10, the statistic S approximately follows a Gaussian distribution. The variance of the sampling distribution of S is then computed by Equation A (8).

$$S = \sum_{i=1}^{n-1} \operatorname{sgn}(x_{i+1} - x_i) \text{ (A5)}$$

where

$$\mathrm{sgn}(\Delta x)=\begin{cases}+1, \Delta x>0\\ 0, \Delta x=0.\\ -1, \Delta x<0\end{cases} \quad \text{(A6)}$$

$$Var(S)=\frac{n(n-1)(2n+5)-\sum_{j=1}^{J} t_j(t_j-1)(2t_j+5)}{18} \quad \text{(A7)}$$

Notation t_j denotes the number of ties to the extent of j. The test p value is estimated from the standard Gaussian value z and is computed using the equation below:

$$z=\begin{cases}\dfrac{S-1}{\left|Var(S)\right|^{1/2}}, S>0\\ \dfrac{S+1}{\left|Var(S)\right|^{1/2}}, S<0\end{cases}. \quad \text{(A8)}$$

If $z > z_{1-\alpha/2}$ or $z < z_{\alpha}$, then the trend is said to be significant. In this research, a p value of $\alpha = 0.1$ is utilized. Therefore, the null hypothesis is rejected if $z > z_{0.95}$ or $z < z_{0.05}$. A positive value of S represents an increasing trend; similarly, a negative represents a decreasing trend. Equations from (A5) to (A8) are cited from Wilks (2011, *pp 166-167*).

Appendix B: Supplementary Tables and Figures

Year	Month	Day	Precip.	Tmax	Tmin
1960	1	1	-99.9	-3.1	-6.8
1960	1	2	-99.9	-4.2	-7
1960	1	3	-99.9	-1.8	-8.4

Table B1: Example input data file that the RClimDex package requires. The format has to be in ASCII.

The daily data were stored as individual files for a given network and variable. For example, all 410 FLNRO_WMB stations of maximum temperature are stored in a large data table, with separate tables for minimum temperature and precipitation variable. Table B2 shows an example for Tmax from FLNRO_WMB.

station_id	vars_id	obs_day	tmax	cell_method
1611	497	1991-10-10 00:00:00	25.5	time: maximum
1611	497	1991-10-10 00:00:00	22.1	time: maximum
1611	497	1991-10-10 00:00:00	17.7	time: maximum
1611	497	1991-10-10 00:00:00	24.1	time: maximum
1611	497	1991-10-10 00:00:00	20.3	time: maximum
1611	497	1991-10-10 00:00:00	7.9	time: maximum
1611	497	1991-10-10 00:00:00	5.6	time: maximum
1611	497	1991-10-10 00:00:00	4.2	time: maximum
1611	497	1991-10-10 00:00:00	4.2	time: maximum
1611	497	1991-10-10 00:00:00	6.1	time: maximum
1611	497	1991-10-10 00:00:00	3.7	time: maximum
1611	497	1991-10-10 00:00:00	-2.0	time: maximum
1611	497	1991-10-10 00:00:00	5.0	time: maximum

Table B2: Sample original maximum temperature data (part) of FLNRO_WMB for all stations

Separating stations for each variable and network is performed by selecting the data with a given station id. Precipitation, maximum temperature and minimum temperature data for each of the individual station is realized in Table B3 that matches the requirement in Table B1.

Year	Month	Day	PRCP	TMAX	TMIN
1991	1	1	-99.9	-99.9	-99.9
1991	1	2	-99.9	-99.9	-99.9
1991	1	3	-99.9	-99.9	-99.9
1991	1	4	-99.9	-99.9	-99.9
1991	1	5	-99.9	-99.9	-99.9
1991	1	6	-99.9	-99.9	-99.9
1991	1	7	-99.9	-99.9	-99.9
1991	1	8	-99.9	-99.9	-99.9
1991	1	9	-99.9	-99.9	-99.9
1991	1	10	-99.9	-99.9	-99.9
1991	1	11	-99.9	-99.9	-99.9

Table B3: Processes data which meets the format requirement of RClimDex package

Year	Month	Day	Value
1967	7	00	28.1
1967	8	00	32.6
1967	9	00	-999.99

Table B4: Monthly series data format requirement for RHTestsV4 (-999.99 is the missing value code)

Table B4 shows the required data format for RHTests V4. The file has to be saved in a separate file in ASCII format. Dates of the input data must be consecutive and in the calendar order. Otherwise, the error message will be given by the software.

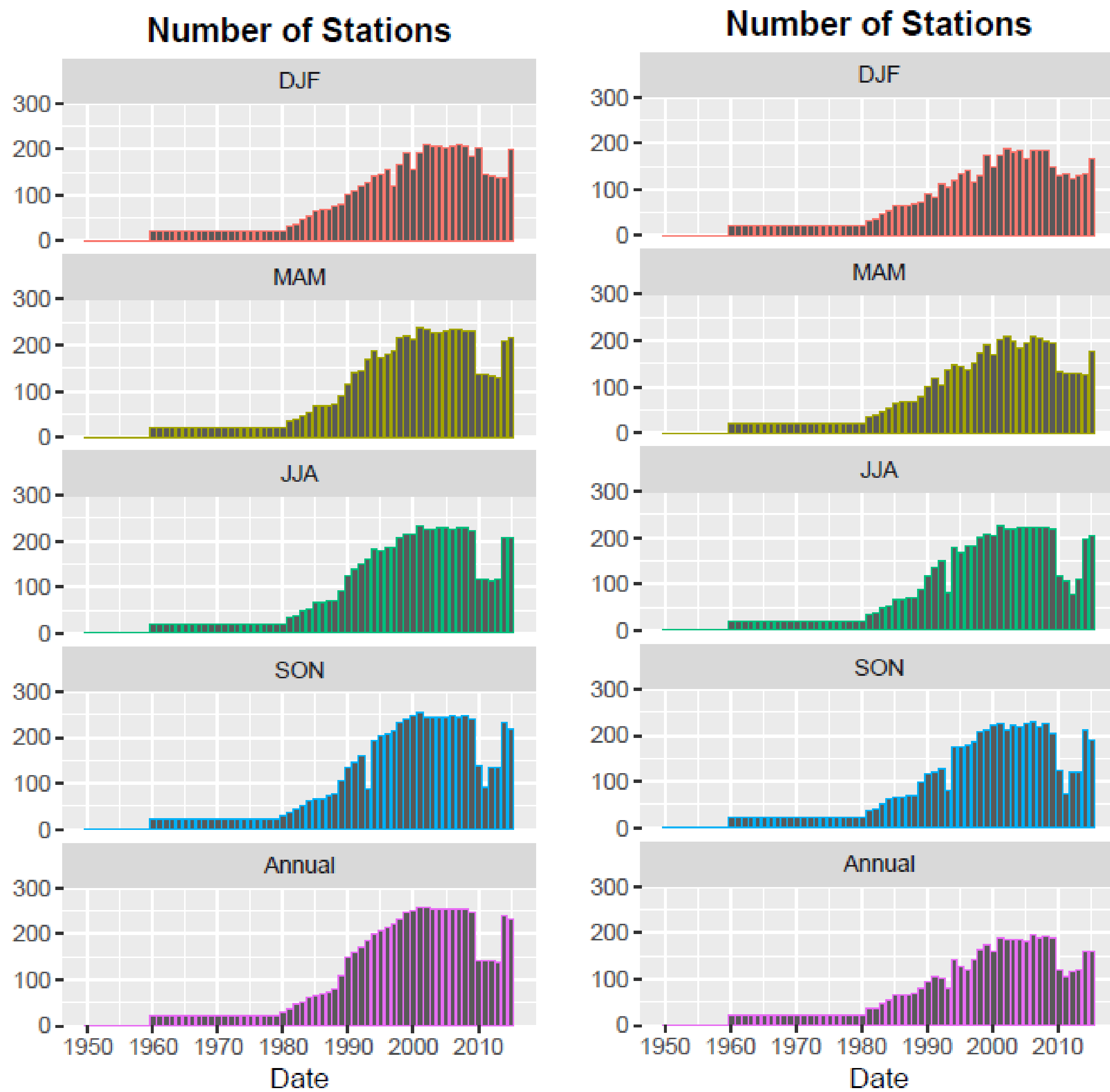


Figure B1: Number of stations per year and per season for all networks. Tmax results are on the left panel, Tmin are on the right.

In order to have a more detailed look at the temporal data coverage for each season, Figure B1 is produced. It shows the number of stations per year for each specific season for both the Tmax (right) and Tmin (left). This figure directly shows that most of the data are centered from 1990 to 2015, which provides helpful information for the selection of trend period. As explained in the main text in SECTION 2.1.1, small numbers of stations within 1960 to 1980 are mostly from the BCH network. The drop of stations from 2010 to 2014 is due to the gaps of FLNRO_WMB.

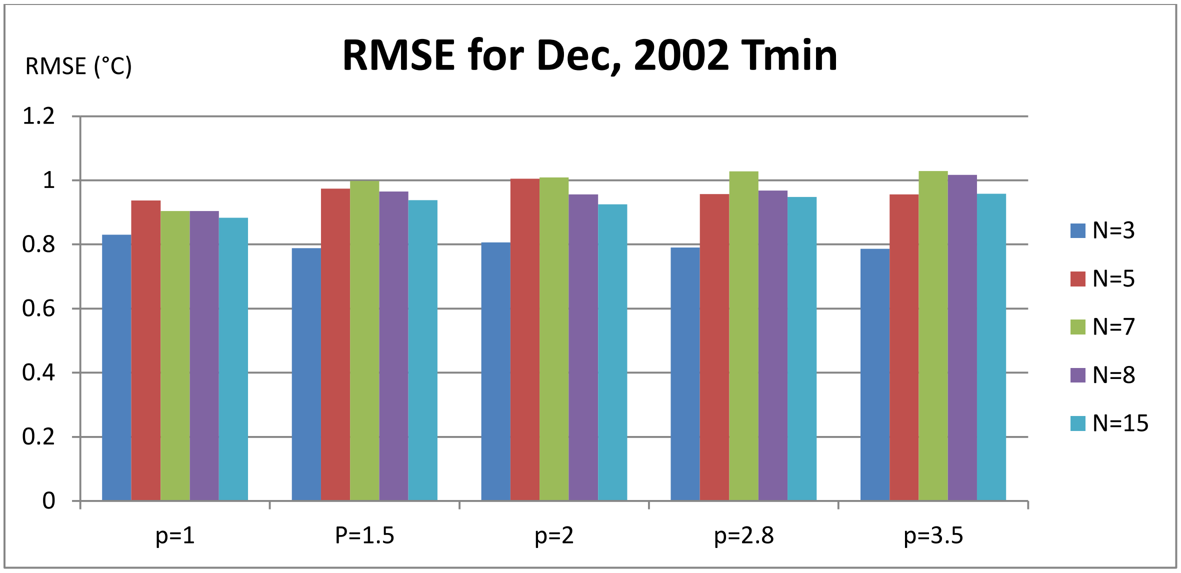


Figure B2: Bar chart shows RMSE for a winter month (December, 2002) for Tmin anomalies based on the PCIC homogenized stations for various N and p.

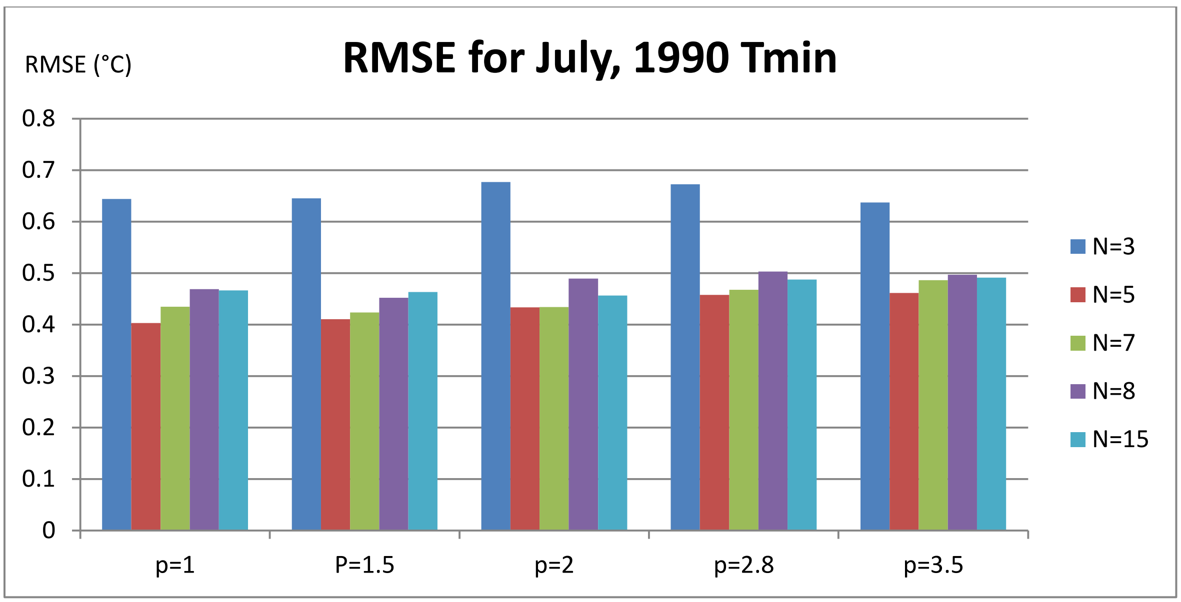


Figure B3: Bar chart shows RMSE for a summer month (July, 1990) for Tmin anomalies based on the PCIC homogenized stations for various N and p.

Figure B2 and Figure B3 show the RMSE values for a specific winter of December 2000 and a specific summer month of July 1990 Tmin anomalies based on the PCIC homogenized stations under various p (1, 1.5, 2, 2.8, and 3.5) and N (3, 5, 7, 8, 15) values. RMSE in these two seasons are quite different. For example, when N=5, RMSE is smaller for July than for December, which means the interpolation performance is better for July. The reason of that is because the spatial variability is large in winter. The inclusion of less correlated stations will lose the representation of the estimation of the local temperature.

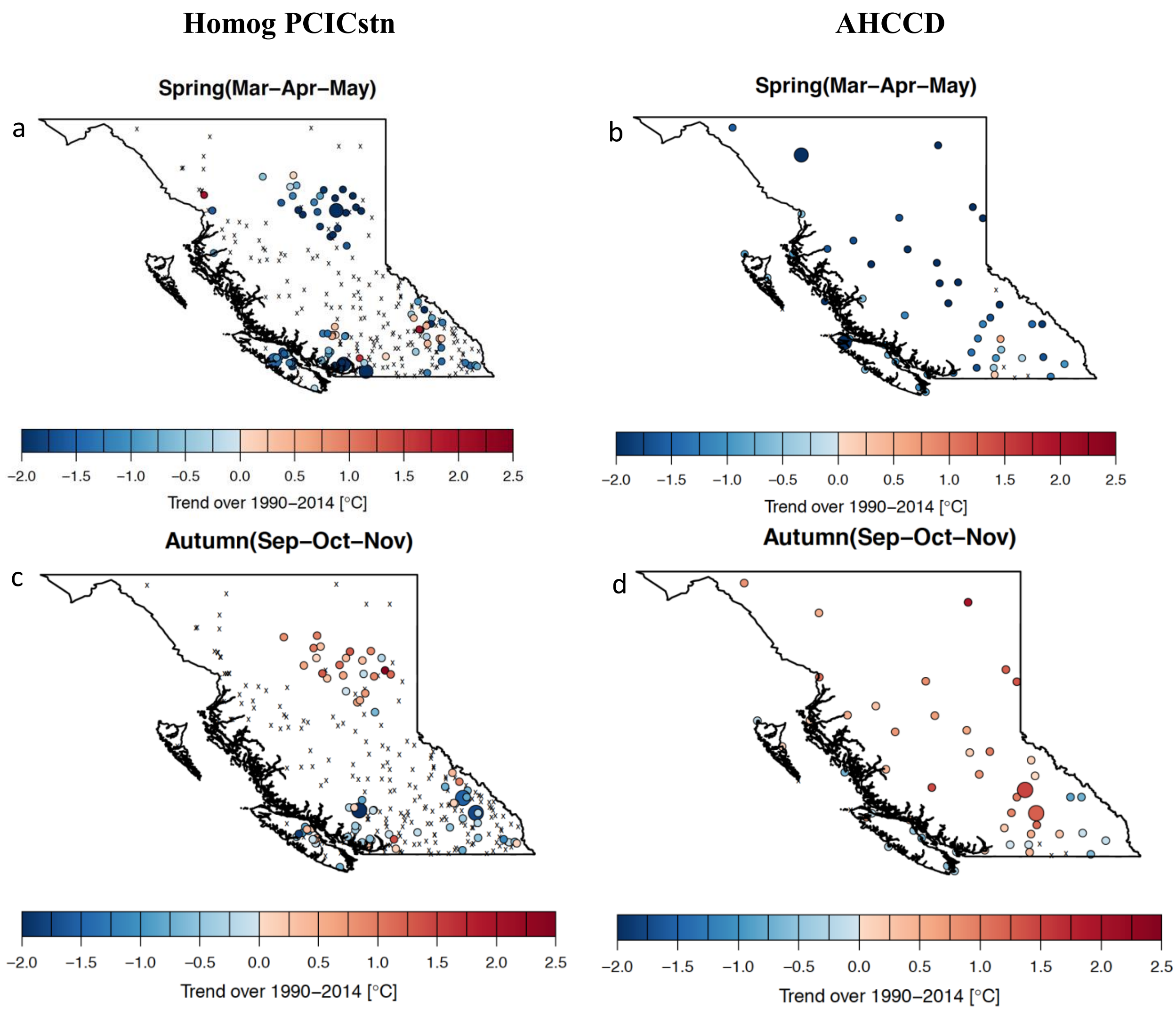


Figure B4: Left panel show the spring and autumn Tmax trends over 1990-2014 for each station based on homog PCICstn; trends based on AHCCD station are listed on the right panel. Warm color indicate positive trends, cool color indicate negative trends. Crosses represent stations with inadequate data, which means stations with than 20 years' data for trend analysis. Large dots mean trends are significant at the 5% significant level; small dots mean trends are insignificant at the 5% significant level.

Single trends results based on homog PCIC stations and AHCCD stations for Tmax are shown in Figure B4. Regarding the trends on spring anomalies (Figure B4 a, b), both of the homogenous data sources illustrate cooling trends in BC. For trend map of homog PCICstn in Figure B4 a, a few stations in Northern BC and several stations in the south do show positive trends. Statistically significant trends are found in the sub boreal interior, Vancouver Island and Geogria depression. In Figure B5 b, distinct negative trends are uniformly distributed across the province, with the exception of only two stations showing positive trends. Two stations have statistically significant trends, one in the northern boreal mountains, the other station located in upper Vancouver Island. Autumn trends in Figure B4 c and d both show warming trend in the northern BC, except few stations with cool trends based on homog PCICstn. Trends in southern BC have opposite sign based on PCIC data and AHCDD in general.

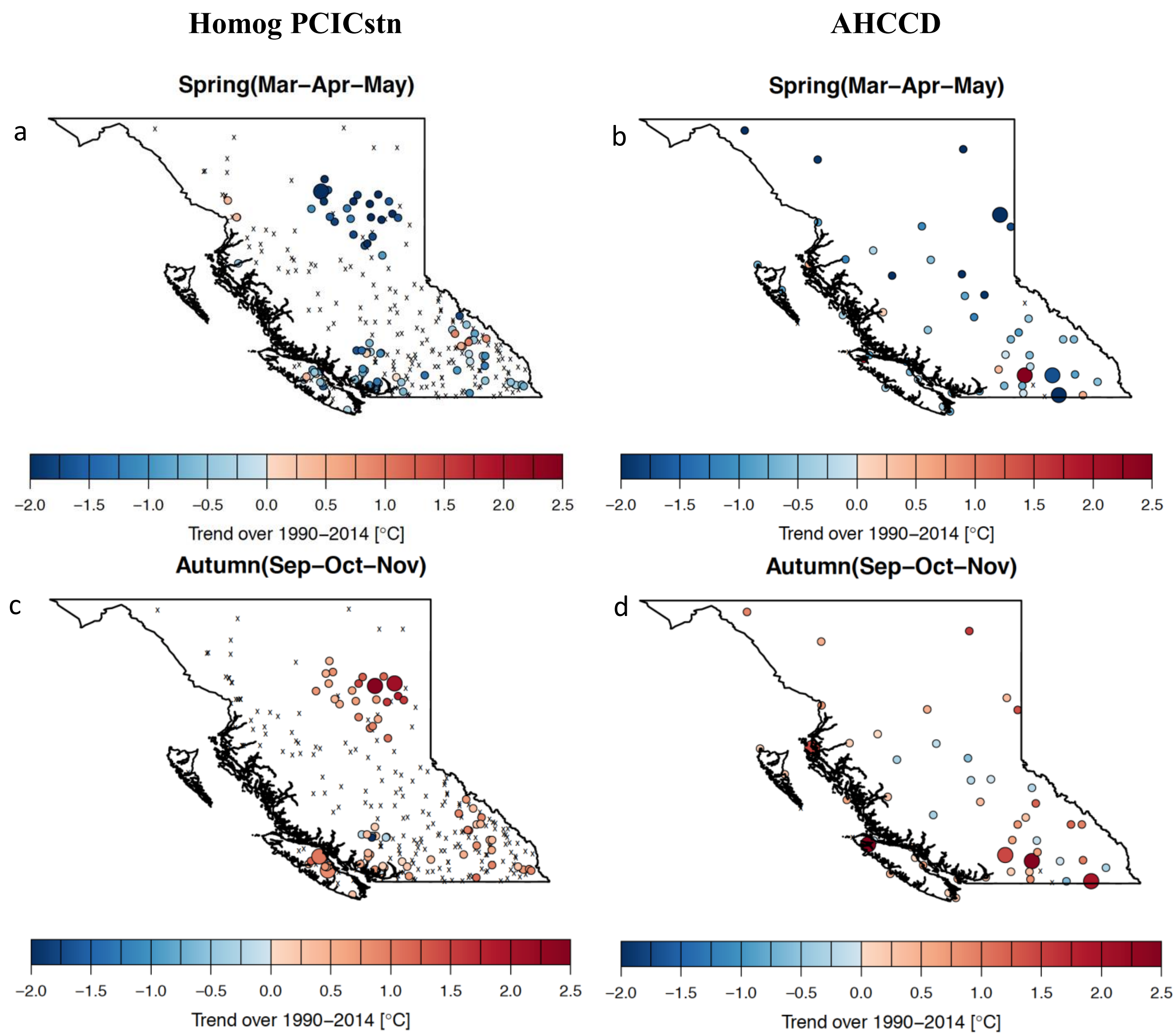


Figure B5: Left panel show the spring and autumn Tmin trends over 1990-2014 for each station based on homog PCICstn; trends based on AHCCD station are listed on the right panel. Warm color indicate positive trends, cool color indicate negative trends. Crosses represent stations with inadequate data, which means stations with than 20 years' data for trend analysis. Large dots mean trends are significant at the 5% significant level; small dots mean trends are insignificant at the 5% significant level.

Spring trends for Tmin are very similar to those for Tmax. Cooling trends are identified across the province for both homog PCICstn and AHCCD stations. Only few stations with warming trends can be identified. In Figure B5 a, displaying homog PCICstn trends in spring a few warming points appear in southern BC, with only two warming stations in the boundary of sub-boreal interior and coast and mountains. Northern parts cool more strongly than southern regions based on the PCIC stations. In Figure B5 b, cooling pattern is uniformly spread out. Three statistically significant trends are showed for AHCCD stations, with two in the southern interior mountains and one in the boreal plains. One exception of a statistically warming trend is identified in southern interior. Generally, the PCIC and AHCCD trends are broadly consistent. For autumn trends, both data sources have dominant warming trends in general with exception of several cooling trends in central interior in Figure B5 d that is based on AHCCD stations.

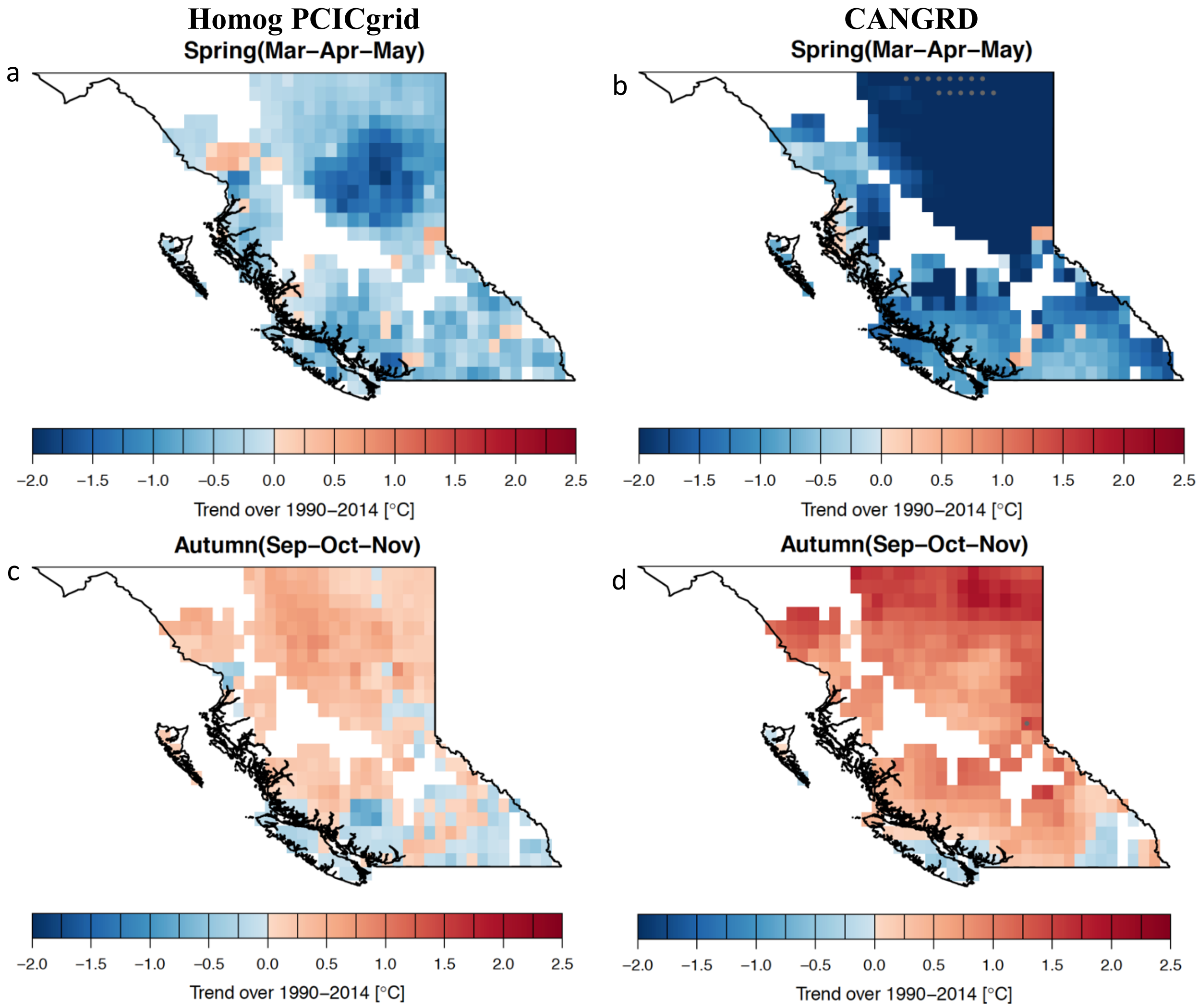


Figure B6: Spring and autumn spatial patterns of trends for Tmax over 1990-2014 based on homogPCIC (left) and CANGRD (right). White areas represent grid boxes with insufficient data. Grey dots indicate that the trend in the grid box is statistically significant at the 5% level.

For the patterns of trends in spring (Figure B6 a, b), dramatic cooling pattern can be seen across the province for both of the data sources. Cooling trends based on CANGRD have stronger magnitude than those based on homog PCICgrid, especially in the northern BC. Regarding the pattern of trends in autumn, it can be seen that northern BC has been warmed distinctively over 1990-2014 based on both homog PCICgrid and CANGRD (Figure B6 c, d). Similarly to the above two seasons, trends based on CANGRD warm faster than those based on PCICgrid, especially in the northern boundary area with the adjacent provinces. The trend pattern is different in the southern BC though, where PCIC data shows dominantly cooling and CANGRD shows warming. Both datasets show cooling, however, in southern Vancouver Island.

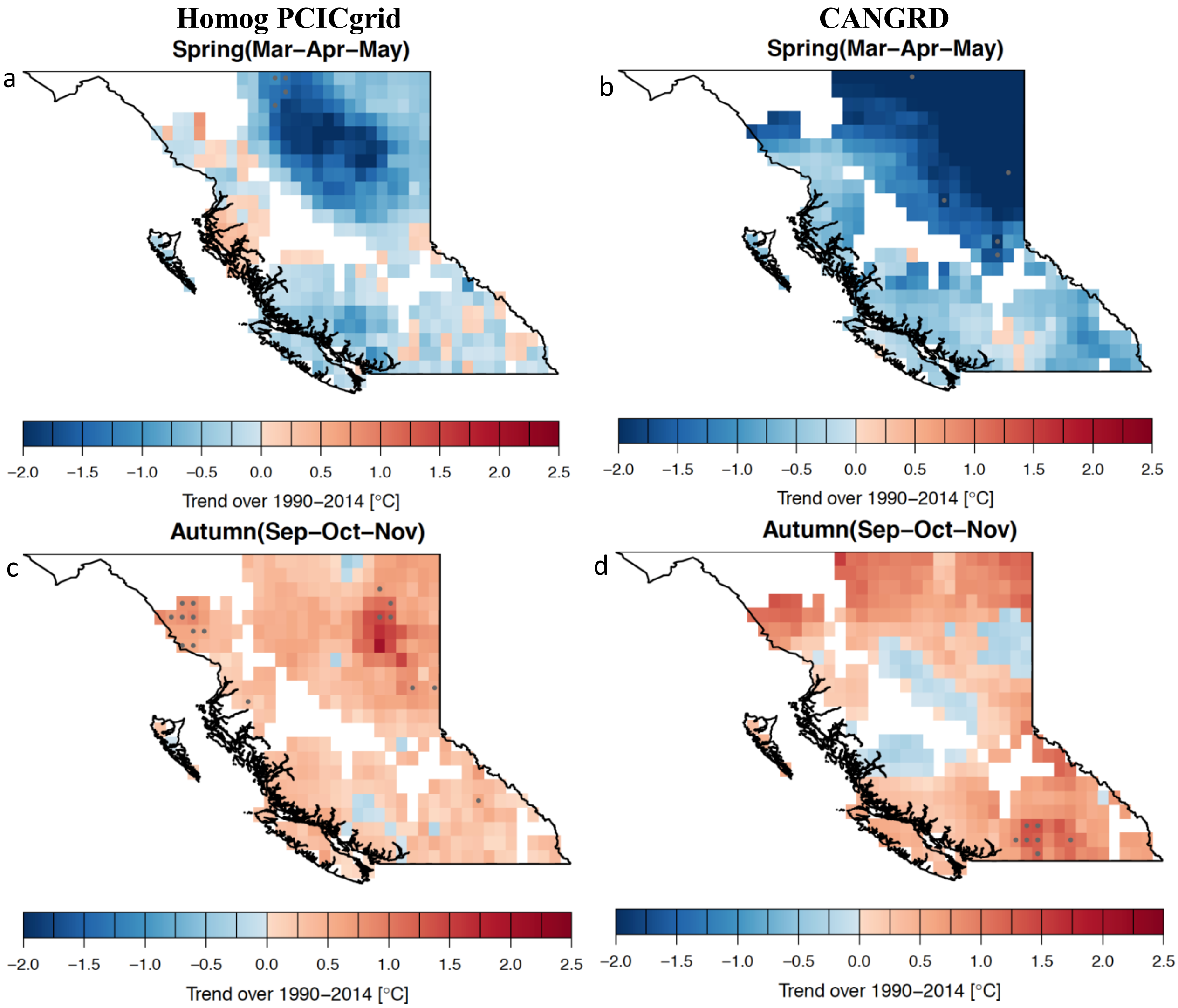


Figure B7: Spring and autumn spatial patterns of trends for Tmax over 1990-2014 based on homogPCIC (left) and CANGRD (right). White areas represent grid boxes with insufficient data. Grey dots indicate that the trend in the grid box is statistically significant at the 5% level.

In terms of the spatial patterns of trends for spring, cooling trends across the province are identified in both maps (Figure B7 a, and B7 b). Opposite patterns in northern BC are found compared with those in winter. Large areas of very strong cooling trends are observed based on both datasets, during which trends based on CANGRD cool more strongly and over larger areas (Figure B7 b) compared with PCICgrid (Figure B7 a). Warming pattern is the pattern for autumn trends as shown in Figure 7 c, d, with exception of small areas of cooling in sub boreal interior and boreal plains based on CANGRD (Figure 7d).