Analysing the effect of different geocomputational techniques on estimating phenology in India

Peter M. Atkinson[1], C. Jeganathan[2] and J. Dash[3]

[1,2,3]School of Geography, University of Southampton, 44 Shackleton Building, Highfield Campus
Southampton-SO17 1BJ, United Kingdom
Telephone: (44) 2380 594617
Fax: (44) 2380 593295
Email: pma@soton.ac.uk
[2]Email: jpc1s07@soton.ac.uk
[2]Email: j.dash@soton.ac.uk

1. Introduction

Vegetation phenology is considered as one of the key indicators that has provided real evidence of climate change (Cleland, 2007). Observation records suggest that spring has advanced globally at the rate of 2 to 5 days per decade (Parmesan and Yohe, 2003; Root et al., 2003) and that senescence has been delayed, mostly as a function of increases in global mean temperature (Menzel, 2003). However, the effects of climate change on vegetation phenology are species-dependent and, therefore, using ground observations of phenological events it is difficult to provide a global-scale picture. On the contrary, satellite-derived vegetation indices provide indirect estimates of vegetation phenological events through repeat coverage over the globe.

From the point-of-view of vegetation phenology, several geo-computational techniques have been used to correct and smooth time series vegetation index data, such as to estimate phenological variables accurately. Some of the frequently used methods in smoothing time-series satellite sensor data are Fourier harmonics based methods (classical and discrete), threshold based methods, curve fitting methods and point of inflection methods (Vivoy et al., 1992, Roerink et al., 2000, Jakubauskas et al., 2001; Wagenseil, 2006; Chen et al., 2001; Chen et al., 2006; Jonsson and Eklundh, 2002; Bradley et al., 2007; Zhang *et al.* 2003). However, the complexity of the method need not determine the accuracy of the method. Every method has its own advantages and disadvantages, and the choice of technique depends upon the real purpose of the study.

These techniques require a fine tuning of many model parameters such as noise-threshold, number of temporal neighbourhood and number of harmonics. The current study assumes that it would be difficult to find a single set of parameters that would be most suitable for all vegetation types of a diverse tropical landscape. Hence, the end user is left to choose a set of parameter values, which may lead to a potential difference of information between users and hence, issues of consistency may become a concern. The current study aims to analyse five techniques (harmonic analysis of time series (HANTS), asymmetric Gaussian, double logistic, Savitzky-Golay and discrete Fourier based methods) for smoothing temporal satellite sensor observations with the ultimate purpose of extracting phenological variables. Hence the current study analyses these techniques from the phenological view-point. The study aims to provide a comparative view and analysis of the capability of these smoothing techniques over multiple landscapes in India.

2. Study area, data and software

The techniques analysed in the study were applied over the complex landscapes of India. The study used the Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyl Index (MTCI). MTCI consists of 46 bands, each corresponding to 8 day composite. The study used the HANTS software ((Verhoef, 1996; Roerink et al., 2000) and TIMESAT software ((Jonsson & Eklundh, 2002, 2004) to execute four techniques (HANTS, Asymmetric Gaussian, Double Logistic and Savitzky-Golay). A separate software routine was developed, in this study, under ARCGIS using Arc-objects and Visual Basic for the discrete Fourier based smoothing.

3. Analysis, Results and Discussion

HANTS provides a choice of number of frequencies, fit error tolerance (FET), direction of outlier, degree of over-determinedness (DOD), damping coefficient (Delta) and scale parameters. The number of frequencies determines the degree of fluctuation in a curve. The first six frequencies were used in this study based on Dash et al., (2009) and Jeganathan et al., (2009). FET is defined by the user, which identifies the absolute acceptable deviation of a value from the fitted curve. Large FET values may eliminate few obvious outlier points and small FET values may remove too many points. DOD refers to number of parameters describing the curve. For a given frequency, say '*n*', the DOD must be greater than or equal to $2n+1$ (Verhoef, 1996) so as to make matrix inversion possible. The process stops if the number of remaining points becomes less than ($DOD+2n+1$). Delta values of 0.1 to 0.5 are recommended to obtain an appropriate damping effect. Scale provides the chance to upscale or downscale the derived data values. Scale was kept constant at 1 in our study. The direction of outlier indicates the direction (high or low) of expected deviation of data during the elimination (cleaning) process. Since there are many different possibilities, a random selection of various model parameters was undertaken and the model fit was checked visually. Qualitative elimination was undertaken first and then a detailed quantitative anlaysis was undertaken.

The HANTS model was run using various model parameters, selected randomly, over a pixel belonging to the Evergreen vegetation type. **Figure** 1 provides the resultant temporal variation from all the runs, along with the original values. It was found, through first level qualitative elimination (as an example using **Figure 1**), that the FET value of 10 and DOD value of 25 to 30 produced acceptable results. Further runs used values within these first level selected ranges, to determine the range of intermediate values with different Delta, to further quantitatively analyse the effect.

Normalised Euclidean Distance (NED) (Eqn. 1) and Spectral Angle Mapper (SAM) (Eqn. 2) measures were calculated between the reference data and the smoothed data to investigate the closeness of the final smoothed data with the original data.

$$NED = \sqrt{\sum_{i=1}^{n} (T(i) - R(i))^2} \quad (1)$$

$$SAM = \frac{\sqrt{\sum_{i=1}^{n} (T(i) - R(i))^2}}{\sqrt{\sum_{i=1}^{n} T(i)^2} \sqrt{\sum_{i=1}^{n} R(i)^2}} \quad (2)$$

T and *R* refer to the target and the reference data, respectively. Target data are the output from HANTS/TIMESAT/FFT using different parameters. The reference data are derived by correcting the original data at the error locations (i.e., deviations from normal phenological trend) through trend mean and utilising expert knowledge of expected temporal phenological influence at those reference locations. **Figure 2** shows a fitted function using optimal model parameters (FET, DOD and Delta) in HANTS and using Fourier (FFT) based fitting over an agricultural landscape. **Figure 3** provides a sample comparison of fitting using Savitzky-Golay (SG), Gaussian, Logistic fit (using TIMESAT) and Fourier (FFT) over another agricultural landscape situated in a different region of India.

It was found in this study that no single set of parameters from HANTS and TIMESAT provide an acceptable fit to all landscapes. Especially when the study area is as large (and complex) as India it is extremely difficult to set suitable parameters. However, the Fourier based approach using four harmonics was suitable for all natural vegetation types and Fourier with six harmonics was suitable for agricultural types. Moreover, the Fourier approach was found to be consistent in terms of algorithm, easy to implement, easy to replicate by different users and the user is required to set only one model parameter. Hence, the current study recommends the Fourier based approach for phenology extraction over vast areas covering complex landscapes. Fourier based fitting was tested over India and the results were validated through existing literature and other published research papers.

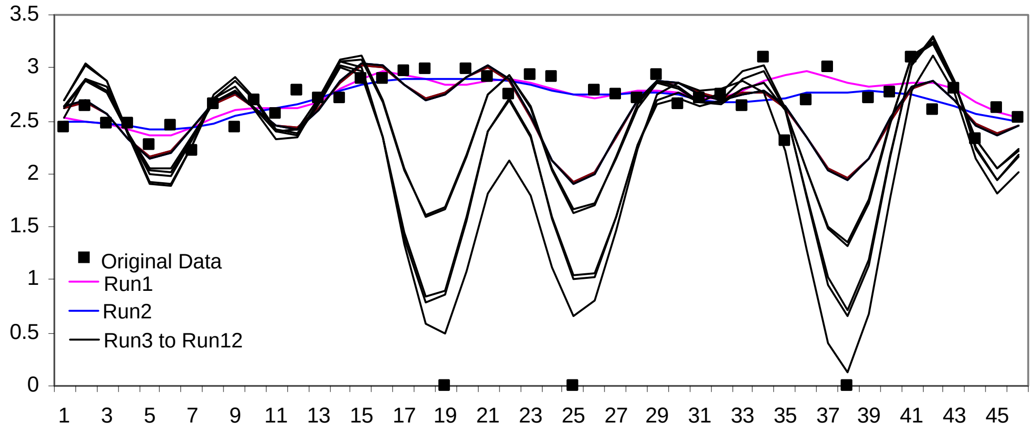

Figure 1. Effect of various HANTS Model Parameters over Evergreen Landscape

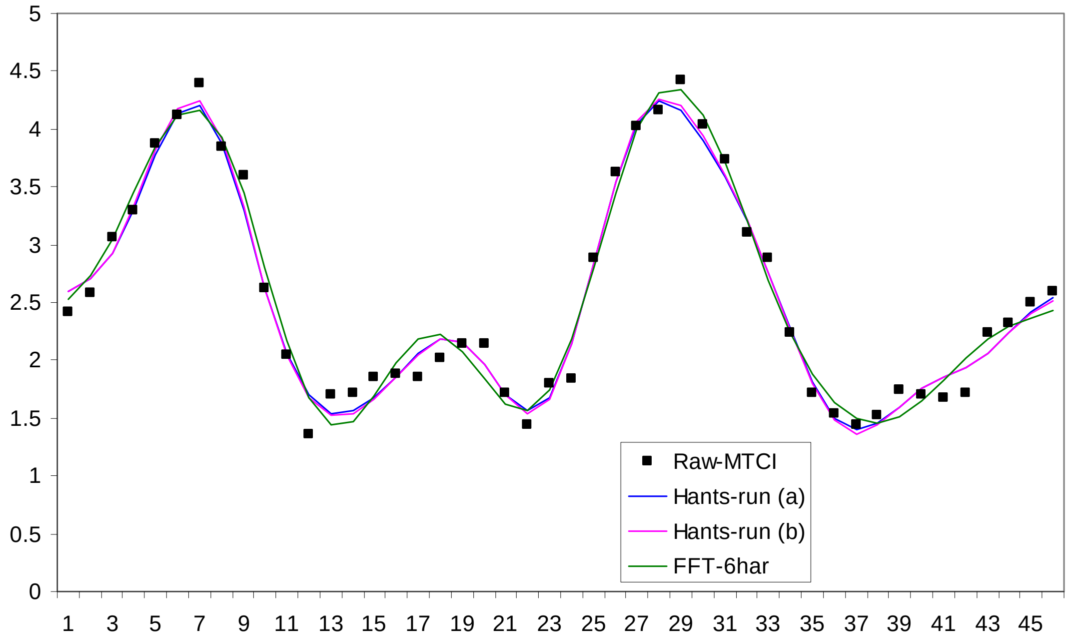

Figure 2. Fitting of HANTS (FET=0.4, DOD>=32) (delta=0.5 for run(a); delta=0.1 for run(b) for an Agricultural Ladscape

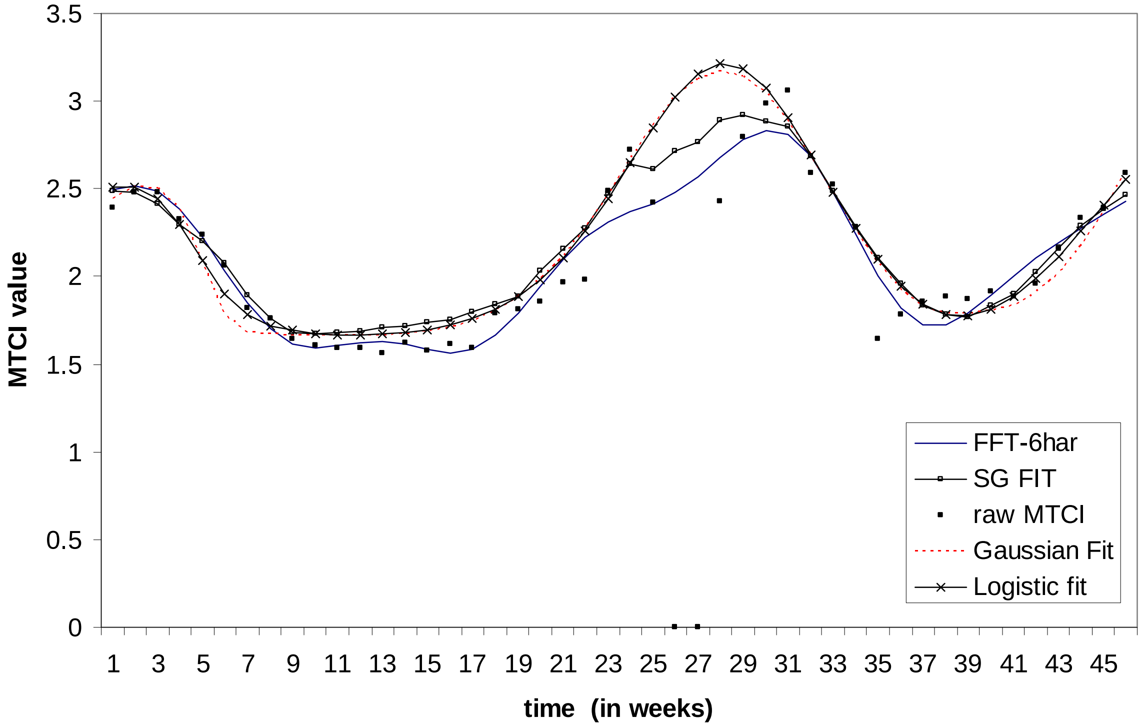

Figure 3**.** Fitting of Savitzky-Golay (SG), Gaussian, Logistic fit (using TIMESAT) and Fourier (FFT) for an Agricultural Ladscape

4. References

Bradley, B.A., Jacob, R.W., Hermance, J.F., Mustard, J.F., 2007, A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment*, 106:137-145**.**

Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A. and Schwartz, M. D., 2007, Shifting plant phenology in response to global change. *Trends in Ecology and Evolution*, 22(7): 357-365.

Chen, X., Xu, C., and Tan, Z., 2001, The analysis of relationships among plant community phenology and seasonal metrics of normalized difference vegetation index on the northern part of the monsoon region of China. *International Journal of Biometeorology*, 45:170–177.

Chen, J.M., Deng, F. and Chen, M.Z., 2006, Locally adjusted cubic-spline capping for reconstructing seasonal trajectories of a satellite-derived surface parameter. *IEEE Transactions on Geoscience and Remote Sensing*, 44:2230-2238.

Dash, J., Jeganathan, C., and Atkinson, P.M., 2009, Satellite derived spatio-temporal variations in vegetation phenology over India. (Submitted and in the final stage of publication).

Jakubauskas, M.E., Legates, D.R. and Kastens, J.H., 2001, Harmonic analysis of time-series AVHRR NDVI data. *Photogrammetric Engineering and Remote Sensing*, 67:461-470.

Jeganathan, C., Dash, J., and Atkinson, P.M., 2009, Mapping phenology of natural vegetation in India using remote sensing derived chlorophyll index. (Submitted and in the final stage of publication).

Jonsson, P., and Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions On Geosciences and Remote Sensing*, 40(8):1824-1832.

Jonsson, P., and Eklundh, L., 2004. TIMESAT-a program for analysing time-series of satellite sensor data. *Computers and Geosciences*, 30:833-845.

Menzel, A., 2003, Plant phenological anomalies in Germany and their relation to air temperature and NAO. *Climatic Change*, 57: 243-263.

Parmesan, C. and Yohe, G., 2003 A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(2):37-42.

Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig , c., and Pounds, J.A., 2003, Fingerprints of global warming on wild animals and plants. *Nature*, 421(2):57-60.

Roerink, G.J., Menenti, M., and Verhoef, W., 2000, Reconstructing cloudfree NDVI composites using Fourier analysis of time series, *International Journal of Remote Sensing*, 21(9):1911–1917

Viovy, N., Arino, O. and Belward, A.S., 1992, The Best Index Slope Extraction (BISE) - A method for reducing noise in NDVI time-series. International Journal of Remote Sensing, 13:1585-1590.

Verhoef, W., Menenti, M., and Azzali, S., 1996, A colour composite of NOAA–AVHRR–NDVI based on time series analysis (1981–1992), *International Journal of Remote Sensing*, 17(2):231–235.

Wagenseil, H., and Samimi, C., 2006, Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: results from a dry savannah environment in Namibia. International Journal of Remote Sensing, 27:3455–3471.

Zhang, X.Y., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C. F., Gao, F., et al., 2003, Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84:471-475.