An Algorithm to Generate Deep-Layer Temperatures from Microwave Satellite Observations for the Purpose of Monitoring Climate Change

MITCHELL D. GOLDBERG AND HENRY E. FLEMING*

NOAA/National Environmental Satellite, Data, and Information Service, Satellite Research Laboratory, Washington, DC

ABSTRACT

An algorithm for generating deep-layer mean temperatures from satellite-observed microwave observations is presented. Unlike traditional temperature retrieval methods, this algorithm does not require a first guess temperature of the ambient atmosphere. By eliminating the first guess a potentially systematic source of error has been removed. The algorithm is expected to yield long-term records that are suitable for detecting small changes in climate.

The atmospheric contribution to the deep-layer mean temperature is given by the averaging kernel. The algorithm computes the coefficients that will best approximate a desired averaging kernel from a linear combination of the satellite radiometer's weighting functions. The coefficients are then applied to the measurements to yield the deep-layer mean temperature. Three constraints were used in deriving the algorithm: 1) the sum of the coefficients must be one, 2) the noise of the product is minimized, and 3) the shape of the approximated averaging kernel is well behaved. Note that a trade-off between constraints 2 and 3 is unavoidable.

The algorithm can also be used to combine measurements from a future sensor [i.e., the 20-channel Advanced Microwave Sounding Unit (AMSU)] to yield the same averaging kernel as that based on an earlier sensor [i.e., the 4-channel Microwave Sounding Unit (MSU)]. This will allow a time series of deep-layer mean temperatures based on MSU measurements to be continued with AMSU measurements. The AMSU is expected to replace the MSU in 1996.

1. Introduction

For long-term monitoring of temperature change, deep-layer mean temperatures derived directly from satellite observations of upwelling radiance have an advantage over traditional operational temperature retrievals. The advantage is that unlike operational retrieval algorithms (Eyre 1989; Fleming et al. 1988; Goldberg et al. 1988; Hayden 1988) an algorithm for deriving deep-layer temperature directly can be made independent of a first guess of the ambient temperature profile. Operational retrievals are dependent on a first guess because the satellite observations alone do not have the vertical resolution to yield pointwise temperatures, which are needed for forecast models. Unfortunately, the error between the first guess and the true ambient condition is systematic and, furthermore, the mean error cannot be entirely removed by the retrieval process (Thompson and Tripputi 1994). Since significant climate change on a global scale can be on the order of only tenths of a degree, temperature products independent of a first guess are a step in the right direc-

* Deceased.

Corresponding author address: Mitchell D. Goldberg, NOAA/NESDIS, Satellite Research Laboratory, Washington, DC 20233.

tion. First guess independency provides certainty that any observed trends in the data are not due to errors in the first guess, which could very well have its own interannual variation. Deep-layer mean temperatures are appropriate for long-term monitoring of temperature trends because nearly all climate models have indicated that climate changes will occur over deep layers and not at isolated levels (Mitchell et al. 1990).

The utilization of measurements from the Microwave Sounding Unit (MSU), on board NOAA's operational polar orbiting satellites, has gained much recognition during the past few years as a measure of deep-layer mean temperature for long-term monitoring of climate change (Spencer and Christy 1992a,b, 1993; Spencer et al. 1990). Because radiance in this spectral region is extremely linear with respect to temperature, the observations can be interpreted as deep-layer mean temperatures for the layer defined by the weighting function. This is not true for the infrared spectral region, where temperature and radiance can be very nonlinear. Microwave observations are usually expressed in brightness temperature, which can be obtained from radiance using the inverse form of the Planck function.

The MSU has four channels measuring outgoing radiation at 50.31, 53.73, 54.96, and 57.95 GHz. Channel 1 (50.31 GHz) has a large surface component and is generally not used for deriving temperature due to un-

certainty in the surface emissivity. The first MSU was launched in 1979, and to date, its replacements have provided nearly complete daily coverage of the earth by scanning across the orbital track at ± 47.35 degrees about nadir at approximately 9.47-degree increments. The MSU's six view angles results in the projection on the earth of 11 fields of view (FOV) for each scan line. The weighting functions for channels 2 through 4 at each of the six view angles are given in Fig. 1. The highest peaking group of weighting functions is for MSU channel 4, followed by MSU channels 3 and 2. The higher peaking weighting functions in each channel grouping are associated with larger off-nadir angles.

Spencer and Christy (1992a), used MSU channel 2 (53.73 GHz) brightness temperatures, adjusted to nadir, to monitor temperature for the layer defined by the channel 2 weighting function on a 2.5° gridpoint scale with a monthly precision of better than 0.1°C in the Tropics and to better than 0.2°C at high latitudes. These estimates of precision were arrived at through intersatellite comparisons and in comparisons with radiosondes. They conclude that "the satellite precision approaches that of individual radiosonde stations in their ability to measure monthly temperature anomalies . . . ." In terms of monthly, zonally averaged temperatures, they estimate their precision is of the order of 0.01°C over a 10-year period.

A deep-layer mean temperature from a single microwave observation has the equivalent vertical resolution of the channel. Improved vertical resolution can be obtained by combining different channels. The layer
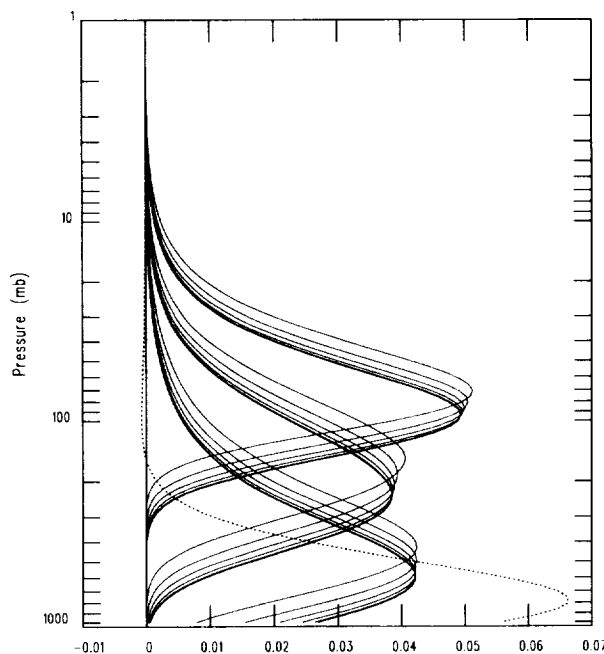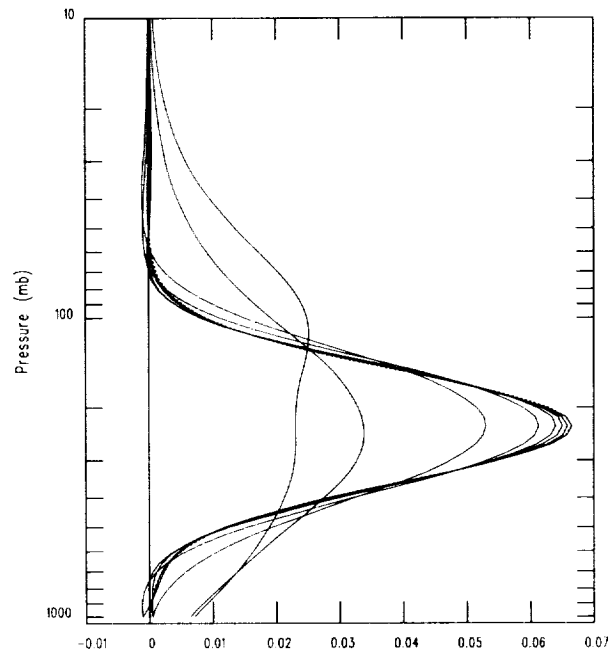

Fig. 2. The influence of the gamma parameter on the shape of the averaging kernel.

is now defined by the averaging kernel, which is simply derived from a linear combination of the weighting functions (using the same coefficients used to combine the measurements). To remove the stratospheric component from MSU channel 2, Spencer and Christy (1992b) combined channel 2 measurements at different viewing angles to create a more narrow averaging kernel, shown as the dotted curve in Fig. 1, than the raw nadir-viewing weighting function. It is interesting to note that the raw channel 2 time series for the period 1979–90 showed a global warming trend of only 0.015°C per decade, while the combined-angle approach yielded an increased global warming trend of 0.032°C per decade. By combining different viewing angles, Spencer was retrieving additional information that a single channel at a common view angle was unable to provide. The only a priori information required was knowledge of the weighting functions, which for the MSU is well known and can be derived from a standard atmosphere. Because the MSU weighting functions are very weakly dependent on temperature and moisture, a fixed set of coefficients can be used globally to derive the deep-layer mean. This is not true for infrared measurements; their weighting functions generally have a much greater dependency on the ambient atmosphere.

Spencer did not use an algorithm to determine the coefficients for his lower-troposphere deep-layer mean temperature. He used trial and error by visual inspection of the averaging kernel to determine the appropriate coefficients. This technique is acceptable when

Fig. 1. MSU weighting functions for channels 2, 3, and 4 at all view angles and Spencer's derived averaging kernel (dotted curve).

considering a very few number of channels or angles. However, as the number of different channels and view angles increases, the determination of the coefficients to yield a desired averaging kernel becomes a formidable task. A quantitative retrieval algorithm is required to optimally solve for the coefficients. The coefficients need to be optimal in the sense that the derived averaging kernel is well behaved and that size of the coefficients are constrained so that the noise of the product does not become large.

The emphasis of this paper is to present an algorithm to derive deep-layer mean temperatures from microwave observations within the band 50–60 GHz. The algorithm, derived in section 2, computes the coefficients needed to combine a set of channel weighting functions into a desired deep-layer mean averaging kernel. The deep-layer mean temperature is obtained by simply applying the coefficients directly to the observed brightness temperatures. Examples of averaging kernels from the MSU are given in section 3. We will also demonstrate that the MSU temperature time series that Spencer pioneered can be continued with the next generation of microwave sounders—the 20-channel Advanced Microwave Sounding Unit (AMSU) (Fischer 1987). The first AMSU is expected to be launched in 1996. This will be accomplished by constraining the averaging kernel associated with the set of measurements from the AMSU instrument to be approximately equal to the averaging kernel associated with the set of measurements from the MSU instrument.

2. Algorithm

Our algorithm for computing deep-layer mean temperatures and its corresponding averaging kernels is a specialized adaptation of the Backus–Gilbert theory discussed in Conrath (1972). The Conrath paper dis-
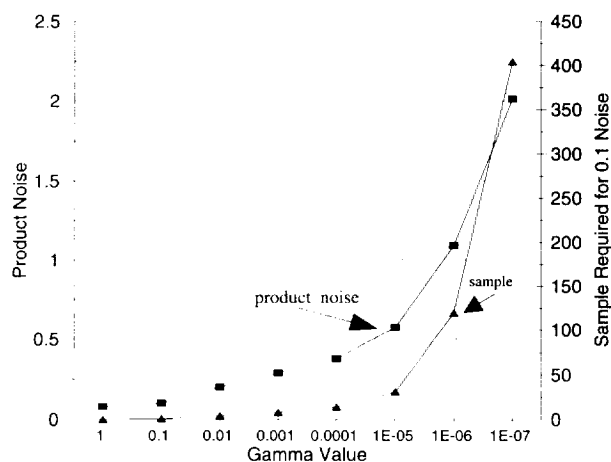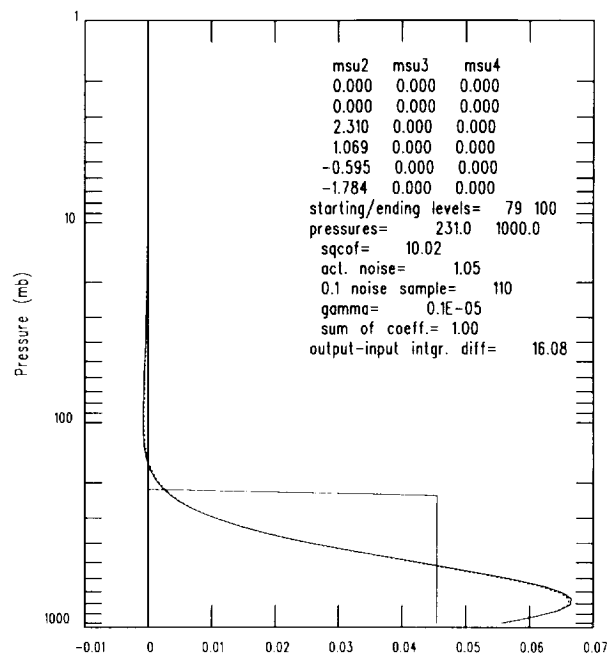

FIG. 4. Comparison of the boxcar-derived averaging kernel, based on MSU channel 2 at view angles 3 through 6 and Spencer's derived averaging kernel (dotted curve). Also shown are the coefficients, the starting and ending levels and pressures of the boxcar function, the sum of the square of the coefficients, the noise of the product, the value of the gamma parameter, the sum of the coefficients, and the integrated difference between the shape constraint and the derived averaging kernel.

cusses the trade-off between instrumental noise and the vertical resolution of the averaging kernel for a given atmospheric level and set of measurements. The derivation of our algorithm begins with the same basic definition of the averaging kernel used by Conrath. However, our approach differs from Conrath with respect to application and constraint. Conrath's constraint is to derive coefficients that, when applied to the weighting functions, attempt to reproduce the ideal dirac delta function. In other words, he is trying to obtain the highest-resolution averaging kernel possible, cognizant of the effects of instrumental noise, for a particular level in the atmosphere. This approach is very useful for comparing the resolving power of current and future sounders. On the other hand, our constraint is to yield coefficients that will reproduce a prespecified averaging kernel. Our averaging kernel, unlike Conrath's, is not associated with a given level. Instead it is "predesigned" to correspond to a desired deep-layer mean temperature t_L derived from a linear combination of n measured brightness temperatures T_i. That is,

$$t_L = c_1 T_1 + \cdots + c_n T_n, \qquad (1)$$

where the c_i are the coefficients of the linear combination.

FIG. 3. The relationship between γ (gamma) and product noise and the required sample to reduce the product noise to 0.1 K.

a. Algorithm constraints

To optimize the coefficients in (1) for a given atmospheric layer and a given set of channels and viewing angles, three constraints have been imposed, which now are explained in detail. The first constraint requires that the sum of the coefficients is unity. Since t_L of (1) can be interpreted as a weighted average of brightness temperatures, the weights must be normalized by constraining the coefficients to have sum one; that is,

$$c_1 + \cdots + c_n = 1. \tag{2}$$

Thus, if all n of the T_i in (1) are identical, then (2) guarantees that t_L will have that same value. Since the T_i are normalized so that a constant shift of one degree in the temperature profile will result in a shift of one degree in the T_i, this constraint will ensure that t_L has the same property.

The second constraint addresses the problem that each of the brightness temperatures used in (1) carries with it a measurement error. Let σ_i^2 be the variance of the error associated with T_i and let σ^2 be the variance of the total error associated with t_L. It is well known that with independence of the individual errors the relationship between the total error variance and the individual error variances is given by

$$\sigma^2 = c_1^2 \sigma_1^2 + \cdots + c_n^2 \sigma_n^2. \tag{3}$$

Consequently, to minimize the magnitude of σ^2 we require as a second constraint that the sum of (3) be a minimum.
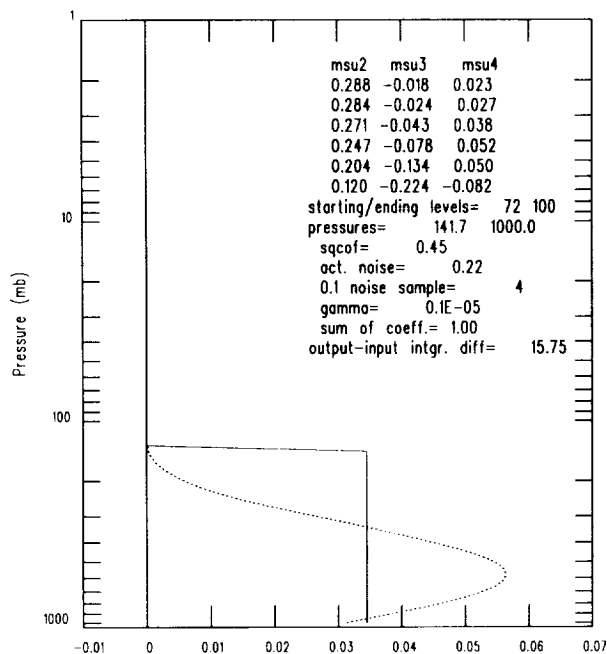

FIG. 5. Boxcar-derived averaging kernels using
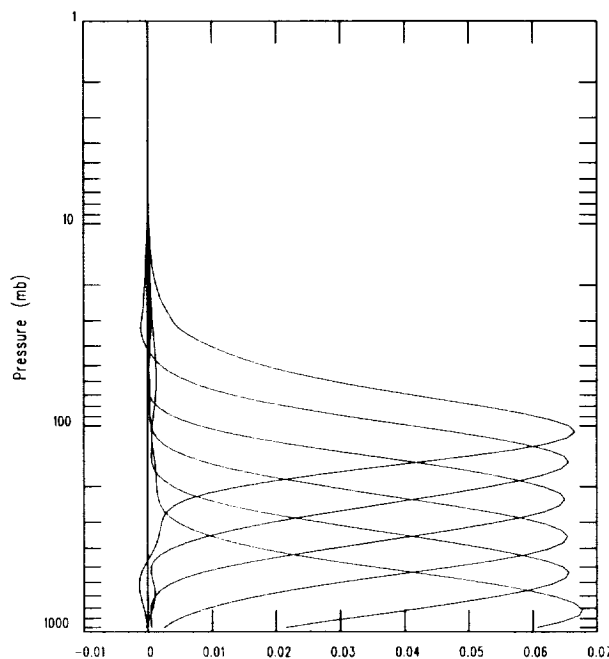MSU 2, 3, and 4 at all view angles.

FIG. 6. Gaussian-derived MSU averaging kernels using
MSU 2, 3, and 4 at all view angles.

For the third constraint one must determine the coefficients c_i of (1) in such a way that the deep-layer mean averaging kernel agrees with the desired averaging kernel as close as possible. The manner in which the averaging kernel is defined is through the weighting functions $w_i(x)$ associated with the ith channel and which are the components of the kernel function in the radiative transfer equation. Thus, the layer over which t_L of (1) is defined is given by the so-called "averaging kernel," given by the linear combination

$$a(x) = c_1 w_1(x) + \cdots + c_n w_n(x). \tag{4}$$

Equation (4) follows directly from (1). Note that x can be any monotonic function of the atmospheric pressure p. The purpose in making w_i a function of x, instead of p directly, is that by judiciously choosing the transformation from p to x, one can shape the weighting function to suit specific needs. It also has the property that the sum of $w_i(x)$ over the range of x is unity. Because of the first constraint, the sum of $a(x)$ over the range of x is also unity. The values of the algorithm-derived averaging kernel represent the true weights of the contribution of the unknown temperature profile to t_L.

Note that the first two constraints were also used by Conrath (the first for a different reason). It is the third constraint and how we treat it that provides the major relevance of this work.

b. Coefficient determination

Determination of the coefficients in the linear combination (1) of brightness temperatures, having the
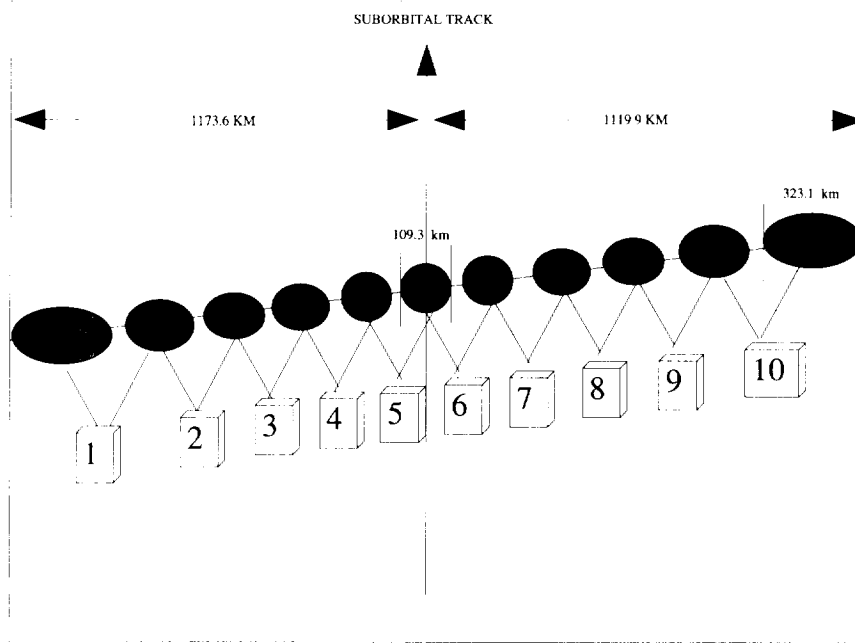
FIG. 7. The relationship between the 11 MSU beam positions
and the ten deep-layer mean temperatures.

three properties discussed above, is now considered.
We begin by letting c and T be the vectors of coefficients
and brightness temperatures in (1), respectively, and
define the n-dimensional vector

$$u = [1, \cdots, 1]^T, \qquad (5)$$

where the transpose superscript T is used because all
vectors are assumed to be column vectors. Then (1)
can be written

$$t_L = c^T T. \qquad (6)$$

and (2) can be written

$$u^T c = 1. \qquad (7)$$

If we let

$$D = \text{diag}(\sigma_1^2, \cdots, \sigma_n^2) \qquad (8)$$

be an n-dimensional diagonal matrix whose diagonal
elements are those indicated, then (3) can be written

$$\sigma^2 = c^T D c. \qquad (9)$$

Furthermore, if we let W be a matrix of weighting
functions with dimensions channel (n) by level (j),
then the averaging kernel $a(x)$ of (4) can be written
as the J-dimensional vector

$$a = W^T c. \qquad (10)$$

Next, a shape vector b of j elements (i.e., the desired
averaging kernel) is defined to constrain the shape
of the resulting averaging kernel. The coefficient

vector c is determined in such a way that the shape
of a, given by (10), approximates the shape vector
as closely as possible. To do this, we minimize the
squared distance between the vectors a and b, while
at the same time satisfying the constraint (7) and
minimizing (9).

We now are ready to determine the coefficient vector
c by optimizing our solution with respect to the three
properties just discussed. This is accomplished by first
establishing a cost, or penalty, function F, which in-
corporates all three constraints. In its most general form
the cost function is

$$F(c) = (W^T c - b)^T S(W^T c - b)$$
$$+ \gamma c^T D c + 2\lambda(1 - u^T c), \qquad (11)$$

where λ and γ are Lagrange multipliers and S
is an arbitrary symmetric, positive definite (usu-
ally diagonal) matrix of dimension $J \times J$. Note
that the three terms on the right-hand side of (11)
represent, respectively, the shape constraint of (10)
minus the shape vector b, the error variance con-
straint (9), and the coefficient normalization con-
straint (7).

To find that vector c, which minimizes F, we dif-
ferentiate F with respect to c and equate the result to
zero. This yields

$$2WS(W^T c - b) + 2\gamma D c - 2\lambda u = 0, \qquad (12)$$

which implies that

$$c = (WSW^T + \gamma D)^{-1}(WSb + \lambda u), \qquad (13)$$

where the inverse matrix is well defined because **W** has dimensions $n \times J$, with $n < J$ and rank n, and **D** and **S** are positive definite.

One solves for the scalar λ by multiplying (13) by $\mathbf{u}^T$ and using (7) to obtain

$$\lambda = [1 - \mathbf{u}^T(\mathbf{WSW}^T + \gamma\mathbf{D})^{-1}(\mathbf{WSb})]/$$

$$[\mathbf{u}^T(\mathbf{WSW}^T + \gamma\mathbf{D})^{-1}\mathbf{u}]. \quad (14)$$

When (14) is applied to (13), one acquires the desired coefficient vector **c**. At this point, all the quantities in (13) and (14) are known except the scalar γ and the matrix **S**. These two quantities are used to provide the averaging kernel with the proper shape.

c. Shaping the averaging kernel

Our ability to accurately fit an averaging kernel vector of (10) to a given shape vector is limited by the number of channels and viewing angles available to us. Ultimately, one would like to fit a boxcar function, since it represents a uniform average of the layer in question. Unfortunately, the limited number of channels available to us prevents us from reproducing the edges of the nonzero portion of the boxcar function as well as the flat portion. Generally, the best one can do is to derive a shape similar to a narrowed weighting function. Other shapes are easier to fit. For example, in the next section we will demonstrate the use of Gaussian functions as well as weighting functions of different sensors. Note that it is immaterial for climate and global change studies that the shape of the derived averaging kernel is not uniform (i.e., flat) over the layer it defines. All that is required is that its shape be known with great accuracy and that the layer in question is well defined (i.e., the boundaries of the layer are clearly delineated with little or no energy leakage contribution from outside the boundaries). The algorithm-derived averaging kernel will display a good deal of "ringing" if the shape function is too narrow (i.e., the boxcar is too narrow or, in the case of a Gaussian function, the value of the standard deviation is too small). Ringing is the undesirable phenomenon that, instead of having a flat zero response outside the nonzero portion of the shape function, one has a set of rapidly decaying positive and negative oscillations. There are three mechanisms that allow one to control ringing: First, the shape function being fitted cannot have its width too narrow; it must have its width at least comparable to the full width at half maximum (FWHM) of the weighting functions. The other controlling variables are the scalar γ and the matrix **S**.

The most important thing to realize about γ and **S** is that they play competing roles (i.e., at all times a trade-off situation exists between them). To see this, consider separately the limiting cases where these quantities are set to zero. First, when $\gamma = 0$ in (13) and (14), the error variance constraint disappears; in
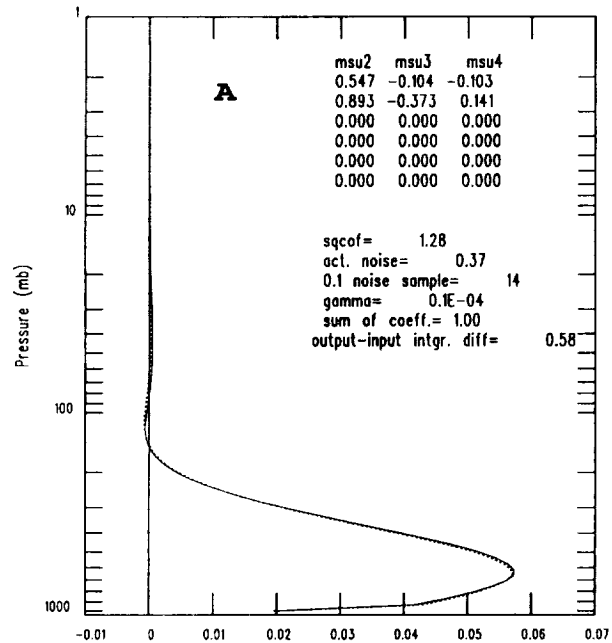

FIG. 8. Averaging kernels for the adjacent field of view deep-layer mean temperatures. Plot B shows the nominal averaging kernel obtained by fitting weighting functions associated with either beam positions 4 and 5 or beam positions 7 and 8 (view angles 3 and 4). Plots A and C–E show the fit of the averaging kernels based on the remaining pairing of beam positions to the nominal averaging kernel.

this case the fit to the desired shape is most realistically (i.e., optimally) determined, subject of course to the required coefficient normalization constraint. To see the role of the matrix **S** in the limiting case $\gamma = 0$, assume the usual practical situation where **S** is a diagonal matrix. Assigning a relatively large value to the ith diagonal element of **S**, will cause the ith point of the averaging kernel to fit the shape vector better at the expense of the other points in the fit. This is equivalent to saying that the ith element of the residual vector $\mathbf{W}^T\mathbf{c} - \mathbf{b}$ in (11) will be relatively smaller. On the other hand, if **S** is set equal to the identity matrix, all the points will be fitted to the shape vector with equal weight.

While the shape-vector fit is best when $\gamma = 0$, the coefficients in the vector **c** of (13) no longer have a constraint on their size, and so σ^2 in (9) can grow without bound. Conversely, assume the other limiting case in which **S** is the matrix having all elements zero (i.e., $\mathbf{S} = 0$). In this case, the shape constraint disappears and so the error σ^2 of (9) is minimized in (11). If the errors of each element in **T** are identical, **D** in (8) becomes the identity matrix multiplied by a constant and the coefficient for each channel is $1/n$. In other words, each channel has equal weight. But now there is no control on the shape or size of the averaging kernel. This also is an untenable situation. Clearly then, neither **S** nor γ can be zero; there must be a trade-off

B

msu2	msu3	msu4
0.000	0.000	0.000
0.741	-0.239	0.033
0.712	-0.272	0.025
0.000	0.000	0.000
0.000	0.000	0.000
0.000	0.000	0.000

starting/ending levels=   77 100
pressures=   200.9   1000.0
sqcof=   1.19
act. noise=   0.36
0.1 noise sample=   13
gamma=   0.1E-03
sum of coeff.= 1.00
output-input intgr. diff=   11.70

C

msu2	msu3	msu4
0.000	0.000	0.000
0.000	0.000	0.000
1.478	-0.661	0.129
-0.010	0.124	-0.060
0.000	0.000	0.000
0.000	0.000	0.000

sqcof=   2.66
act. noise=   0.54
0.1 noise sample=   29
gamma=   0.1E-04
sum of coeff.= 1.00
output-input intgr. diff=   -0.49

D

msu2	msu3	msu4
0.000	0.000	0.000
0.000	0.000	0.000
0.000	0.000	0.000
2.366	-0.895	0.074
-0.927	0.396	-0.015
0.000	0.000	0.000

sqcof=   7.42
act. noise=   0.90
0.1 noise sample=   81
gamma=   0.1E-04
sum of coeff.= 1.00
output-input intgr. diff=   -0.96

E

msu2	msu3	msu4
0.000	0.000	0.000
0.000	0.000	0.000
0.000	0.000	0.000
0.000	0.000	0.000
3.143	-0.449	0.001
-1.900	0.154	0.050

sqcof=   13.71
act. noise=   1.23
0.1 noise sample=   150
gamma=   0.1E-04
sum of coeff.= 1.00
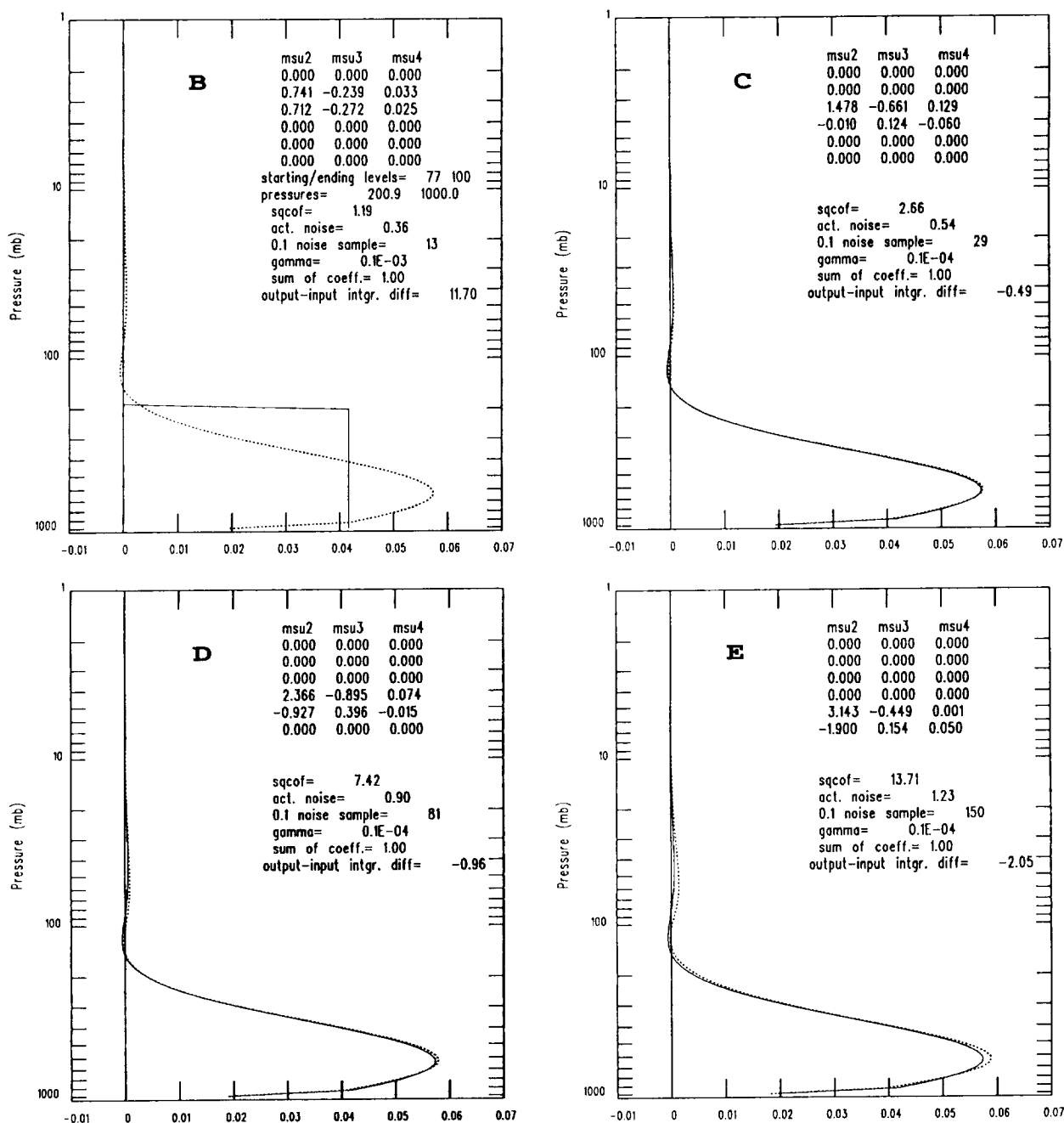output-input intgr. diff=   -2.05

FIG. 8. (Continued)

between their magnitudes in order to achieve a satisfactory balance between an acceptable averaging kernel shape and an acceptable error level in the deep-layer mean temperature.

The best strategy we found when using a boxcar constraint is to define **S** as a diagonal matrix with values of zero in the nonzero elements of the boxcar and values of one elsewhere. This will tend to force the averaging kernel to be zero outside the boxcar. For a Gaussian function, we found that simply defining **S** to be the identity matrix produced desirable results. The reason **S** is not critical for a Gaussian function is probably due to the smooth transition to zero from the Gaussian's maxima. Figure 2 demonstrates the influence of the γ parameter for fitting a Gaussian shape constraint (dashed curve) from MSU channels 2, 3, and 4 weighting functions at all view angles. There are six averaging kernels for six different values of γ. The
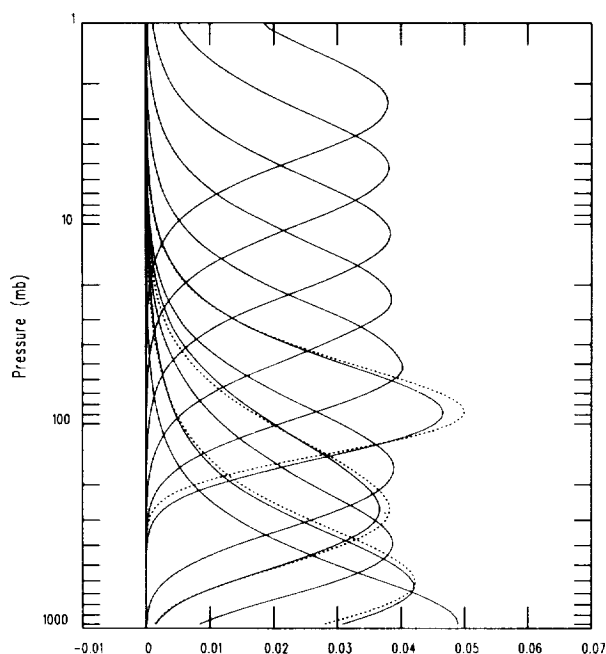
FIG. 9. Comparison of AMSU and MSU (dotted curves) weighting functions.

γ parameter was given a value of unity to produce the curve with the least Gaussian shape and then reduced for each of the remaining five permutations by an order of magnitude. The curve best approximating the Gaussian used a value of 1.0×10^{-7} for γ. Equation (3) was used to compute the noise of the deep-layer mean associated with each averaging kernel. In this study, the noise for all channels is always assumed to be 0.33 K. The deep-layer mean temperature noise as a function of γ as well as the sample size required to reduce the noise to 0.1 K are shown in Fig. 3 (since the noise is assumed to be random, the noise will be reduced by the square root of the sample size). Clearly, there is a trade-off between the goodness of fit and the product noise. The best fit cannot be used because the noise is too large. The averaging kernel associated with the lowest noise is too broad. A good compromise is the averaging kernel associated with a 0.0001 value for γ, because 1) the noise is at an acceptable level, 2) the shape is similar to the shape constraint, and 3) there is no ringing.

There is no exact recipe to derive the optimum averaging kernel and its associated coefficients. The following is the methodology that we use. First, begin by selecting a desired width and mean height for the shape function. By trial and error we found that 0.001 is a good maximum value for γ. If ringing occurs, gradually widen the function until the ringing subsides. Then fine tune γ by reducing it until a more desirable shape of the averaging kernel is achieved. Remember, if γ is chosen too small, the elements of the coefficient vector

of (13) will be too large, which will amplify the total error σ^2 in (9). The user would need to decide an acceptable error level. Note that if the deep-layer means are averaged over large temporal and/or spatial domains, then some allowance can be made in the individual error level since averaging reduces the noise by the square root of the sample size.

3. Application

In this section, we will demonstrate the use of this algorithm to derive MSU averaging kernels of deep-layer mean temperatures that we believe would be optimal for climate studies. AMSU averaging kernels will also be shown; however, the emphasis will be to demonstrate that AMSU measurements can be used to continue time series based on MSU measurements.

a. MSU averaging kernels

The Spencer and Christy (1992b) averaging kernel shown in Fig. 1 was produced by combining MSU channel 2 weighting functions associated with view angles 3 through 6. The Spencer product is widely known as MSU-2R. The MSU-2R averaging kernel was computed using coefficient values of 2.0 for angles 3 and 4, and -1.5 for angles 5 and 6. The first experiment with our algorithm was to adjust the size of the boxcar and the γ parameter so as to yield an averaging kernel most similar to the one produced by Spencer. As was discussed in the previous section, there are not enough channels to reproduce a boxcar. Of course, for global
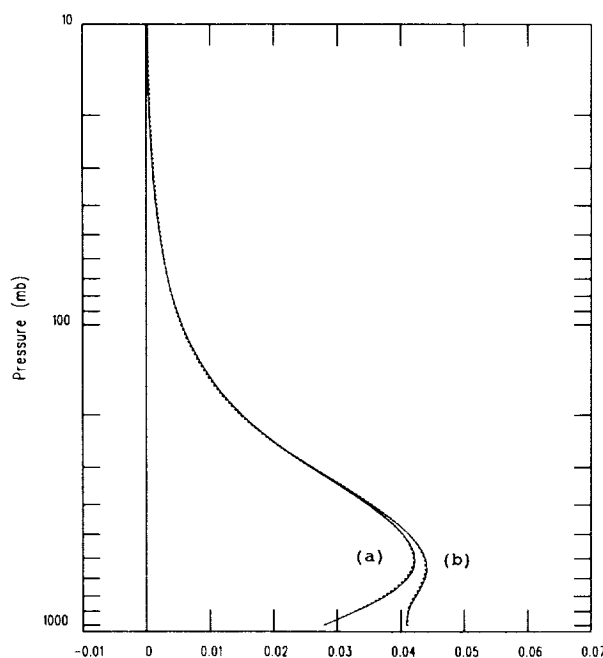

FIG. 10. Actual and AMSU-derived MSU channel 2 weighting functions for surface emissivities of (a) 1.0 and (b) 0.5.
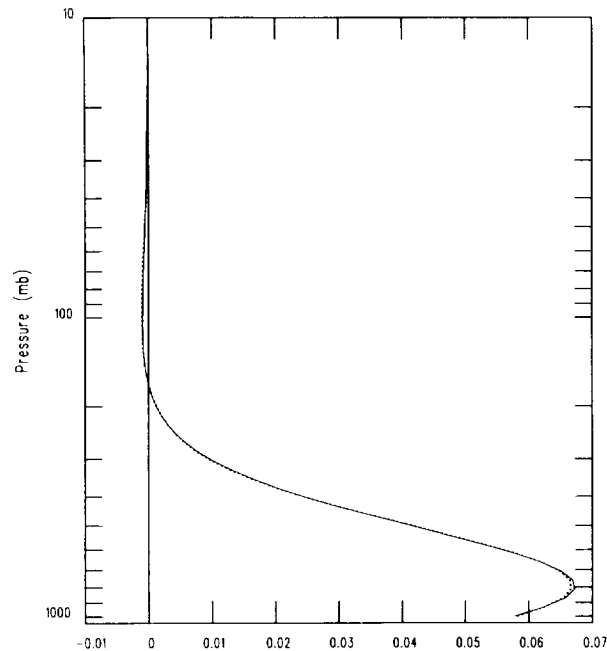
FIG. 11. Actual and AMSU-derived MSU-2R averaging kernels for surface emissivity of 1.0.

change purposes the averaging kernel does not have to be a boxcar; all that is necessary is for the averaging kernel be known. Our averaging kernel is shown along with Spencer's (dotted curve) in Fig. 4. These averaging kernels are similar. The difference is that one was determined subjectively and the other quantitatively using the algorithm of section 2. The numerical values in the columns labeled msu2, msu3, and msu4 are the derived coefficients (i.e., the c_i). Each channel is associated with six coefficients, one for each view angle, beginning with view angle 1 (nadir). Hence, there is a potential maxima of 18 channels. It is seen that the only nonzero coefficients are at view angles 3 through 6 for channel 2. Also shown in the figure is the sum of the square of the coefficients, the noise of the product, the required sample size to reduce the noise of the product to 0.1 K, and the integrated difference in degrees Kelvin. The integrated difference is scalar product of the difference between the shape vector and the derived averaging kernel vector and a standard midlatitude temperature profile (vector). Note that the shape constraints and weighting functions used by the algorithm are defined at 100 levels equally spaced in log pressure.

The advantage of using an algorithm to objectively determine the coefficients becomes quite clear if instead of using four measurements to produce an averaging kernel, all 18 effective channels are considered. For example, the averaging kernel given by the solid curve in Fig. 5 was obtained by using all view angles of MSU channels 2, 3, and 4, and hence all coefficients are non-

zero. This kernel is more desirable for monitoring temperature in the lower troposphere than the other averaging kernels shown in Fig. 4 since there is far less signal from the surface.

As participants in the NOAA/NASA Pathfinder program (Ohring and Dodge 1992), we are planning to construct time series of two types of deep-layer mean temperatures covering the entire MSU archive. We will provide additional atmospheric layers to the ones given by Spencer and colleagues. The first type is referred to as scan line products, since observations from all view angles are to be used simultaneously. For the second type, observations from adjacent FOVs are used. The remainder of this section is devoted to a discussion of these two product types.

The scan line products are derived from using Gaussian shape constraints. There will be a total of six different deep layers, their averaging kernels are shown in Fig. 6. The values of the product noise associated with the six averaging kernels beginning with the highest peaking one are 0.89, 1.36, 0.76, 1.13, 0.62, and 0.81 K. Each Gaussian function had a standard deviation of six pressure levels, beginning at pressure level 67 (100 mb) and were separated by six levels. The use of Gaussian curves for the shape vector has the advantage that one can better dictate the location and shape of the derived averaging kernel. Because of the limited number of channels on the MSU sounder, averaging kernels confined solely to the stratosphere cannot be provided. Note that these averaging kernels are relatively narrow in comparison with the raw weighting
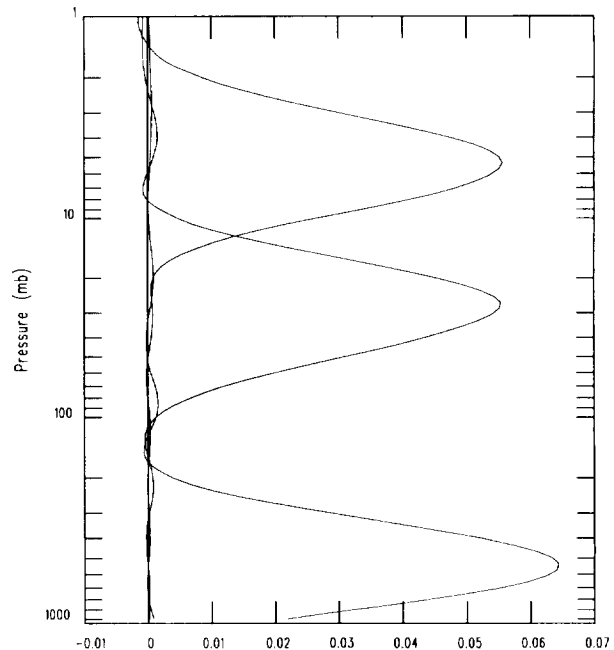

FIG. 12. Example of nadir Gaussian-derived AMSU averaging kernels in the troposphere and lower and upper stratosphere.

functions and could have been centered anywhere in the troposphere without excessive ringing.

The scan line product's poor horizontal resolution, which is on the order of 1100×150 km^2 (the mean area of scan line projected on the earth on either side of nadir), results in relatively poor sampling. To monitor small temperature fluctuations, these products will need to be averaged over relatively large spatial and temporal domains to reduce the product noise. One suggestion is to average over 10° latitude bands and for time intervals on the order of a month. The sample size will be about 10 000 (assumes 2 products per scan line, 220 scan lines per orbit, 14 orbits per day). The single largest product noise of 1.36 K will be reduced to a precision of 0.0136 K.

For regional climate monitoring the horizontal resolution of the product needs to be much smaller. To do this we want to derive products that are ideally based on a single FOV. In other words, 11 products for the 11 FOVs along the MSU scan line. There are different ways to do this. To apply the same set of coefficients to all view angles, the off-nadir measurements need to be adjusted to look as if they were observed at nadir (i.e., limb correct the measurements). For example, one can collect a large ensemble of measurements for all FOVs and compute regression coefficients using the measurements observed at a given FOV as the predictors and measurements observed at nadir as the predictands (Wark 1993). The MSU-2 and MSU-4 time series given in Spencer and Christy (1992a, 1993), respectively, were limb corrected.

Our approach is not to use a statistical method to limb correct, since we believe it is undesirable to adjust the measurements based on historical data. An attempt was made to physically limb correct the MSU by using the algorithm to compute coefficients for combining weighting functions at a particular off-nadir view angle to fit the nadir-viewing weighting functions. Unfortunately, this technique did not work well at the larger view angles. We also tried to compute a different set of coefficients for each view angle in order to fit a common averaging kernel. However, a combination based on only three channels was insufficient to maintain the same averaging kernel along the scan line. The solution was to use information from a pair of adjacent FOVs, which provides a total of six weighting functions to fit the desired averaging kernel. To better visualize this approach, the MSU scan line geometry and the adjacent FOVs used to yield the ten deep-layer mean temperature products across each scan is given in Fig. 7. So instead of a two products per scan line, this technique yields ten products. The nominal averaging kernel (dotted curve) is shown in plot B of Fig. 8. This averaging kernel was derived from a boxcar constraint and used weighting functions from view angles 2 and 3, which corresponds either to field of view 4 and 5 or 7 and 8. The nominal averaging kernel was then used as a shape constraint for other angular combinations
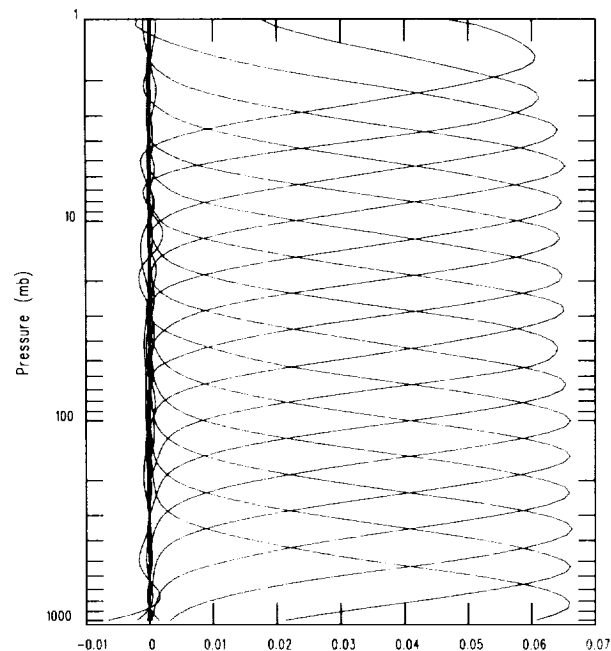

FIG. 13. Gaussian-derived AMSU averaging kernels using channels 4–14 at all view angles.

given in plots A and C–E of Fig. 8. Notice that the fit is good for all combinations with the exception of the largest off-nadir angles. However, even though the integrated difference for that combination is 2.0 degrees we found that the frequency distribution of each of the ten products is very similar and they differ only by an offset. For each of the ten products its mean condition will be subtracted when the data is analyzed. Hence, the bias for a given combination is removed.

b. Continuing the MSU time series with AMSU

Nadir-viewing weighting functions for AMSU channels 4 through 15, along with those from the MSU (dotted curves), are given in Fig. 9. Note that for AMSU there is actually no nadir position, the nearest off-nadir angle is 3.33 degrees. Of utmost concern is the ability to continue the record of tropospheric temperature trends established with MSU channel 2 and other linear combinations of the MSU channels. The AMSU channel that is most similar to MSU channel 2 is AMSU channel 5 (53.596 GHz). However, this channel is slightly more sensitive to the lower atmosphere and has a larger surface contribution, thereby producing a different signal. A change in the deep-layer mean temperature weighting function could conceivably produce a spurious signal in the time series based solely on MSU channel 2 and AMSU brightness temperatures.

The question to be answered is can this algorithm reproduce the MSU channel 2 weighting function from

the AMSU channel weighting functions? The answer to this question is yes. In Fig. 10 there are four curves. The curves to the left are the actual MSU channel 2 (solid curve) and the AMSU reconstructed MSU channel 2 (dotted curve) for a surface emissivity of 1.0. The other set of curves is for an emissivity of 0.5. The shape is different because the emissivity enters into the computation of the weighting functions. Only the "nadir" AMSU channels 4 through 7 weighting functions were used in deriving the coefficients. The selected emissivities are the extremes values for the surface emissivity in the 50-GHz band. The actual and reconstructed MSU channel 2 weighting functions are virtually identical. It is very important to note that the coefficients, based on an emissivity of 1.0, were used for reconstructing the weighting functions for an emissivity of 0.5. In other words, the reconstruction of MSU channel 2 is insensitive to surface emissivity, which is very important since the estimation of surface emissivity would add uncertainty to the final product. By using the appropriate γ, the integrated error between the real and reconstructed MSU channel 2 weighting function can be forced to be virtually zero. If we did nothing and simply used AMSU channel 5 to continue MSU channel 2, there would be a sizable airmass dependent bias. For a summer midlatitude atmosphere, the bias would be about 5.6 K. The linear combination of AMSU to yield an equivalent MSU channel 2 measurement is

$$T_{msu2} = -0.0488T_{amsu4} + 0.932T_{amsu5}$$
$$+ 0.208T_{amsu6} - 0.466T_{amsu7}. \quad (15)$$

Spencer's MSU-2R product can also be reproduced from the AMSU. AMSU channels 4 through 8 using 10 angles ranging from 18.66 to 49.55 degrees were combined to fit the MSU-2R averaging kernel. The AMSU equivalence of MSU-2R is given in Fig. 11. The coefficients are obtainable from the author.

Even though the accuracy of fitting AMSU to MSU appears to be high, the underlying assumption is that the weighting functions are known exactly. In practice, we know this is not true. Therefore, in conjunction with this algorithm, overlap of MSU and AMSU will be needed to adjust for the component that is left over after the "known" physics have been accounted for.

The AMSU by itself will be a very important sensor for monitoring temperature trends throughout the atmosphere. Its numerous channels will enable one to monitor temperature in three important regions of the atmosphere: the upper and lower stratosphere and the troposphere. Figure 12 shows examples of AMSU averaging kernels in these three regions. All were derived from initial Gaussian curves using only nadir measurements. Narrower averaging kernels can be achieved by utilizing off-nadir measurements. The technique used to generate the six averaging kernels, shown in Fig. 6, was applied to AMSU channels 5 through 14

weighting functions at all view angles. The result, shown in Fig. 13, clearly demonstrates that the ability to derive these averaging kernels is no longer restricted to the troposphere. It is also important to mention that the lowest six averaging kernels in Fig. 13 are virtually identical to the six averaging kernels shown in Fig. 6. Therefore, in addition to Spencer's time series of MSU, we will be able to extend our own time series with AMSU.

4. Summary

An algorithm for deriving deep-layer mean temperatures from microwave sensors has been developed. The algorithm, in conjunction with the microwave channels considered in this study, is completely independent of a priori information. Independence from ancillary data is critical for high-precision monitoring of climate trends, so that any observed trends in the deep-layer mean temperatures are attributed only to trends in the sensor's measurements. The algorithm also has been shown to be capable of combining measurements from next-generation microwave sensors to reconstruct measurements from current sensors. This enables one to generate continuous time series of satellite-derived temperature trends accurately, regardless of changes in satellite instrumentation.

The next step is to produce the actual MSU time series from the two types of deep-layer mean temperatures we plan to derive as part of the TOVS Pathfinder project. The first type will yield six different atmospheric deep-layer mean temperatures; their averaging kernels were shown in Fig. 6. The second type uses adjacent angular combinations to yield a single atmospheric averaging kernel. The important feature of the second type is that for each adjacent combination, the averaging kernel along the scan line is preserved so that limb correcting the measurements can be avoided.

REFERENCES

Conrath, B., 1972: Vertical resolution of temperature profiles obtained from remote sensing measurements. *J. Atmos. Sci.*, **29**, 1261–1272.

Eyre, J. R., 1989: Inversion of cloudy radiances by nonlinear optimal estimation I: Theory and simulation for TOVS. *Quart. J. Roy. Meteor. Soc.*, **115**, 1001–1026.

Fischer, J. C., 1987: Passive microwave observing form environmental satellites, A status report based on NOAA's June 1–4, 1987,

NOAA Tech. Rep. NESDIS 35, conference in Williamsburg, VA, 292 pp.

Fleming, H. E., M. D. Goldberg, and D. S. Crosby, 1988: Operational implementation of the minimum variance simultaneous method. Preprints, *Third Conf. on Satellite Meteorology and Oceanography*, Anaheim, CA, Amer. Meteor. Soc., 16-19.

Goldberg, M. D., J. M. Daniels, and H. E. Fleming, 1988: A method for obtaining an improved initial approximation for the temperature/moisture retrieval problem. Preprints, *Third Conf. on Satellite Meteorology and Oceanography*. Anaheim, CA, Amer. Meteor. Soc., 20-23.

Hayden, C. M., 1988: GOES-VAS simultaneous temperature-moisture retrieval algorithm. *J. Appl. Meteor.*, **27**, 705-733.

Mitchell, J. F. B., S. Manabe, V. Meleshko, T. Tokioka, 1990: Equilibrium climate change—and its implications for the future. *Climate Change. The IPCC Scientific Assessment, World Meteorological Organization*, Houghton, J. T., G. J. Jenkins, and J. J. Ephrams, Eds, Cambridge University Press, 131-172.

Ohring, G., J. C. Dodge, 1992: The NOAA/NASA pathfinder program. *Int. Radiation Symp.*, Keevallik and Karner, Eds., Deepak, 405-408.

Spencer, R. W., and J. R. Christy, 1992a: Precision and radiosonde validation of satellite gridpoint temperature anomalies, Part 1: MSU channel 2. *J. Climate*, **5**, 847-857.

——, and ——, 1992b: Precision and radiosonde validation of satellite gridpoint temperature anomalies, Part 2: A tropospheric retrieval and trends during 1979-90. *J. Climate*, **5**, 858-866.

——, and ——, 1993: Precision lower stratospheric temperature monitoring with the MSU: Technique, validation, and results 1979-1991. *J. Climate*, **6**, 1194-1204.

——, ——, and N. C. Grody, 1990: Global atmospheric temperature monitoring with satellite microwave measurements: Method and results 1979-1984. *J. Climate*, **3**, 1111-1128.

Thompson, O. E., and M. T. Tripputi, 1994: NWP initialized satellite temperature retrievals using statistical regularization and singular value decomposition. *Mon. Wea. Rev.*, **122**, 897-926.

Wark, D. Q., 1993: Adjustment of TIROS operational vertical sounder data to a vertical view. NOAA Tech. Rep. NESDIS 64, 36 pp.