Developing a Mathematical Model For Wheat Yield Prediction Using Landsat ETM+ Data

M. Aboel Ghar
CEReS, Chiba Univerdity
1-33 Yayoi cho, Inage ku, Chiba 263-8522 Japan
ghar@graduate.chiba-u.jp

A. Shalaby
CEReS, Chiba Univerdity
1-33 Yayoi cho, Inage ku, Chiba 263-8522 Japan
adel@graduate.chiba-u.jp

R. Tateishi
CEReS, Chiba Univerdity
1-33 Yayoi cho, Inage ku, Chiba 263-8522 Japan
tateishi@ceres.cr.chiba-u.ac.jp

Abstract: Quantifying crop production is one of the most important applications of remote sensing in which the temporal and up-to-date data can play very important role in avoiding any immediate insufficiency in agricultural production. A combination of climatic data and biophysical parameters derived from Landsat7 ETM+ was used to develop a mathematical model for wheat yield forecast in different geographically wide Wheat growing districts in Egypt. Leaf Area Index (LAI) and fraction of Absorbed Photosynthetically Active Radiation (fAPAR) with temperature were used in the modeling. The model includes three sub-models representing the correlation between the reported yield and each individual variable. Simulation results using district statistics showed high accuracy of the derived correlations to estimate wheat production with a percentage standard error (%S.E.) of 1.5% in El- Qualyobia district and average (%S.E.) of 7% for the whole wheat areas.

Key Words: Mathematical Model, Wheat Yield Prediction, Landsat7 ETM+

1. Introduction

Estimation of agricultural production is a key element of rural development and an indicator for national food security. Crop yield forecasts a few months before harvest can be of a paramount importance for timely initiation of food trade to secure the national demand and timely organization of food transport within countries [3]. Wheat is one of the most important stable foods in both local and regional scales. In Egypt, Wheat represents approximately 1.01 Million Hectare (about 30% from the whole agricultural areas during winter season). Therefore, there is a high need to develop an early warning system to help the strategists and decision makers to keep away from any unexpected lack in the final crop production. Basically, estimation of the total agricultural production could be done through yield estimation and crop area determination. In the aspect of yield estimation, statistical relationships between Biophysical factors derived from remote sensing data and crop yield have been frequently used as the main component of yield modeling approach. Also, meteorological conditions have been considered as a limit factor for the crop yield because of the strong influence of climate on the growing cycle and physiological conditions.

Many research works have been done in order to figure out the most suitable method to predict crop production with maximum level of accuracy. [4]Combined Perpendicular Vegetation Index (PVI) with meteorological data to establish wheat yield prediction model using NOAA-AVHRR data. In many cases, LAI was used as a key factor in crop growth and yield prediction models that were built up using remote sensing. Among these efforts, [1], developed a statistical correlation between LAI derived from Landsat TM data and dry biomass of Corn crop in the upper Rhine valley. Satellite derived LAI; meteorological and agronomic variables with statistical regression analysis and analytical capabilities of GIS were used to develop a model for Tea yield estimation [8]. [5] Suggested a linear relation between the yield of Corn and Wheat and various weighted and summed spectral quantities. Absorbed Photosynthetically Active Radiation (APAR) was also used worldwide as a main element in yield modeling. Lobell et el., 2003 inserted fraction of Absorbed Photsynthetically Active Radiation (fAPAR) to create a simple model based on crop light-use efficiency to predict yield and planting dates of Wheat.

In this study, both LAI, fAPAR derived from Landsat ETM+ data and air temperature as a climatic factor together with GIS approach were utilized to make simple models representing the relation between Wheat yield and each individual factor. Different combinations between these sub-models were tested in order to find the most suitable model form estimating Wheat yield production.

2. Materials and Methods

1- Study area

East of Nile Delta is one of the centers of Wheat production in Egypt. Wheat yield has increased gradually in this area in the last few years. As a test area, a region in Eastern Nile Delta south of Zagazig city was selected to analyze the potential of yield prediction for Wheat with remote sensing data. Figure (1) shows the location of the study area, situated in the north east of Egypt. The cultivation is very intensive in this area. The average Wheat farms areas varied between 4.6 and 21 Hectares. Due to low winter rainfall of Egypt, Wheat farms have to be irrigated for optimal production.

2- Meteorological data

Climatic stations of the Ministry of Agriculture in Egypt were used for information on the mean monthly temperature. Interpolation process according to [9] was carried out to convert monthly climatic data to ten-day composite data.

$$D2 = \frac{-M1 + 29M2 - M3}{27} \qquad (1)$$

Where D2 is the average of air temperature in the second decade of the month, M1 is average of air temperature in the previous month, M2 is the average of air temperature in the target month and M3 is the average of air temperature in the next month

Since the cultivation in this area depends mainly on a net of irrigation canals as the only resource of irrigation water, the precipitation data are not a necessary input for the applied yield formation model.

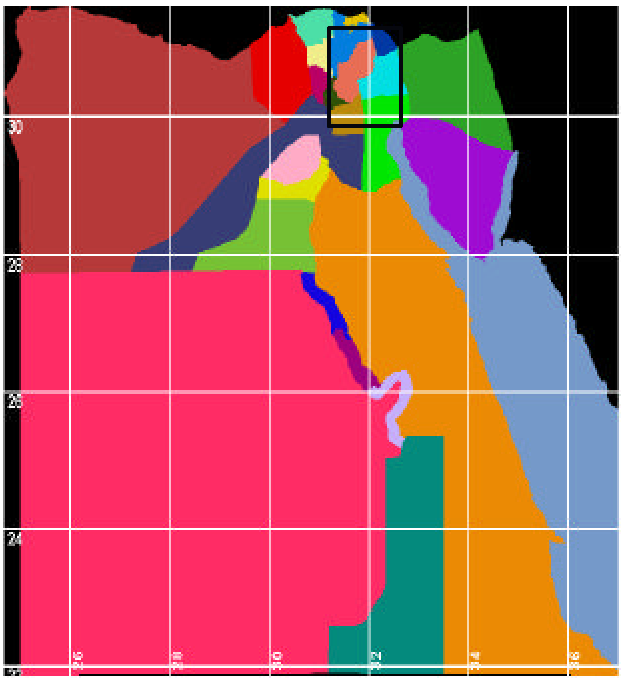

Figure (1): A map of Egypt explaining the study area

3- Satellite data processing

A time series of Satellite images acquired by Landsat ETM+ was used to determine plant development of the Wheat fields under investigation in order to use this information as input in an agrometeorological model for yield prediction. For the 2001 scenario, two ETM+ scenes with cloud free conditions were available during the vegetation period of Wheat (January 30 and March 19).

The geometric correction using ground control points enables the overlay of multi temporal images and combinations with map of GIS information [1]. The digital images were geo-referenced by nearest neighbor resampling algorithm with RMS error less than 0.5 pixel using digital 1:50,000 digital topographic maps. In this data set, the locations of 35 wheat fields were determined in the images.

Among the various spectral vegetation indices, the NDVI is the most common one. Many results can already be found in the literature relating NDVI and plant parameters for various agricultural species [10], [11], [1]. NDVI was calculated for each scene. LAI was calculated from NDVI based on a methodology created by [7]. Fraction of Absorbed Photosynthetically Active Radiation was also generated from NDVI using a methodology created by [2]. Generation of LAI and fAPAR was carried out in two different ways: using the real values of NDVI extracted from 35 Wheat points to calculate LAI and fAPAR. The second way is using GIS and the reflectance values from Wheat plants at the different spectral bands together with the real NDVI values to create Wheat LAI and fAPAR maps.

The temperature data collected from different climatic stations in all wheat production centers in Egypt, LAI, fAPAR and reported yield were used as main inputs for Wheat Yield prediction.

3. Results and Discussion

The modeling process was carried out in two main steps; generating simple relationships between reported yield and each individual factor and combining the models in more complicated model.

1- The relationship between Temperature and yield

The temperature of second decade of March was extracted from monthly temperature data and used with reported Wheat yield as two factors. Exponential Association was found to be the best formula to present the relationship between temperature and reported yield with (0.9) correlation coefficient

$$Yield = 6.46(1 - 2.7^{-0.35Temp}) \qquad (2)$$

The validation of this model was performed using the reported Wheat yield in 2002. (7.2%) was the percentage standard error (%S.E.) of the derived correlation.

2- The relationship between LAI and yield

LAI could be defined as the leaf area per unit ground area. LAI is a factor that indicates how many leaf (or photosynthetically active) surfaces are in column extended form, the ground area under the canopy diameter, up through the canopy. It was found that Exponential association is the best formula describing the relationship between LAI and Wheat yield. The form of the formula is as follows:

$$Yield = 8.01(1 - 2.7^{-1.32LAI}) \quad (3)$$

The correlation coefficient of this relationship was (0.7). The validation of this formula was carried out using the reported yield by generating (8%) as the % of standard error.

3- The relationship between fAPAR and yield

Rational Function was found to be the best formula describing the relationship between fAPAR and Wheat yield with (0.68) correlation coefficient.

$$Yield = \frac{3.49 * 10^{-7} + 52766073\, fAPAR}{1 + 8760142\, fAPAR - 2090808.5\, fAPAR^2} \quad (4)$$

The validation approach generated (7%) standard error

4- Combining sub-models

There are many forms of combined models using the three derived sub-models. The percentages of standard error were around (7%) for all cases. Among these combined models, the best model showed relative high accuracy is the combination between temperature and fraction of Absorbed Photosynthetically Active Radiation as follows:

$$2Yield = 6.46(1 - 2.7^{-0.35Temp.}) + \frac{3.49 * 10^{-7} + 52766073\, fAPAR}{1 + 8760142\, fAPAR - 2090808.5\, fAPAR^2} \quad (5)$$

4. Summary

Air temperature, fraction of Absorbed Photosynthetically Active Radiation and Leaf Area Index together with the reported yield of the years 2001 and 2002 were used to develop a mathematical model for Wheat yield prediction, which could be applied in large Wheat areas in Egypt. NDVI values of 35 Wheat points were used to calculate both LAI and fAPAR based on reviewed experimental equations. These two sub-models were validated using other points collected from LAI and fAPAR maps. Climatic data from different climatic stations in the year 2001 were used to create a sub-model between air temperatures and yield. This sub-model was validated using the climatic data of the year 2002. The model could be applied in almost all over the country. The different sub-models showed around 7% as average percentage of standard error. A trial to combine the three sub-models showed almost the same level of accuracy. The relative low accuracy in Wheat estimation may be due to the difference in climatic conditions and soil characteristics between the different parts of Egypt. Further work is necessary to increase the accuracy of the derived model by inserting other factors related to the final crop yield.

References

[1] Bach H., Yield estimation of corn based on multitemporal LANDSAT-TM data as input for agrometeorological model. Pure Appl. Opt., 7, 809 – 825, 1998.

[2] Baret, F., Guyot, G., and Major, D.J., Crop biomass evaluation using radiometric measurements, Photogrammetria (PRS), 43, 241 – 256, 1989.

[3] Bastiaanssen and Ali, A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan, Agriculture, Ecosystem and Environment 94, 321-340, 2003.

[4] Fuqin L. and Guoliang T., Research on remote sensing meteorological model for Wheat yield estimation, Proc. ACRS 1991, Singapore.

[5] Hamar D., Ferncz C., Lichtenberger J., Tarcsai G., Ferncz A., Yield estimation for corn and Wheat in the Hungarian great plain using Landsat MSS data, International Journal of Remote Sensing, 17, 9, 1689 – 1699, 1996.

[6] Lobell D., Anser G., Monasterio J., Benning T., Reomte sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties, Agriculture, Ecosystem and Environment 94, 205 – 220, 2003.

[7] Nemani, R. R. and Running S.W., Testing a theoretical climate-soil-leaf area hydrological equilibrium of forests using satellite data and ecosystem simulation, Agric. Forest Meteorol., 44, 245 – 260, 1989.

[8] Rajapakse R., Tripathi N., Honda K., Modelling tea (Camellia L O.Kuntze) yield using satellite derived LAI Landuse and Meteorological data, Proc. ACRS 2000, Taipei, Taiwan.

[9] Sys. C., Ranst. V and Debaveye. J. Land Evaluation Part I, Agricultural publication No. 7, ITC Ghent, 1991.

[10] Wiegand C L, Gerbermann A.H., Gallo K.P., Blad B. L. and Dusek D., Multisite analyses of spectral – biophysical data for corn Remote Sensing of Environment, 33, 1-16, 1990.

[11] Wiegand C.L., Mass S.J., Aasse J.K., Hartfield J. L., Pinter P.J., Jackson R.D., Kanemasu E. T., and Lapitan R. L., Multisite analyses of spectral-biophysical data for Wheat, Remote Sensing of Environment, 1992.