4-29-2013

Refining Environmental Satellite Data Using a Statistical Approach

Md Zahidur Rahman
CUNY La Guardia Community College

Leonid Roytman
CUNY City College

Abdelhamid Kadik
CUNY La Guardia Community College

Refining Environmental Satellite Data Using a Statistical Approach

Md Z Rahman[1], Leonid Roytman[2] , and Abdel Hamid Kadik[3]
[1,3]Department of Mathematics, Engineering, and Computer Science, LaGuardia Community College of the City University of New York, 31-10 Thomson Avenue, Long Island City, NY-11101
[2] Department of Electrical Engineering, The City College of the City University of New York, 138 Street & convent Avenue, New York, NY-10031

ABSTRACT

The proposed approach in this article applies an efficient and novel statistical technique to accurately describe radiometric data measured by Advanced Very High Resolution Radiometers (AVHRR) onboard the National Oceanic and Atmospheric Administration's (NOAA) Polar Orbiting Environmental Satellites (POES). The corrected data set will then be applied to improve the strength of NOAA Global Vegetation Index (GVI) data set for the 1982-2003 period produced from AVHRR. The GVI is used extensively for studying and monitoring land surface, atmosphere and recently for analyzing climate and environmental changes. The POES AVHRR data, though useful, cannot be directly used in climate change studies because of the orbital drift in the NOAA satellites over the lifetime of the satellites. This orbital drift causes inaccuracies in AVHRR data sets for some satellites. The main goal is achieved by implementing a statistical technique that uses an Empirical Distribution Function (EDF) to produce error free long-term time-series for GVI data sets. This technique permits the representation of any global ecosystem from desert to tropical forest and to correct deviations in satellite data that are due to orbital drifts and AVHRR sensor degradations. The primary focus of this research is to generate error free satellite data by applying the EDF technique for climatological research.

Keywords: statistical, satellites, vegetation, error, data, degradation

1. INTRODUCTION

This paper investigates the Brightness Temperature (BT) data stability calculated from the AVHRR observations for the period 1982-2003 [3]. AVHRR weekly data for the five NOAA afternoon satellites NOAA-7, NOAA-9, NOAA-11, NOAA-14, and NOAA-16 are used for the China dataset. It was found that data for the years 1988, 1992, 1993, 1994, and 2000 are not stable enough compared to other years because of satellite orbit drift, AVHRR sensor degradation, and electronic and mechanical satellite systems deterioration/failure etc. For our research the data for NOAA-7(1982, 1983), NOAA-9 (1985, 1986), NOAA-11(1989, 1990), NOAA-14(1996, 1997), and NOAA-16 (2001, 2002) are assumed to be standard due to the fact, that equator crossing time of satellite between 1330 and 1500, which maximized the value of coefficients[2,8]. In other years, quality of satellite observations significantly deviates from the standard. In this research, we apply Empirical Distribution Function (EDF) which is most effective statistical technique to correct this deficiency of data for the affected years.

For almost two decades, the Advanced Very High Resolution Radiometer (AVHRR) on NOAA polar-orbiting satellites have observed radiances, which were collected, sampled, and stored for the entire world. These data were intensively used by the global community for studying and monitoring land surface, atmosphere, and recently for analyzing climate and environmental changes [1] [2]. AVHRR data, though informative, can not be directly used in climate change studies because of the orbit drift in the NOAA satellites (particularly, NOAA-9, -11, and -14) over these satellites life time [5] [6]. Price 1991 attributed this drift to the selection of a satellite orbit designed to avoid direct sunshine on the instruments. These goals were seriously challenged by performance of NOAA satellites, which orbit degraded over time, and by deterioration of AVHRR sensor. As the result long-term time series of the Brightness Temperature (BT) calculated from the thermal infrared channel (CH4) experienced some trend. Brightness temperature is defined as thetemperature a blackbody would be in order to produce the radiance perceived by the sensor. Brightness Temperature is a descriptive measure of radiation in terms of the temperature of a hypothetical blackbody emitting an identical amount of radiation at the same wavelength. The brightness temperature is obtained by applying the inverse of thePlanck function to the measured radiation. Depending on the nature of the source of radiation and any subsequent absorption, the brightness temperature may be independent of, or highly dependent on, the wavelength of the radiation. The Method of Matching empirical distribution functions

ern,
Proc. of SPIE Vol. 8723, 87231F · © 2013 SPIE · CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2015592

were applied for improving stability of BT data for AVHRR observations. The purpose of this paper is to test stability of BT, corrected CH4 value for all over china region. We can apply this methodology globally to generate error-free vegetation index data sets in BT to develop the climatology.

1.1 Land Targets

The land targets of 28^0 N to 43^0 N in latitude and 75^0 E to 123^0 E in longitudes were selected from China (Figure 1).

Figure 1. Geographical Map of China with the area study (boardered).

We attempted to select relatively small uniform areas using common knowledge of geography, climate, ecosystem, and human activities. The main cover types are desert, forest, and grassland. The information collected by the ground truth survey is useful for the classification of NOAA AVHRR weekly-composted BT data.

2. DATA AND PROCESSING

The Ch4 were obtained from the NOAA/NESDIS Global Vegetation Index (GVI) data set [5]. The GVI is produced by sampling and mapping the 4 km AVHRR- derived radiance to a 16 km map daily. The daily maps are composited over a seven-day period by saving the channel and BT values for that day that has the largest BT value during the seven days for each map cell. We used weekly Ch4 data from 1982-2003, which were collected from five satellites NOAA –7, -9, -11, -14 and -16. Multi-temporal NOAA –7, -9, -11, -14 and -16 dataset from the NOAA Satellite Archive (SA) were used for this study. The calibration of AVHRR CH4 (10.3-11.3 µm) utilizes a set of time varying calibration coefficients to correct for sensor degradation, as there is no on-board calibration for these channels. The seven-day composited CH4 values were corrected for using the matching of empirical distribution functions and BT was calculated. High frequency (from week to week) fluctuations in the corrected weekly BT time series were smoothed with a statistical filter.
NOAA-NASA AVHRR pathfinder program, which has been entrusted with the task of the reprocessing and rehabilitation of the AVHRR records for the period 1981-1990 (which has been termed the pathfinder period) for the purpose of the production of long- term records for climate studies, set up the AVHRR Pathfinder Calibration Working group in March 1991 to assess the degradation of channels 1 and 2 on the NOAA-7, -9, –11, and –14 spacecraft (referred to as the afternoon satellites since they cross the equator traveling north-wards in the afternoon) and to make recommendations for implementing appropriate corrections for the observed degradation .The report of this working group has now been published [10].

3. METHODOLOGY

There is no physical method used for stability of BT. We used statistical analysis for the correction of BT in 1988, 1992, 1993, 1994, and 2000. The matching of empirical distribution function (EDF) is a statistical technique which is applied to improve the stability of BT time series. EDF is a non-decreasing function of BT value and its maximum value is unity (Figure 2).

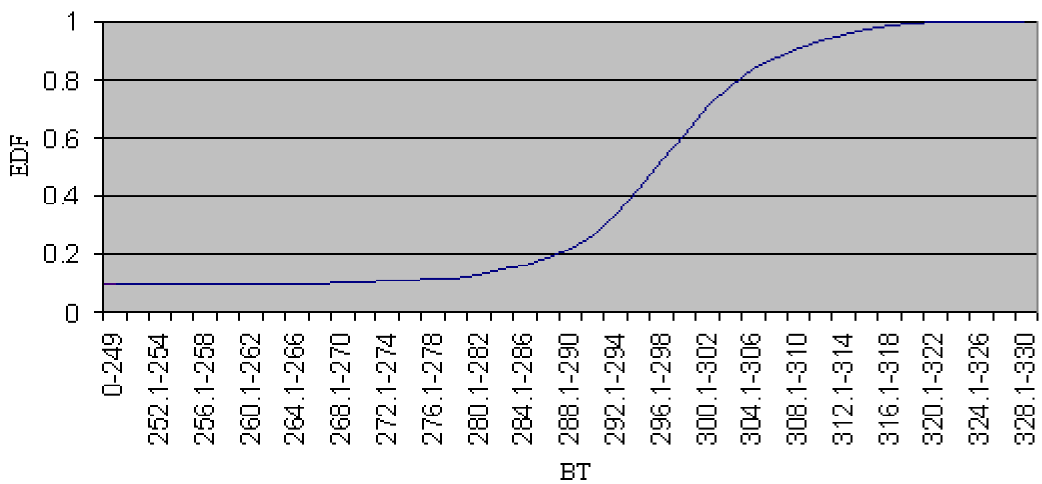

Figure 2. Empirical distribution functions of BT data

We assume that data from NOAA-7(1982, 1983), NOAA-9 (1985, 1986), NOAA-11(1989, 1990), NOAA-14(1996, 1997), and NOAA-16 (2001, 2002) are the best suited for analysis because of their equator crossing time between 1330 and 1500. As this period of the day is the best time (hottest temperature) for satellite observation, we consider data for these years as standard .Then the BT values of other years (1988, 1992, 1993, 1994, and 2000) are adjusted so that their distributions are the same as that of the standard.

To find the normalized value corresponding to the unnormalized value, the following is the procedure:

i) For the BT value in unnormalized EDF year, find the EDF value from the EDF of unnormalized year

ii) Then find the point on the standard year's EDF with the same EDF value.

iii) Finally, use the EDF of the standard year to find the normalized BT value. Since the data are actually discrete, we will need to interpolate within the EDF of the standard year to find the new value of BT.

4. RESULTS AND DISCUSSION

We have produce BT time series of five NOAA satellites are illustrated in Figure 3 which shows BT data for 1988 (NOAA-9), 1992, 1993, 1994 (NOAA-11), and 1995, (week #1-8), 2000 (NOAA-14) are not enough stable compare to other years because of satellite orbital drift, Mt Pinatubo, and sensor degradation. Therefore, we need to correct the BT data for affected years.

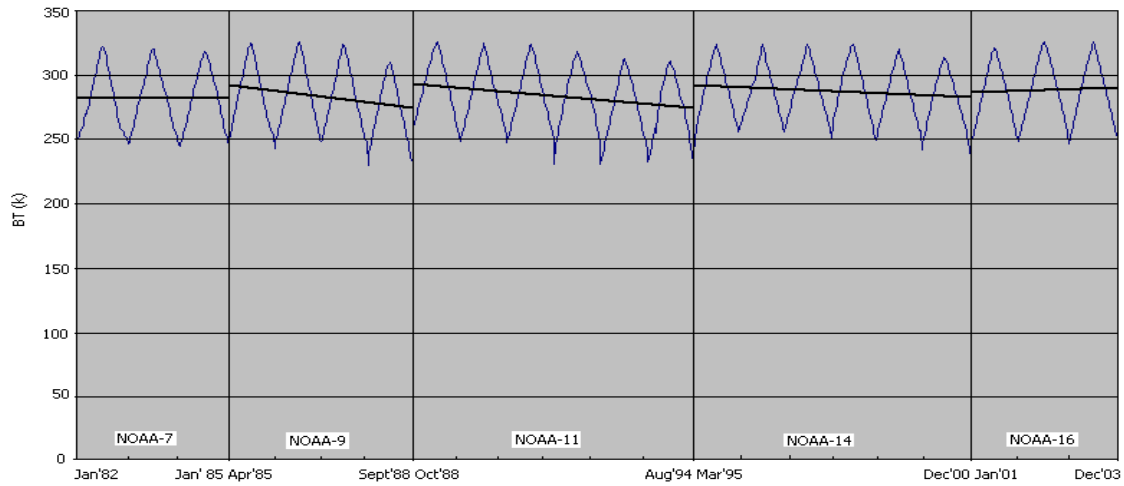

Figure 3 BT time series

For this correction, we apply the method of matching empirical distribution functions for affected years. We normalize data by the method of matching of EDF's compared with standard according section 3. Figure 4 depicts how the procedure is applied in actual practice to generate normalization BT values. The figure depicts idealized EDF's for the standard year and 1988 year. In the figure the EDF's are continuous, but in practice they are discrete, being specified only at integer values of BT.

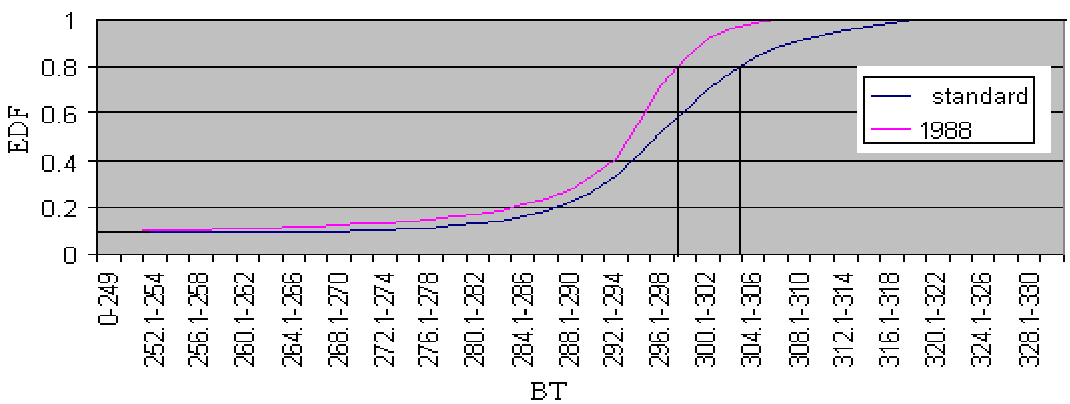

Figure 4 Illustration of procedure to generate normalization BT data

For example, for the BT value 300 in year 1988 find the EDF value from the EDF of year 1988. In the illustration it is EDF_{88} is 0.8. Then find the point on the standard year's EDF with the same EDF value. Therefore, EDF value can also be expressed as $EDF_{standard}$ is 0.8. Finally, use the EDF of the standard year to find the normalized count value 306. Since the data are actually discrete, we will need to interpolate within the EDF of the standard year to find the value of 306, which must then be rounded to the nearest integer. We do interpolate with in the EDF of the standard year to find the normalized count value. After interpolated, we found the normalized BT value 300 in year 1988 is 306.

Therefore, 1988 new 300 = $300_{1988} + (306_{standard} - 300_{1988}) = 306$

Using this technique, EDF's produce to normalize or correct data for the years 1988, 1992, 1993, 1994, and 2000 compared with standard which are illustrated in following Figures.

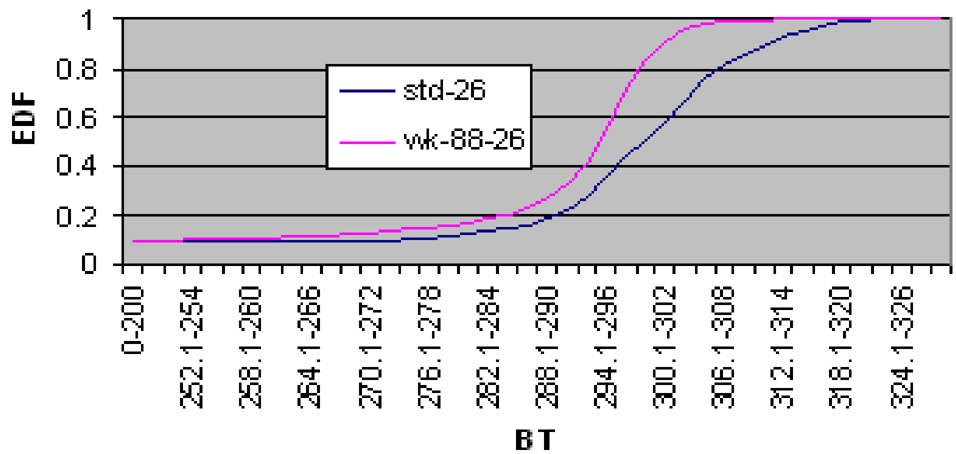

Figure 5 (a). Empirical distribution functions for unnormalized BT data of 1988 (26th week) compared with standard

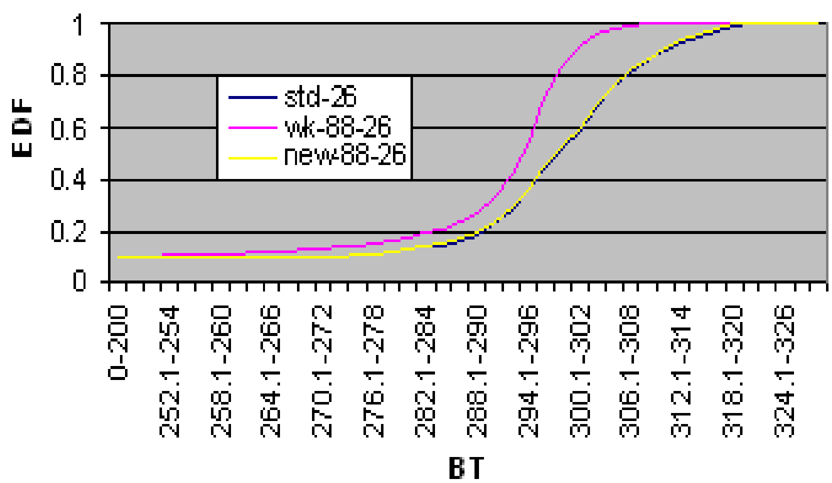

Figure 5 (b) Empirical distribution functions for normalized BT data of 1988 (26th week) compared with standard

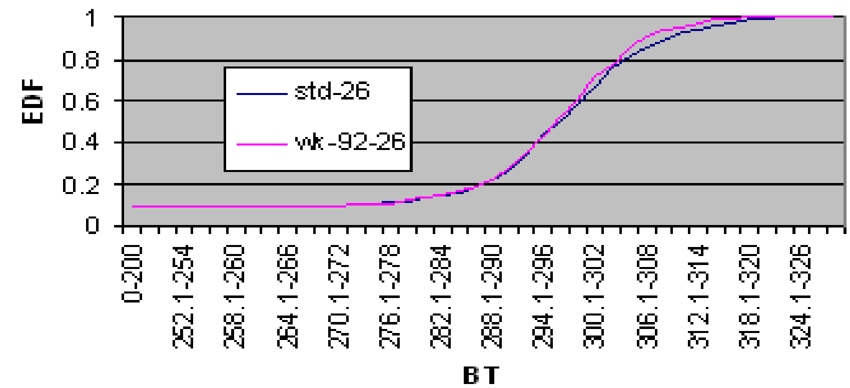

Figure 6(a). Empirical distribution functions for unnormalized BT data of 1992 (26th week) compared with standard

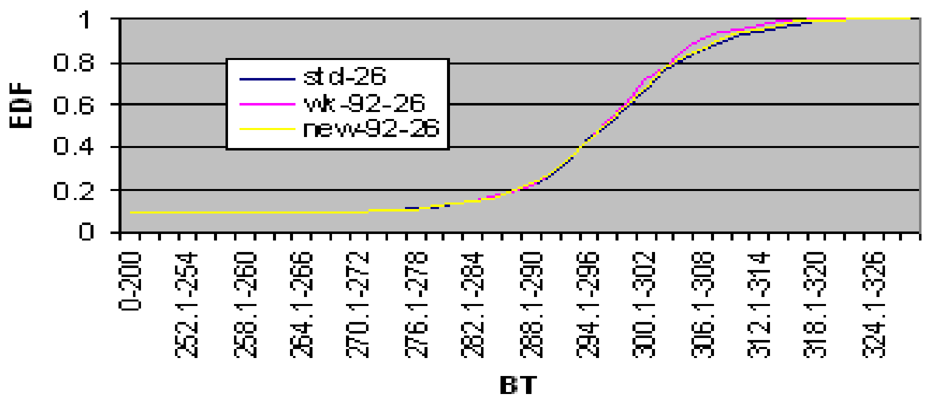

Figure 6(b).Empirical distribution functions for normalized BT data of 1992 (26th week) compared with standard

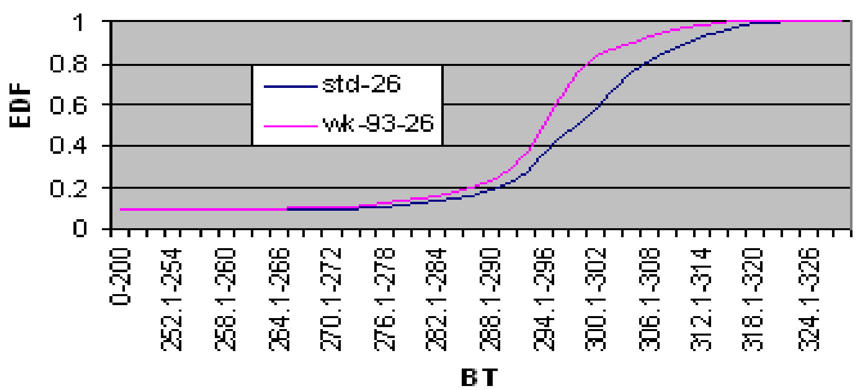

Figure 7(a). Empirical distribution functions for unnormalized BT data of 1993 (26th week) compared with standard

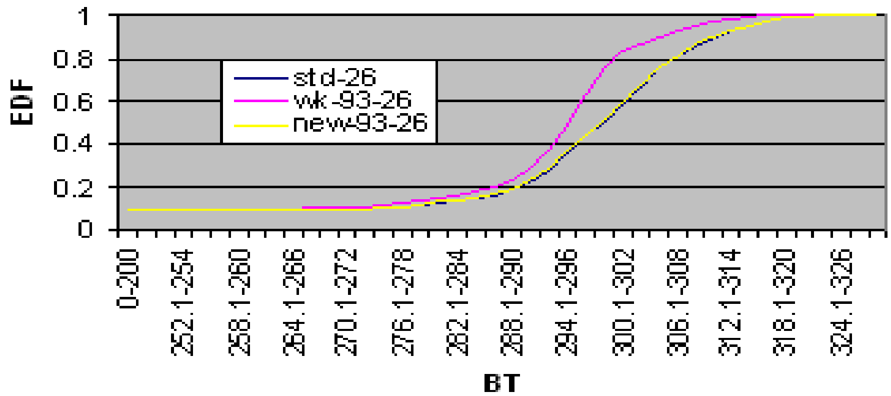

Figure 7 (b). Empirical distribution functions for normalized BT data of 1993 (26th week) compared with standard

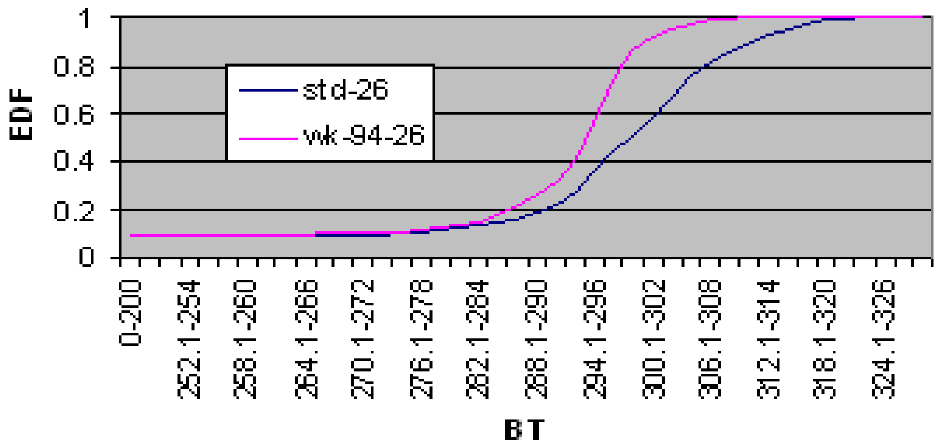

Figure 8(a). Empirical distribution functions for unnormalized BT data of 1994 (26th week) compared with standard

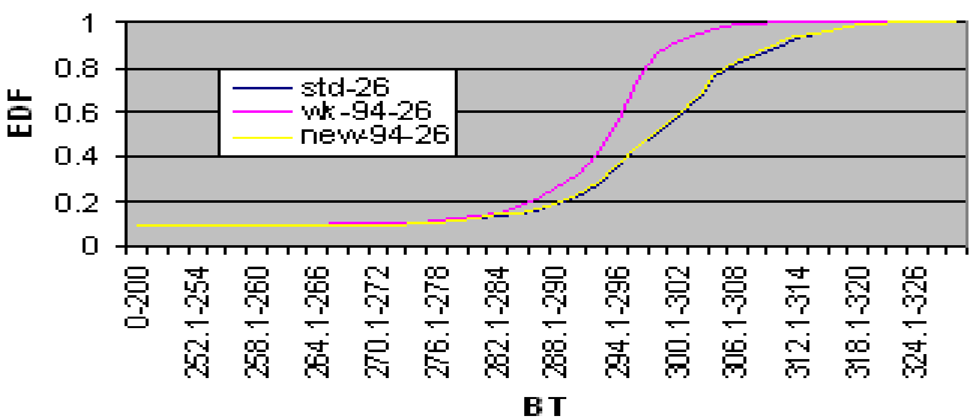

Figure 8(b). Empirical distribution functions for normalized BT data of 1994 (26th week) compared with standard

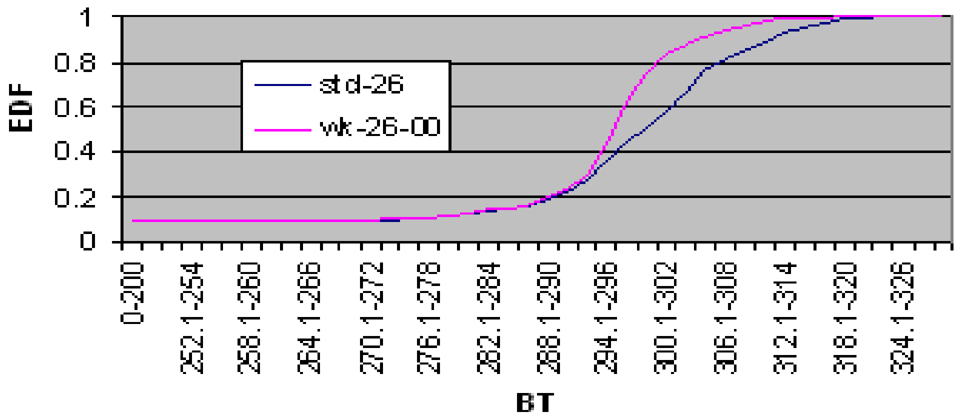

Figure 9 (a). Empirical distribution functions for unnormalized BT data of 2000 (26th week) compared with standard

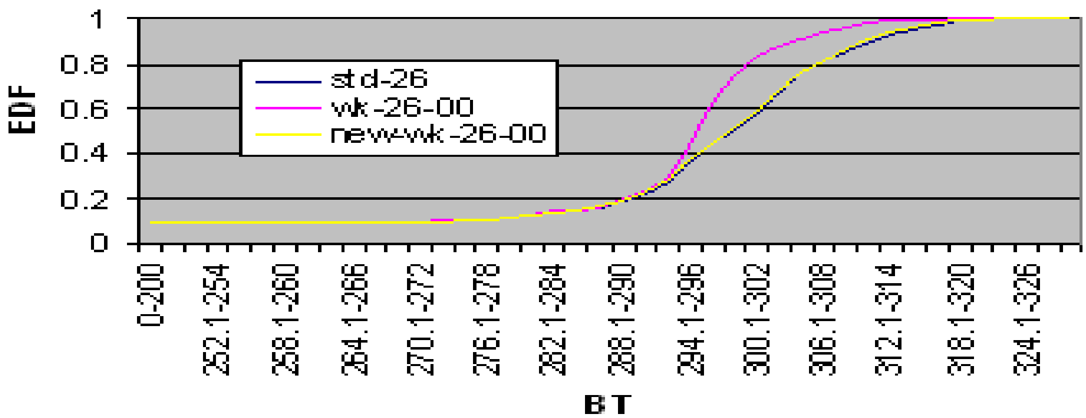

Figure 9 (b) Empirical distribution functions for normalized BT data of 2000 (26th week) compared with standard

Figure 5- 9 (part a) shows the EDF's of unnormalize data for the years 1988, 1992, 1993, 1994, and 2000 as well as the EDF of the standard. The abscissa labeled "BT" value and ordinate labeled "EDF" (0 to 1). Figure 5-9 (part b) shows the EDF's of the normalize data for each of those years and indicates that the normalization was successful in making the EDF's of the two years nearly identical. This implies that the relationships between the EDF's remained essentially the same between two years. Those relationships, in fact, depend only on the relative function between two years. As long as the relative functions remain in the same, the normalization data remain effective. Using normalized value, we produce new BT time series as shown in Figure 10 which shows improve the BT data (pink line) of the year of 1988, 1992, 1993, 1994, and 2000.

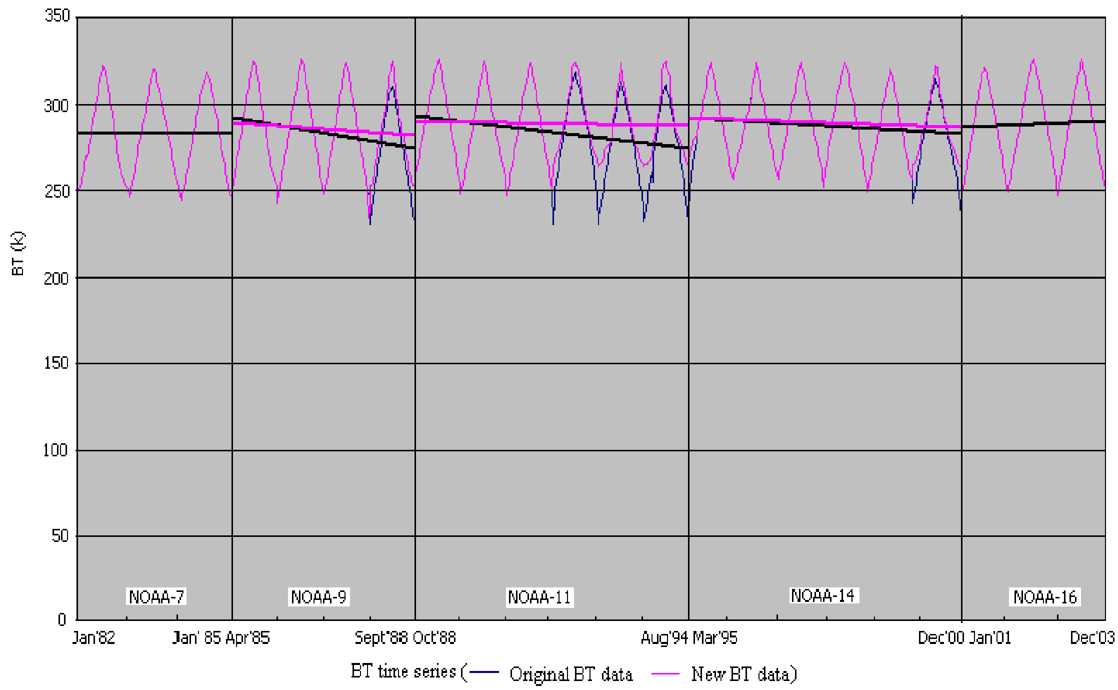

Figure 10. Corrected BT time series

BT trends for china and jumps between the satellites are illustrated in Figure 3 and 10 and the errors are estimated in Table 4.1 Figure 3 shows some BT trends for each satellite except NOAA-16 and jump from one satellite to the next one. Considering old BT trend (Table 4.1), for china, NOAA-9, -11, and -14 have negative trend and NOAA-7, -16 have positive or no trend. Therefore, NOAA-7, and -16 shows clear tendency to BT increase during its three years in operation. However, important is trend rate. Analysis shows that high rate of BT change for NOAA-9,-11, and -14 by reduction of BT in 1988, 1992-1994, and 2000 due to elevated amount of stratospheric aerosol from Mt Pinatubo and considerable degradation of satellite orbit.

Regarding BT jump from one satellite to the next in Table 4.1 (B), general tendency is a reduction of BT between beginning of NOAA-9 and the end of NOAA-7, between beginning of NOAA-16 and the end of NOAA-14. An increase in BT is observed only during satellite change from NOAA-9 to NOAA-11, NOAA-11 to NOAA-14, and NOAA-14 to NOAA-16 due to already mentioned sharp stratospheric aerosols increase and orbit drift of satellite.
After correction of BT, we also estimate errors in Table 4.1 for new BT. Figure 10 shows improve the BT trends for each satellite and jump from one satellite to the next one. But there remain other potential sources of error in BT such as an incomplete drift correction, inaccurate BT calculation, and influence from Mt. Pinatubo. The EDF method is designed to reduce errors due to orbit drift, the dominant uncertainty in temperature variation during the satellite life time [11]. However, it may be difficult to accurately and completely remove this effect and thus orbit remains as an error source, though at a reduced level. Another large uncertainty lies in BT calibration and Mt Pinatubo which includes all errors such as incomplete atmospheric corrections, surface corrections, sensor degradation and volcanic eruptions.

Table 4.1 : Estimation of Errors in (A) BT trend at the End of a Satellite Life and (B) Jumps between the Satellites (% to the beginning level)

Ecosystem		A			B			
		N-9	N-11	N-14	N-7/9	N-9/11	N-11/14	N-14/16
Whole China	Old BT	-6	-6.5	-3	3.3	7	6.5	5.1
	New BT	-2	-1	-1.7	2	3	1.7	0

5. CONCLUSIONS

Empirical distribution function improves the time related stability of NDVI for all satellites, especially NOAA-9, -11, and -14 environmental satellites. This is strong evidence that normalization by EDF matching is an effective method for improving stability of long-term BT time series. For climate and global change studies, BT time series are not stable enough. Following BT data distortion due to external forcing (stratospheric aerosols and satellite orbit degradation), BT data for 1988 (NOAA-9) and 1993, 1994 (NOAA-11), and 2000 (NOAA-14) will likely distort mean values and other statistics and should be tested comprehensively before they are used for monitoring the environment.

REFERENCES

1. Holben BN "Characteristics of maximum-value composite image from temporal AVHRR data",Int. J. Remote Sens. Vol. **7**: 1417-1434, 1986.
2 . Kogan, F.N., and X. Zhu, "Evolution of long-term errors in NDVI time series", Adv. Space Res. Vol. **28**, No 1, pp. 149-153, 2001.
3. Kogan, F.N., "Global Drought Detection and Impact Assessment from space in Drought",A Global Assessment Ed. By D.A Wilhite, vol.**1**, Hazard and Disaster Series, pp. 196-210, Routledge, London-New York, 2000
4. Kogan, F.N., "Global Draught Watch from Space",Bull, Amer. Meteor. Soc., **78**, 621-636, 1997.
5. Kidwell, K.B. (Ed), "Global Vegetation Index User's Guide",U.S Department of Commerce, Technical Report, Washington, D.C., 1997.
6. M.P. Weinreb, R.Xie, J.H. Lienesch, and D.S Crosby; "Removing Striepes in GOES Images by Matching Empirical Distribution Functions",NOAA Technical Memorandum NESDIS 26, Washington, D.C, May 1989.
7. P. Cracknell, "*The Advanced Very High Resolution Radiometer*",Taylor & Francis, page 90-97,1997.
8. Rao, C.R.N. and J.Chen, "Revised post-launch Calibration of the visible and Near-infrared Channels of the Advanced Very High Resolution Radiometer on the NOAA-14 Spacecraft",Int. J. Remote Sensing ***20***, 3485-3491, 1999.
9. Rao, C.R.N. and J.Chen, "Inter-satellite Calibration Linkages for the Visible and Near-infrared Channels of the Advanced Very High Resolution Rdiometer on the NOAA –7, -9, and –11 Spacecrafts",Int. J. Remote Sensing **16**, 1931-1942, 1995.
10. Rao, C.R.N. , "Degradation of the visible and near infrared channels of the Advanced Very High Resolution Radiometer on the NOAA –9 spacecraft", assessment and recommendations for corrections. NOAA Technical Report NESDIS 70 (Washington DC: US Department of Commerce), 1993b.\
11. Price, **J. C.,** "**T**iming of NOAA afternoon passes",***Int. J. Remote Sens.*, 12,** 193-198, 1991.