TAO, Vol. 8, No. 1, 95-110, March 1997

Systematic Error Correction of GMS-4 IR Data

Gin-Rong Liu[1] and Wann-Jin Chen[2]

(Manuscript received 5 August 1996, in final form 23 October 1996)

ABSTRACT

About 7 to 10 anomalous peaks are found in the histogram analysis of a Geostationary Meteorological Satellite (GMS-4) infrared (IR) image. These extra peaks, called systematic errors, need to be corrected before the GMS-4 image data are applied in the study of cloud climatology, rainfall estimation and hydrologic cycles. Because an Advanced Very High Resolution Radiometer (AVHRR) IR image is free of such systematic errors, it serves as a reference for GMS-4 IR data correction in this study. It is noted that when comparing the pixels without systematic errors in GMS-4 IR image data with the corresponding National Oceanic and Atmospheric Administration (NOAA) AVHRR Channel 4 data, a good linear relationship emerges. In view of the effects caused by different observation times in order to obtain a stable relationship between the GMS-4 IR brightness temperature (Tb) and AVHRR IR Tb, it is necessary that scattering criteria be set when selecting the samples used to determine the regression equation. The results show that if the histogram's anomalous peak samples located outside the 95% confidence interval are corrected by applying the regression equation, the histogram's anomalous peaks almost disappear.

(Key words: AVHRR data, GMS-4 data, Systematic error)

1. INTRODUCTION

One geostationary weather satellite can monitor the weather of about one-third of the Earth's full disk. Therefore, combining all the geosychronous weather satellite observations can explicitly monitor weather on a global scale. Since July 1983, the International Satellite Cloud Climatology Project (ISCCP, Rossow *et al.*, 1985) has collected global satellite-measured radiance expecting to investigate global cloud climatology and the effect of cloud on the Earth's radiation budget (Schiffer and Rossow, 1983).

The quality of satellite measurements influences the accuracy of data applications, which means that before radiation data are used, it is necessary that they be checked in advance to ensure that such they are reasonable. Lin and Lin (1993) found five anomalous peaks on the

[1]Center for Space and Remote Sensing Research, National Central University, Chung-Li, Taiwan, R.O.C.
[2]Department of Applied Physics., Chung Cheng Institute of Technology, Tashi, Tao Yuan, Taiwan, R.O.C.

histogram of GMS-4 IR image, when they used those satellite data to analyze the cloud patterns of some weather systems. In this study, it was found that there are about 7 to 10 anomalous peaks on the GMS-4 IR histogram observations. These anomalous peaks are referred to as systematic errors and may be explained as discontinuities arising from the conversion of the detector voltages to digital counts.

The correction of systematic errors is more urgent and important than the investigation of the causes in that these errors distort research results, such as those pertaining to rainfall estimation, cloud climatology and earth-atmospheric energy budget studies that apply GMS-4 IR data. In this paper, the AVHRR data consistent with GMS-4 IR image location and observation time are used to correct systematic errors.

2. DATA COLLECTION AND PROCESSING

GMS-4's Visible and Infrared Spin Scan Radiometer (VISSR) has 2 spectral channels, visible (0.5-0.75 μm) and infra-red (10.5-12.5 μm), with 1.25 km and 5km resolution at the nadir, respectively. The NOAA series' AVHRR has 5 spectral bands (0.58-0.68, 0.725-1.1, 3.55-3.93, 10.3-11.3 and 11.5-12.5 μm) with 1.1 km resolution at the nadir.

The geographic coordinates of a GMS-4 image can be calculated by using orbit and attitude prediction data contained in the raw data for the VISSR, and IR image calibration obtained by using the black-body shutter count, space shutter, black-body shutter temperature and scanner temperature (Murayama, 1989). The spectral response function is considered in the calculation of GMS-4 IR radiance (Sasaki *et al.*, 1990). The geographic coordinates of an AVHRR image are calculated using geometrical theory with given ephemeris prediction data (Tseng, 1988). The AVHRR IR image calibration is estimated by using the temperature of space and the internal black body target (Lauritson *et al.*, 1979).

AVHRR Channel 4 observations are selected as references for the correction of systematic errors in a GMS-4 IR image. Owing to the differences in satellite altitude, attitude and spatial resolution, for the purposes of comparison, it is necessary that the original images be redone so that they have the same spatial resolution and registration. The size of the re-sampled images is chosen as 10 degrees latitude and longitude, ranging from 15 to 25 degrees north and from 115 to 125 degrees east. The pixel size is 0.05 by 0.05 degree's latitude and longitude, thereby giving the re-sampled images 200 lines with 200 pixels to a line. The Tb of the GMS-4 re-sampled image is obtained using the bi-linear interpolation method, while that of the AVHRR re-sampled image is averaged from 5 by 5 pixels of the original image centered on the position of the re-sampling grid. Figures 1 and 2 are sample images of GMS-4 and NOAA AVHRR Channel 4 IR images. Figures 3 and 4 are the images restructured from those shown in Figures 1 and 2 mentioned above. On the whole, position matching is satisfied by viewing the cloud pattern but some fine deviations can still be found between the resampled images when the two IR images are superimposed.

Such deviations may be caused by the misregistration errors between different satellites. This problem can be solved by imposing the AVHRR IR image on the GMS-4 IR image and horizontally shifting until the minimum root mean square error (RMSE) of the difference between the corresponding Tb is identified. The location shift estimated by this registration

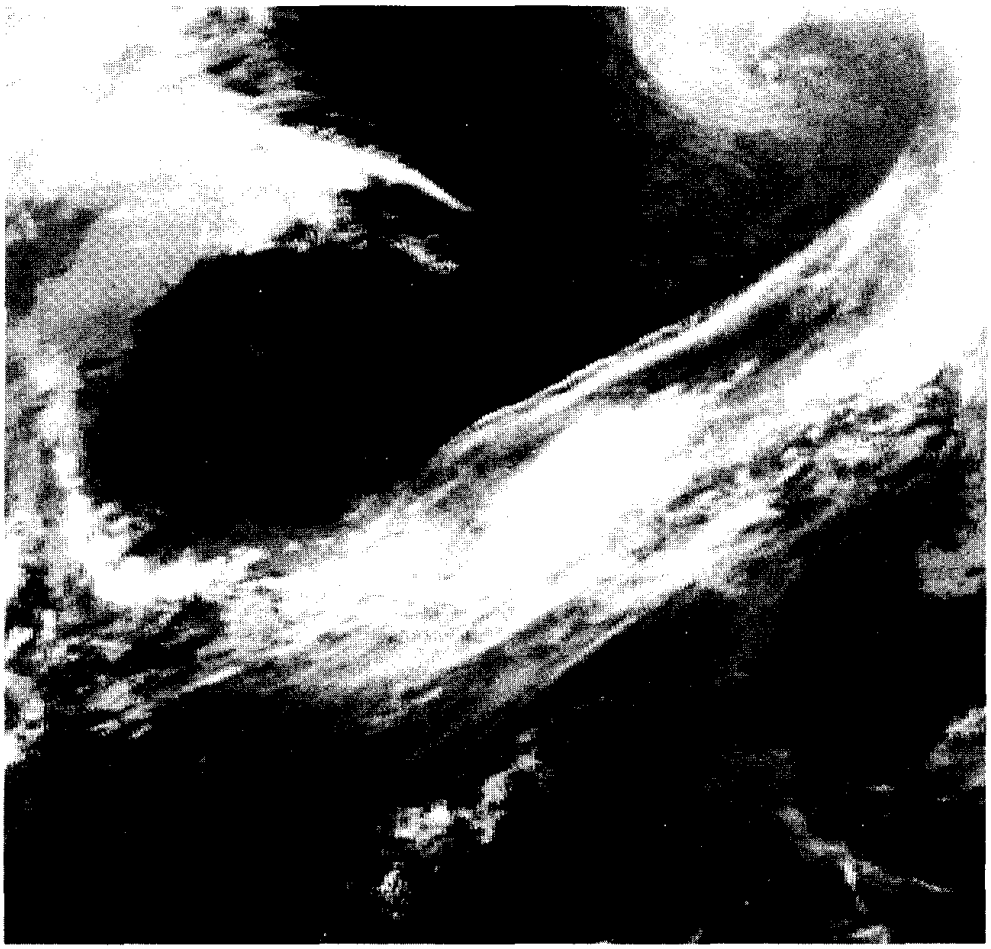

Fig. 1. GMS-4 IR image observed at 00:32 GMT June 2, 1993.

technique should be added to the re-sampled images to correct the registration of the two images. In general, the shift is very small.

3. CHARACTERISTICS OF SYSTEMATIC ERROR AND A CORRECTION METHOD

The anomalous peaks found on the histogram of a GMS-4 IR image is defined as follows: the Tb of the GMS-4 IR image is divided into 256 grey levels, from digital count 0 to 255, and each represents a radiance or Tb. On the histogram of the GMS-4 IR image digital count , if a digital count whose pixel numbers divided by the total number of pixels is larger than 0.001 and more than one and a half times its nearby digital count numbers, then this particular digital count is called an anomalous peak, its ratio is called the peak ratio, and its Tb is identified as the peak Tb.

Table 1 shows the locations of anomalous peaks on the histogram of a GMS-4 IR image for 4 regions representing different weather conditions at an observation time. It is evident that the locations of peak count are all the same for different observed regions at the same observation time. Table 2 is the same as Table 1 but for the same observation regions at different

Fig. 2. AVHRR Channel 4 IR image (see text) observed at 00:40 GMT June 2, 1993.

observation times. Unlike Table 1, the locations of peak counts change with time. From Tables 1 and 2, it is found that the location of an anomalous peak changes from time to time.

Figures 5 (a)-(h) represent the distribution of the peaks Tb, 291.5K, 286.4K, 280.0K, 273.8K, 266.0K, 257.5K, 246.0K and 228.0K, respectively. The peak counts are distributed structurally instead of randomly as noise. For this reason, the anomalous peak counts cannot simply just be picked out from an image as noise instead of correcting those anomalous peak pixels.

Because the histogram of an AVHRR IR image is free of these anomalous pixels (see Figure 6), it can serve as a reference for the correction of GMS-4 observations. At first, to determine their relationship using the linear regression technique the normal IR Tb data pairs are selected as samples, i.e. one GMS-4 IR Tb and its corresponding AVHRR Channel 4 Tb representing one pair. From the derived regression equation, the GMS-4 IR Tb can be estimated from their corresponding AVHRR IR Tb. Using this relationship, those GMS-4 pixels whose digital count is located far from the regression line are regarded as bad data and are, therefore, corrected. A certain confidence interval, such as 95%, of the linear regression analysis is chosen as the threshold to distinguish bad data. The equation for the confidence interval can be written as :

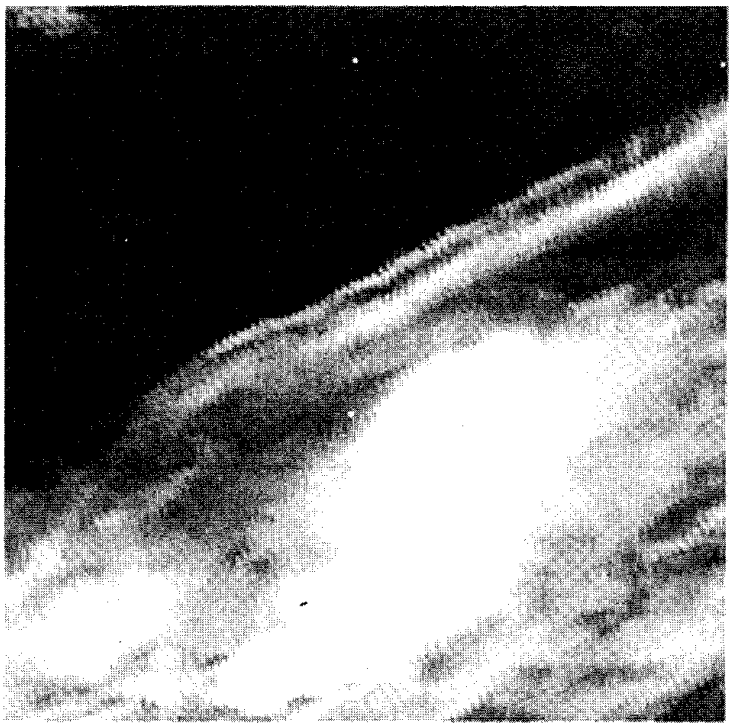

Fig. 3. Same as Figure 1, but re-sampled. Image size is from 115°E to 125°E and from 15°N to 25°N.

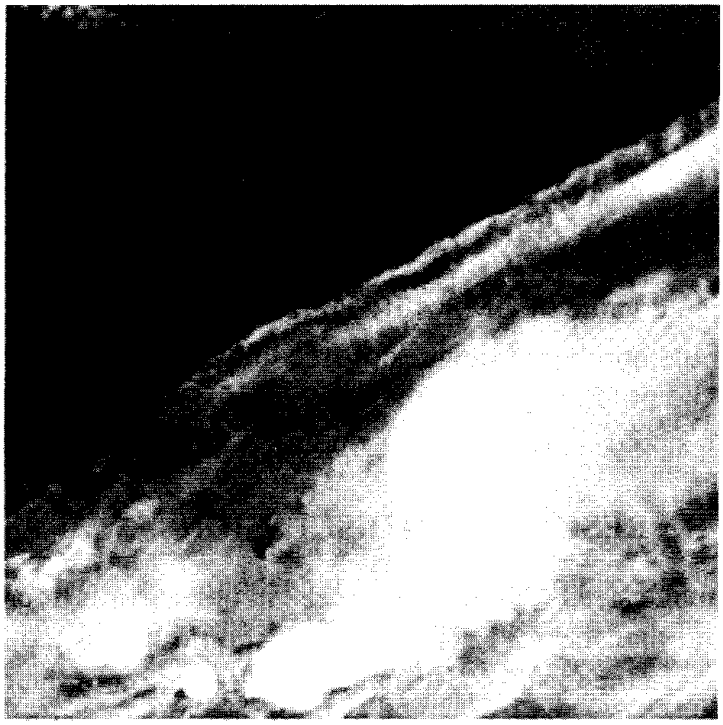

Fig. 4. Same as Figure 2, but re-sampled. Image size is the same as Figure 3.

Table 1. Peak Tbs for different regions but with the same observation time (00:32 GMT June 2, 1993).

	PEAK Tb(K)								
REGION	1	2	3	4	5	6	7	8	9
A	296.4	291.5	286.4	280.0	273.8	266.0	257.5	246.0	228.8
B		291.5	286.4	280.0	273.8	266.0	257.5	246.0	228.8
C		291.5	286.4	280.0	273.8	266.0	257.5	246.0	228.8
D		291.5	286.4	280.0	273.8	266.0	257.5	246.0	228.8

Table 2. Peak Tbs for same regions but with different observation dates.

	PEAK Tb(K)								
DATE	1	2	3	4	5	6	7	8	9
93/05/28	293.3	288.2	282.0	275.4	268.3	259.4	247.5	231.5	
93/05/31	295.5	290.6	284.4	278.0	270.5	261.9	251.0	235.9	
93/06/02	296.4	291.5	286.4	280.0	273.8	266.0	257.5	246.0	228.8

$$\forall_p \pm t(v,\alpha)s\sqrt{1+1/n+(x_p-\sum x/n)^2/(\sum x^2-(\sum x)^2/n)}$$

where $\forall_p$ is the predicted value of GMS-4 IR Tb; t is the value of the student's t-distribution which is a function of the number of degrees of freedom, v; the confidence interval is α; s is the standard error of the prediction; n is the number of samples; x represents all of the observed AVHRR IR Tb's; and x_p is the corresponding AVHRR IR Tb's as a predictor.

4. ANALYSIS OF THE RESULTS AND DISCUSSION

The re-sampled image size is 200 by 200 pixels; in other words, there are 40,000 pixels on each image. It is found that the AVHRR IR image should be shifted 2 lines upward and 9 pixels leftward relative to the GMS-4 IR image for registration in the case (Case A) used in this study (see Figures 1 and 2). The numbers of pixels that can therefore be matched between GMS-4 and AVHRR images are reduced from 40,000 to 37,816. Figure 7 shows the scatter diagram of GMS-4 IR Tb versus AVHRR IR Tb of the normal data pairs (33,675 in total). Several horizontal blank lines are detected on the figure, but these are due to the exclusion of

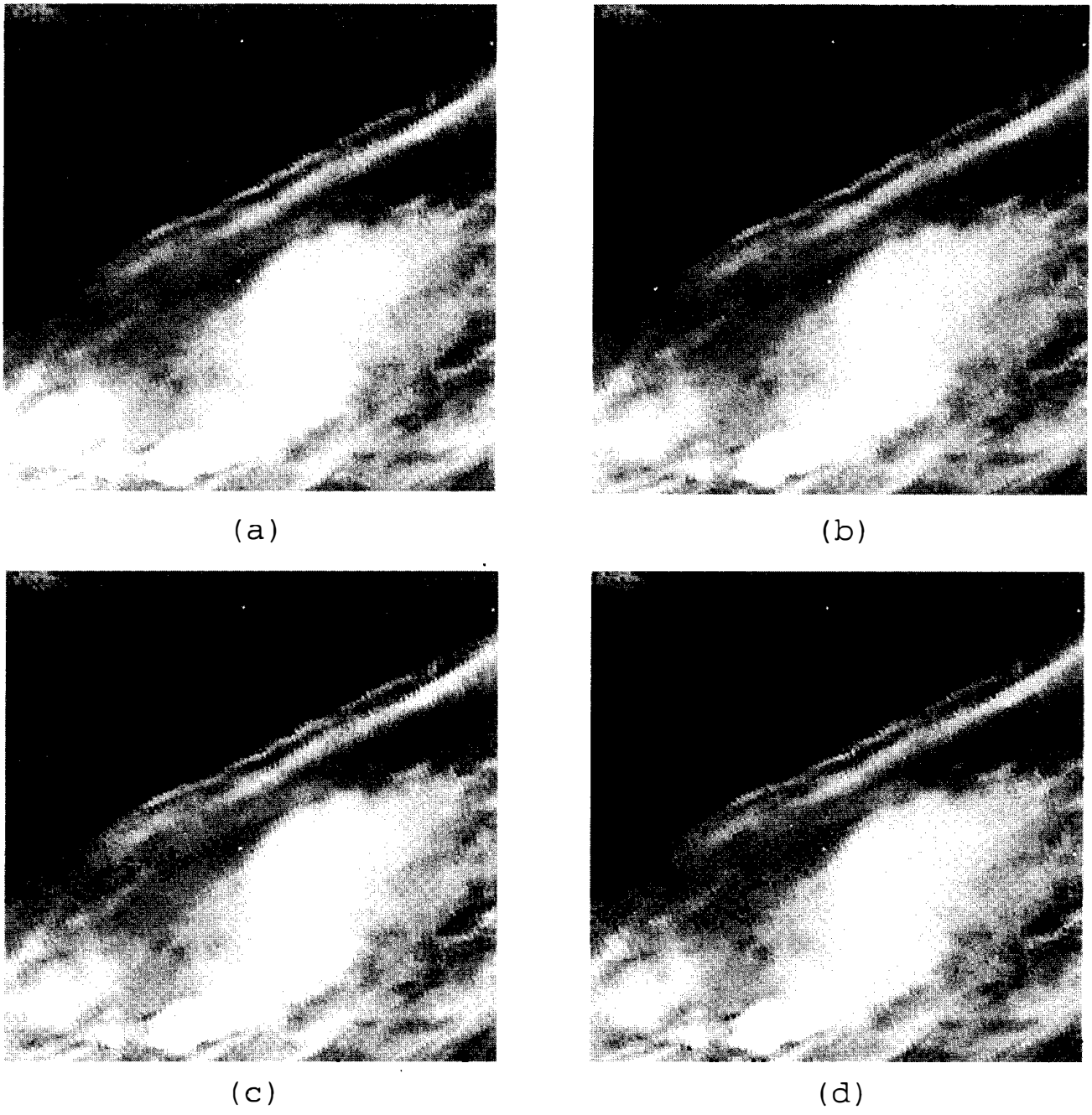

Fig. 5. Distribution of peaks Tb (black points on the image) (a) Tb 291.5 K. (b) Tb 286.4K. (c) Tb 280.0 K. (d) Tb 273.8 K.

those pixels whose digital counts are located on the anomalous peaks. The regression line of all the samples is shown in the centered solid line on the figure, and the upper and lower 95% confidence interval bands are shown in lines above and below the regression line. Figure 8 shows the same scatter diagram as in Figure 7, but only the peak pixels are shown. There are 505 sample points located outside the 95% confidence interval band that should be corrected. Figure 9 presents the Tb histogram of a GMS-4 IR image after correction and compares it to that of the AVHRR image. It can be seen that although the sharpness of the peak ratio de-

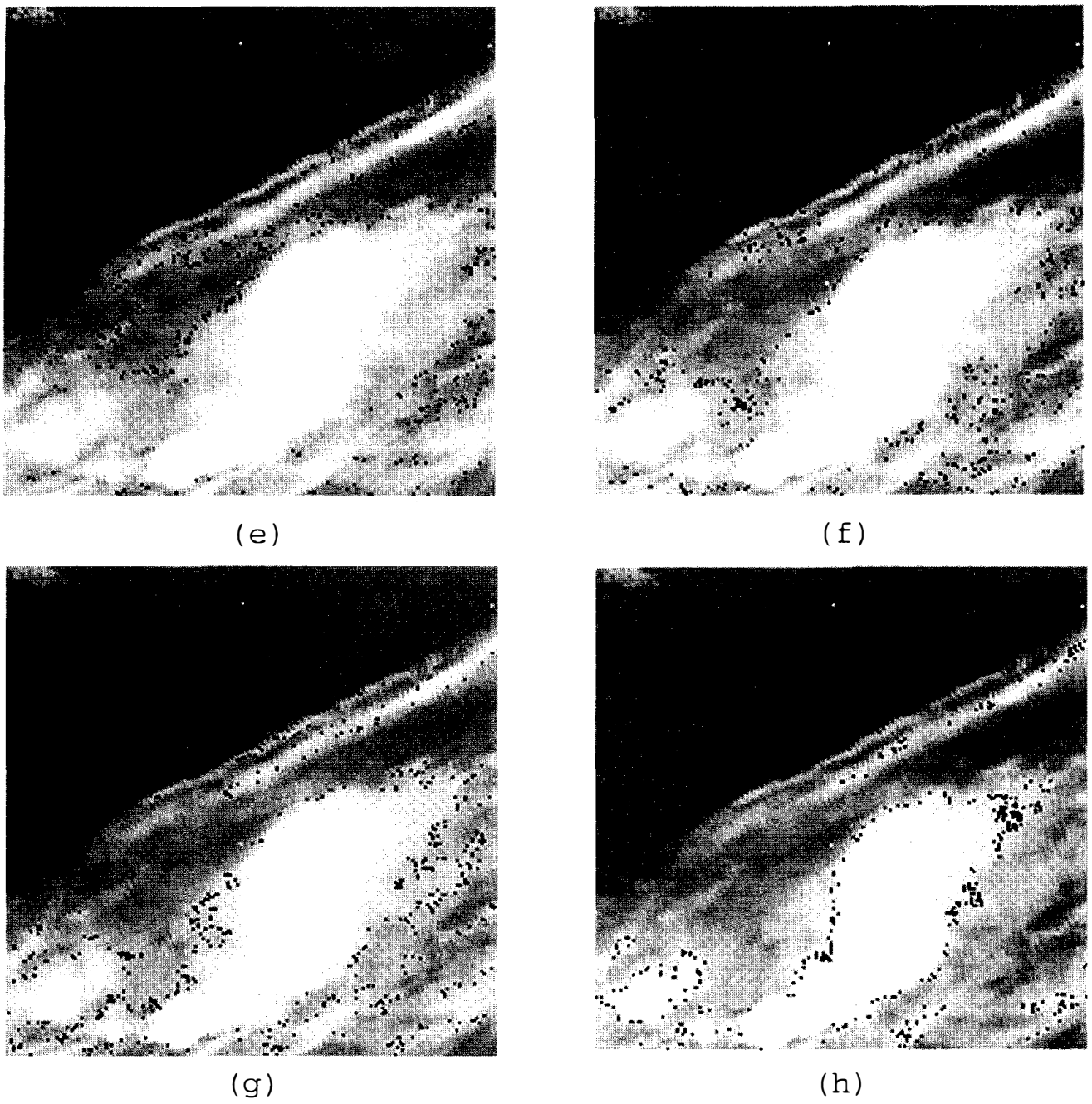

Fig. 5. (e) Tb 266.0 K. (f) Tb 257.5 K. (g) Tb 246.0 K. (h) Tb 228.0 K.

creases, it remains larger than that of its neighbors. It is evident that more of the peak pixels should be corrected in order to smooth out the sharpness of the peak ratio. Figures 10 and 11 are also the same as Figures 8 and 9 but for an 80% confidence interval. When this confidence interval is used, the corrected pixels increase to 1,727, while the sharpness of the peak ratio almost disappears. From a series of comparisons from 95% to 70% confidence intervals the 80% interval may be chosen as the optimal criteria which distinguishes bad GMS-4 IR data provided that all data samples fit the regression line. In addition to the correction of peak pixels located outside the confidence interval, even those not in the peak count but located

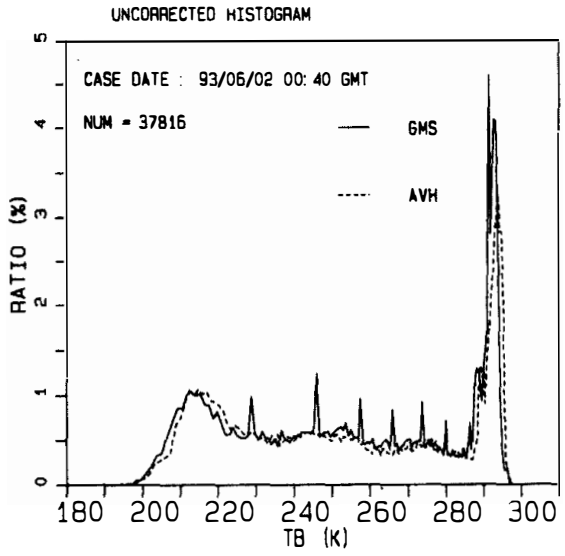

Fig.6. Histogram analysis of IR Tb for a resampled GMS-4 IR image of Figure 3 (in solid line) and AVHRR IR image of Figure 4 (in dashed line).

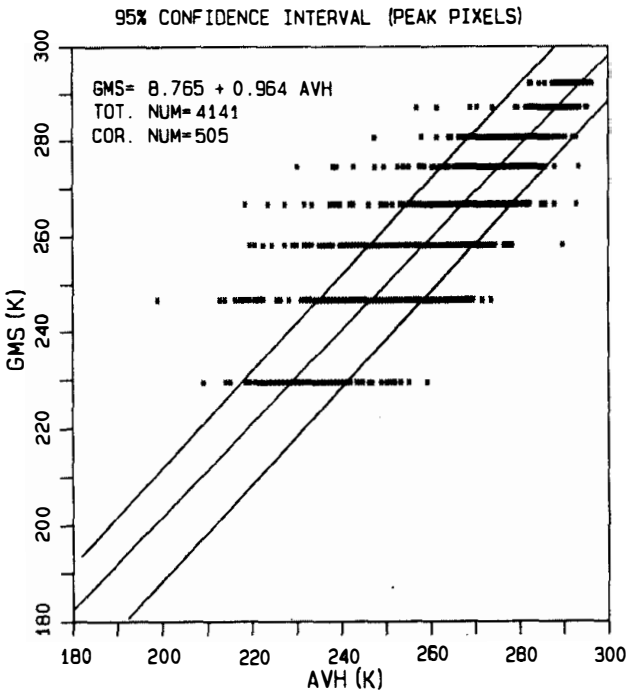

Fig.8. Same as Fig.7, but for peak pixels.

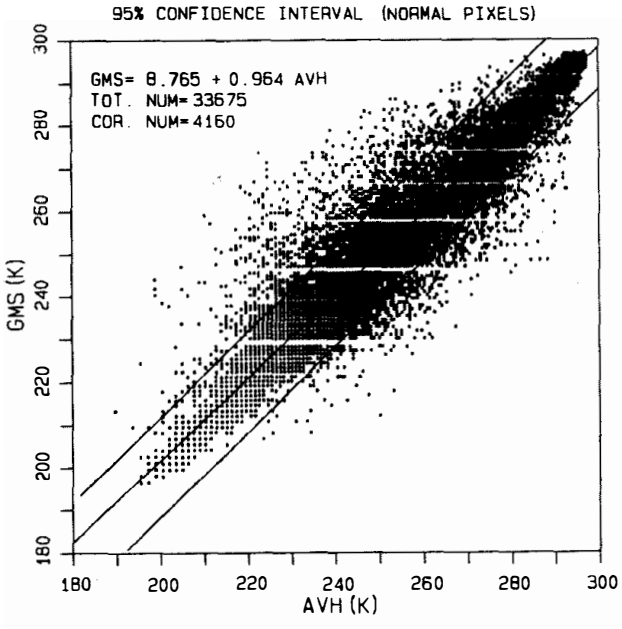

Fig.7. Scatter diagram of GMS-4 IR Tb versus AVHRR IR Tb of the normal data pairs. The lines are the regression line (center) and the 95% confidence interval band (upper and lower lines), respectively. The COR. NUM is the number of pixels located outside the confidence interval band.

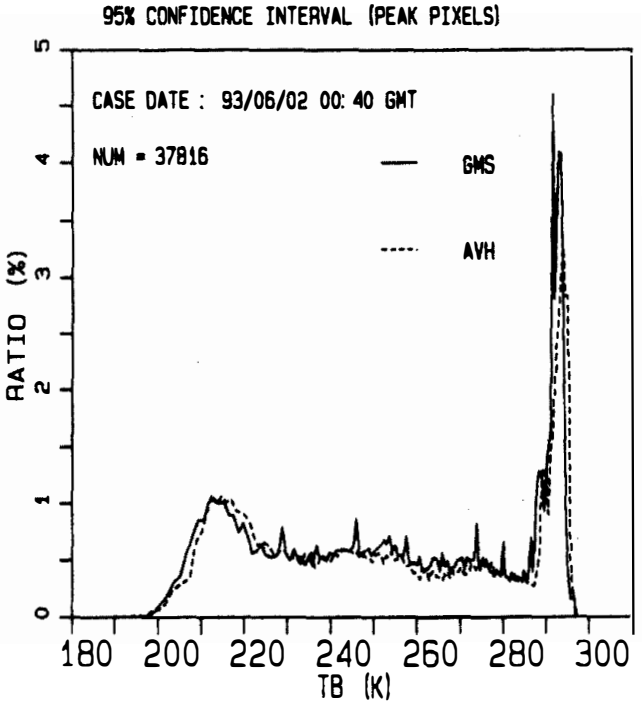

Fig.9. Same as Fig.6, but only the peak pixels ouside the 95% confidence interval band are corrected.

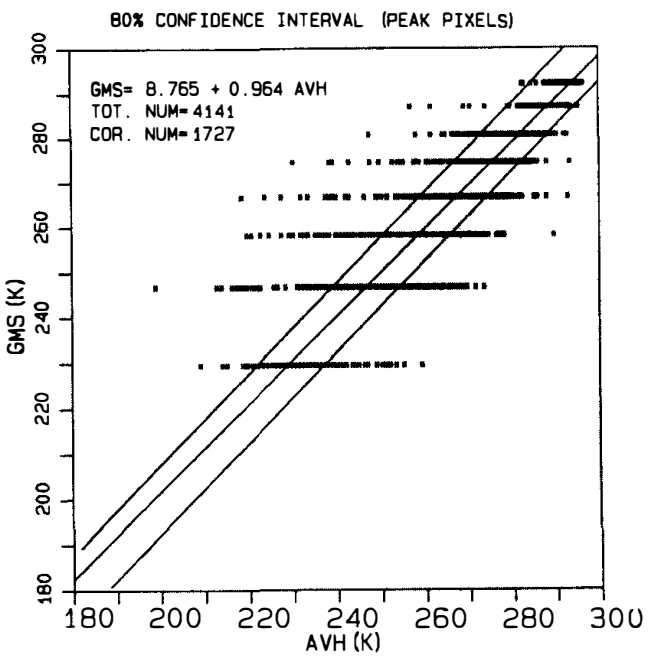

Fig. 10. Same as Fig. 8, but for an 80% confidence interval.

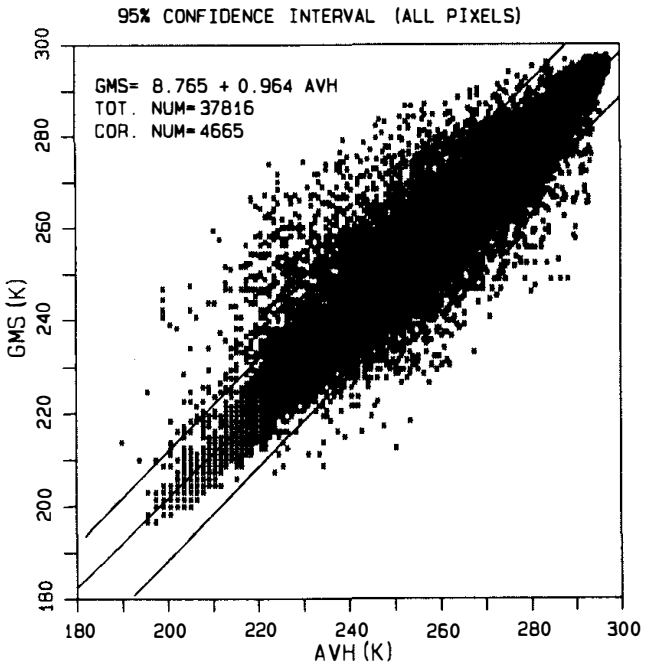

Fig. 12. Same as Fig. 8, but all pixels outside the 95% confidence interval band are corrected,

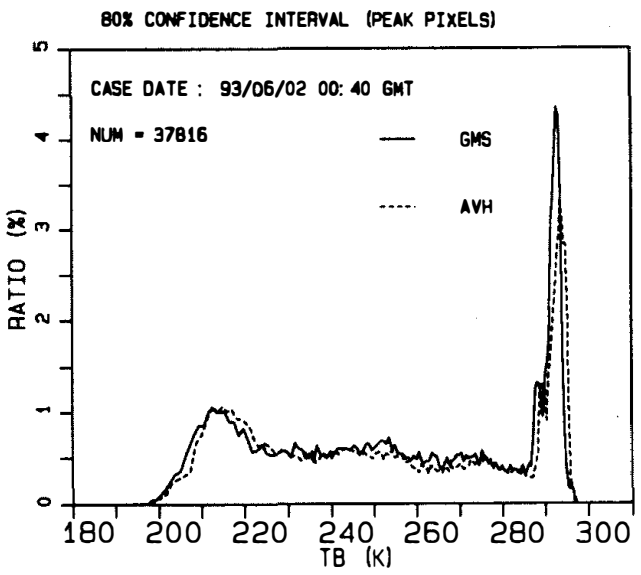

Fig. 11. Same as Fig. 9, but for an 80% confidence interval.

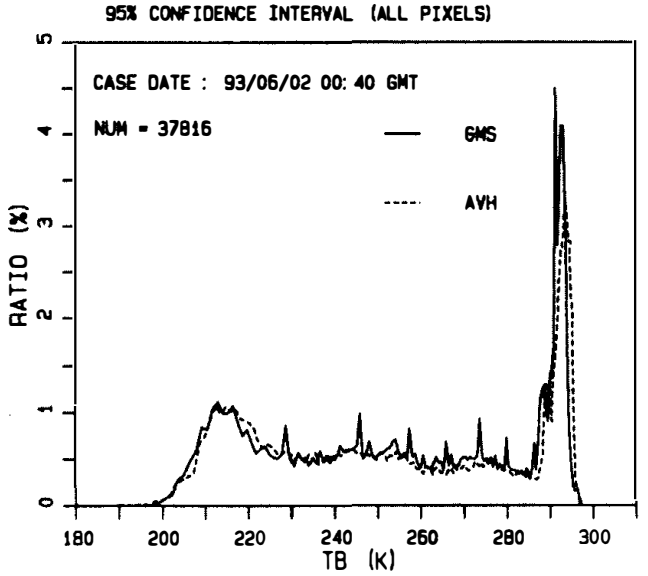

Fig. 13. Same as Fig. 9, but all pixels outside the 95% confidence interval band are corrected,

outside the confidence interval can be corrected by the same method.

Figures 12 to 15 are the same as Figures 8 to 11 but include all pixels. The ratio of the corrected pixel number to the total number is 12.2% (505:4,141) in Figure 8, while the ratio is 12.3% (4,665:37,816) in Figure 12 using a 95% confidence interval. Although both ratios are almost the same, the correction result for only peak count pixels is a little better than for all pixels when comparing Figures 9 and Figure 13. Due to the narrowness of the confidence

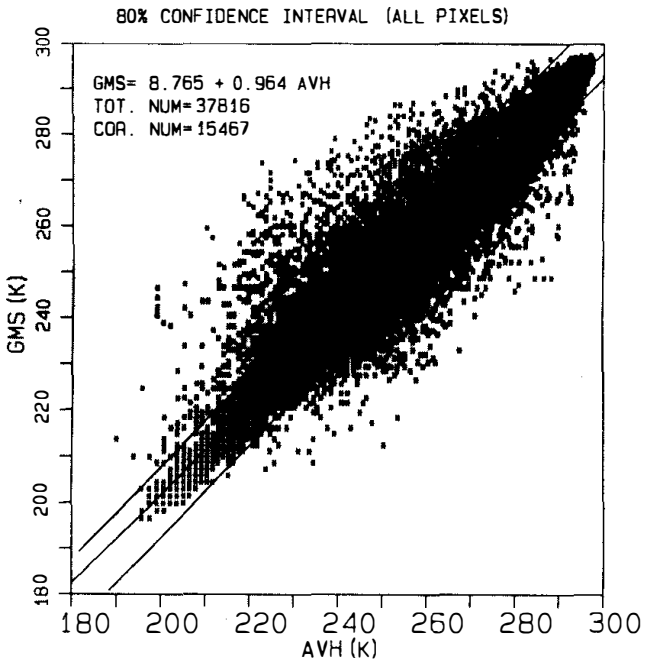

Fig. 14. Same as Fig. 12, but for an 80% confidence interval.

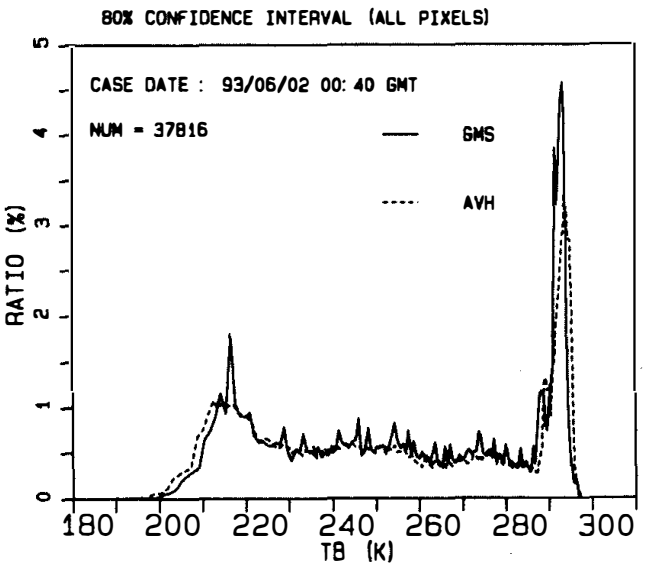

Fig. 15. Same as Fig. 13, but for an 80% confidence interval.

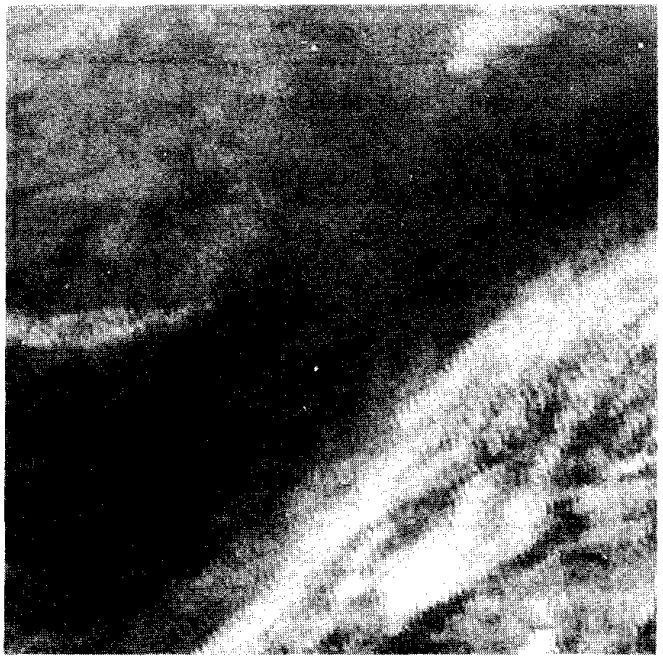

Fig. 16. Same as Fig.3, but the GMS-4 IR image is observed at 00:33 GMT May 28, 1993 (Case B).

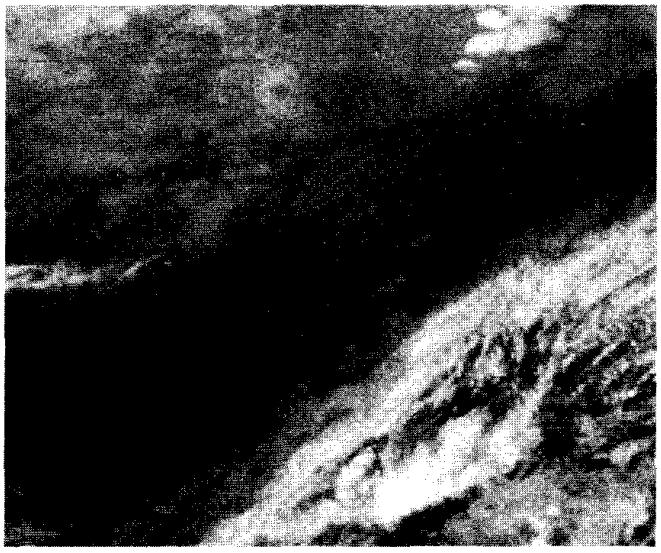

Fig. 17. Same as Fig.4, but the AVHRR IR image is observed at 00:04 GMT May 28, 1993 (Case B).

interval band, the ratio of the corrected number increases to 41.7% in Figure 10 and to 40.9% in Figure 14 using the 80% confidence interval. Comparing Figures 11 and Figure 15, it is found that the correction results for all pixels are clearly worse than those for only peak pixels. The situation is the same at the 70% confidence interval. Correction results are all worse when correcting an all-pixel case than when merely correcting the peak-count-pixel case when results are compared with the AVHRR data. The reason for this is that the Tb of the pixels

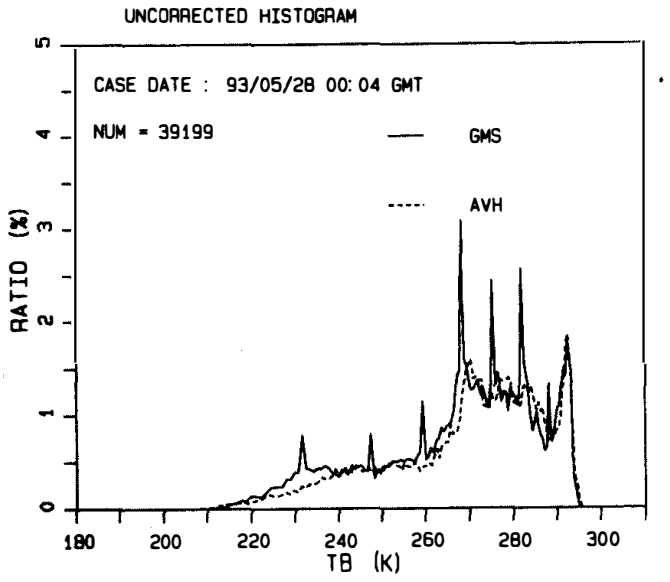

Fig. 18. Same as Fig. 6, but the satellite images are observed at 00:33 GMT May 28, 1993 (Case B).

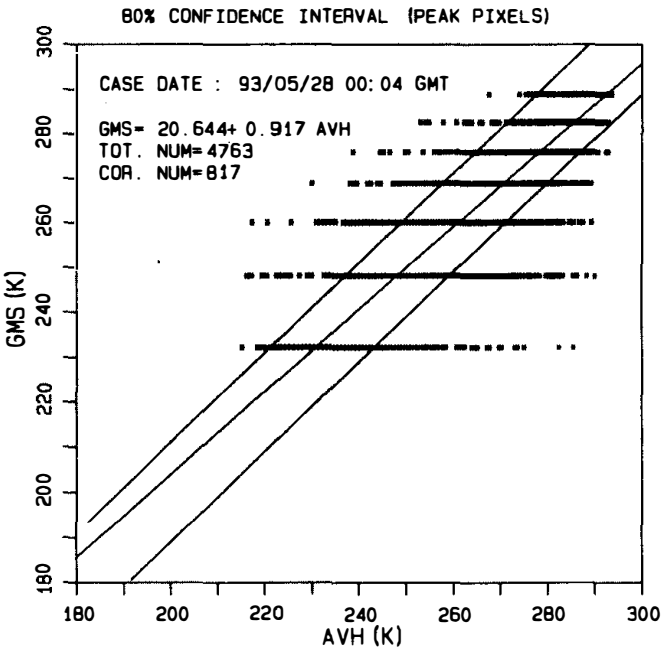

Fig. 19. Same as Fig. 10, but for Case B.

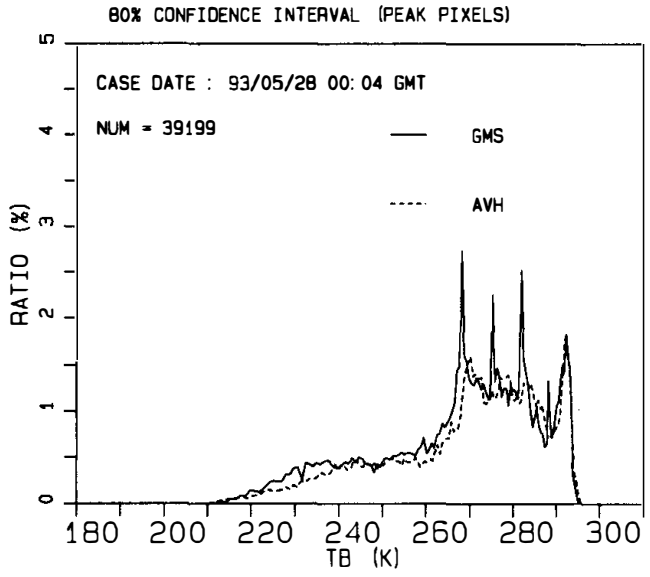

Fig. 20. Same as Fig. 13, but for Case B.

outside the confidence interval band is restructured to their corresponding AVHRR Tb. If pixels, including original normal data, are re-sampled, the Tb values become overestimated, and small peak counts occur, but owing to insufficient information, it is difficult to assign a real Tb value to the unreasonable GMS-4 pixels. It is, therefore, better to correct the peak count pixels outside the 80% confidence interval instead of correcting each and every pixel.

Figures 16 and 17 are the restructured GMS-4 IR image and AVHRR IR image at 00

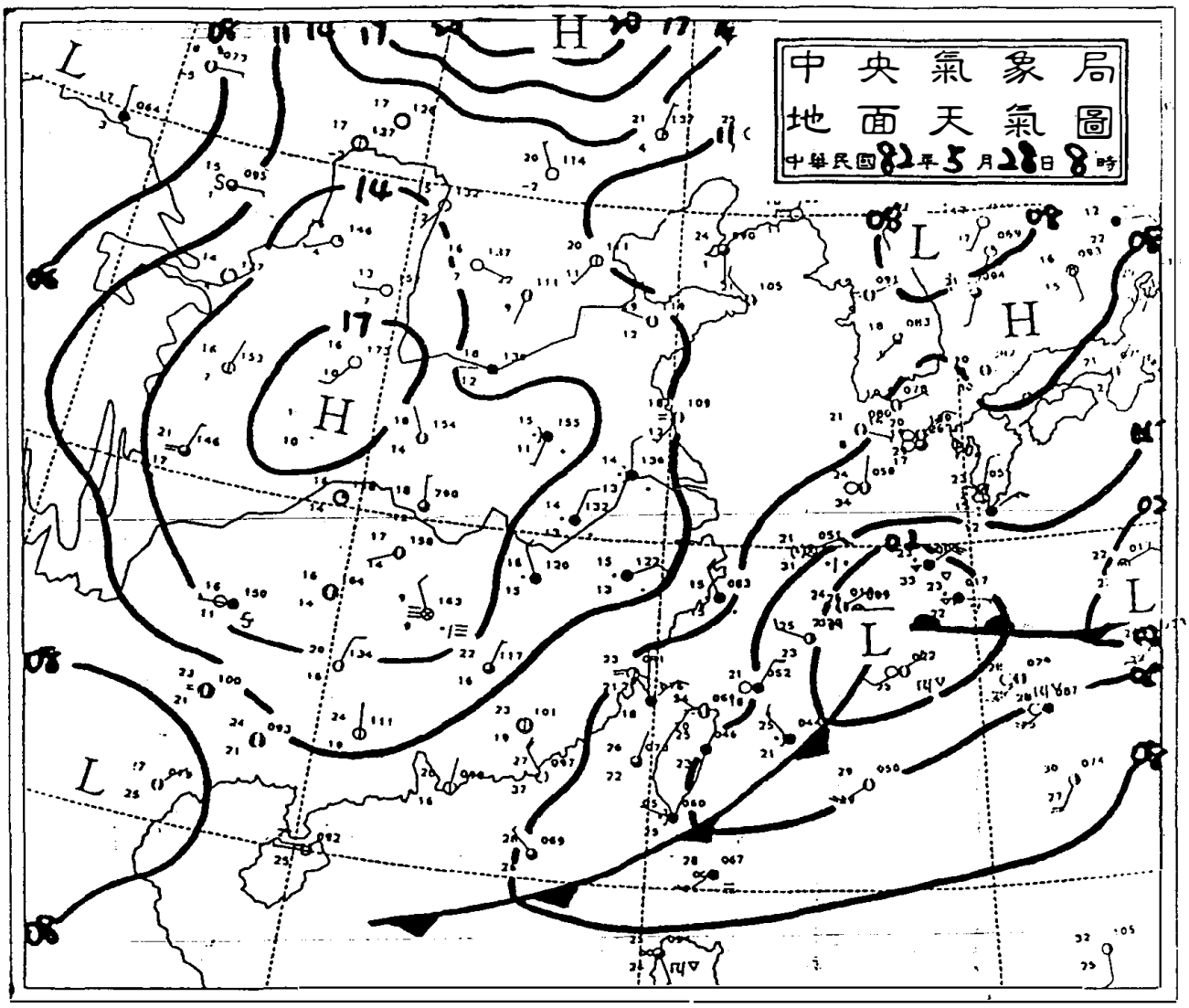

Fig. 21. Surface weather map of 00:00 GMT May 28, 1993.

GMT May 28, 1993 (Case B), respectively. From visual inspection, these two images seem to match each other very well, and by calculating the root mean square error (RMSE) for Tb between the two images, it is found that if the AVHRR image is lifted 2 lines upward and 3 pixels leftward, the RMSE is its minimum for this case. Figure 18 shows the Tb histograms. There exist 7 peak counts that differ from those in Figure 6. Figure 19 is the scatter diagram of the two IR Tbs for this case. Figure 20 illustrates the results after correction using the 80% confidence interval. It is evident that 3 peak counts with Tb less than 260K disappear, whereas 4 other peak counts with Tbs larger than 260 K still exist.

Comparing Figures 10 and 19, three differences can be found between them. First, the slope and intercept of the regression lines are different. Second, the 80% confidence interval band for Case B is wider than that for Case A, and finally, the corrected pixel number in Case A is 1727, which is much more than that in Case B (817 pixels). This may imply that the fitted regression line is not suitable for this correction technique. In general, under stable weather conditions the value of Tb for a particular location does not change too much over a short time,

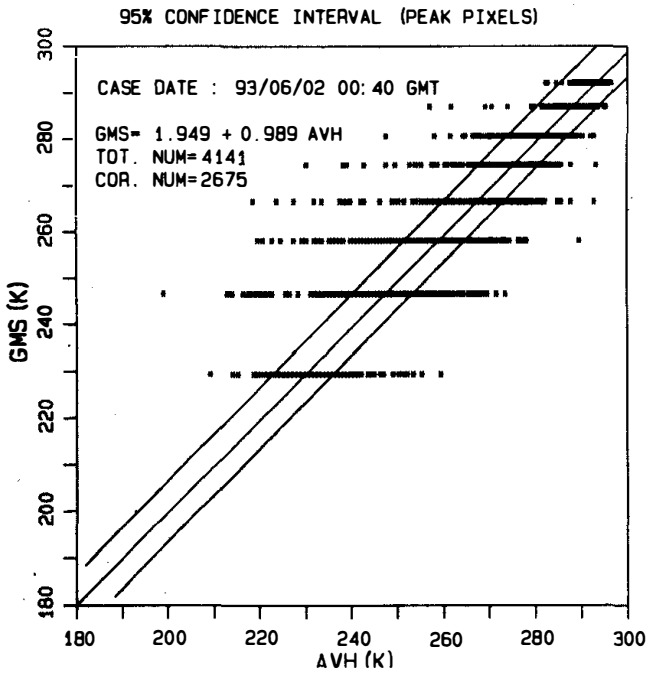

Fig. 22. Same as Fig. 8, but just those samples whose Tb difference is less than 10 K are used.

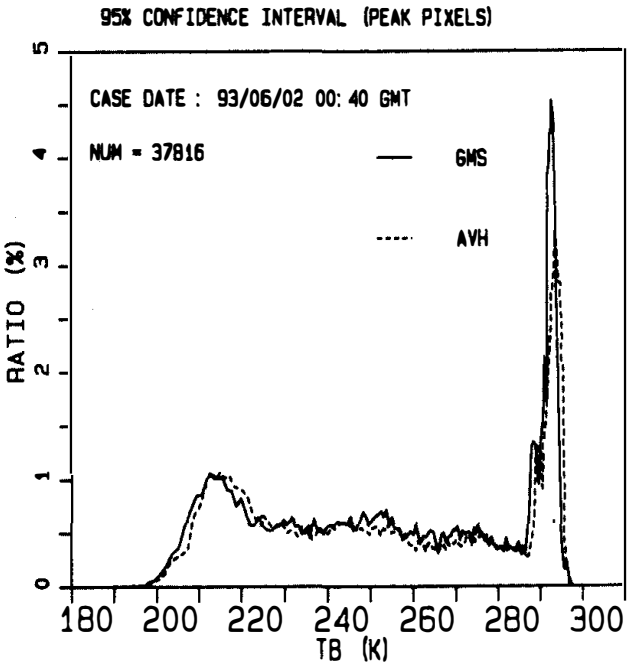

Fig. 23. Same as Fig. 9, but just those samples whose Tb difference is less than 10 K are used.

but it may well indeed vary quite a lot under unstable weather conditions, especially on the cloud edge. Thus, if the weather conditions are unstable, the difference in measurements between two satellites whose observation times are little different are large for some pixels. During the observation time period of Cases A and B, there existed a Mei-Yi front extending from the south of Japan, through the south of Taiwan and on to the south China Sea (see Figure 21).

In order to exclude unsuitable samples caused by different satellite observation times, only the samples whose Tb difference between the GMS-4 and AVHRR IR data is less than 10 K are used to fit the regression line. Figure 22 depicts the scatter diagram of peak count pixels and the new regression line and its 95% confidence interval band for Case A. It is noted that the confidence interval band narrows when compared to that in Figure 8. From the slope and intercept of the regression line in Figure 22, a better correspondence between GMS-4 and AVHRR IR radiance is obvious. Figure 23 shows the Tb histogram of AVHRR and resampled GMS-4 IR data by using a 95% confidence interval. There is no doubt that the 95% confidence interval is enough to correct the GMS-4 systematic error if selected samples are used to fit the regression line. The same situation is found in Case B. In fact, in a comparison of Figures 22 and Figure 10 it can even be seen that the 95% confidence interval band is narrower than its 80% counter part.

5. CONCLUSIONS

From the results and discussion in the previous section, it is readily understood that the

systematic errors in a GMS-4 IR image can be corrected by using an AVHRR IR image as a correction reference. That is, a linear regression relationship between GMS-4 IR Tb and AVHRR Channel 4 Tb is established, and this relationship is then applied to estimate the Tb of a GMS-4 IR channel if the measurement of the pixel is diagnosed as systematic error data.

When the effects caused by different observation times for two satellites are taken into consideration, a constraint on the sample selection is required so as to obtain a more stable and accurate regression relationship between GMS-4 IR Tb and AVHRR Channel 4 Tb. Results show that if the GMS-4 peak count data located outside the 95 % confidence interval band of the regression line are corrected using the regression relationship, the peak counts should almost disappear, and the data histogram should correspond to that of the AVHRR data.

Notwithstanding the fact that using an AVHRR IR image can successfully correct the systematic errors of a GMS-4 IR image, the applicable area is limited to a region that the NOAA series of satellites passes by. If all the NOAA satellite receiving stations within the coverage of a GMS-4 image can be utilized to correct for station coverage, the corrected area may be greatly enlarged using 4 NOAA satellite AVHRR IR images (including NOAA-10,11,12 and 14), but the problem of different observation times making the correction inaccurate still exists. In the future, a more profound examination and full verification of abnormal GMS-4 IR data correction are needed to identify the reliability of this correction technique. Finding a method that can use only a single AVHRR IR image to correct a full GMS-4 IR image or other correction technique suitable for all GMS-4 observations, whether there are NOAA AVHRR data or not, would be worthy of study. Although the GMS-5 was launched on June 14, 1995 and its IR data are not found to have any such systematic errors, it is still valuable for those who use GMS-4 IR images to apply this correction technique.

Acknowledgments The authors are grateful to the Center for Space and Remote Sensing for providing satellite data. The anonymous reviewers have greatly improved the manuscript.

REFERENCES

Lauritson, L., G. J. Nelson and F. W. Porto, 1979: Data extraction and calibration of TIROS-N/NOAA radiometers. NOAA Tech. Memo. NESS, 107, National Environmental Satellite Service, Washington, D. C., U.S.A., 58pp.

Lin Ho and P. H. Lin, 1993: The Systematic errors of geosynchronous satellite data . *TAO*, **4**, No., 331-337.

Murayama N., 1989: The GMS user's guide. Meteorological Satellite Center, Tokyo, Japan, 222pp.

Rossow W. B., F. Mosher, E. Kinsella, A. Arking, M. Desbois, E. Harrison, P. Minnis, E. Ruprecht, G. Seze, C. Simmer and E. Smith, 1985: ISCCP cloud algorithm intercomparison. *J. Appl. Meteor.*, **24**, 877-903.

Sasaki H., Y. Shirakawa and Y. Takeuchi, 1990: Estimation of sea surface temperatures using MOS-1, GMS and NOAA, Proc. MOS-1 Data Evaluation, 109-118.

Schiffer, R. A. and W. B. Rossow, 1983: The international satellite cloud climatology project (ISCCP): The first project of the World Climate Research Programme. *Bull. Am. Meteor. Soc.*, **64**, 779-784.

Tseng, C.-Y., 1988: Remote sounding of the atmosphere from satellites. Bo Hai Town, Taipei, Taiwan, 630pp. (in Chinese)