Systematic Biases in Manual Observations of Daily Maximum and Minimum Temperature

JON M. NESE

Department of Environmental Sciences, The Pennsylvania State University, Monaca, Pennsylvania

(Manuscript received 10 August 1992, in final form 10 September 1993)

ABSTRACT

The authors demonstrate that manual observations of daily maximum and minimum temperature are strongly biased toward temperatures ending in certain digits. The nature and severity of these biases are quantified using standard statistical methods. Temperatures ending in "0", "2", "5", and "8" are overrepresented in the data, with the bias toward multiples of ten being most statistically significant.

Inconsistencies in the distribution of the data by final digit suggest that biasing toward a temperature T may result not only from misobservations of temperatures $T \pm 1$ but also from misobservations of temperatures $T \pm 2$. Although changes adopted by the U.S. Weather Bureau in 1950 in the rules governing the rounding of temperature observations improved several of the biases, all biases remained statistically significant after the rule revision.

To estimate the potential effect of these biases on the mean and standard deviation of a temperature distribution, biasing simulations were performed on various normal distributions. In addition, it is shown that these biases can affect other relevant climatic statistics, such as the number of days that certain temperature thresholds are reached.

1. Introduction

Psychologists have recognized for years that humans are biased toward the base of a system of measurement. Meteorological observations also exhibit these biases, as demonstrated by Quayle (1971) and Reverdin (1988). Court (1951), in a statistical analysis of hourly temperature data from 1935 to 1939, noted that even-numbered temperatures were reported more frequently than odd-numbered temperatures, with the exception of temperatures ending in "5." This even-numbered bias likely resulted from the observational rule in place prior to 1950, which specified that temperatures ending in 0.5 were to be rounded to the nearest even degree (U.S. Department of Commerce 1941). To eliminate this bias, the U.S. Weather Bureau revised its rules on 1 January 1950, to require that all fractions up to and including 0.5 be rounded to the next lower integer (Court 1951). Since 1950, the observational rule has again been revised so that fractions ending in 0.5 are now rounded to the next higher integer (U.S. Department of Commerce 1988). (Although not the purpose of this paper, it can be argued that this second revision introduced an average warming of 0.1°F into the temperature data.)

In this paper, we quantify the nature and statistical significance of biases toward temperatures ending in certain digits by evaluating daily maximum and minimum temperature data from the period 1901–1986. We also evaluate the subperiods 1901–1949 and 1950–1986 to assess the effect of the 1950 revision in the observational rule. The most statistically significant and most consistent bias, both among stations and over time, is toward temperatures that are multiples of ten, while other highly significant biases exist toward temperatures ending in "5," "2," and "8." The exact secondary biases vary from station to station. The observational rule change in 1950 had some of the desired effect, most notably in reducing the bias toward multiples of ten, but significant biases persist in the data even after the rule change. We investigate whether these biases have any effect on climatic means and standard deviations, as well as on temperature threshold statistics.

2. Data sources

Daily maximum and minimum temperatures for the period 1901–1986 from 20 locations in the United States were used in the study. Only locations with less than 2.5% missing data during this period were used. These restrictions guarantee a total sample size N of more than 61 000 temperatures for each location. An attempt was made to use geographically diverse stations, but the principal criterion was completeness of

Corresponding author address: Dr. Jon M. Nese, Department of Environmental Sciences, The Pennsylvania State University, Beaver Campus, Brodhead Road, Monaca, PA 15061.

© 1994 American Meteorological Society

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC

the data record. The total number of daily maximum and minimum temperature observations from all 20 locations in the period 1901–1986 is 1 248 792: 711 128 observations from 1901–1949 and 537 664 observations from 1950–1986. The data was acquired from optical disks produced by U.S. West Optical Publishing, from data supplied by the National Climatic Data Center. The locations are listed in Table 1, along with the station identification number, total number of temperature observations, and the percent of missing data.

TABLE 1. Observing stations used in the analysis, station identification numbers, total number of maximum and minimum temperature observations, and percentage of data missing for each station in the period 1901–1986.

Station name	Station ID number	Total number of observations	Percent missing data
Baltimore, Maryland	0470	62 816	0.010%
Clarinda, Iowa	1533	62 395	0.680%
Crookston, Minnesota	1891	62 516	0.487%
Fort Collins, Colorado	3005	62 534	0.458%
Hartington, Nebraska	3630	62 074	1.191%
Helena, Montana	4055	62 748	0.118%
Independence, Kansas	3954	62 677	0.231%
Jacksonville, Illinois	4442	62 694	0.204%
Kaufman, Texas	4705	62 315	0.807%
Kingfisher, Oklahoma	4861	61 359	2.329%
Little Rock, Arkansas	4248	62 818	0.006%
Marion, Indiana	5337	62 690	0.210%
Napoleon, North Dakota	6255	62 273	0.874%
New York City, New York	5801	62 777	0.072%
Olga, Washington	6096	62 521	0.479%
Saginaw, Michigan	7227	62 399	0.673%
Salisbury, North Carolina	7615	61 439	2.201%
Tucson, Arizona	8815	62 288	0.850%
University Park, Pennsylvania	8449	62 774	0.076%
Wooster, Ohio	9312	62 685	0.218%

3. Results

We quantify the digit biases by first computing the observed proportions of temperatures ending in each digit and then by computing standard statistical measures of significance. We estimate the effect of the biases on the mean and standard deviation of various temperature distributions and also consider the effect of the biases and the 1950 observational rule revision on temperature threshold statistics.

We use $p(i)$, for $i = 0, 1, \cdots 9$, to represent the observed proportion of temperatures ending in the digit i. As a basis for comparison, we assume that an observed temperature is equally likely to end in any of the ten digits. Although this assumption is not necessarily true, for temperature distributions as large ($N > 61\,000$) and with as wide a range as those considered here it is very nearly true, especially if the distribution is close to being symmetric [this can be demonstrated by assuming that temperature is normally distributed and computing the $p(i)$ for various realistic values of the mean and standard deviation of the distribution].

a. Distribution of temperatures by final digit

The presence of significant biases in the daily maximum and minimum temperature data is qualitatively evident in histograms. As an example, we show in Fig. 1a,b histograms for Napoleon that vividly demonstrate the biases, particularly the bias toward multiples of ten (which are darkened). Although in some cases more pronounced relative maxima occur at temperatures ending in digits other than "0," the bias toward multiples of ten is the most statistically significant and also the most consistent throughout the datasets (as we will show).

This consistency is apparent in Table 2, which shows the $p(i)$, $i = 0, 1, \cdots 9$ for each station, for the period 1901–1986. Both Fig. 1 and Table 2 show that in the Napoleon data, temperatures ending in "5" are also highly overrepresented. The nature and magnitude of these secondary biases vary from station to station: for example, the bias in Baltimore and New York City is toward even numbers only, while in Crookston and Independence the secondary biases are toward temperatures ending in "2" and "5."

An additional feature of these biases deserves note. As an example, consider $p(9) = 0.0796$, $p(0) = 0.1378$, and $p(1) = 0.0909$ for Napoleon. Clearly, temperatures ending in "0" are highly favored: a reasonable assumption is that the "excess" in $p(0)$ should be accounted for by a "deficit" in $p(9)$ and $p(1)$. However, the total proportion by which temperatures ending in "9" and "1" are underrepresented (0.0295) is less than the proportion by which temperatures ending in "0" are overrepresented (0.0378). This problem also appears in Napoleon's data with $p(4)$, $p(5)$, and $p(6)$, as well as in the data for several other locations, including Clarinda, Kingfisher, and Olga.

A similar "accounting" problem occurs when more than two consecutive digits are underrepresented or overrepresented. For example, in Kaufman, Kingfisher, and Little Rock we find $p(i) < 0.1$ for four consecutive digits (6, 7, 8, 9), while in Fort Collins and Wooster, $p(i) > 0.1$ for three consecutive digits (4, 5, 6). If these apparent inconsistencies are bias related, then it is likely that not only are temperatures $T \pm 1$ being mistakenly observed as T but in some instances so are temperatures $T \pm 2$ and possibly $T \pm 3$.

Of course, these "accounting" problems may simply be inherent in the climatology of the station, and unrelated to biasing. This explanation is plausible for datasets with a small number N of observations but unlikely for large datasets. The experiments mentioned earlier, performed using a normal distribution and $N = 61\,000$, suggest that this effect is at least an order of magnitude too small to explain the inconsistencies seen in the observed proportions in Table 2.

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC

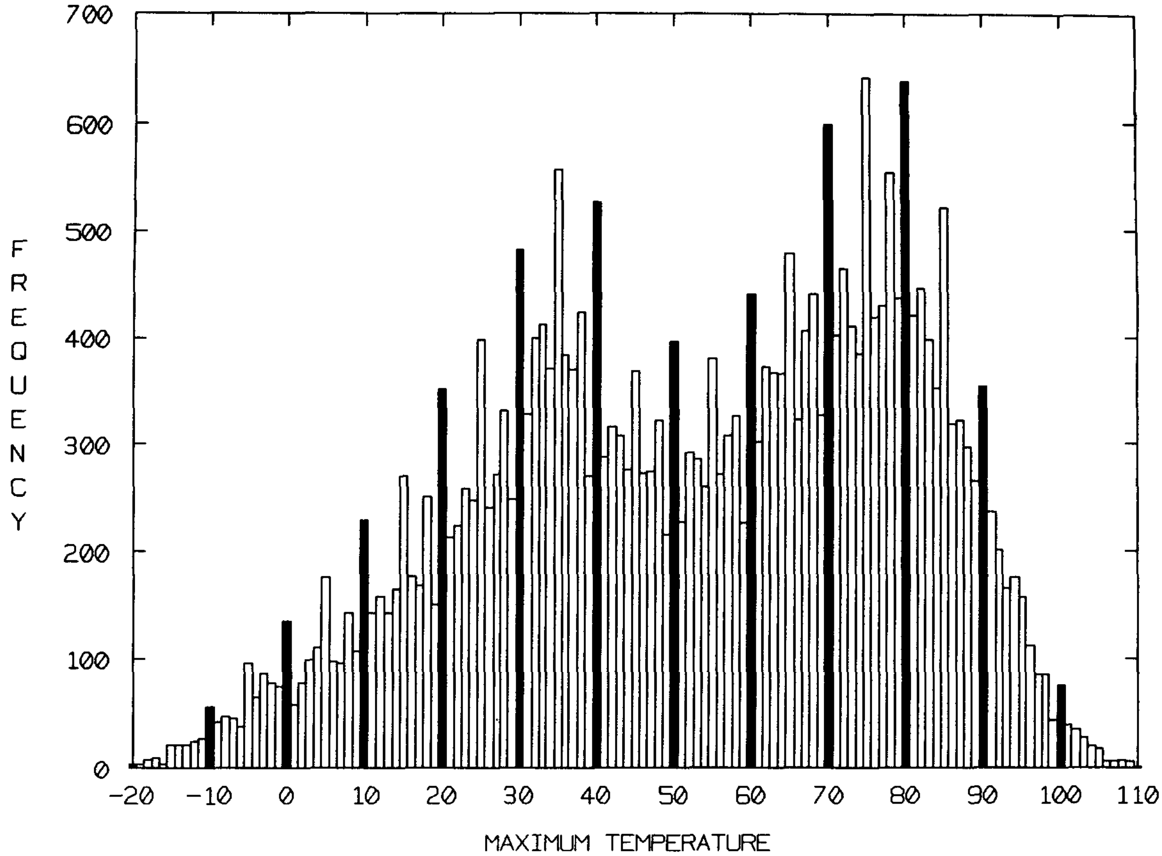


FIG. 1. Histograms of (a) maximum temperature data, and (b) minimum temperature data, from Napoleon, North Dakota.

We next consider the 1 248 792 observations from all 20 stations as one dataset, and group the observations based on the last digit of the daily maximum or minimum temperature. In Fig. 2 we show the $p(i)$ for each digit $i = 0, 1, \cdots 9$ for the periods 1901–1949 (black bar on left), 1950–1986 (white bar in middle), and 1901–1986 (black bar on right). In all periods, temperatures ending in "0", "5", "2," and "8" are clearly overrepresented. To assess the statistical significance of these biases, we use the total number of observations for each station, found in Table 1, and the numerical proportions used to construct Fig. 2, to compute the test statistic χ^2 for multinomial experiments, defined as

$$\chi^2 = \sum_{i=0}^{9} \frac{[O(i) - E(i)]^2}{E(i)}, \qquad (1)$$

where $O(i)$ and $E(i)$ are the observed and expected frequencies of temperatures ending in the digit i; each $E(i)$ is assumed to be one-tenth of the total number of observations in the period of interest (Panofsky and Brier 1968). We hypothesize that the $O(i)$ are not significantly different than the $E(i)$. The nominal critical value of χ^2 for accepting this hypothesis at the 0.1% significance level is 27.9. Applying Eq. (1) separately to the periods 1901–1949, 1950–1986, and 1901–1986, we find values of χ^2 of 5391, 2634, and 7870, respectively, thus rejecting the null hypothesis at extraordinarily high levels of significance for all three time periods.

The significance of individual observed proportions with respect to the theoretical probabilities can be evaluated using a z-test statistic (Triola 1986),

$$z = \frac{O(i) - E(i)}{[(1 - p)E(i)]^{1/2}}, \qquad (2)$$

where $O(i)$ and $E(i)$ are as defined earlier and p is the theoretical probability that an observed temperature will end in a particular digit (we assume $p = 0.1$ for all digits). For the digits "0", "2", "5," and "8," Eq. (2) was used to evaluate the null hypothesis that the observed frequency $O(i)$ is not significantly different than the expected frequency $E(i)$, for the periods 1901–1949, 1950–1986, and 1901–1986. For $i = 0, 2, 5, 8$, for all periods, the null hypothesis is rejected at very high levels of significance: the bias toward multiples of ten is most significant in all three periods, followed in order of significance by the biases toward "5", "2," and "8."

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC

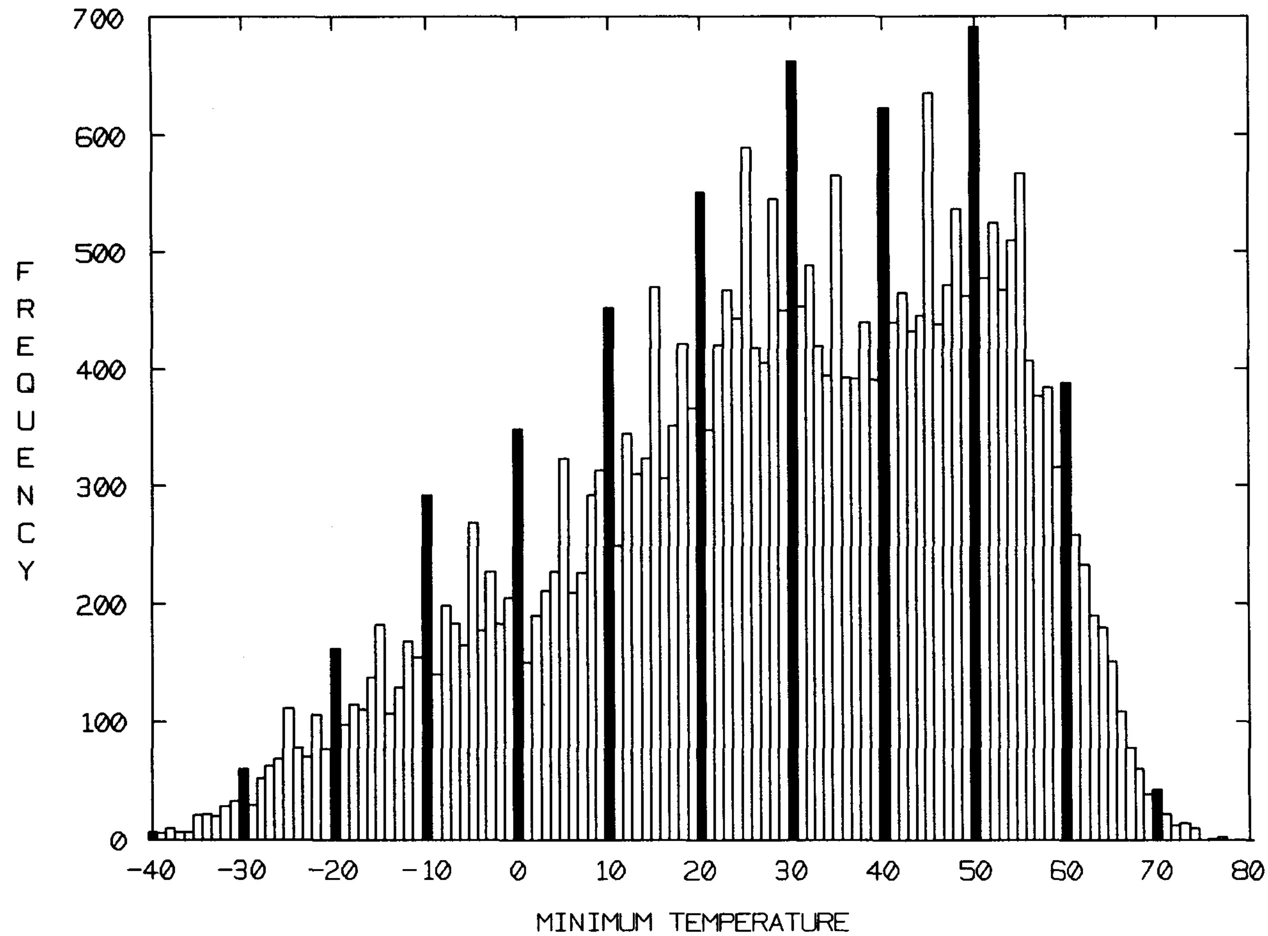


FIG. 1. (*Continued*)

Although the biases in the 1950–1986 data remain highly significant, Fig. 2 suggests that the observational rule revision in 1950 produced some of the desired effect, most notably the decreases in $p(0)$ and $p(8)$ and the increases in $p(9)$ and $p(3)$. Figures 3a,b,c give a year-to-year accounting of $p(\text{even}) = p(0) + p(2) + p(4) + p(6) + p(8)$, $p(0)$, and $p(5)$ from 1901 to 1986; the thicker curves are 5-yr centered means. Although no dramatic improvements in the biases coincide exactly with the 1950 rule revision, in general $p(0)$ and $p(\text{even})$ decrease toward 0.1 and 0.5, respectively, through the period, indicating an improvement in these biases. More quantitatively, we test the hypothesis that the sample means of $p(0)$ and $p(\text{even})$ in the period 1901–1949 (0.1151 and 0.5280, respectively) are not significantly different than the respective sample means in the period 1950–1986 (0.1090 and 0.5078, respectively). Using a z-test statistic, both hypotheses are rejected at less than a 10^{-7} level of significance.

Whereas $p(\text{even})$ and $p(0)$ generally decrease with time, $p(5)$, on average, increases slightly, also an improvement in that odd numbers are being observed more frequently (as intended by the rule revision) but a worsening of the "5" bias in that $p(5) - 0.1$ increases. The hypothesis that the sample mean of $p(5)$ in the period 1901–1949 (0.1047) is not significantly different than the sample mean in the period 1950–1986 (0.1080) is rejected at the 10^{-3} level of significance.

b. Effect of biases on climatological quantities

An important issue is whether these biases can significantly impact derived climatological quantities, such as the mean or standard deviation (or some higher-order moment) of a temperature distribution, or temperature threshold statistics.

1) MEANS AND STANDARD DEVIATIONS

We estimate the effect of these biases on the mean μ and standard deviation σ of a distribution of N temperatures by assuming that temperature is normally distributed. Then, a biased version of the normal dis-

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC

tribution is created using the observed proportions for a particular location (a row of Table 2) as a "guide" to biasing. The biasing procedure is described in detail in the Appendix. The mean and standard deviation of the biased distribution are compared to those of the original distribution, giving an estimate of the change in $\mu(\Delta\mu)$ and the change in $\sigma(\Delta\sigma)$ introduced by the biasing. Figure 4 shows a normal distribution with $N = 36\,000$, $\mu = 50$, and $\sigma = 20$, biased using the observed proportions from Napoleon as a guide. At least qualitatively, Fig. 4 resembles Figs. 1a,b (we could further reproduce the skewness of the observed data, but the results presented below change little).

A set of biasing experiments was performed using each of the 20 rows of observed proportions in Table 2. To simulate the effect of the biasing on monthly, yearly, and century-long datasets, the following values for N were tested: 30, 60, 360, 720, 36 000, and 72 000. For each value of N, values of μ from 50 to 59 were tested every 1 and values of σ from 4.0 to 29.9 were tested every 0.1 (for the small sample sizes $N = 30$ and $N = 60$, only $4.0 \leqslant \sigma \leqslant 9.9$ was tested). Thus, for each location for $N = 30$ and $N = 60$, 600 different (μ, σ) combinations were considered; for the larger values of N, 2600 different (μ, σ) combinations were tested for each location. Because the biasing procedure requires sampling, 50 simulations were performed for each location for each combination of N, μ, and σ to enhance the statistical significance of the results.

The results of these biasing experiments are summarized in Table 3, which shows for each location, for $N = 60$, $N = 720$, and $N = 72\,000$, the average magnitude of the change and the maximum change (both in °F) in the mean and the standard deviation of the sample found in all experiments performed for that location for that value of N. Not surprisingly, because small sample means and standard deviations are more variable than their large sample counterparts, the largest changes in μ and σ usually occur in the smallest datasets. Also, the changes in μ and σ are generally larger at locations such as Napoleon, Hartington, and Baltimore, where the biases are most pronounced.

For $N = 60$ and $N = 720$, the average magnitude of $\Delta\mu$ is in the range 0.07°F–0.26°F; the maximum change is on the order of ±1°F. The average magnitude of $\Delta\sigma$ for $N = 60$ and $N = 720$ is in the range 0.05°F–0.15°F; the maximum change is approximately ±0.75°F. For $N = 72\,000$, the effect of the biasing is substantially less, with the average magnitude of $\Delta\mu$ and $\Delta\sigma$ in the range 0.008–0.019°F, with maximum change on the order of 0.1°F.

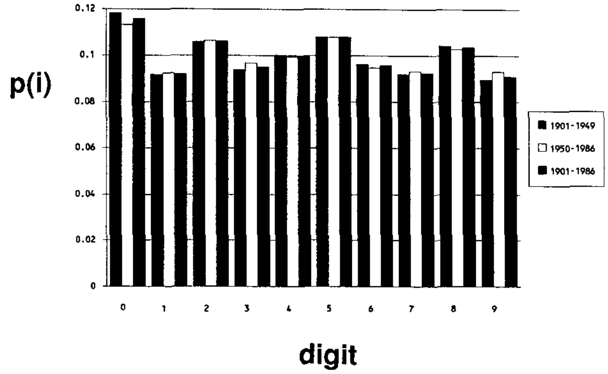
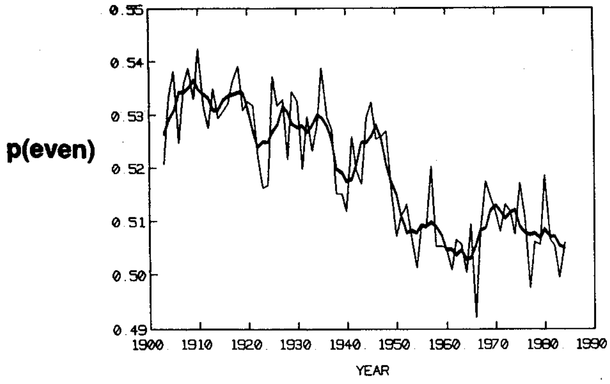


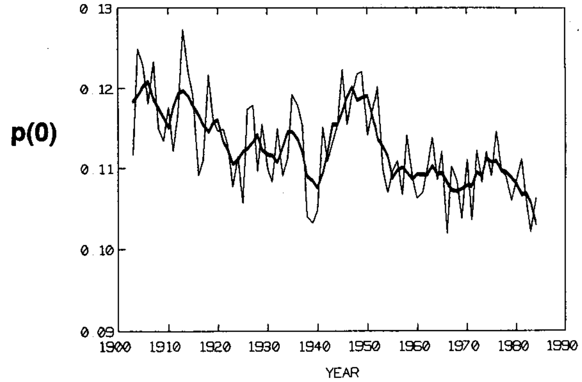
FIG. 2. Observed proportions of temperatures ending in the digits 0–9, for the periods 1901–1949, 1950–1986, and 1901–1986.

TABLE 2. Observed proportions of temperatures ending in the digits 0–9 for each location for the period 1901–1986.

	Final digit of temperature									
Station	0	1	2	3	4	5	6	7	8	9
Baltimore	0.1193	0.0814	0.1148	0.0887	0.1119	0.0887	0.1102	0.0858	0.1142	0.085
Clarinda	0.1232	0.0919	0.1032	0.0934	0.095	0.1172	0.0926	0.0914	0.1038	0.0883
Crookston	0.1301	0.0889	0.1096	0.0992	0.0931	0.1163	0.0873	0.0918	0.1003	0.0835
Fort Collins	0.1053	0.0965	0.1036	0.0963	0.1028	0.1024	0.1021	0.0967	0.0992	0.0951
Hartington	0.1182	0.0873	0.1197	0.0883	0.1026	0.1114	0.0867	0.0824	0.1239	0.0795
Helena	0.1097	0.0958	0.1052	0.0975	0.1023	0.1027	0.0984	0.0948	0.1024	0.0912
Independence	0.1136	0.0913	0.1058	0.0974	0.0968	0.1103	0.0948	0.096	0.1002	0.0938
Jacksonville	0.1076	0.0954	0.101	0.0991	0.1001	0.1066	0.0945	0.095	0.1023	0.0984
Kaufman	0.1211	0.0874	0.1162	0.0962	0.0962	0.123	0.0844	0.0902	0.0961	0.0892
Kingfisher	0.1151	0.0917	0.1025	0.0948	0.0962	0.1133	0.0936	0.0965	0.0992	0.0971
Little Rock	0.1091	0.0968	0.1095	0.0999	0.1054	0.1026	0.0975	0.0909	0.0983	0.09
Marion	0.1111	0.0942	0.1059	0.099	0.0974	0.1101	0.0962	0.0899	0.1026	0.0936
Napoleon	0.1378	0.0909	0.1008	0.0956	0.0923	0.1282	0.0857	0.0878	0.1013	0.0796
New York City	0.1084	0.0954	0.1032	0.0954	0.1033	0.0984	0.1026	0.0961	0.1035	0.0937
Olga	0.129	0.0875	0.102	0.0879	0.0921	0.1192	0.088	0.0929	0.1067	0.0947
Saginaw	0.1117	0.0946	0.1033	0.0986	0.1011	0.102	0.099	0.0964	0.1021	0.0912
Salisbury	0.1212	0.0892	0.1032	0.0868	0.1044	0.0997	0.1011	0.0912	0.1099	0.0933
Tucson	0.1106	0.0923	0.1072	0.0947	0.1023	0.0996	0.0999	0.0929	0.1061	0.0944
University Park	0.1073	0.0936	0.1048	0.0979	0.1009	0.1056	0.0976	0.0962	0.1004	0.0957
Wooster	0.1141	0.0944	0.1042	0.0923	0.1023	0.1057	0.1011	0.0923	0.1019	0.0917

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC





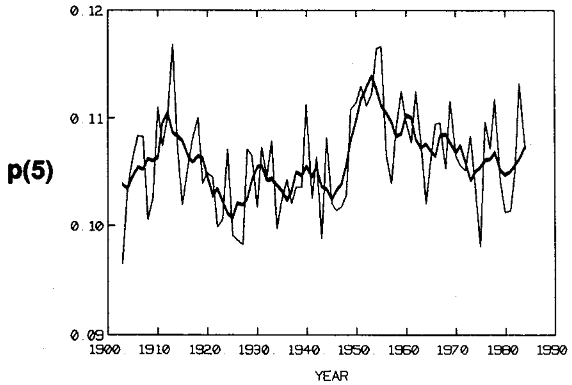


FIG. 3. Observed proportions of temperatures ending in (a) even digits, (b) "0," and (c) "5," for each year during the period 1901–1986. The thicker curves are 5-yr centered means.

We note that the averages of $\Delta\mu$ and $\Delta\sigma$ (as opposed to average magnitudes) were very close to zero for almost all locations. Thus, these simulations suggest that although the overall effect of these biases when averaged over many months and many locations is likely inconsequential in an absolute sense, it is possible for these biases to have a nonnegligible effect on means and standard deviations for individual monthly and yearly datasets at some locations.

2) THRESHOLDS

Also of interest to the climate community (and some in the engineering and agricultural communities) is whether temperature threshold statistics, such as the number of days the daily maximum or minimum temperature reaches a particular value, could be affected by these digit biases and by the 1950 revision in observational rules. As examples, we consider the commonly used thresholds 90°F for maximum temperatures and 32°F for minimum temperatures, although these arguments can easily be adapted to threshold values ending in digits other than "0" and "2."

Because all stations show a bias toward temperatures ending in "0," it is likely that some daily maximum temperatures that should have been observed as 89°F were instead observed as 90°F; thus, the *actual* number of days on which the daily maximum reaches 90°F at a particular station is likely *less* than the *observed* number. This effect would be most noticeable at locations where daily maximum temperatures near 90°F are common; for example, at Little Rock, 90°F is observed as the daily maximum 10.6 days yr^{-1}, on average, but 89°F and 91°F are observed as the daily maximum 8.4 days yr^{-1} and 8.8 days yr^{-1}, respectively (a similar discontinuity occurs at Hartington, Independence, and Crookston, when 80°F is used as the threshold value).

Similarly, all stations show a bias toward temperatures ending in "2," suggesting that some daily minimum temperatures that should be observed as 33°F are instead observed as 32°F; thus, the *actual* number of days on which the minimum temperature reaches 32°F is likely *less* than the *observed* number (the argument regarding the 32°F threshold is complicated by the physical reasoning that may tend to favor 32°F in the data).

Perhaps more importantly, the observational rule revision in 1950 may complicate the interpretation of changes in temperature threshold statistics over time. For example, after the rule revision, the biases toward temperatures ending in "0" and "2" lessened at most locations; thus, misobservations such as 89°F being observed as 90° or 33°F being observed as 32°F should occur less frequently in the 1950–1986 data, artificially *decreasing* the number of days on which the 90°F and

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC

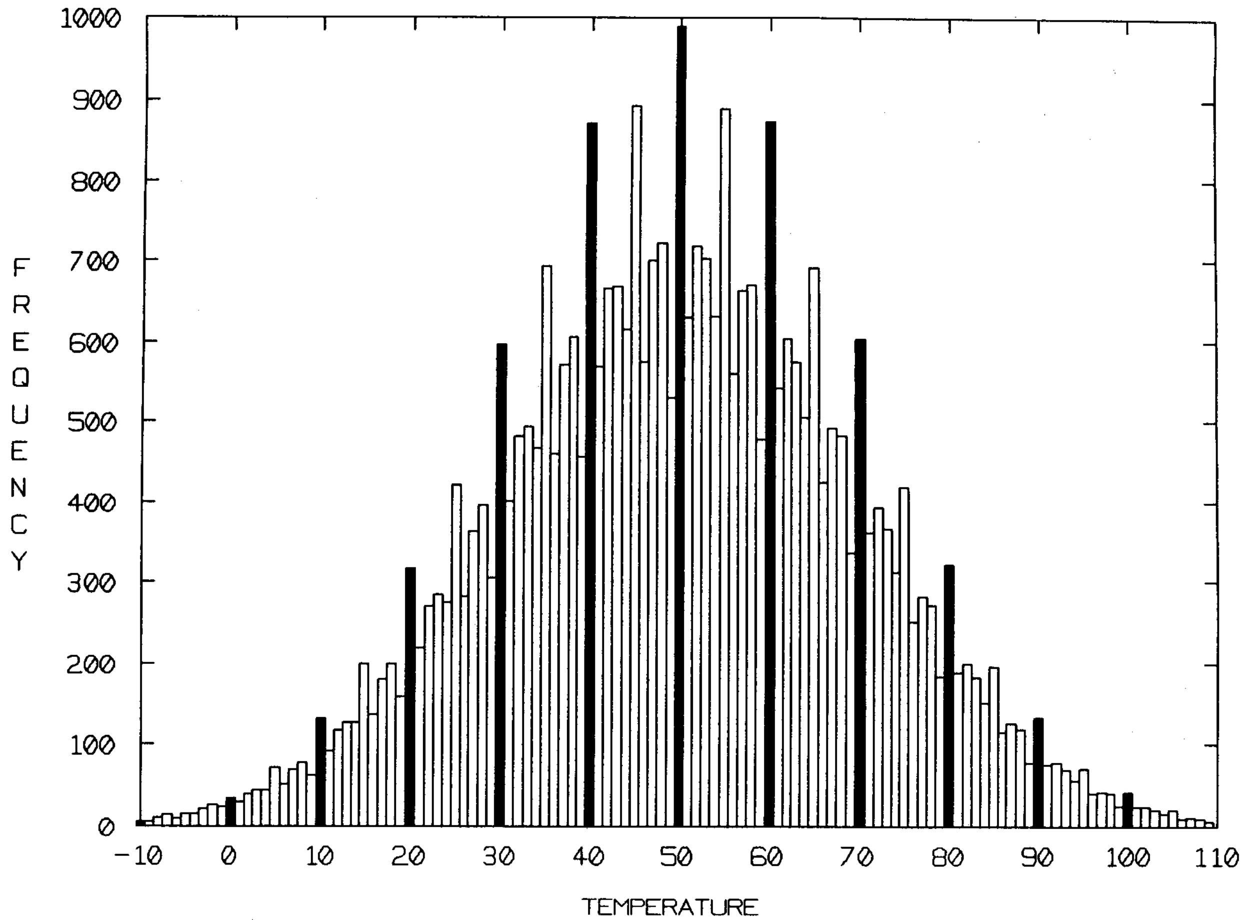


FIG. 4. A normal distribution of 36 000 observations, with mean $\mu = 50$ and standard deviation $\sigma = 20$, biased using the $p(i)$, $i = 0, 1, \ldots, 9$, for Napoleon, North Dakota, as a guide to biasing.

32°F thresholds were reached in the 1950–1986 period. Similar arguments can also be applied to threshold values ending in other digits, depending on the biases at a particular station.

For example, Crookston, Hartington, Marion, Napoleon, Salisbury, and Tucson do show a marked increase (decrease) in the number of days per year on which 89°F (90°F) was observed as the daily maximum temperature in the 1950–1986 period. Indeed, the averages of $p(9)$ and $p(0)$ for these stations for the period 1901–1949 are 0.0793 and 0.1267, respectively, while these averages for the period 1950–1986 are 0.0934 and 0.1103. Similar examples abound in the data for other threshold values. Of course, other factors may contribute to these changes, but it is at least possible that the observational rule revision is partly responsible.

4. Conclusions

Although results for only 20 locations are presented, these analyses demonstrate that highly significant biases exist in manual observations of daily maximum and minimum temperature. We emphasize that the stations used in this analysis were chosen based on the observational criteria presented in section 2, and not because these data were the best available to demonstrate the bias.

We note that in other countries where temperature reporting procedures differ and where temperature is reported in degrees Celsius, the digit biases in the daily maximum and minimum temperature data may be different. In addition, in countries such as Canada where the units used to report temperatures have been changed (from degrees Fahrenheit to degrees Celsius) and the digital archive has been converted (thus changing the resolution of the temperature data), the digit biasing problem may be further complicated.

The most statistically significant and most consistent bias among stations is toward temperatures that are multiples of ten. Secondary biases vary from station to station, although temperatures ending in "5" and "2" are generally favored. A revision of the observational rule in 1950, intended to eliminate biases toward

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC

TABLE 3. Average magnitude (in degrees Fahrenheit) of the change introduced to the mean and standard deviation of a normal distribution, computed by simulating the "digit" bias for each location, for N = 60, 720, 72 000. The maximum, or worst case change, is given in parentheses.

	Average magnitude of change in mean (worst case)			Average magnitude of change in standard deviation (worst case)		
	N = 60	N = 720	N = 72 000	N = 60	N = 720	N = 72 000
Baltimore	0.26 (−1.25)	0.17 (−1.20)	0.018 (0.12)	0.15 (−0.72)	0.11 (0.64)	0.014 (−0.07)
Clarinda	0.17 (−1.02)	0.15 (−0.90)	0.015 (−0.09)	0.10 (−0.56)	0.10 (−0.54)	0.012 (0.06)
Crookston	0.21 (−0.98)	0.16 (1.02)	0.017 (0.12)	0.12 (−0.54)	0.10 (0.51)	0.013 (−0.08)
Fort Collins	0.07 (−0.50)	0.07 (0.48)	0.009 (−0.05)	0.05 (−0.42)	0.06 (−0.44)	0.008 (0.06)
Hartington	0.24 (−1.02)	0.19 (−1.37)	0.019 (−0.12)	0.15 (−0.70)	0.12 (−0.73)	0.015 (0.09)
Helena	0.14 (−0.82)	0.10 (−0.65)	0.010 (−0.07)	0.08 (−0.43)	0.06 (−0.43)	0.009 (−0.06)
Independence	0.15 (0.73)	0.11 (0.74)	0.012 (0.09)	0.10 (−0.56)	0.08 (0.55)	0.010 (0.06)
Jacksonville	0.08 (−0.50)	0.09 (−0.49)	0.010 (0.06)	0.05 (−0.42)	0.05 (−0.42)	0.008 (0.05)
Kaufman	0.23 (−0.93)	0.17 (0.65)	0.018 (−0.15)	0.15 (0.62)	0.11 (0.65)	0.014 (0.10)
Kingfisher	0.16 (0.73)	0.12 (−0.73)	0.013 (0.08)	0.10 (−0.65)	0.08 (−0.52)	0.010 (−0.06)
Little Rock	0.19 (0.77)	0.11 (0.70)	0.012 (−0.08)	0.10 (−0.53)	0.07 (−0.55)	0.010 (0.08)
Marion	0.15 (−0.70)	0.12 (0.70)	0.012 (−0.08)	0.09 (−0.52)	0.08 (0.48)	0.010 (0.07)
Napoleon	0.23 (−1.25)	0.18 (−1.20)	0.019 (−0.12)	0.15 (−0.67)	0.12 (−0.77)	0.015 (0.08)
New York City	0.08 (−0.50)	0.09 (0.59)	0.010 (0.08)	0.05 (−0.53)	0.07 (0.50)	0.009 (−0.05)
Olga	0.22 (0.92)	0.17 (0.90)	0.018 (0.11)	0.12 (−0.62)	0.10 (0.72)	0.014 (−0.08)
Saginaw	0.14 (0.68)	0.10 (0.68)	0.010 (0.08)	0.08 (0.43)	0.06 (−0.40)	0.009 (0.06)
Salisbury	0.20 (0.78)	0.13 (0.86)	0.015 (0.11)	0.14 (0.56)	0.09 (0.50)	0.012 (−0.07)
Tucson	0.10 (−0.50)	0.11 (0.66)	0.011 (0.08)	0.05 (−0.42)	0.07 (−0.48)	0.009 (0.06)
University Park	0.08 (−0.50)	0.10 (−0.58)	0.010 (0.06)	0.07 (−0.53)	0.06 (−0.41)	0.008 (0.06)
Wooster	0.11 (0.60)	0.12 (−0.60)	0.012 (0.08)	0.08 (−0.42)	0.08 (0.45)	0.010 (0.06)

even numbers, did decrease the frequency of observation of even-numbered temperatures, particularly multiples of ten, but the bias toward temperatures ending in "5" worsened. Despite the improvement, the biases toward temperatures ending in "0", "2", "5," and "8" remained highly statistically significant even after the rule revision.

A reasonable assumption is that an excess in observations of a temperature T may be accounted for by a deficit in observations of temperatures $T \pm 1$. However, our results suggest that the overrepresentation of a particular temperature T may also involve underrepresentation of temperatures $T \pm 2$ and even possibly $T \pm 3$.

Experiments performed by biasing a normal distribution suggest that, in the worst case, these biases can alter the mean and standard deviation of a distribution by several tenths of a degree for small datasets, such as a set of monthly or yearly observations, but the effect on large datasets over long time periods is most likely negligible. The biases and the observational rule revision in 1950, however, may have a nonnegligible effect on temperature threshold statistics. More qualitatively, the presence of these biases in the data and the fact that the strength of the biases has varied with time makes one wonder in how many other as yet undiscovered or unquantified ways the record of temperature is contaminated.

Acknowledgments. I wish to thank George Young and Fred Gadomski for numerous suggestions and helpful discussions. The recommendations of two anonymous reviewers contributed significantly to the final version of this paper. This research was supported by a Research Development Grant from The Pennsylvania State University.

APPENDIX

Description of the Biasing Procedure

Given a normal distribution of N observations with mean μ and standard deviation σ, we use the following procedure to bias the distribution, using a station's $p(i)$, $i = 0, 1, \cdots 9$, (a row of Table 2) as a guide to biasing.

For each $i = 0, 1, 2, \cdots 9$, the following three steps were performed:

1) Simulate an observation T from the specified Gaussian distribution.
2) If the last digit of T is the digit i, include the observation T in the biased distribution. Otherwise, discard T.
3) Repeat steps 1 and 2 until $N * p(i)$ observations ending in the digit i have been accumulated.

This procedure generates a biased distribution having the same $p(i)$ as the station used to guide the biasing. Comparing the mean and standard deviation of the biased distribution with μ and σ gives estimates $\Delta\mu$ and $\Delta\sigma$ of the effect of the biasing.

Because this procedure requires sampling of the original normal distribution, some of $\Delta\mu$ and $\Delta\sigma$ is undoubtedly related to the manner in which the dis-

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC

tribution is sampled. The contributions $\Delta\mu_s$ and $\Delta\sigma_s$ of sampling to $\Delta\mu$ and $\Delta\sigma$ were estimated in the following way: for each biasing simulation, a parallel simulation was also performed using the $p(i)$ of the original Gaussian distribution as the guide to biasing. Using these $p(i)$, "perfect" sampling would reproduce the original normal distribution, giving $\Delta\mu = 0$ and $\Delta\sigma = 0$. Thus, any changes in μ and σ generated by the parallel experiment were assumed to be due to sampling, and $\Delta\mu$ and $\Delta\sigma$ were reduced by the amounts $\Delta\mu_s$ and $\Delta\sigma_s$. The results reported in Table 3 are these "corrected" values.

REFERENCES

Court, A., 1951: Temperature frequencies in the United States. *J. Meteor.*, **8**, 367–380.

Panofsky, H., and G. W. Brier, 1968: *Some Applications of Statistics to Meteorology.* The Pennsylvania State University Press, 224 pp.

Quayle, R. G., 1971: Hints to the observer: Wind direction. *Mar. Wea. Log.* **March,** p. 74.

Reverdin, G., 1988. Comments on wind reports in marine decks—equatorial Atlantic Ocean (1982–1984). *Tropical Ocean-Atmos. Newslett.,* **27,** 22–24.

Triola, M., 1986: *Elementary Statistics.* 3d ed., Benjamin/Cummings, 581 pp.

U.S. Department of Commerce, 1941: *Instructions for Airway Meteorological Service, Circular N.* United States Department of Commerce, Weather Bureau. 5th ed., 201 pp.

——, 1988: *Federal Meteorological Handbook No. 1: Surface Observations.* 3d ed. Office of the Federal Coordinator for Meteorological Services and Supporting Research, U.S. Department of Commerce/National Oceanic and Atmospheric Administration, 226 pp.

Unauthenticated | Downloaded 12/09/23 02:31 AM UTC