Stationarity analyses of historical Canadian prairie weather data
Why do statistical tests give differing answers?

Kevin Shook and John W. Pomeroy

SGI Canada Hydrometeorology Research Program, Centre for Hydrology, University of Saskatchewan,

117 Science Place, Saskatoon, Saskatchewan, Canada, S7N 5C8

Introduction

- Modeling the hydrology of future climates requires consideration of current climate data. It is important to know if the current climate variables can be considered to be stationary, or if it is necessary to separate more-recent data values from older data, in order to obtain stationary data to act as a baseline for simulations of future climates.
- Weak stationarity requires that the first and second moments (mean and variance) do not vary with respect to time. Strict stationarity also requires the invariance of higher moments such as the skewness and kurtosis.
- Tests of long term (> 100 years) historical precipitation and temperature datasets should determine if the datasets are stationary.
- This research is intended to determine the presence or absence of weak stationarity in data sets of precipitation and temperatures measured in the Canadian prairies.

Selection of data

- The data selected for testing were obtained from Envionment Canada's 2002 CDWest CD-ROM, which collection contains data for 4,442 stations in Western Canada until the end of 2002, and into 2003.
- The data were extracted from the CD image for six locations: Calgary, Medicine Hat, Saskatoon, Regina, Indian Head and Brandon, because these had spatially continuous records (the station did not change location during the period of record) and data preceding the 20th century.
- The variables analyzed were daily measurements of air temperature (min, max, and mean daily values) and precipitation (rainfall, snowfall, and total precipitation).
- Each dataset had gaps in its daily records; the missing data were not filled.
- As windspeed and changes in shielding are known to have affected historical snowfall (and therefore total precipitation) values, rehabilitated snowfall data were obtained from the Historical Adjusted Climate Database for Canada (Mekis and Hogg, 1999).
- Rehabilitated precipitation values were created by adding the rehabilitated daily snowfall data to the measured rainfall at each site.
- The analyses were performed using the tseries package (time series analyses) of the statistical program R, described by Ihaka and Gentleman (1996).
- The program may be downloaded at www.r-project.org.

Analyses

- The ADF (Augmented Dickey Fuller) test is a well known test for the presence of weak stationarity in time series data. Its null hypothesis is that the data set is not stationary.
- The KPSS (Kwiatkowski, Phillips, Schmidt, Shin) test is a more recently developed test of stationarity (Kwiatkowski, et. al, 1992). It is known to be fairly conservative as it has rejected stationarity in datasets which other tests have determined to be stationary.
- The KPSS test evaluates the existence of stationarity about a mean value or about a linear trend.
- The null hypothesis of both types of KPSS test is that the data are stationary.

Results

- All data sets were determined to be stationary by the ADF test, at the 1% level.
- The null hypotheses of stationarity about a mean value could not be rejected by the KPSS test, even at the 10% level, for any of the data sets except for the Indian Head minimum daily air temperatures, which were found to be stationary about a linear trend.

A paradox?

- Many researchers including Akinremi et al. (1999), Skinner and Gullet (1993), and Vincent and Mekis (2005) have found trends in rainfall and precipitation in the Canadian prairies.
- When the previously analyzed data sets are bisected, the halves frequently show large differences in their respective means, as shown in Figure 1.
- Such changes in the mean appear to be due to large-scale trends, the existence of which contradicts the results of the ADF and KPSS analyses, resulting in a paradox.

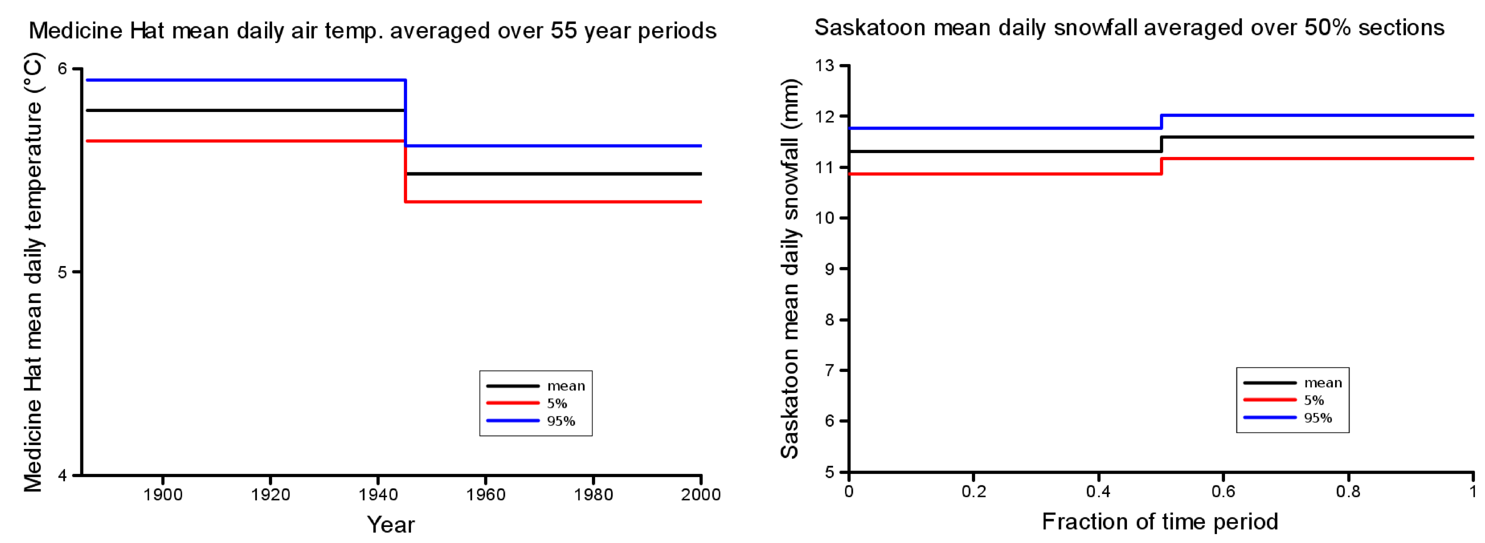

Figure 1:Means of bisected data sets of mean daily air temperature and mean daily snowfall. Confidence levels (5% and 95%) were established through bootstrapping. The temperature data were bisected at year boundaries. The snowfall data were divided into two equal-length segments.

Solution of the paradox

- The solution of the paradox lies in the fact that the data are neither deterministic nor random in their temporal distribution.
- When the data sets are split into smaller sub-sets, the mean value displays frequent significant shifts at smaller scales, as shown in Figure 2.

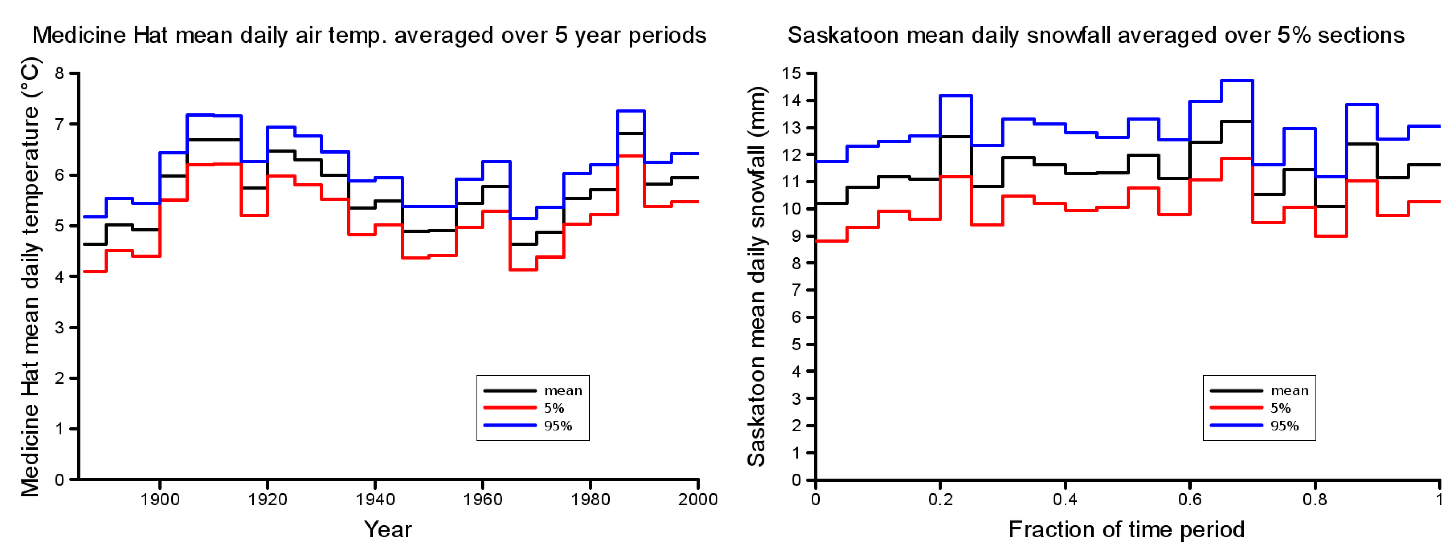

Figure 2: Means of sub-divided data sets of mean daily air temperature and mean daily snowfall. Confidence levels (5% and 95%) were established through bootstrapping. The temperature data were bisected at year boundaries over 5-year intervals. The snowfall data were divided into 20 equal-length segments.

- This behavior is characteristic of fractal datasets, which display trends at all scales.
- The Hurst exponent (H), of a time series indicates if data are randomly distributed or display a long-term memory (Hurst, 1951).
- Values of $H > 0.5$ are an indication that the data are fractal in their distribution.
- The mean H value of the precipitation data sets is 0.55 (standard deviation = 0.02), and of the temperature data sets is 0.63 (standard deviation = 0.02).
- The values of H confirm that the data sets are *weakly* fractal.
- As consequence of their weak fractal structure, the autocorrelation functions (ACF) of these datasets decline very rapidly with lag distance, as shown in Figure 3. Note that autocorrelation is present at daily and yearly scales, which is also a property of fractal data.
- Time series having monotonic trends have ACF values that decline slowly. Truly random data are not autocorrelated.
- As the ACF values of the measured datasets decline very rapidly, the ADF and KPSS tests conclude that the data sets are stationary.

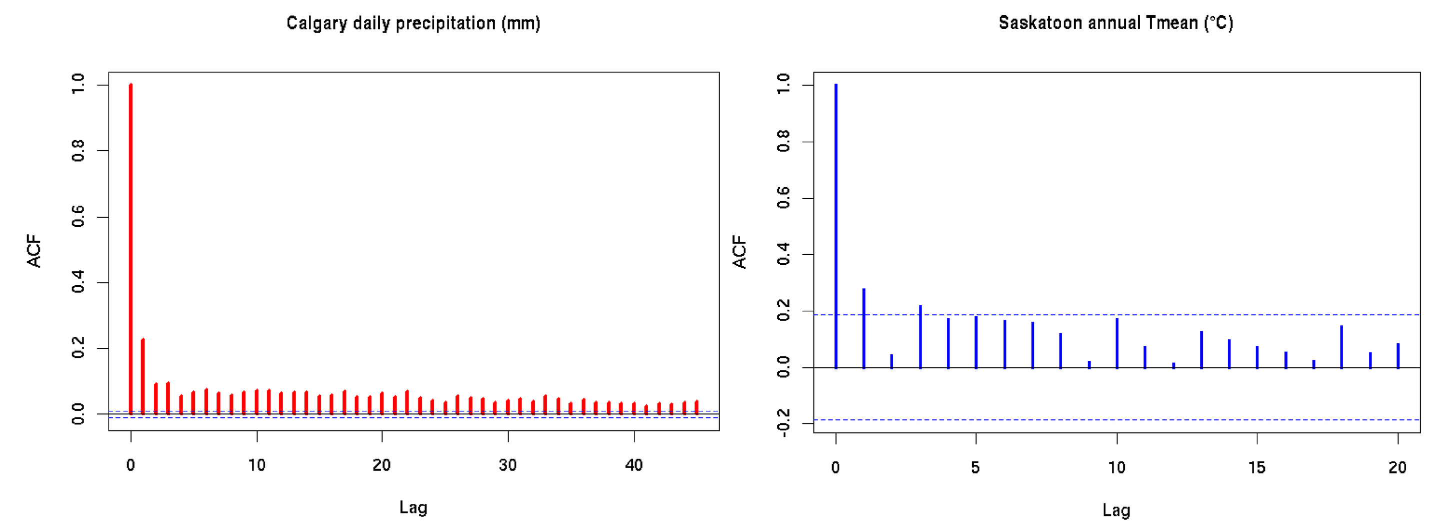

Figure 3: Autocorrelation function vs. lag (days) for Calgary daily precipitation.

Autocorrelation function vs. lag (years) for Saskatoon annual mean air temperature.

Conclusions

- The weak fractal structure of the data sets causes several effects:
  1. Because the data sets are only weakly autocorrelated, their ACF values decline very quickly with lag distance, causing the ADF and KPSS tests to conclude that the data are stationary. Had the data been more strongly fractal, with more slowly declining ACF values, it is likely that the tests would have rejected stationarity.
  2. Because the data sets are non-random, trends do exist – at all scales.
  3. The analyzed data sets are non-stationary – at all scales.
- Stationarity tests such as the KPSS and ADF tests are of little value for weakly fractal datasets.
- Further research is required to determine if the observed weak fractal structure of the data series is characteristic of the prairies, perhaps due to the cold and semi-arid climate, or is a more general property of Canadian hydrometeorological variables,
- Further research is also required to fully characterize and parameterize the observed weak fractal structure of the datasets. When these structures have been characterized, it should be possible to statistically downscale future climate scenarios.

References

Hurst, H. E. (1951) Long term storage capacity of reservoirs. Trans. Amer. Soc. Civil Engrs. 116, 770-779

Kwiatkowski, D, Phillips, P. C. B., Schmidt, P., and Y. Shin. 1992. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. Journal of Econometrics 54, 159–17.

Mekis. E. and W.D. Hogg, 1999. Rehabilitation and Analysis of Canadian Daily Precipitation Time Series. Atmosphere-Ocean, 44(2), 53-85.

Acknowledgments

The authors wish to acknowledge the support of the Drought Research Initiative (DRI), the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) and SGI Canada.

This research was performed entirely with Free Open Source Software (F.O.S.S.).

open source