Using AVHRR-Based Vegetation Health Indices for Estimation of Potato Yield in Bangladesh

Rahman A[1]*, Kaisar Khan[2], Krakauer NY[3], Roytman L[3] and Kogan F[4]

[1]ICET/SETAS, Centennial College, Toronto,ON, M1G 3T8, Canada
[2]School of Applied Technology, Humber College, Toronto, ON M9W 5L7, Canada
[3]NOAA–CREST, CCNY, 138th Street and Convent Ave, New York, NY 10031, USA
[4]NOAA/NESDIS, 5200 Auth Rd. Camp Spring, MD 20746, USA

Abstract

We developed a model correlating Advanced Very High Resolution Radiometer (AVHRR)-based Temperature Condition Index (TCI) and Vegetation Condition Index (VCI) with potato yield in Bangladesh. Weekly TCI and VCI Indices for 1993–2005 along with official potato yield statistics were used for model development. We found a strong correlation between inter annual variation in potato yield and TCI and VCI during the critical period of potato growth (vegetative phase and tuber initiation), mid-December to mid-January. Principal component regression (PCR) was used to construct a model to predict potato yield as function of TCI and VCI. The model explained about 75% of year-to-year variation in potato yield. Remote sensing offers valuable and readily obtainable information on potato yield well in advance of harvest.

Keywords: Remote Sensing; AVHRR; Principal component regression; Vegetation condition index; Temperature Condition Index (TCI)

Introduction

Potato is an important vegetable crop in Bangladesh as well as nearby countries including India and Indonesia. Potato is the third largest crop in Bangladesh after rice and wheat and therefore contributes much to the total food supply of the country [1]. Potato cultivation in Bangladesh has significantly increased over the past years, so that variation in yield has become a growing concern. Potato, as a plant native to a temperate climate, thrives best under certain climatic, soil and physiographic conditions. Adaptation of the plant to tropical and subtropical climate has been successfully carried out through plant breeding and plant propagation methods.

Bangladesh is situated between 20.70°–26.80° N latitude and 88.01°–92.75° E longitude. Four out of Bangladesh's 64 districts – Rangpur, Bogra, Dhaka and Comilla – produce 60% of the national potato crop and account for 50% of the land planted in potato. The average yields of potato in Bangladesh is only around 15 ton/hectare although the yield potential, or achievable yield, has been reported to be as high as 40 ton/ hectare [2,3].

Satellite remote sensing is one of the best tools available to derive accurate and timely information on the spatial distribution of crops and growing conditions over large areas and could be used for crop growth monitoring and yield estimation. Bala and Islam [4] attempted to estimate potato yield in and around Munshiganj district area using TERRA MODIS 8-day composite surface reflectance. Only two years of data were used to develop a regression model, and the statistical significance of the obtained model was not investigated. The current study is the first to attempt to estimate countrywide potato yield from remote sensing. We use AVHRR-based vegetation health indices and 13 years of potato yield statistics to develop a statistical model for estimating inter annual variation in potato yield from remote sensing information.

Data and Methods

Potato yield statistics and satellite data for Bangladesh were used in this study. Potato production data were collected from the Bangladesh Bureau of Statistics, which estimates potato production and cultivated area from sampling surveys. Yield was calculated by

dividing total potato production by the cultivated area. The 1993–2005 Bangladesh potato yields (Y) time series is shown in Figure 1a. Yield has increased by almost 40% over this period, which is attributable primarily to improved breeding, fertilizer use, and other changes in
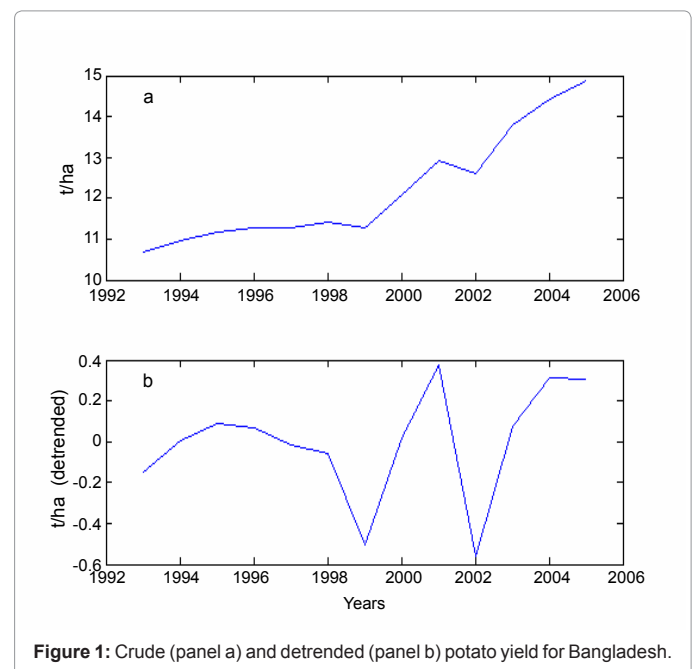

Figure 1: Crude (panel a) and detrended (panel b) potato yield for Bangladesh.

growing practices and agricultural technology. Here, our focus is in the high-frequency inter annual variation in potato yield (for example, the relatively lower yields in 1999 and 2002), which is in part attributable to variation in climate conditions before and during the growing season and therefore might be expected to correlate with remotely sensed land surface state. To extract the high-frequency component from the long-term increasing trend, we estimate the trend component of potato yield from a five-year boxcar moving average:

$$Y_{trend}(t) = (Y(t-2) + Y(t-1) + Y(t) + Y(t+1) + Y(t+2))/5 \quad (1)$$

Near the beginning and end of the time series when values of Y for earlier or later years were not available, we substituted the first or last year's yield (respectively) in the moving average in place of the missing values. We subtracted the trend component from the yield statistics to obtain a time series (Figure 1b) of interannual fluctuation in yield (DY):

$$DY = (Y - Y_{trend}) \quad (2)$$

The satellite data used were from the NOAA Global Vegetation Index (GVI) data set. The GVI data set was developed by aggregating the 4 square km Global Area Coverage (GAC) daily AVHRR product to 16 square km spatial resolution and seven–day composite [5,6]. GVI is based on measured intensities from the visible (VIS, 0.58–0.68 µm, Ch1), near infrared (NIR, 0.72–1.00 µm, Ch2) and infrared (IR, 10.3–11.3 µm, Ch4) AVHRR channels. Post-launch-calibrated VIS and NIR intensities were converted to reflectance [5] and used to calculate the Normalized Difference Vegetation Index (NDVI). The Ch4 counts were converted to brightness temperature (BT) [6,7].

The TCI and VCI indices were calculated from NDVI and BT. Details of the algorithm are presented in [6]. The Vegetation Condition Index (VCI) characterizes plant greenness and the Temperature Condition Index (TCI) characterizes surface temperature. The TCI and VCI indices are scaled to range from 0 (severe vegetation stress) to 100 (exceptionally favorable conditions) [6,8].

Satellite data were collected over land pixels for 6 administrative divisions of Bangladesh. In each administrative division spatial average values of NDVI and BT were calculated for each week during 1993–2005. VCI and TCI for the entire Bangladesh were calculated as averages over the six administrative divisions [9]. For consistency, weekly TCI and VCI were detrended by subtracting a five-year moving average in exactly the same way as the potato yield time series was. Thus, we are looking at the correlation between interannual anomalies in TCI and VCI and interannual anomalies in potato yield.

Environmental Influence on Potato Yield

Bangladesh has a tropical monsoon climate: the seasons are pre-monsoon (March to May), monsoon (June to October), and winter (November to February). The annual rainfall ranges from 1500 mm (60 inches) to 3000 mm (120 inches), monthly mean minimum temperature varies between 11°C to 28°C, and monthly average humidity varies from 59% to 89%. Potato growing season starts in November and ends in April (the Rabi agricultural season). Potatoes grow best in relatively cool, moist conditions. They are quite sensitive to moisture stress over much of the growing season, so they need relatively high soil moisture levels (60–80% of field capacity) to achieve high yields and good quality. The main source of water for potato production is the moisture stored from the monsoon rains [10]. The development of potatoes can be broken down into four distinct growth stages: a) Vegetative growth, b) Tuber initiation, c) Tuber growth, and d) Maturing.

Results and Discussion

Figure 2 shows dynamics of correlation coefficients for potato yield versus VCI and TCI for Bangladesh. These results were analyzed first, to interpret the response of potato productivity to moisture (VCI) and thermal (TCI) conditions expressed by the vegetation health indices. As seen in Figure 2, in the early period of the potato growing season, which is vegetative phase (weeks 50–51, December), the correlation for TCI is 0.52–0.55. Potato yield also correlates highly with VCI (r = –0.45 – –0.52) during mid-December to mid-January (weeks 51–52, 1–2) which corresponds to the critical vegetative growth and tuber initiation stages of the potato crop. Figure 2 confirms that positive correlation with TCI indicates that above average yield is associated with higher TCI (cooler thermal condition) during weeks 50–51 and negative correlation with VCI indicates that below average yield is associated with higher VCI (high rainfall) during weeks 51–52, 1–2.

While the observed positive correlation of potato yield with TCI (i.e. cooler surface temperature corresponds to higher yield) is consistent with our understanding of the plant's physiology, the observed negative correlation of potato yield with VCI (i.e. less green surface corresponds to higher yield) is unexpected. We hypothesize that denser vegetation during the early winter is associated with relatively warm climatic conditions, and thus with less favorable conditions for potatoes, which do better under cool conditions.

The results of correlation analysis in (as shown in Figure 2) were used to develop regression equations of DY versus VCI. Equation 3 is the multiple regressions between DY and VCI and TCI. The result of fitting the multiple regression model given by Equation 3 is shown in Table 1.

$$DY = b_0 + b_1 TCI_{50} + b_2 TCI_{51} + b_3 VCI_{51} + b_4 VCI_{52} + b_5 VCI_1 + b_6 VCI_2 \quad (3)$$

The linear regression model of Equation 3 is not significant at $p < 0.05$, with an F value of 3.39. This is because too many variables are included in the regression model. One strategy would thus be to delete some of the weekly TCI and VCI indices from the predictive model, even though they all show substantial correlation with yield. Alternatively, because TCI and VCIs for consecutive weeks are highly correlated, we can condense the information content of the weekly TCI and VCI values into a smaller number of principal components that can be used as predictor variables instead. Adopting this approach, we used
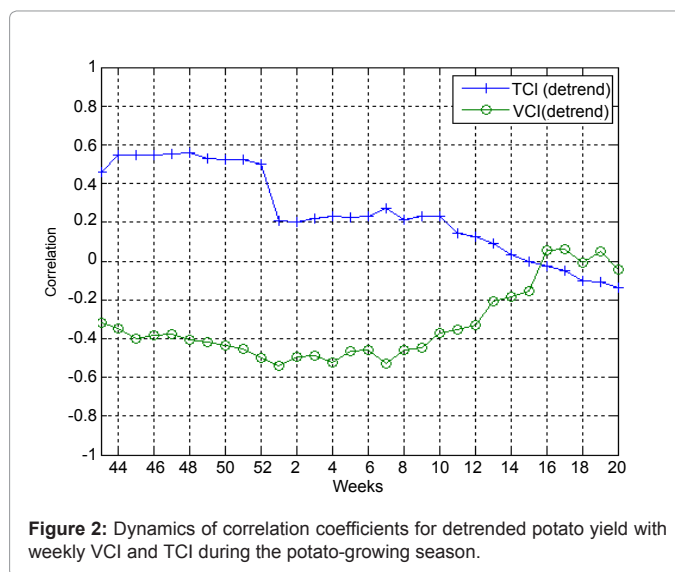

Figure 2: Dynamics of correlation coefficients for detrended potato yield with weekly VCI and TCI during the potato-growing season.

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-0.00117	0.05275	-0.02	0.983
TCI_{50}	1	0.12918	0.05588	2.31	0.0601
TCI_{51}	1	-0.11974	0.05468	-2.19	0.0711
VCI_{51}	1	0.18121	0.06152	2.95	0.0258
VCI_{52}	1	-0.22478	0.09882	-2.27	0.0633
VCI_1	1	-0.08181	0.05906	-1.39	0.2153
VCI_2	1	0.11916	0.03853	3.09	0.0213

R^2 =0.77, RMSE=0.18, F= 3.39, P=0.08

Table 1: Results of multiple linear regression (OLS) of DY on the equation (3).

an alternative method of estimation, principal component regression (PCR), which is known to often result in more robust estimation and prediction than OLS [11].

Using PCR methodology, the variables corresponding to weekly TCI and VCI was transformed into new orthogonal or uncorrelated variables, the principal components (PCs) of the correlation matrix. In PCR, PCs were sequentially tested for their contribution to improve the regression model for potato yield, keeping only those that resulted in a significant (at the 0.05 level) reduction in residual variance [9]. The first PC, corresponding closely to the average TCI over weeks 50–51 and VCI over weeks 51–52, 1–2, was the best predictor, with PC 4 and 6, corresponding to more detailed patterns in TCI and VCI, also significant predictors. The PCR model altogether explained 75% of the variance in determining potato yield. We then re-expressed the regression coefficients from the PCR analysis in terms of the TCI and VCI variables in Equation 3 to develop a PCR-based yield estimation model:

$$DY=100+0.12*TCI_{50} - 0.12*TCI_{51}+0.19*VCI_{51} - 0.25*VCI_{52} - 0.06*VCI_1+0.11*VCI_2 \quad (4)$$

Conclusion

This globally applicable technique for monitoring vegetation health, including drought, from AVHRR data was applied for statistical modeling of potato yield in Bangladesh. Correlation and principal component regression analysis relating potato yield deviation from the technological trend (DY) with the vegetation condition index (VCI and TCI) during 1993–2005 showed strong correlation during the critical period of potato growth (vegetation and tuber set/initiation phase). Thus, whether relatively high or low yield is likely can be determined well in advance of harvest based on remotely sensed climate conditions. Weekly sampled vegetation health data are available in real time at http://orbit.nesdis.noaa.gov/smcd/emcb.

References

1. Wennergren EB, Antholt CH, Whitaker MD (1984) Agricultural development in Bangladesh. USA: Westview Press.

2. Azad AK, Nabi MN (1984) Agro-economic analysis of cropping pattern involving jute and potato. Bangladesh J Jute & Fibre Research **7:** 129-138.

3. Elias M, Hossain I, Mondal NI, Islam N (1984) Potato production in Bangladesh: Socio-economic assessment of improved technology and identification of constraints to its higher production. Joydebpur: AED/BARI. Report No 84-87.

4. Bala SK, Islam AS (2009) Correlation between potato yield and MODIS derived vegetation Indices. Int J Remote Sens 30: 2491-2507.

5. Kidwell KB (1997) Global Vegetation Index Users Guide NOAA. Technical Report.

6. Kogan FN (1997) Global drought watch from space. Bull Am Meteorol Soc 78: 621-636.

7. Kogan FN (2001) Operational space technology for global vegetation assessment. Bulletin of the American Meteorological Society 82: 1949-1964.

8. Rahman A, Kogan F, Roytman L (2006) Analysis of malaria cases in Bangladesh with remote sensing data. Am J Trop Med Hyg 74: 17-19.

9. Rahman A, Roytman L, Krakauer NY, Nizamuddin M, Goldberg M (2009) Use of Vegetation Health Data for Estimation of Aus Rice Yield in Bangladesh. Sensors 9: 2968-2975.

10. Rashid MM, Hussain MM (1985) Development of potato in Bangladesh–policies and programs Annual Report of Joydebpur. Potato Research Centre, BARI.

11. Draper N, Smith H (1981) Applied Regression Analysis, New York, Wiley.