Detecting Changes through Time in the Variance of a Long-Term Hemispheric Temperature Record: An Application of Robust Locally Weighted Regression*

ANDREW R. SOLOW

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

(Manuscript received 18 May 1987, in final form 15 November 1987)

ABSTRACT

Historic records of global or hemispheric temperature are an important source of information about climate change. In order to analyze such records statistically, it is necessary to have some knowledge about the behavior of their variances through time. A modified version of robust locally weighted regression, a nonparametric regression procedure, is used to study the behavior of the variance for a record of Southern Hemisphere temperature deviations. Sampling considerations suggest that the variance should decrease through time, as new recording stations are added to the sampling network. Surprisingly, the variance is found to remain virtually constant through time.

1. Introduction

Historic records of air temperature are an important source of information about global climate change (Ellsaesser et al., 1986). Such records are useful, for example, in establishing empirical relationships between air temperature and other variables such as precipitation, the level of atmospheric carbon dioxide, or sea level. Perhaps more importantly, such records provide a relatively long-term baseline against which recent behavior can be measured.

A number of attempts have been made to construct long-term global or hemispheric temperature records. Many of these are discussed in the comprehensive review by Ellsaesser et al. (1986). In this paper, we will focus on the annual record of Southern Hemisphere temperature deviations constructed by Jones et al. (1986) and updated by Jones (1985). This record, which is described further in the next section, runs from 1858 to 1985 and is plotted in Fig. 1. Let Y_t be the value of this record in year t. A very general model for Y_t is

$$Y_t = M_t + e_t, \quad t = 1, \cdots, n \qquad (1)$$

where M_t is a smoothly varying, or systematic, component and e_t is an irregular component with mean zero and variance V_t. To estimate M_t efficiently, or to test an hypothesis about M_t, it is necessary to have some knowledge about V_t. For example, to test for climate change, Solow (in press) used a two-phase regression model that assumed constant, unknown V_t. As discussed in the next section, the behavior of V_t through time depends on both natural climate variability and the sampling used to construct the record. The purpose of this paper is to describe and apply a method for studying the behavior of V_t. This approach, which is based on robust locally weighted regression (Cleveland, 1979), is nonparametric in the sense that it is not necessary to specify a parametric form for M_t. In fact, one of the uses of this approach is to guide the choice of a parametric form for M_t.

In section 2, we discuss briefly the construction of the record shown in Fig. 1. In particular, we show that changes in sampling over time should act to reduce V_t. In section 3, we describe the approach that we will take to study the behavior of V_t through time. This approach is applied in section 4, where we find the surprising result that V_t remains virtually constant through time. This result is discussed in section 5.

2. The data

The annual record of Southern Hemisphere temperature deviations shown in Fig. 1 was compiled by Jones et al. (1986). The way in which the record (denoted as SH60) was constructed was described in some detail in the original paper. For the purposes of this paper, the important points are these. For a fixed month, the value of SH60 is formed by taking a weighted average of a number of station monthly averages. The weights in this average arise from a scheme by which the station data are interpolated onto a regular grid. Annual values of SH60 were found by averaging monthly values. The grid covers the Southern Hemisphere from 5° to 60°S, with a unit grid spacing of 5° latitude by 10° longitude. The number and spatial dis-

* Woods Hole Oceanographic Institution Contribution Number 6646.

Corresponding author address: Dr. Andrew Solow, Woods Hole Oceanographic Institution, Woods Hole, MA 02543.

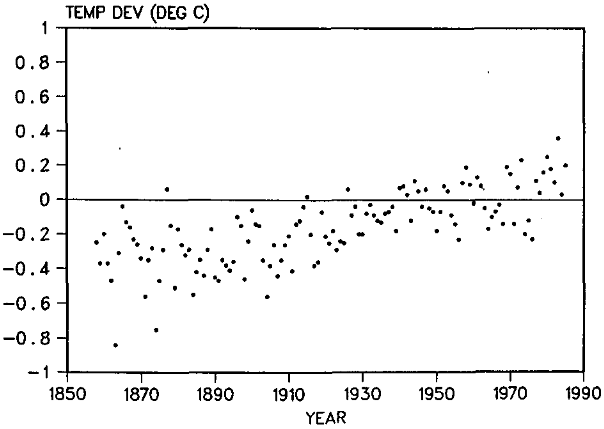

FIG. 1. Annual Southern Hemisphere surface air temperature deviations, 1858–1985 (Jones et al., 1986; Jones, 1985).

tribution of stations change with time. For example, only seven stations are available for 1858. By 1920, however, eighty stations are available. Rather than reporting the number of stations available in each year, Jones et al. (1986) reported the percentage of the area of the Southern Hemisphere accounted for by grid nodes onto which a value is interpolated. For example, the seven stations in 1858 account for only 1% of the Southern Hemisphere, while the eighty stations in 1920 account for 12%. The percent coverage over the entire period is shown in Fig. 2. We denote by C_t the percent coverage in year t.

Let Y_{it} be the annual average for station i in year t. Then:

$$Y_t = \sum_{i=1}^{p} w_{it} Y_{it} \quad t = 1, \cdots, n \tag{2}$$

where p is the total number of stations ever used to form the overall record and w_{it} is zero if station i is not included in the overall record in year t. The weights in (2) satisfy the side condition:

$$\sum_{i=1}^{p} w_{it} = 1, \quad t = 1, \cdots, n. \tag{3}$$

If we use the same general representation as (1) for the individual station records—that is,

$$Y_{it} = M_{it} + e_{it}, \quad i = 1, \cdots, p; \quad t = 1, \cdots, n, \tag{4}$$

then (2) can be written as

$$Y_t = \sum_{i=1}^{p} w_{it} M_{it} + \sum_{i=1}^{p} w_{it} e_{it} \quad t = 1, \cdots, n. \tag{5}$$

Note that this is equivalent to (1) with

$$M_t = \sum_{i=1}^{p} w_{it} M_{it}$$

$$e_t = \sum_{i=1}^{p} w_{it} e_{it}.$$

Finally, the variance of e_t is given by

$$V_t = \sum_{i=1}^{p} \sum_{j=1}^{p} w_{it} w_{jt} \operatorname{cov}(e_{it}, e_{jt}) \tag{6}$$

where cov denotes covariance. We stress that this representation of SH60, while exact, was not used explicitly by Jones et al. (1986) in constructing SH60.

The behavior of V_t through time depends on two factors: 1) the number and spatial distribution of stations receiving nonzero weights and 2) the covariances among the station records. The first of these factors is controlled by sampling, while the second is controlled by natural climatic variability. Without referring to the original station records, and the way in which they enter SH60, it is not possible to predict or explain fully the behavior of V_t. In broad terms, however, the substantial increase in coverage over the course of the record suggests that V_t should decline over time, provided the sampling effect is not offset by a change in climatic variability operating in the opposite direction.

3. Methodology

The general approach that we will take to study the behavior of V_t through time is to estimate M_t and to use the residuals to estimate V_t. Three complications arise, however. First, we do not have a parametric model for M_t, at least at the outset. Second, even if we had such a model, the efficient estimation of the parameters would require some knowledge of V_t. Third, even if M_t is known, in all generality it would not be possible to estimate V_t with any precision. This is because, without further restrictions on V_t, only a single term (i.e., e_t) would contain information about V_t. To address the first complication, we will use a robust, nonparametric method to estimate M_t. This method is called robust locally weighted regression (RLWR), and was developed by Cleveland (1979). Walden and

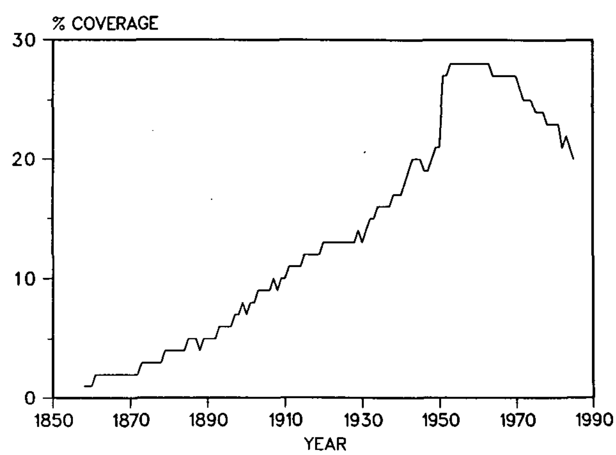

FIG. 2. Percent coverage, 1858–1985 (Jones et al., 1986; Jones, 1985).

Prescott (1983) used RLWR to estimate trends in annual maximum sea levels for several locations along the English coast. As its name suggests, RLWR has two important features. First, it is robust, in the sense that the influence of outlying observations (which are internally identified) is reduced. Second, like similar kernel smoothers (e.g., Gaussian filtering), RLWR allows the estimate of M_t to respond to the local behavior of the data, and does not require the choice of a model for M_t that is valid over the entire period of observation. As we have said, if such a global model is appropriate, RLWR can be useful in identifying a suitable parametric form. To address the second complication, we will take an iterative approach. Under this approach, the residuals from the current estimate of M_t are used to estimate V_t, and this estimate of V_t is used to update the estimate of M_t. To address the third complication, we will assume that any changes in V_t occur smoothly over time. In broad terms, the behavior of C_t shown in Fig. 2 suggests that this assumption is justified, at least in terms of the sampling contribution to V_t.

To describe RLWR, it is useful to suppose for the moment that V_t is constant through time. In that case, RLWR proceeds as follows:

(A1) For each t, fit a linear regression of Y_t on t by weighted least squares. The weight given to the observation $(t', Y_{t'})$ in fitting the regression at the point t is given by the tricube function:

$$w_{t'}(t) = \begin{cases} (1 - |(t - t')/h|^3)^3, & \text{if } |(t - t')/h| < 1 \\ 0, & \text{if } |(t - t')/h| \geqslant 1 \end{cases} \tag{7}$$

where h is called the span (or half-span). The estimate of M_t is the fitted value of the regression at t.

(A2) Let:

$$r_t = Y_t - \hat{M}_t, \quad t = 1, \cdots, n$$

where $\hat{M}_t$ is the estimate of M_t found from step (1), and let s be the median of the absolute values of the r_t. Define robustness weights by

$$d_t = \begin{cases} (1 - (r_t/6s)^2)^2 & \text{if } |r_t/6s| < 1 \\ 0 & \text{if } |r_t/6s| \geqslant 1 \end{cases} \tag{8}$$

$t = 1, \cdots, n$, and compute new estimates of M_t, $t = 1, \cdots, n$, by weighted linear regression with new weights $d_{t'}w_{t'}(t)$.

(A3) Iterate steps (A1) and (A2) a total of T times.

The use of the tricube weight function guarantees that, in estimating the value of the systematic component at t, observations near t receive relatively high weight, while observations greater than h away from t receive zero weight. This allows the estimate of M_t to respond to local changes in behavior. The use of the robustness weights reduces the influence of any outlying observations on the estimate of M_t. To apply RLWR, it is necessary to choose values for T and h. One way to choose T is to define a convergence criterion and to iterate until this criterion is satisfied. Cleveland (1979) suggested that this is needlessly complicated in most cases, and suggested that two iterations are usually sufficient. In any case, it is probably inadvisable to automate the procedure completely, and some degree of judgement should be used in its application. The choice of h is rather more important. Small values of h tend to attribute more of the detailed behavior of Y_t to the systematic component, while larger values of h tend to attribute more to the irregular component. Cleveland (1979) suggested the use of the PRESS procedure (Allen, 1974; Cook and Weisberg, 1982) to choose h. This procedure is based on cross-validation, and works in the following way:

(B1) Let $M_t^*(h)$ be the locally weighted regression estimate of M_t, excluding the observation (t, Y_t) and excluding the robust weighting step, for span h. The initial value of h—say, h_0—is chosen by minimizing

$$\sum_{t=1}^{n} (Y_t - M_t^*(h))^2 \tag{9}$$

over h.

(B2) Let d_t, $t = 1, \cdots, n$, be the robustness weights for the residuals from the locally weighted regression fit with $h = h_0$, and let $M_t^{**}(h)$ be the RLWR estimate of M_t, excluding the observation (t, Y_t) and using the d_t for robustness weights, for span h. The next value of h—say, h_1—is chosen by minimizing

$$\sum_{t=1}^{n} d_t(Y_t - M_t^{**}(h))^2. \tag{10}$$

(B3) Iterate steps (B1) and (B2) a total of T' times.

Again, the number of iterations, T', can be chosen either automatically, by applying a convergence criterion, or by a simpler ad hoc procedure.

We have assumed so far that V_t is constant through time. Of course, our primary concern is with detecting and estimating nonconstant V_t. While it would be possible to devise a procedure for doing this that is based on the residuals from the final RLWR estimate of M_t, the possibility that V_t changes through time has two implications for RLWR itself. First, terms like $(Y_t - M_t^*)^2$ appearing in minimization steps should be rescaled to take account of changes in V_t. Failure to rescale the residuals in this way tends to give periods of high V_t undue influence in choosing h. Second, the scale measure s used in the robustness step should be adapted to reflect local changes in scale. Failure to rescale s locally tends to underweight nonoutliers in periods of high V_t and overweight outliers in periods of low V_t.

Before describing a modification of RLWR that allows for nonconstant V_t, we make the following remark. At the completion of a full iteration in RLWR, the estimate of the systematic component at time t—which we will refer to generically as $\hat{M}_t$—has the form:

$$\hat{M}_t = \sum_{i=1}^{n} a_{it} Y_i, \quad t = 1, \cdots, n. \tag{11}$$

The weights in this linear combination depend on the tricube weights, the robustness weights, and the position of t among the points receiving nonzero weight in the estimation of M_t. Details are given in the Appendix. If $\hat{M}_t$ is close to M_t, then the residual:

$$r_t = Y_t - \hat{M}_t$$

$$\approx e_t - \sum_{i=1}^{n} a_{it} e_i, \quad t = 1, \cdots, n$$

has variance

$$\text{var} r_t \approx V_t(1 - 2a_{tt}) + \sum_{i=1}^{n} a_{it}^2 V_i \tag{12}$$

provided that the e_t are uncorrelated. That is, the variance of r_t depends not only on V_t, but on V_i, where the observation Y_i enters $\hat{M}_t$ with nonzero weight. It can be shown, however, that

$$\sum_{i=1}^{n} a_{it}^2 V_i \approx \sum_{i=1}^{n} a_{it}^2 V_t$$

and consequently,

$$\text{var} r_t \approx V_t(1 - 2a_{tt} + \sum_{i=1}^{n} a_{it}^2). \tag{13}$$

Since the a_{it} are known, the r_t can be rescaled according to

$$r_t^* = r_t/(1 - 2a_{tt} + \sum_{i=1}^{n} a_{it}^2)^{1/2} \tag{14}$$

so that

$$\text{var} r_t^* \approx V_t.$$

We note in passing that the denominator in (14) tends to one fairly rapidly as h increases, so that in many cases the rescaling step is not needed. This same remark applies when the e_t are serially correlated.

We next describe a modification of RLWR to allow for nonconstant V_t. This modification is made by including the following steps at the conclusion of each iteration of RLWR:

(C1) Let $\hat{M}_t$, $t = 1, \cdots, n$, be the current estimate of M_t. Form the absolute values of the rescaled residuals, r_t^*, $t = 1, \cdots, n$, defined in (14).

(C2) Apply the RLWR algorithm to $|r_t^*|$, $t = 1, \cdots, n$. Let u_t, $t = 1, \cdots, n$, be the estimate of the systematic component of $|r_t^*|$, $t = 1, \cdots, n$.

(C3) In the next iteration of RLWR applied to the original data:

(i) replace s by u_t in step A2

(ii) divide $(Y_t - M_t^{**})$ by u_t in step B2.

We make two remarks about this procedure. First, for moderate departures from constant V_t, the final estimate of M_t from the modified procedure is unlikely to be very different from the final estimate from the unmodified procedure. The importance of taking account of nonconstant V_t arises when we attempt to construct confidence intervals or test an hypothesis. For this reason, in exploratory analysis, it may be sufficient to use the unmodified procedure to estimate M_t and to apply RLWR to the absolute values of the rescaled residuals from the final estimate (Cleveland, 1980). If V_t is found to change through time, the modified procedure can then be employed. Second, instead of applying RLWR to the absolute values of the rescaled residuals, we could apply it to the squares of the rescaled residuals. We prefer the former for reasons of robustness (e.g., see Huber, 1981). Because the absolute value of the rescaled residual is a scale measure with the same units as Y_t, it will behave like $V_t^{1/2}$, and not like V_t itself. For example, for normally distributed rescaled residuals, the expected value of u_t is (asymptotically) $(V_t\pi/2)^{1/2}$.

Summarizing, the modified RLWR algorithm proceeds by the following steps:

1) choose h_0 by the PRESS procedure
2) find the initial RLWR estimate of M_t, $t = 1, \cdots, n$
3) apply RLWR to the absolute values of the residuals
4) iterate using the current u_t, $t = 1, \cdots, n$, to define the robustness weights and to scale the squared error terms used in the PRESS procedure.

4. Application

In this section, we apply the modified RLWR procedure to SH60. The procedure converged after two iterations, with only slight differences between the final results and the results of the first iteration. The span chosen by the PRESS procedure for the temperature data was 29 years. Only three observations received robustness weights less than 0.5. These were 1863 with robustness weight 0.098, 1874 with robustness weight 0.307, and 1877 with robustness weight 0.446. The final estimate of M_t, $t = 1, \cdots, n$, is plotted in Fig. 3. In broad terms, this estimate declines slightly from the beginning of the record to around 1890. From around 1890 to around 1960, the estimate increases at a decreasing rate. After around 1960, the estimate increases at an increasing rate until the end of the record. Although we are only peripherally concerned with estimating M_t in this paper, we note that care should be taken in interpreting the behavior of the estimate at the ends of the observation period, where the data neighborhood is asymmetrical.

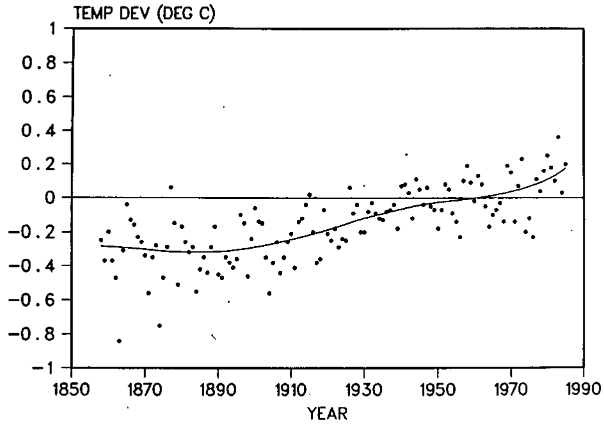

FIG. 3. Modified robust locally weighted regression estimate of M_t (h = 29).

Once the final modified RLWR estimate of M_t was found, RLWR was applied to the absolute value of the rescaled residuals. Again, the procedure converged after two iterations. The span chosen by the PRESS procedure was 34 years. Four observations received robustness weights less than 0.5. These were 1863 with zero weight, 1874 with robustness weight 0.034, 1877 with robustness weight 0.189, and 1976 with robustness weight 0.444. The observation from 1976 received robustness weight 0.645 in the final estimation of M_t. The final estimate of the systematic component in the absolute rescaled residuals is plotted in Fig. 4. In light of the substantial variation of coverage through time shown in Fig. 2, what is most remarkable about Fig. 4 is the essentially constant scale of the residuals through time. It is possible to discern a slight increase in scale from the beginning of the record to around 1900, a slight decrease to around 1935, and a slight increase to the end of the record. However, the magnitude of this variation is so small that the assumption of constant scale is unlikely to pose any problems for further analysis. On the other hand, the identification of a small number of observations as outlying suggests that care be taken in applying nonrobust statistical procedures to this data.

The analysis so far assumes that the final estimate of M_t is close to the true systematic component of Y_t. One possible explanation of the result presented here is that the modified RLWR estimate of M_t is too smooth. In that case, part of the systematic behavior of Y_t is incorrectly attributed to the residuals. If the behavior of the residuals is dominated by this component, then it is possible that the true behavior of V_t is masked. Because the amount of smoothing depends on h, to investigate this possibility we applied the procedure with $h = 5$. The final estimate of M_t is shown in Fig. 5. As expected, this estimate is not as smooth as that shown in Fig. 3. In particular, there is evidence of a periodic component in M_t. Incidentally, note that the apparent acceleration at the end of SH60 shown in Fig. 3 vanishes in Fig. 5. The absolute rescaled residuals from the estimate of M_t shown in Fig. 5 are shown in Fig. 6, along with the RLWR estimate of their systematic component. As expected, there is a tendency for these residuals to be smaller in magnitude than those shown in Fig. 4. However, the behavior through time of the RLWR estimate is essentially identical to that shown in Fig. 4. Although this experiment is not conclusive, it strongly suggests that the main result of this section is not due to oversmoothing.

It would be interesting to repeat this analysis for a similar record of Northern Hemisphere temperatures. Although we will not present such an analysis here, in preliminary work using other methods Grotch (personal communication) found a similar result.

5. Discussion

The main result of the previous section—namely, that V_t remains approximately constant through time—

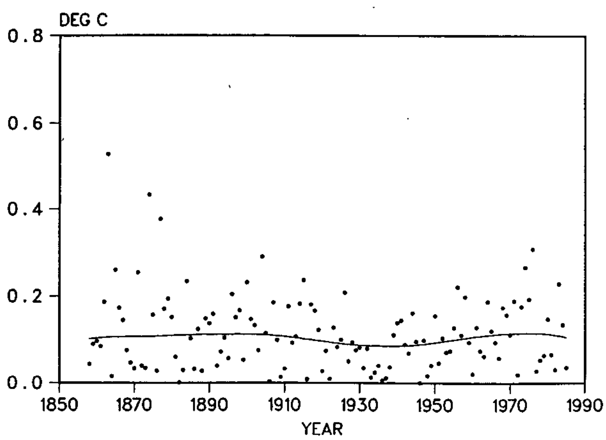

FIG. 4. Robust locally weighted regression estimate of the systematic component of the absolute rescaled residuals for the estimate of M_t shown in Fig. 3.

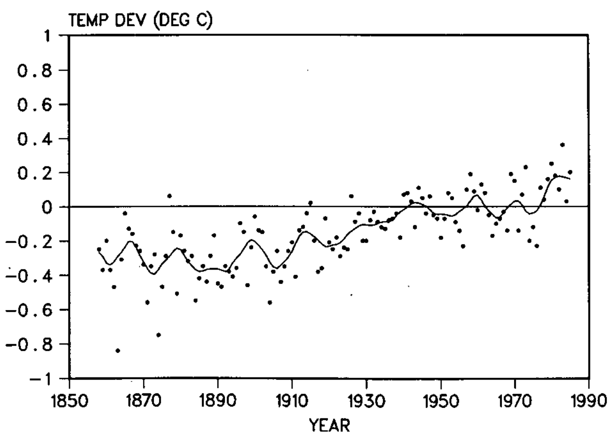

FIG. 5. Modified robust locally weighted regression estimate of M_t (h = 5).

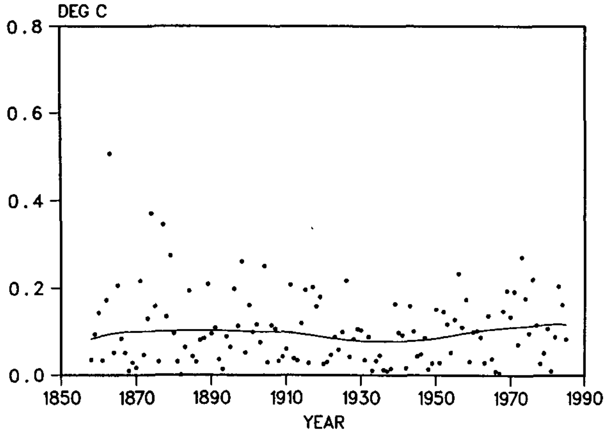

FIG. 6. Robust locally weighted regression estimate of the systematic component of the absolute rescaled residuals for the estimate of M_t shown in Fig. 5.

is quite surprising in light of the substantial increase in network coverage through time. This result is important by itself, since it supports the assumption of constant variance made in statistical analyses of SH60 [e.g., Solow (in press)]. On the other hand, because the result is so surprising, it calls for an explanation. As we have said, it is not possible to explain fully the behavior of V_t without referring to the original station records and the way in which they enter SH60. Because we will not attempt such an analysis here, the following remarks, which are based on variants of a simple model, are necessarily speculative. Nevertheless, it is instructive to consider the implications for the behavior of V_t of such a model, if only to guide a later, more comprehensive analysis.

In section 2, we showed that

$$V_t = \sum_{i=1}^{p} \sum_{j=1}^{p} w_{it} w_{jt} \operatorname{cov}(e_{it}, e_{jt}), \quad t = 1, \cdots, n$$

where w_{it} is the weight received by station i in calculating Y_t. These weights satisfy the side condition:

$$\sum_{i=1}^{p} w_{it} = 1 \quad t = 1, \cdots, n.$$

It is useful to express the covariance terms as

$$\operatorname{cov}(e_{it}, e_{jt}) = (V_{it} V_{jt})^{1/2} c_{ij,t} \tag{15}$$

where V_{it} is the variance of e_{it} and $c_{ij,t}$ is the correlation between e_{it} and e_{jt}.

We remarked in section 2 that, broadly speaking, V_t is determined by two factors: the sampling and interpolation scheme, which determine the weights, and natural climatic variability, which determines the covariance terms. Again in broad terms, the effect of the change in sampling through time, as reflected in C_t, is to reduce V_t. Since C_t increases from 1% to 27%, before declining to 20%, this effect should be substantial. Therefore, to explain the results of the previous section, it is necessary to identify an offsetting effect in natural climatic variability.

Under the simplest model:

1) The individual station variances are all equal and do not change through time: i.e., $V_{it} = V_{jt} = V$.
2) The individual station records are uncorrelated: i.e., $c_{ij,t} = 0$, $i \neq j$.
3) The number of stations receiving nonzero weight in Y_t, which we denote by p_t, is proportional to the coverage, C_t, and the station weights are all equal (and, therefore, proportional to C_t^{-1}).

For this simple model, V_t is proportional to C_t^{-1}. From the behavior of C_t shown in Fig. 2, u_t (which should behave like $V_t^{1/2}$) should decrease by a factor of approximately five over the course of the record. Clearly, the behavior of u_t shown in Fig. 4 is inconsistent with this simple model.

The simple model can be made more complicated (and more realistic) in a number of ways. We will briefly consider three of these.

First, the individual station variances can be allowed to differ. To explain the results of the previous section, however, it would be necessary to assume that the more variable station records were added to the sampling network later. In fact, to the extent that marine stations were added later, the opposite effect would be expected.

Second, the individual station variances can be allowed to change through time. Again, to explain the results of the previous section, it would be necessary to assume that these variances increased with time. This would be an interesting result in itself.

Third, the station records can be allowed to be correlated. This is quite likely. In fact, in a certain sense the gridding strategy used by Jones, et al. (1986) implicitly assumes some kind of positive spatial correlation that dampens with distance. The existence of spatial correlation of this type implies that V_t should decrease less rapidly than p_t^{-1}. The strength of this effect depends on the strength of spatial correlation. For example, in the case where $c_{ij,t} = 1$, for all i, j, and t, V_t will not change through time (provided the individual station records have the same variance and this variance does not change through time). This extreme case implies the existence of a single, hemispherewide noise process. On the other hand, to the extent that $c_{ij,t}$ declines with the distance between stations i and j, and that the substantial increase in C_t indicates that the distances between stations are large, this effect is limited.

The interplay among p_t, C_t and V_t is another area for further study. The question of the existence of spatial correlation becomes important here, too. For example, in the absence of spatial correlation, the weights on the individual station records should depend only on their variances (i.e., gridding the data is not sensible), and p_t is more important than C_t in determining V_t.

On the other hand, in the presence of spatial correlation, the weights on the individual station records should take account of their geographical configuration (i.e., gridding the data is sensible), and C_t is more important than p_t in determining V_t.

Finally, throughout this discussion we have assumed that any variability associated with measurement error is dominated by natural climatic variability. This is perhaps an unrealistic assumption. It may be, for example, that variability due to measurement error has changed through time, as new methods and technologies become available. If this is so, however, it seems reasonable to suppose that improvements in temperature measurement would tend to reduce variability, and consequently this effect would operate in the same direction as the sampling effect.

Acknowledgments. This work was supported by The Pew Memorial Trust and the Marine Policy Center of the Woods Hole Oceanographic Institution. The author wishes to thank J. M. Broadus of the Marine Policy Center for his help.

APPENDIX

Derivation of the Weights in Expression (11)

In this Appendix, we outline the derivation of the weights, a_{it}, in (11). Suppose that we are estimating M_t—that is, the value of the systematic component at time t. This involves fitting a weighted linear regression. The weight on observation $(t', Y_{t'})$ in this regression is given by the product of the current robustness weight, $d_{t'}$, and the current localization weight, $w_{t'}(t)$. Let:

$$w^*_{t'}(t) = d_{t'} w_{t'}(t), \quad t' = 1, \cdots, n$$

and define the following matrices:

$$\mathbf{X}^{\mathrm{T}} = \begin{cases} 1 & 1 \cdots 1 \\ 1 & 2 \cdots n \end{cases}$$

$$\mathbf{Y}^{\mathrm{T}} = (Y_1 Y_2 \cdots Y_n)$$

$$\mathbf{W}^*(t) = \mathrm{diag}(w^*_{t'}(t)) \quad t' = 1, \cdots, n$$

where the superscript T denotes transpose. The two parameter estimates for this linear regression are given by

$$b(t) = (\mathbf{X}^{\mathrm{T}}\mathbf{W}^*(t)\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}^*(t)\mathbf{Y}$$

where the first element of $b(t)$ is the estimated intercept and the second is the estimated slope. Let

$$\mathbf{M}(t) = \mathbf{X}b(t).$$

That is $\mathbf{M}(t)$ is the n-vector of regression fits. Note that the tth element of $\mathbf{M}(t)$ is used as the estimate of M_t. The other elements of $\mathbf{M}(t)$ are discarded. Finally, let

$$\begin{aligned} r(t) &= Y - \mathbf{M}(t) \\ &= (\mathbf{I} - \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{W}^*(t)\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}^*(t))\mathbf{Y} \\ &= (\mathbf{I} - \mathbf{H}(t))\mathbf{Y} \end{aligned}$$

be the n-vector of residuals from the regression fits, where $\mathbf{I}$ is the $n \times n$ identity matrix. The $n \times n$ matrix $\mathbf{H}(t)$ is called the hat matrix associated with the regression (Cook and Weisberg, 1982). The elements of the tth row of $(\mathbf{I} - \mathbf{H}(t))$ give the weights, a_{it}, $i = 1, \cdots, n$, used in (11).

REFERENCES

Allen, D. M., 1974: The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, **16**, 125–127.

Cleveland, W. S., 1979: Robust locally weighted regression and smoothing scatterplots. *J. Amer. Stat. Assoc.*, **74**, 829–836.

——, 1980: Challenging assumptions in time series models. *Technometrics*, **22**, 479–481.

Cook, R. D., and S. Weisberg, 1982: *Residuals and Influence in Regression.* Chapman and Hall, 230 pp.

Ellsaesser, H. W., M. C. MacCracken, J. J. Walton and S. L. Grotch, 1986: Global climatic trends as revealed by the recorded data. *Rev. Geophys.*, **24**, 745–792.

Huber, P. J., 1981: *Robust Statistics.* Wiley and Sons, 308 pp.

Jones, P. D., 1985: Southern Hemisphere temperatures: *Climate Monitor*, **14**, 132–140.

——, S. C. B. Raper and T. M. L. Wigley, 1986: Southern Hemisphere surface air temperature variations: 1851–1984. *J. Climate Appl. Meteor.*, **25**, 1213–1230.

Solow, A. R., 1988: Testing for climate change: An application of the two-phase regression model. *J. Climate Appl. Meteor.*, in press.

Walden, A. T., and P. Prescott, 1983: Identification of trends in annual maximum sea levels using robust locally weighted regression. *Estuar. Coastal Shelf Sci.*, **16**, 17–26.