

# Daily temperature and pressure series for Salem, Massachusetts (1786–1829)

G. van der Schrier · P. D. Jones

Received: 18 November 2005 / Accepted: 10 May 2007 / Published online: 12 September 2007  
© Springer Science + Business Media B.V. 2007

**Abstract** The digitization and homogenization of a record with four daily thermometer and two daily barometer readings is described for the meteorological journal of Dr. E. A. Holyoke of Salem (Massachusetts). These records begin in January 1786 and span the period to March 1829 for temperature, and the period to December 1820 for pressure. The records are reconstructed and some inhomogeneities are identified and corrected for. The temperature data compares favourably with monthly temperature data from New Haven (Connecticut) and a comparison with independently reconstructed daily pressure data for nearby Cambridge (1780–1789) show that the temporal variations of the data agree very well. It is shown that the number of extremely cold days was considerably greater during 1786–1829 than recent times, while the number of warm days in the early and modern records are comparable. A probability distribution of daily winter pressure values shows a mean of the distribution which is 3 to 4 hPa lower during 1786–1820.

## 1 Introduction

Daily meteorological observations were made by Dr. Edward Augustus Holyoke of Salem, MA. (42.32°N, 70.53°W) during the period 1786 to 1829. These measurements include two daily barometer readings and four daily temperature readings. Dr. Holyoke's meteorological journal was published at the time in 1833 (Hale 1833). The purpose of this study is to report on the digitization of the data, describe

---

G. van der Schrier (✉) · P. D. Jones  
Climatic Research Unit, School of Environmental Sciences,  
UEA, Norwich NR4 7TJ, UK  
e-mail: schrier@knmi.nl

*Present Address:*  
G. van der Schrier  
KNMI, P.O. Box 210, 3730 AE De Bilt, The Netherlands

the procedure to homogenize the pressure and temperature record, and to give a preliminary analysis.

Dr. Holyoke, born August 13th 1728, started his meteorological journal when he was over 50 years old and continued it until a few weeks before his death at age 100 at March 31st 1829. His first observations are from January 1786, the last one is from March 1st 1829. Dr. Holyoke was one of the founding members of the American Academy of Arts and Sciences in 1780 and was its president from 1814 until 1820. More information on the personal background of Dr. Holyoke can be found at: [www.salemweb.com/tales/holyoke.shtml](http://www.salemweb.com/tales/holyoke.shtml).

A detailed description of Dr. Holyoke's measurements and the dataset is given in Section 2. Quality control and homogenization of the records is described in Section 3 and a preliminary analysis of the daily temperature and pressure records is given in Section 4. The study is concluded in Section 5.

## 2 Description of two meteorological journals

### 2.1 Dr. Holyoke's meteorological journal

Dr. Holyoke's journal begins with the year 1786 and continues to the end of 1821, after which Dr. Holyoke continued his measurements in a less regular and extensive manner until 1829. The hand-written volumes of the journal are located at the Houghton Library of Harvard University. A selection of this data has been published in the *Memoirs of the American Academy of Arts and Sciences* by E. Hale in 1833. Since the authors of the present study did not have access to the original hand-written manuscript, we focused on the publication of Hale (1833). Some details regarding the original manuscript are given in the introduction to this study, many of which are reproduced here.

During the first 36 years of observations, Dr. Holyoke made two daily barometer readings, eight temperature readings (four indoors, four outdoors), four observations of the course of the winds, four of the state of the weather and for a large part of this period, four observations with a hygrometer. Occasionally, Dr. Holyoke made additional daily observations and continued this for a few months. Interspersed in his journal are observations of atmospheric phenomena and the state of the season as indicated by the progress and state of the vegetation.

Of this abundance of data, only the two daily barometer readings, the four temperature readings (outdoors) and the four observations of the state of the weather are included in the Hale (1833) publication. From 1821 onwards, the two barometer readings were discontinued and replaced with an additional temperature observation at night, probably from a minimum-temperature recording thermometer. The minimum-temperature readings were discontinued after January 1824.

Originally, Dr. Holyoke distinguished the several states of the weather with great detail: serene, fair, pleasant; slight cloudiness, cloudy, flying clouds; drizzly, shower, rain, much rain; sprinkling of snow, snow, much snow, and many other terms. Hale (1833) reduced these categories to fair, cloudy, rain, hail and snow.

It must have been impossible for Dr. Holyoke to make the observations every day at precisely the same times. In the hand-written manuscript, the observation time for the first measurement was generally noted down as 'eight o'clock', but sometimes

‘from eight to nine’. The second observation (for temperature and the state of the weather) is generally made around noon, but sometimes ‘at 1 P.M.’, ‘at 2 P.M.’ or ‘between 1 and 2 P.M.’. Hale concludes that the objective seems to have been to make an observation at the hottest part of day. The third measurement (for temperature and the state of the weather) is at sunset, and the fourth is uniformly at ten in the evening.

The make of the thermometer and barometer Dr. Holyoke used are unknown. There are indications that initially a Hoxbee alcohol thermometer was used which was replaced with a mercury thermometer of unknown make (McAllistar, personal communication). The make of the barometer is unknown.

The town of Salem had around 9,000 residents around AD 1800, with the downtown area, where Dr. Holyoke lived throughout this period, densely packed with houses. Dr. Holyoke’s original house no longer exists.

The record, as it is published by Hale, is remarkably complete; only 3.4% of the barometer readings are missing, 1.5% of the temperature readings and 0.1% of the descriptions of the weather are absent. There are days in the record for which the barometer readings at 8 A.M. and 10 P.M. are missing, but all four temperature readings and weather characterizations were present. This suggests that others, apart from Dr. Holyoke, made observations.

Monthly values of the Holyoke data have previously been used by Lamb and Johnson (1959) and van der Schrier and Barkmeijer (2005) for their pressure reconstructions. Ludlum (1966) also used the data for his accounts of the Early American Winters.

Lamb and Johnson (1959) qualify this record with the remark that it ‘appears more homogeneous and trustworthy than that of most official observatories of that time.’

## 2.2 Prof. Wigglesworth’s meteorological journal

The meteorological journal by Edward Wigglesworth from Cambridge (MA) is digitized to compare the Holyoke daily pressure and temperature data with an independent record. The Wigglesworth data will also be used to establish the relation between temperatures outdoors and indoors, which will be needed in Section 3.3 when the reduction of the Holyoke record to 0°C is discussed.

Edward Wigglesworth (1732–1794) was Hollis professor of Divinity from 1765 to 1791 at Harvard University and acting president of the Faculty of Philosophy, Natural Philosophy and Theology during 1780–1781. Interestingly, in Wigglesworth’s Harvard years, he and John Wintrop, who was Hollis professor of Mathematics and experimental Philosophy at the time, were colleagues. In the preface to his meteorological journal, Wigglesworth writes that his thermometer ‘is that used by the late (...) Professor Wintrop, in making his Observations’. John Wintrop made a large contribution to the documentation of weather in the Boston area (Ludlum 1966).

Wigglesworth and Holyoke were contemporaries, lived ca. 25 km from each other and had a similar social status, which validates the assumption that their housing conditions were comparable. The Wigglesworth record spans the period July 1780–December 1789 and the year 1793. Wigglesworth made three daily barometer and three daily thermometer readings (outdoors), observed the direction of the winds and characterized the weather in similar terms as Holyoke. For the period January

1786–December 1789, this was supplemented with three daily indoor temperature measurements. The measurements were made at 8 A.M., 1 P.M. and 9 P.M. Unfortunately, no barometer readings were made during the year 1793.

The Wigglesworth data is digitized from photographs of the original hand-written manuscripts. The photographs are in the collection of files from Prof. Lamb at the Climatic Research Unit and were presumably ordered for his SLP-reconstruction (Lamb and Johnson 1959).

### 3 Quality control

#### 3.1 Position of the instruments

Hale (1833) does not give any information on the position of the thermometer and barometer Holyoke used. However, Brown (1940) reports that Holyoke writes in his original hand-written diary that the temperature measurements were read from ‘accurate instruments made in London and suspended in the open air in the shade on the north side of my house in the main street of Salem about 7 ft above the ground.’

In the preface to his journal, Wigglesworth writes that “The Barometer and Thermometer are hung in a staircase, by the side of a North Window.” We have not been able to deduce information on the position of the outside thermometer, apart from the statement that “the Thermometer was actually exposed to the external Air”. Wigglesworth adds that the barometer and thermometer used in making the observations is made by James Champney of London. The barometer has a nonius, or vernier, scale is fixed to it.

#### 3.2 Detection of random data errors

Random data errors are likely to have been introduced in the publication of the Holyoke data by Hale (1833) and in the digitization. To minimize these types of errors, the difference in the pressure readings at 8 A.M. and 10 P.M. is calculated for each day, and the readings are treated as suspect if the pressure difference is larger than  $\approx 1$  in. Temperature readings were treated as suspect if  $T_{8\text{am}} - T_{\text{noon}} \geq 15^\circ\text{F}$  (ca.  $8.3^\circ\text{C}$ ) or if  $T_{10\text{pm}} - T_{\text{noon}} \geq 15^\circ\text{F}$ . The source was inspected for all these flagged values. Only a small percentage of flagged temperature data turned out to be flawed.

When it was impossible to deduce a reasonable value for evident outliers, a missing value was substituted. Table 1 lists the flagged years with replaced data.

#### 3.3 Reduction to $0^\circ\text{C}$

Reduction to  $0^\circ\text{C}$  and the conversion to pressure in hPa was made using the standard equation (Middleton 1947; Moberg et al. 2002), which adjusts for the temperature of the room in which the barometer was kept. These temperatures are not available, so they had to be estimated from the outdoor temperatures. To establish a relation between temperatures indoors and outdoors, the Wigglesworth meteorological journal was used. The assumption is made that the relation between indoor and outdoor temperatures is robust and can be applied to the Holyoke record as well.

**Table 1** Observation in the Holyoke record which most probably are typographical mistakes in the Hale-manuscript. Replaced values are in the right-most column

Observation	Year	Month	Day	Obs. time	Replaced with
76°F	1789	July	31	10 P.M.	Missing
29.01 in.	1793	September	7	10 P.M.	30.01 in.
30.83 in.	1795	October	20	8 A.M.	29.83 in.
39.77 in.	1797	October	30	8 A.M.	29.77 in.
72°F	1797	November	21	sunset	27°F
3029. in.	1798	May	27	10 P.M.	30.29 in.
59.84 in.	1798	August	30	10 P.M.	29.84 in.
20.00 in.	1798	December	22	10 P.M.	30.00 in.
59.90 in.	1800	August	29	10 P.M.	29.90 in.
57°F	1803	July	31	noon	75°F
≈ 1 in. increase between 8 A.M. and 10 P.M.	1805	May	17–23		Missing
30.52 in.	1810	May	4	10 P.M.	29.52 in.
29.08 in.	1810	May	10	8 A.M.	30.08 in.
29.03 in.	1812	May	20	8 A.M.	30.03 in.
30.83 in.	1812	July	26	8 A.M.	29.83 in.
≈ 1 in. increase between 8 A.M. and 10 P.M.	1815	September	4–5		Missing
72°F	1818	February	27	10 P.M.	27°F
98°F	1818	June	15	8 A.M.	Missing
86°F	1818	November	27	10 P.M.	Missing
P	1820	July	25	noon	F
27.73 in.	1820	September	24	10 P.M.	29.73 in.
84°F	1822	October	6	10 P.M.	48°F
33°F	1823	August	23	noon	Missing
93°F	1823	November	11	sunset	Missing
52°F	1824	June	26	noon	Missing
53°F	1827	August	31	noon	Missing

Following Moberg et al. (2002), we estimated barometer temperatures at times  $i$ ,  $T_i$  from outdoors temperatures  $T_j^{\text{out}}$  by applying first a low-pass filter and then a linear model:

$$T_i = a_0 + a_1 \sum_{j=1}^N w_j T_{i+1-j}^{\text{out}}. \quad (1)$$

The weights  $w_j$  are determined from a one-tailed Gaussian weighted running mean with an  $N$ -length window. The model parameters were determined from the calibration years 1786 and 1787, and validated in the years 1788 and 1789. The best performing model, in terms of closeness of fit, was determined by calculating indoors temperatures by Eq. 1 for various choices of  $N$ . Optimal values of the coefficients  $a_0$ ,  $a_1$  were determined with a least-squares calculation.

The final choice of  $N$  was four for both the evening and the morning temperatures, indicating that the evening indoors temperature reading is optimally explained by a combination of the simultaneous outdoors temperature and the preceding

temperatures at 1 p.m., in the morning and the evening temperatures of the previous day (analogues for the morning temperatures). This finding indicates that the thermal inertia of the building was rather low, which reflects the abundant use of wood in house building. The regression parameters are  $a_0 = 5.41448$  and  $a_1 = 1.34950$  for the morning values, and  $a_0 = 6.37673$  and  $a_1 = 1.34227$  for the evening values. The temperature readings are still in their Fahrenheit guise for these calculations.

The validation period gave mean and median values of the difference between observed and estimated morning temperatures of 0.08 and 0.14°C, while the 5th and 95th percentiles are at −1.83 and 2.22°C. For the evening indoor temperatures, the mean and median are −0.32 and −0.26°C respectively, while the 5th and 95th percentiles are −2.27 and 1.83°C. The adjustment to the pressure readings associated with these extreme percentiles is ca.  $\pm 0.4$  hPa.

The reduction to 0°C and the conversion to pressure in hPa for the Wigglesworth record is done in a similar fashion, with the exception that for the period January 1786–December 1789 the actual indoor temperature measurements are used rather than an estimate from outdoor temperatures.

### 3.4 Reduction to normal gravity and sea level

The reduction to normal gravity and sea-level for the pressure records from Salem and Cambridge were made using standard theory (Middleton 1947; Moberg et al. 2002). Since both towns are located close to 45°N and near sea-level, these adjustments amounted to ca. 0.2 and ca. 0.6 hPa respectively.

### 3.5 Homogenization of the records

#### 3.5.1 Temperature

The Holyoke record is homogenized using a combination of the two-phase regression method introduced by Vincent (1998) and visual inspection. Homogenization of the temperature measurements is done using the monthly mean daily temperatures from New Haven (Connecticut, 40.8°N, 72.9°N) as a reference series. The scarcity of instrumental data for this period and for this area limits the choice for neighbouring series to the New Haven record only. New Haven is ca. 220 km away from Salem. This records spans 1781–1970 and it has been subject to some quality control (Jones et al. 1985). The four temperature measurements of the Salem record, made at 8 a.m., noon, sunset and 10 p.m., are tested for homogeneity separately.

The procedure for homogeneity testing and data adjustment consists of the application of two linear regression models to the series. The first model determines whether the time series is homogeneous for the tested interval of time. If the timeseries fail to pass this test, a second model is applied which locates the position and size of a potential step in the time series. The latter quantities are determined by optimally fitting, in a least-squares sense, a straight line through the data points on either side of a potential step. If these fits do not match well enough, following a criterion explained below, then the position and size of the step is identified. Every data point in the tested time interval, with the exception of three values at either extremity of the interval, is checked for a step.

Monthly means of the daily New Haven temperatures are subtracted from monthly means of the 8 A.M., noon, sunset and 10 P.M. temperature measurements of the Salem record. This yields difference records with a non-zero mean and a seasonal cycle. The seasonal cycle in this difference record is related to the fact that *daily* temperature values from New Haven are subtracted from temperature values from *specific times of the day* from Salem. The observations at 8 A.M. and 10 P.M. are generally far from the daily minimum in the warm season (Moberg et al. 2002) but they are much closer to the daily minimum temperature in the winter season. The seasonal cycle of noon or sunset temperatures is not expected to parallel the seasonal cycle of daily averaged temperature either.

Both the offset and the seasonal cycle in the difference records are removed prior to the homogenization. The offset and a mean seasonal cycle is determined by calculating average January, February etc. conditions over the nearly 43 years of data available.

The first regression model, which tests for homogeneity, was applied to 60 month chunks of data. If the autocorrelation exceeds the  $2n^{-1/2}$  level, with  $n$  the window length, or if the Durbin–Watson statistic is below 1.6, then the other regression model, description of a step, was applied to this chunk of data. The threshold for the autocorrelation relates to the 95% confidence level, with values outside this level are significantly different from zero at the 5% level. The Durbin–Watson statistic is a test for the correlation of regression residuals (Wilks 1995, §6.2.6). It tests the null-hypothesis that the residuals are serially independent against the alternative that they are consistent with a first-order autoregressive process. The threshold for the Durbin–Watson statistic relates to the 5% level where the null-hypothesis of zero serial correlation can either be rejected or where this statistic is indeterminate.

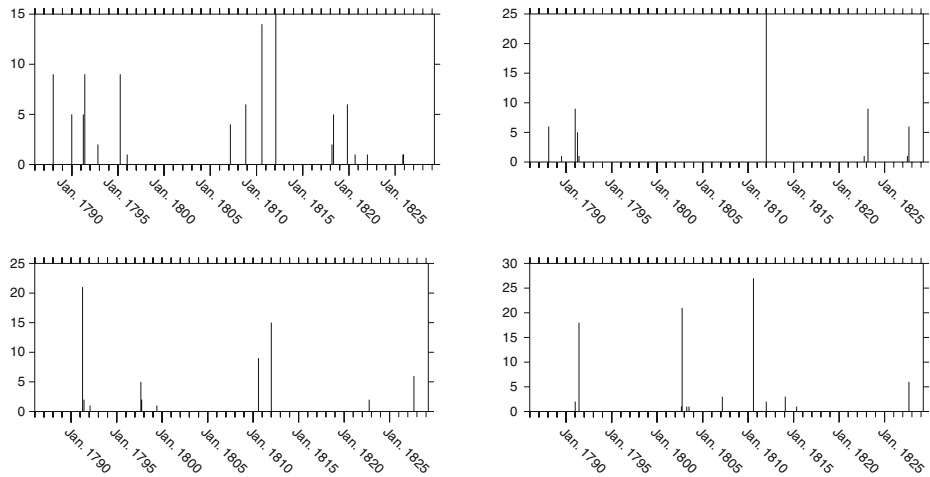
A potential step is considered ‘significant’ if the associated  $F_{max}$  statistic exceeds 6.9. This latter value is from the literature (Lund and Reeves 2002, their table 1) and corresponds to data of 50–75 samples. This level corresponds to the 95% confidence level that the null-hypothesis of no change can be rejected.

The 60-month window was then moved with steps of 2 months, and the procedure is repeated on the new chunk of data until the end of the record is reached. If a potential step really is robust, then this step should be detected in subsequent applications of the regression models until the potential step has moved out of this window.

The detected changepoints in the difference timeseries are shown in Fig. 1, the most robustly detected changepoints are tabulated in Table 2. The large number of detected changepoints for the 8 A.M. measurement is perhaps due to the fact that the morning measurement was occasionally made ‘from eight to nine’ (Section 2.1). This must have had a large impact on the recorded temperature since the 8 A.M. measurement is at a steep part of the diurnal temperature cycle.

A visual inspection of the difference records (Fig. 2) indicates that the basis for some detected discontinuities is lacking. The jumps in the records which can be detected visually are included in Fig. 2 as dashed vertical lines.

Concluding, Fig. 2 shows that the period between January 1788 and April 1791 is clearly anomalous for all four measurements, although the January 1788 changepoint is not detected in the sunset and 10 P.M. measurement. For the sunset measurement, this is likely to be related to the absence of data in the year 1786. For the 10 P.M. measurement, a discontinuity seems to be present at January 1788, but the associated



**Fig. 1** Detected change points, following the criteria explained in the text. The *horizontal axis* denotes the position of a changepoint in the record, *vertical* is the number of times a changepoint at this location is detected by subsequent applications of the regression models in the moving window. The *upper left and right panels* show the detected change points for the 8 A.M. and noon records respectively. The *lower left and right panels* show the detected change points for the sunset and 10 P.M. records

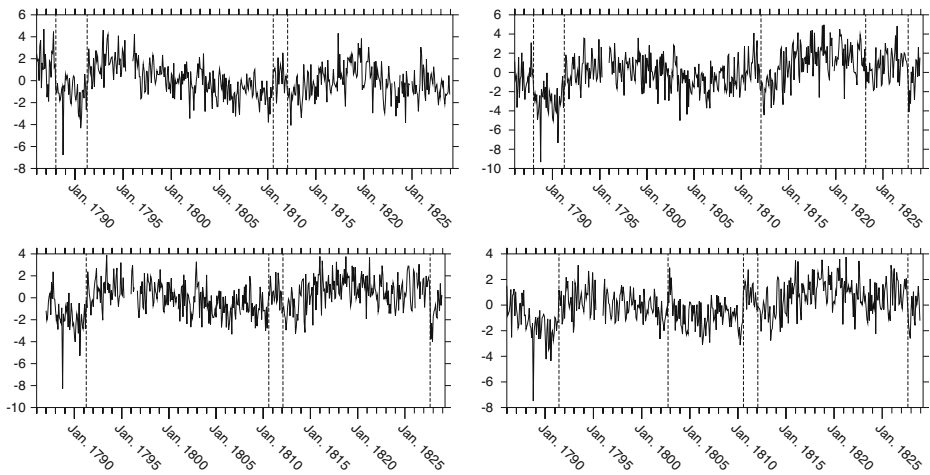
step (just) fails to pass the significance test. The jumps at August 1810 and February 1812 are shared by all four measurements, barring the February 1812 step in the noon measurement. Finally, a changepoint is present at October 1827, except for the 8 A.M. measurement.

A changepoint is detected in the noon measurement around March 1823, which seems to be related more to the fact that near this point a negative temperature trend is followed by a positive trend. This slow variation in the difference temperature

**Table 2** Detected changepoints in the Salem temperature record

Date	Step size (°F)	Date	Step size (°F)
<b>8 A.M.</b>		<b>noon</b>	
January 1788	−2.7	January 1788	−3.8
April 1791	+2.7	April 1791	+2.6
April 1795	−2.1	October 1802	+1.6
November 1789	+2.1	February 1812	−3.4
August 1810	+2.6	March 1823	−2.2
February 1812	−2.6	October 1827	−3.3
May 1818	+2.1		
November 1819	−2.1		
<b>sunset</b>		<b>10 P.M.</b>	
April 1791	+2.2	April 1791	+2.6
September 1797	+1.4	October 1802	+1.5
August 1810	+2.6	August 1810	+2.8
February 1812	−2.5	February 1812	−1.9
October 1827	−3.5	October 1827	−2.7





**Fig. 2** Uncorrected difference temperature (Salem record minus New Haven record) for the 8 A.M. and noon measurement (*top left and right* respectively), and the sunset and 10 P.M. measurement (*bottom left and right* respectively). The *vertical axis* gives the difference temperature in °F. The *vertical lines* in the diagrams denote the most robustly detected changepoints

is evident in all four temperature measurements which motivates us to reject this changepoint. A changepoint at October 1802 in the 10 P.M. is detected, but this seems to be related more to the brief positive excursion of the difference temperature, which is seen in all four temperature measurements, than to a ‘true’ changepoint. This jump is rejected too.

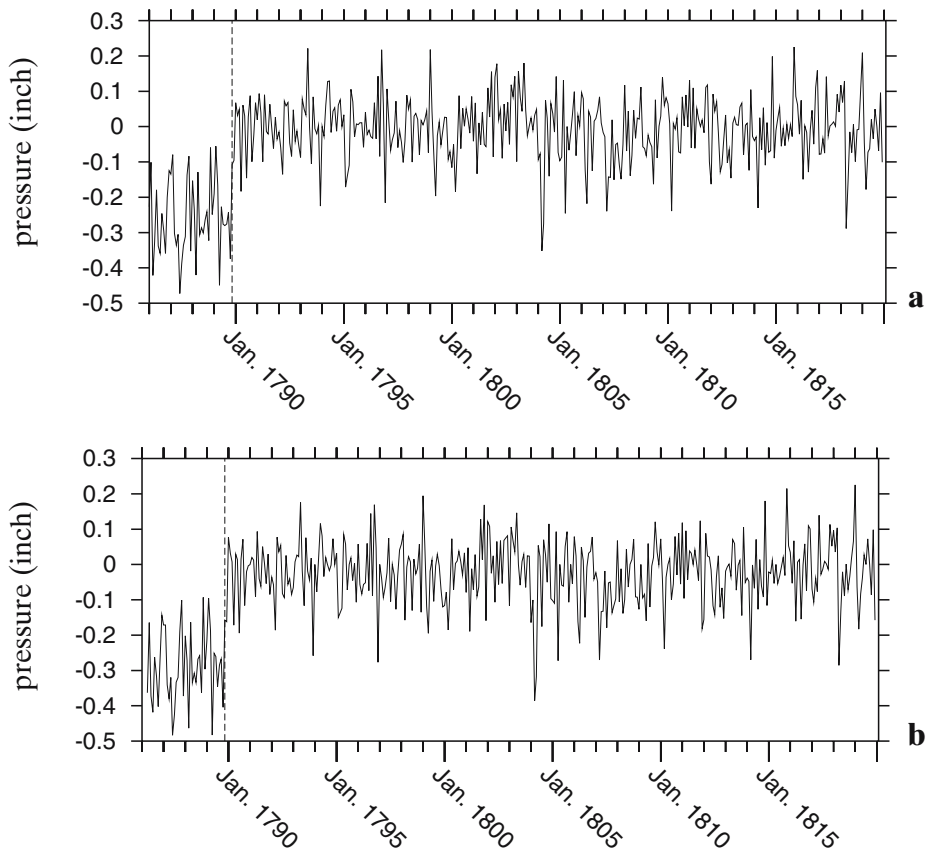
### 3.5.2 Pressure

The timeseries of the pressure measurements made at 8 A.M. and 10 P.M. required one rather obvious correction. The time series are tested separately for homogeneity, but the location of the identified changepoint are identical for the two timeseries. Moreover, the adjustments required to make the timeseries homogeneous are very similar too (Table 3). The uncorrected pressure series, as monthly averaged values and the location of the identified changepoint are shown in Fig. 3.

Dr. Holyoke was unable to make any barometer readings during most of January 1794. The readings are replaced by the comment “Barometer deranged during this time” and he adds “Probably some of the irregularities of several preceding months are owing to the imperfection of the Barometer”. Apparently, the barometer was repaired in this period, and the measurements were continued on the evening of

**Table 3** Estimated bias (inch) of the pressure measurements in the Holyoke record. The position and value of the identified changepoint are given

8 A.M. Date	Step (inch)	10 P.M. Date	Step (inch)
November 1789	0.241	November 1789	0.243



**Fig. 3** Uncorrected pressure (anomalous monthly values) with the positions of the located change-point, for **a** the 8 A.M. measurement and **b** the 10 P.M. measurement

January 28<sup>th</sup> 1794. The homogenization procedure does not indicate any step changes in this month.

It seems likely that the barometer was brought to the Rev. John Prince, rector of the First Church in Salem, for its maintenance. In Holyoke's time, John Prince was the one instrument maker of special ability in Massachusetts (Middleton 1964).

### 3.6 Further adjustment to the pressure record

The annual cycle of the Holyoke and Wigglesworth records has also been compared to modern data from Barnstable Municipal Airport which is very near to Hyannis (MA). These data are from the 1973–2004 period. The reason for choosing this station is threefold. It has both temperature and pressure data available at similar observation times as the Holyoke record. Furthermore, following DeGaetano (1996), data from a rural area is chosen to avoid an urban influence as much as possible. Finally, similar to Salem, the modern station is located near the Atlantic seaboard.

**Fig. 4** Annual cycle in pressure for the Holyoke data (*lower curve*) and the modern data (*upper curve*) before adjustment of the mean of the older record, with their running means (*dashed*). Pressure is an arithmetic mean of the 8 A.M. and 10 P.M. readings for both records

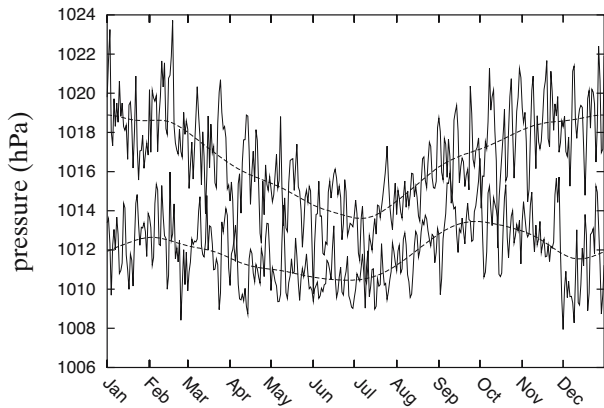


Figure 4 shows the annual cycle of the Holyoke and modern pressure readings (a simple arithmetic mean of the readings of 8 A.M. and 10 P.M.). It shows systematically lower pressure values in the older record. The difference is ca. 3.2 hPa in summer (June–July–August) and ca. 6.7 hPa in winter (December–January–February; Table 4).

An additional argument to adjust the pressure readings so that they match the average summer readings of the modern data, is that the annual cycle in temperature for the early and the modern data is similar for the summer period (Fig. 5a). Large differences between the average seasonal cycles of the early and modern data exist in the winter season.

A systematic difference in the annual cycle between the old and modern pressure records is not to be expected; the weight of the air column over north-eastern USA at the turn of the 18th century should not be different throughout the year than it is two centuries later. Figure 4 suggests that from spring to early autumn the behaviour of the annual cycle is similar for the modern and old data, the late autumn and winter season, though, are clearly anomalous. A difference in the annual cycle of sea-level pressure could then be related to dynamic processes in winter; the atmospheric circulation is more energetic in the cold season than that in the warm season.

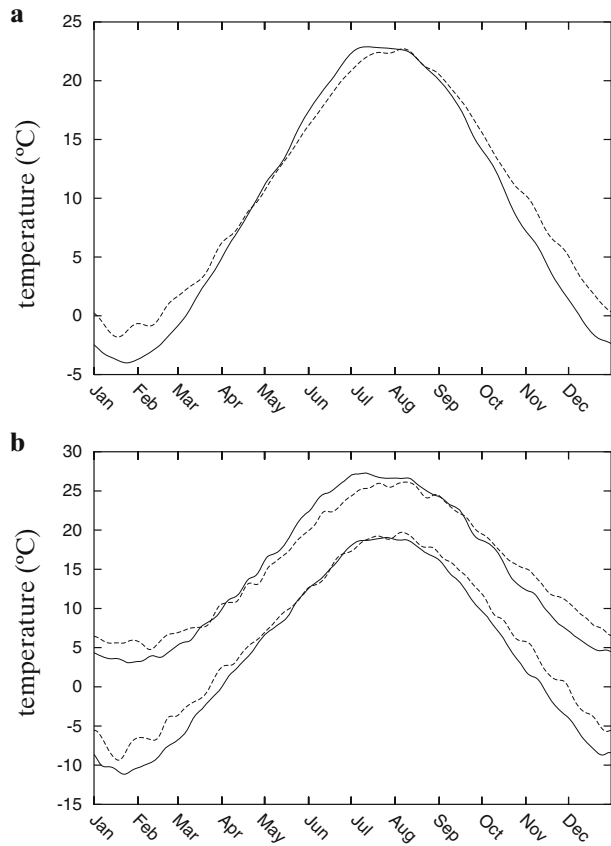
Following this reasoning, all pressure readings in the Holyoke record are adjusted by adding 3.23 hPa to ensure that the average *summer* pressure is similar to that of the modern record.

A systematic offset in the seasonal cycle is also found when comparing the Wigglesworth data with the modern data; it is even larger than the offset found

**Table 4** Mean values of the pressure readings as a function of the season (for the Holyoke and modern data: arithmetic mean of 8 A.M. and 10 P.M. readings, for the Wigglesworth data: 9 A.M. and 10 P.M.)

Season	Holyoke	Wigglesworth	Modern
DJF	1011.97	1006.32	1018.73
MAM	1011.37	1004.85	1015.93
JJA	1011.01	1005.12	1014.24
SON	1013.21	1007.71	1017.72

**Fig. 5** **a** Annual cycle for of temperature for the modern data (1973–2004, *dashed*) and for the Holyoke data (1786–1828, *solid*). **b** The annual cycle of the values of the 90th and 10th percentiles for the modern data (*dashed*) and the Holyoke data (*solid*)



between the Holyoke record and the modern data. The difference in seasonal cycle between the Holyoke and Wigglesworth pressure records is ca. 5 hPa and more-or-less constant over the year, with the Wigglesworth record showing the lower values.

Middleton (1964, p. 342) also noticed the low pressure readings from the Champney barometer Wigglesworth used, and argues that the barometer must have suffered in coming across the Atlantic. It may be that a similar argument applies to the barometer used by Holyoke, which had gone through a similar journey.

### 3.7 Comparison with other records

Only a few meteorological records exist for the period under consideration and from the vicinity of Salem. Some early instrumental data from the USA, dating from the mid-eighteenth to early nineteenth century, are compiled by Havens (1956). One of these records is a monthly temperature record from New Haven (briefly discussed in Section 3.5.1). Temperature values from the Holyoke record and the New Haven record (with the seasonal cycle removed) show a high correlation of 0.8. For the Holyoke record, a simple arithmetic mean of the temperatures at 8 A.M., noon and 10 P.M. is taken to produce a “daily mean.” The weight of the three measurements

in calculating the “daily mean” is equal and does not vary throughout the year. The good correspondence between the two records lends further confidence to the Holyoke record.

Daily records of the arithmetic mean of the temperature readings at 8 A.M., noon and 10 P.M. of the Holyoke record and daily values of the arithmetic mean of temperature readings of the Wigglesworth record of 8 A.M., 1 P.M. and 9 P.M. are compared over the period January 1786–December 1789. The records correlate very well at 0.99, although the means of these records differ by 0.83°C. This discrepancy is only partly related to the difference in observation times of the two records. Replicating the difference in observation times, averaged over a 4 year period, in a modern record, indicates that this effect accounts for ca. 0.3°C.

The arithmetic mean of the pressure readings at 8 A.M. and 10 P.M. of the Holyoke record and the arithmetic mean of the Wigglesworth record of 8 A.M. and 9 P.M. are compared over the period January 1786–December 1789. The correlation between these records is high at 0.91.

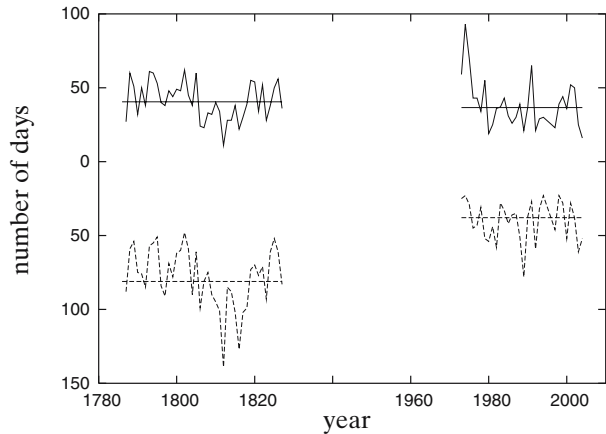
#### 4 Analysis of the daily records

The long-term study of extremes has usually been limited to the last few decades where digitized data are available. The early instrumental measurements from Salem discussed in this study allow for a comparison of late twentieth century extremes in temperature and sea-level pressure with those in the late eighteenth to early nineteenth century in the north-eastern USA. A preliminary analysis of daily extremes of temperature and seasonal values of pressure for Salem is compared against modern temperature and pressure measurements.

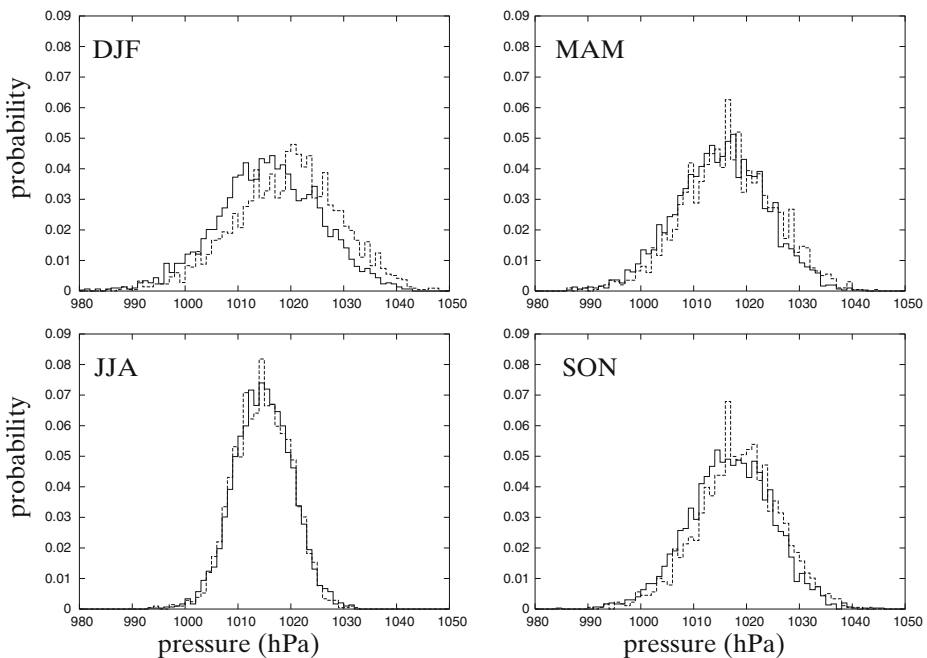
##### 4.1 Temperature

To construct the annual cycle, an average temperature value for each day of the year is calculated for both the Holyoke-record and the modern record. A ‘daily’ temperature is calculated by taking a simple arithmetical mean of the measurements made at 8 A.M., noon and 10 P.M. Simply averaging the values for each day of the year does not produce a smooth annual cycle. A more realistic annual cycle is produced by smoothing the values with a 31-point Gaussian-weighted running mean, over a 14-month long period, starting with December and ending with January. Figure 5a shows the annual cycles for the old and modern period and it demonstrates that the average summer temperatures of the 1786–1829 period were comparable to modern summer temperatures, but the winter temperatures were ca. 3°C cooler than in the modern period. Figure 5b shows the annual cycle in the values of the 90th and 10th percentiles of the daily temperatures. These latter curves have been smoothed with a 21-point Gaussian-weighted running mean to show more clearly periods with a higher variability in the daily temperature for the time of year, as argued by Jones and Lister (2002). Here the 90th and 10th percentiles of the data are calculated by fitting a gamma distribution using maximum likelihood methods. The reason for using the gamma distribution rather than a Gaussian distribution is that the temperature values are significantly negatively skewed during most of the year (Jones and Lister 2002). This figure demonstrates that in summer the temperature in the earlier period was

**Fig. 6** Extreme day counts for the Salem daily data. The *upper part* shows the annual count of warm days greater than the 90th percentile (*solid*) for the data from the early and modern periods with their mean values. The values of the 90th and 10th percentiles are calculated over the period 1973–2004. The lower part gives the annual counts of cold days less than the 10th percentile (*dashed*). Note the inverted scale for the cold days and the discontinuity in the horizontal time axis



similar to that in the modern period. The values of the 90th percentile of summer temperatures may have even been a little higher in the earlier period. In winter, however, both the values of the 10th and the 90th percentiles of the temperatures in the earlier period were considerably below that of the modern period.



**Fig. 7** Probability distribution for daily pressure. The *solid line* refers to daily pressure for the period 1786–1820, the *dashed line* for 1973–2004. Subpanels give the probability distribution for winter (DJF), spring (MAM), summer (JJA) and autumn (SON)

Figure 6 shows the number of warm and cold days for the modern and old periods. A warm (cold) day is a day with a temperature above (below) or equal to the 90th (10th) percentile, where the percentiles are calculated over the modern period. This figure shows that the number of warm days during 1786–1829 is more-or-less similar, or even slightly higher, compared to the number of warm days in the modern period. The number of cold days, however, were much more numerous during 1786–1829 at ca. 81 year<sup>-1</sup> than in the modern period where the number of cold days is ca. 38 year<sup>-1</sup>.

#### 4.2 Pressure

An average pressure value for each day in the record is calculated for both the Holyoke-record and the modern record by taking a simple arithmetical mean of the measurements made at 8 A.M. and 10 P.M. From these 'daily' pressure values, a probability distribution for 1786–1820 and 1973–2004 is calculated, in bins of 1 hPa (Fig. 7). This figure shows that the pressure in winter (DJF) during the earlier period was considerably lower than it is during the recent period. The shape of the probability distribution for the older record is similar to that of the recent period, but it is shifted towards lower values. The difference in the mean of the two distributions is 3.6 hPa. In spring and autumn, this difference reduces to 1.5 hPa.

### 5 Summary

This paper has documented the digitization, adjustments and homogenization of a daily temperature and pressure record for Salem (Massachusetts). These records are from a meticulously kept meteorological diary of Dr. Edward A. Holyoke, which was previously published by Hale (1833).

The temperature measurements, made at 8 A.M., noon, sunset and 10 P.M., span the period January 1st 1786 to March 31st 1829 and the pressure measurements, made at 8 A.M. and 10 P.M., span the period January 1st 1786 to December 31st 1820. In addition to the temperature and pressure readings, there are four daily characterizations of the weather.

In addition to the Holyoke record, a much shorter meteorological journal of Prof. Edward Wigglesworth from Cambridge (MA) is digitized and adjusted. This record spans the period 1780–1789 and 1793, with three daily temperature (outdoors) and three daily pressure readings. For the period January 1786–December 1789 three temperature readings (indoors) were made too. The pressure readings were discontinued in 1793.

The Holyoke temperature record has been compared to a monthly temperature record from New Haven (Connecticut, ca. 220 km from Salem in a SW direction) and shows a strong correlation. A strong correlation ( $> 0.9$ ), both for temperature and pressure, is also found between the daily Holyoke and Wigglesworth records for the period they overlap, lending further confidence to the Holyoke record. A systematic offset between the Holyoke and Wigglesworth records exists of ca. 5 hPa.

The Holyoke record was compared with modern measurements from Barnstable Municipal Airport, near Hyannis (MA). This showed that the Holyoke pressure data

showed a offset with modern pressure measurements of ca. 3 hPa in summer. This offset has been corrected by assuming that the mean summer pressure values were similar in the older and the modern period.

A probability distribution for daily pressure values shows that during winter the means of these distributions were 3 to 4 hPa apart, with pressure values in the older period lower than in the modern period. The two available reconstructions of sea-level pressure for the North Atlantic sector which include this period (Lamb and Johnson 1959; van der Schrier and Barkmeijer 2005) show an anomalously low pressure over the north-eastern USA. However, these reconstructions make use, amongst others, of the Holyoke record, so this latter argument has a degree of circularity in it. Nevertheless, for the van der Schrier and Barkmeijer (2005) reconstruction holds that an anomalous low pressure area over the north-eastern USA was still found (albeit weaker) when they did not make use of the Holyoke record in their reconstruction.

The annual cycle in temperature for the Holyoke record matches that of the modern record for the summer season. Average winter temperatures were ca. 3°C lower in the early period than in the modern period. The daily Salem temperature record was analyzed for extremes by counting days  $\geq$  90th and  $\leq$  10th percentile for each year. The series shows a drastic increase in the number of annual cold extremes in the older record, while the number of annual warm extremes is similar for the two periods.

Concluding, with four daily temperature measurements over the period 1786–1829 and two pressure measurements for the period 1786–1820, and with only 1.5 and 3.4% of data missing respectively, the Holyoke record can be used to study day-to-day variability in the 1786–1829 period.

The Holyoke and Wigglesworth pressure and temperature records are made available through web-site of the Climatic Research Unit ([www.cru.uea.ac.uk](http://www.cru.uea.ac.uk)).

**Acknowledgements** Herman van der Schrier, the father of GvdS, is thanked for his help in digitizing the Holyoke data and Jim McAllistar ([www.salemweb.com](http://www.salemweb.com)) is thanked for his information on historic Salem. Thanks also goes to Thomas C. Peterson and John M. Kobar for the Hyannis data and Lucie Vincent and Malcolm Haylock for their assistance with the homogenization. We thank John Kington, Rob Allan and Hendrik Wallbrink for their constructive comments. GvdS is funded by the UK Natural Environment Research Council (NERC) through the RAPID Climate Change programme.

## References

- Brown RH (1940) The first century of meteorological data in America. *Mon Weather Rev* 68: 130–134, May
- DeGaetano AT (1996) Recent trends in maximum and minimum temperature threshold exceedences in the northeastern United States. *J Climate* 9:1646–1660
- Hale E (1833) A meteorological journal from the year 1786 to the year 1829, inclusive, by Edward A. Holyoke, M.D., A.A.S. *Mem Am Acad Arts Sci* 1:107–216
- Havens JM (1956) An annotated bibliography of meteorological observations in the United States, 1731–1818. Technical Report No. 5, Florida State University, Department of Meteorology. Prepared under project NR-082-071, contract NONR-1600(00) with the Office of Naval Research
- Jones PD, Lister DH (2002) The daily temperature record for St. Petersburg (1743–1996). *Clim Change* 53:253–267
- Jones PD, Raper SCB, Santer B, Cherry BSG, Goodess C, Kelly PM, Wigley TML, Bradley RS, Diaz HF (1985) A grid point surface air temperature data set for the Northern Hemisphere.



- Technical Report TRO22, U.S. Dept. of Energy Carbon Dioxide Research Division, prepared under contract No. DE-AC02-79EV10098
- Lamb HH, Johnson AI (1959) Climatic variation and observed changes in the general circulation. Part I and Part II. *Geogr Ann* 41:94–134
- Ludlum DM (1966) Early American winters 1604–1820. The history of American weather. American Meteorological Society, Boston
- Lund R, Reeves J (2002) Detection of undocumented changepoints: a revision of the two-phase regression model. *J Climate* 15:2547–2554
- Middleton WEK (1947) Meteorological instruments. The University of Toronto Press, Toronto, Canada
- Middleton WEK (1964) The history of the barometer. Johns Hopkins University Press, Baltimore, MD
- Moberg A, Bergström H, Krigsman JR, Svanered O (2002) Daily temperature and pressure series for Stockholm (1756–1998). *Clim Change* 53:171–212
- van der Schrier G, Barkmeijer J (2005) Bjerknes' hypothesis on the coldness during 1790–1820 AD revisited. *Clim Dyn* 24:355–371, doi:[10.1007/s00382-004-0506-x](https://doi.org/10.1007/s00382-004-0506-x). Erratum: *Clim Dyn* doi:[10.1007/s00382-005-0053-0](https://doi.org/10.1007/s00382-005-0053-0)
- Vincent LA (1998) A technique for the identification of inhomogeneities in Canadian temperature series. *J Climate* 11:1094–1104
- Wilks DS (1995) Statistical methods in the atmospheric sciences. Academic, New York