Pattern Recogn. Phys., 1, 135–142, 2013
www.pattern-recogn-phys.net/1/135/2013/
doi:10.5194/prp-1-135-2013

Pattern Recognition in Physics
Open Access

Characterizing change in the variability of surface air temperature records: a comparative approach

C. Suteanu

Department of Geography and Department of Environmental Science, Saint Mary's University,
923 Robie St., Halifax, Nova Scotia, B3H 3C3, Canada

Correspondence to: C. Suteanu (cristian.suteanu@smu.ca)

Received: 31 July 2013 – Revised: 28 October 2013 – Accepted: 3 November 2013 – Published: 26 November 2013

Abstract. An accurate and comprehensive characterization of surface air temperature (SAT) variability is important for numerous purposes: studies on climate and climate change, the evaluation of climate model outputs, research on the impacts of changes in variability, etc. In this paper, SAT variability is considered from two different points of view: a measure of dispersion referring to the values in the time series, but ignoring their temporal sequence (the standard deviation); and a measure of persistence, for which the succession of the values in the time series is important (the exponent H established with Haar wavelet analysis). This paper uses daily minimum and maximum temperature records from Canadian stations in the Atlantic region and finds that: (i) SAT pattern variability can be assessed with the help of distinct methods applied together, in ways that might not be possible with any of the applied methods used separately; (ii) SAT pattern variability changes significantly over time; (iii) oscillations on scales from years to decades in both standard deviation S and the H exponent take place; (iv) the temporal change in SAT variability is reflected differently by the two applied methods: general statements concerning increases or decreases in variability should not be made without specifying the applied measure of variability.

1 Introduction

Characterizing surface air temperature (SAT) variability in relevant ways is particularly important for multiple purposes, from verifying properties of climate model outputs to assessing the different aspects and impacts of climate change (Rohr and Raffel, 2010; Wheeler et al., 2000). Assessing variability in a comprehensive way is, however, problematic. No single method seems to exhaustively address the richness of such natural patterns. The conclusions of studies dedicated to this subject are often inconsistent (Walsh et al., 2011). Finding a quantitative confirmation of the reported human perception of increasing variability in weather patterns (Nakashima et al., 2012) has been challenging. Multiple discrepancies between statistical variability estimators and human perception of change have been identified (Walsh et al., 2005, 2011). Variability in time series can take many different forms. In numerous studies, variability is evaluated by using variance or standard deviation (Boer, 2010; Esau et al., 2012). On the other hand, variability can also be assessed in other ways. While measures of dispersion do not refer to the succession of the data in the time series, the "smoothness" of time series depends on the sequence in which the various values are ordered. The resulting degree of "persistence" (Bunde and Havlin, 2002) in such patterns is also related to variability, but in a way that is different from those captured by statistical moments (Hsu et al., 2011; Rybski et al., 2008). Unlike the latter, persistence critically depends on the actual succession of the time series values on different timescales. Higher persistence involves a "smoother" signal, with less abrupt change on the scale range over which such persistence is established.

In other words, we can have, for instance, one and the same time series characterized by "strong" variability due to a high value of the measure of dispersion, and by "weak" variability due to signal smoothness, as reflected in its high degree of persistence due to the ordering of the values in the times series; of course, the reverse situation (lower dispersion of signal values but a less smooth signal due to decreased

Table 1. Meteorological stations used in this study (data from Vincent et al., 2002).

Station	N Latitude	W Longitude	Elevation [m]
Yarmouth, Nova Scotia	43.83	66.08	43
Sable Island, Nova Scotia	43.93	60.02	5
Saint John, New Brunswick	45.32	65.88	109
Fredericton, New Brunswick	45.87	66.53	20
Moncton, New Brunswick	46.12	64.68	71
Sydney, Nova Scotia	46.17	60.05	62
Charlottetown, Prince Edward Island	46.25	63.13	23
Miramichi, New Brunswick	47.02	65.47	33

persistence) is also possible. In this paper we define persistence by the presence of long-range correlations (details are provided in Sect. 3). Long-range correlations have been extensively investigated in relation to atmospheric processes (Efstathiou and Varotsos, 2010; Fraedrich et al., 2009; Kiralyi et al., 2006; Varotsos et al., 2009). Studies concerning long-range correlations applied to SAT data from Arctic stations have shown that pattern persistence varies on yearly to multi-decadal scales, and regions characterized by common temporal change in pattern persistence can be established (Suteanu and Mandea, 2012).

In this study we address the problem of SAT variability in a comparative framework, by applying two distinct approaches to SAT records and confronting their results concerning pattern change. Our main objectives are to assess the change in pattern variability occurring over time, and to evaluate the extent to which the different approaches agree with each other with respect to temporal change in pattern characteristics. For this purpose, the following questions will be addressed in the paper:

a. Does variability change over time to a significant extent?

b. Do the two applied methods identify similar changes in variability occurring over time?

c. Has a consistent increase in SAT variability taken place over the last century?

d. Are the answers to these questions station-specific, or independent of the analyzed station?

The paper is organized as follows: Sect. 2 presents the data sources and their location, Sect. 3 describes the methods applied in this study, results are presented in Sect. 4, and the last section is dedicated to conclusions.

2 Analyzed data

We analyze minimum and maximum daily SAT records from meteorological stations in the Maritime Provinces of Canada (Fig. 1). This region is interesting from the point of view of

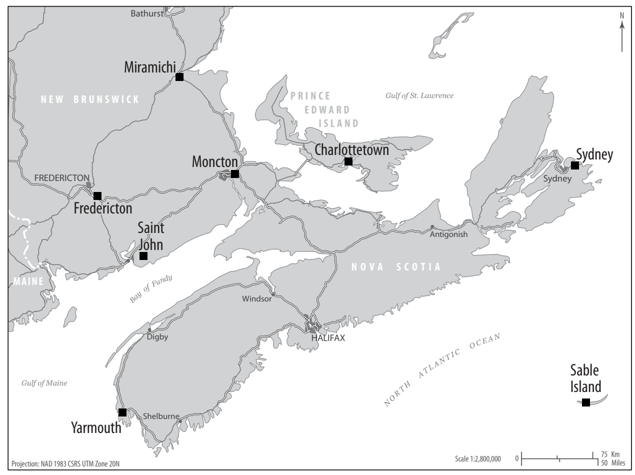

Figure 1. Stations included in this study (see Table 1 for more details).

SAT pattern variability, since it is subject to a wide spectrum of processes operating on different spatial and temporal scales such as the dominant westerly wind belt (which also carries slow- as well as fast-moving disturbances), blocking highs affecting westerly disturbances for days to weeks, continental cyclones arriving through central and eastern Canada from the south, Atlantic coast systems from the south and southeast, warm humid air penetrating through the Mississippi Valley and central Canada, and cold air flowing from the Arctic, as well as the Atlantic Ocean, with its moderating effect (which is reduced in the presence of sea ice) and its warm and cold currents exerting a temporally variable impact on SAT (a discussion on sources of SAT variability in this region can be found in Suteanu, 2011).

To address the objective of assessing pattern change over time, we choose long-term records (> 100 yr; Table 1) from the available homogenized station records (Vincent et al., 2002). Data homogeneization was performed by Vincent et al. (2002) for 210 stations; they found that 60 % of the initial data sets included up to four steps, caused mainly by changes in station location (e.g., stations moved to an airport site, which happened mainly in the 1940s), modifications in observation procedures, and the joining of different record sets (e.g., changes from human to automated observations occurred mainly in the last decade of the 20th century). They found that the applied homogeneization procedures had a significant impact on the mean values of the distributions, while the standard deviation suffered minimal changes; on the other hand, temperature trends decreased in many cases. Overall, the authors concluded that changes in daily variability can be more reliably assessed on homogenized data.

All the studied data sets start in 1895. The record length of the available data spans 113 yr for all stations except Charlottetown and Miramichi (which have lengths of 107 and 110 yr, respectively).

3 Methods

In the context presented in Sect. 1, to assess changes in variability occurring over time we choose two different paths for the evaluation of pattern variability.

One path simply addresses a measure of dispersion that does not take in consideration the succession of the data in the time series: standard deviation, a statistic often applied to this type of data (Boer, 2010; Esau et al., 2012; Przybylak and Vizi, 2005; Walsh et al., 2005). The other path depends on the actual sequence of the time series values, and assesses signal persistence: in some way, one may illustrate persistence by the degree to which there is a tendency for a signal to keep increasing its values when it has been rising (or decreasing its values when it has been falling), rather than changing its direction of change. Most importantly, in many natural time series, this tendency is expressed on a wide range of scales. One can thus identify and evaluate long-range correlations, i.e., the fact that the autocorrelation function decays slowly, not according to an exponential, but to a power law. It is thus important to underline that persistence is assessed on a wide interval of timescales – in this case, from weeks to years and decades.

The results obtained with the two different methodological approaches are then compared to each other, in order to check to what extent they lead to convergent conclusions with respect to changes in pattern variability.

Signal persistence is quantified here with the help of Haar wavelet analysis (Lovejoy and Schertzer, 2012). Wavelet analysis has been found particularly effective in geophysics, where its capability of addressing non-stationary signals and localized structures is important (Mallat and Hwang, 1992). For time series like the ones analyzed here, singularities may be significant, but they are invariant with respect to translation, and we are not interested in localized features: it is thus advantageous to consider averages over different time intervals in order to strengthen the statistical relevance (Lovejoy and Schertzer, 2012). Compared to other wavelets (Bacry et al., 1989; Holschneider, 1995), Haar wavelets are simple and easy to apply. They prove to represent a useful tool for

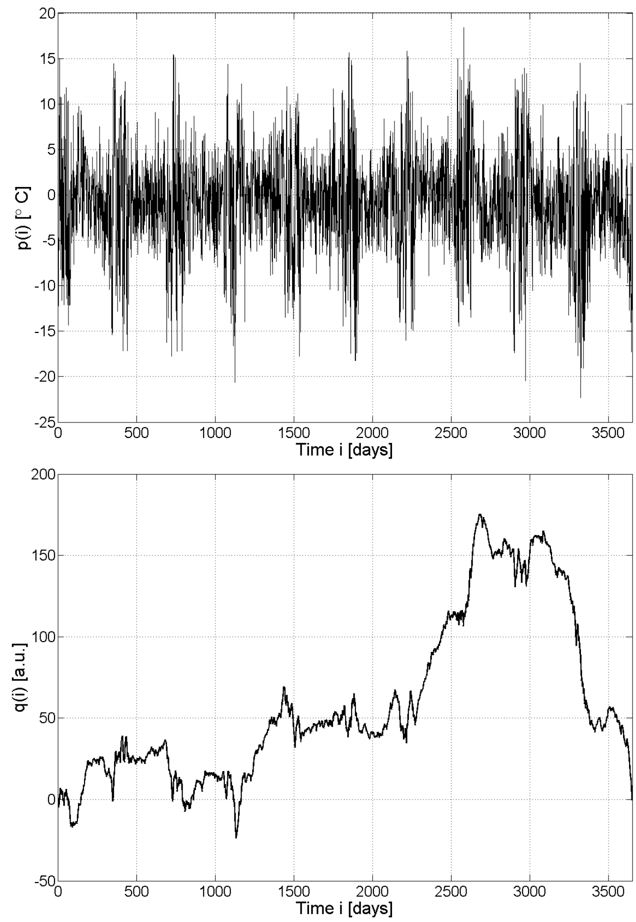

Figure 2. Time series $p(i)$ and $q(i)$ for ten years of daily minimum temperature records from Miramichi.

the analysis of atmospheric variables on timescales, which are labeled "macroweather" by Lovejoy and Schertzer (2013, p. 350–356).

The methodology is applied to daily minimum and maximum temperature time series and involves the following steps. The time series $t(i)$ is first subject to an operation by which the yearly variation is removed by subtracting the average temperature of every day of the year from each day's value:

$$p(i) = t(i) - < t_i > . \tag{1}$$

An example is shown in Fig. 2a. Unlike other methods dedicated to the assessment of signal persistence, such as detrended fluctuations analysis, this analysis method can be applied directly to the time series $p(i)$, and not only to its integral. However, analyzing the signal integral leads to narrower confidence intervals for the established exponent than addressing the time series $p(i)$. Therefore, we produce the normalized time series $w(i)$ by subtracting the average of $p(i)$ from the values of the $p(i)$ time series, and divide them by the standard deviation of the time series, and then we build the cumulative sum $q(i)$ (Fig. 2b):

$$q(i) = \sum_{j=1}^{i} w(j). \tag{2}$$

To apply this wavelet approach to the time series, one evaluates the way the size of the fluctuation F is scaling with the timescale or "lag" s; for this purpose, it is possible to choose from a variety of "mother" wavelets. Lovejoy and Schertzer (2012) show that the Haar wavelet, proposed as early as 1910 (Haar, 1910), is both easy to implement and to interpret, and leads to accurate results; it is defined by

$$\Psi(i) = \begin{cases} 1; & 0 \le i \le 1/2 \\ -1; & -1/2 \le i \le 0 \\ 0; & \text{otherwise.} \end{cases} \tag{3}$$

This leads to a simple analysis procedure, according to which one divides the time series in sections of size s, with s spanning a range of timescale values. This is followed by finding the fluctuation size F, based on the computation – for each lag s and for each section of size s – of the mean square difference between values in $(x+s/2)$ and $(x-s/2)$. The exponent H is then determined from the power law obtained for the relation between the lag s and the fluctuation size F:

$$F(s) \propto s^H. \tag{4}$$

For the temporal scales analyzed here, we have $-1 < H < 0$ for time series $p(i)$: they reflect the way in which deviations decrease when data are averaged over increasingly wide timescales (Fig. 2a). On the other hand, we have $0 < H < 1$ for time series $q(i)$: they show the fact that deviations get larger when wider scales are taken into consideration (Fig. 2b). While the theoretical relation $H[p(i)]+1 = H[q(i)]$ determined with Haar wavelets holds for the analyzed time series, we apply the analysis to the integral $q(i)$ rather than the signal $p(i)$, due to its robustness over a wide scale range and its narrow confidence intervals. An example for the $F(s)$ relationship is shown in Fig. 3; the linear relationship in log–log space is strong and extends over more than three orders of magnitude for both minimum and maximum temperature time series. As it is often the case, minimum temperature records are characterized by higher persistence than maximum temperature records (Kiraly et al., 2006; Suteanu, 2011; Suteanu and Mandea, 2012).

The methods described above are applied to successive windows of different lengths (from 3 to 7 yr). The window size does not have a major influence on the outcomes of the analysis, which is in agreement with other studies (Walsh et al., 2005; Suteanu and Mandea, 2012). The change in persistence over time is compared to the uncertainty affecting exponent values. This is accomplished by representing the 95 % confidence intervals in the determination of the exponent H (Chatterjee and Hadi, 1986). An example is presented

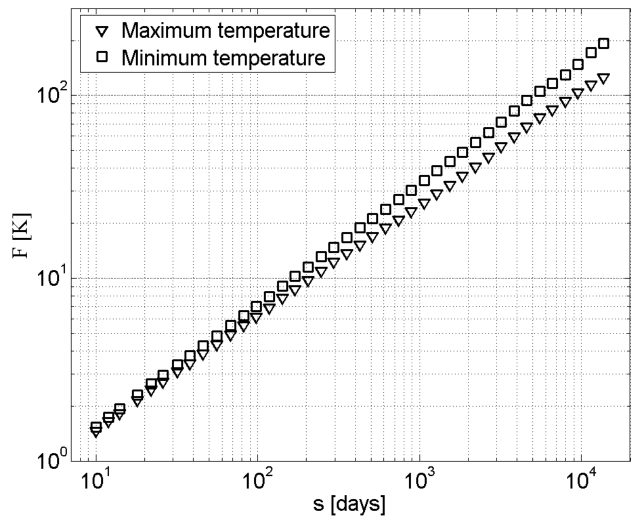

Figure 3. Example of Haar wavelet analysis applied to daily SAT records: Fredericton, Canada.

in Fig. 4: the size of the confidence intervals is low compared to the amplitude of the exponent variation.

In order to compare the outcomes of the two approaches regarding changes in pattern variability, we also consider the cross-correlation between the temporal variation in H and in S.

4 Results

Minimum and maximum daily temperature data from the stations specified in Sect. 1 are analyzed based on the two approaches described above. For all data sets, H and S vary over time, showing fluctuations on scales of years to decades (examples are presented in Fig. 5). Similarities between the variation patterns corresponding to the different stations are visible for both H and S.

A small size of confidence intervals compared to the amplitude of H (as shown in Fig. 4) is found in all cases. To improve graph readability, symbols for confidence intervals are thus not included in the other graphs shown in Fig. 5. In spite of the differences among the stations, strong maxima and well-pronounced minima in terms of persistence are found for the analyzed stations over the same or similar time intervals (Fig. 5a, c, e, g). S graphs also emphasize oscillations occurring on time periods that are comparable among the different stations (Fig. 5b, d, f, h).

A detailed comparison of the variations in H and S yields striking similarities over certain periods – for instance, maxima during the middle of the first decade of the 20th century, in the beginning of the 1920s, and the beginning to the mid-1930s. However, the noticeable decrease in S occurring in the mid-1950s has no correspondent in H graphs, which is also the case for other minima and maxima of S during later time periods.

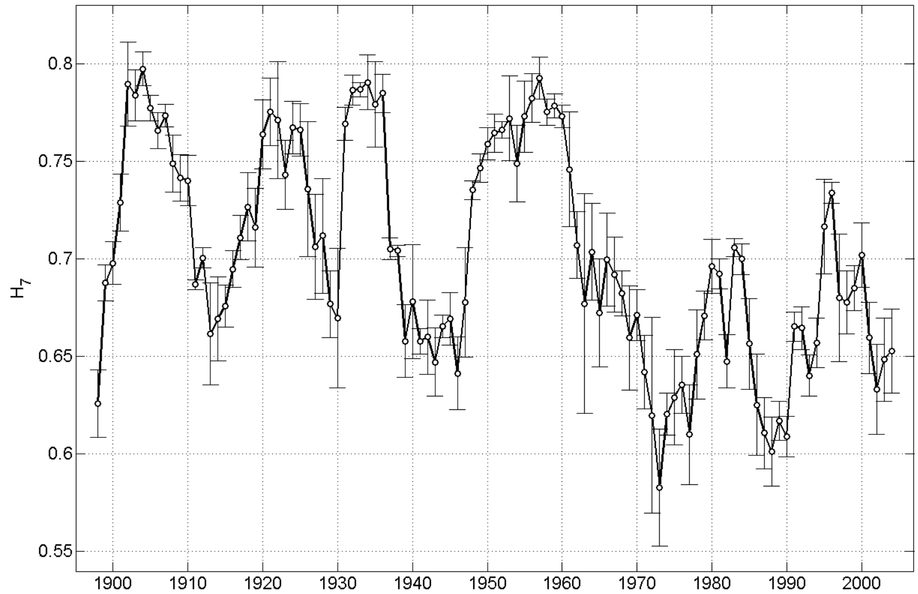

Figure 4. Results of Haar wavelet analysis applied to sliding windows of daily minimum SAT records from Sydney; window width: 7 yr; sliding step: 1 yr. Error bars indicate 95 % confidence intervals.

The H and S graph pairs for the analyzed stations indicate the possibility that they might be indeed correlated, but the correlation seems to change over time. Therefore, to better compare these temporal variations, we also consider the cross-correlation among their successive segments. The time length L chosen for the sliding windows used in cross-correlation does not play a determinant role with respect to the identified relation between H and S, for which the temporal resolution is one year (Fig. 5); however, longer intervals yield fewer points for correlation results, decreasing the temporal resolution of the study. We choose L to be an odd number, in order to assign the value of each window-by-window correlation to the middle sample of each window. Here we show results obtained for $L = 19$ yr. The correlation graphs $C_{H,S}$ obtained in this way are shown in Fig. 6.

Two groups of stations can be identified based on $C_{H,S}$ results. The first group includes five stations: Miramichi, Charlottetown, Sydney, Moncton, and Fredericton (Fig. 6a). Remarkably, while their H and S variations considered separately emphasize strong changes in time (including pronounced minima and maxima) on timescales of years to decades, the $C_{H,S}$ graph offers a different picture. One can distinguish two broad domains: one is characterized by strong positive correlation, with small fluctuations around 0.7 *to* 0.8, extending from the beginning of the records to the 1940s; the second regime, extending from that interval onwards, is dominated by negative correlation, around −0.7 *to* −0.8. A significant peak towards less negative correlation can be noticed around 1960.

The stations in the second group (Sable Island, Yarmouth, Saint John) are presented in Fig. 6b. They also have many elements in common. The most different among them is Sable Island, which shows high amplitude oscillations during the several first decades of the century, when the other stations have a more stable and strong positive correlation between H and S. If we compare the two groups of stations, we can notice in group 1 that Sydney, although it is similar to the other group 1 time series, shares the oscillatory behavior of group 2 stations before the 1940s, and thus seems to represent a transition between the two groups. The spatial relationships between the stations in the two groups can be seen on the map in Fig. 1 (group 2 stations occupy the southernmost positions).

The temporal change in the relationships among H and S highlights the extent to which they grasp different aspects of the temperature pattern variability. The two distinct regimes – characterized by positive and negative H vs. S correlation, respectively, cannot be identified from the separate analyses performed with the two methods. It is by looking through the "lens" of two distinct methodologies together that new aspects of pattern change can be uncovered.

5 Conclusions

Based on this comparative approach, the questions in Sect. 1 with respect to the temporal change in SAT variability can be answered as follows:

a. SAT pattern variability can be assessed with the help of distinct methods applied together, in ways that might not be possible with any of the applied methods used in separation.

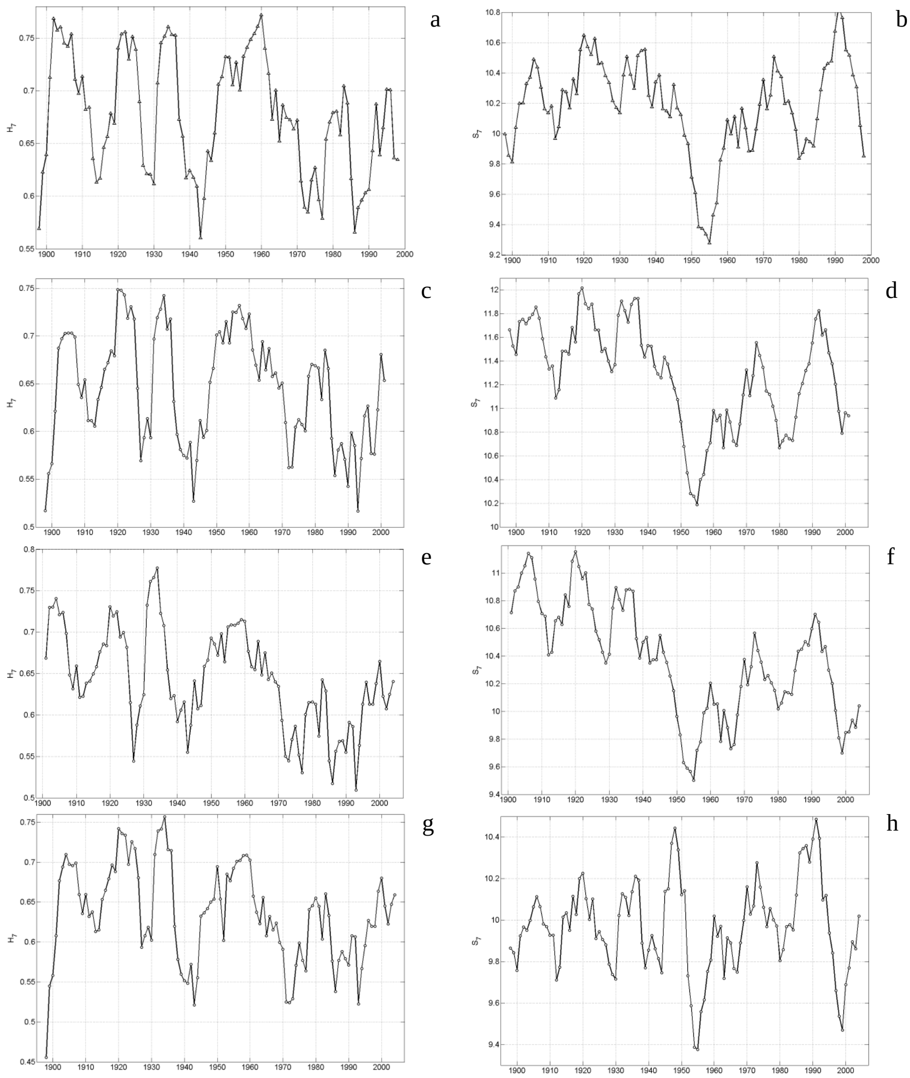

Figure 5. Examples of results obtained for H (left column) and S (right column): **(a, b)** Charlottetown; **(c, d)** Miramichi; **(e, f)** Moncton; **(g, h)** Saint John. Window width: 7 yr; sliding step: 1 yr.

b. SAT pattern variability is subject to change. Whether assessed with the help of statistical moments such as standard deviation, or by evaluating long-range correlations using Haar wavelet analysis, SAT pattern variability changes over time.

c. The temporal change in SAT variability is not the same for the two approaches. Strong positive correlation between the variation for S and H is followed, at some point in time, by strong negative correlation. General statements concerning increases or decreases in variability should therefore not be made, unless the applied measure of variability is specified.

d. Oscillations in both standard deviation S and the H exponent take place on scales from years to decades. This study does not find a consistent, simple trend of increasing SAT pattern variability in the analyzed stations. Human perception of growing variability may rely on other variables than temperature, or on the occurrence of storms and other extreme events (Walsh et al., 2005, 2011); even a combination of meteorological factors (such as precipitation, wind, cloud cover, etc.) and possibly other aspects of the environment, which may not be documented by historical weather records, may be the source of the perceived increasing variability.

e. All the answers provided to the questions above are valid for every station included in this study. While variability patterns are station-specific, groups of stations have important elements in common.

The lack of consensus concerning pattern variability and its temporal change can be partially explained by the fact that,

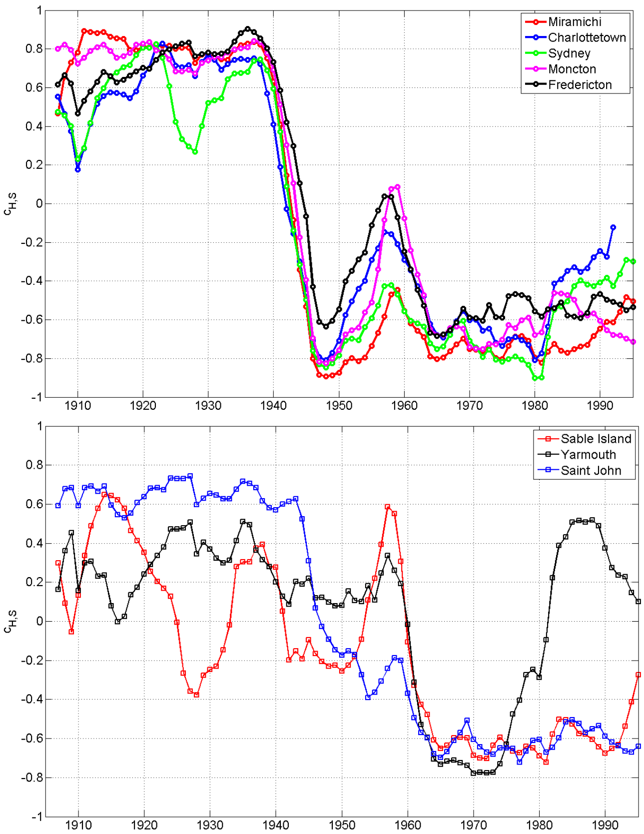

Figure 6. Correlation between S and H for the two groups of stations: **(a)** the short-term variations in S and H are replaced by two broad regimes, one with strong positive correlation and one with negative correlation, plus a remarkable increase in decade 6; and **(b)** while they have certain elements in common with each other and with the stations in **(a)**, these stations are characterized by more distinct patterns of temporal change.

like many natural patterns, SAT records cannot be fully characterized by one specific analysis method. Each method only captures some aspects of the patterns. Variability is expressed in SAT time series in many ways and on many timescales. Therefore, the often asked question about variability increasing over the last several decades cannot receive a simple answer; given the nature of variability, this situation is not expected to change, even if new, more powerful analysis tools will reveal novel and interesting aspects of pattern change.

This investigation showed that considering together the two approaches presented here can lead to the identification of regimes of pattern variability, which are not detectable with the individual methods themselves. Properly combining selected analysis methods may lead to more helpful methodologies for pattern characterization.

Acknowledgements. The author would like to thank Lucie Vincent (Climate Research Division, Environment Canada) for making available the homogenized temperature data sets analyzed in this paper, and Will Flanagan (Geography Department, Saint Mary's University) for providing Fig. 1. This work was supported by the research grant "Characterizing change in atmospheric temperature variability patterns", Saint Mary's University, Halifax, Canada, and the research grant 21/5.10.2011, Program TE, Romania.

Edited by: N.-A. Mörner
Reviewed by: H. Jelbring and one anonymous referee

References

Bacry, A., Arneodo, A., Frisch, U., Gagne, Y., and Hopfinger, E.: Wavelet analysis of fully developed turbulence data and measurement of scaling exponents, edited by: Lessieur, M. and Metais, O., Turbulence and coherent structures, Kluwer, 703–718, 1989.

Boer, G. J.: Changes in interannual variability and decadal potential predictability under global warming, J. Climate, 22, 3098–3109, doi:10.1175/2008JCLI2835.1, 2010.

Bunde, A. and Havlin, S.: Power-law persistence in the atmosphere and in the oceans, Phys. A, 314, 15–24, 2002.

Chatterjee, S. and Hadi, A. S.: Influential observations, high leverage points, and outliers in linear regression, Stat. Sci., 1, 379–416, 1986.

Efstathiou, M. N. and Varotsos, C. A.: On the altitude dependence of the temperature scaling behaviour at the global troposphere, Int. J. Remote Sens., 31, 343–349, 2010.

Esau, I., Davy, R., and Outten, S.: Complementary explanation of temperature response in the lower atmosphere, Environ. Res. Lett., 7, 044026, doi:10.1088/1748-9326/7/4/044026, 2012.

Fraedrich, K., Blender, R., and Zhu, X.: Continuum climate variability: long-term memory, extremes, and predictability, Int. J. Mod. Phys. B, 23, 5403–5416, 2009.

Haar, A.: Zur Theorie der orthogonalen Funktionsysteme, Math. Ann., 69, 331–371, 1910.

Holschneider, M.: Wavelets : An Analysis Tool, Clarendon Press, 1995.

Hsu, H.-m., Lin, C.-Y., Guenther, A., Tribbia, J. J., and Liu, S. C.: Air-chemistry "turbulence": power-law scaling and statistical regularity, Atmos. Chem. Phys., 11, 8395–8413, doi:10.5194/acp-11-8395-2011, 2011.

Kiraly, A., Bartos, I., and Janosi, I. M.: Correlation properties of daily temperature anomalies over land, Tellus A, 58, 593–600, 2006.

Lovejoy, S. and Schertzer, D.: Haar wavelets, fluctuations and structure functions: convenient choices for geophysics, Nonlin. Processes Geophys., 19, 513–527, doi:10.5194/npg-19-513-2012, 2012.

Lovejoy, S. and Schertzer, D.: The Weather and Climate, Emergent Laws and Multifractal Cascades, Cambridge University Press, 2013.

Mallat, S. and Hwang, W.: Singularity detection and processing with wavelets, IEEE Trans. Info. Theory, 38, 617–643, 1992.

Nakashima, D. J., Galloway McLean, K., Thulstrup, H. D., Ramos Castillo, A., and Rubis, J. T.: Weathering Uncertainty: Traditional Knowledge for Climate Change Assessment and Adaptation. Paris, UNESCO, and Darwin, UNU, 120 pp., 2012.

Przybylak, R. and Vizi, Z.: Air temperature changes in the Canadian Arctic from the early instrumental period to modern times, Int. J. Climatol., 25, 1507–1522, 2005.

Rohr, J. R. and Raffel, T. R.: Linking global climate and temperature variability to widespread amphibian declines putatively caused by disease, P. Natl. Acad. Sci. USA, 107, 8269–8274, 2010.

Rybski, D., Bunde, A., and Von Storch, H.: Long-term memory in 1000-year simulated temperature records, J. Geophys. Res., 113, D02106, doi:10.1029/2007JD008568, 2008.

Suteanu, C.: Detrended fluctuation analysis of daily atmospheric surface temperature records in Atlantic Canada, Can. Geogr., 55, 180–191, doi:10.1111/j.1541-0064.2010.00323.x, 2011.

Suteanu, C. and Mandea, M.: Surface air temperature in the Canadian Arctic: scaling and pattern change, Meteorol. Atmos. Phys., 118, 179–188, doi:10.1007/s00703-012-0206-8, 2012.

Varotsos, C., Efstathiou, M., and Tzanis, C.: Scaling behaviour of the global tropopause, Atmos. Chem. Phys., 9, 677–683, doi:10.5194/acp-9-677-2009, 2009.

Vincent, L. A., Zhang, X., Bonsal, B. R., and Hogg, W. D.: Homogenization of daily temperatures over Canada, J. Climate, 15, 1322–1334, 2002.

Walsh, J. E., Shapiro, I., and Shy, T. L.: On the variability and predictability of daily temperatures in the Arctic, Atmos. Ocean, 43, 213–230, doi:10.3137/ao.430302, 2005.

Walsh, J. E., Overland, J. E., Groisman, P. Y., and Rudolf, B.: Ongoing Climate Change in the Arctic, Ambio, 40, 6–16, 2011.

Wheeler, T. R., Craufurd, P. Q., Ellis, R. H., Porter, J. R., and Vara Prasad, P. V.: Temperature variability and the yield of annual crops, Agriculture, Ecosystems and Environment, 82, 159–167, 2000.