

DAILY TEMPERATURE EXTREMES AND PRECIPITATION ACROSS INDIANA OVER A 120-YEAR PERIOD

John E. Frederick¹: Department of the Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, IL 60637 USA

ABSTRACT. A database of daily maximum and minimum temperatures recorded at 17 sites across Indiana from 1 September 1897 through 31 August 2018 shows relatively warm and cool periods alternating over decadal timescales. Average values in all 17 stations were computed over 20-year intervals in order to smooth out year-to-year variability. These intervals reveal an underlying pattern in daily extreme temperatures. The spread from the coolest 20-year period to the warmest is 1.1–1.2° C for both the mean daily maximum and minimum temperatures. Relatively warm daily maxima and minima exist in the 40-year interval 1918–1958, followed by cooler temperatures during the next two decades ending in 1978. Annual mean maxima and minima over the years 1998–2018 are warmer than in the previous 20-year period, 1978–1998, but still cooler than the peak reached earlier in the 20th century. However, when viewed over three-month intervals, daily temperature minima for March–May and June–August during the most recent 20-year period are the warmest in the record. The same is true for daily temperature maxima in March through May. However, 20-year mean daily temperature maxima for June through August have remained stable over the past 60 years. Time-integrated precipitation amounts for the 17-station composite increased during the second half of the 20th century. Twenty-year mean annual precipitation for the period 1998–2018 is 12–13% above corresponding values for 1938–1978.

Keywords: Climate, precipitation, temperature, weather

INTRODUCTION AND OBJECTIVES

The concept of climate encompasses all of the quantities required to specify the state of the Earth's atmosphere (e.g., Wilson et al. 1971; Barry & Hall-McKim 2014). Among these quantities, temperature and precipitation have a clear influence on society. Some of the processes that influence climate operate on a planetary spatial scale. These include absorption of solar radiation, emission and absorption of longwave thermal radiation by polyatomic gases such as carbon dioxide and water vapor, and exchanges of latent heat when liquid water evaporates from the Earth's surface and condenses in the atmosphere to form clouds and precipitation (e.g., Frederick 2008). Considerable attention has focused on changes in climate with emphasis on the role of human activities, particularly as they influence temperature (IPCC 1996, 2013). Analysis of the surface temperature dataset provided by stations throughout the world (Jones et al. 2012; Osborn & Jones 2014) shows an increase in mean surface temperature from the start of the 20th century up to about 1945, a flattening or a small decrease

until approximately 1975, and an increase thereafter. The rate of change during the first 10 to 15 years of the 21st century was smaller than during the period from 1975 to 2000. Based on these data the warming in global mean temperature from 1900 to 2015 was roughly 1° C (Met Office Hadley Centre 2018), although the change was not a simple linear trend over time. While global averages provide useful measures of large-scale climate, changes in the atmosphere need not be spatially uniform. For example, weather patterns that influence regional temperatures and precipitation respond to varying surface topography, properties of local ground cover such as reflectivity and moisture content, and proximity to large bodies of water (e.g., Sutton 1953). These factors mesh with large-scale variables such as the latitudinal temperature gradient to produce characteristic regional climates.

This work focuses on the regional climate of the state of Indiana using temperature and precipitation data collected over a period of 121 years. The geographic region of interest includes major agricultural activities and several metropolitan areas with populations greater than 100,000, including one with approximately two million people (STATS Indiana 2018). Agricultural

¹ *Corresponding author:* John E. Frederick, 1950 E. Greyhound Pass, Box 18-199, Carmel, IN 46033 USA.

Table 1.—Stations measuring daily temperature extremes and precipitation: percent data availability for the period 1 September 1897 through 31 August 2018.

Station (s): nearest town	Latitude (°N)	Longitude (°W)	T _{max} availability (%)	T _{min} availability (%)	Precipitation availability (%)
1. Mount Vernon	37.93	87.90	97.2	97.2	97.9
2. Princeton	38.36	87.57	98.2	98.2	98.8
3. Paoli	38.56	86.47	97.6	97.7	95.9
4. Vincennes	38.68	87.53	98.6	98.6	98.9
5. Madison	38.74	85.38	96.4	95.8	95.3
6. Oolitic	38.89	86.53	87.4	87.4	87.2
7. Columbus	39.12	85.92	99.4	99.4	99.2
8. Bloomington	39.17	86.53	96.7	97.5	97.3
9. Greensburg	39.35	85.48	97.9	97.6	96.5
10. Shelbyville	39.53	85.78	94.1	93.7	93.0
11. Rockville	39.76	87.23	97.2	97.1	97.0
12. Anderson	40.11	85.69	98.7	98.5	97.6
13. Farmland	40.25	85.13	83.3	83.3	93.9
14. Marion	40.57	85.66	99.5	99.5	98.8
15. Winamac	41.03	86.60	88.1	88.0	87.3
16. Angola	41.66	85.00	85.6	85.6	85.9
17. South Bend	41.71	86.25	96.4	96.4	96.4

productivity and water supplies depend on regional precipitation amounts, while air temperatures influence the length of the growing season and the demand for energy to heat or cool interior living space. Motivated by these issues, the objective of this work is to characterize changes that have occurred in daily maximum and minimum temperatures and in precipitation across Indiana over timescales of years to more than a century. This will provide a useful baseline against which to measure future behavior. Assessments of climate-related influences on activities throughout the state would benefit from knowledge of atmospheric changes that have actually occurred over the past several decades.

THE DATASETS AND DATA HANDLING

The datasets consist of daily maximum temperatures (T_{\max}), daily minimum temperatures (T_{\min}), and 24-hour integrated precipitation (P) at each of 17 sites across Indiana. The time frame covers 121 years from 1 September 1897 through 31 August 2018. All data were obtained via the web site of the NOAA National Centers for Environmental Information (CDO 2018). The stations are part of the Global Historical Climatology Network and have undergone established quality assurance checks before being made available for public access. Table 1 lists the stations from South to North across the state with their latitudes and longitudes. The number of

stations that have made weather observations in Indiana since the latter part of the 20th century is far larger than the subset in Table 1. The 17 stations selected for this work are those that were in operation over most or all of the time from 1897 through the present. The requirement for ongoing observations throughout the entire 20th century greatly limits the number of sites available.

It is important that the datasets be as complete as possible. Yet, gaps ranging from three months to several years exist in the record acquired since the year 2000 at several sites. In these cases data from one or more stations, not listed in Table 1, were used to fill the gaps and create a continuous record through 31 August 2018. Details of these special cases appear in the Appendix. Data gaps prior to the year 2000, when the number of observing stations was relatively small, were treated as described later. Table 1 lists the percentage of days during the 121 years on which data exist, where these values include data merged from nearby sites. Only stations operating over the time frame of interest where entries exist on 83% or more of the days enter the analysis. The average data availabilities for T_{\max} , T_{\min} , and P , computed over all 17 sites, are 94.8%, 94.8%, and 95.1%, respectively.

Data handling: temperatures.—The original data consist of the quantities $T_{\max}(s,y,d)$, $T_{\min}(s,y,d)$, and $P(s,y,d)$ where s labels a site in Table 1, y labels a year, and d labels the day

number of the year, $d=1, 2, \dots, 365$ or 366 where $d=1$ is 1 January. The multi-year mean daily extreme temperatures for each site s and day d are:

$$T_{\max}^C(s, d) = [\sum_y T_{\max}(s, y, d)]/N_x(s, d) \quad (1)$$

$$T_{\min}^C(s, d) = [\sum_y T_{\min}(s, y, d)]/N_n(s, d) \quad (2)$$

where data are available at site s on day-of-year d during $N_x(s, d)$ years for T_{\max} and $N_n(s, d)$ years for T_{\min} . The summations extend over all years, a maximum of 121, where each one-year period begins on 1 September and $y=1$ labels 1 September 1897 through 31 August 1898. The quantities $T_{\max}^C(s, d)$ and $T_{\min}^C(s, d)$ expressed as functions of d define climatological annual cycles in the daily maximum and minimum temperatures and serve as references for defining relatively warm and cool periods. The daily “temperature deviations” are:

$$\delta T_{\max}(s, y, d) = T_{\max}(s, y, d) - T_{\max}^C(s, d) \quad (3)$$

$$\delta T_{\min}(s, y, d) = T_{\min}(s, y, d) - T_{\min}^C(s, d) \quad (4)$$

Equations 3 and 4 characterize daily deviations from the long-term climatology at each site, although such high temporal resolution is neither necessary nor desirable when the focus is on long-term behavior. It is useful to consider average values of $\delta T_{\max}(s, y, d)$ and $\delta T_{\min}(s, y, d)$ computed over 3-month-long seasons, corresponding to specific ranges of d , or over entire years, corresponding to all values of d . Then, further averaging over all sites, corresponding to all values of s , can reveal the annualized behavior of temperature and identify extended warm and cool periods on a statewide basis from the late 19th century to the summer of 2018.

This work divides a year into four “meteorological seasons” based on the day-number d . Meteorological Autumn, labelled by $m=1$, encompasses $d=244-334$ and coincides approximately with September through November. Winter, $m=2$, being approximately December-February, refers to $d=335-365$ or 366 plus $d=1-60$ of the following calendar year. Meteorological Spring, $m=3$, defined by $d=61-151$ approximately spans March-May, and Summer, $m=4$, refers to June-August with $d=152-243$. The average values of $\delta T_{\max}(s, y, d)$ and $\delta T_{\min}(s, y, d)$ over season m at

each site are:

$$\delta T_{\max}^M(s, y, m) = [\sum_d \delta T_{\max}(s, y, d)]/M_x(s, y, m) \quad (5)$$

$$\delta T_{\min}^M(s, y, m) = [\sum_d \delta T_{\min}(s, y, d)]/M_n(s, y, m) \quad (6)$$

where the sums extend over the number of days in season m of year y on which data exist at site s , labelled as $M_x(s, y, m)$ and $M_n(s, y, m)$. When data are available on every day of the season $M_x(s, y, m) = M_n(s, y, m) = 91$ or 92 . Alternately, annual-average values, labelled $\delta T_{\max}^A(s, y)$ and $\delta T_{\min}^A(s, y)$ result from extending the summations in Eqs. 5 and 6 over all values of d in each 12-month period and dividing by the appropriate number of days on which data exist. As used in this work, one year begins on $d=244$, 1 September in non-leap years, and extends through $d=243$, 31 August of the following calendar year, where time is labelled by the calendar year that contains September.

Finally, the mean over all sites of the seasonal or annual values gives a state-wide picture of the behavior of temperature extremes. The seasonal values are:

$$\delta T_{\max}^N(y, m) = [\sum_s \delta T_{\max}^M(s, y, m)]/S_{x1}(y, m) \quad (7)$$

$$\delta T_{\min}^N(y, m) = [\sum_s \delta T_{\min}^M(s, y, m)]/S_{n1}(y, m) \quad (8)$$

where $m=1, 2, 3, 4$, and the annualized values are:

$$\delta T_{\max}^S(y) = [\sum_s \delta T_{\max}^A(s, y)]/S_{x2}(y) \quad (9)$$

$$\delta T_{\min}^S(y) = [\sum_s \delta T_{\min}^A(s, y)]/S_{n2}(y) \quad (10)$$

where $S_{x1}(y, m)$ and $S_{n1}(y, m)$ are the number of sites, up to 17, that have seasonal mean data in season m of year y . Similarly, $S_{x2}(y)$ and $S_{n2}(y)$ are the number of sites that have annual mean data in year y .

If sites throughout the state experience relatively warm or relatively cool years, each of the $\delta T_{\max}^A(s, y)$ -values will have a significant positive correlation with the multi-station mean $\delta T_{\max}^S(y)$ in Eq. 9. The same applies to the $\delta T_{\min}^A(s, y)$ and $\delta T_{\min}^S(y)$ in Eq. 10. If, however, the annual-mean deviations vary randomly from site-to-site, there will be a tendency for positive and negative values

Table 2.—Statistical properties of the temperature deviations δT_{\max} and δT_{\min} : correlations of annual means across stations and standard deviations of daily values.

Station (s): nearest town	Correlation coefficient	Correlation coefficient	Standard deviation	Standard deviation
	r_{\max}	r_{\min}	σ_{\max} (°C)	σ_{\min} (°C)
1. Mount Vernon	0.84	0.86	5.4	4.9
2. Princeton	0.89	0.85	5.4	5.1
3. Paoli	0.91	0.84	5.4	5.6
4. Vincennes	0.93	0.87	5.5	5.1
5. Madison	0.80	0.77	5.3	4.9
6. Oolitic	0.88	0.82	5.5	5.3
7. Columbus	0.91	0.72	5.5	5.2
8. Bloomington	0.92	0.63	5.4	5.2
9. Greensburg	0.85	0.58	5.5	5.4
10. Shelbyville	0.96	0.92	5.4	5.2
11. Rockville	0.68	0.89	5.4	5.3
12. Anderson	0.93	0.72	5.5	5.3
13. Farmland	0.94	0.80	5.5	5.3
14. Marion	0.92	0.92	5.5	5.3
15. Winamac	0.91	0.84	5.4	5.3
16. Angola	0.94	0.81	5.3	5.0
17. South Bend	0.87	0.74	5.4	5.1

to cancel in the summations of Eqs. 9 and 10. Correlation coefficients $r_{\max}(s)$ that relate $\delta T_{\max}^S(y)$ to each of the $\delta T_{\max}^A(s,y)$, $s=1, 2, \dots, 17$, measures the extent to which the entire state of Indiana experiences a similar temperature deviation in a given year. The analogous correlation coefficient relating $\delta T_{\min}^S(y)$ and the $\delta T_{\min}^A(s,y)$ is $r_{\min}(s)$. Table 2 lists r_{\max} and r_{\min} for each site. All values are positive and lie in the range 0.58 to 0.96; the average values of r_{\max} and r_{\min} computed over the 17 stations are 0.89 and 0.80, respectively. This shows a high degree of spatial coherence in the annual-mean temperature deviations across Indiana.

The mathematical development given above is straightforward, but the summations assume that temperature data exist for each day of each 12-month period at each station. Yet, Table 1 shows that, on average, measurements are not available on about 5% of the days. This work applies two methods to account for gaps in the temperature record. Method 1 computes the means in Eqs. 1 and 2 using the existing measurements, with $N_{\max}(s,d)$ and $N_{\min}(s,d)$ being the number of years that actually enter the summations for each site and day-number. Subsequent calculations based on Eqs. 3–10 use all available data in the summations with appropriate adjustment of the divisors. Method 1 effectively assumes that

missing data have no effect on the computed temperature deviations when averaged over a season, a year, or over multiple sites.

Method 2 for handling gaps in the record uses existing data to estimate values of missing temperatures. This approach is based on statistical regression and utilizes the fact that there is a significant positive correlation between temperature deviations across all stations, where the values in Table 2 illustrate this correlation on an annualized basis. For each day of the 121-year record, the average T_{\max} and T_{\min} values computed across all stations are:

$$T_{\max}^S(y, d) = [\sum_s T_{\max}(s, y, d)]/S_{x3}(y, d) \quad (11)$$

$$T_{\min}^S(y, d) = [\sum_s T_{\min}(s, y, d)]/S_{n3}(y, d) \quad (12)$$

where $s=1, 2, \dots, S_{x3}(y,d)$ in Eq. 11 and $s=1, 2, \dots, S_{n3}(y,d)$ in Eq. 12. Here $S_{x3}(y,d)$ and $S_{n3}(y,d)$ are the number of stations, up to 17, where values of T_{\max} and T_{\min} , respectively, exist on day number d of year y . Each day of the entire 121-year data record has a computed value of $T_{\max}^S(y,d)$ and of $T_{\min}^S(y,d)$; no gaps exist. The next step in Method 2 involves relating the daily values from each site to the multi-station means. This is done by determin-

ing the regression coefficients $a_0(s)$, $a_1(s)$, $b_0(s)$, and $b_1(s)$ in the expressions:

$$T_{\max}(s, y, d) = a_0(s) + a_1(s)T_{\max}^S(y, d) + \varepsilon_{\max}(s, y, d) \quad (13)$$

$$T_{\min}(s, y, d) = b_0(s) + b_1(s)T_{\min}^S(y, d) + \varepsilon_{\min}(s, y, d) \quad (14)$$

where $\varepsilon_{\max}(s, y, d)$ and $\varepsilon_{\min}(s, y, d)$ are residuals which approximate normal distributions with means of 0.0. The regression coefficients for each site, derived by minimizing the sum of squares of the residuals, have very high statistical significance. The t-statistics for a_1 and b_1 , defined as the ratio of the estimated value to its standard error, are near 1000. Equation 13 typically explains 94–98% of the variance in the $T_{\max}(s, y, d)$, with similar results for $T_{\min}(s, y, d)$ in Eq. 14. Equations 13 and 14, with the derived regression coefficients and residuals set to 0.0, are the basis for estimating values of $T_{\max}(s, y, d)$ and $T_{\min}(s, y, d)$ on all days when measurements are not available. The percentage variance explained by Eqs. 13 and 14 provides high confidence in the estimates provided by the approach. These computed values then replace the gaps in the daily record for each site.

The 121-year data record covers 44,194 days at each of 17 stations for a total of 751,298 station-days, and actual temperature measurements exist on 94.8% of these. Methods 1 and 2 are different approaches to cope with the 5.2% of station-days when data are missing, although the quantity of estimated data is substantial for Oolitic, Farm-land, Winamac, and Angola. Method 2 captures day-to-day fluctuations in temperature that extend over the entire state via use of $T_{\max}^S(y, d)$ and $T_{\min}^S(y, d)$, while the regressions scale this behavior to conditions encountered at each individual station. Furthermore, when missing data are filled in using Method 2 the final datasets that enter the various averages are effectively “complete” with no remaining gaps. Given its relative sophistication, this work regards Method 2 as the preferred approach to account for missing data. However, differences between the seasonal and annual-mean temperature deviations produced by Methods 1 and 2 can provide an index of the accuracy of the results.

Data handling: precipitation.—The precipitation data for each site consist of total daily liquid precipitation values. For site s in year y

on day d this is $P(s, y, d)$ expressed in cm. As with temperature deviations, the year is divided into four meteorological seasons labelled by m . The total precipitation for season m in year y at site s is the sum over the corresponding daily values:

$$P^M(s, y, m) = \sum_d P(s, y, d) \quad (15)$$

where the summation extends over all day-of-year numbers encompassed by the season. A summation analogous to Eq. 15 over all days of each 12-month period, extending from 1 September through the next 31 August, produces annual total precipitation values, $P^A(s, y)$, for site s and year y . Application of Eq. 15 requires that one account for data gaps, and as with temperature, this must involve certain assumptions. Table 1 shows that, depending on the site, data were missing on from 0.8% (Columbus) to 14.1% (Angola) of the days from 1 September 1897 through 31 August 2018. This work considers two methods for handling the data gaps and examines the sensitivity of the conclusions to the different approaches.

Method 1 makes the extreme assumption that $P(s, y, d) = 0.0$ on each day when no measurement exists. This produces a seasonal sum for each station and 12-month period via Eq. 15 that is less than or equal to the true total precipitation, where the magnitude of the error depends on the number of days on which precipitation actually occurred but for which the data are missing.

Method 2 for filling data gaps is based on regression and is analogous to that used with the temperature datasets. Let $P_S(y, d)$ be the multi-station mean precipitation for day-number d of year y computed via:

$$P_S(y, d) = [\sum_s P(s, y, d)] / N_{ps}(y, d) \quad (16)$$

where s ranges from 1 to $N_{ps}(y, d)$ and where $N_{ps}(y, d)$ is the number of stations with data on day-number d of year y . Each day of the entire 121-year time frame has a value of $P_S(y, d)$. The multi-station mean computed via Eq. 16 is the basis for estimating values of missing data at individual stations. Consider station s whose precipitation record has gaps to be filled in. The existing data from station s and the $P_S(y, d)$ -values for the same days are the basis of a regression:

$$P(s, y, d) = c_0(s) + c_1(s)P_S(y, d) + \varepsilon_p(s, y, d) \quad (17)$$

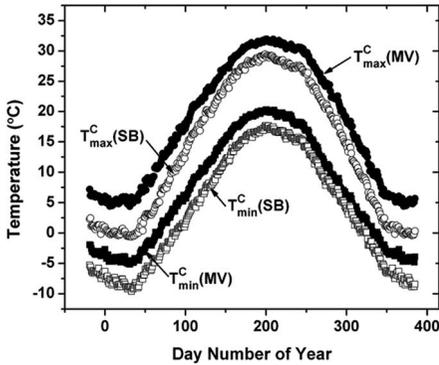


Figure 1.—Climatological annual cycles in daily maximum T_{\max}^C and daily minimum T_{\min}^C temperatures for Mount Vernon (MV, solid symbols) and South Bend (SB, open symbols). Days 0 and 365 refer to 31 December in adjacent non-leap years. The horizontal scale extends from 20 days before Day 0 to 20 days after Day 365.

where $\varepsilon_p(s,y,d)$ is the residual, and the best-fit coefficients $c_0(s)$ and $c_1(s)$ for the site are derived by least-squares methods. In practice, $c_0(s)$ varied from -0.020 (Oolitic) to $+0.041$ (South Bend), while $c_1(s)$ varied from 0.525 (South Bend) to 1.243 (Oolitic). All values of $c_1(s)$ had very high statistical significance, with the t -statistic in the range $t = 93$ to 250 . However, the percent variances explained by Eq. 17 are much smaller than the corresponding regressions for temperature, ranging from only 16.4% for Oolitic to 61.3% at South Bend with an average over all stations of 44.1%. This indicates a larger degree of spatial and temporal variability in precipitation than in temperature. Equation 17 with the residual set to 0.0 allows estimating missing $P(s,y,d)$ -values for each station. On days when $P_m(y,d) = 0.0$, missing data were filled in with $P(s,y,d) = 0.0$ regardless of the value of $c_0(s)$. Method 2 has the advantage of using information from multiple sites to fill in missing data, where the regression coefficients customize the estimates to each individual station. Method 2 should produce accurate estimates of total precipitation over a season or a year, although the daily estimates will contain larger errors than in the case of temperature.

RESULTS: TEMPERATURE

Figure 1 presents the climatological annual cycles in daily extreme temperatures, $T_{\max}^C(s,d)$ and $T_{\min}^C(s,d)$ as functions of d for the southernmost and northernmost stations in Table 1,

Mount Vernon and South Bend, respectively. These curves illustrate the magnitude of the latitudinal gradient across Indiana and the contrast between stations. The warmest daily maxima occur in the period 12–28 July, several weeks after the summer solstice. The coldest daily minima appear over an extended period from approximately 2 January to 12 February. Averaged over the annual cycle, the daily maximum at Mount Vernon is 4.1°C warmer than at South Bend, while the daily minimum averages 3.3°C warmer. The relevant quantities for this work are temperature deviations defined relative to a climatological annual cycle, as in Fig. 1, computed separately for each station. Variations in regional climate that affect the entire state of Indiana will appear as systematic changes in the temperature deviations δT_{\max} and δT_{\min} imposed on a background that includes latitudinal variations and site-specific influences.

The temperature deviations defined by Eqs. 3 and 4 are well-approximated by normal distributions with means of 0.0 and standard deviations that vary little with location. Figure 2 presents histograms of $\delta T_{\max}(s,y,d)$ and $\delta T_{\min}(s,y,d)$ for Shelbyville where the solid lines indicate best-fit normal distributions. The standard deviations are 5.4°C for δT_{\max} and 5.2°C for δT_{\min} . Small deviations from a normal distribution appear as small positive values of δT_{\min} , but the wings of both distributions are well-represented by a Gaussian curve. The distributions of δT_{\max} and δT_{\min} for all stations are very similar to those for Shelbyville; Table 2 lists the standard deviations of the best-fit normal distributions to δT_{\max} and δT_{\min} for each site, labelled σ_{\max} and σ_{\min} , respectively. All of the standard deviations lie in the range $\sigma_{\max} = 5.3\text{--}5.6^\circ\text{C}$ and $\sigma_{\min} = 4.9\text{--}5.6^\circ\text{C}$. The following analyses are based on spatial averages of the original $\delta T_{\max}(s,y,d)$ and $\delta T_{\min}(s,y,d)$ over all 17 stations as well as temporal averages over seasons or the entire year.

Figure 3 presents the multi-station annual mean values, $\delta T_{\max}^S(y)$ and $\delta T_{\min}^S(y)$, computed via Eqs. 9 and 10, versus year for a 120-year period beginning in 1898. Data gaps were filled using the regression-based Method 2 described previously. Each point refers to 12 months beginning on 1 September and ending with 31 August of the following year. Horizontal line segments are means of these annual results over six 20-year periods beginning with 1 September of year y_1 and ending with 31 August of year y_2 extending from $(y_1,y_2) = (1898, 1918)$ to $(y_1,y_2) = (1998, 2018)$.

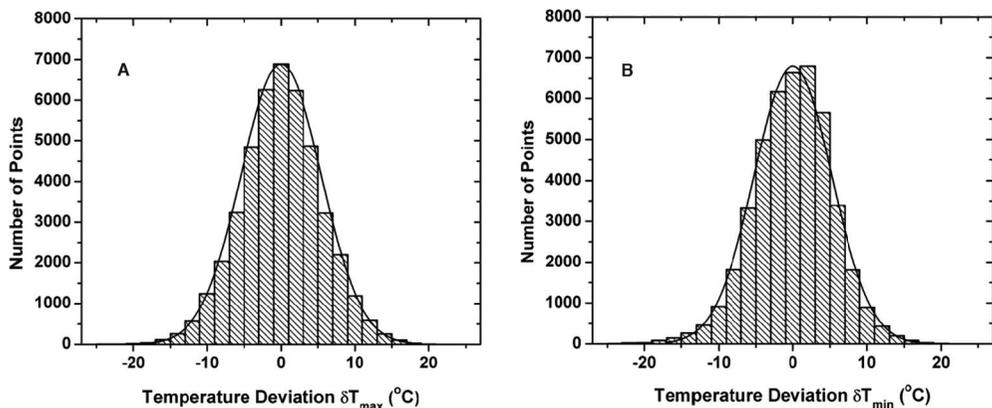


Figure 2.—Histograms of daily temperature deviations for Shelbyville over the period 1 September 1897 through 31 August 2018. The smooth curve indicates the best-fit normal distribution. (A) Deviations of daily maximum temperature δT_{\max} ; (B) Deviations of daily minimum temperature δT_{\min} .

Results for individual years span the range $\delta T_{\max}^S(y) = -2.1^\circ\text{C}$ to $+2.3^\circ\text{C}$ and $\delta T_{\min}^S(y) = -2.0^\circ\text{C}$ to $+2.1^\circ\text{C}$, but the erratic year-to-year variations lead to a substantial degree of cancellation in the 20-year means. The 20-year means of $\delta T_{\max}^S(y)$ vary from a minimum of -0.57°C for 1978–1998 to a maximum of $+0.55^\circ\text{C}$ for 1918–1938. Corresponding 20-year means of $\delta T_{\min}^S(y)$ are -0.70°C for 1958–1978 and $+0.48^\circ\text{C}$ for 1918–1938. The differences among the 20-year means are quite small with a spread from coldest to warmest being only $1.1\text{--}1.2^\circ\text{C}$ for both the daily maxima and minima.

Figure 3 shows drops in the 20-year mean temperature deviations from 1938–1958 to 1958–

1978 for both the maximum and minimum temperatures. The data imply a cooling between 0.9°C and 1.0°C , and these are the largest changes between two consecutive 20-year periods in the record. This raises the concern that an instrument-related shift might have influenced multiple sensors somewhere near the boundary between the two periods. Instrument history files provided with the datasets show that sensors were replaced at more than half of the sites between 1952 and 1956, although the old sensors were exchanged for new ones of the same type. An examination of annual-mean data from the individual stations shows no obvious shift in the records that coincide in time with the instrument changes. In most

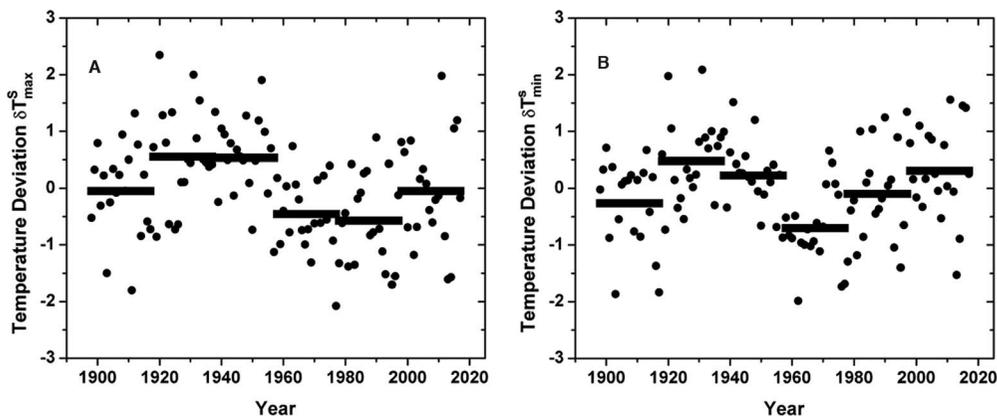


Figure 3.—Mean temperature deviations for one-year intervals extending from 1 September through 31 August for 1898 through 2018. Points refer to 12-month means averaged over 17 sites. Line segments are averages of 20 consecutive 12-month periods. (A) Deviations of daily maximum temperature δT_{\max}^S ; (B) Deviations of daily minimum temperature δT_{\min}^S .

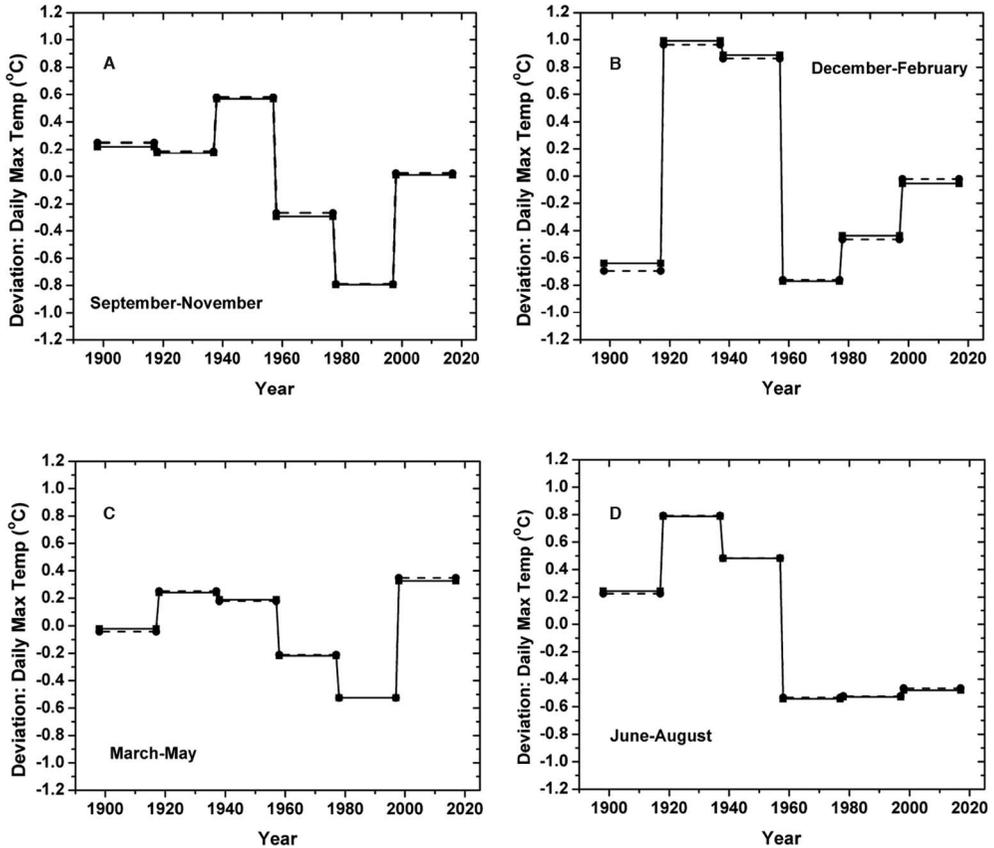


Figure 4.—Deviations in daily maximum temperature averaged over 17 sites and 20-year periods from 1 September 1898 through 31 August 2018. The panels refer to different meteorological seasons. (A) Autumn, September–November; (B) Winter, December–February; (C) Spring, March–May; and (D) Summer, June–August. Dashed lines use Method 1 for filling in missing data; solid lines use Method 2.

cases, no immediate change appeared or the decline in temperature did not develop until several years after a change in sensor. Based on the information available, there is no compelling basis to claim that the apparent cooling across Indiana from 1938–1958 to 1958–1978 is an artifact of the instrumentation, although a detailed assessment of events that occurred roughly 60 years ago is not possible. Any attempt to alter the archived data to remove hypothesized biases must involve subjective assumptions and would be open to serious criticism. Clearly, a geophysical origin for the patterns shown in Fig. 3 is predicated on a stable long-term data record. It is relevant to note that datasets published by IPCC (2013) show a slight decline in global-mean temperature between approximately 1955 and 1975 after reaching a peak in the 1940s. This global behavior is qualitatively consistent with

that observed across Indiana, although changes in the globally-averaged data are less pronounced.

Figure 4 presents 20-year means values of $\delta T_{\max}^N(y,m)$ computed using Eq. 7 for A = Autumn, $m=1$, B= Winter, $m=2$, C=Spring, $m=3$, and D = Summer, $m=4$, where these meteorological seasons are defined according to the day-numbers of the year given previously. The data record extends from September 1898 through August 2018. Solid lines account for data gaps via Method 2, while the dashed lines use Method 1. No significant differences exist between results based on these two different approaches, and this consistency supports the claim that missing data have negligible impact on the derived temporal behavior. A feature common to all seasons is a drop in δT_{\max}^N from 1938–1958 to 1958–1978, although the change in Spring is less prominent than in other seasons. As described above, there is

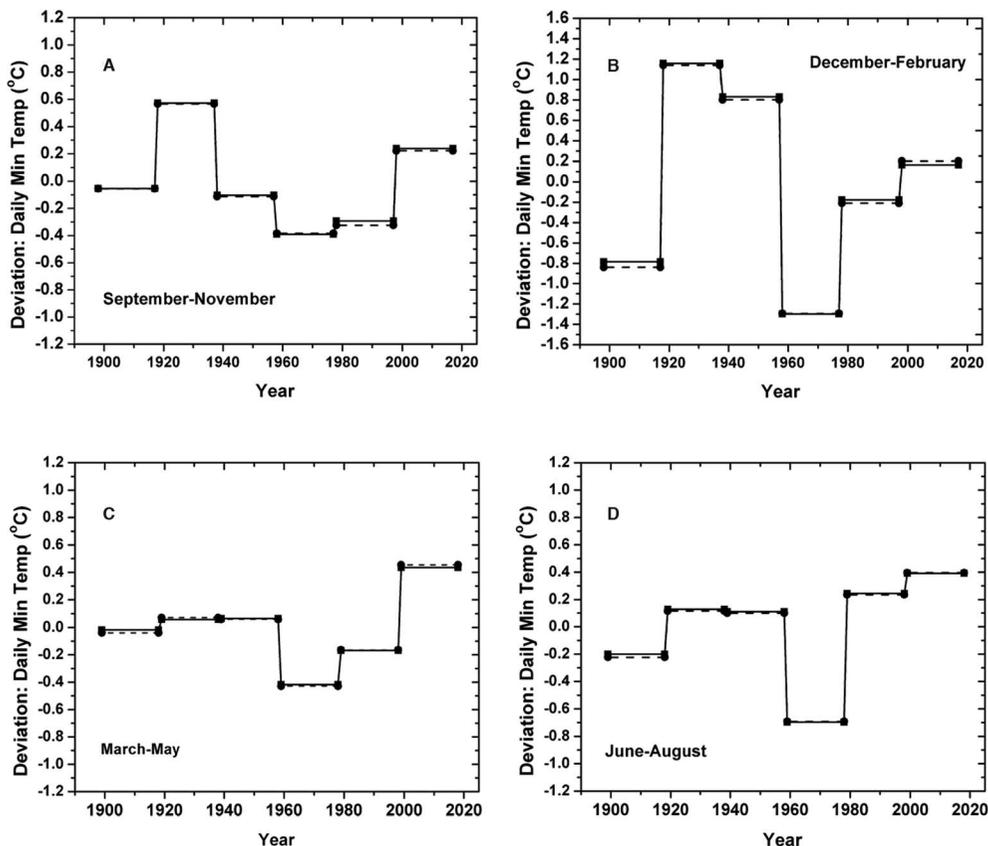


Figure 5.—Deviations in daily minimum temperature averaged over 17 sites and 20-year periods from 1 September 1898 through 31 August 2018. The panels refer to different meteorological seasons: (A) Autumn, September–November; (B) Winter, December–February; (C) Spring, March–May; and (D) Summer, June–August. Dashed lines use Method 1 for filling in missing data; solid lines use Method 2.

no convincing evidence in the historical records that this behavior is an instrument-related artifact.

The 20-year mean δT_{\max}^N -values are smallest in 1978–1998 for Autumn and Spring and in 1958–1978 for Winter and Summer, while larger values existed earlier in the 20th century. The daily maximum temperature deviations increase from 1978–1998 to the most recent interval, 1998–2018, in all four seasons, although the change in the summer maxima is negligible over the past 60 years. The most recent 20-year interval had the highest δT_{\max}^N -values of the 120-year record during Spring, while the most recent 20-year interval for Autumn, Winter and Summer ranked as the fourth, third, and fourth highest δT_{\max}^N of the six intervals for these seasons respectively. Although the temporal patterns in Fig. 4 are of interest, all of the values lie in the narrow range -1°

C to $+1^\circ$ C. When averaged over 20-year periods, changes in daily maximum temperatures across Indiana have been small. The greatest temporal change appears in meteorological winter where the contrast between the smallest and largest values in Fig. 4B is 1.7 – 1.8° C.

Figure 5 presents 20-year mean values of $\delta T_{\min}^N(y,m)$ computed via Eq. 8 for each meteorological season. Note that the vertical scale for Winter in Fig. 5B is expanded relative to those for other seasons. As with Fig. 4, the dashed lines based on Method 1 for filling data gaps differ insignificantly from the solid lines based on Method 2. A drop in the 20-year mean δT_{\min}^N from 1938–1958 to 1958–1978 occurs in all seasons, but the change in Spring is much less than in Winter and Summer. The corresponding change in Autumn is very small and, by itself, would not arouse concern over a possible

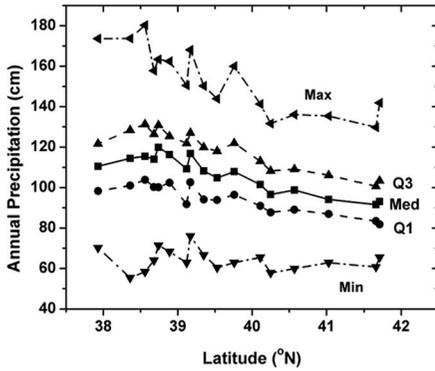


Figure 6.—Measures of annual total precipitation at each station expressed as functions of latitude based on 121 years of data. Curves denote the median (Med), the upper limit of the smallest quartile of values (Q1), the lower limit of the largest quartile of values (Q3), the minimum (Min) and the maximum (Max) at each site. Missing data were filled in using Method 2.

instrument artifact. All seasons show a monotonic increase in δT_{\min}^N over the most recent 60 years, from 1958–1978 to 1998–2018. Autumn and Winter of the most recent 20-year period have the second warmest δT_{\min}^N -values of the 120 years, while Spring and Summer of the most recent period have the highest δT_{\min}^N -values in the record. As with the δT_{\max}^N -values in Fig. 4, Winter is the most variable season for which the largest 20-year mean δT_{\min}^N -value is 2.4–2.5° C warmer than the minimum.

RESULTS: PRECIPITATION

Figure 6 summarizes the total 12-month precipitation received at each site over the 121-year period 1 September 1897 – 31 August 2018 with values expressed as functions of latitude across Indiana. Missing data were filled in using Method 2. The quantities presented are the median annual precipitation, the upper limit of the smallest quartile of values, the lower limit of the largest quartile, the smallest annual precipitation recorded at each site and the largest. One fourth of the annual precipitation amounts recorded at a site lie between each pair of adjacent curves. The annual median values cover the range from 91.6 cm for Angola to 119.8 cm at Madison. Mean precipitation amounts lie from 3.6 cm below to 2.3 cm above the medians depending on location. There is an obvious dependence on latitude across the state with smallest annual precipitation amounts occurring at the most

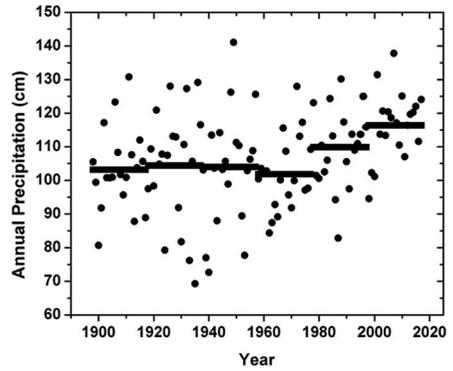


Figure 7.—Total annual precipitation for 12-month intervals from 1 September 1898 through 31 August 2018. Points refer to 12-month totals averaged over 17 stations. Line segments are averages of 20 consecutive 12-month periods. Missing data were filled in using Method 2.

northern locations. The six southernmost sites have median amounts in the range 110–120 cm, while values for the five northernmost stations lie between 90–100 cm. The minimum annual precipitation over the period is typically 55–65% of the median value while the maximum is 135–150% of the median.

The next issue involves temporal changes in precipitation over the state on time scales of years to over a century. Figure 7 presents values of total annual precipitation averaged over all 17 stations in Table 1, with data gaps filled via Method 2. The annual multi-station composites show a large scatter from year-to-year around a long-term average of 106.6 cm, varying from a minimum of 69.3 cm in 1935 to a maximum of 141.1 cm in 1949. The horizontal line segments in Fig. 7 denote 20-year averages beginning on 1 September 1898 and ending with 31 August 2018. Despite the large interannual variability, the 20-year means are remarkably stable over the first 80 years of the observing period followed by an increase in recent years. The 20-year mean annual value for the period 1 September 1958 – 31 August 1978 was 101.8 cm, rising to a maximum of 116.4 cm for 1 September 1998 – 31 August 2018.

Figure 8 presents 20-year mean seasonal precipitation amounts for A = Autumn, B = Winter, C = Spring, and D = Summer. Dashed lines denote results obtained using Method 1 which sets daily precipitation to 0.0 on days when data are missing; solid lines use the regression-based Method 2 for filling the gaps. The only major disagreement between results from the two

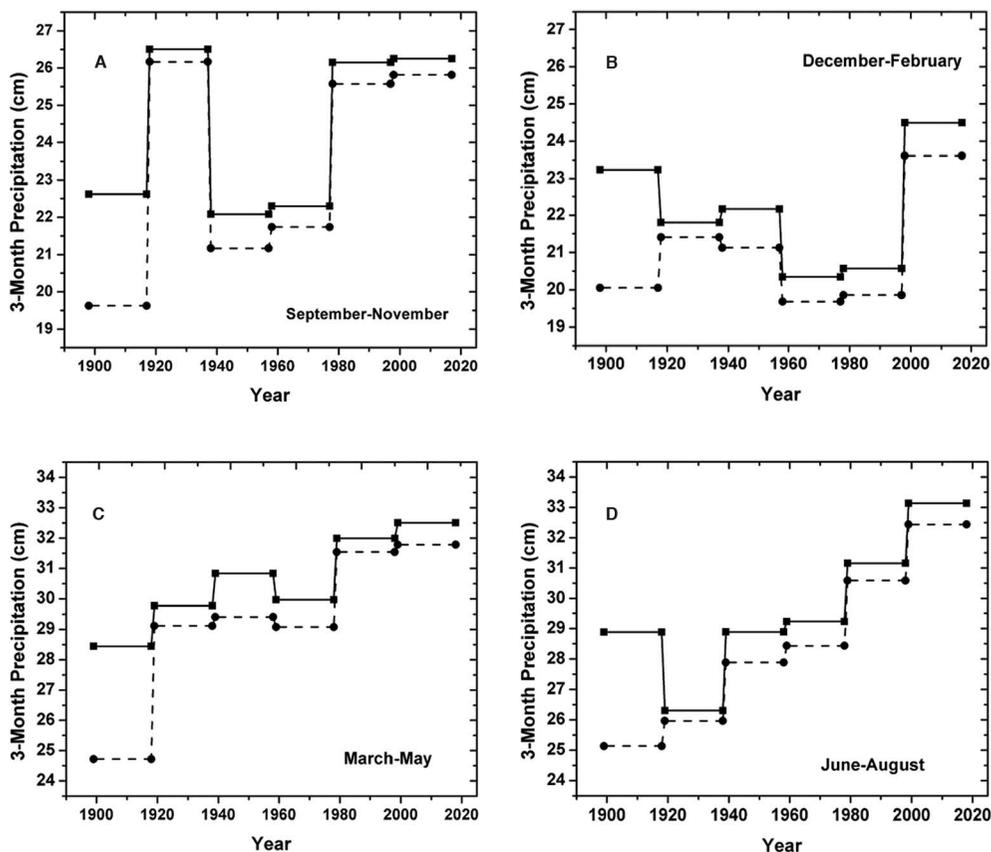


Figure 8.—Total precipitation for three-month periods averaged over 17 sites and 20-year intervals from 1 September 1898 through 31 August 2018. The panels refer to different meteorological seasons: (A) Autumn, September–November; (B) Winter, December–February; (C) Spring, March–May; and (D) Summer, June–August. Dashed lines use Method 1 for filling in missing data; solid lines use Method 2.

approaches appears in the earliest interval, 1 September 1898 – 31 August 1918, during which the datasets were less complete than in later decades. In this case, Method 1 provides a substantial underestimate of the true precipitation. Still, as of 1 September 1918 and afterwards, the same temporal pattern appears regardless of the method used to fill in missing data. All values quoted below are based on Method 2.

The largest annual precipitation may occur in Spring or Summer. Note the wider range of precipitation in the y-axis for those seasons. When averaged over the entire record, amounts received in Spring are slightly larger than in Summer, 30.6 cm versus 29.6 cm. Smaller long-term amounts occur in Autumn and Winter, 24.3 cm and 22.1 cm, respectively. The most striking feature of Fig. 8 is the growth in precipitation over time. Monotonic increases occur over the three most

recent 20-year periods in Winter and Spring, over the last four periods in Autumn, and over the last five 20-year periods in Summer. Twenty-year-mean annual precipitation for 1998–2018 exceeded that for 1958–1978 by 17.7% in Autumn, 20.4% in Winter, 8.5% in Spring, and 13.3% in Summer. In three of the four seasons, the largest total precipitation exists in the most recent 20-year period, while in Autumn the total for 1998–2018 is negligibly less, 0.25 cm, than the record set earlier in the 20th century.

DISCUSSION

The 12-month mean temperature deviations $\delta T_{\max}^S(y)$ and $\delta T_{\min}^S(y)$ averaged over all stations vary in the range -2.1°C to $+2.3^\circ\text{C}$ over the 121-year period 1 September 1897 through 31 August 2018. This interannual variation is largely random, and the long-term behavior is not well-

described by a simple trend line. However, when averaged over 20-year intervals a pattern consisting of alternating warmer and cooler periods emerges, where the range from coolest to warmest is less than 1.2°C . The 20-year-mean $\delta T_{\text{max}}^{\text{S}}$ and $\delta T_{\text{min}}^{\text{S}}$ values both have maxima in the interval 1 September 1918 to 31 August 1938, with $\delta T_{\text{max}}^{\text{S}}$ dropping to relatively low values in the 40-year span 1 September 1958 to 31 August 1998. The minimum for $\delta T_{\text{min}}^{\text{S}}$ is in the earlier of these two 20-year intervals. The two most recent 20-year periods for $\delta T_{\text{max}}^{\text{S}}$ and the three most recent for $\delta T_{\text{min}}^{\text{S}}$ show monotonic increases over time, although the 20-year mean temperature deviations for the period 1998–2018 do not reach the levels that prevailed for 1918–1938. Given the small magnitude of the 20-year-mean temperature changes, the observing network must provide a long-term dataset stable to a tolerance of less than 1°C in order to detect variations of geophysical origin unambiguously.

A similar overall pattern in the temperature deviations emerges when results are sorted according to meteorological seasons, although there are differences in detail. All seasons show a monotonic increase in $\delta T_{\text{min}}^{\text{N}}$ over the past three 20-year periods. In Spring and Summer the largest $\delta T_{\text{min}}^{\text{N}}$ -values of the record appear in the most recent 20-year period, indicating a tendency toward warmer daily minimum temperatures which typically occur near sunrise. The temperature minima depend on prevailing atmospheric water vapor amounts which modulate surface heating by longwave radiation and control latent heat release in condensation (e.g., Petterssen 1940). In the case of $\delta T_{\text{max}}^{\text{N}}$, the value for Spring shows a maximum in the most recent 20-year interval. In contrast, the mean value of $\delta T_{\text{max}}^{\text{N}}$ for Summer has remained nearly flat over the most recent three 20-year periods. These maximum values typically occur in mid-afternoon and are coupled to the prevailing level of sunlight as modified by cloudiness.

The 17-station composite implies an increase in seasonal and annual precipitation over time. In Winter, Spring, and Summer the largest seasonally-integrated precipitation values on record occurred in the most recent 20-year period, while in Autumn the most recent 20-year result was only 0.2–0.3 cm below the maximum established early in the 20th century. Viewed in an annualized sense, yearly total precipitation in the most recent 20-year period was 12–13% above the average value for 1 September 1938 to 31 August 1978. The

measurement of precipitation is straightforward and is not subject to issues of long-term stability that potentially influence the temperature record.

Long-term changes in daily maximum and minimum temperatures across Indiana show a temporal pattern similar to, but not identical to, global mean temperatures presented by IPCC (2013). A period of relative global-scale warmth existed early in the 20th century, followed by a period of constant-to-lower temperatures. A warming from the 1970s into the 21st century is the most recent feature of the global data. Globally-averaged data of necessity obscure day-night contrasts and regional-scale effects that can be substantial. The temperatures and precipitation experienced in the state are functions of prevailing weather patterns which include cloudiness, water vapor amounts, and air flow from other latitudes. Imposed on this regional variability are small systematic changes in the surface radiation balance in response to the buildup of long-lived greenhouse gases. Viewed on a year-to-year basis or as 20-year means, the archived meteorological records are likely to reflect regional influences more than the global-scale effects. There is no expectation that long-term changes confined to Indiana will precisely mimic those derived for a much larger geographic area or the entire globe.

ACKNOWLEDGMENTS

The author thanks the University of Chicago for financial and computing support as an emeritus faculty member during the course of this work.

Just prior to publication, the author passed away.

APPENDIX: THE TREATMENT OF GAPS IN DATA ACQUIRED SINCE 2000

The datasets from Vincennes, Madison, and Winamac contain gaps after the year 2000 in which no data exist for periods of three continuous months or longer. Such extended gaps prior to the year 2000 were treated by Methods 1 and 2 as described in the text. However, the enlarged number of stations in recent years allows one to substitute measured values taken from locations close to, or at least at the same latitude as, the long-established sites in Table 1. Such an approach was not possible in the early years of the data records due to the relative sparsity of stations.

For the period 1 July 2013 through 31 August 2018 temperature and precipitation data associated with the station near Vincennes in Table 1 came from the site

labelled “Vincennes 5NE” by the National Oceanic and Atmospheric Administration. This substitute location is at latitude 38.74°N, about 8 km northeast of Vincennes, Indiana. Similarly, for the three-month period 1 June through 31 August 2018, temperature and precipitation data associated with the site at Winamac in fact came from Woodburn, Indiana at latitude 41.16°N, essentially the same latitude as Winamac. The assembly of a complete dataset for Madison covering recent years proved more difficult than for Vincennes and Winamac. Temperature measurements for Madison ended in the summer of 2011, while daily precipitation measurements continued through 30 June 2018. Temperature data taken from nearby Big Oaks Wildlife Refuge, latitude 38.93°N, served as a substitute during the period 1 June 2011 through 31 July 2018, when this dataset ended. The final month of substitute temperatures, 1–31 August 2018, came from Crittenden, Kentucky, latitude 38.77°N. No precipitation data exist for either Big Oaks or Crittenden during the period 1 July through 31 August 2018. The final summer of precipitation data used for Madison comes from measurements at Seymour, Indiana, latitude 38.96°N, about 50 km to the northwest. These substitutions have negligible effect on the 17-station composite values on which the conclusions of this work are based.

LITERATURE CITED

- Barry, R.G. & E. Hall-McKim. 2014. *Essentials of the Earth’s Climate System*. Cambridge University Press, Cambridge, United Kingdom. 259 pp.
- CDO (Climate Data Online). 2018. NOAA National Centers for Environmental Information. Asheville, North Carolina. At: <https://www.ncdc.noaa.gov> (Accessed 18 October 2018).
- Frederick, J.E. 2008. *Principles of Atmospheric Science*. Jones and Bartlett Publishers, Sudbury, Massachusetts. 211 pp.
- IPCC. 1996. *Climate Change 1995: The Science of Climate Change*. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change. (J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg & K. Maskell, Eds.). Cambridge University Press, Cambridge, United Kingdom. 572 pp.
- IPCC. 2013. *Climate Change 2013: The Physical Science Basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex & P.M. Midgley, Eds.). Cambridge University Press, Cambridge, United Kingdom. 1535 pp.
- Jones, P.D., D.H. Lister, T.J. Osborn, C. Harpham, M. Salmon & C.P. Morice. 2012. Hemispheric and large-scale land-surface air temperature variations: an extensive revision and an update to 2010. *Journal of Geophysical Research* 117:D05127.
- Met Office Hadley Center. 2018. HadCRUT.4.6.0.0 (current version). At: https://www.metoffice.gov.uk/hadobs/hadcrut4/data/versions/HadCRUT.4.6.0.0_release_notes.html (Accessed 19 October 2018).
- Osborn, T.J. & P.D. Jones. 2014. The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. *Earth System Science Data* 6:61–68.
- Petterssen, S. 1940. *Weather Analysis and Forecasting*. McGraw-Hill Book Co., New York, New York. 503 pp.
- STATS Indiana. 2018. Indiana Business Research Center, Indiana University Kelly School of Business, Bloomington, Indiana. At: <http://www.stats.indiana.edu> (Accessed 18 October 2018).
- Sutton, O.G. 1953: *Micrometeorology*. McGraw-Hill Book Co., New York, New York. 333 pp.
- Wilson, C.L., W.H. Matthews, W.W. Kellogg & G.D. Robinson (Eds.). 1971. *Inadvertent Climate Modification - Report of the Study of Man’s Impact on Climate (SMIC)*. The MIT Press, Cambridge, Massachusetts. 308 pp.

Manuscript received 25 October 2018, revised 8 March 2019.