

Robustness of Tropospheric Temperature Trends from MSU Channels 2 and 4

CELESTE M. JOHANSON AND QIANG FU

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

(Manuscript received 9 May 2005, in final form 19 December 2005)

ABSTRACT

Tropospheric temperature trends based on Microwave Sounding Unit (MSU) channel 2 data are susceptible to contamination from strong stratospheric cooling. Recently, Fu et al. devised a method of removing the stratospheric contamination by linearly combining data from MSU channels 2 and 4. In this study the sensitivity of the weights of the two channels in the retrieval algorithm for the tropospheric temperatures to the choice of period of record used in the analysis and to the choice of training dataset is examined. The weights derived using monthly temperature anomalies are within about 10% of those obtained by Fu et al. irrespective of the choice of analysis period or training dataset. The trend errors in the retrieved global-mean tropospheric temperatures tested using two independent radiosonde datasets are less than about $0.01 \text{ K decade}^{-1}$ for all time periods of 25 yr or longer with different starting and ending years during 1958–2004. It is found that the retrievals are more robust if they are interpreted in terms of the layer-mean temperature for the entire troposphere, rather than the mean of the 850–300-hPa layer. Because large spurious jumps remain in the reanalyses, especially prior to 1979, one should be cautious when using them as training datasets and in testing the trend errors.

1. Introduction

Estimates of tropospheric temperature trends from the satellite-borne Microwave Sounding Units (MSU) require removal of the stratospheric influence from the MSU channel 2. Fu et al. (2004a, hereafter FJWS) used a linear combination of MSU channels 2 (T_2) and 4 (T_4) to derive the tropospheric temperature trends. Since the effective weighting function for the retrieval of tropospheric temperatures in the tropical region (30°N – 30°S) is near zero throughout the stratosphere (Fig. 1a), the FJWS method allows a physical removal of the stratospheric contamination in the tropical region (Fu and Johanson 2005). Note that a purely physical retrieval of the tropospheric temperature should have an effective weighting function that is zero in the stratosphere and has positive contributions from the troposphere. Although the effective weighting function in the Tropics has some negative/positive contribution from the stratosphere, it is very near to zero and the retrieval is essentially physical.

For the global-mean condition, the removal is accomplished through a statistical cancellation of the negative and positive portions of the contributions from the stratosphere (Fig. 1b). These results have been validated through both observational data and data from climate model simulations. Fu and Johanson (2004) used stratospheric trend profiles based on an independent set of observations to estimate the stratospheric contribution to the trend of MSU channel-2 brightness temperatures and found a contribution of $-0.08 \text{ K decade}^{-1}$, which is consistent with the results of FJWS. Kiehl et al. (2005) alleged that, regardless of the success of a model simulation in predicting the state of the current climate, each simulation provides a realization of the climate system and is therefore useful in testing satellite retrieval algorithms. The FJWS algorithms were tested with climate model simulations, both by Kiehl et al. (2005) and Gillett et al. (2004). In these studies, the FJWS method was successful in capturing the models' tropospheric temperature trends.

Despite these validations, debate continues over the ability of the statistically based global-mean FJWS retrieval to accurately remove the stratospheric contamination (e.g., R. W. Spencer and J. R. Christy 2005, personal communication; NRC 2005; Tett and Thorne

Corresponding author address: Dr. Qiang Fu, Dept. of Atmospheric Sciences, Box 351640, University of Washington, Seattle, WA 98195.
E-mail: qfu@atmos.washington.edu

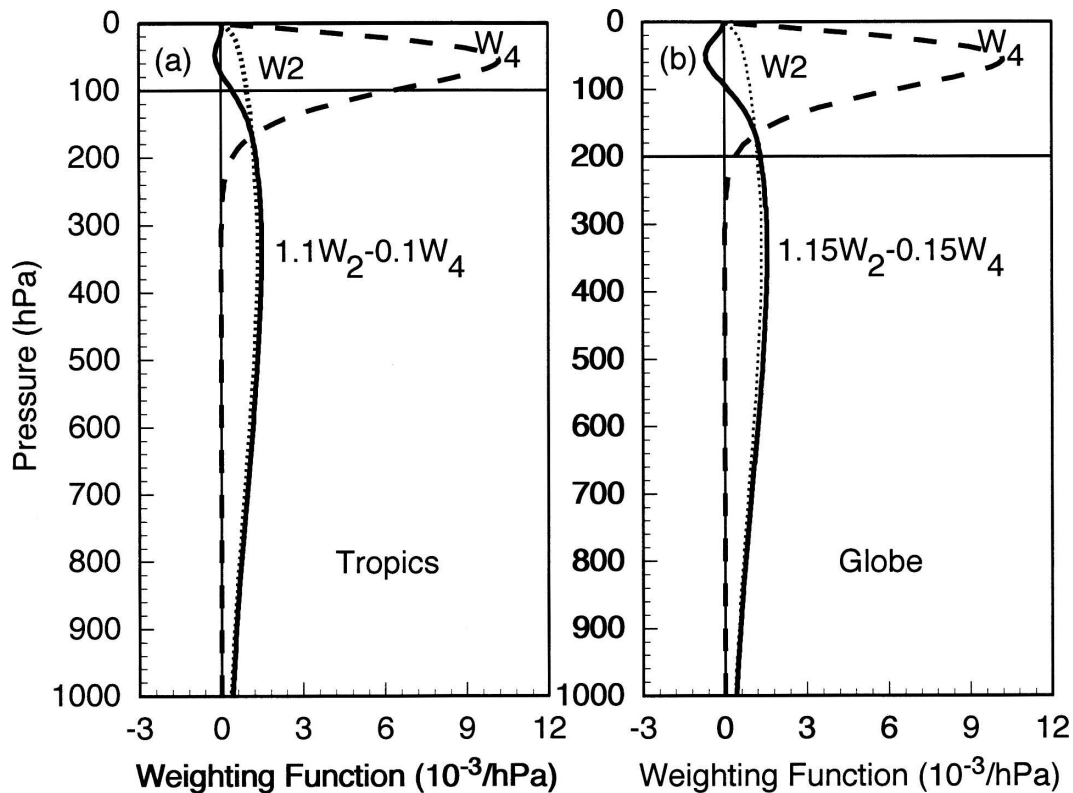


FIG. 1. MSU weighting functions for channels 2 (W_2) and 4 (W_4) and the effective weighting functions for the retrievals of tropospheric temperatures for (a) the Tropics (FJWS; Fu and Johanson 2005) and (b) the globe (FJWS). The horizontal line is located at the mean tropopause pressure.

2004; Fu et al. 2004b). In this study, we will address the issues associated with the MSU-derived global-mean tropospheric temperature trends that are based on the FJWS technique. In addition, we will investigate the application of the FJWS method to the extratropical regions in both the Northern ($>30^\circ\text{N}$) and Southern ($>30^\circ\text{S}$) Hemispheres, which has not been done before.

The tropospheric temperature monthly anomaly (T_{TR}) for the global mean is estimated in FJWS's method by $\tilde{T}_{\text{TR}} = 1.156T_2 - 0.153T_4$. The coefficients, or weights, of the two channels were obtained from a regression relating globally averaged monthly temperature anomalies for the 850–300-hPa layer to the corresponding simulated T_2 and T_4 anomalies, based on the Lanzante–Klein–Seidel (LKS) radiosonde dataset for 1958–97 (Lanzante et al. 2003). One legitimate question regarding the FJWS method is how these regression coefficients, and thus the tropospheric trends based upon them, depend on the choice of training dataset and time period used in the analysis. The interpretation of the FJWS retrievals in the Tropics as the layer-mean temperature for the entire troposphere (Tett and Thorne 2004; Fu et al. 2004b) also raises the question

on the use of the 850–300-hPa layer as the target predictand for the regression. Herein, we will use two independent radiosonde datasets to examine the sensitivity of the regression coefficients and derived tropospheric temperature trends to the choice of training dataset and the period of record used in the analysis. The global-mean, tropical, and extratropical regions will be considered. We will also clarify the implication of the tropospheric target layer used in the development of the retrieval scheme and, thus, the vertical extent of the layer represented by the inferred time series of T_{TR} . Finally, we will discuss the limitations of the reanalyses datasets in deriving satisfactory coefficients and in testing retrieved trend errors due to large spurious jumps.

This paper is organized as follows. Section 2 describes the data and methods used. Section 3 shows the impact of atmospheric variability and the training dataset on the regression coefficients of the retrieval scheme. The errors in retrieved tropospheric temperature trends are examined in section 4. The discussion on the use of reanalyses datasets is given in section 5. The summary and conclusions are given in section 6.

2. Methods and data

Following FJWS, the retrieved tropospheric temperature monthly anomalies using observed brightness temperatures of MSU channels 2 and 4 are in the form

$$\tilde{T}_{\text{TR}} = a_0 + a_2 T_2 + a_4 T_4, \quad (1)$$

where the regression coefficients, a_2 and a_4 (i.e., the weights of the two channels), are computed from a linear least squares fit of the simulated T_2 and T_4 monthly anomalies to a tropospheric target layer temperature (T_{TR}) using a training dataset (e.g., radiosonde data). Simulated T_2 and T_4 are equivalent to those that would be observed by the MSU for a given atmospheric profile. They are simulated by applying the static weighting functions from Christy et al. (1998) to the given vertical anomaly profile. Note that the coefficient a_0 in Eq. (1) is near zero and irrelevant in trend analysis, so we do not consider it. In this study, an additional constraint that $a_2 = 1 - a_4$ is used, which simplifies the analysis with little impact on the results, and leads to a normalized effective weighting function. The trends of T_{TR} , T_2 , and T_4 are removed prior to the regression.

FJWS used the mean temperature between 850 and 300 hPa ($T_{850-300}$) as the tropospheric target layer temperature. The derived global-mean effective weighting function is shown in Fig. 1b. Since the T_4 signal comes mainly from the stratosphere, the shape of the effective weighting function in the troposphere must be very similar to the tropospheric portion of the T_2 weighting function (Fig. 1) regardless of the choice of the target layer temperature. Thus, for the purpose of removing the stratospheric contamination in T_2 , it is natural to define the tropospheric target layer temperature in the form

$$T_{\text{TR}} = \frac{W_s T_s + \int_{1000}^{P_t} W_2(p) T(p) dp}{W_s + \int_{1000}^{P_t} W_2(p) dp}, \quad (2)$$

where W_2 is the T_2 weighting function, W_s the relative contribution of the surface temperature (T_s), $T(p)$ the atmospheric temperature profile, and P_t the tropopause pressure; P_t is 200 hPa for the global mean, 100 hPa for the Tropics (30°S–30°N), and 250 hPa for the extratropics (>30°S or 30°N). (Defining the extratropical tropopause at 300 hPa introduces a negligible change in the results. We used 250 hPa based on the NCEP data, which actually does have a global tropopause at about 175 hPa.)

In the regression of T_{TR} versus T_2 and T_4 in Eq. (1), the large natural variability in the upper troposphere

tends to increase the effective weighting function's relative weight there. This increase also increases the weight somewhat in the lower stratosphere, so the regression may become a trade-off between maximizing the upper-tropospheric information and minimizing stratospheric contamination. The aim of our retrieval is to remove the stratospheric contamination rather than capture the upper-tropospheric variability. Hence, we use $P_t + 25$ hPa instead of P_t in Eq. (2) to derive the target-layer temperatures for the regression. This would affect the regression coefficients in the Tropics but not in the extratropics. Note that the retrieved temperature using Eq. (1) with derived coefficients should be simply interpreted as a deep-layer mean temperature for the entire troposphere weighted by the normalized tropospheric portion of W_2 (or more precisely the effective weighting function of $a_2 W_2 + a_4 W_4$; see Fig. 1).

As will be shown in section 3, the regression coefficients derived from this study are within about 10% of those given in FJWS and these different weights result in a difference of less than 0.01 K decade⁻¹ in the derived tropospheric temperature trends. As we will discuss later in this paper, interpreting the retrieved results as the mean temperatures for the entire troposphere is important in the trend analyses. This is especially true when the retrieval scheme is tested using datasets that have very different trends between the upper troposphere and mid and lower troposphere (Tett and Thorne 2004; Fu et al. 2004b).

We use two independent radiosonde temperature products to examine the dependence of our retrieval methods and trends on the particular training datasets. The LKS dataset, which is derived from individual radiosonde soundings in the Comprehensive Aerological Radiosonde Dataset (CARDS) archive (Eskridge et al. 1995), contains a select global network of 87 stations and terminates in 1997. It has subjectively derived bias adjustments throughout the length of its record. The LKS dataset has recently been extended beyond 1997 by applying a different bias adjustment technique (Free et al. 2004). This updated dataset is the RATPAC (Radiosonde Atmospheric Temperature Products for Assessing Climate: Free et al. 2005). Herein we use RATPAC for the period from 1958 to 2004. The global and regional-mean monthly temperature anomalies are derived following Seidel et al. (2004).

Another radiosonde dataset that we use for 1958–2004 is HadAT, which has recently been constructed by Thorne et al. (2005). HadAT uses a neighbor comparison approach to build spatial as well as temporal consistency throughout the record. This dataset consists of 676 radiosonde stations and is a gridded dataset. We

derive the global- and regional-mean products with the data weighted in proportion to the surface area. Since HadAT includes many more radiosonde stations than the RATPAC and the adjustment and quality control procedures of the two groups sufficiently differ from one another, the resulting datasets are considered to be essentially independent.

Neither the 40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40; Uppala et al. 2005) nor the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) 40-Year Reanalysis Project (Kalnay et al. 1996) datasets will be used as training datasets because large spurious discontinuities are present in these analyzed record, especially prior to 1979. Related issues will be discussed in section 5.

We consider the data time series from 1 January 1958 to 31 December 2004. To test the atmospheric variability/stationarity impact using a given training dataset, we derive the regression coefficients for all time periods at least 25 yr long with different starting and ending years (e.g., 1958–82, 1958–83, . . . , 1958–2004, 1959–83, . . . , 1980–2004). Thus, we have a total of 253 time segments for both datasets. Although these time segments overlap and are therefore not independent of one another, this analysis is statistically meaningful noting that the starting and ending points may have a significant impact on the regression coefficients and trends. [For example, the trend for 1979–2004 is very different from that for 1975–2004 due to the climate regime shift (Trenberth 1990), and the sensitivity to starting and ending points can be also related to El Niño events and major volcanic eruptions.] Deriving regression coefficients from all of these time segments provides a comprehensive check on how the regression coefficients and retrieved-trend errors depend on the stationarity of the atmosphere. Note that our estimated uncertainty range may be biased high due to more sampling for shorter time periods.

3. Impact on regression coefficients

Figure 2 shows the range and mean of a_4 for the globe, Tropics, and extratropics in both the Northern and Southern Hemispheres using the two independent training datasets. The range and mean of a_4 are derived from the regressions as described in section 2 over all 253 time segments for both HadAT and RATPAC.

The range of a_4 in Fig. 2 reflects the extent of the FJWS method's sensitivity to the length and time period of the time series considered, which all have different starting and ending years. For each dataset over a given region, the relative difference between the

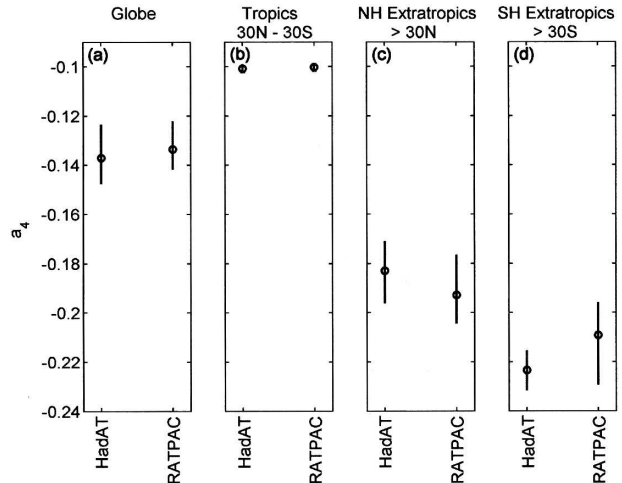


FIG. 2. The range (minimum to maximum) and mean of regression coefficient a_4 based on all time periods at least 25 yr long during 1958–2004 for the retrieval of tropospheric temperatures over (a) the globe, (b) the Tropics (30°N–30°S), and the extratropics for (c) the Northern Hemisphere (>30°N) and (d) the Southern Hemisphere (>30°S).

maximum and minimum a_4 values is less than 15%, indicating little dependence on the atmospheric temporal variability. Note that this relative difference is only about 2% in the Tropics where a physical retrieval is achieved (Fig. 1a). The channel 4 weight also has little dependence on the training dataset. The mean values of a_4 from the HadAT and RATPAC are within 5% of each other.

The longest time segment for deriving the regression coefficients is from 1958 to 2004. Since using a shorter period may increase the uncertainty in the coefficients, we recommend the regression coefficients based on the time segment of 1958–2004 for the tropospheric temperature retrievals. The uncertainties associated with these coefficients should be smaller than the ranges shown in Fig. 2.

Table 1 shows the comparison of a_4 coefficients using HadAT (1958–2004) and RATPAC (1958–2004) from the present study with those from Fu and Johanson (2005) and FJWS. The HadAT and RATPAC coefficients are the same over the globe and Tropics but have a relative difference of ~5% over the extratropical regions. These coefficients are within about 10% of the FJWS values, which is due to the use of different tropospheric target layer temperature in the regression. In the Tropics, the present coefficient is only 2% different from that derived by Fu and Johanson (2005) who used physical/mathematical arguments rather than a training dataset.

Using Eq. (1) and noting that $a_2 = 1 - a_4$, we can estimate the retrieved-trend uncertainty due to the re-

TABLE 1. Comparison of the a_4 coefficients from the present study with those from Fu and Johanson (2005) and Fu et al. (2004a).

	Present study		Fu and Johanson (2005)	FJWS
	HadAT (1958–2004)	RATPAC (1958–2004)		
Globe	−0.141	−0.141		−0.153
Tropics (30°S–30°N)	−0.102	−0.102	−0.10	−0.112
NH extratropics (>30°N)	−0.181	−0.194		
SH extratropics (>30°N)	−0.223	−0.214		

gression coefficient uncertainties by $-a'_4 T_2 + a'_4 T_4$, where a'_4 is the uncertainty in a_4 and T_2 and T_4 are the observed MSU trends. We use the maximum range of a_4 relative to its mean derived over all time periods to estimate the uncertainty in the regression coefficients, which is on the order of 10%. When this uncertainty is applied to observed MSU T_2 and T_4 [e.g., 0.043 and $-0.452 \text{ K decade}^{-1}$, respectively, from the University of Alabama in Huntsville (UAH) (Christy et al. 2003) for the global mean], it leads to an uncertainty of at most $0.01 \text{ K decade}^{-1}$ in the derived tropospheric temperature trends. A similar result is obtained using the Remote Sensing Systems (RSS) T_2 and T_4 trends. Considering observed global-mean MSU trend uncertainties of $0.05 (T_2)$ and $0.1 \text{ K decade}^{-1} (T_4)$ (Christy et al. 2003), along with the coefficient in Table 1, we derive the total uncertainty in the retrieved global-mean tropospheric temperature trend as $[0.01^2 + (1.141 \times 0.05)^2 + (0.141 \times 0.1)^2]^{1/2} = 0.06 \text{ K decade}^{-1}$, which is similar to the observed MSU T_2 trend uncertainty. Thus, the uncertainties in both the regression coefficients and the MSU T_4 make a small contribution to the uncertainty in the retrieved tropospheric trends. In the above calculation, we assume that the T_2 and T_4 errors are independent of each other. Since the errors in T_2 and T_4 tend to be of the same sign [e.g., by comparing the RSS (Mears et al. 2003) and UAH T_2 and T_4 for the structural uncertainties], our retrieved tropospheric trend uncertainty should be slightly smaller than the T_2 trend uncertainty. Note that the retrieved RSS and UAH global tropospheric trends from 1979 to 2004 are 0.113 and $0.194 \text{ K decade}^{-1}$, respectively.

We obtain similar regression coefficients (e.g., -0.141 versus -0.153 for the global mean) when regressing T_2 and T_4 versus quite different target predictands (T_{TR} defined in section 2 versus $T_{850-300}$). This is because the T_{TR} and $T_{850-300}$ monthly anomalies are highly correlated. As shown in Fig. 3 from the LKS dataset, the correlation coefficient is 0.99. In general, $T_{850-300}$ is a good indicator of the tropospheric deep-layer mean temperature. But since T_{TR} from Eq. (2) represents the actual layer mean estimated by the retrieval, it has more of a physical basis for the purpose of

removing the stratospheric influence in T_2 . Thus, our retrievals should be interpreted as the deep-layer mean temperature for the entire troposphere (also see Tett and Thorne 2004; Fu et al. 2004b). Finally, it is found that the correlation coefficients between the retrieved tropospheric temperature monthly anomalies and the target predictands [Eq. (2)] are 0.99 for both radio-sonde datasets.

4. Tropospheric trend errors

We apply the regression coefficient values shown in Table 1 to simulated T_2 and T_4 anomaly time series to obtain tropospheric temperature trends for all different time periods described in section 2. Differences between the retrieved trends and those of the target layer provide an indication of the error in the retrieval. We test the RATPAC 1958–2004 coefficients using both the HadAT and RATPAC datasets. The distributions

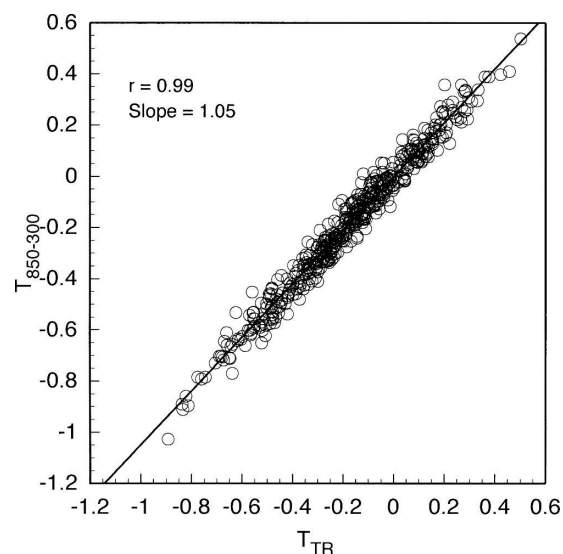


FIG. 3. Global-mean $T_{850-300}$ monthly anomalies vs the T_{TR} monthly anomalies from LKS for 1958–97 [T_{TR} is defined in Eq. (2)]. The slope is 1.05. The high correlation ($r = 0.99$) between the 850–300-hPa layer and the entire troposphere (T_{TR}) results in similar regression coefficients as derived from either layer.

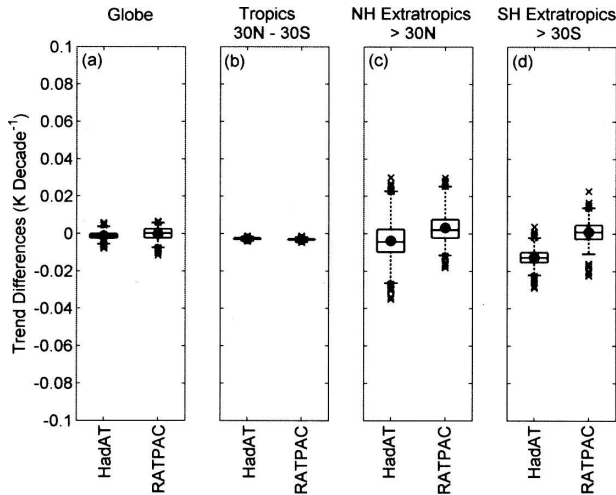


FIG. 4. Distribution of trend differences between retrievals and tropospheric target layer temperatures from all time periods at least 25 yr long during 1958–2004. The vertical boxes represent the interquartile range of each distribution. The horizontal line dividing each box represents the median values, and the mean is represented by a dot. Whiskers extend from the 2.5 percentile to the 97.5 percentile. The “x” represents all cases outside the 2.5 to 97.5 percentile range. The regression coefficients derived from RATPAC for 1958–2004 (see Table 1) are used in all the retrievals. The tropospheric target layer temperature is defined as a deep-layer mean temperature for the entire troposphere weighted by the tropospheric portion of the effective weighting function and normalized.

of trend differences are shown in Fig. 4. Similar results are obtained by applying the HadAT coefficients (not shown here).

We first note that the retrieved-trend errors using the RATPAC coefficients do not depend on whether we use the independent dataset (HadAT) or the same dataset (RATPAC) to test coefficients. [We also find that the trend error distributions associated with the RATPAC 1958–2004 coefficients tested using the HadAT dataset are similar to those obtained by applying the HadAT coefficients from each period to the same HadAT time segment.] This is because 1) the regression coefficients are not sensitive to the trends (remember that we derive regression coefficients using detrended monthly anomalies of T_2 , T_4 and tropospheric target temperatures) and 2) the derived coefficients have little dependence on the training datasets and time periods considered.

The mean errors shown in Fig. 4 are -0.001 (HadAT) and -0.000 (RATPAC) over the globe, -0.003 (HadAT) and -0.003 (RATPAC) over the Tropics, -0.004 (HadAT) and 0.003 (RATPAC) over the NH extratropics, and -0.013 (HadAT) and 0.001 K decade $^{-1}$ (RATPAC) over the SH extratropics. The maximum trend errors over the globe and Tropics are less than

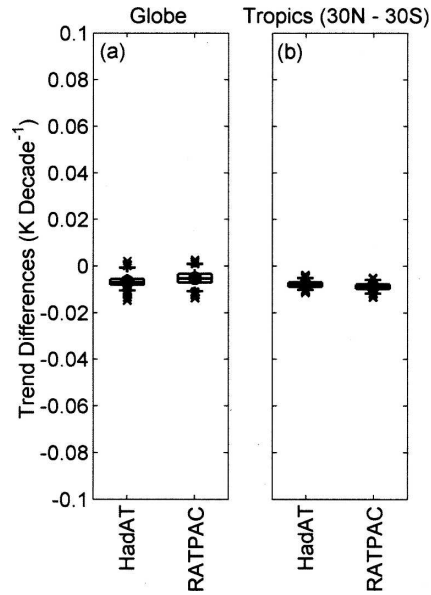


FIG. 5. As in Figs. 4a and 4b but using the weights from Fu et al. (2004a) for the retrievals.

about 0.01 K decade $^{-1}$ while the maximum errors over the extratropical regions are less than about 0.03 K decade $^{-1}$. The small differences between the retrieval trends and the target layer trends shown in Fig. 4 indicate that the retrieval method can effectively capture the tropospheric temperature trends by accurately removing the stratospheric influence in T_2 .

The trend errors shown in Fig. 4 are smallest in the Tropics, become larger in the global mean, and are largest in the extratropics. This result can be explained as follows. For a tropical tropopause of about 100 hPa, a significant portion of the T_4 weighting function arises from the lower stratosphere (Fig. 1a). Thus, it has a significant overlap with the T_2 weighting function, which guarantees a physical removal of the stratospheric contribution in T_2 . As tropopause height decreases, the overlap between the T_2 and T_4 weighting functions in the lower stratosphere also decreases and at the same time more stratospheric information is contained within T_2 . At a tropopause for the global-mean and extratropical regions, a statistical cancellation of the negative and positive portions of the contributions from the stratosphere is then required. But, if the temporal variations of stratospheric temperatures are vertically coherent and hence well correlated with T_4 , the reduced overlap between the two channels' weighting functions is still sufficient to remove the stratospheric contribution in T_2 .

Figure 5 is the same as Figs. 4a and 4b except for using the weights of $(1.156, -0.153)$ over the globe and $(1.123, -0.112)$ over the Tropics from FJWS. The target

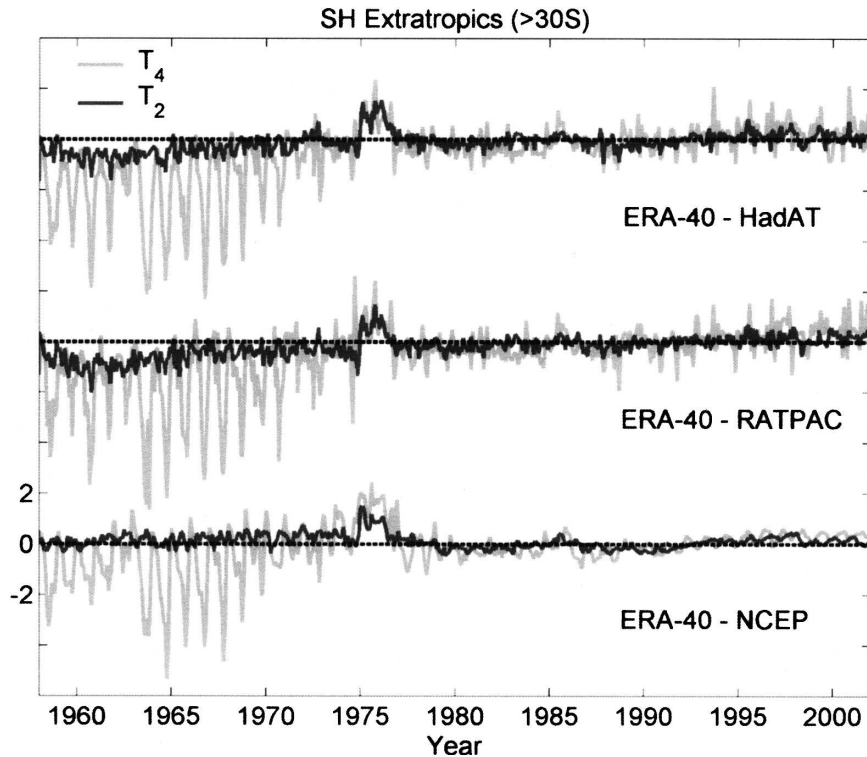


FIG. 6. Difference time series in T_4 and T_2 between ERA-40 and other three datasets including HadAT, RATPAC, and NCEP-NCAR over the Southern Hemisphere extratropical region ($>30^\circ\text{S}$).

predictand is as defined in section 2. The mean and maximum trend errors are within about 0.01 and 0.015 K decade^{-1} , respectively. Note here that although $T_{850-300}$ may usually represent the global-mean tropospheric layer well (FJWS; Gillett et al. 2004; Kiehl et al. 2005), FJWS retrievals should be more precisely interpreted as the T_{TR} that reflects the entire troposphere. This is particularly true when the retrieval scheme is tested using datasets whose trend for the upper troposphere differs substantially from the trend for the mid and lower troposphere (e.g., Tett and Thorne 2004; Fu et al. 2004b).

5. Impact of spurious jumps in reanalyses datasets

The FJWS method unconditionally removes the stratospheric contamination in the tropical region (see Fig. 1), which is physically based and can be tested with any dataset including the reanalyses (e.g., Tett and Thorne 2004; Fu et al. 2004b). For the global-mean condition or over the extratropical regions, the removal is of a statistical nature. It is accomplished through a cancellation of the negative and positive portions of the contributions from the stratosphere. This requires that

the temporal variations of stratospheric temperatures are vertically coherent and well correlated with variations in T_4 . Such a cancellation has been tested in the present study using two independent radiosonde datasets with different time segments. However we did not use the reanalyses datasets for this purpose since artificial jumps of different magnitudes at different altitudes violate the vertical coherency in the stratosphere.

As an example, Fig. 6 shows the difference time series of ERA 40 – HadAT, ERA 40 – RATPAC, and ERA 40 – NCEP-NCAR for the simulated T_4 and T_2 over the extratropics in the Southern Hemisphere. Relative to other datasets, the ERA-40 T_4 has large annual oscillations (>4 K) prior to 1970s. In the mid-1970s, its T_4 temperature anomalies dramatically increased, plateaued for a few years, and then dropped to a higher level than the anomalies prior to the sudden jump. This leads to an increase of about 2 K between the pre-1973 and post-1977 anomalies in the ERA-40 simulated T_4 . These ERA-40 behaviors are very likely artificial because they are aliased into the difference time series between the ERA-40 and all other datasets. The jump also occurs in the ERA-40 T_2 relative to the

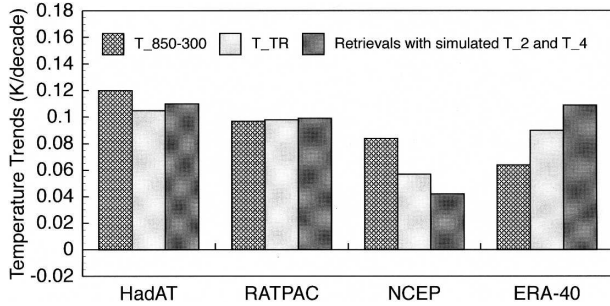


FIG. 7. Trends in global-mean monthly temperature anomalies of the layer between 850 and 300 hPa ($T_{850-300}$), the layer for the entire troposphere [T_{TR} as defined in Eq. (2)], and the MSU retrievals using the simulated T_2 and T_4 with the weights ($a_2 = 1.141$, $a_4 = -0.141$).

other three datasets, but its magnitude is much smaller. By checking the difference time series at different pressure levels in the stratosphere, we find that the magnitude of the jump in temperature anomalies varies from pressure level to pressure level. This indicates that the vertical structure of the ERA-40 anomalies may not be vertically coherent any more in the stratosphere due to such artificial jumps. We conclude that reanalysis datasets are less reliable for both deriving regression coefficients and testing the retrieved-trend errors because large different artificial jumps remain in these datasets.

J. R. Christy (2004, personal communication) showed large global-mean trend differences by comparing FJWS retrievals and $T_{850-300}$ using the NCEP-NCAR and ERA-40 data for 1979–2001. Similar results are shown in Fig. 7 for 1979–2004. We do see the large trend difference between $T_{850-300}$ and our retrievals for each reanalysis data. However, more than half of these differences can be explained by the large cooling between the tropopause and 300 hPa in the NCEP-NCAR dataset and large heating of this layer in the ERA-40 dataset (see Fig. 8), which leads to trends of the entire troposphere substantially smaller and larger than the $T_{850-300}$ trends, respectively. [Of course for 1979–2004 the strong upper-tropospheric cooling in the NCEP-NCAR data and the strong heating in the ERA-40 data cannot both be correct, and they may be both wrong although these reanalyses become much more reliable after 1979.] As shown in Fig. 7, the trend differences between our retrievals with simulated T_2 and T_4 and the deep-layer mean temperatures for the entire troposphere are less than $0.005 \text{ K decade}^{-1}$ for HadAT and RATPAC and about $0.02 \text{ K decade}^{-1}$ for NCEP-NCAR and ERA-40. The larger trend differences between the retrievals and tropospheric target temperatures for the reanalyses are due to their unrealistic

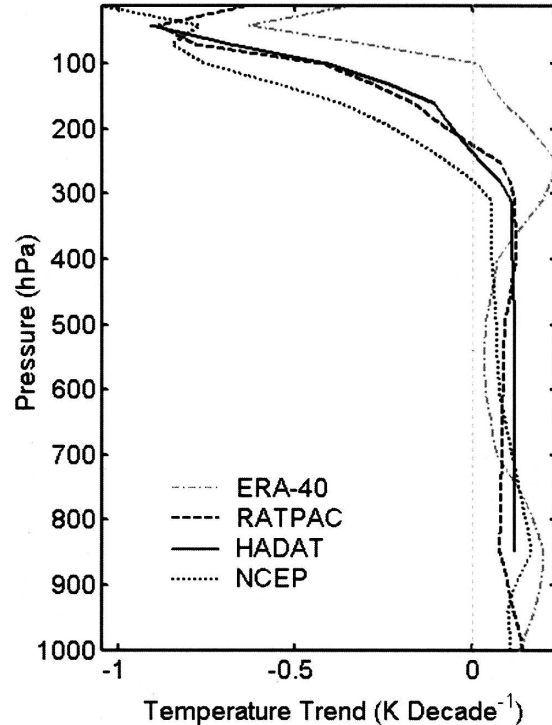


FIG. 8. Global-mean vertical temperature trend profiles for 1979–2004 from HadAT, RATPAC, and NCEP-NCAR and for 1979–2001 for ERA-40.

vertical trend structures above 200 hPa, which can be caused by the huge heating and cooling trends in the tropical upper troposphere associated with the ERA-40 and NCEP-NCAR datasets, respectively.

6. Summary and conclusions

FJWS used a linear combination of MSU channels 2 and 4 to derive the tropospheric temperatures. This method relies on the ample overlap between MSU channels 2 and 4 to remove stratospheric contamination in T_2 . Although a physical removal of the stratospheric influence is achieved in the Tropics, for the global-mean and extratropical tropospheric temperatures a statistical relationship must be derived based upon ancillary training datasets.

In this study, we use two independent radiosonde datasets to examine the sensitivity of the regression coefficients to the choice of analysis dataset and the period of record used in the analysis. To test the sensitivity to the period of record, we consider all time periods of 25 yr or longer with different starting and ending years for each dataset during 1958–2004. We find that the derived weights of the two channels are insensitive to both the time period considered and the choice of

dataset. For the global mean, the weights (a_2 , a_4) obtained using the longest time segment from the present study are (1.141, -0.141); they are within 10% of those from FJWS (1.156, -0.153).

Applying regression coefficients (1.141, -0.141) to the two datasets for all the different time periods considered, the global-mean trend errors are found to be less than about $0.01 \text{ K decade}^{-1}$. The small differences in the trends between the retrievals and the target layer temperatures for a wide range of vertical trend structures considered in this analysis demonstrate that the FJWS method can accurately remove the stratospheric contribution from T_2 . By analyzing the ERA-40 over the SH extratropical region as an example, we suggest that reanalyses datasets are less reliable for both deriving the retrieval regression coefficients and testing the retrieved-trend errors. This is because the large artificial jumps in reanalyses violate the vertical coherency in the stratosphere.

We have clarified the implication of the tropospheric target layer used in the development of the retrieval scheme. We find that the mean temperature of the entire troposphere weighted by the tropospheric portion of the T_2 weighting function is more physically based for the regression to remove the stratospheric contribution in T_2 . The FJWS retrievals should be interpreted as the deep-layer mean temperature representing the entire troposphere. Because the average differences between the retrieval trends and the trends for such a layer are near zero, we have confidence that the retrieved tropospheric trend is free of stratospheric contamination.

Acknowledgments. We thank J. M. Wallace, K. E. Trenberth, D. L. Hartmann, S. G. Warren, and D. J. Seidel for valuable discussions, and we specifically thank Wallace, Trenberth, Warren, and Seidel for useful comments and suggestions on the manuscript. We thank M. Free and P. Thorne, respectively, for providing the RATPAC and HadAT datasets. The comments of S. Tett and an anonymous reviewer were helpful in improving the manuscript. This study is supported by NOAA Grant NA17RJ1232.

REFERENCES

- Christy, J. R., R. W. Spencer, and E. S. Lobl, 1998: Analysis of the merging procedure for the MSU daily temperature time series. *J. Climate*, **11**, 2016–2041.
- , —, W. B. Norris, W. D. Braswell, and D. E. Parker, 2003: Error estimates of version 5.0 of MSU-AMSU bulk atmospheric temperatures. *J. Atmos. Oceanic Technol.*, **20**, 613–629.
- Eskridge, R. E., O. A. Alduchov, I. V. Chernykh, Z. Panmao, A. C. Polansky, and S. R. Doty, 1995: A Comprehensive Aerological Reference Data Set (CARDS): Rough and systematic errors. *Bull. Amer. Meteor. Soc.*, **76**, 1759–1775.
- Free, M. J., J. K. Angell, I. Durre, J. Lanzante, T. C. Peterson, and D. J. Seidel, 2004: Using first differences to reduce inhomogeneity in radiosonde temperature datasets. *J. Climate*, **17**, 4171–4179.
- , D. J. Seidel, J. K. Angell, J. Lanzante, I. Durre, and T. C. Peterson, 2005: Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new dataset of large-area anomaly time series. *J. Geophys. Res.*, **110**, D22101, doi:10.1029/2005JD006169.
- Fu, Q., and C. M. Johanson, 2004: Stratospheric influences on MSU-derived tropospheric temperature trends: A direct error analysis. *J. Climate*, **17**, 4636–4640.
- , and —, 2005: Satellite-derived vertical dependence of tropical tropospheric temperature trends. *Geophys. Res. Lett.*, **32**, L10703, doi:10.1029/2004GL022266.
- , —, S. G. Warren, and D. J. Seidel, 2004a: Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, **429**, 55–58.
- , D. J. Seidel, C. M. Johanson, and S. G. Warren, 2004b: Reply. *Nature*, **432**, doi:10.1038/nature03210.
- Gillett, N. P., B. D. Santer, and A. J. Weaver, 2004: Stratospheric cooling and the troposphere. *Nature*, **432**, doi:10.1038/nature03209.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kiehl, J. T., J. M. Caron, and J. J. Hack, 2005: On using global climate model simulations to assess the accuracy of MSU retrieval methods for tropospheric warming trends. *J. Climate*, **18**, 2533–2539.
- Lanzante, J. R., S. A. Klein, and D. J. Seidel, 2003: Temporal homogenization of monthly radiosonde temperature data. *J. Climate*, **16**, 224–262.
- Mears, C. A., M. C. Schabel, and F. J. Wentz, 2003: A reanalysis of the MSU channel 2 tropospheric temperature record. *J. Climate*, **16**, 3650–3664.
- NRC, 2005: Review of the U.S. Climate Change Science Program's synthesis and assessment product on temperature trends in the lower atmosphere. National Research Council Rep., National Academies Press, 74 pp.
- Seidel, D. J., and Coauthors, 2004: Uncertainty in signals of large-scale climate variations in radiosonde and satellite upper-air temperature datasets. *J. Climate*, **17**, 2225–2240.
- Tett, S. F. B., and P. Thorne, 2004: Tropospheric temperature series from satellites. *Nature*, **432**, doi:10.1038/nature03208.
- Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper-air temperatures from 1958 to 2002. *J. Geophys. Res.*, **110**, D18105, doi:10.1029/2004JD005753.
- Trenberth, K. E., 1990: Recent observed interdecadal climate changes in the Northern Hemisphere. *Bull. Amer. Meteor. Soc.*, **71**, 988–993.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 reanalysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.