Review of "Jens Esmark's Christiania (Oslo) meteorological observations 1816-1838: The first long term continuous temperature record from the Norwegian capital homogenized and analysed"
Geir Hestmark and Øyvind Nordli
Clim. Past Discuss., doi:10.5194/cp-2016-60, 2016

General remarks:

This paper presents a very interesting and valuable data set which extends our knowledge of early 19th century climate in north western Europe. It is an important paper which I recommend to be published following some additional calculations and revisions. While the authors have made extensive efforts to account for data quality and to homogenize the readings for long term climatic analysis in the face of sparse metadata, I am particularly uneasy about the lack of information concerning the observation times, and it is my opinion that further analyses may help reduce this uncertainty.

In particular, the authors could make use of frequency analysis as exemplified by the work of Bergström and Moberg (2002) and Slonosky (2014) to compare Esmark's daily morning, afternoon and evening observations to the nearly 25 years of modern hourly observations mentioned in Table 1 and possibly obtain an approximate idea of the times of observation. It may be necessary to sub-divide the historical record for suspected changes in observation time derived from the SNHT analysis and to consider the possibility of observation times, especially in the morning, changing with the season, if this is supported by other metadata (e.g the statement of observation times quoted on lines 188-189). If probable times of observation can be established, the entire analysis will stand on much firmer ground. As it stands, there are many adjustments made on a statistical basis which add to the uncertainty of the final values of the observations, particularly given the differences seen when compared with other nearby observations.

The accounting of the adjustments due to inhomogeneities detected by the SNHT and other intra-series comparisons is extremely thorough and to be commended, but as is presented leaves the reader confused. A plethora of monthly adjustments is proposed in Tables 2-5, but it is not clear which adjustments were finally applied to which observations, the sequence of the adjustments nor whether the adjustments were applied to the daily data or to the monthly means. If daily, there will be artificial jumps between the end of one month and the beginning of the next - see Vincent et al (2002). In general, more use might be made of the advantages gained by having daily, rather than monthly, observations to analyze; much work had been done in the field of historical climatology in the past decade or two on analyzing daily observations directly.

The fairly large differences shown between these data and other nearby stations, less than 1 km away, also give reasons for concern about the final quality of the data. Comparisons with other series, such as Uppsala -Bergström and Moberg (2002) - and Stockholm –Moberg and Berström (1997), although a considerable distance away, may still give valuable indicators as to the character of each month and help decide which series in the comparisons are the more reliable.

Finally, all the data, including the raw data, should be placed in an online archive.

Specific remarks:
Introduction, line 45 and thereafter: “protocol” usually refers to a method; it would be less confusing if the authors could use a word like “logbook” or “weather registers” if they mean the actual physical records of Esmark’s weather observations.

Introduction and Section 2: An interesting and important synopsis of the observer, his location and environment, and his instruments.

Line 350: The authors should take note of Gauvin’s 2012 article on the Réaumur thermometer:

The authors should be aware from Gauvin’s work that theoretical adjustment of 1.25 for Réaumur to Celsius may not be accurate. This could help explain some of the large differences seen when comparing Esmark’s values to the nearby observations in section 5.

Section 4.1: The SHNT results seem somewhat ambiguous. Can the SNHT be run on all the 7665 days of observations, rather than dividing up into months and seasons? This might give a clearer indication of the actual break date. If this is too large a number for computational purposes, the series could be tested on running sub-portions (i.e first six months, move forward three months and test next six month period, and so on). Testing on other variables such as pressure might also give a potential indication of a change in the positioning of the instruments. It may help to further divide section 4.1 into subsections dealing with all of the adjustments to each of the three observation times separately and consecutively.

Line 380: A synopsis of the shifts and dates for each of the observation series would make these clearer. What were the final adjustments made to each series? A table summarizing the actual adjustments applied and the order in which they were applied would be helpful.

Line 381: The authors appear to be postulating a replacement of an hourly observation in the morning with a minimum thermometer. Hourly temperature observations and minimum temperature observations are not the same entity. If the authors think that a minimum thermometer was in use, a new series labelled “minimum temperature” should be analyzed. Rather than an inhomogeneity, this is a new variable.

Line 390-2/871: This reasoning needs to be better explained, especially given the actual observation times are unknown. What does the description in the title of Table 4 “minimum temperature at 0800 UTC” mean?

Line 405: More specifics are needed to explain this conclusion: 26% of interpreted “minimum” values being higher than the evening temperature is a high proportion. This unusual temperature trend is a situation which could occur with the passage of frontal systems overnight. What is the proportion of such unusual diurnal temperature trends in the modern record?

Table 4/ Line 872: What is the authors' interpretation of the negative summer differences for 1816-1828 and 1822-1828, compared to the modern differences? How are these differences changed with the selection of different observation times in the modern period (e.g. 0700, 0600, and sunrise?). Why is the period 1816-1821 corrected but not 1822-1828? Are these results from before or after the application of the adjustment of the 1821 inhomogeneity?

Line 442: How does this weakened diurnal temperature wave affect the reasoning section 4.1 concerning the minimum thermometer?

Page 14, lines 447-451, Figure 5: Adjusting from one postulated unknown time to a second unknown time is a procedure beset with uncertainty, particularly as the linear trend does not appear to apply as well in the middle of the period, 1833-1836, when the points would give a much less steep slope. Have the authors explored regressions and residuals for other, finer time resolutions than the three-month period shown in Fig 5? What is the value of the sum of squares error? If better estimates of actual times of observation can't be made, some portion of the data may just have to be classified as unusable.

Section 4.5 Again, how and why are these adjustment values derived? This is not clear.

Line 457, Figure 8: These adjustments should be presented in a Table separate from the Figure.

Line 482: Section 4.6 should be in the discussion section, while section 5.1 would perhaps be better placed as a summary in section 4. The comparisons with other observers and discussion of the thermometer error would be better placed in a data quality and comparison section, with the climatic discussion in a separate section.

Line 541: Again, if we don't know the observation times, it's impossible to attribute the difference between the observers to a specific cause such as instrument location.

Line 550: is 2100 UTC after sunset in summer?

Figure 12: This would seem to suggest that the unadjusted values for Esmark are closer to the Observatory than the adjusted values.

Check grammar: line 273; 416; 419; 432… The grammar is in numerous locations (insufficient time to enumerate here) somewhat awkward, for example putting a place indicator before a time indicator, when often in English the time clause precedes the location clause.

References:

Bergström, Hans, and Anders Moberg. "Daily air temperature and pressure series for Uppsala (1722–1998)." Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources. Springer Netherlands, 2002. 213-252

Gauvin, J. F. (2012). The instrument that never was: inventing, manufacturing, and branding Réaumur's thermometer during the enlightenment. Annals of Science, 69(4), 515-549.

Moberg, A., Bergström, H., Krigsman, J. R., & Svanered, O. (2002). Daily air temperature and pressure series for Stockholm (1756–1998). In Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources (pp. 171-212). Springer Netherlands.

Slonosky, V. (2014). Historical climate observations in Canada: 18th and 19th century daily temperature from the St. Lawrence Valley, Quebec. Geoscience Data Journal, 1(2), 103-120.

Vincent, L. A., Zhang, X., Bonsal, B. R., & Hogg, W. D. (2002). Homogenization of daily temperatures over Canada. Journal of Climate, 15(11), 1322-1334.)