UNIVERSITY OF MINNESOTA
ST. ANTHONY FALLS LABORATORY
Engineering, Environmental and Geophysical Fluid Dynamics

Project Report No. 442

STATISTICAL ESTIMATION OF AN UPPERBOUND ON WEEKLY STREAM TEMPERATURES

by

Jonathan A. Othmer, Omid Mohseni,
And Heinz G. Stefan

October 2001
Minneapolis, Minnesota

UNIVERSITY OF MINNESOTA
ST. ANTHONY FALLS LABORATORY
Engineering, Environmental and Geophysical Fluid Dynamics

Project Report No. 442

STATISTICAL ESTIMATION OF AN UPPER BOUND ON WEEKLY STREAM TEMPERATURES

by

Jonathan A. Othmer, Omid Mohseni,
and Heinz G. Stefan

October 2001
Minneapolis, Minnesota

Prepared for: Research Record
Last Revised: October 2001
Disk Locators: Stefan Zip Disk #24; Stefan Compendium.pdf

Abstract

There is interest in how global climate change will affect the earth'ssy stems. One of the systems that will certainly be affected is surface water, in particular, streams. This report addresses, in three parts, change in stream temperatures as a result of global climate change. Work has already been done on the effects of climate change on stream temperatures (Erickson and Stefan, 2000; Mohseni and Stefan, 1999; Mohseni et al., 2001; Mohseni and Stefan, 1998, 1999; Sinokrot and Stefan, 1993; Stefan et al., 2001).

The first part of this report is an extension of work done by Erickson et al. (1998). A method of Hershfield (1961) is employed where a maximum series of data is analyzed and a standard enveloping deviate, K, is determined. We analyze stream temperature data from 993 USGS stream gauging stations and compute K-values for each station. We looked at possible trends and patterns in the K-values, suggest a "reasonable" K, and discuss the significance and applicability of K.

The second part of the report continues the analysis of 993 stream gauging stations. We analyzed the stations to look for any trends in the data during the periods 1980 – 1990 and 1970 – 1990. We found no significant trends.

The final part of this report is an extension of work done by Mohseni et al. (1997) to fit a curve to air temperature/stream temperature data. We use a transformation to linearize the equation fit by Mohseni et al.. Then we use our linearized equation to estimate α, the maximum stream temperature, again using the data from 993 stream gauging stations.

Acknowledgments

The work herein was supported by the Minnesota Supercomputing Institute for Digital Simulation and Advanced Computation at the University of Minnesota as part of a summer research program for undergraduates.

Contents

List of Tables

List of Figures

1 Statistical Estimation of an Upper Bound on Weekly Stream Temperatures

1.1 Introduction

It is now widely accepted that global climate change is taking place on the earth (Houghton et al., 1996). One of many effects of climate warming is a rise in air temperatures which in turn leads to an increase in stream temperatures (Mohseni et al., 1999). This increase will have a profound effect on the availability and distribution of habitats for fish.

Observations show that as air temperatures increase, stream temperatures begin to level off. This is justified theoretically, because as air temperature rises, heat loss through evaporation begins to counter the effects of heating by convection and radiation from the warmer air. We also see a leveling off at the cold end of the water temperature range. As air temperatures drop below $0°C$ water temperatures level out near $0°C$. This gives an S-shaped curve with possible asymptotes at 0 and some α.

Deterministic differential equation models for the heat-transfer process in streams can be developed but they are difficult to apply to large numbers of streams. They require many parameters as input variables, which must be recalibrated for each stream. This means that predictions can only be made for streams with some past record of data. The calibration coefficients are related to conditions of the past.

To determine if there is actually an upper bound on stream temperatures, we turn to statistical methods. We have a large amount of stream temperature data from stream gauging stations located throughout the United States. We also have air temperature data from nearby weather stations. There are 993 streams for which we have records; the records range in length from 3 to 30 years. They give daily temperature readings which for the purpose of this study have been averaged to weekly values. This length is suitable for fish habitat studies (Eaton and Scheller, 1996; Stefan et al., 2001). Our data set contains 507,072 weekly stream readings.

Mohseni and Stefan (1998) did a curve fit for each of the stream gauging stations. The curve:

$$T_s = \mu + \frac{\alpha - \mu}{1 + e^{\gamma(\beta - T_a)}} \tag{1}$$

was fit to the data for each stream. $\alpha = T_s^*$ represents the maximum stream temperature; β is the air temperature at the inflection point of the curve; T_a is the air temperature; $\mu = T_s^{**}$ is the estimated minimum stream temperature; and γ is a function of the slope at the point of inflection. For many of the streams the temperatures were never high enough for the data to reach the leveling-off part of the curve. In these cases many α's fit equally well and determining the "correct" one was impossible.

We are left with two questions:

i. Is there a single maximum value, α, for all streams, or is the maximum stream temperature stream-dependent?

ii. For either a single stream or all streams, what is the absolute maximum, T_s^*, or fuzzy maximum, T_a^* such that $\frac{\partial T_s}{\partial T_a} < \epsilon$ for $T_a > T_a^*$?

1.1.1 Heat Transfer in Streams

The process of heat transfer between streams and its surroundings is complex and involves many components. Sinokrot and Stefan (1993) developed a deterministic model for heat exchange in streams. We will not discuss it in depth but we will mention the processes that are involved. The first is heat exchange between the water and the atmosphere. This is comprised of short-wave solar radiation, net long-wave radiation, evaporative heat transfer (Mohseni and Stefan, 1999), and convective heat transfer. The second component is heat transfer between the stream bed and the water. A third component is point and non-point heat sources such as ground-water, industrial cooling water releases, and hypolimnetic reservoir releases. Stream temperatures are also affected by upstream water temperatures, especially near snow-melt regions where temperatures can be close to $0°C$.

1.1.2 Statistical Methods for Stream Data Analysis

Erickson et al. (1998) performed statistical analysis on weekly stream temperature data. He employed a method devised by Hershfield (1961). Hershfield's method is concerned only with analysis of the extreme values. The values for the analysis are a maximum series taken from the data. For this maximum series the mean, $\overline{T}$, and the standard deviation, s, are calculated. When calculating the mean and the standard deviation, the highest observed value is omitted. $\overline{T}$ and s are used to calculate the enveloping standard deviate, K. K represents the number of standard deviations that must be added to the mean to obtain the maximum value. The highest observed value is included in the computation of K. This is equivalent to computing a mean and standard deviation for a stream *before* the maximum is observed. Thus:

$$T_{\text{MAX}} = \overline{T} + sK \tag{2}$$

Hershfield used the largest K to forecast the maximum expected rainfall for a given return period. We will use the same method for determining, $\overline{T}$, s, and K and apply it to stream temperatures. It is important to note that unless specified otherwise, $\overline{T}$ and s denote the mean and standard deviation of the *maximum series*, not the data set as a whole.

1.2 Preliminary Analysis

For our preliminary analysis we examined 19 records whose lengths were 25 years or longer. We began by inspecting the records to find sharp discontinuities that might suggest human interference with the stream. We plotted the temperatures versus time and looked for significant trends over the years and sharp changes from one year to the next. We discarded the records that were not "clean" by visual inspection (see Table 1.1).

To determine whether there was any correlation between record length and the calculated parameters we analyzed 10, 15, 20, 25, and 30 year sliding segments taken from the stream records. For each segment-length we took a partial maximum series with the equivalent of 2 readings per year. That is, for an x-year segment we took the largest $2x$ values regardless of which year they occurred in. This gives us about the top 4% of data. We used a partial maximum series because an annual maximum series with short records ($\leq$ 30 years) would not give enough data points. The partial series gives us twice as many points. For each series we computed $\overline{T}$, s, K, and the coefficient of skewness, C_s.

For the most part, our results showed little correlation (see Table 1.2). Further, the wide distribution of correlation coefficients suggests that instances which show a high correlation are coincidental and do not correspond to an actual correlation. In particular we were unable to find clear correlations between record length, RL, and s and K.

Figure 1.1 shows all of the K values plotted against the record length. There is not an apparent trend although we can see that the variability in K decreases as record length increases. This may, however, be a result of the fact that there are fewer data points for the long records. Appendix A shows similar plots for each of the twelve stations that we considered. One weak correlation that we did find is between the mean temperature of the sample and the standard deviation. Figure 1.2 shows a plot of the data and a trend line. Although there appears to be a trend, the R^2 value is quite small, 0.1616, due to the large amount of scatter.

We decided to look at the distribution of K values for short records to see if we could interpolate values for longer records. However, there was not not enough data in the small sample of records to see clear distributions. So we turned to the data set of 993 streams.

1.3 Analysis of all Stations

For our statistical analysis of the large data set we used the same method employed in our preliminary analysis. In addition, however, we performed an autocorrelation test with $k = 1, 2, 3$ to determine whether the partial series method biased the data towards a particular year. The results were evaluated at a 95% confidence level . Those stations that did not pass the test were discarded. 273 stations, or 27.5% were discarded.

1.3.1 Characterization of the Data

The 720 remaining gauging stations used in our analysis are located across the 48 contiguous states. Figure 1.3 shows the locations of the stations. Record lengths varied from 3 to 30 years.

The maximum observed temperature varied widely from stream to stream. Figure 1.4 shows the distribution of maximum observed temperatures for all stations, and also stations with records 15 years or longer, 20 years or longer, and 25 years or longer. Figure 1.5 shows the cumulative frequencies of the observed maxima for various record lengths. The main difference between the curves occurs in the upper tail. The highest observed temperatures occurred more often in the larger data sets.

1.3.2 Statistical Analysis

Our analysis of the larger data set did not uncover any correlations between variables. As in the preliminary analysis, we computed correlations between different parameters. These correlations are summarized in Table 1.3.

The variation in parameters calculated from the statistical analysis was great. Table 1.4 shows parameters for the stations with the ten highest and five lowest K values. Table 1.5 shows the station information for the extreme stations.

A high K value of 14.09 is quite extraordinary and may suggest the presence of human intervention in the stream. A record length of 10 years is not particularly short; also, a drainage area of 1387mi^2 suggests that the stream is not particularly small. Figure 1.6 shows the record for station 07324200. A spike in temperature near the beginning of the record has a value $10°C$ higher than any other observation in the record. That much variation between a single measurement and its surrounding measurements is unlikely and may suggest a measurement error or extremely low flow. Figure 1.7 shows the stream flow record for station 07324200. The flow is highly irregular which would contribute to the high variability in stream temperatures. For the week with the maximum temperature reading average flow was around 0.003ft^3/s. Also, the record is missing many data points. [The small number of dots for a given year on the diagram indicates the presence of holes in the data.] The remaining water temperature records of high and low-K streams can be found in Appendix B.

Although there seemed to be no relation between record length and K we decided to look further. We partitioned the data according to record length: 0-5, 6-10, . . . , 26-30 years. For each partition we determined the average K value and plotted it at the upper bound of the partition. The results, shown in Figure 1.8, indicate an upward trend in K in relation to Record Length. We performed the same operations with our s values and found a similar result, shown in Figure 1.9. These trends indicate two things. The trend in $\overline{K}$ suggests that the farthest outliers gets farther from the mean as record length increases. The trend in $\overline{s}$ suggests that the number of outliers also increases with record length. Neither of these results is surprising. As record length increases it would be expected to encounter several points that lie far out on the probability distribution.

Thus far, in our discussion, we have not looked at the geographic distribution of streams. However, geographic location would seem to affect the amount of temperature variation in a stream. We plotted the K values on a map of the United States to look for geographic patterns. The results show little pattern. To the contrary, the plot shows that streams with widely varying K-values can be found in a small geographic region. Figure 1.10 shows plots with only the high K-values and only the low K-values. Both high and low values show a wide geographic distribution.

We have looked for trends or patterns in our data. If we want to apply a statistic to every stream, even those which are not in our data set, we need a "universal" K. We could find such a K by taking the largest observed value, or the second or third largest. We could also choose the value at 95% or 99% of the observed distribution. It might also make sense to establish some minimum record length for the data to be considered. Since there are no clear trends that allow us to interpolate from short records the value of a longer record, we may just drop shorter records. Figures 1.11, 1.12, 1.13, and 1.14 show the predicted maximum value for all streams in the data set given various choices of K. In every case a single K was used for all of the streams. The predicted distributions look very similar to the observed distributions. However, the largest three K's from the pool of all streams predict temperatures above 45 and even $50°C$, which seems unlikely given the range observed water temperature data.

1.3.3 Theoretical Predictions

Instead of merely examining the observed distribution of K we can fit a theoretical probability distribution function (PDF) to the calculated K values derived from the observed data. The observed distribution of K's has a lower bound at 0 and is skewed to the right. Thus, we chose to fit an Extreme Value Type-I and a Pearson Type-III Distribution. Parameters were estimated using the method of moments. As before, we applied a variety of cutoffs for record length. A χ^2 goodness-of-fit test was applied to the curve and evaluated at an $\alpha = 0.05$ confidence level. For 25 years or longer there were not enough data points to fit a curve. For the entire set of 720 stations both the Extreme-Value and Pearson distributions failed the χ^2 test. We got the same result applying the fit and test to 207 stations with records 15 years or longer. With the 88 stations whose records were 20 years or longer both distributions passed the χ^2 test; we used the Extreme Value Type-I because it seemed to be a better fit under visual inspection. Figure 1.15 shows the observed values, a theoretical PDF and the theoretical cumulative density function (CDF).

With a theoretical distribution we can arbitrarily choose a desired percentile and determine a K-value from that. Table 1.6 shows some K-values from the fitted EV-I curve. However,bec ause our original data comes from a partial maximum series we cannot translate these probabilities into estimated return times.

1.4 Choosing a K

By choosing a single large K-value to apply to all streams one hopes to find a statistical upper bound on stream temperatures. We have presented several methods for determining K-values but we have not yet offered "the correct" K. Our analysis of the record of station 07324200 that had the largest K suggested that the stream did not meet our desire for a free-flowing stream that had not been interfered with. Thus, we might not wish to use the highest K. We can choose a K from our theoretical distribution but we must remember that because we used a partial maximum series in our analysis we cannot translate probabilities on the distribution into return times. Since we are looking for an upper bound, the highest K at 14.09 is not likely to be exceeded. If we assume that conditions will be similar in the future to those today we can choose a K based on how well the predicted values align with observed values. If we compare Figures 1.4, 1.11, 1.12, 1.13, and 1.14 we find that a K in the range of 6 to 7 gives predicted values that fall in well with the observed maximum values in terms of where the extremes of the distributions lie. Figures 1.16 and 1.17 show predicted maximum stream temperatures with $K = 6.5$. Figure 1.18 shows a map of the United States with contours based on predicted maximum temperatures. Figure 1.19 shows the computed S values plotted on a map of the United States.

1.5 Conclusion

In this study we have avoided the problems of deterministic stream temperature models in favor of statistical computation of an enveloping standard deviate which gives the variability in the maximum series of stream temperatures. We present several methods for going from observed K-values to a "universal" K which would give us a statistical upper bound the amount of variability in maximum stream temperatures.

Our discussion was motivated by an interest in the effects of climate warming. It is only natural to ask what our results can tell us about the effects of climate change on stream temperature. Since our method was statistical we cannot use our results to establish a theoretical upper bound on free-flowing stream temperatures. Also, our method, like statistical methods in general, is predicated on future conditions closely resembling past conditions. In the case of climate warming, we are assuming that this is not the case. We might conjecture that K-values are constant with relation to climate change, but we have no evidence in the data or theoretically to back such a claim up.

Table 1.1. Stations (7 of the 19) with 25-year or longer records that were not considered in the preliminary analysis.

Station	Reasons for Dropping
11407000	sharp drop observed in maximum temperature around 1970
11445500	sharp drop in temperature in the mid 1960s with spike in 1978
14148000	sharp drop in temperature in 1961
14151000	steady decline in maximum stream temperatures
14159500	drop in stream temperatures in 1963
14178000	upward trend in maximum temperatures throughout the years
14238000	sharp jump in temperature in 1920

Table 1.2. Correlation coefficients for various pairings of parameters for each station.

Station	RL, K	RL, s	s, K	$\overline{T}$, s
01127000	0.395747063	0.315480103	-0.043222073	0.828225153
01466500	-0.069353	0.236382696	-0.071455282	-0.479125964
01595200	-0.058379386	0.551500312	-0.192441283	0.039953457
04208000	0.415011152	0.227628584	-0.104527026	0.789897259
05331000	0.300834096	0.350250844	-0.063023207	0.616022373
11187000	0.316436109	0.307151708	-0.052736554	-0.057464901
11447650	0.687412325	0.478308831	0.85737352	0.804651989
12113000	0.144453484	-0.079869834	-0.7812732	-0.511815115
12117500	0.521794803	0.395564306	0.873599449	0.40464099
12179000	0.769068926	-0.048534391	0.363366467	0.131564331
13334300	-0.15526106	0.578918839	-0.724516617	0.556873316
14150000	0.498673175	0.250467836	0.433700642	0.732890167

Table 1.3. Correlation coefficients for various pairings of parameters for 720 stream gauging stations.

Parameters	Correlation Coefficient
RL, $\overline{T}$	-0.0539
RL, K	0.1221
RL, s	0.1016
RL, Max	-0.0091
s, K	-0.05983
$\overline{T}$, s	0.1075
$\overline{T}$, K	-0.0015
$\overline{T}$, Max-$\overline{T}$	0.1050
$\overline{T}$, Max:$\overline{T}$	-0.2070

Table 1.4. Statistical parameters for the stations with the 10 highest and 5 lowest K values.

Site No.	$\overline{T}$(°C)	s(°C)	Max(°C)	K	C_s	Max-$\overline{T}$	Max:$\overline{T}$
Highest K							
07324200	26.95	0.783	37.04	14.09	0.479	11.04	1.409
10237000	20.89	0.597	26.35	9.49	0.195	5.67	1.271
02430000	26.28	0.463	30.50	9.13	1.474	4.22	1.160
03274600	28.72	0.558	33.34	8.28	1.673	4.62	1.161
05437632	20.83	0.286	23.18	8.21	-0.377	2.35	1.112
01192370	22.78	0.511	26.89	8.03	0.123	4.11	1.180
10350405	22.91	0.682	28.00	7.47	0.628	5.09	1.222
05112000	23.66	0.526	27.49	7.29	2.089	3.83	1.162
07053825	18.25	0.754	23.48	6.95	0.882	5.24	1.287
01449360	20.85	1.195	28.93	6.77	1.041	8.09	1.388
Lowest K							
02084558	31.89	0.543	32.53	1.19	-0.099	0.64	1.020
02382500	24.09	0.464	24.60	1.11	-0.341	0.51	1.021
08171000	30.54	0.419	30.99	1.11	-1.604	0.46	1.015
13185000	20.72	0.787	21.58	1.10	-1.001	0.87	1.042
02326512	28.74	0.275	28.99	0.93	-0.653	0.26	1.009

Table 1.5. Information for the extreme records drawn from the pool of 720 stations.

Site No.	Location	RL(yrs)	Lat.(°)	Long.(°)	Drainage(mi^2)
07324200	Washita River near Hammon, OK	10	35.65	99.30	1387
10237000	Beaver River at Adamsville, UT	7	38.25	112.76	303
02430000	Mackeys Creek near Dennis, MS	7	34.52	88.32	66.9
03274600	Miami River at New Baltimore, OH	17	39.26	84.66	3814
05437632	Spr Crk at Rock Vly Coll at Rockford, IL	3	42.29	88.98	2.81
01192370	Porter Brook near Manchester, CT	6	41.76	72.50	2.20
10350405	Truckee R Rt Bk Bl Tracy, NV	10	39.56	119.51	1590
05112000	Roseau River near Caribou, MN	4	48.98	96.46	1420
07053825	Lake Taneycomo at Forsyth, MO	3	36.69	93.12	
01449360	Pohopoco Creek at Kresgeville, PA	20	40.89	75.50	49.90
02084558	Albemarble Canal near Swindell, NC	4	35.63	76.72	68
02382500	Coosawatte River at Carters, GA	4	34.60	84.69	521
08171000	Blanco River at Wimberley, TX	3	29.99	98.08	355
13185000	Boise River near Twin Springs, ID	3	43.65	115.72	830
02326512	Aucilla River near Scanlon, FL	3	30.23	83.91	805

Table 1.6. K-values computed for various percentiles from the theoretical Extreme Value Type-I distribution.

Percentile	99	95	97.5	99	99.9	99.99
K	4.221	4.673	5.117	5.697	7.145	8.591

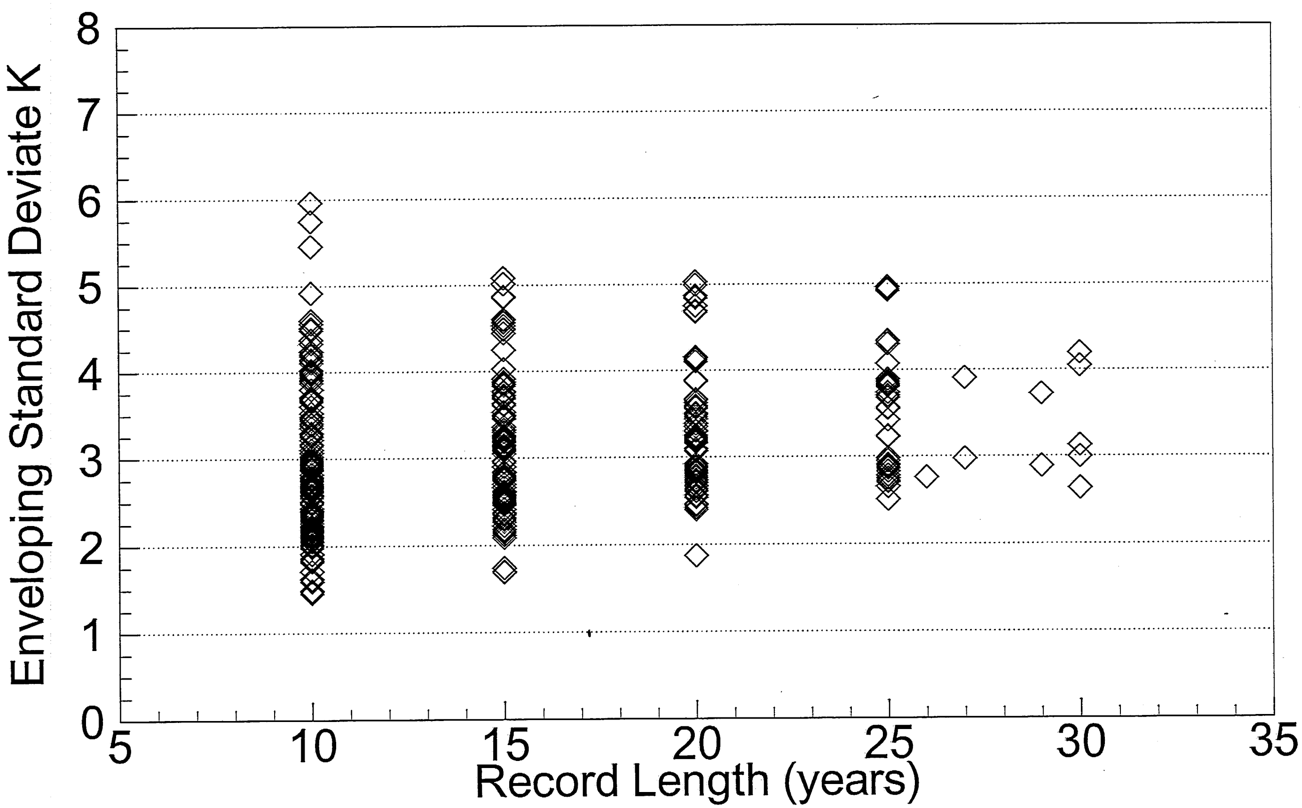


Figure 1.1. A scatter plot showing the enveloping standard deviate, K, versus record length for all 12 stations.

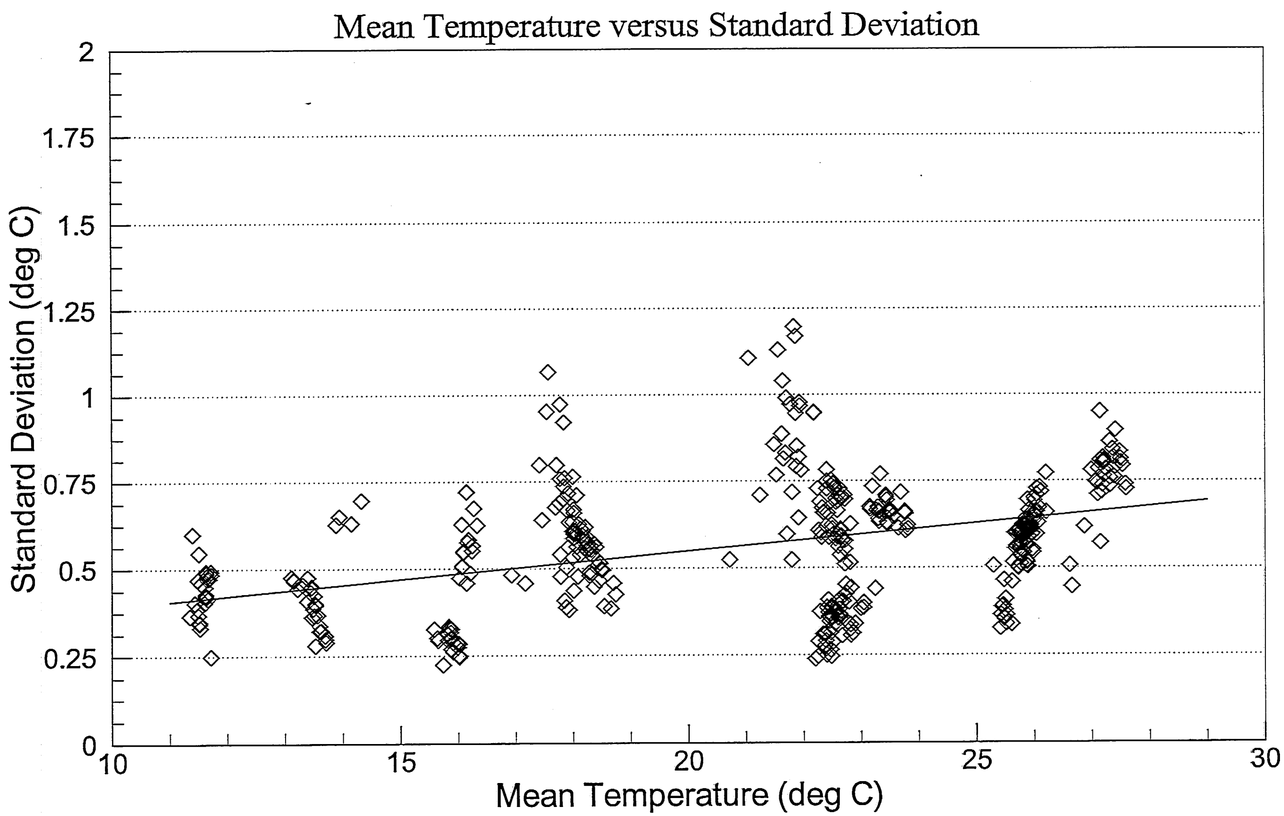


Figure 1.2. The standard deviation of a sample versus the mean of the sample for 12 stations with record length of 25 years or longer. The line plotted is $s = 0.0159\overline{T} + 0.2314$. $R^2 = 0.1616$.

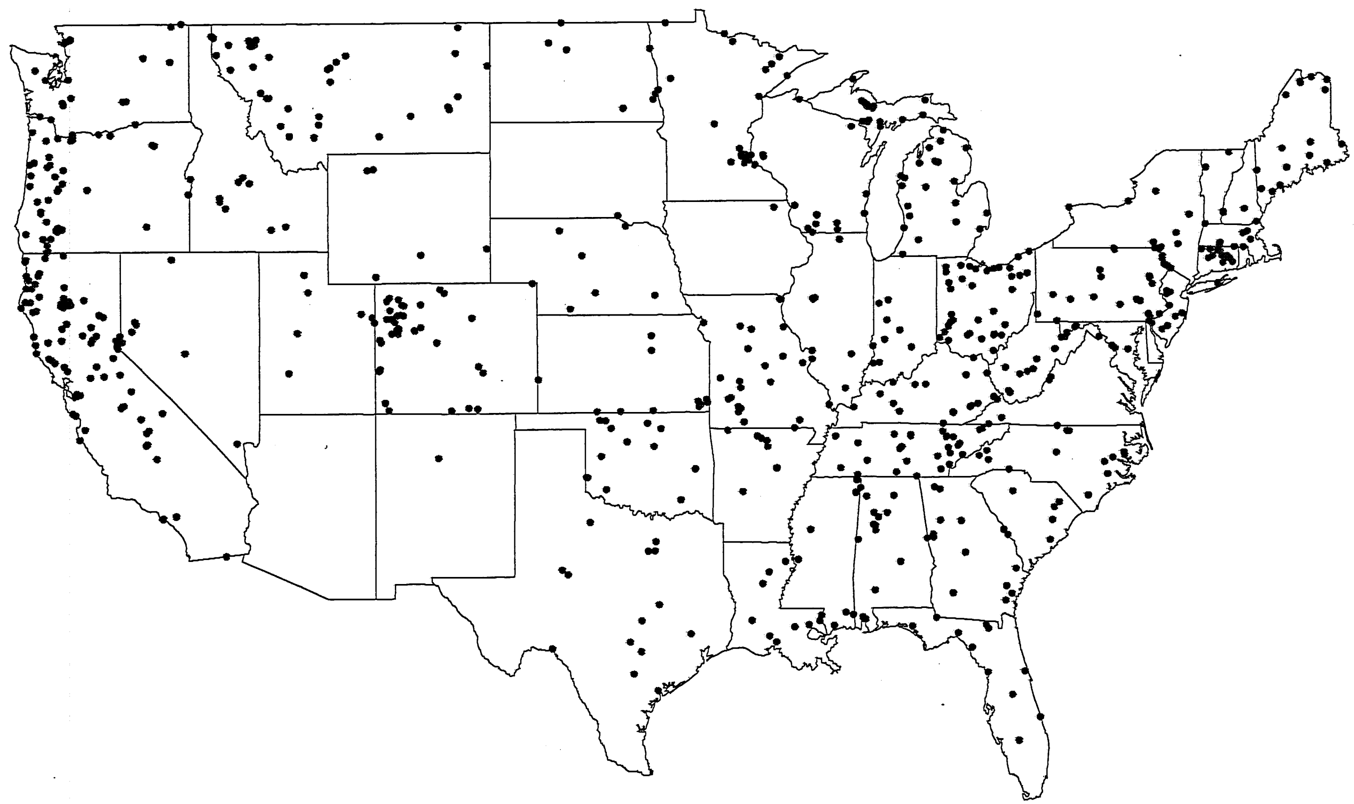


Figure 1.3. The location of the 720 stream gauging stations in the United States.

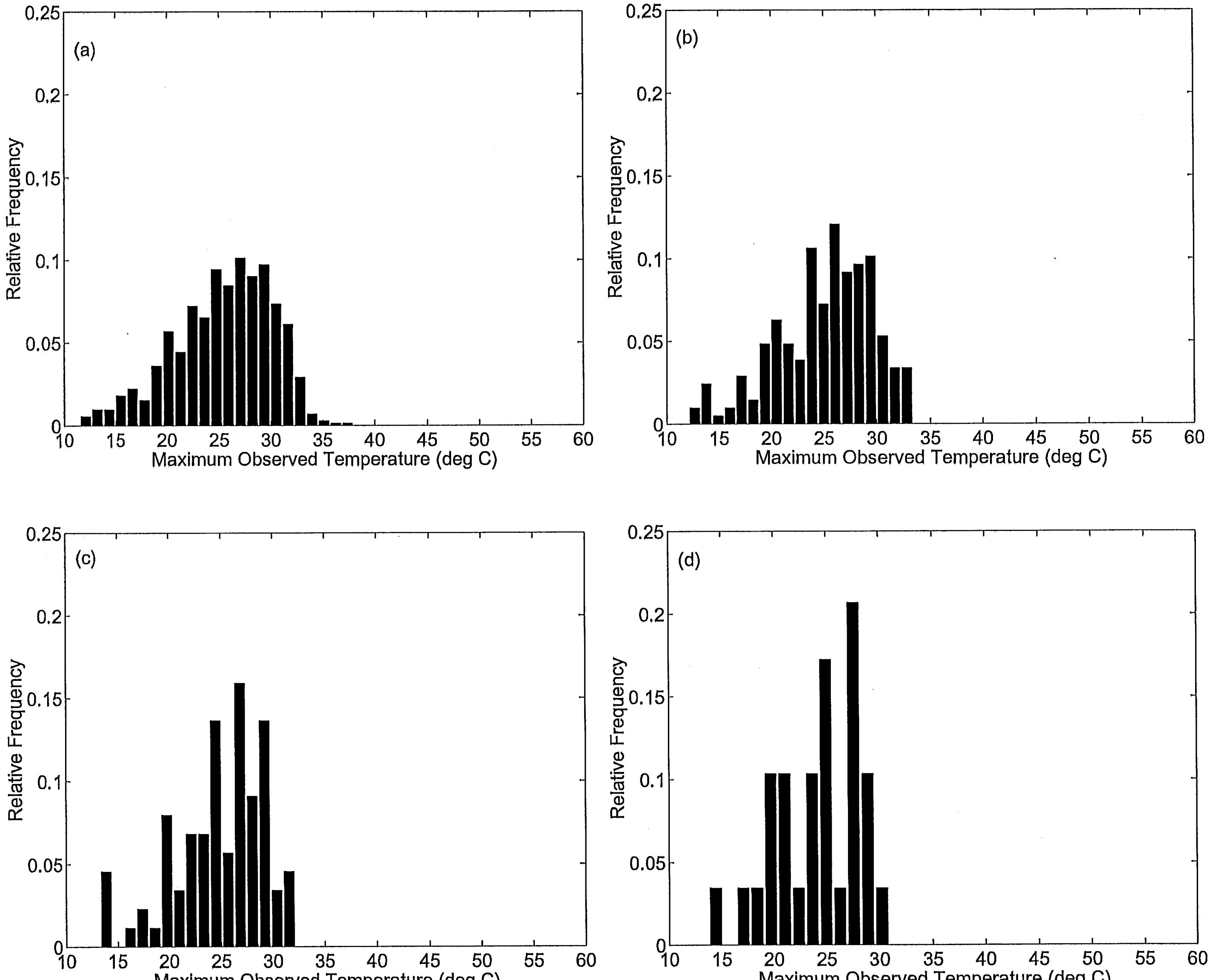


Figure 1.4. Histograms of the maximum observed temperatures stations. (a) all 720 stations. (b) 207 stations with records 15 years or longer. (c) 88 stations with records 20 years or longer. (d) 29 stations with records 25 years or longer.

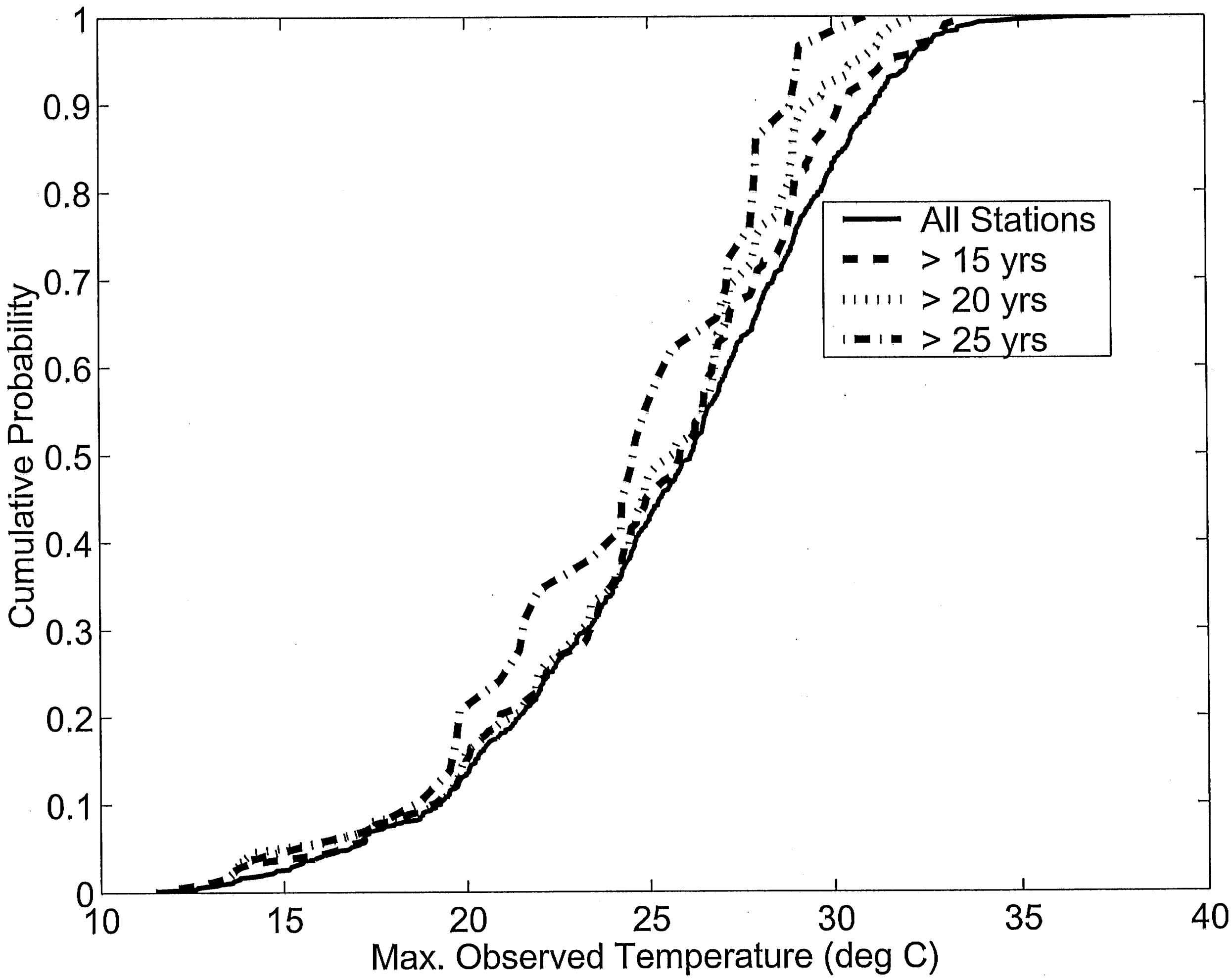


Figure 1.5. Cumulative distribution function for the observed maximum stream temperatures.

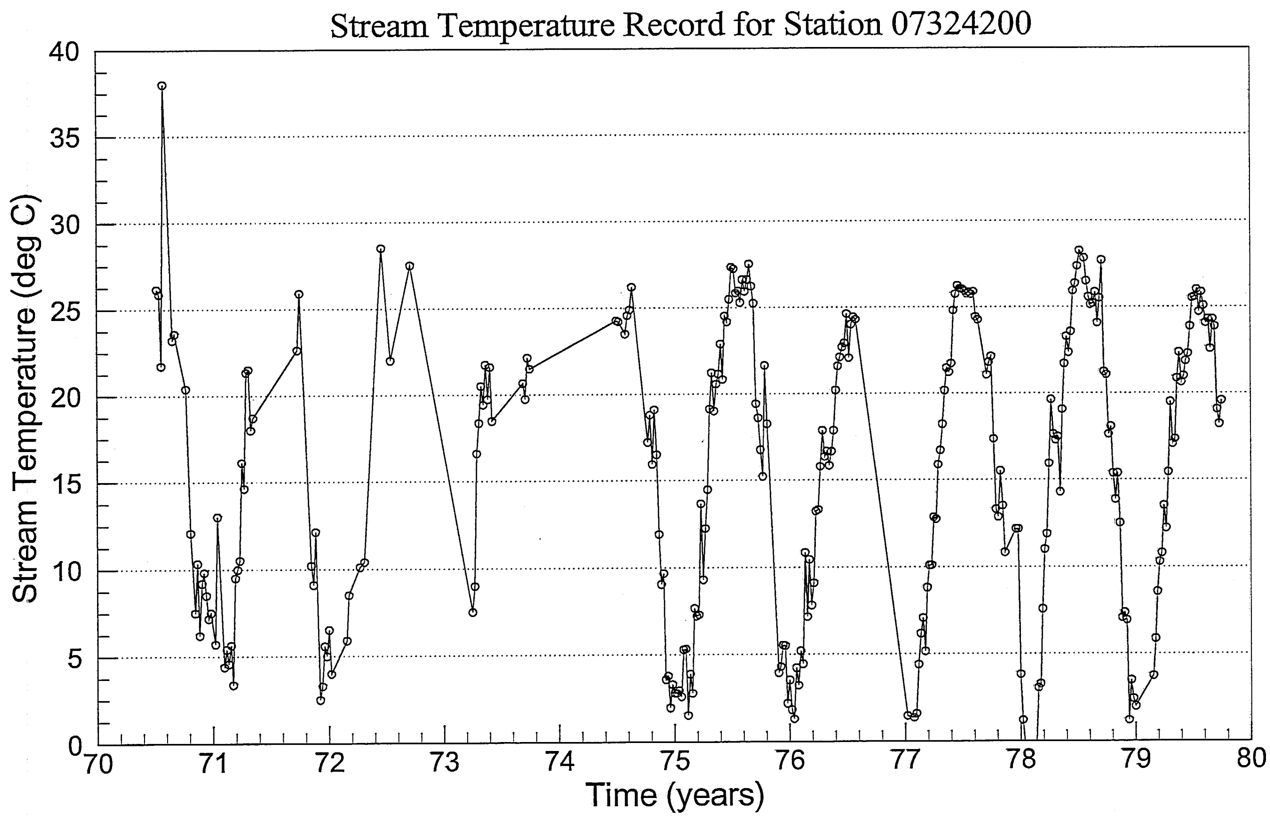


Figure 1.6. The stream temperature record for station 07324200. Measurements are represented by dots.

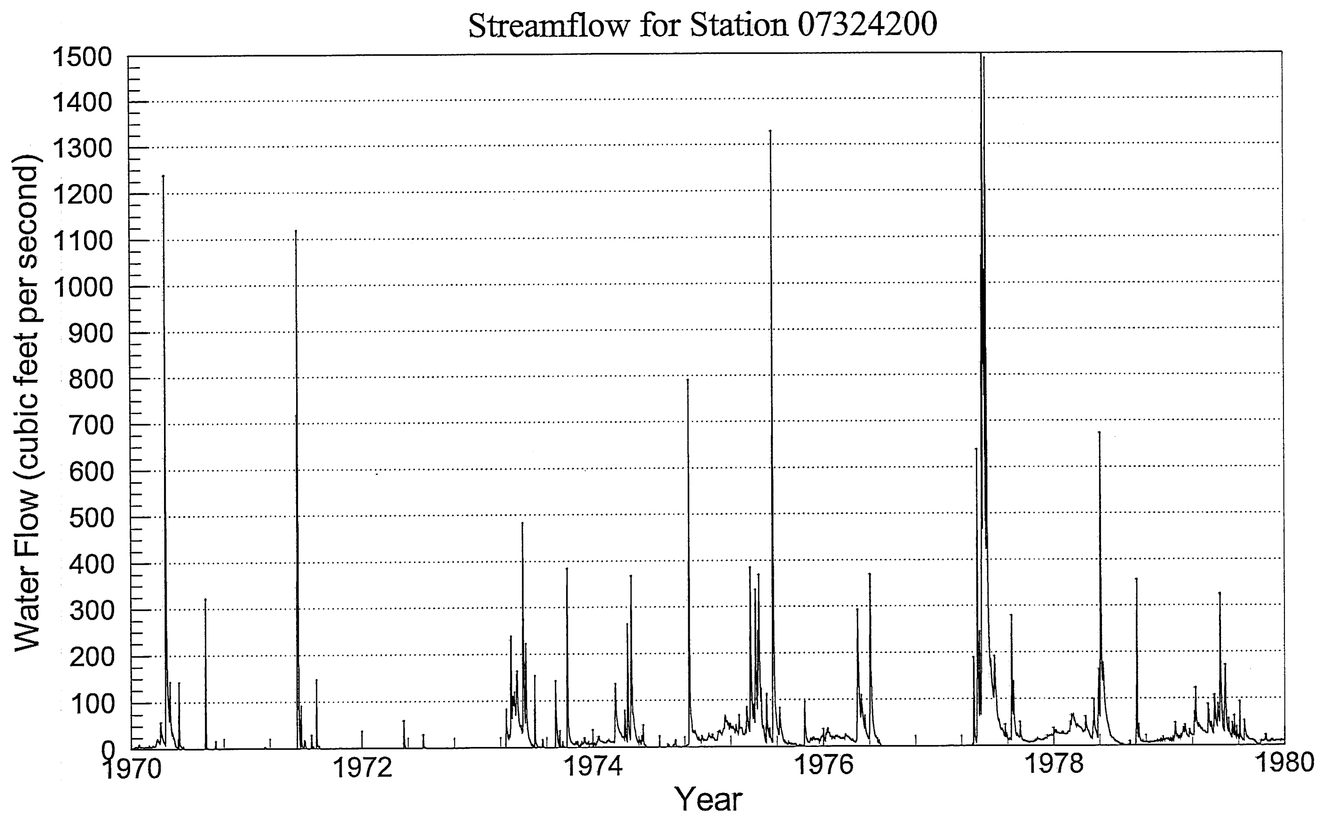


Figure 1.7. The stream flow record for station 07324200.

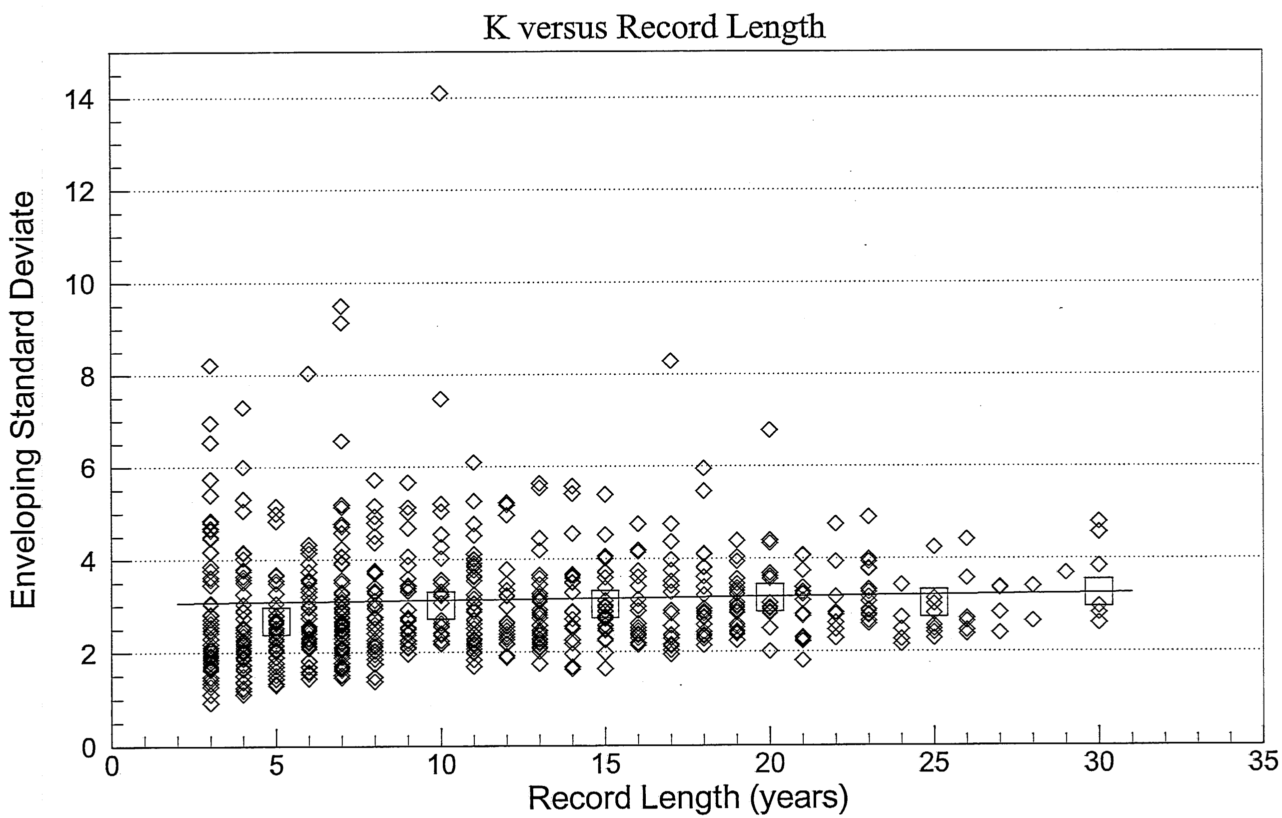


Figure 1.8. Computed K values as a function of Record Length. The average K values from the partitioning described in the text are shown by the large squares. A line fitted to the average values is shown. $K = 0.017\mathrm{RL} + 2.72$. $R^2 = 0.719$.

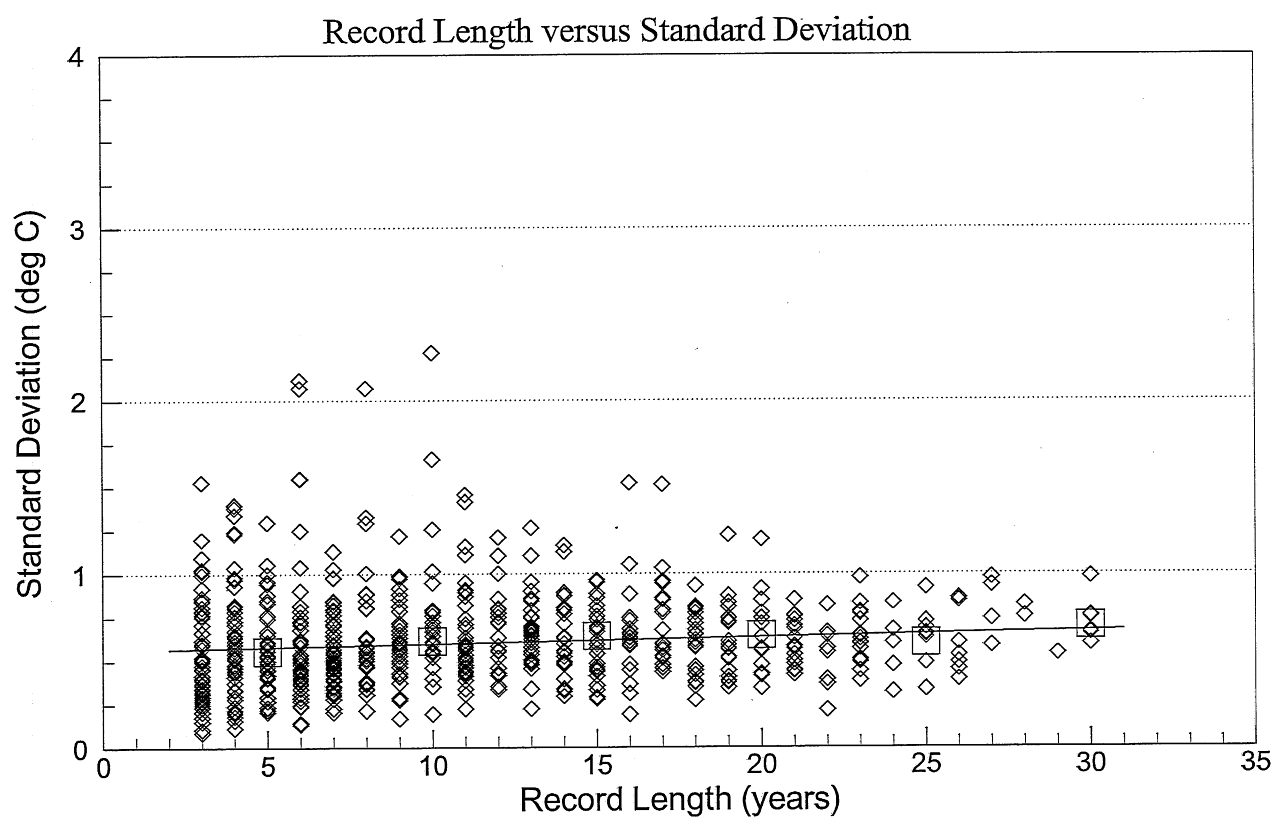
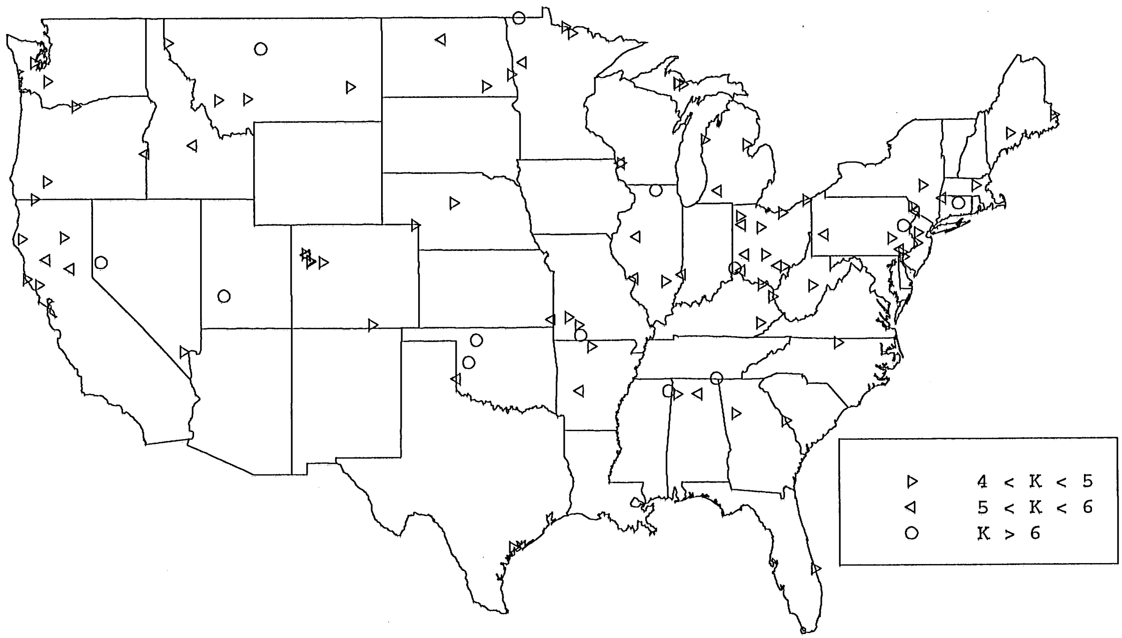


Figure 1.9. A plot of the computed s values as a function of Record Length. The average s values from the partitioning described in the text are shown with the large squares. A line fitted to the average values is also shown. $s = 0.0038\text{RL} + 0.559$. $R^2 = 0.578$.



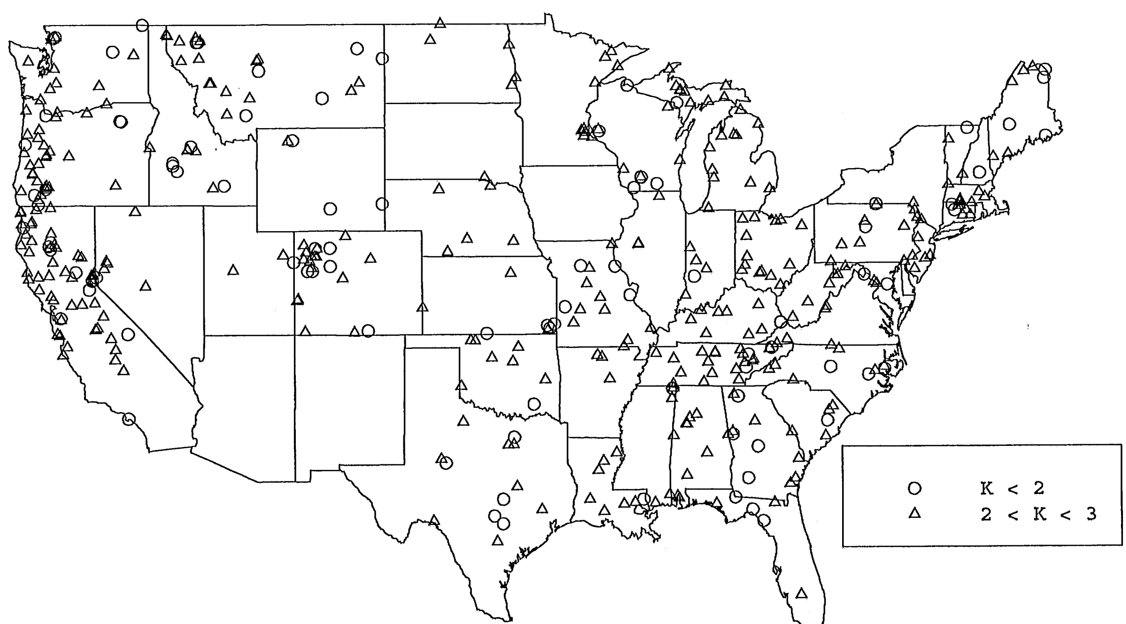


Figure 1.10. Top: A map showing the location of stream gauging stations in the United States with K-values larger than 4. Bottom: A map showing the location of stream gauging stations in the United States with K-values smaller than 3.

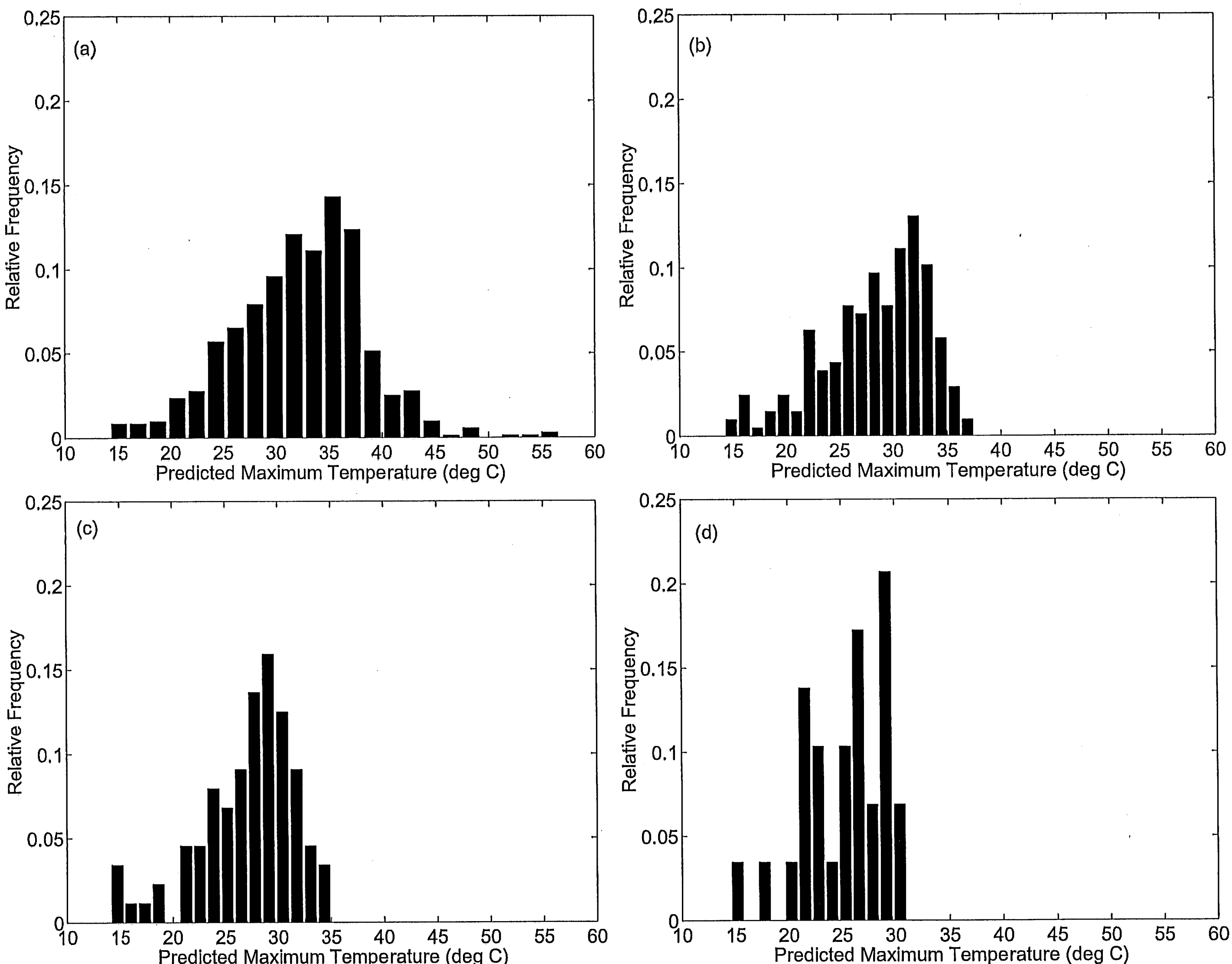


Figure 1.11. Maximum temperatures predicted by assuming the highest K-value for all streams. (a) All 720 stations; $K = 14.094$. (b) 207 stations with records 15 years or longer; $K = 8.282$. (c) 88 stations with records 20 years or longer; $K = 6.772$. (d) 29 stations with records 25 years or longer; $K = 4.790$.

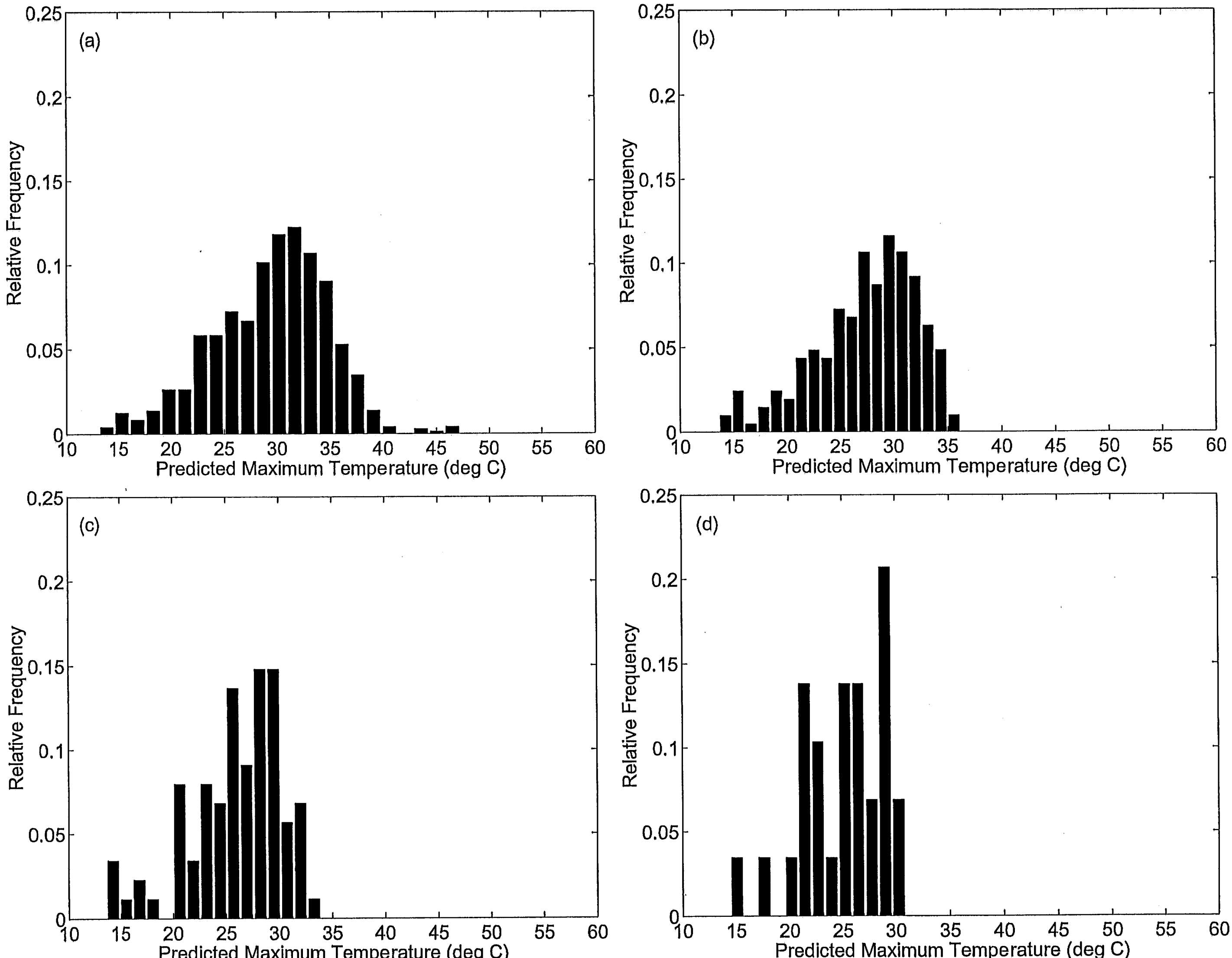


Figure 1.12. Maximum temperatures predicted by assuming the second highest K-value for all streams. (a) All 720 stations; $K = 9.494$. (b) 207 stations with records 15 years or longer; $K = 6.772$. (c) 88 stations with records 20 years or longer; $K = 4.8890$. (d) 29 stations with records 25 years or longer $K = 4.5670$.

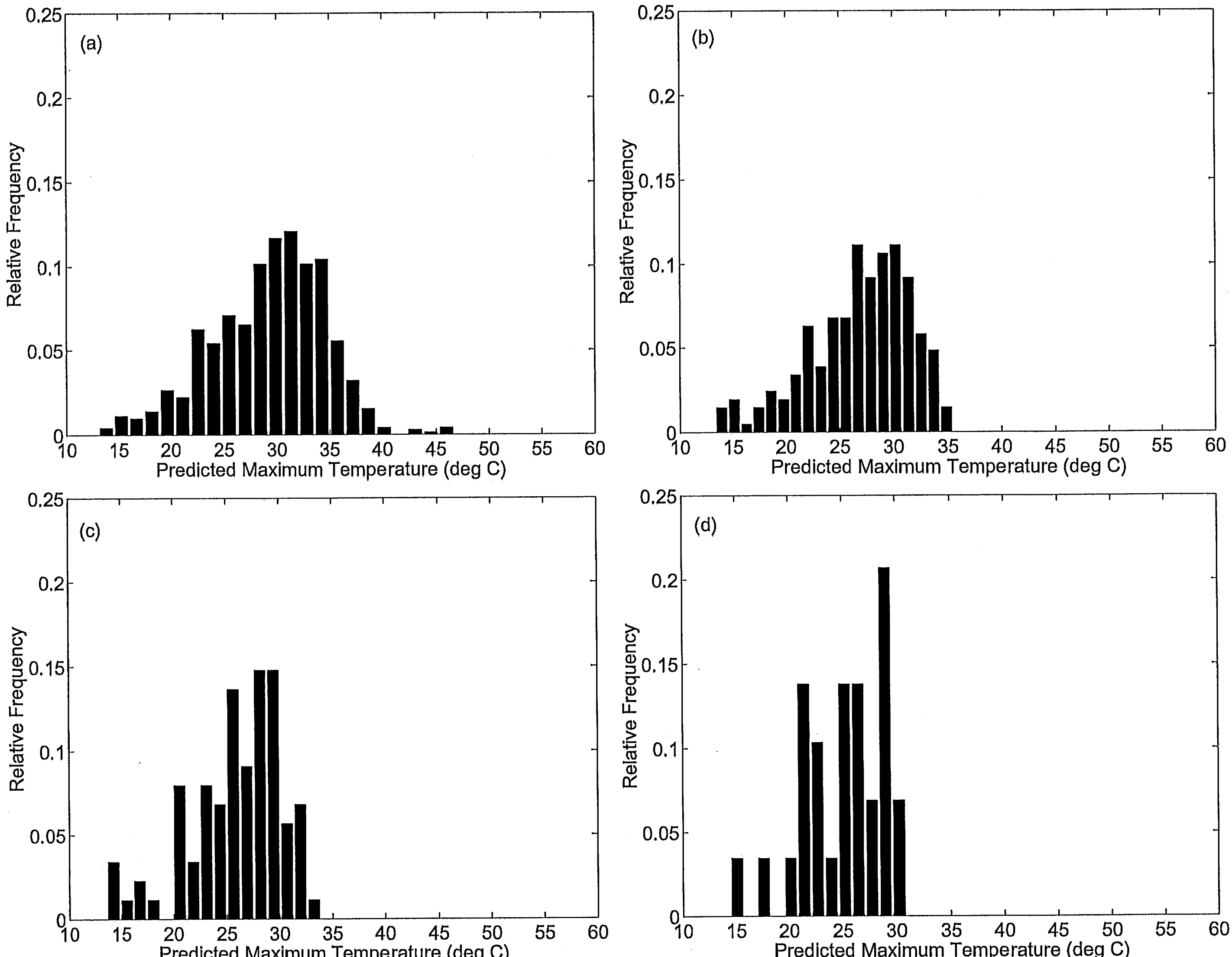


Figure 1.13. Maximum temperatures predicted by assuming the third highest K-value for all streams. (a) All 720 stations.; $K = 9.134$. (b) 207 stations with records 15 years or longer; $K = 5.941$. (c) 88 stations with records 20 years or longer; $K = 4.790$. (d) 29 stations with records 25 years or longer; $K = 4.562$.

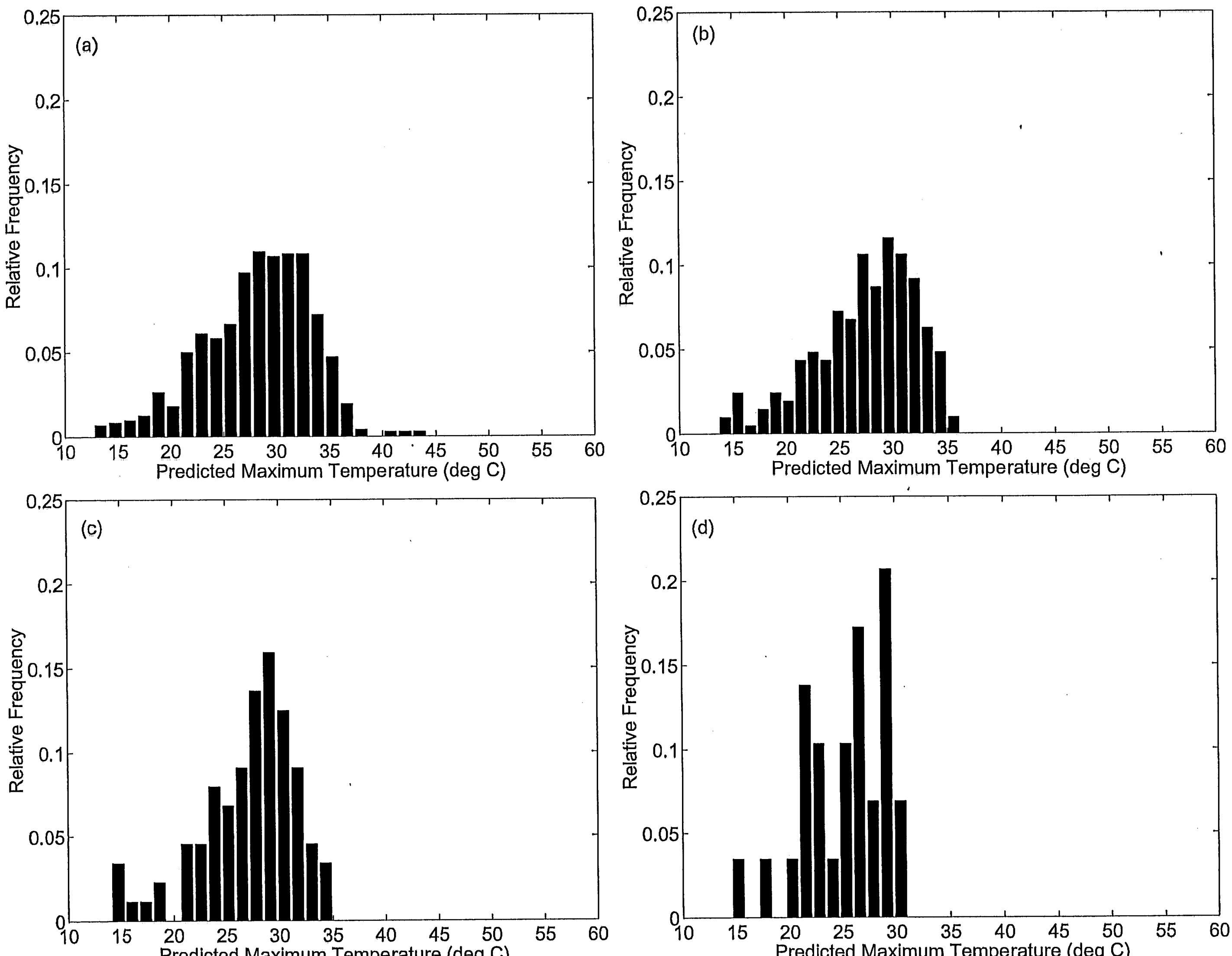


Figure 1.14. Maximum temperatures predicted by assuming the K value found at 99% in the observed distribution for all streams. (a) All 720 stations; $K = 7.468$. (b) 207 stations with records 15 years or longer; $K = 6.772$. (c) 88 stations with records 20 years or longer; $K = 6.772$. (d) 29 stations with records 25 years or longer; $K = 4.790$.

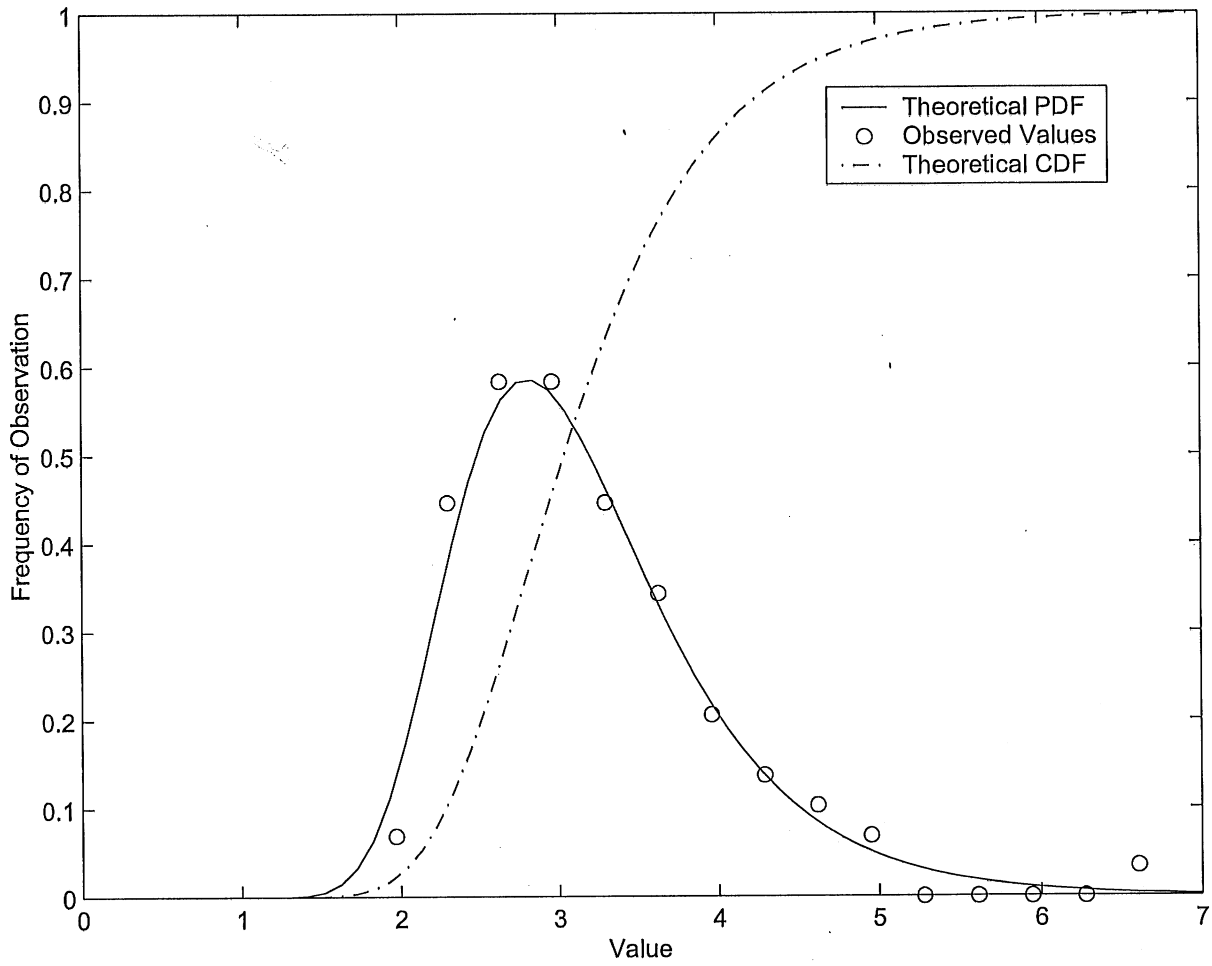


Figure 1.15. An Extreme-Value Type-I PDF and CDf fitted to the observed K-values from 88 stations with records 20 years or longer.

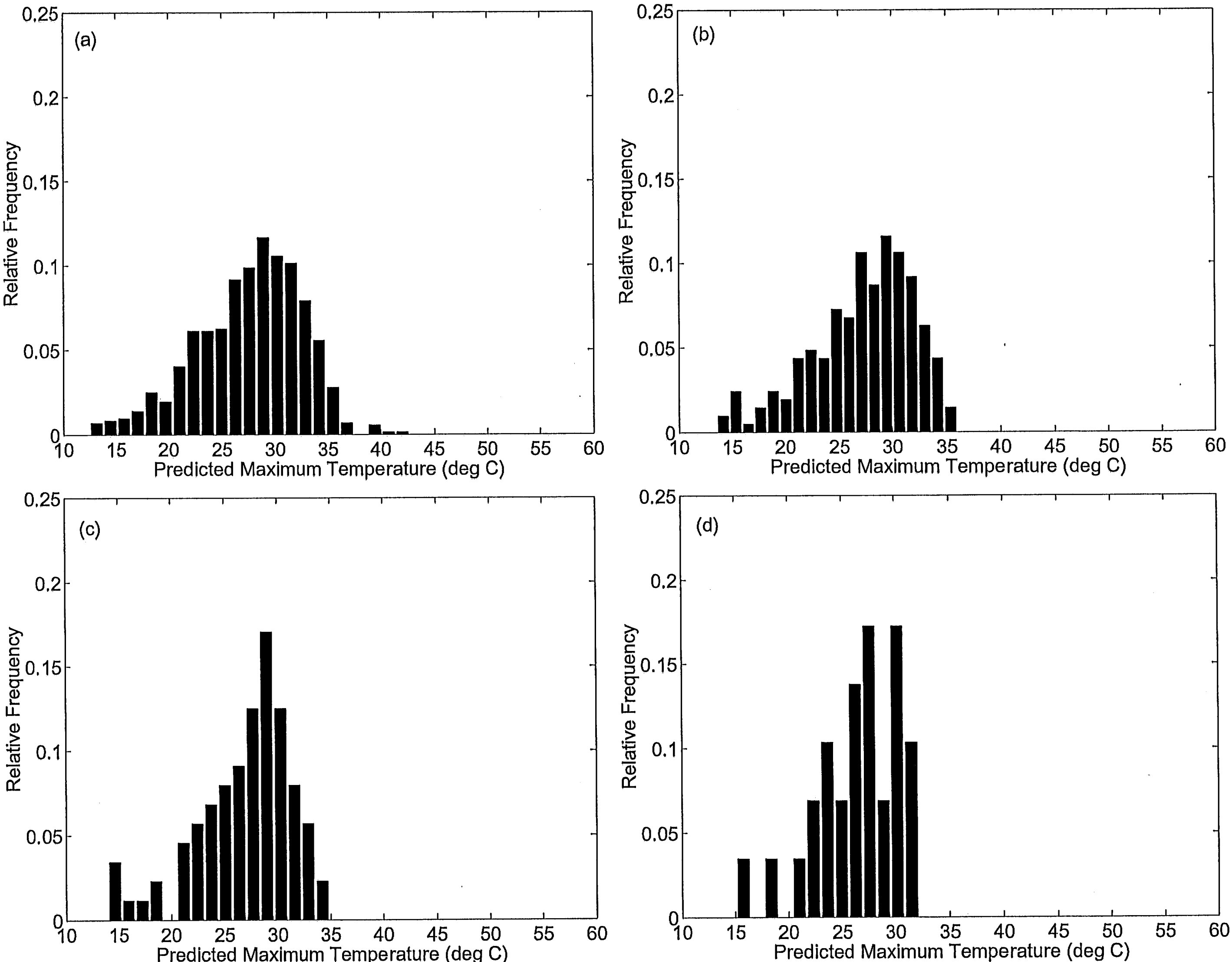


Figure 1.16. Maximum temperatures predicted by assuming the $K = 6.5$. (a) All 720 stations. (b) 207 stations with records 15 years or longer. (c) 88 stations with records 20 years or longer. (d) 29 stations with records 25 years or longer.

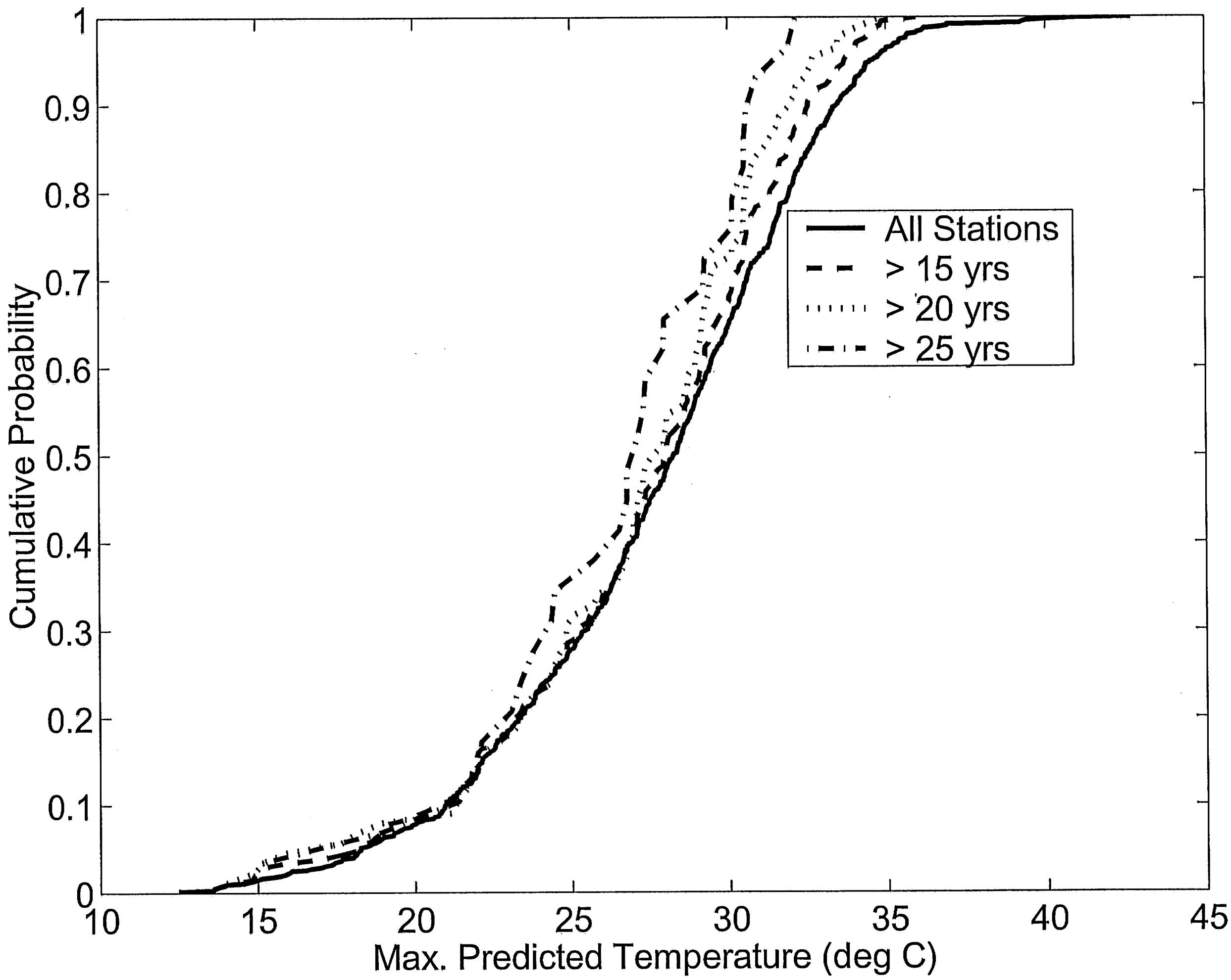


Figure 1.17. A cumulative distribution function for the predicted maximum stream temperatures with $K = 6.5$.

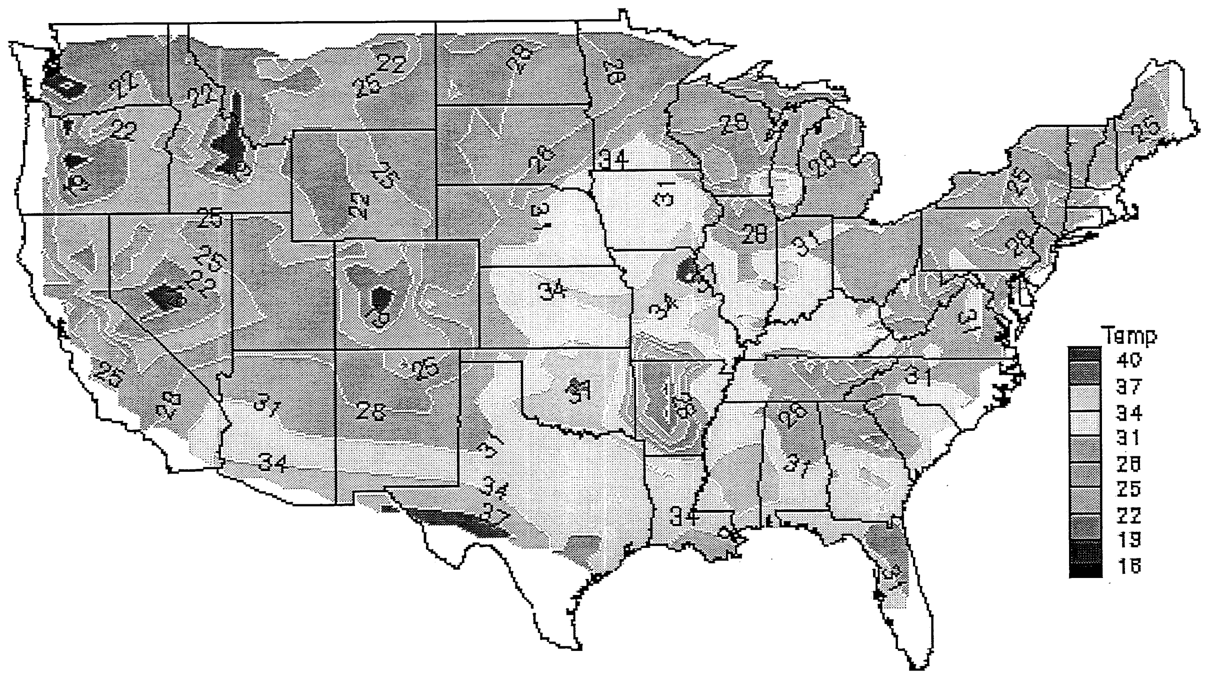


Figure 1.18. A map of the United States showing predicted maximum temperatures with $K = 6.5$.

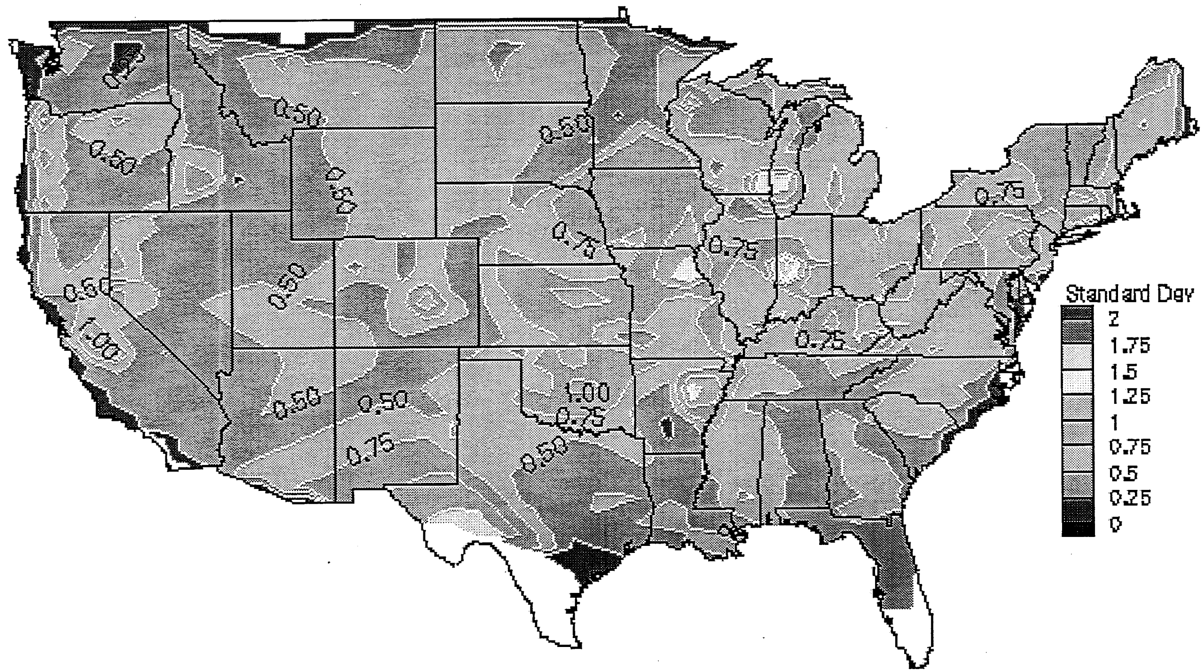


Figure 1.19. A map of the United States showing computed standard deviations.

2 Trend Analysis of Weekly Stream Temperature Data

2.1 Introduction

There is a strong indication that global climate change is taking place (Houghton et al., 1996). Warming trends have been found in various sets of data including air temperatures, soil temperatures and one set of water temperatures from Lake Zurich, Switzerland (Livingstone, 2001). We have examined weekly stream temperature data from 993 USGS stream gauging stations located throughout the contiguous United States. This report summarizes our trend analysis of this data.

2.2 Analysis

For our analysis we considered two subsets of our data. We examined 189 records that covered the period 1980 – 1990 and 79 records that covered the period 1970 – 1990. For each record we found annual maximum, minimum and mean temperatures for each year. We then fit a line through each series (maximum, minimum and mean) and examined the slopes. This gives us a general idea whether there are upward trends or not. Figures 2.1 and 2.2 show the results. There is not a significant bias towards positive slopes which suggest that either many streams are not getting warmer or that the trends we are looking for are more subtle than can be found in weekly temperature readings.

A possible explanation for the wide variety of slopes might be that streams in one region might have experienced warming while others did not. For example, streams in the northwest are heavily influenced by hypolimnetic dam releases and might not show significant changes in temperature. To test this hypothesis and look for geographical patterns we plotted the calculated slopes from the annual maximum temperatures on a map of the US (Figures 2.3 and 2.4). No geographical patterns were evident. Stations with a wide variety of slopes could be found in a small geographical area. The absence of a geographic patterns is also evident in the listing of the stream gauging stations with the largest positive or negative slopes (Table 2.1). The eleven stations are in nine different states.

We hypothesized that a trend might also be evident in the length of winter if climate warming had impacted the streams significantly. To test this hypothesis we looked at the number of weeks that a stream was "cold" per year. Once we had a count of "cold" weeks for each year we fit a line to the numbers and looked for downward trends. To count a week as "cold" we chose a threshold value of $1°C$. We did this for years covering the period 1980 – 1990 and 1970 – 1990. The results show little preference for negative slopes. For the 10 year period there were 60 stations with $m < 0$ and 52 stations with $m > 0$ (Figure 2.5). For the 20 year period there were 25 stations with $m < 0$ and 18 stations with $m > 0$ (Figure 2.6). The high number of zero slopes indicates that many stations never reached near-freezing temperatures (i.e. the count of cold weeks for all years was 0). This is not surprising since many streams were in warmer regions of the country. Also, gauging stations below dams are not likely to register cold temperatures.

2.3 Conclusion

Our motive in the analysis we performed was to detect any warming trends that might exist in weekly stream temperature records in the contiguous United States. Although some streams showed upward trends in temperatures we found no clear warming pattern. This may mean that as of 1990 any climate warming effects were overshadowed by other factors affecting stream temperatures and no warming of streams had occurred. Our method of detection for trends, linear regression, is somewhat crude and it may be that noise in the data (random variation) is enough to throw off our linear regression slopes. For our analysis we also used weekly temperature readings which may have obscured extreme highs and low through the averaging process. Livingstone (2001) found trends in monthly profiles of a lake in Switzerland, but weekly readings might be too infrequent for trend analysis in streams. Future analysis of this problem should include data more recent than 1990. Livingston found the clearest trends in the most recent data and the same situation is likely to be the case with stream temperatures.

Table 2.1. The station locations for the 9 stations with regression slope $m < -0.1$ and the 2 stations with $m > 0.1$.

Station No.	Location	Lat.	Lon.	Slope, m
01454720	Lehigh River At Easton, PA	40.7	-75.2	-0.333
01482800	Delaware River At Reedy Island Jetty, DE	39.5	-75.6	-0.275
03575000	Flint River Near Chase, AL	34.8	-86.5	-0.269
01362198	Esopus Creek At Shandaken, NY	42.1	-74.3	-0.234
12037400	Wynoochee River Above Black Cr Nr Montesano, WA	47.0	-123.7	-0.162
11302000	Stanislaus R Bl Goodwin Dam Nr Knights Ferry, CA	37.9	-120.6	-0.134
01426500	West Branch Delaware River At Hale Eddy, NY	42.0	-75.3	-0.115
03099510	Mahoning R At OH-PA St Line Bl Lowellville, OH	41.0	-80.5	-0.113
06775500	Middle Loup River At Dunning, NB	41.8	-100.1	-0.101
14173500	Calapooia River At Albany, OR	44.6	-123.1	0.111
11303500	San Joaquin R Nr Vernalis, CA	37.7	-121.3	0.279

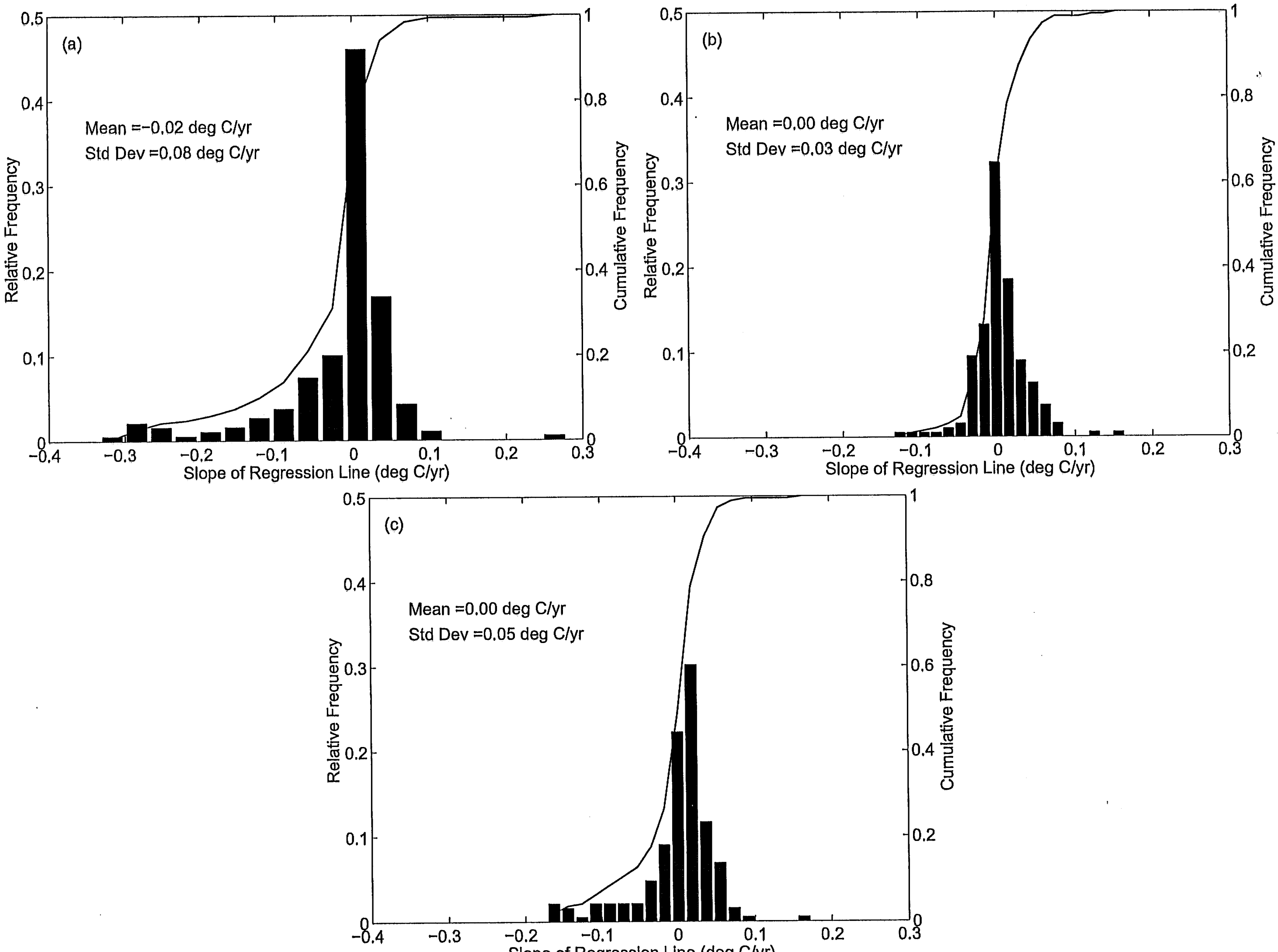


Figure 2.1. Histograms of slopes calculated for regression lines for stream temperature data from 1980 – 1990 (189 stations). (a) annual maximum ($103 > 0$, $86 \leq 0$). (b) annual minimum ($109 > 0$, $80 \leq 0$). (c) annual average ($142 > 0$, $47 \leq 0$).

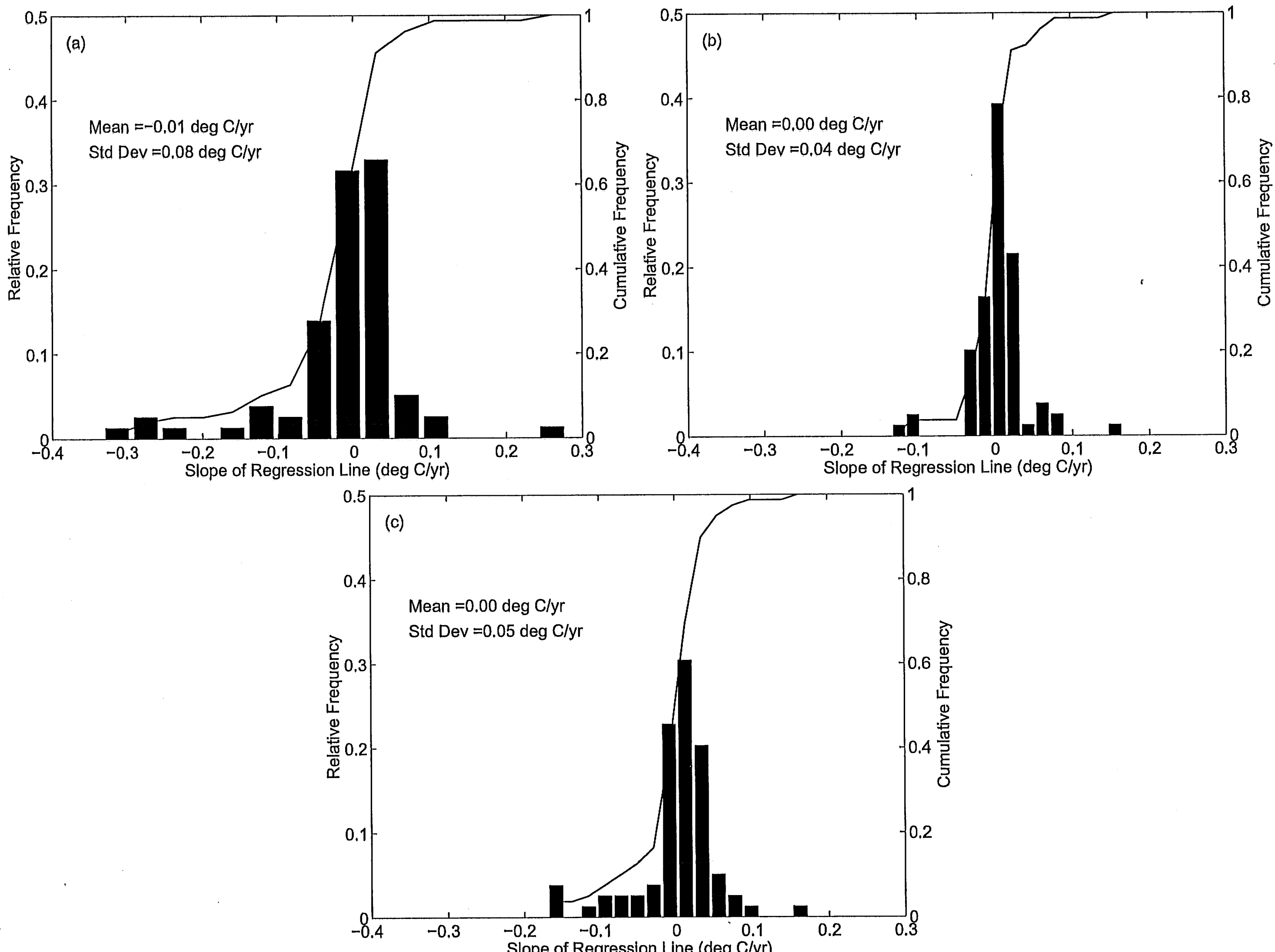


Figure 2.2. Histograms of slopes calculated for regression lines for stream temperature data from 1970 – 1990 (79 stations). (a) annual maximum ($38 > 0, 44 \leq 0$). (b) annual minimum ($49 > 0, 33 \leq 0$). (c) annual average ($51 > 0, 31 \leq 0$).

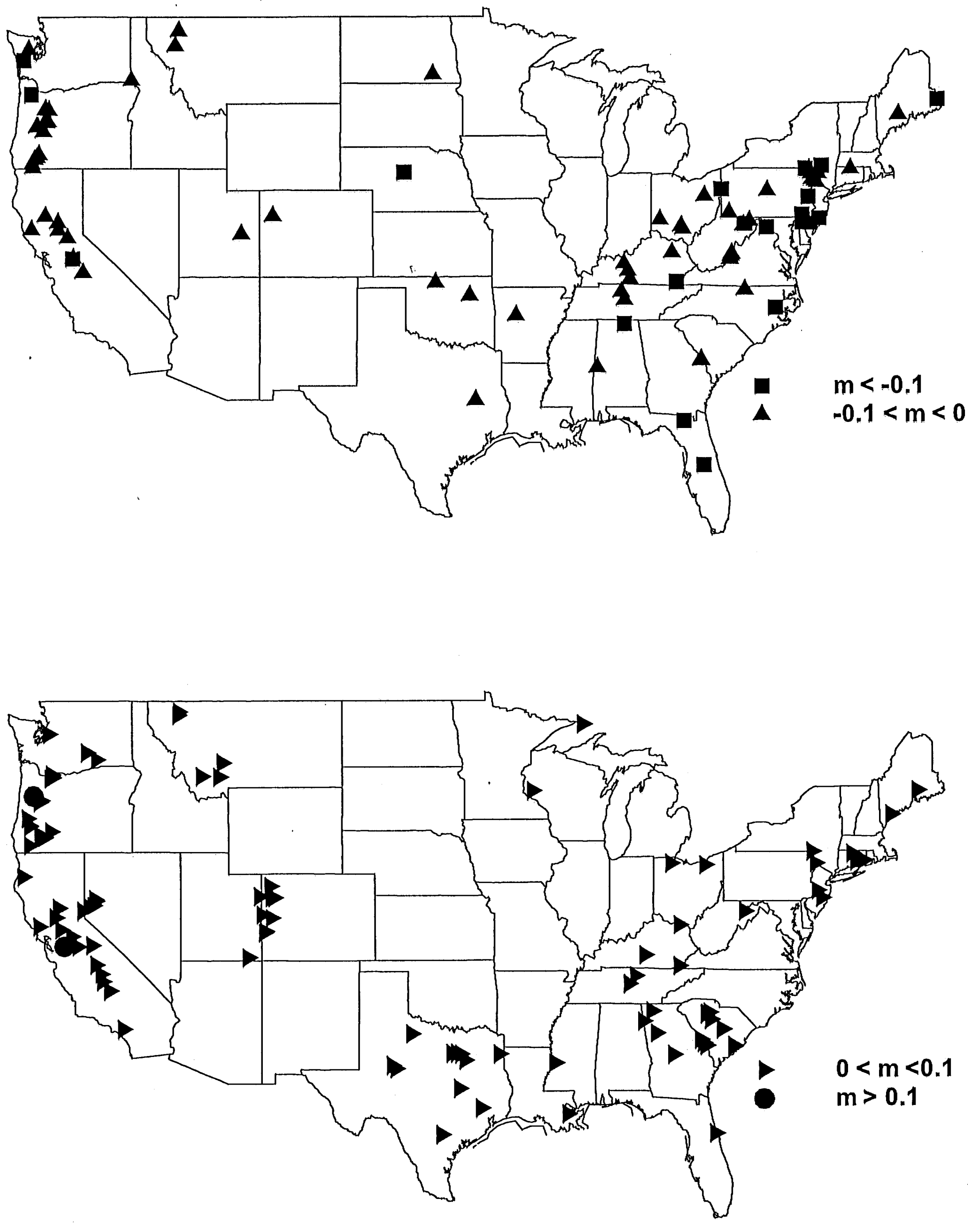
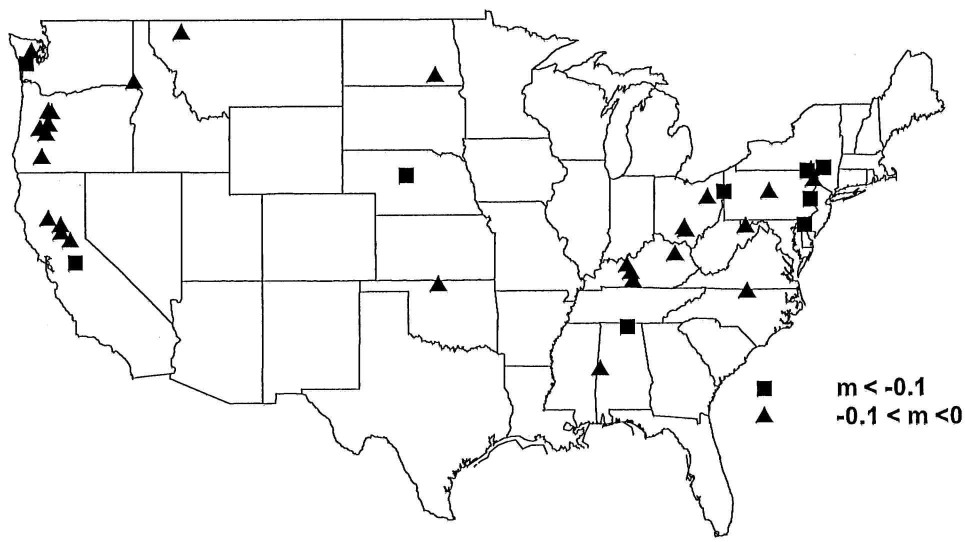


Figure 2.3. Top: Plot of stations in the US with slopes less than 0 for 1980 – 1990. Bottom: Plot of stations with slopes greater than 0 for 1980 – 1990.



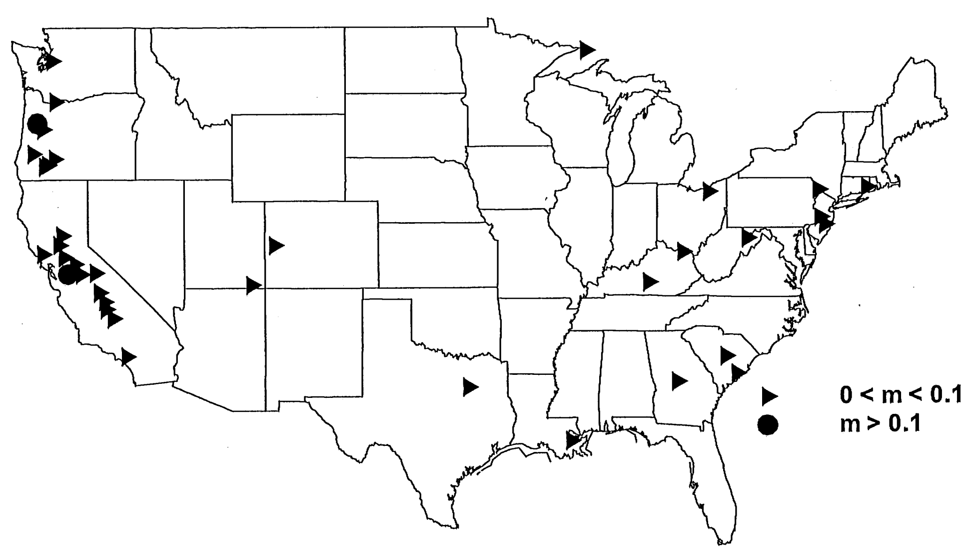


Figure 2.4. Top: Plot of stations in the US with slopes less than 0 for 1970 – 1990. Bottom: Plot of stations with slopes greater than 0 for 1970 – 1990.

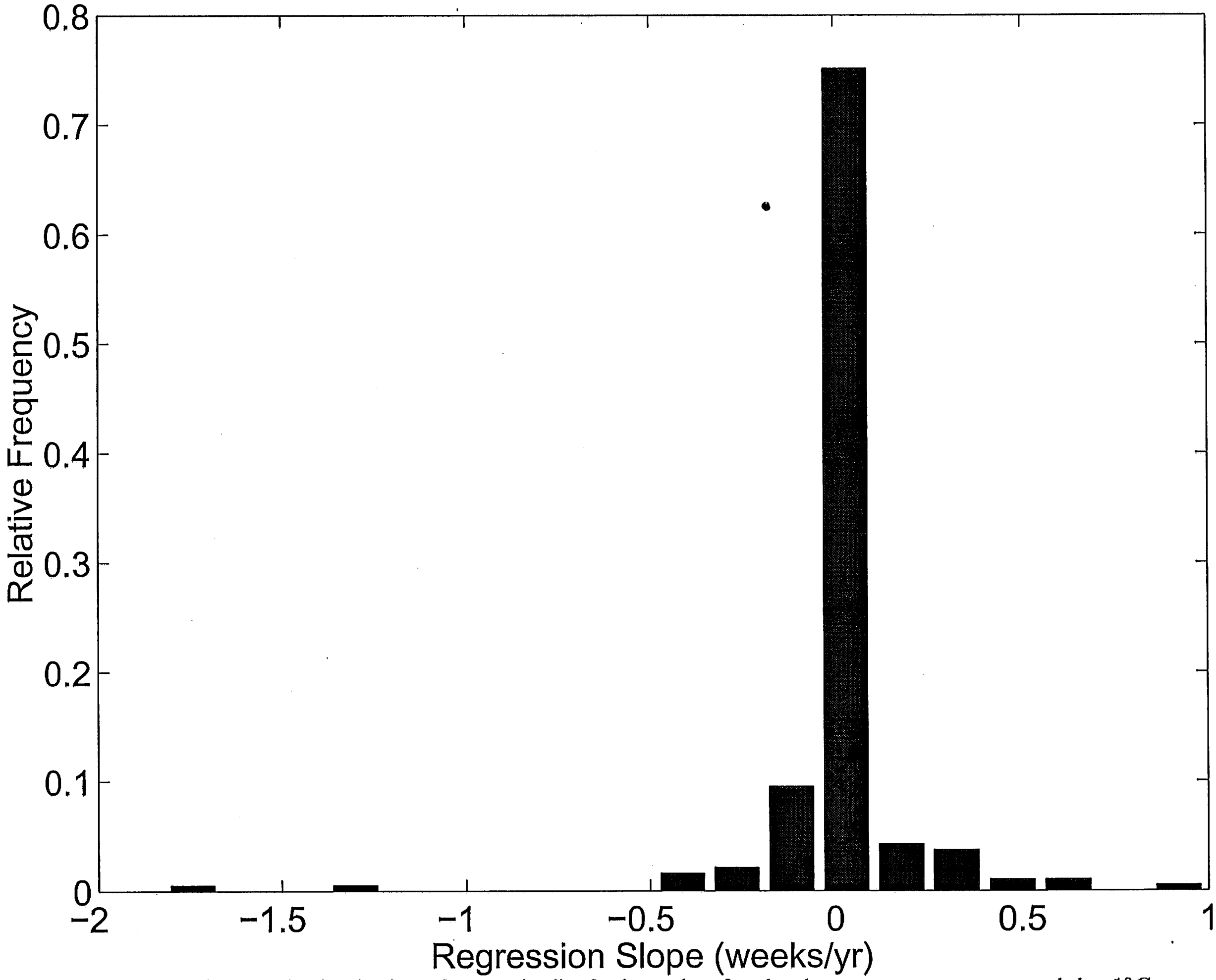


Figure 2.5. A histogram showing the slope of a regression line for the number of weeks where stream temperatures were below $1°C$ for the period 1980 – 1990.

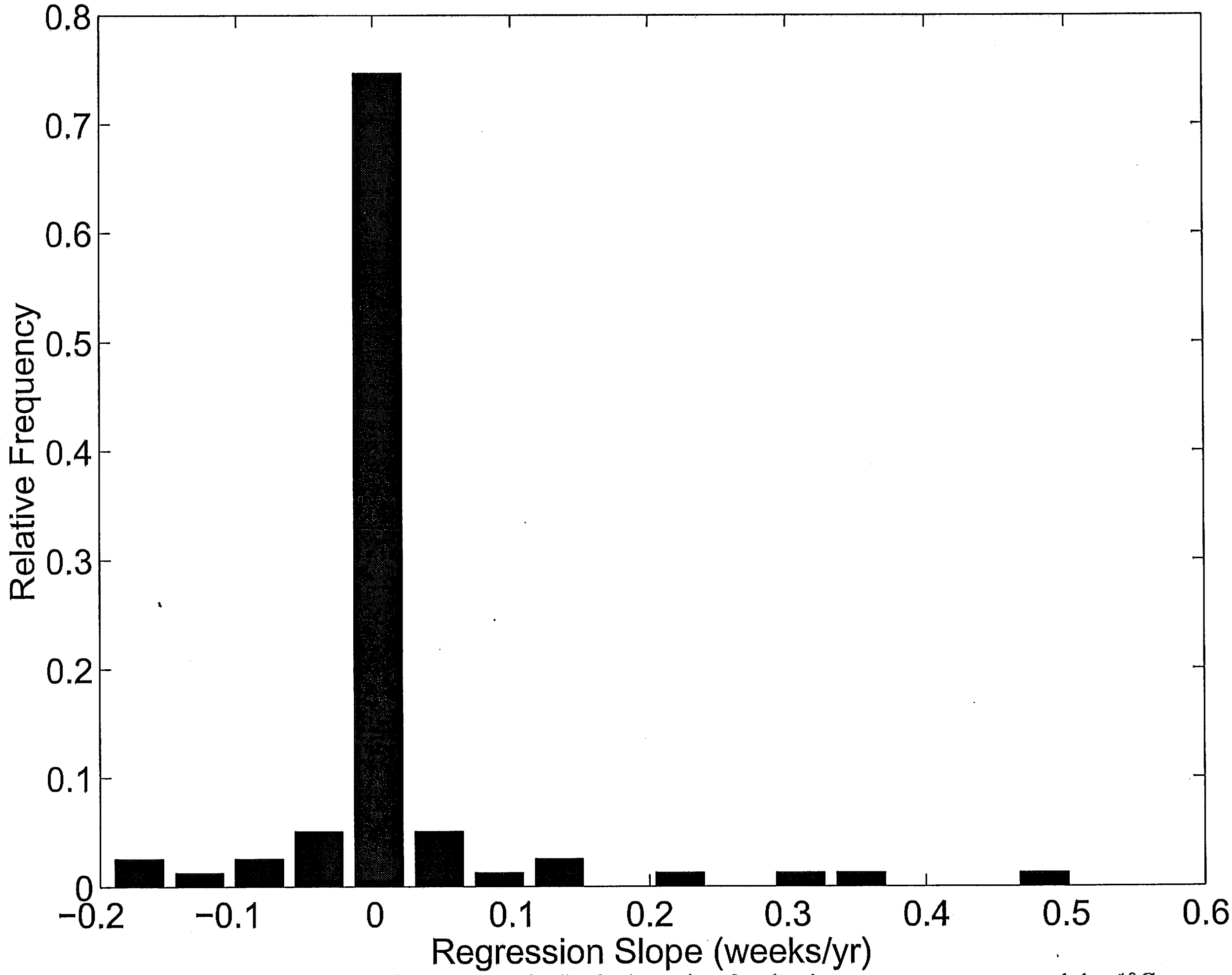


Figure 2.6. A histogram showing the slope of a regression line for the number of weeks where stream temperatures were below $1°C$ for the period 1970 – 1990.

3 Estimation of maximum stream temperatures through linearization and regression

3.1 Introduction

Observations of the air/stream temperature relationship show that as air temperature rises stream temperatures begin to level off. Stream temperatures also level off near $0°C$. The result is an S-shaped curve. Mohseni and Stefan (1998) did a curve fit for stream temperature/air temperature data for USGS water gauging stations throughout the United States. The curve they fit was:

$$T_s = \mu + \frac{\alpha - \mu}{1 + e^{\gamma(\beta - T_a)}} \tag{3}$$

where α is the maximum stream temperature, β is the air temperature at the inflection point of the curve, μ is the estimated minimum stream temperature, γ is a function of the slope at the point of inflection on the curve, T_s is the stream temperature, and T_a is the air temperature. The process used for the curve fit was a least-squares optimization of parameters. Unfortunately there were many stations with insufficient data in the upper regions of the curve to do a reliable fit. For that reason we will revisit the problem with a different method for parameter estimation.

3.2 Linear Transformations

Scientists studying chemical kinetics are frequently faced with a similar problem of parameter estimation. The relationship between reaction speed at substrate density is given by the Michaelis-Menten equation:

$$v = \frac{V[S]}{K_s + [S]} \tag{4}$$

Fortunately there are several transformations that can be applied to Equation 4 to more easily estimate the parameters (Brezonik, 1994; Lehninger, 1975). The first transformation is the Lineweaver and Burk equation:

$$\frac{1}{v} = \frac{1}{V} + \frac{K_s}{V}\frac{1}{[S]} \tag{5}$$

There is also the Hanes equation:

$$\frac{[S]}{v} = \frac{K_s}{V} + \frac{1}{V}[S] \tag{6}$$

And the Eadie-Hofstee equation:

$$v = -K_s\frac{v}{[S]} + V \tag{7}$$

All of these equations allow estimation of V through a linear regression process. Equation 4 does not give the S-shape that we are looking for in a curve. We could get the desired S-Shape by substituting $[S] = e^{\gamma(T_a - \beta)}$ into Equation 4. Plots of the Michaelis-Menten equation(Equation 4) with the substitution and Equation 3 show slightly different shapes. It turns out that we can apply a similar transformations to Equation 3 to linearize it. The result is:

$$\frac{e^{\gamma T_a}}{(T_s - \mu)} = \frac{1}{(\alpha - \mu)}e^{\gamma T_a} + \frac{e^{\gamma\beta}}{(\alpha - \mu)} \tag{8}$$

From this we can easily extract α once we have performed a linear regression on $e^{\gamma T_a}$ and $e^{\gamma T_a}/(T_s - \mu)$.

3.3 Regression

When we performed the transformation we were unable to remove the γ in the T_a term and the μ in the T_s term. This means that they must be estimated beforehand. To begin with we estimated μ from the smallest data point found. In Mohseni and Stefan (1998) we find the following expression for γ

$$\gamma = \frac{\tan\theta}{\alpha - \mu} \tag{9}$$

Mohseni and Stefan (1998) estimated $\tan\theta$ to be about 3.6. To obtain $\alpha - \mu$ we estimated α and μ from the largest and smallest data points, respectively.

As an example, we show three streams from three different states, Minnesota, Washington, and Oklahoma (Figures 3.1 – 3.3). The top plot in all three figures show the actual air/stream temperature data (air temperature data are from weather stations closest to the stream gauging sites). Since we are interested in high-temperature behavior we removed all points where either air temperature or stream temperature were below $5°C$. We did this because there tends to be a lot of "noise" in the data around $0°C$ which can throw off our estimation. After removing these points we applied the transformation to the data. The bottom plots show the transformed data and the fitted line.

We applied this method to data from 993 stream gauging stations. To examine the sensitivity to our choice of minimum temperature cutoff we performed the analysis with a minimum temperature w where $w = 2, 5, 10\ °C$. Histograms of the Results are shown in Figure 3.4. Figure 3.5 shows the distribution of r-values from the linear regressions. Overall the variability in distributions from our choice of w was not large. Variability of the distribution itself was much larger. Although we expected to see some variability in α from stream to stream, the actual variability we found is extremely wide and suggests that we may need to refine our method.

Although most of the lines showed a fairly close fit, there was a wide distribution of r-values. We established minimum r-values of 0.9, 0.95, and 0.99 and plotted only those stations whose fit was at least that good. The results are shown in Figure 3.6. We extracted maximum stream temperatures at different percentiles of the cumulative frequency functions (Table 3.1). Only those from minimum r-values of 0.99 seem meaningful. We also plotted the computed maximum values on a map of the United States (Figure 3.7). This values show more range than the values obtained earlier using a statistical estimation of upper bounds.

3.4 Conclusion

In using a linearization method for estimating α we sought to simplify the estimation process. We applied our method to many more streams than Mohseni and Stefan (1998) (993 versus 573), but the distribution of our results was similar to that of Mohseni and Stefan and the bounds were also similar. Although this seems to be a good method for estimation, it can still be refined. Although most streams showed a good linear fit (Figure 3.5) there were still many with a poor fit. One source of error might be our estimation of γ We estimated γ by using the range of our data for a particular stream. If the stream is from a cold region the range of the data may be much smaller than the actual range of temperatures for that stream. This suggests that a better method for estimating γ is needed. One method might be to employ an iterative method which optimizes γ for the best fit in terms of minimizing error. This would diminish the simplicity of the linearization method, but it might provide better estimations of α.

Table 3.1. Table showing α in $^\circ C$ at various percentiles of the distributions in Figure 3.6.

	Percentile			
Minimum r-value	90	95	99	99.9
.90	36.78	39.13	42.46	45.19
.95	34.93	36.41	40.60	42.87
.99	30.17	33.86	37.73	37.73

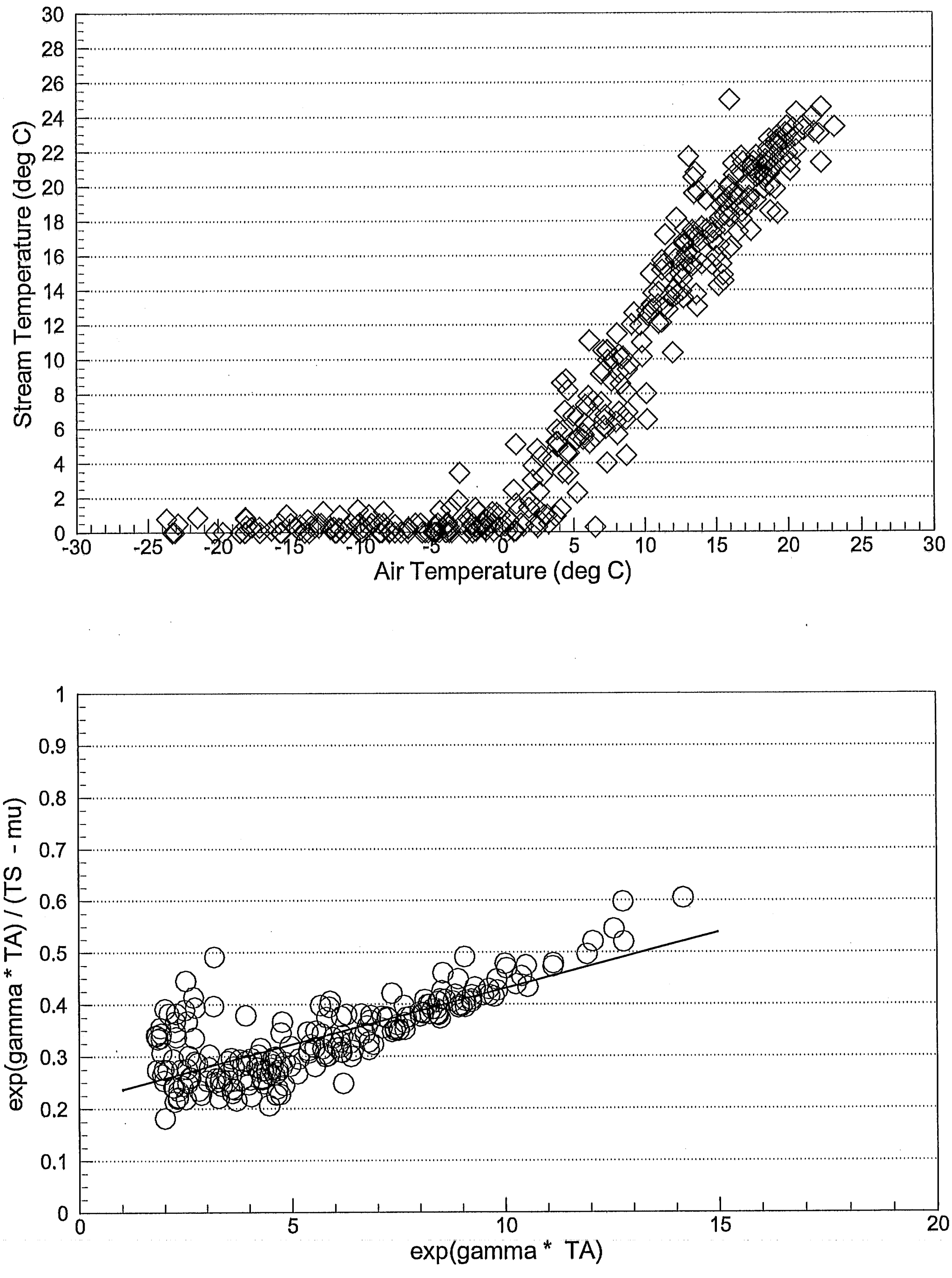


Figure 3.1. An example of stream temperature data and the transformation for station 04015475 at Partridge River above Colby Lake at Hoyt Lakes, MN. Top: All stream/air temperature data. Bottom: Transformed data and fitted line. $y = 0.022x + 0.215$, $r = 0.767$, where x and y are the horizontal and vertical axes on the lower plot, respectively.

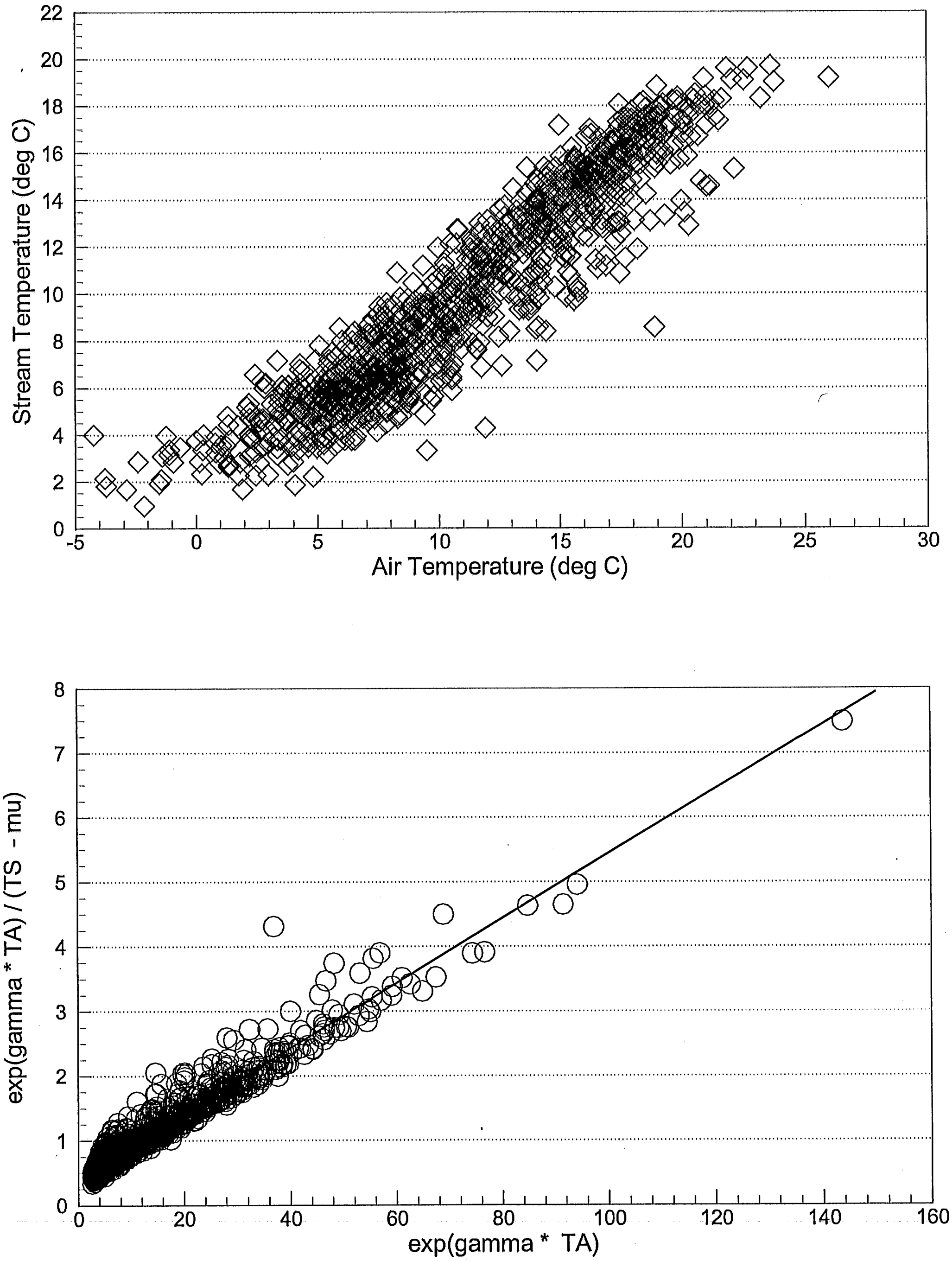
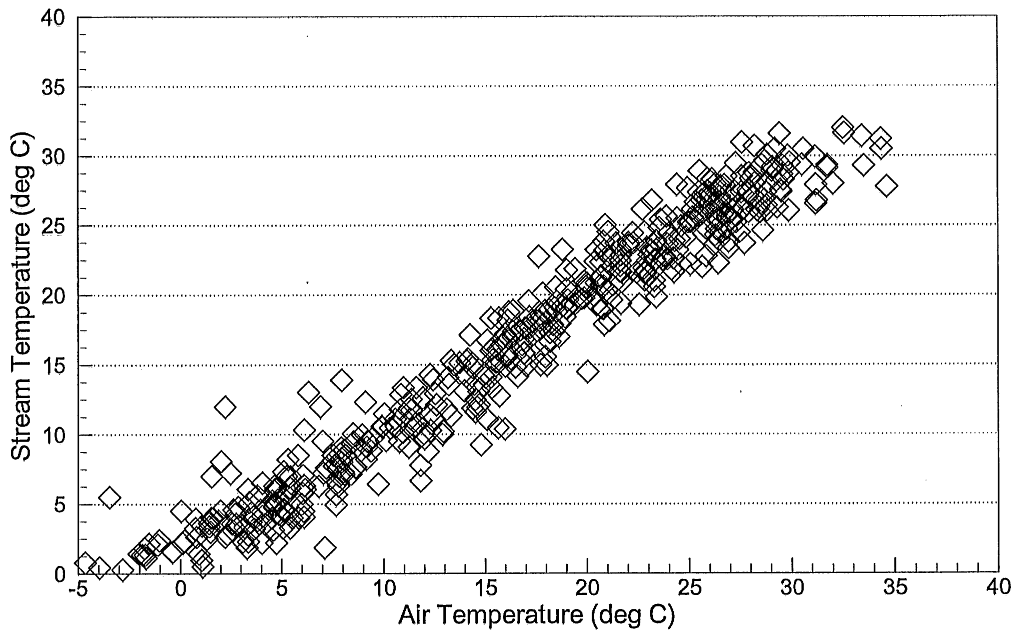


Figure 3.2. An example of stream temperature data and the transformation for station 12113000 at Green River near Auburn, WA. Top: All stream/air temperature data. Bottom: Transformed data and fitted line. $y = 0.050x + 0.428$, $r = 0.969$.



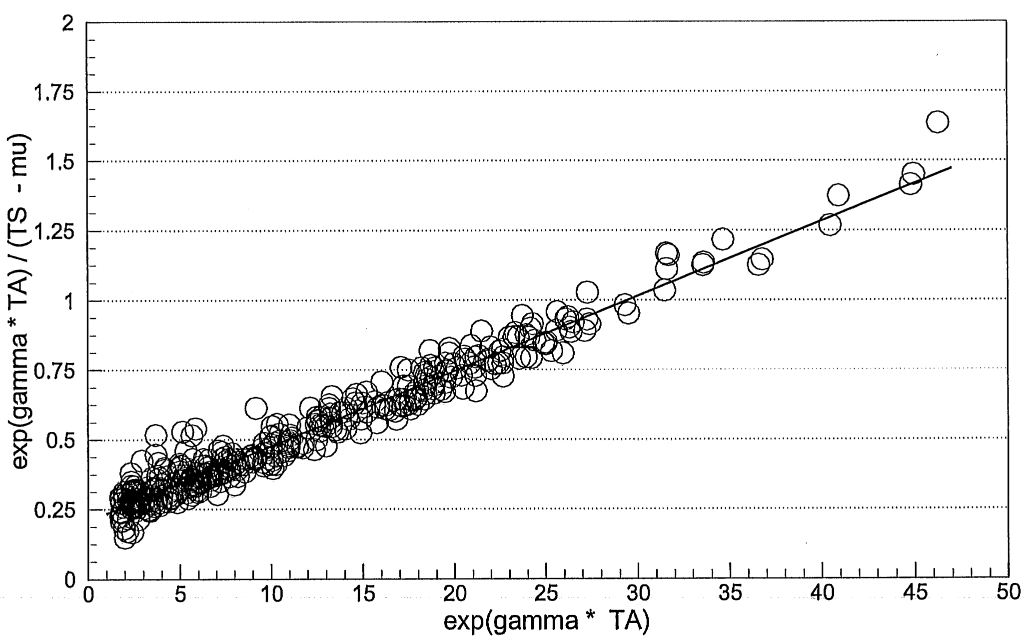


Figure 3.3. An example of stream temperature data and the transformation for station 07152500 at Arkansas River at Ralston, OK. Top: All stream/air temperature data. Bottom: Transformed data and fitted line. $y = 0.027x + 0.207$, $r = 0.979$.

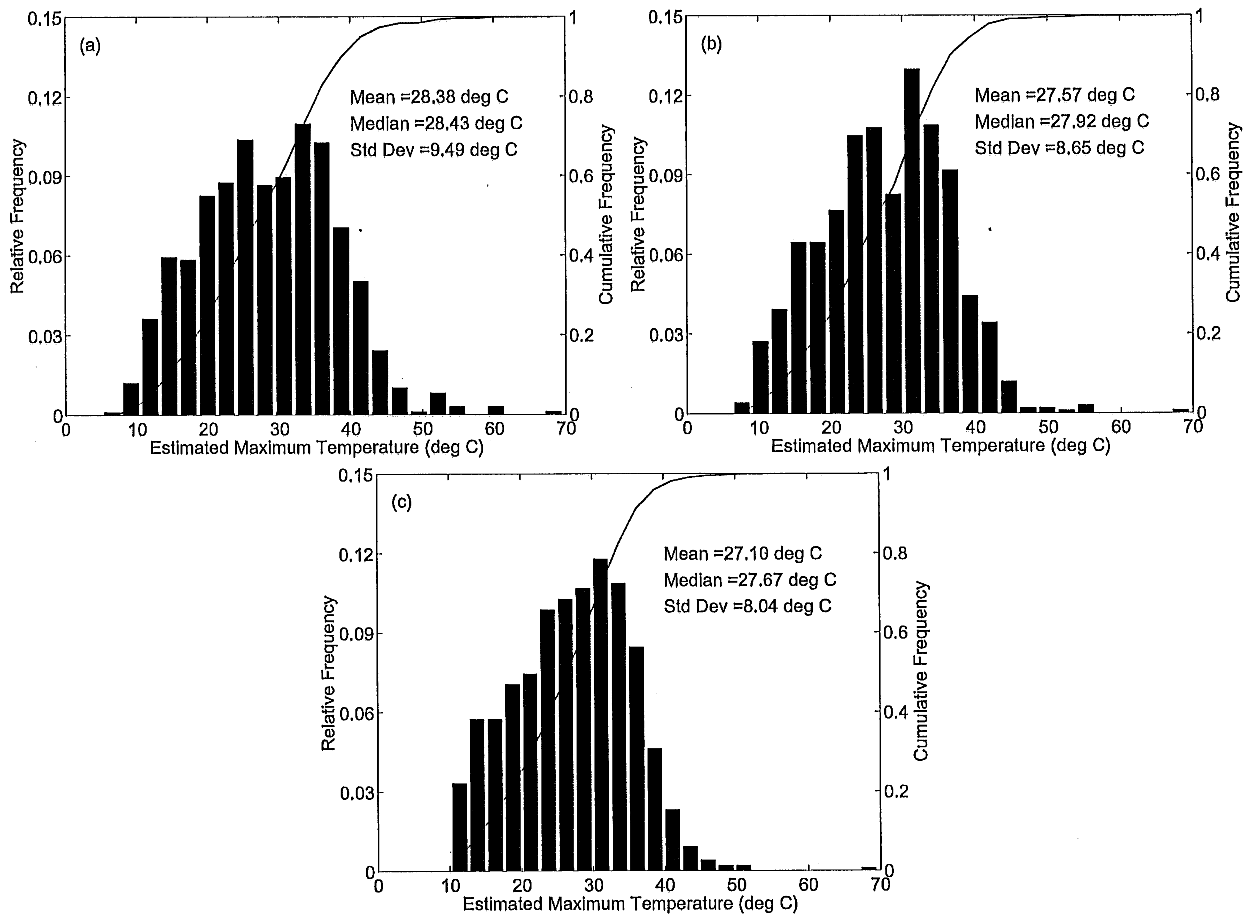


Figure 3.4. A histogram of the predicted α values from 993 streams. (a) $w = 2°C$. (b) $w = 5°C$. (c) $w = 10°C$.

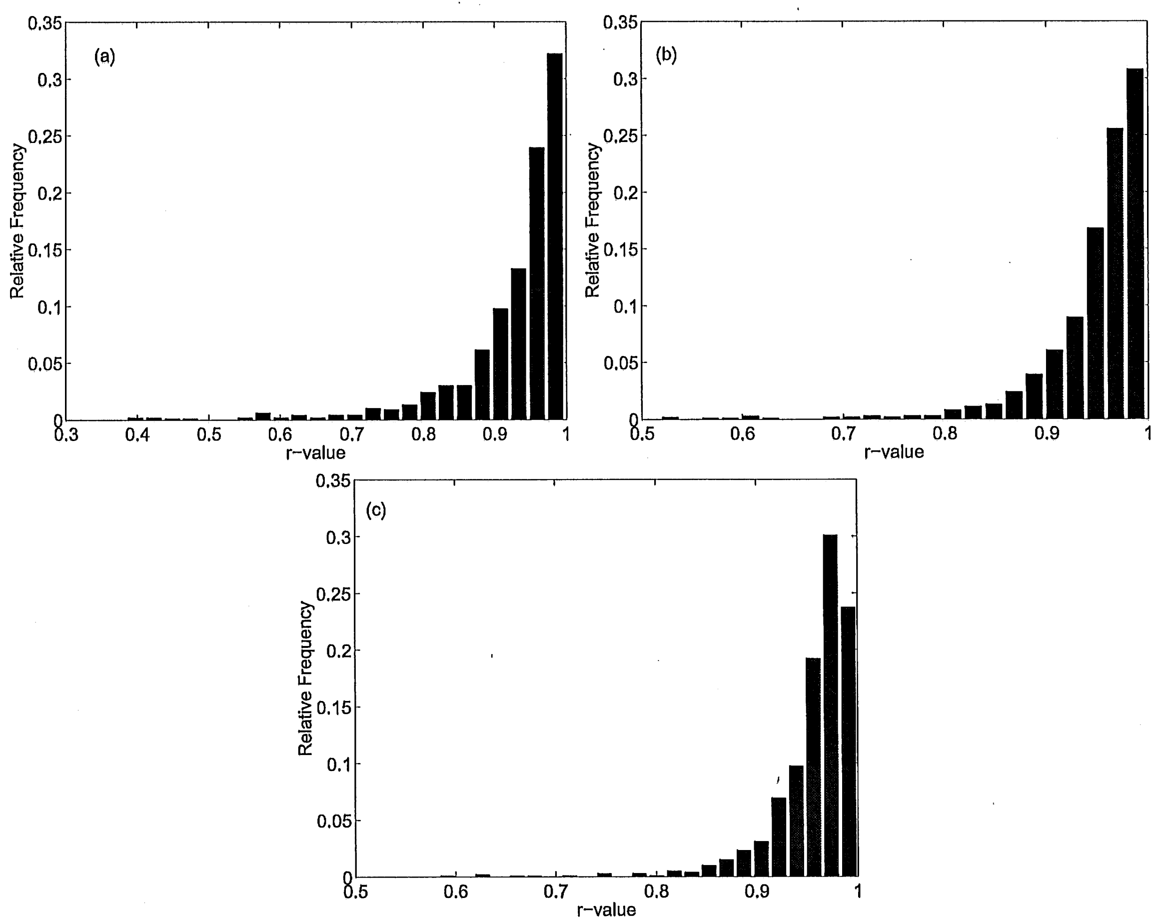


Figure 3.5. A histogram of the r-values from the linear regressions on 993 stations. (a) $w = 2°C$. (b) $w = 5°C$. (c) $w = 10°C$.

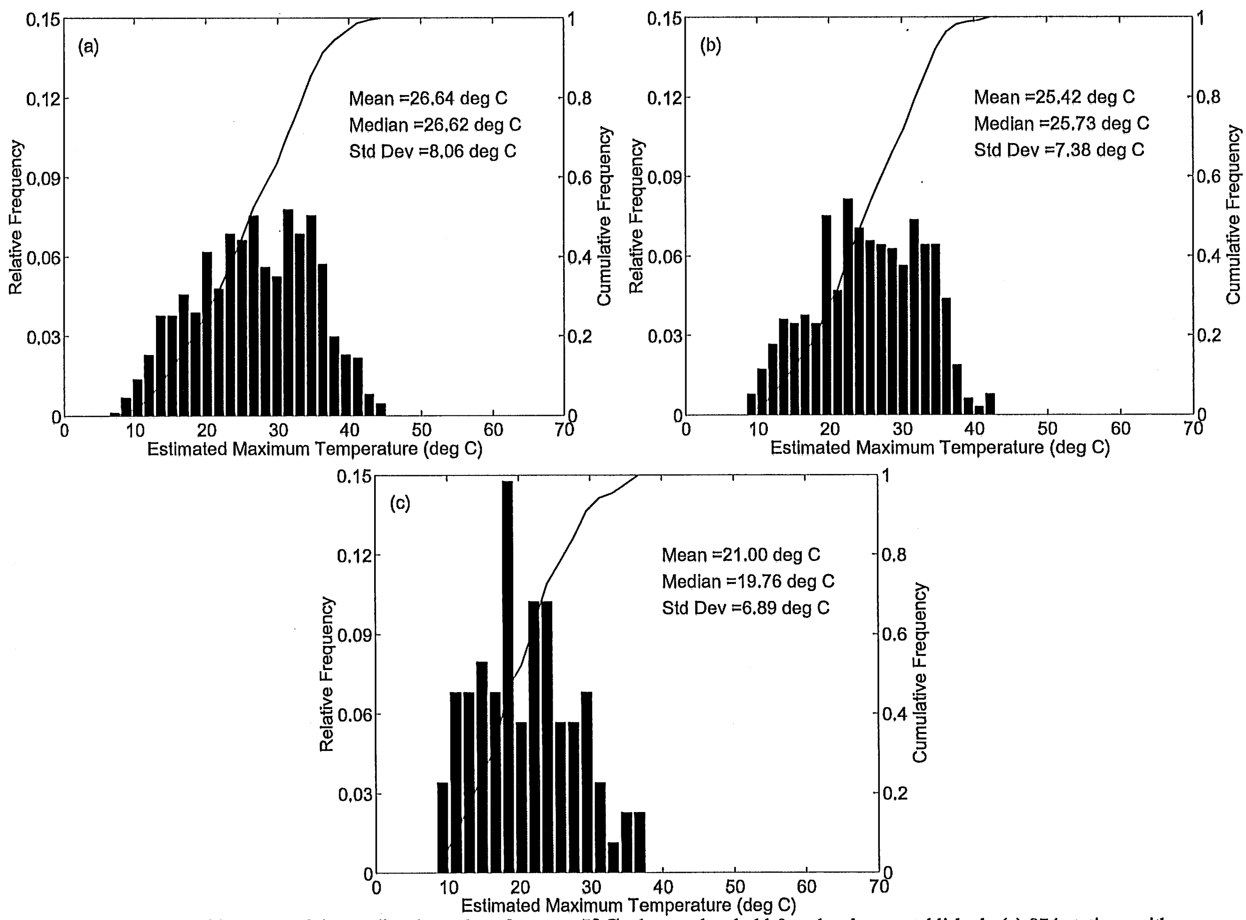


Figure 3.6. A histogram of the predicted α values for $w = 5°C$ where a threshold for r has been established. (a) 874 stations with $r \geq 0.90$. (b) 638 stations with $r \geq 0.95$. (c) 88 stations with $r \geq 0.99$.

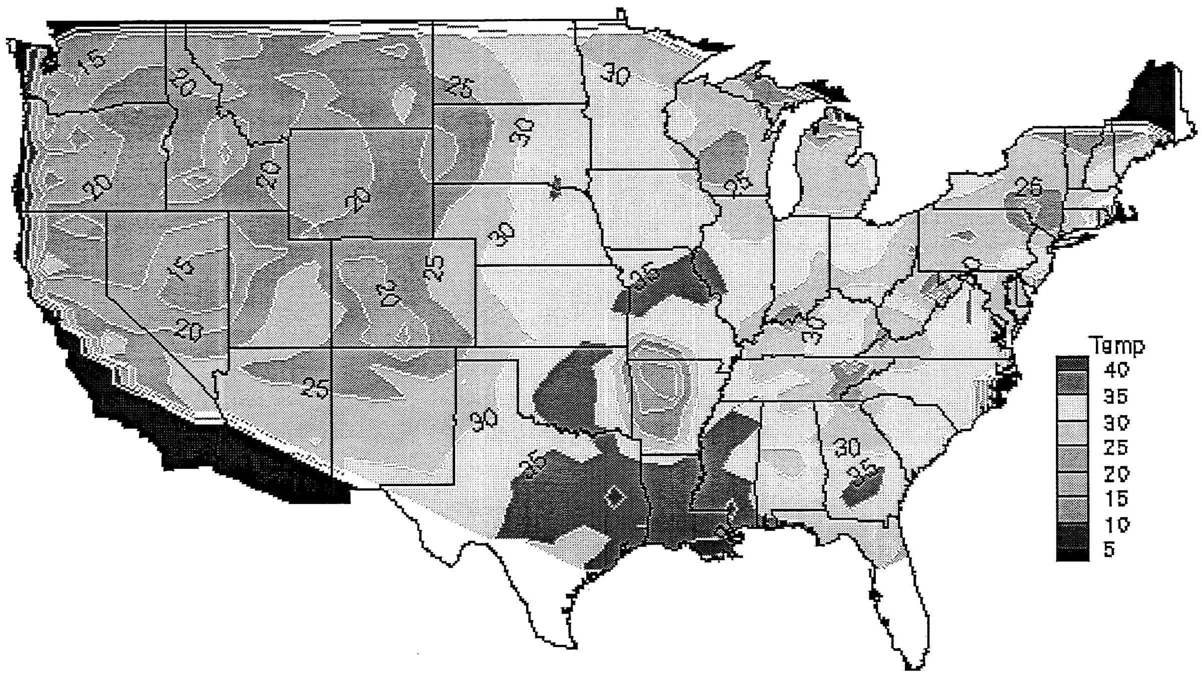


Figure 3.7. A plot showing predicted α's from Figure 3.6 (b) for the United States.

References

Brezonik, P. L. (1994). *Chemical Kinetics and Process Dynamics in Aquatic Systems*. Lewis Publishers, Ann Arbor, MI.

Eaton, J. G. and Scheller, R. M. (1996). Effects of climate on fish thermal habitat in streams of the united states. *Limnology and Oceanography*, 41:1109–1115.

Erickson, T. and Stefan, H. G. (2000). Linear air/water temperature correlations in streams during open water periods. *Journal of Hydrologic Engineering, ASCE*, 5(3):317–321.

Erickson, T. R., Mohseni, O., and Stefan, H. G. (1998). Estimation of an upper bound for weekly stream temperatures. Project Report No. 424, St. Anthony Falls Laboratory, Minneapolis, Minnesota.

Hershfield, D. M. (1961). Estimating the probable maximum precipitation. *Journal of the Hydraulics Division, ASCE*, 87(HY5):99–116.

Houghton, J., Filho, L. G. M., Callander, B., Harris, N., Kattenberg, A., and Maskell, K., editors (1996). *Climate Change 1995: The Science of Climate Change*. Cambridge University Press, Cambridge.

Lehninger, A. L. (1975). *Biochemistry*. Worth Publishers, New York.

Livingstone, D. M. (2001). Impact of secular climate change on the thermal structure of a large temperate central european lake. submitted to *Climactic Change*.

Mohseni, O., Erickson, T. R., and Stefan, H. G. (1997). A non-linear regression model for weekly stream temperatures at 585 gaging stations in the U.S. Project Report No. 401, St. Anthony Falls Laboratory, University of Minnesota, Minneapolis, MN.

Mohseni, O., Erickson, T. R., and Stefan, H. G. (1999). Sensitivity of stream temperatures in the u.s. to air temperatures projected under a global warming scenario. *Water Resources Research*, 35(12):3723–3733.

Mohseni, O., Erickson, T. R., and Stefan, H. G. (2001). Upper bounds of weekly stream temperatures. *Journal of Environmental Engineering, ASCE*. accepted.

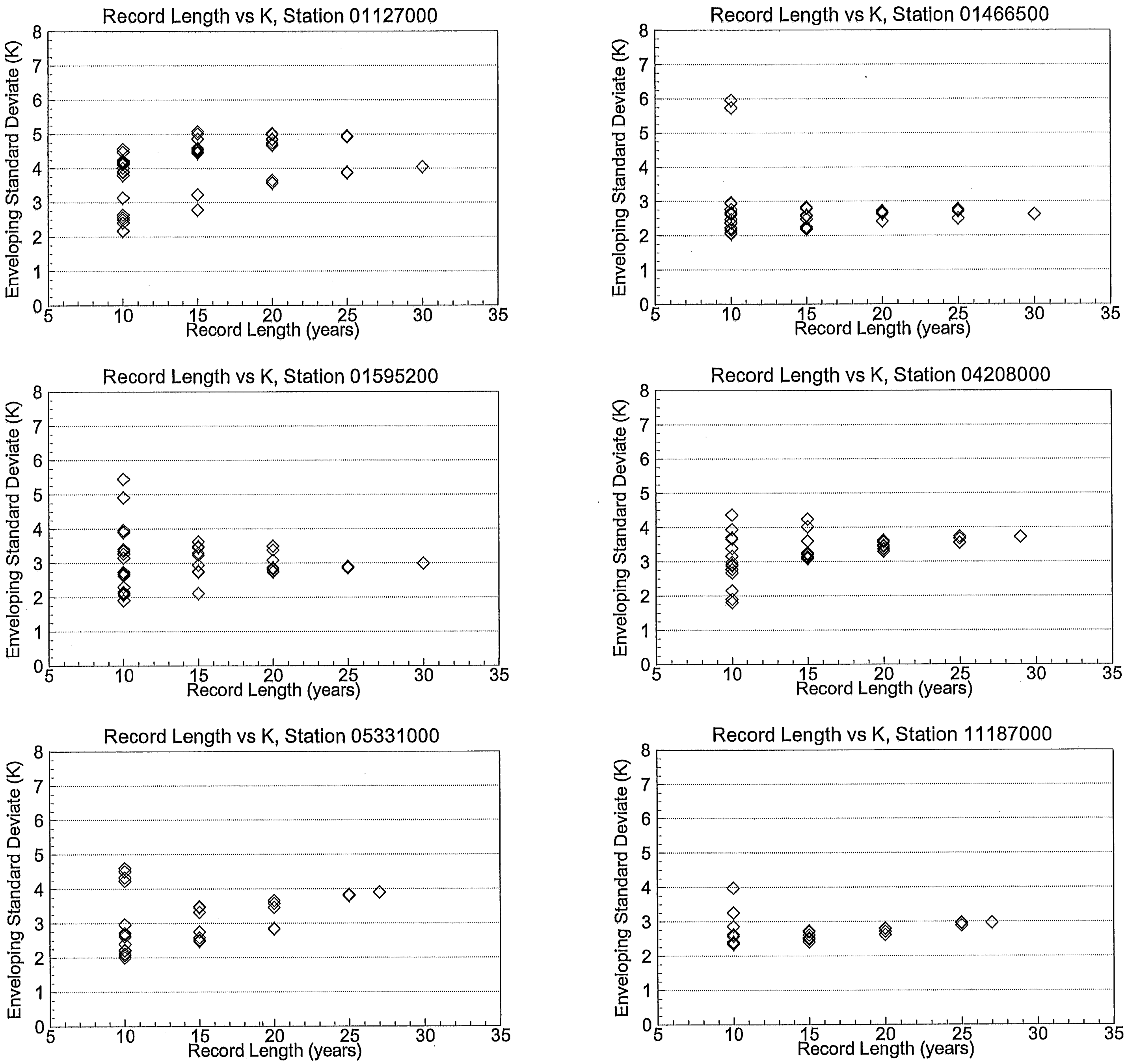
Mohseni, O. and Stefan, H. G. (1998). A non-linear regression model for weekly stream temperatures. *Water Resources Research*, 34(10):2685–2692.

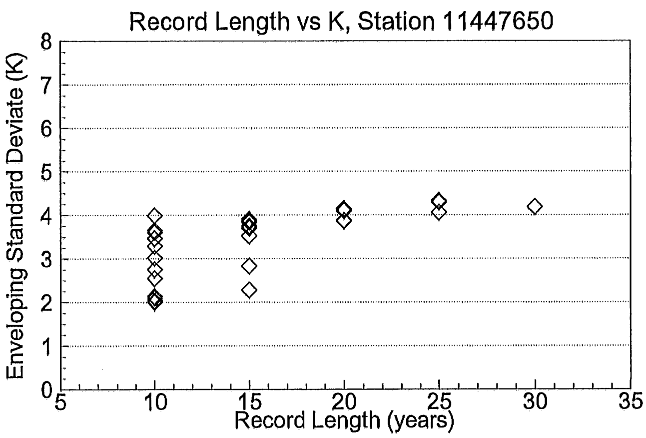
Mohseni, O. and Stefan, H. G. (1999). Stream temperature/air temperature relationship: A physical interpretation. *Jour. of Hydrology*, 218:128–141.

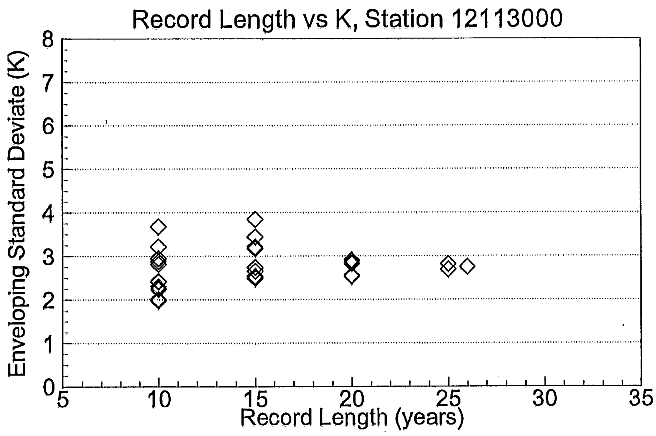
Sinokrot, B. A. and Stefan, H. G. (1993). Stream temperature dynamics: Measurements and modeling. *Water Resources Research*, 29(7):2299–2312.

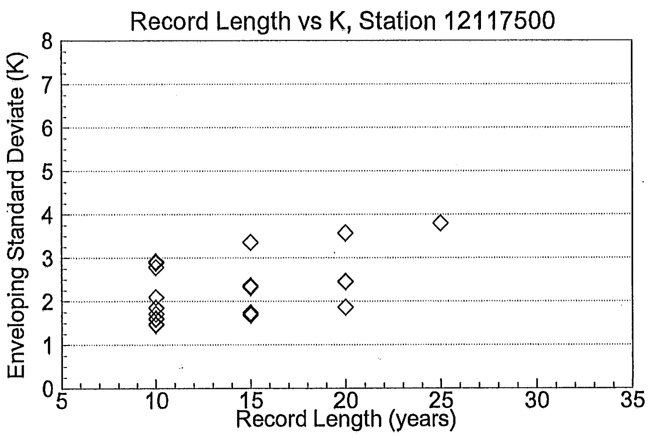
Stefan, H. G., Fang, X., and Eaton, J. G. (2001). Simulated fish habitat changes in north american lakes in response to projected climate warming. *Transactions of the American Fisheries Society*, 130:459–477.

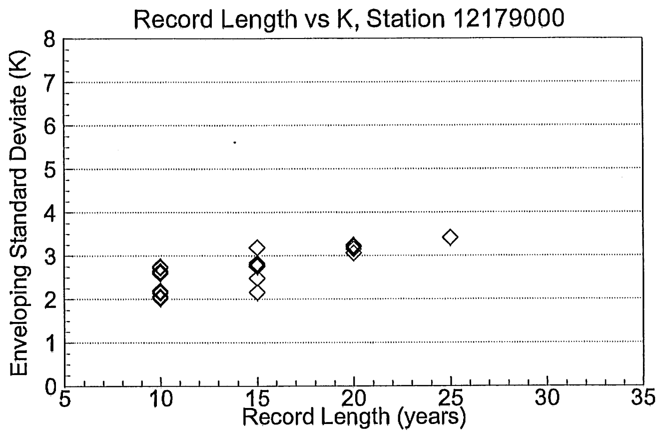
Appendix A: Plots of RL vs K for the Preliminary Analysis

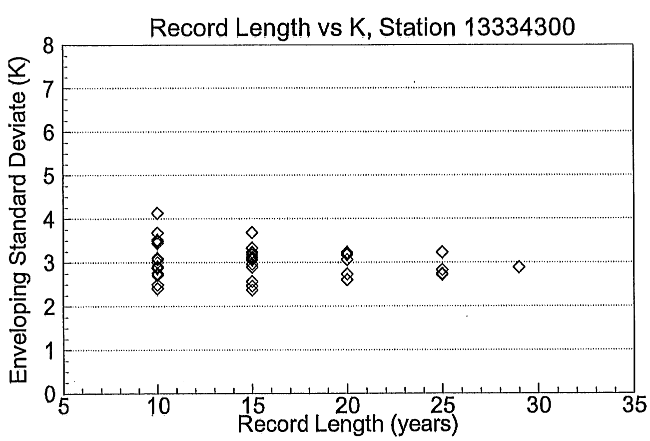


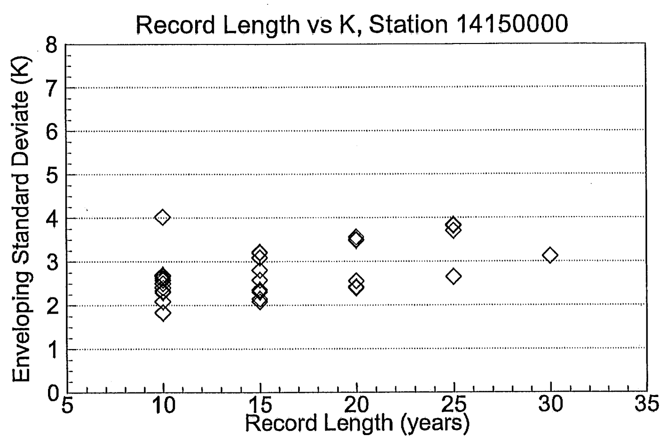

Record Length vs K, Station 11447650
Enveloping Standard Deviate (K)
Record Length (years)


Record Length vs K, Station 12113000
Enveloping Standard Deviate (K)
Record Length (years)

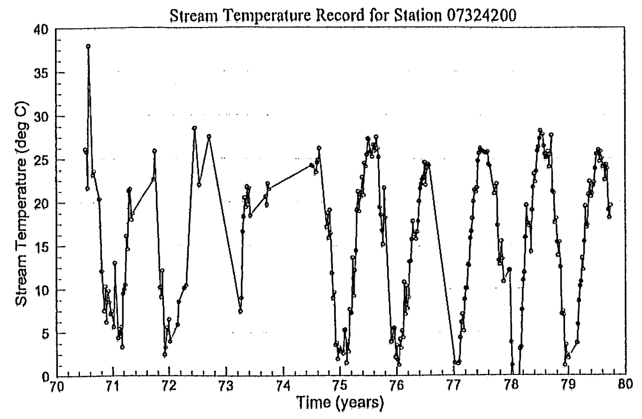

Record Length vs K, Station 12117500
Enveloping Standard Deviate (K)
Record Length (years)

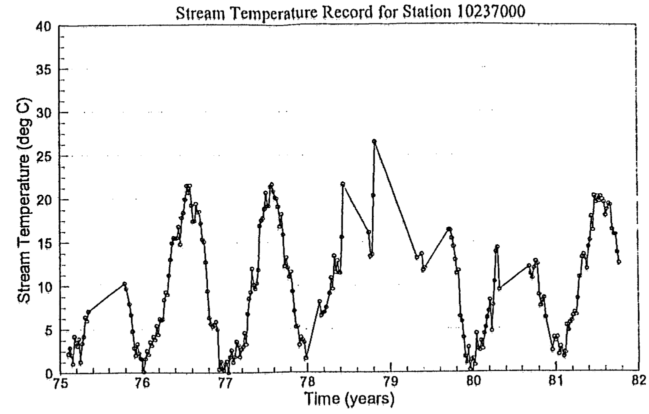

Record Length vs K, Station 12179000
Enveloping Standard Deviate (K)
Record Length (years)

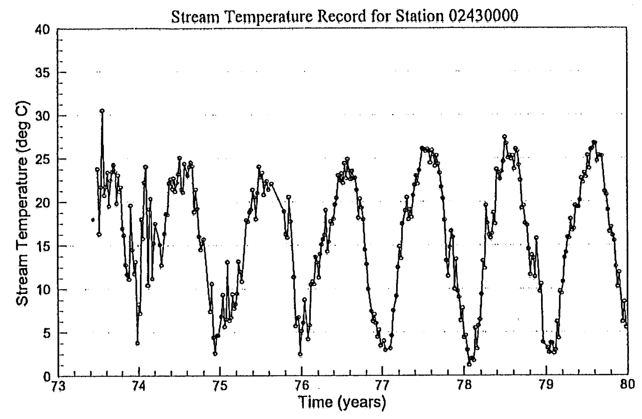

Record Length vs K, Station 13334300
Enveloping Standard Deviate (K)
Record Length (years)

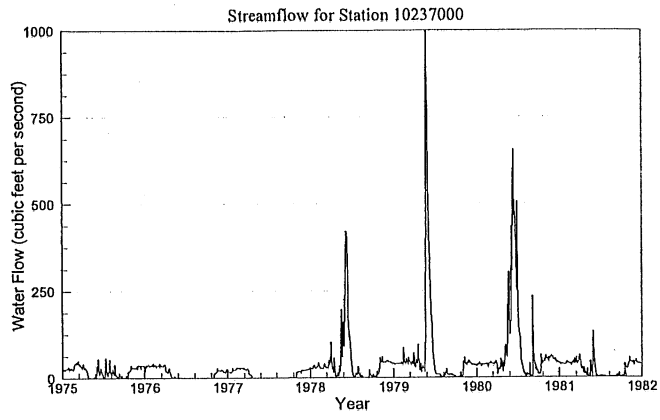

Record Length vs K, Station 14150000
Enveloping Standard Deviate (K)
Record Length (years)

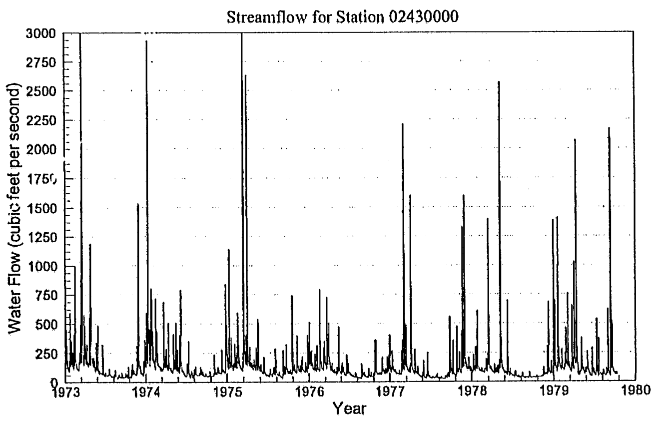
Appendix B: Time Series and Streamflows for Extreme Records

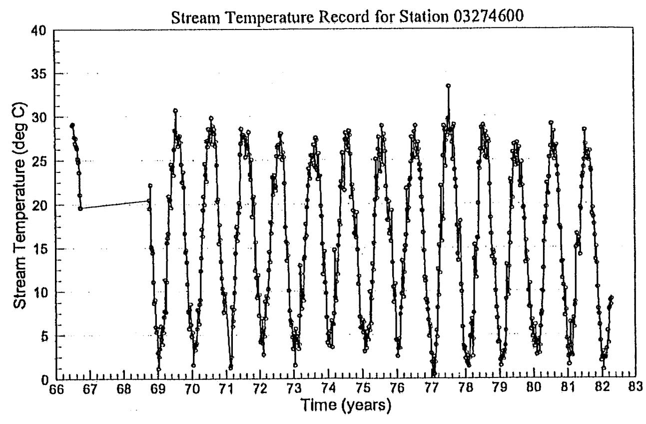


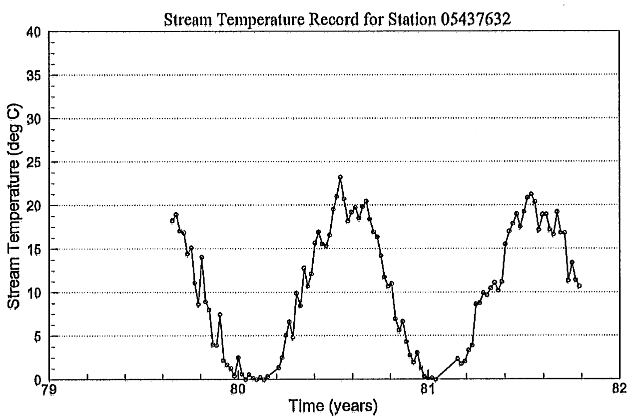


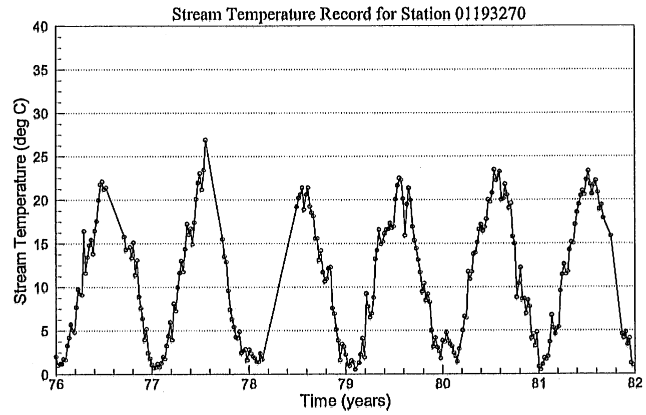


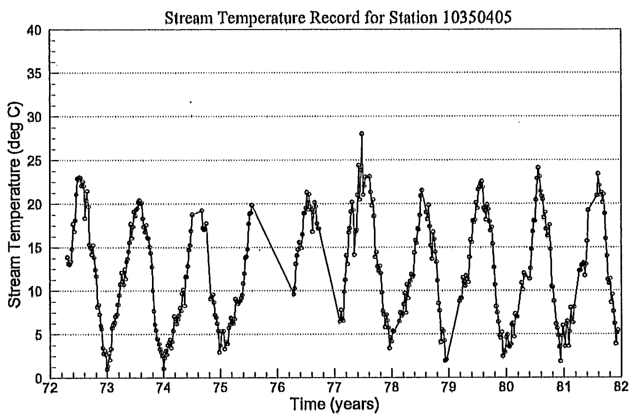


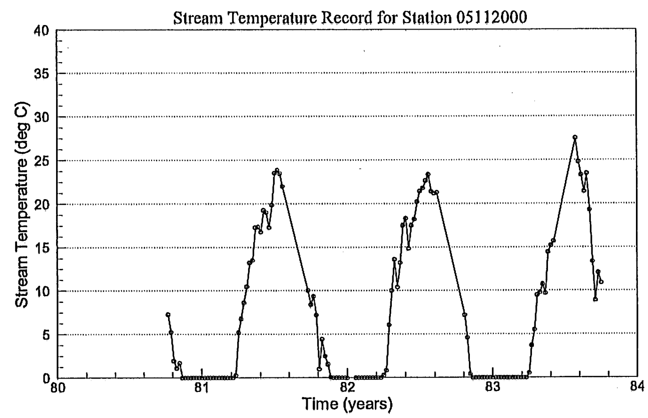


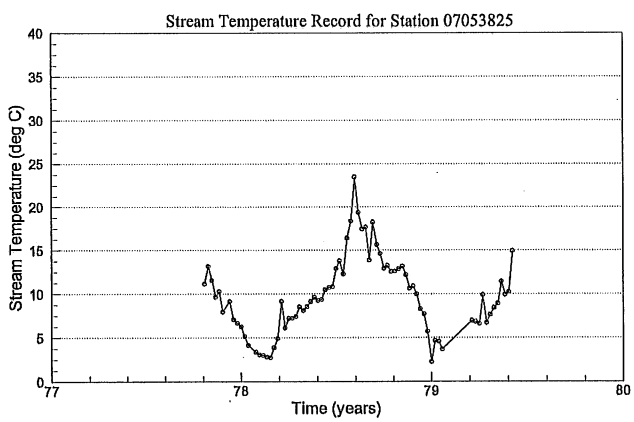


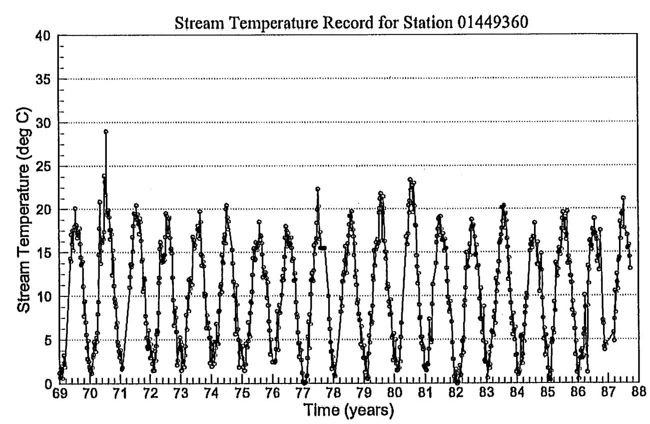

Stream Temperature Record for Station 05437632
Stream Temperature (deg C)
0 5 10 15 20 25 30 35 40
79 80 81 82
Time (years)

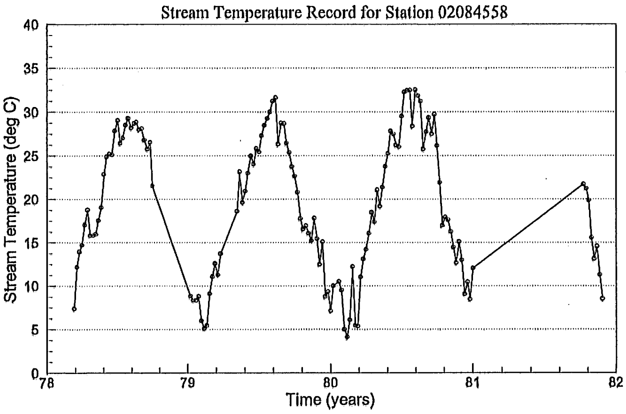

Stream Temperature Record for Station 01193270
Stream Temperature (deg C)
0 5 10 15 20 25 30 35 40
76 77 78 79 80 81 82
Time (years)

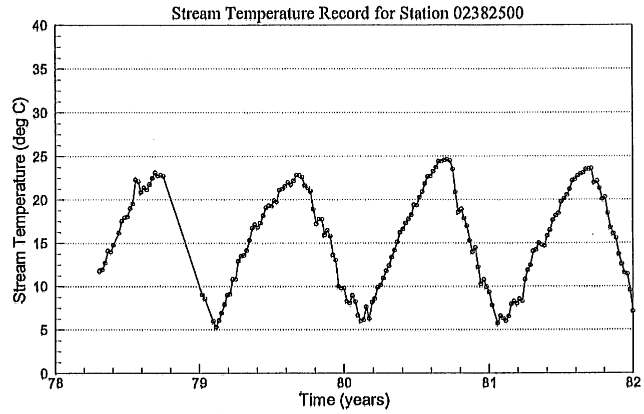

Stream Temperature Record for Station 10350405
Stream Temperature (deg C)
0 5 10 15 20 25 30 35 40
72 73 74 75 76 77 78 79 80 81 82
Time (years)

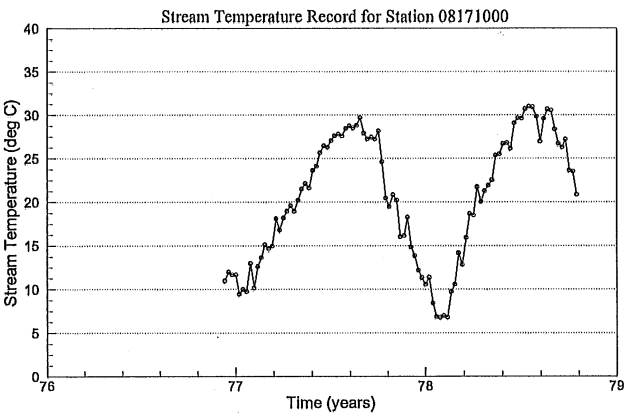

Stream Temperature Record for Station 05112000
Stream Temperature (deg C)
0 5 10 15 20 25 30 35 40
80 81 82 83 84
Time (years)

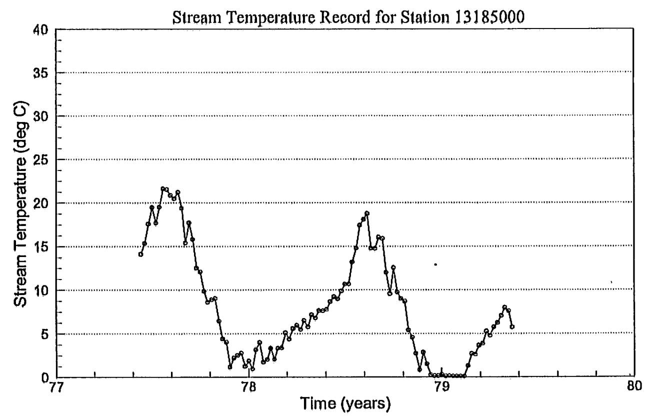

Stream Temperature Record for Station 07053825
Stream Temperature (deg C)
0 5 10 15 20 25 30 35 40
77 78 79 80
Time (years)

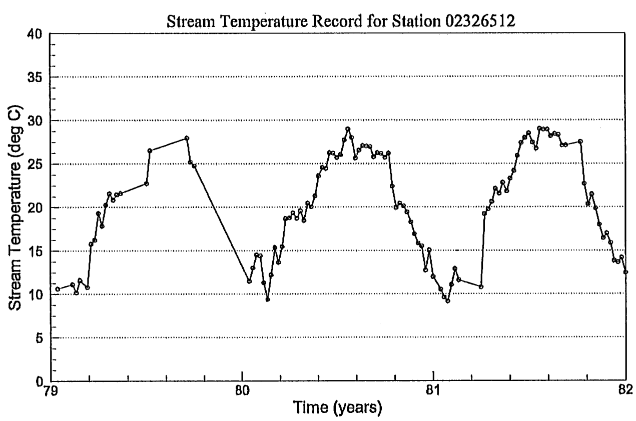

Stream Temperature Record for Station 01449360
Stream Temperature (deg C)
0 5 10 15 20 25 30 35 40
69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88
Time (years)


Stream Temperature Record for Station 02084558
Stream Temperature (deg C)
Time (years)
0
5
10
15
20
25
30
35
40
78
79
80
81
82


Stream Temperature Record for Station 02382500
Stream Temperature (deg C)
Time (years)
0
5
10
15
20
25
30
35
40
78
79
80
81
82


Stream Temperature Record for Station 08171000
Stream Temperature (deg C)
Time (years)
0
5
10
15
20
25
30
35
40
76
77
78
79


Stream Temperature Record for Station 13185000
Stream Temperature (deg C)
Time (years)
0
5
10
15
20
25
30
35
40
77
78
79
80


Stream Temperature Record for Station 02326512
Stream Temperature (deg C)
Time (years)
0
5
10
15
20
25
30
35
40
79
80
81
82