Covariance Matrices and Means of Atmospheric Planck Function Profiles for Application to Temperature Sounding from Satellite Measurements

D. S. CROSBY,[1] H. E. FLEMING AND D. Q. WARK

National Environmental Satellite Service, NOAA, Washington, D. C. 20233

(Manuscript received 12 June 1972, in revised form 8 September 1972)

ABSTRACT

The statistical minimum-rms inversion method used to obtain temperature profiles, requires estimates of covariance matrices and means for Planck function profiles of the atmosphere. In order to obtain these estimates over the pressure range of 1000 to 0.01 mb, it was necessary to combine data from temperature measurements by radiosondes, rocketsondes and grenadesondes. Radiosonde data reaching the 10-mb level were extended to higher levels by means of a modified regression technique. Matrices and means have been obtained by this method for seasonal and geographical groupings in the Northern Hemisphere and the tropics. Details of the geographical and time changes in the matrices and the means are presented.

1. Introduction

The practical use of infrared (IR) measurements from satellites to infer vertical temperature profiles has been demonstrated with the Satellite Infrared Spectrometers (SIRS) on the Nimbus 3 and 4 by Wark (1970) and Smith *et al.* (1970). Encouraged by these results, the Global Atmospheric Research Project (GARP) and the World Weather Watch (WWW) have adopted this means for obtaining global measurements of the free air temperatures. Requirements for coverage and accuracy have been set forth by GARP (1969).

A standard technique for obtaining vertical atmospheric temperature profiles from the satellite measurements of IR radiances is the statistical minimum-rms inversion method. Details are given by Alishouse *et al.* (1967), Foster (1961), Rodgers (1970), Strand and Westwater (1968), and Wark and Fleming (1966). It is shown by these authors that it is necessary to have estimates of the covariance matrices and means in terms of the equivalent Planck radiance profiles. The emphasis of this paper is on this estimation problem and some of the results from the techniques used.

The inversion method requires knowledge of the temperature profiles, and hence Planck profiles, over the pressure range of 1000 mb to 0.01 mb (surface to about 80 km). However, direct measurements of temperature can be achieved only over limited portions of this range by a single instrument. In order to obtain a complete profile, temperature measurements by radiosondes, rocketsondes, and grenadesondes must be pieced together. Another complication immediately arises because measurements by these devices are seldom coincident in space and time. Nevertheless, it is shown that one can make a reasonable estimate of the covariance matrix by a modified regression technique, while the mean is pieced together by translation.

Individual temperature profiles are continuous functions of pressure, but they can be approximated by vectors whose elements are temperatures at prescribed pressure levels. The pressure slicing is in equal increments with respect to pressure to the two-sevenths power, which has been selected because of the optical characteristics of the atmosphere. Serial numbering of these incremental pressures is from 1 at 0.01 mb to 100 at 1000 mb. The final step in the preliminary preparation of the data is to transform the temperature vectors into Planck radiances at the arbitrary frequency of 700 cm^{-1}.

With 100 pressure levels, it is necessary to estimate a 100×100 covariance matrix and a 100×1 mean vector. If we let the Planck radiances be represented by $\mathbf{X}$, then estimates are needed for $E[\mathbf{X}]=\boldsymbol{\mu}$, the mean vector, and $E[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T]=\boldsymbol{\Sigma}$, the covariance matrix. Throughout, T denotes the transpose.

2. Estimation of the covariance matrices and means

Radiosonde data are generally available from the surface (the 100th pressure level) to 10 mb (the 25th pressure level); rocketsonde data are relatively complete from 25 mb (level 33) to 0.7 mb (level 10); and grenadesonde data are available from 2.6 mb (level 16) to 0.01 mb (the 1st level). The problem is to devise some technique for combining these data to obtain estimates for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

Let $\hat{\boldsymbol{\Sigma}}=[\sigma_{ij}]$ be the estimate for the covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{X}$. In order to obtain the estimators

[1] The American University, Washington, D. C.

Unauthenticated | Downloaded 03/29/23 12:17 AM UTC

σ_{ij}, $i,j=25,\ldots,100$, standard techniques are applied to the radiosonde data. That is,

$$\sigma_{ij}=\frac{1}{N-1}\sum_{k=1}^{N}[x_i(k)-\bar{x}_i][x_j(k)-\bar{x}_j], \tag{1}$$

where $25\leqslant i\leqslant 100$, $25\leqslant j\leqslant 100$, N is the sample size of those vectors that have both the ith and jth levels, and $\bar{x}_i$ and $\bar{x}_j$ are the usual estimators for the means. Let $\mathbf{\Sigma}^*$ be the matrix of these elements.

For notational convenience, let the rocketsonde data be represented by sub-vectors of the form

$$\mathbf{Z}(k)=[x_{10}(k),\ldots,x_{24}(k)]^T,$$
$$\mathbf{Y}(k)=[x_{25}(k),\ldots,x_{33}(k)]^T,$$

so that $[\mathbf{Z}(k)^T,\mathbf{Y}(k)^T]^T$ is the kth vector of rocketsonde data. Furthermore, let $\mathbf{B}$ represent the estimated regression coefficients of $\mathbf{Z}$ on $\mathbf{Y}$, and R_{ZY} the estimated residual variance of $\mathbf{Z}$, given $\mathbf{Y}$.

These matrix components are now used to compute the next set of entries in the estimator $\hat{\mathbf{\Sigma}}$, beyond those already given by (1). Define the sub-matrices $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$, $\mathbf{C}_4$ as follows:

$$\mathbf{C}_1=[\sigma_{ij}]_{i,j=10}^{24},\qquad \mathbf{C}_2=[\sigma_{ij}]_{i,j=25}^{33}$$
$$\mathbf{C}_3=[\sigma_{ij}]_{i=10,\ j=25}^{24,100},\qquad \mathbf{C}_4=[\sigma_{ij}]_{i=25,\ j=25}^{33,100}$$

The elements of the sub-matrices $\mathbf{C}_2$ and $\mathbf{C}_4$ are given by (1), but $\mathbf{C}_1$ and $\mathbf{C}_3$ are determined by the products

$$\mathbf{C}_1=\mathbf{BC}_2\mathbf{B}^T+\frac{N-10}{N-1}R_{ZY},\quad \mathbf{C}_3=\mathbf{BC}_4.$$

Since $\hat{\mathbf{\Sigma}}$ should be symmetric, this property may be used to obtain some of the remaining elements:

$$\mathbf{C}_3{}^T=[\sigma_{ij}]_{i=25,\ j=10}^{100,24}.$$

Entries σ_{ij} of $\hat{\mathbf{\Sigma}}$ have now been determined for $i,j=10,\ldots,100$.

Completion of the matrix $\hat{\mathbf{\Sigma}}$ is accomplished by using the grenadesonde data from the first level to level 16. Let

$$\mathbf{W}(k)=[x_1(k),\ldots,x_9(k)]^T,$$
$$\mathbf{V}(k)=[x_{10}(k),\ldots,x_{16}(k)]^T,$$

so that $[\mathbf{W}(k)^T,\mathbf{V}(k)^T]^T$ is the kth vector of grenadesonde data. As before, let $\mathbf{D}$ represent the estimated regression coefficients of $\mathbf{W}(k)$ on $\mathbf{V}(k)$, and R_{WV} the estimated residual variance of $\mathbf{W}$, given $\mathbf{V}$. Define sub-matrices $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{H}_3$, $\mathbf{H}_4$ as follows:

$$\mathbf{H}_1=[\sigma_{ij}]_{i,j=1}^{9},\qquad \mathbf{H}_2=[\sigma_{ij}]_{i,j=10}^{16}$$
$$\mathbf{H}_3=[\sigma_{ij}]_{i=1,\ j=10}^{9,100},\qquad \mathbf{H}_4=[\sigma_{ij}]_{i=10,\ j=10}^{16,100}$$

Calculations for $\mathbf{H}_2$ and $\mathbf{H}_4$ have already been made, but $\mathbf{H}_1$ and $\mathbf{H}_3$ are given by

$$\mathbf{H}_1=\mathbf{DH}_2\mathbf{D}^T+\left(\frac{N-8}{N-1}\right)R_{WV},\quad \mathbf{H}_3=\mathbf{DH}_4.$$

The remaining elements, by the symmetry property, are given by

$$\mathbf{H}_3{}^T=[\sigma_{ij}]_{i=10,\ j=1}^{100,9},$$

thereby completing the entire covariance matrix $\hat{\mathbf{\Sigma}}$.

If there are no missing data in any of the radiosondes from 1000 mb to 10 mb, then the matrix

$$\mathbf{\Sigma}^*=[\sigma_{ij}]_{i,j=25}^{100},$$

defined by (1), is positive semi-definite, and consequently, the entire estimator $\hat{\mathbf{\Sigma}}$ will be semi-definite. Note that the grenadesonde and rocketsonde data are assumed to be complete over the pressure range used.

In practice some of the radiosonde data were incomplete over part of the required pressure interval. However, the sample sizes were so large that inconsistent (i.e., at least one negative eigenvalue) estimators were not a problem.

Other techniques for estimating the covariance matrix are less suitable for these kinds of data because they may produce inconsistent estimators. They also require that the segmented means coincide in the regions of overlap. This is not necessary in the present estimation procedure for the covariance matrix. Ultimately, however, smooth means are needed to obtain physically realistic solutions.

The mean profile $\boldsymbol{\mu}$ is estimated from the separate means of the three data bases. From 1000 to 10 mb the elements of this estimate, $\bar{\mathbf{X}}^*$, are

$$\bar{x}_i=\frac{1}{N}\sum_{k=1}^{N}x_i(k),\quad i=25,\ldots,100.$$

The mean vector $\bar{\mathbf{Z}}$ is fitted to $\bar{\mathbf{X}}^*$ by translating $\bar{\mathbf{Z}}$ until the element $\bar{x}_{24}$ and the element $\bar{x}_{25}$ of the vector $\bar{\mathbf{X}}^*$ are almost coincident, differing only by that amount which causes the slope at the juncture to be smooth. The mean vector $\bar{\mathbf{W}}$ is fitted to the composite vector $[\bar{\mathbf{Z}}^T,\bar{\mathbf{X}}^{*T}]^T$ in the same way.

3. Sample areas and time periods

Covariance matrices and means must be established for different geographical areas and times of the year because of the differing statistical characteristics of the atmosphere. Each of the three data bases are considered separately. Because radiosonde data are the most abundant, and because they cover the part of the atmosphere which is of primary concern in the inversion, they form the basis for the selection of the areas and the time periods.

Unauthenticated | Downloaded 03/29/23 12:17 AM UTC

For the radiosonde data taken over a five-year period, the area from 15N to 90N is divided into maritime and continental regions in the latitudinal bands: 15–30N, 30–45N, 45–60N, 60–90N. The tropical band, 15S–15N, includes both maritime and continental regions. Representative stations were selected from each area. Estimates of the means and covariance matrices are based on 12 bimonthly groupings of the data to cover the full year.

Rocketsonde data are grouped into the five latitudinal bands, without regard to topography. In order to obtain samples of sufficient size, it is necessary to use time periods of three months. There are twelve such periods for each of the five latitudinal bands, and each period begins on the fifteenth of a month and continues for three months.

Rocketsonde and radiosonde data from the same latitudinal bands are combined, using the technique of Section 2. The two-month period of radiosonde data is centered in time within the three-month period of rocketsonde data. For example the February–March radiosonde data were used with the 15 January to 15 April rocketsonde data.

Grenadesonde data are very sparse, and were therefore grouped into three data clusters: a full-year grouping for the area 15S to 30N, and two six-month groupings for 30–90N. The two six-month groupings are for the periods 15 April to 15 October, and 15 October to 15 April.

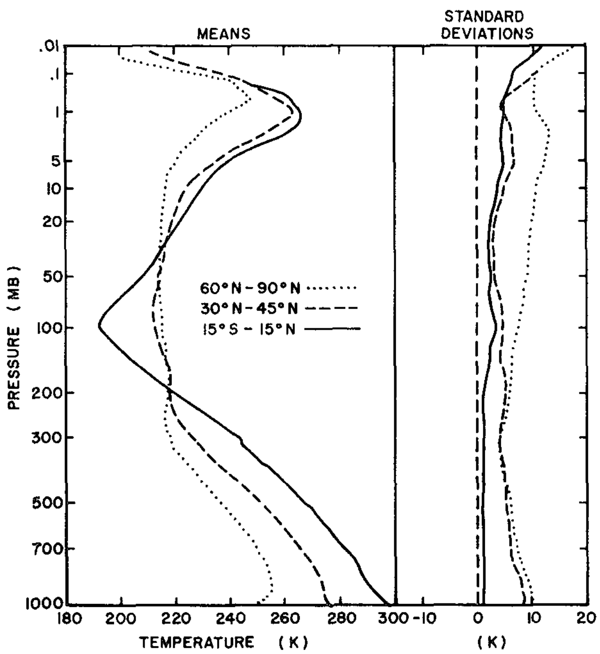

FIG. 1. Means and standard deviations for the tropics and two continental latitudinal bands for January–February.

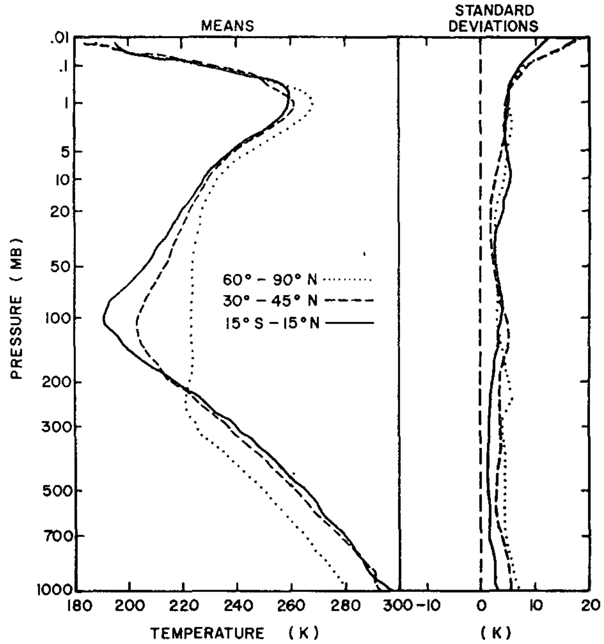

FIG. 2. Means and standard deviations for the tropics and two continental latitudinal bands for July-August.

4. Results

The means and standard deviations at the various pressure levels for the continental January–February data set are displayed in Fig. 1, which includes the tropical region from 15S to 15N, the temperate region 30–45N, and the polar region 60–90N. Fig. 2 shows the same information for the July–August data set. For ease of interpretation, the Planck radiance data have been converted to the temperature equivalents. Note the pronounced increase in the extratropical regions of the temperature variation about the mean in the winter period as opposed to the summer period. As might be expected, both the mean and the deviations about it are almost invariant in the tropics.

Covariance matrices and corresponding means have been prepared for each of the nine regions by month for the entire year.

REFERENCES

Alishouse, J. C., L. J. Crone, H. E. Fleming, F. L. Van Cleef and D. Q. Wark, 1967: A discussion of empirical orthogonal functions and their application to vertical temperature profiles. *Tellus*, **19**, 477–482.

Foster, M., 1961: An application of the Wiener-Kolmogorov smoothing theory to matrix inversion. *J. Soc. Ind. Appl. Math.*, **9**, 387–392.

GARP Joint Organizing Committee, 1969: The planning of the First GARP Global Experiment. GARP Publ. Ser. No. 3, 36 pp.

Rodgers, C. D., 1970: Remote sounding of the atmospheric temperature profile in the presence of cloud. *Quart. J. Roy. Meteor. Soc.*, **96**, 654–666.

Smith, W. L., H. M. Woolf and W. J. Jacob, 1970: A regression method for obtaining real-time temperature and geopotential height profiles from satellite spectrometer measurements and

Unauthenticated | Downloaded 03/29/23 12:17 AM UTC

its application to Nimbus 3 "SIRS" observations. *Mon. Wea. Rev.*, **98**, 582–603.

Strand, O. A., and E. R. Westwater, 1968: Minimum-rms estimation of the numerical solution of a Fredholm integral equation of the first kind. *SIAM J. Num. Anal.*, **5**, 287–295.

Wark, D. Q., 1970: SIRS: An experiment to measure the free air temperature from a satellite. *Appl. Opt.*, **9**, 1761–1766.

——, and H. E. Fleming, 1966: Indirect measurements of atmospheric temperature profiles from satellites: I. Introduction. *Mon. Wea. Rev.*, **94**, 351–362.

Unauthenticated | Downloaded 03/29/23 12:17 AM UTC