City University of New York (CUNY)
CUNY Academic Works

Publications and Research | LaGuardia Community College

5-1-2012

Cost Effective Malaria Risk Control Using Remote Sensing and Environmental Data

Md Zahidur Rahman
CUNY La Guardia Community College

Leonid Roytman
CUNY City College

Abdel Hamid Kadik
CUNY La Guardia Community College

How does access to this work benefit you? Let us know!

Follow this and additional works at: http://academicworks.cuny.edu/lg_pubs


Part of the Diseases Commons

Recommended Citation

Rahman, Md Zahidur; Roytman, Leonid; and Kadik, Abdel Hamid, "Cost Effective Malaria Risk Control Using Remote Sensing and Environmental Data" (2012). *CUNY Academic Works.*
http://academicworks.cuny.edu/lg_pubs/54

This Article is brought to you for free and open access by the LaGuardia Community College at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.

Cost Effective Malaria Risk Control Using Remote Sensing and Environmental Data

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

Malaria transmission in many part of the world specifically in Bangladesh and southern African countries is unstable and epidemic. An estimate of over a million cases is reported annually. Malaria is heterogeneous, potentially due to variations in ecological settings, socio-economic status, land cover, and agricultural practices. Malaria control only relies on treatment and supply of bed networks. Drug resistance to these diseases is widespread. Vector control is minimal. Malaria control in those countries faces many formidable challenges such as inadequate accessibility to effective treatment, lack of trained manpower, inaccessibility of endemic areas, poverty, lack of education, poor health infrastructure and low health budgets. Health facilities for malaria management are limited, surveillance is inadequate, and vector control is insufficient. Control can only be successful if the right methods are used at the right time in the right place. This paper aims to improve malaria control by developing malaria risk maps and risk models using satellite remote sensing data by identifying, assessing, and mapping determinants of malaria associated with environmental, socio-economic, malaria control, and agricultural factors.

Keywords: malaria, epidemic, vector, control, health

1. INTRODUCTION

Malaria is a major public health problem in many developing countries, including Bangladesh and southern African countries. The World Health Organisation (WHO) has estimated that over one million cases of malaria are reported each year in Bangladesh [1]. This paper studies the risk factors for malaria in Bangladesh. The country has an area of 147,570 sq km and a population of 143.8 million. Out of total 64 administrative districts, 13 are in the high malaria endemic area located along the border areas with India and Myanmar. About 98% of the total malaria morbidity and mortality occuring in Bangladesh each year is reported from these districts and more than 75% of the cases are *P. falciparum* malaria . In 2004 the official annual number of laboratory confirmed malaria cases was almost 60,000 with more than 500 deaths [1]. However, the number of unreported and clinically diagnosed malaria cases is estimated to range between 400,000 and 1 million [1].

Seven of the 34 malaria mosquito species found in Bangladesh have been incriminated in malaria transmission [2] Four species are important malaria vectors: Anopheles dirus, An. philippinensis, An. sundaicus and An. minimus. The other three (An. aconitus, An. annularis,and An. vagus) have been incriminated during malaria outbreaks. An. dirus is the primary vector. It is associated with the hilly forested and foothill areas in the eastern part of the country and largely rests outdoors. *An. minimus* is also an important vector in these areas. The peak hours of biting are 22:00-24:00 for *An. dirus* and 20:00-22:00 for An. minimus. An. philippinensis is a malaria vector of the central flood plain and deltaic regions of the country and *An. sundaicus* is found in the coastal areas .

The malaria disease burden and malaria risk factors are poorly studied in Bangladesh. Malaria was almost under control in 1971-1972. Malaria incidence has been increasing since 1973 with periodic epidemic outbreaks (Figure 1). Like other vector borne diseases it is of paramount importance to investigate the bionomics of the malaria vectors and their relation with ecological factors for any given situation. Malaria control is not possible if risk factors are not identified and controlled.

Sensing Technologies for Global Health, Military Medicine, Disaster Response, and Environmental Monitoring II; and Biometric Technology for Human Identification IX, edited by Sárka O. Southern, et al., Proc. of SPIE Vol. 8371, 83711I
© 2012 SPIE · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.918814

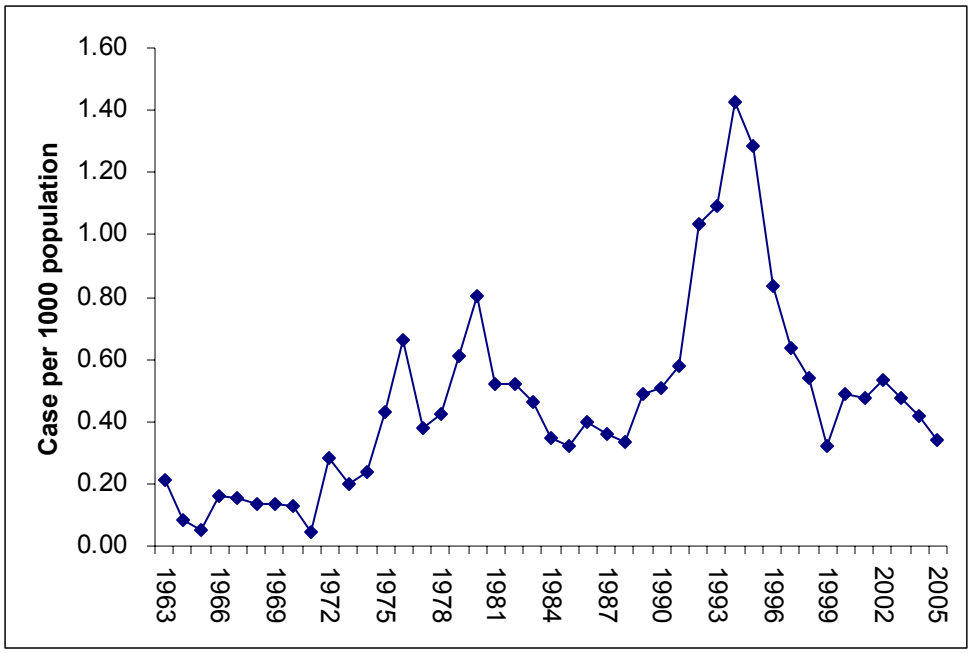

Fig. 1. Incidence of malaria in Bangladesh from 1963 to 2005. Data from Directorate General Health Services, Bangladesh.

Malaria control in Bangladesh faces many formidable challenges, such as inadequate accessibility to effective treatment, lack of trained manpower, inaccessability of endemic areas, the population at risk for malaria is poor, lacks education, and often belong to ethnic minorities. Health facilities for the management of severe malaria are limited, surveillance is inadequate, and vector control is insufficient.

To improve malaria control interventions it is important to know where resources will be spent most efficiently. By applying spatial tools such as Remote Sensing (RS), Geographical Information Systems (GIS) and spatial statistics this task will be made easier. The use of RS and GIS for investigating spatial relations between malaria risk and risk factors is rare in Bangladesh. However, a recent study showed correlations between malaria cases and satellite-based vegetation health proxies (moisture and thermal conditions) during the rainy season in Bangladesh [16]. Furthermore, investigations on malaria mosquitoes in the flood plains of the country found that vector densities were more affected by large-scale environmental changes and population growth than to flood-control activities [3]. Malaria investigations in the eastern forests of Bangladesh during the late 1970s showed a seasonal and holoendemic malaria transmission pattern and determined several aspects of the biology of the primary vector *An. dirus* [14],[15].

However, no study has yet been carried out to investigate the combined effects of socioeconomic, entomological, and environmental risk factors for malaria in Bangladesh. Some studies of this type have been done in other epidemiological settings in Asia, Africa, and the Americas. Early studies successfully used RS and GIS technologies to detect mosquito breeding habitats [7], predict densities of *Anopheles* vectors [26], [17], and classify the risk of malaria transmission [4]. A malaria risk analysis in Sri Lanka using GIS found increased risk in areas characterized by higher than average rainfall, greater forest coverage, slash and burn cultivation, presence of abandoned irrigation reservoirs, and poor socio-economic status [10]. Surface slope and wetness indices derived from spatial modeling (GIS, RS and Digital Elevation Models, DEM) were positively associated with the abundance of several major malaria vectors in Thailand [23]. Such associations permit real-time monitoring and prediction of vector distributions, facilitating improved control before mosquitoes emerge and transmit disease. In coastal Kenya, [11] used Radar remote sensing to identify environmental factors associated with malaria risk. A risk map was generated to show populated areas within a 2 km buffer zone around wetland areas associated with high larval breeding. By using 1 m spatial resolution Ikonos images and computer modelling based on topographic land cover features it was possible to identify 40% of anopheline larval habitats in the highlands of Kenya [13]. Panchromatic aerial photos identified only 10% and Landsat TM 7 images was not able to identify any breeding habitats. In Gambia [24], and Kenya [7], RS-derived NDVI (Normalized Difference Vegetation Index) lagged by one month showed significant correlations with malaria cases. NDVI is often used as a proxy for monitoring vegatation development as a response

to regional rainfall distribution and can be used for malaria forecasting.
This paper investigates Normalized Difference Vegetation Index (NDVI) stability in the NOAA/NESDIS Global Vegetation Index (GVI) data for the period 1982-2003 [9]. AVHRR weekly data for the five NOAA afternoon satellites NOAA-7, -9, -11, -14, and NOAA-16 for dataset is studied. It was found that data for the years 1988, 1992, 1993, 1994, 1995, and 2000 are not sufficiently stable compared to other years because of satellite orbit drift, and AVHRR sensors degradation. For our research the data for 1982, 1983, 1985, 1986, 1989, 1990, 1996, 1997, 2001, and 2002 are assumed to be standard due to the fact, that equator crossing time of satellite fall within 1330 and 1500, and hence maximizing the value of coefficients. The crux of the proposed correction procedure is to slice the standard year's data set into two subsets. The subset 1 (1982, 1985, 1989, 1996, 2001) called standard data correction sets is used for correcting unstable years and then corrected data for this years compared with the standard data in the subset 2 (1983, 1986, 1990, 1997, 2002) that are used for data validation. The main objective of this paper is to correct the NDVI data for unstable years by the method of Empirical Distribution Functions (EDF). The corrected dataset can be applied to improve public health through reduced malaria burden, cost-effective malaria risk control using remote sensing and environmental data, and contributions to development of a Malaria Early Warning System for Bangladesh which represented globally.

2. DATA AND PROCESSING

Satellite data were collected from the NOAA/NESDIS Global Vegetation Index (GVI) data set [8] [9] which is one of the most widely used satellite products worldwide. The GVI is produced by sampling and mapping the 4-km daily radiance in the VIS (Ch1, 0.58-0.68 μm), NIR (Ch2, 0.72-1.1 μm) in Fig. 2 measured onboard NOAA polar-orbiting satellites, to a 16-km map. To minimize cloud effects, these maps, including the NDVI, solar zenith angle, and satellite scan angle, are composited over a 7- day period by saving those values that have the largest difference between VIS and NIR reflectance for each map cell. The weekly GVI data from January 1982 through January 1985 for NOAA-7, from April 1985 through September 1988 for NOAA-9, from October 1988 through August 1994 for NOAA-11, from March 1995 through December 2000 for NOAA-14, and from January 2001 through December 2003 for NOAA-16 were used here.
During 1985-2000, the performance of the channel 1 and 2 differed between NOAA-9, NOAA-11, and NOAA-14 satellites and most importantly, degraded over time for each satellite differently. Since there is no in-flight calibration of channel 1 and 2 of the AVHRR, the question arises as to the validity of the pre-launch calibration coefficients, both in the early days after launch and, perhaps more seriously, after the AVHRR has been in space for a long time. There is a clear evidence in several environmental products, such as the NDVI , global cloud morphology, and earth radiation budget [8] , that are generated from channel -1 and channel-2 AVHRR data to indicate that the performance of the instrument these two channels has deteriorated after launch. The need to correct for this in orbit degradation has been keenly felt recently since it is now being proposed to use the long-term records of AVHRRderived environmental products in climate and global change studies [18] [19] [20] [21] [5] and the degradation of the instrument with time is clearly illustrated by the results shown in Fig. 3. Therefore, the standard data preparation procedure for the 7-day composite time series now includes a correction of Ch1 (VIS) and Ch2 (NIR) values following Rao and Chen [20].

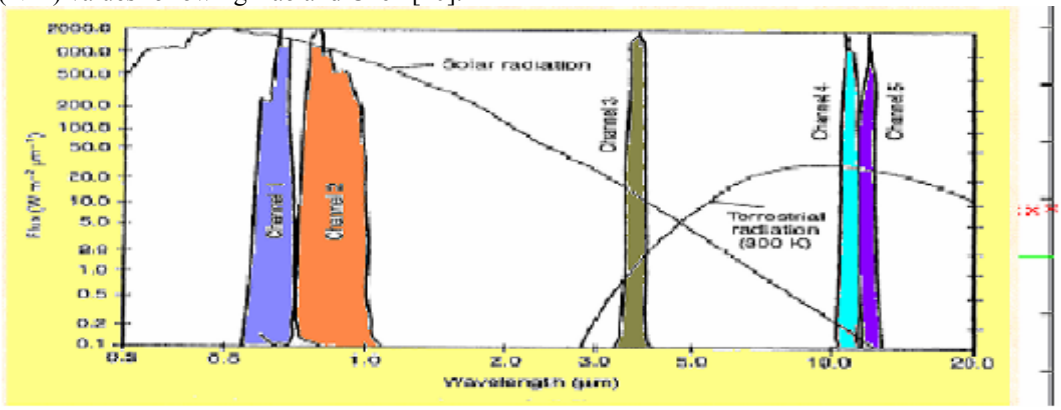

Fig. 2. Normalized spectral response of AVHRR

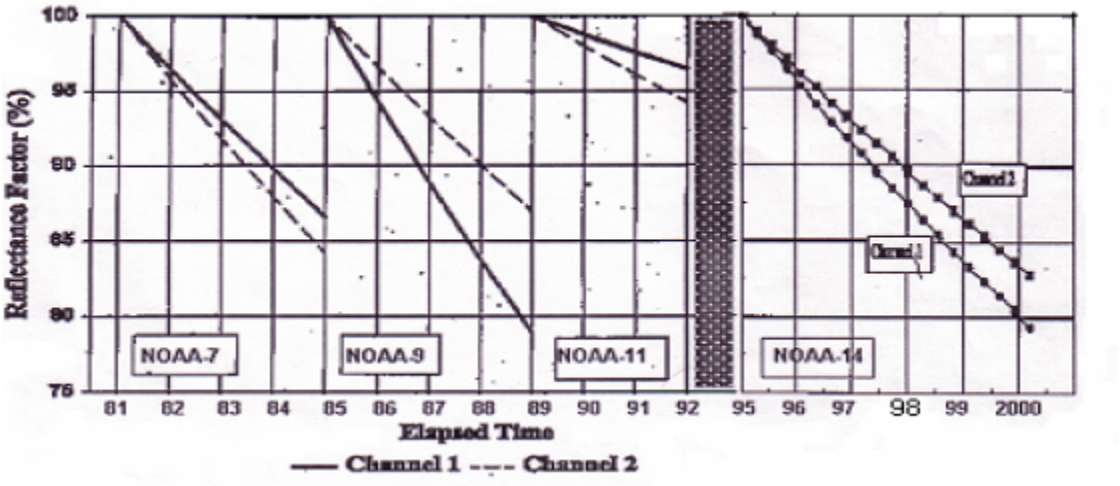

Fig. 3. Degradation of AVHRR Channels 1 and 2 [Rao, 1994]

3. METHODOLOGY

For each satellite, we construct the NDVI time series and also approximate linear trend using least square technique. From trend equation, we estimate two values: the largest difference ((dNt) between NDVI at the beginning (Nb) and the end (Ne) of satellite life and difference (dNs) between NDVI at the beginning of the next (n) satellite (Nbn) and at the end of the previous (p) one (Nep). [12]

$$dN_t = 100^* (N_e - N_b)/(N_b);\ dN_s = 100^*(N_{bn} - N_{ep})/N_{ep} \quad (1)$$

If the dNt values are positive then the NDVI time series upward trend and downward for negative value; positive dNs indicate larger NDVI at the end of the previous satellite and smaller NDVI in the opposite case.

There is no available physical method that can be used to correct for the stability of NDVI. Therefore, we developed a statistical model for the correction of NDVI. The empirical distribution function (EDF) is a statistical technique which is used to generate a normalization data of the years 1988, 1992, 1993, 1994, 1995 and 2000 compared with standard Empirical distribution function (EDF) approach is based on the physical reality, that each ecosystem may be characterized by very specific statistical distribution, independent of the time of observation. It is the best available technique to normalize satellite data. It allows us to represent global ecosystem from desert to tropical forest and to correct extreme distortions in satellite data related to technical problem. To generate the normalization data, we begin by selecting samples of unnormalized earth-scene data covering as much of the range intensities as possible. For NOAA satellites, the area will be rectangular, extending several thousand pixels from desert to tropical forest (both east to west and north to south). Corresponding to the incoming radiance from any pixel, the instrument will respond with an output, x . One can compile the discrete density function, i.e., the histogram, describing the relative frequency of occurrence of each possible count value, for each year. For year **i**, which is the year to be normalized, let the histogram be P**i**(x). An EDF, P**i**(x) can then be generated; viz.[25],

$$P_i(x) = \sum_{t=0}^{x} p_i(t) \quad (2)$$

The EDF is also known as a cumulative histogram of relative frequency. It is a non-decreasing function of x, and its maximum value is unity.
For convenience, however, we have chosen the maximum value to be 1; i.e., if the maximum possible output in counts is x, then $P_i(x) = 1$, as shown in Figure 4

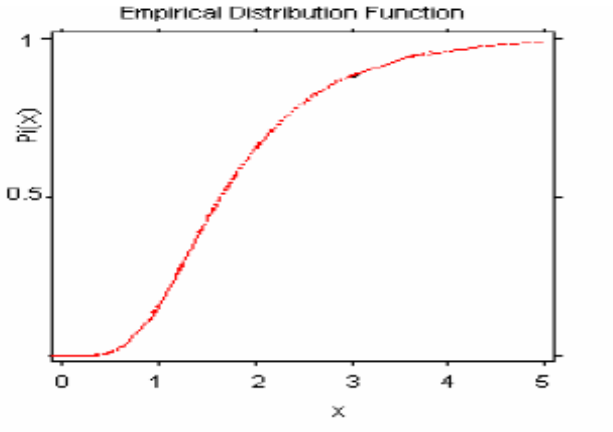

Fig. 4. Empirical Distribution Function

In these terms, the basic premise of normalization is that for each output value x in year i, the normalized value x' should satisfy [25]

$$P_s(x') = P_i(x), \quad (3)$$

Where the subscript **s** refers to the standard year. In practice, not only is P_s non-decreasing, but it is also monotonically increasing as a function of x' in the domain of x' where there are data. Therefore, it can be inverted, yielding the solution for x', [25]

$$x' = P_{s-1}(P_i(x)) \quad (4)$$

When it is applied sequentially for every possible count value x, equation 4 generates the normalization data relating each x to an x'. Fig. 5 shows how the procedure is applied in actual practice to generate the normalization data [25]. The figure shows idealized EDF's for the standard and unnormalized years i. In the figure the EDF's are continuous, but in practice they are discrete, being specified only integer values of x.

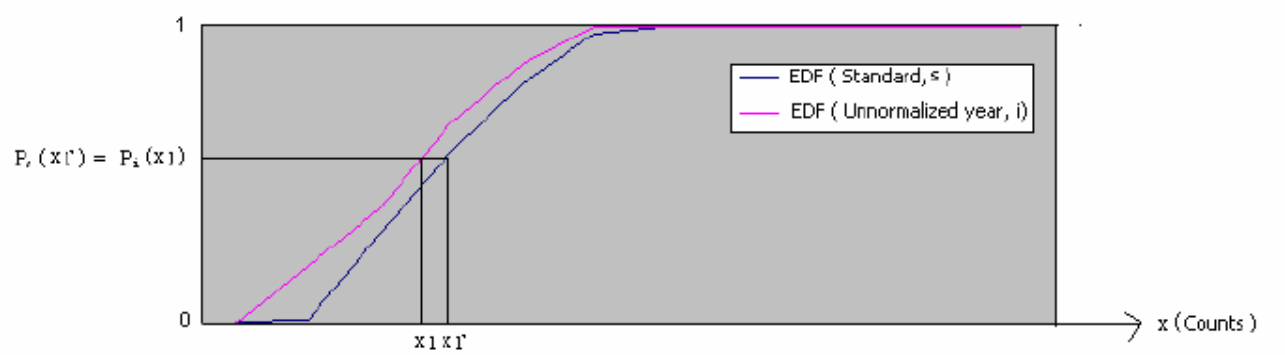

Fig. 5. Example of procedure to generate normalization data

To find x'_1, the normalized count value corresponding to the unnormalized count value of x_1, the following is the procedure: First, for the count value x_1 in unnormalized year i, , find the decimal or percentage value from the EDF of year i. In the illustration it is Pi (x_1). Then find the point on the standard year's EDF with the same decimal or percentage value. According to equation 3, that decimal or percentage can also be expressed as $P_s(x'_1)$. Finally, use the EDF of the standard year to find the normalized count value x'_1. Since the data are actually discrete, we will need to interpolate within the EDF of the standard year to find the value x'_1. Using this technique, we can generate the normalization data. Therefore, we choose EDF method for the normalization of satellite data to improve the cost effective malaria risk control .

4. RESULTS AND DISCUSSION

We produce NDVI time series of five NOAA satellites, which is illustrated in Fig. 6.

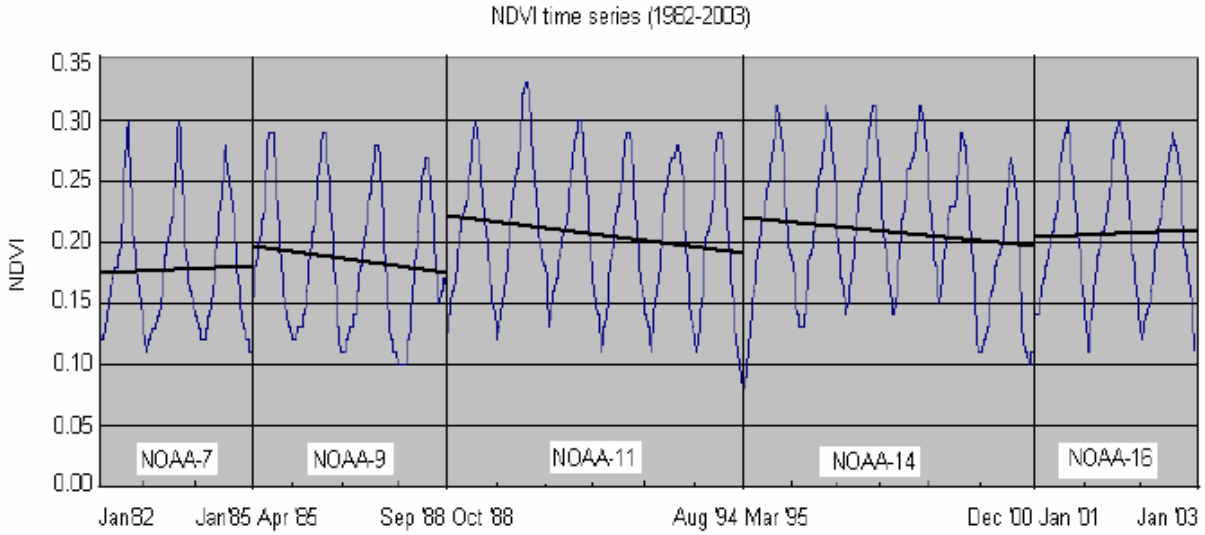

Fig. 6. NDVI time series (yearly old NDVI data).

Data from the afternoon polar orbiters is preferred for producing the NDVI time series because of the high sun elevation angle (low solar zenith angle). However, the equator crossing time drifts to a later hour as the satellites age [27]. Satellite orbit drift results in a systematic change of illumination conditions which is one of the main sources of non-uniformity in multi annual NDVI time series. Fig. 6 Shows that the NDVI data of 1988, 1992, 1993, 1994, 1995 and 2000 are nonuniform compared to other years because of satellite orbital drift, and sensor degradation. Therefore, we need to correct the data of those years. We apply EDF for the correction of data of those years. First, EDF construct for unnormalized data and than generate the normalize data compared with standard. Figure 8 shows how the procedure is applied in actual practice to generate normalization NDVI value [10].The figure shows idealized EDF's for the standard and the year of 1988. As EDF are based on cumulative histogram, they are discrete. But in Figure 7 they are shown as continuous function.

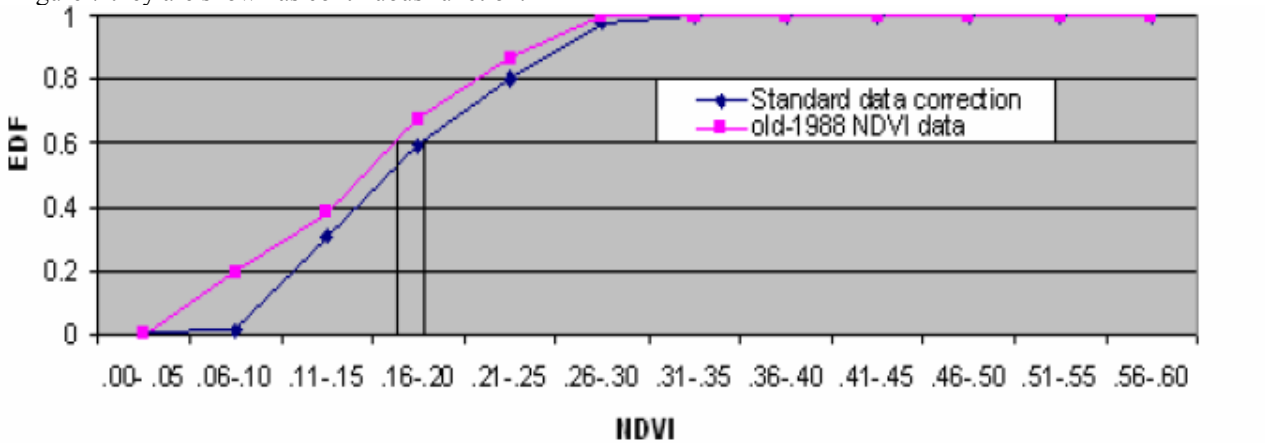

Fig. 7. Illustration of procedure to generate normalization NDVI data.

For example, for the NDVI value 0.16 in year 1988 (Fig. 7) find the value from the EDF of year 1988. In the illustration it the EDF_{88} is 0.6. Then find the point on the standard year's EDF with the same EDF value. According to equation 3, that the EDF value can also be expressed as the $EDF_{standard}$ is 0.6. Finally, use the EDF of the standard year to find the normalized count value 0.18. Since the data are actually discrete, we will need to interpolate within the EDF of the standard year to find the value of 0.18. Therefore,

New NDVI value for 1988 = $NDVI_{1988}$ + ($NDVI_{standard}$ -$NDVI_{1988}$) or

New NDVI value for 1988 = 0.16 + (0.18-0.16) = 0.18

Using this technique, we corrected data for the years 1988, 1992, 1993, 1994, 1995, and 2000 and a produced new NDVI time in Fig. 8 which shows improvement of NDVI data (pink line) for the year 1988, 1992, 1993, 1994, 1995,

and 2000 which can be used for cost-effective malaria risk control and contributions to development of a Malaria Early Warning System for Bangladesh. This model also can be used globally for prediction in vector born diseases such as malaria, dengue etc.

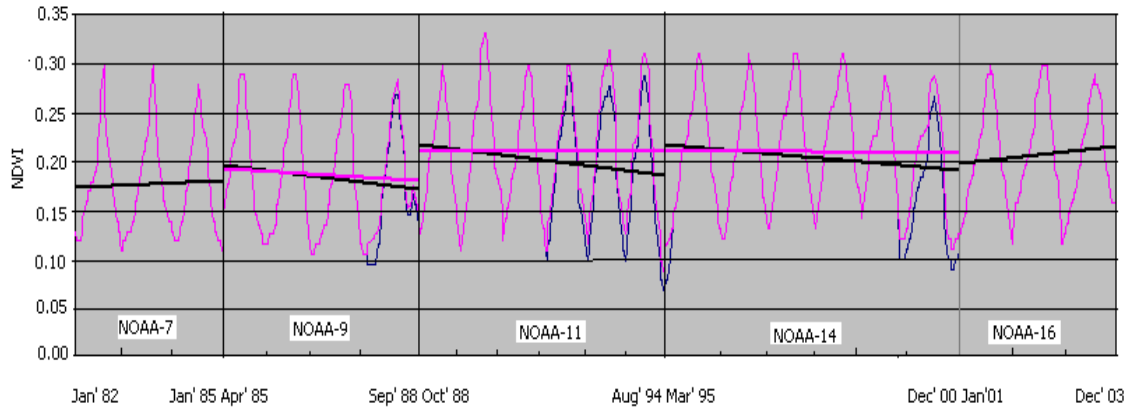

Fig. 8 New NDVI time series (weekly) (old NDVI data —— , and new NDVI data ——).

5. CONCLUSIONS

Empirical distribution function approach proposed here can be used to correct GVI and similar satellite date sets for improving public health reduced malaria burden, cost-effective malaria risk control and contributions to development of a Malaria Early Warning System for Bangladesh which represented globally.It should also be noted, that resolution of the EDF technique is limited by the available representative sample. For years when EDF's have steps,or flat spots, it results in errors. These are caused by the intensity levels with a zero frequency of occurrence in the histograms and are, therefore, an artifact of digitization or normalization. Typically, it should be determined by simulation. In addition, EDF's are only applicable to continuous distributions.

REFERENCES

[1] WHO/UNICEF. World Malaria Report. Roll Back Malaria partnership. World Health Organization/United Nations Children's Fund. http://rbm.who.int/wmr2005/, (2005)
[2] Ahmed TU. Checklist of the mosquitoes of Bangladesh. Mosquito systematic. 19(3):187-200, (1987)
[3] Birley, MH., An historical review of malaria, kala-azar, and filariasis in Bangladesh in relation to the Flood Action Plan. Ann Trop Med Parasitol. 87: 319-334., (1993).
[4] Beck, L. R., M. H. Rodriguez, S. W. Dister, A. D. Rodriguez, E. Rejmankova, A. Ulloa, R. A. Meza, D. R. Roberts, J. Paris, R. K. Washino, C. Hacker, and L. J. Letgers. Remote sensing as a landscape epidemiological tool to identify villages at high risk for malaria transmission. Am. J. Trop. Med. Hyg. 51: 271-280., (1994)
[5] Brest, C. L., and W. B. Rossow, "Radiometric calibration and monitoring of NOAA AVHRR data for ISCCP,"Int. J. Remote Sens., 13, 235-273, (1992)
[6] Gutman, G. Garik, "On the use of long-term global data of land reflactances and vegetation indices derived from the AVHRR," J. Geophys. Res., 104, 6, 6241-6255, (1999).
[7] Hay SI, Snow RW, and Rogers DJ. Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. Trans R Soc Trop Med Hyg. 92: 12-20., (1998)
[8] Kidwell, K., "Global Vegetation User's Guide, NOAA," Natl. Environ. Satellite Data and Inf. Serv. Natl. Clim. Data Cent., Dep. Of Commer., Washington, D. C., 1997.
[9] Kidwell, K. B., NOAA Polar Orbiter Data Users Guide. U.S. Department of Commerce. National Oceanic and Atmospheric Administration, Satellite Data Services Division, Washington, D.C., (1991).
[10] Klinkenberg E, van der Hoek W, and Amerasinghe FP. A malaria risk analysis in an irrigated area in Sri Lanka. Acta Trop. 89: 215-25.,(2004)
[11] Kaya S, Pultz TJ, Mbogo CM, Beier JC, Mushinzimana E,. The use of Radar remote sensing for identifying environmental factors associated with malaria risk in coastal Kenya. IGARSS, Toronto, June 24-28. Submitted to

the International Geoscience and Remote Sensing Symposium (IGARSS '02), Toronto, June 24-28, (2002).
[12] Kogan, F.N., and X. Zhu, "Evolution of long-term errors in NDVI time series," Adv. Space Res. Vol. 28 (1), page149-153(2001)
[13] Mushinzimana E, Munga S, Minakawa N, Li L, Feng CC, Bian L, Kitron U, Schmidt C, Beck L, Zhou G, Githeko AK, Yan G. (2006). Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands. Malar J. 16;5:13.
[14] Rosenberg R, and Maheswary, NP., Forest malaria in Bangladesh. I. Parasitology.Am J Trop Med Hyg. 31: 175-82., (1982a)
[15] Rosenberg R, and Maheswary, NP., Forest malaria in Bangladesh. II. Transmission by Anopheles dirus. Am J Trop Med Hyg. 31:183-91., (1982b).
[16] Rahman, A., Kogan, F. and Roytman, L., Analysis of malaria cases in bangladesh with remote sensing data. Am J Trop Med Hyg 74: 17-19., (2006).
[17] Rodriguez, A. D., M. H. Rodriguez, J. E. Hernandez, S. W. Dister, L. R. Beck, E. Rejmankova, and D. R. Roberts.,Landscape surrounding human settlements and Anopheles albimanus (Diptera: Culicidae) abundance in Southern Chiapas, Mexico. J. Med. Entomol. 33: 39-48, (1996).
[18] Rao, C.R.N. and J.Chen, "Nonlinearity corrections for the thermal infrared channels of the Advanced Very High Resolution Radiometer": assessment and recommendations. NOAA Technical Report NESDIS 69 (Washington DC: US Department of Commerce, (1993a)
[19] 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)
[20] Rao, C.R.N. and J.Chen, "Post-launch calibration of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on NOAA-7, -9, and -11 spacecraft". NOAA Technical Report NESDIS 78 (Washington, DC: US Department of Commerce).,(1994).
[21] Rao, C.R.N. and J.Chen, "Postlaunch calibration of the visible and near-infrared channels of the advanced very high resolution radiometer on the NOAA 14 spacecraft", 1, Int. J. Remote Sens., 7, 2743-2747, (1996).
[22] P. Cracknell, The Advanced Very High Resolution Radiometer, Taylor & Francis, page 90-97, (1997)
[23] Sithiprasasna R, Linthicum KJ, Liu GJ, Jones JW, and Singhasivanon P., Use of GIS-based spatial modeling approach to characterize the spatial patterns of malaria mosquito vector breeding habitats in northwestern Thailand. Southeast Asian J Trop Med Public Health. 34: 517., (2003)
[24]Thomson MC, Connor SJ, D'Alessandro U, Rowlingson B, Diggle P, Cresswell M, Greenwood B., Predicting malaria infection in Gambian children from satellite data and bed net use surveys: the importance of spatial correlation in the interpretation of results. Am J Trop Med Hyg. 61: 2-8., (1999)
[25] Weinreb, M.P, RXie, J.H.Lienesch, D.S Crosby . "Removing stripes in GOES images by Matching Empirical Distribution Functions," NOAA Technical Memorandum NESDIS,(1989)
[26] Wood, R., R. K. Washino, L. Beck, K. Hibbard, M. Pitcairn, D. R. Roberts, E. Rejmankova, J. Paris, C. Hacker, J. Salute, P. Sebesta, and L. Legters., Distinguishing high and low anopheline-producing rice Þelds using remote sensing and GIS technologies. Prev. Vet. Med. 11: 277-288., (1991).
[27] Jin, Menglin, R. E. Trendon., Correcting the orbit drift effect on AVHRR land surface skin temperature measurements, , International Journal of Remote Sensing 23, 4543-4558, (2003)