A STUDY OF TIME SERIES NOISE REDUCTION TECHNIQUES IN THE CONTEXT OF LAND COVER CHANGE DETECTION

XI CHEN[†]*, VARUN MITHAL[†]*, SRUTHI REDDY VANGALA[†]*, IVAN BRUGERE[†]*, SHYAM BORIAH*, AND VIPIN KUMAR*

The purpose of this study is to introduce concepts relevant to performance of (i) change detection algorithms within (ii) various regional contexts with differing noise characteristics according to (iii) differing strategies of noise reduction. The relevant interrelations of these three elements are presented, and focused analysis is presented from the perspective of varying (i) and (iii) for a comparative analysis across (ii).

Six smoothing methods has been studied in this work: Savitzky-Golay (SG) method [7], The Savitzky-Golay method iterated to upper envelope (SG-Itr) [3], Harmonic Analysis of Time Series (HANTS) [6], Double Logistic function fitting method (DL) [1], Data Assimilation method(DA) [5]and a naive outlier identification and imputation scheme (SO).

In this work, we enumerate three general data characteristics, especially relevant in the MODIS EVI data, which a given noise reduction technique may take advantage of: neighborhood coherence, quality annotation and background model.

For a noise reduction technique we identify the following two questions to be of relevance:

- Which observations in the time series should be imputed?
- How are these observations to be imputed?

Based on the first question, the reviewed methods can then be organized into (1) *selective* and (2) *non-selective* imputation methods. Selective methods identify some observations that they consider noisy and ought to be imputed. On the other hand, in the non-selective methods every observation is imputed. We consider the selective methods to be more conservative as they modify fewer observations as opposed to the non-selective methods which modify every observation and therefore no processed data value corresponds to the real observation. Intuitively, if an observation is not clearly anomalous and is annotated as a high quality observation, the value reported by the MODIS is as trustworthy as can be ascertained. Time series smoothing methods should thus be considered the most aggressive because generally every observation of the original time series is modified without identifying trustworthy observations. Note that typically the imputations of selective methods will modify the observation by large magnitude because large outlier values are imputed in this case. The non-selective methods are less conservative and modify each value but the total modification in the value itself is generally of smaller magnitude for most observations.

Imputation is done primarily based on the three characteristics of neighborhood coherence, quality annotation, and background model. Most of the non-selective methods rely only on neighborhood coherence and use function fitting on temporal neighbors to eliminate the noise. In contrast, some of the selective methods such as DA does not account for the temporal coherence. While all three properties play an important role when removing noise yet there is no method that uses them all.

In our study, we present two noise characteristics in varying degress in different regions. The effectiveness of noise reduction for change detection methods is closely related to the susceptibility of these methods to these characteristics. First, *unbiased noise* of relatively small amplitude exists as a component of each observation due to variations in atmospheric conditions or instrument imprecision. This noise causes neighboring observations to be arbitrarily different from each other due

[†] These authors contributed equally to this work.
*University of Minnesota, `<chen,mithal,vangala,ivan,sboriah,kumar>@cs.umn.edu`.

to phenomena other than vegetation growth, where no land cover change has occurred. Second, the presence of relatively large, positively or negatively *biased noise* produces anomalous observations which do not follow the phenological trend of the time series. Often these observations are annotated with a low quality flag (QA) but sometimes may not be recorded accurately in the QA annotation.

Change Detection methods are impacted by both biased and unbiased noise in the data. A naive algorithm using observation-wise comparisons between the same months in two years will be severely impacted by biased noise and raise many false alarms. Therefore, most algorithms, including those used in this study, consider a more robust statistic like the average over an entire year for change detection. The Manhattan Delta [4] and Yearly Delta algorithms [2] are impacted by biased noise as it can increase the distance considerably and give a false appearance of change in EVI. The Yearly Delta algorithm is robust to unbiased noise in the data as averaging of it tends to be approximately zero. However, the Manhattan Delta algorithm is additively impacted by the unbiased noise.

In this paper we have shown that the interrelations between noise characteristics endemic to differing data regions, change detection methods, and noise reduction methods. We have provided contrasts between selective and non-selective imputation methods and their effects on biased and unbiased noise characteristics. We conclude that less conservative, non-selective noise reduction methods generally follow more conservative, selective methods to improve results. Conversely, we conclude that non-selective methods tend to perform poorly in the presence of positively or negatively biased noise. Depending on the susceptibility of the change detection method to each of these noise characteristics, either smoothing or outlier detection may not be necessary.

For an extended version of this study which includes a detailed discussion on noise reduction algorithms, noise characteristics of vegetation index data, as well as a comprehensive experimental evaluation, we refer the reader to [4].

Acknowledgements

The research described here is supported by the Planetary Skin Institute, NSF Grant IIS-0713227, NSF Grant IIS-0905581, NASA grant NNX09AL60G and the University of Minnesota MN Futures Program. Access to computing facilities was provided by the University of Minnesota Supercomputing Institute.

References

[1] P. S. Beck, C. Atzberger, K. A. Høgda, B. Johansen, and A. K. Skidmore. Improved monitoring of vegetation dynamics at very high latitudes: A new method using modis ndvi. *Remote Sensing of Environment*, 100(3):321 – 334, 2006.

[2] S. Boriah. *Time Series Change Detection: Algorithms for Land Cover Change*. PhD thesis, University of Minnesota, Twin Cities, April 2010.

[3] J. Chen, P. Jonsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh. A simple method for reconstructing a high-quality ndvi time-series data set based on the savitzky-golay filter. *Remote Sensing of Environment*, 91(3-4):332 – 344, 2004.

[4] X. Chen, V. Mithal, S. R. Vangala, I. Brugere, S. Boriah, and V. Kumar. A study of time series noise reduction techniques in the context of land cover change detection. Technical Report 11-016, University of Minnesota, Department of Computer Science and Engineering, 2011.

[5] J. Gu, X. Li, C. Huang, and G. S. Okin. A simplified data assimilation method for reconstructing time-series modis ndvi data. *Advances in Space Research*, 44(4):501 – 509, 2009.

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

[7] A. Savitzky and M. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 336(8):1627 – 1639, 1964.