A GLOBAL EVALUATION OF HARMONIC ANALYSIS OF TIME SERIES UNDER DISTINCT GAP CONDITIONS

Jie Zhou[1, 2], Li Jia[1, 3], Guangcheng Hu[1] and Massimo Menenti[4, 1]

1. *State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Beijing 100101, China*
2. *Graduate University of Chinese Academy of Sciences, Beijing 100049, China*
3. *Alterra, Wageningen University and Research Centre, Wageningen, The Netherlands*
4. *Delft University of Technology, Delft, The Netherlands*

ABSTRACT

Reconstruction of time series of satellite image data to obtain continuous, consistent and accurate data for downstream applications is playing a crucial role in remote sensing applications such as vegetation dynamics, land cover changes, land-atmosphere interactions and climate changes. Among the numerous methods and models developed to reconstruct time series of satellite observations in recent decades, Harmonic ANalysis of Time Series (HANTS) is one of the mostl widely used. Many studies based on time series reconstructed with HANTS documented the excellent performance of this method.

In the view of this study, the HANTS algorithm can be divided into two sub-processes, i.e. contaminated data identification and series reconstruction based on valid data. This study was dedicated to the evaluation of the performance of the latter sub-process. A simulated reference series dataset was constructed firstl, and then random gaps were introduced to these reference series. We built a look up table for distinct gap conditions by doing statistics on the deviation between the reference series and series reconstructed from gapped reference series. The look up table was used to evaluate the performance of a global NDVI time series dataset processed by HANTS.

The results show that the size of maximum gap (MGS), the number of loss (NL) and the number of gaps (NG) were significant factors in the reconstruction. When NDVI time series were rebuilt by HANTS, most of the region north than 40^0N and mountainous areas of earth show bad reconstruction performance, that is, the root mean square deviation (RMSD) could exceed 0.25. this can be attributed to the periodical snow cover in these regions.

INTRODUCTION

The long term satellite based earth observations have accumulated massive remotely measured time series since 1970s, which are playing a more and more irreplaceable role in earth science research ranging from ecology (1) to climatology (2) and agriculture (3). However, these time series datasets derived from visible and infrared channels are always contaminated by undetected cloud, snow or poor atmospheric conditions, which makes the temporal continuity, consistency and reliability of the time series of land surface observations less than optimal to meet the requirements of applications (4). In this context, researchers have developed numerous methods and models to address the problem of reconstructing gap free time series from irregularly spaced observations (5-9), among which the Harmonic Analysis of Time Series (HANTS) method is used very widely (7, 8, 10, 11).

The reconstruction algorithms aim at constructing a time series close to the “true” time series, while a reference time series, which can approximate the “true” time series, is needed to evaluate the accuracy. So the reference time series become the key of accuracy evaluation. Several works have evaluated the performance of several popular time series reconstruction algorithms in remote

sensing. Atzberger (12) proposed a range of quality indicators based on real time series which can be a relative evaluation method to illustrate the improvements of rebuilt time series compared to original ones but cannot quantify the accuracy of the reconstruction method. On the contrary, Hird(13) modelled NDVI time series, which were assumed to be reference series, and then added noise at different levels. The performance of filters was assessed as the absolute difference between the rebuilt time series and reference time series. Compared to Atzberger's method, Hird' method constructs a synthetic time series dataset and give the absolute accuracy of algorithms, which may be more appropriate for addressing the problem of accuracy evaluation for a specific algorithm like HANTS.

The objective of this study is to evaluate the performance of the HANTS algorithm to reconstruct global NDVI data under different gap conditions. In this paper, we give the theory of harmonic analysis and describe the two sub-processes of HANTS in the first place, and then the gap issue of remotely sensed temporal data is discussed. This is done by using a simulated reference time series NDVI dataset. Finally, the accuracy of a time series rebuilt from a MODIS global NDVI dataset (MOD13C1) is quantified based on the evaluation results.

THEORETICAL BASIS

Theory of harmonic analysis

Harmonic analysis is the representation of functions or signals as the superposition of basic waves, and the study of and generalization of the notions of Fourier series and Fourier transforms. When it is used in remote sensing to reconstruct the time series with unexpected biases or gaps, the basic formula is

$$\tilde{y}(t_j) = a_0 + \sum_{i=1}^{nf} [a_i \cos(2\pi f_i t_j) + b_i \sin(2\pi f_i t_j)] \qquad (1)$$

$$y(t_j) = \tilde{y}(t_j) + \varepsilon(t_j) \qquad (2)$$

$$(j=1,2,\cdots,N)$$

Here, y, $\tilde{y}$, ε and t_j is the original series, the reconstructed series, the error series and the time that y is obtained (observed), respectively. For example, time series of MODIS vegetation products, e.g. MOD13 and MOD15, are composited to a 4-day, 8-day, 16-day or monthly period, so the number of observations in a year, i.e. N, can be 92, 46, 23 or 12. The number of waves with different frequency is nf. a_i and b_i are coefficients of trigonometric components with frequency f_i. a_0 can be viewed as the coefficient at zero frequency which is the average of the series. All the coefficients of trigonometric components mentioned above are derived by minimizing $|\varepsilon|^2$ globally using least square method. The global minimum of $|\varepsilon|^2$ is referred as $|\varepsilon|^2_{\min,g}$. Next, the reconstructed series, i.e., $\tilde{y}$, can be determined.

HANTS is just a harmonic analysis based algorithm. The detail of HANTS has been given by Menenti et al. (1993) and Verhoef et al. (1996). In fact, there are two sub-processes involved in the HANTS algorithm. One is contaminated data identification process during which the outliers are flagged as invalid data during iteration. The other is the process of time series reconstruction based on remaining valid data. Not only HANTS, nearly all reconstruction algorithms face these two sub-processes. In this context, the accuracy of the reconstruction results for the same original series dependent on the accuracy of the two sub-processes. This study focused on the accuracy of the second sub-process and we assume that the first sub-process can identify the outliers successfully.

Gaps in temporal series of image data

Figure.1 gives an illustration of a NDVI time series at Qianyanzhou station in China (mainly grassland and farmland) in 2009 and the reconstructed time series. As one can see, the MODIS quality assessment (QA) information can identify some of the contaminated data (P1 and P2), but not all (P3). In this case, HANTS identifies all the anomalous observations, so QA information can contribute a lot but not replace reconstruction algorithms like HANTS to identify anomalous observations..

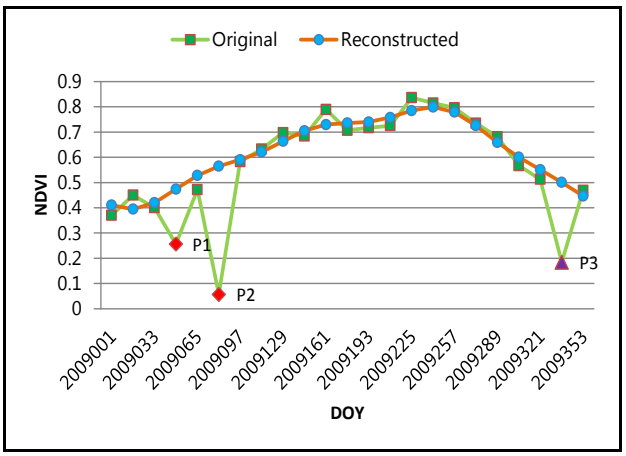

Figure 1:An original NDVI (MOD13Q1) series and reconstructed series by HANTS at Qianyanzhou station (26.7°N, 115.1°E) . P1, P2, P3 data points are biased data when assessed visually. But only P1 and P2 are flagged as "Pixel produced, but most probably cloudy" , P3 is flagged as "VI produced, good quality". When processed by HANTS, all P1, P2 and P3 are identified as contaminated data.

Generally speaking, the time interval of remotely sensed temporal series is constant, i.e. $t_{j+1}-t_j$ is a const. However, considering the effects of clouds, snow or ice cover, bad geometry and uncertainty existing in algorithms, only the valid values which are free of all the effects can be used to reconstruct series in normal harmonic analysis based algorithms. This is true for HANTS algorithm where the final time series is reconstructed from part of the raw time series data after a number of iterations. We make an assumption that the valid values are or most approximate to the "true" values. For simplicity, we refer the series without and with bad quality values as "GVTS" (abbreviation for "Global Valid Time Series") and "PVTS" (abbreviation for "Partly Valid Time Series"), respectively. When PVTS is used to reconstruct series, a zero weight will be assigned to bad quality observations, which means the bad quality values have no contribution to the resulting time series. In this context, the $|\varepsilon|^2_{\min,g}$ is, in fact, just partially minimized $|\varepsilon|^2_{\min,p}$. In contrast, the series reconstructed from GVTS reaches a definite global minimum $|\varepsilon|^2_{\min,g}$. Taking $|\varepsilon|^2_{\min,p}$ as an approximation of $|\varepsilon|^2_{\min,g}$ underlies the problem of using PVTS to rebuild time series. As one can expect, the distribution of gaps in a series can have significant impacts on the accuracy of the reconstruction as documented by ε in Eq. (2).

In the process of accuracy evaluation, all the attention has been concentrated on ε. More specifically, we want to investigate the dependence of ε on gap conditions for different types of time series. To this end, and considering above discussion, one of the most immediate means is to compare the time series rebuilt from PVTS with true time series in the same pixels. The true time series can be approximated by series rebuilt from GVTS following the assumption mentioned above. In practice it is very difficult for a remotely sensed time series to be entirely free of biases or gaps when taking into account the fact that 50 to 60% of the global land surface is covered by clouds every day. So it is reasonable to design a time series generator to simulate GVTS based on the features observed in real time series of variables like NDVI and FAPAR derived from satellite observations. Next, gaps are introduced in the simulated GVTS according to different configurations to simulate a PVTS. Then the difference, i.e., ε between rebuilt time series from GVTS and PVTS is used to evaluate performance of harmonic analysis statistically. Specifically, the Root Mean Square Deviation (RMSD) value of ε is used to describe the variance of ε. In addition, RMSD is normalized to the average of GVTS, which is the Coefficient of Variation of the RMSD, CV (RMSD).

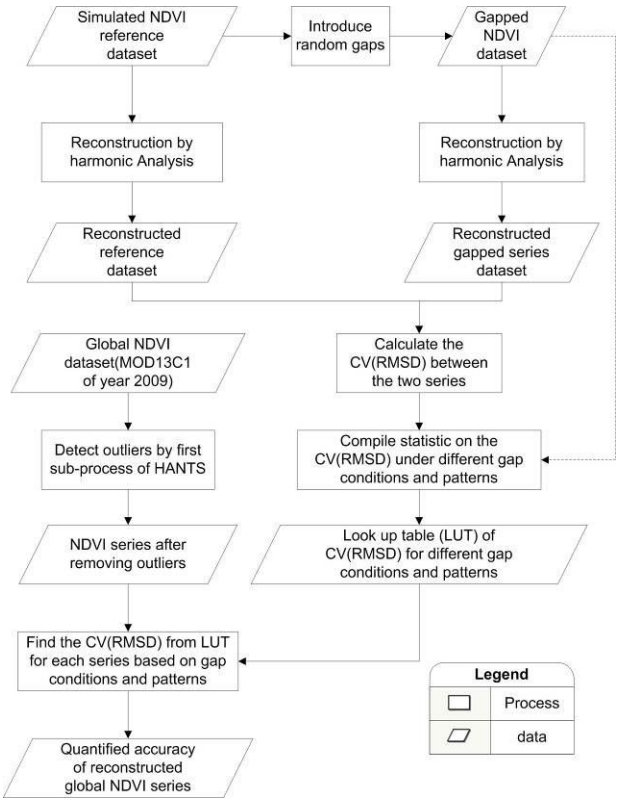

Figure 2: Flow diagram of this study.

METHODS

The flow diagram to evaluate the global accuracy of HANTS is given in Figure.2. The work could be divided into two parts: firstly, an accuracy Look Up Table (LUT) was constructed based on the simulated reference NDVI series and gap configurations. Then the LUT was used to quantify the global accuracy of HANTS.

Design of the time series generator

The phenological signatures of major land cover type can be parameterized using the average, variance, phase and primary frequency (the number of peaks in one year period). Then, such phenological signatures can be classified into four distinct patterns: evergreen forests, grasslands/scrublands, double-cropping agricultural land, and barren land/desert. The main characteristics and corresponding abbreviations for each pattern are given in Table 1.

Without loss of generality, the NDVI signals are simulated as:

$$y_{GVTS} = A_0 + \sum_{i=1}^{4} A_i \cos(\frac{2\pi t}{T_i} + \varphi_i) + e \qquad (3)$$

where

$$\begin{aligned} & e \sim N(0, 0.01 \times A_0^2) \\ & (T_1, T_2, T_3, T_4) = (360, 180, 120, 90) \\ & A_i \sim U(w_i^L A_0, w_i^u A_0) \\ & \varphi_i \sim U(0, 2\pi) \\ & A_0 \sim U(B_L, B_U) \end{aligned} \quad .$$

Here N and U are normal distribution and uniform distribution, respectively. The parameters e, A_i, φ_i and A_0 have random values with the distribution given in Eq. (3) during the generation of the synthetic GVTS dataset. This simulated time series is a combination of several low frequency harmonic components and a white noise. The value of parameters in Eq. (3) for different shape patterns are listed in Table 2. In this context, we assume that numerous simulated GVTS of NDVI can fully represent the synoptic character of true global NDVI.

Simulate the series under different gap conditions

As mentioned above, the objective of the study is to evaluate the impacts on the performance of harmonic analysis under different gap configurations. The problem is how to design the gap distribution quantitatively, i.e., how to simulate PVTS under controlled conditions.

In this study, the size of maximum gap (MGS), the number of loss (NL) and the number of gaps (NG) are used to parameterize the gap distributions. Specifically, gaps are introduced in simulated GVTS to generate corresponding PVTS by varying these parameters. The number of samples is 46 in one simulated series in this study. Also in this study, five low frequency harmonic components (including the zero frequency) are involved in reconstructing series, so the minimum

number of the remaining valid values should be no less than 14, i.e. compatible with the nine parameters describing the curve and with five being the degree of overdeterminedness.

Table 1: Characteristics and abbreviation for each pattern.

Pattern	Characteristics	Abbreviation
evergreen forests	**H**igh annual average value, **L**ow seasonal variance	H-L
grasslands/scrublands	**M**oderate annual average, **S**ingle seasonal variance	M-S
double-cropping agricultural land	**M**oderate annual average, **B**i-seasonal variance	M-B
barren land/desert	**L**ow annual average, **L**ow seasonal variance	H-L

Table 2: Value of Parameters in (3) for different patterns.

Pattern	w_1^L, w_1^u	w_2^L, w_2^u	w_3^L, w_3^u	w_4^L, w_4^u	B_L, B_U	Coefficient of variation	Mean
H-L*	0.04,0.06	0.04,0.06	0.01,0.02	0.01,0.02	0.60,0.80	0.055	0.70
M-S	0.40,0.60	0.00,0.30	0.00,0.15	0.00,0.05	0.30,0.60	0.370	0.45
M-B	0.00,0.20	0.20,0.30	0.00,0.15	0.00,0.05	0.30,0.60	0.260	0.45
H-L	0.08,0.10	0.06,0.08	0.04,0.06	0.02,0.04	0.00,0.30	0.110	0.20

To allow a proper statistical assessment of performance, a GVTS dataset for each pattern, which contains 100,000 time series simulated by the time series generator, is built. Next, gaps with random position and random size are introduced in each time series, from which the corresponding PVTS are simulated. For each PVTS, MGS, NL and NG can be counted. Then for each shape pattern, a new dataset containing 100,000 records is derived. The fields of each record contain CV(RMSD) of ε , MGS, NL and NG.

RESULTS

The accuracy of harmonic analysis under distinct gap condition

In this study, the number of samples in a time series is 46 and the simulated MGS and NL of each time series are in the range from 1 to 31. The average and coefficient of variation of four patterns are given in Table 2. The mean CV(RSMD) corresponding to different MGS levels and different patterns are given in Figure 3 (left panel), and the results for different NL are given in Figure 3 (middle panel). The CV (RMSD) quantifies the deviation of the rebuilt time series from the reference series and NDVI value mainly ranges from 0 to 1.0, so 0.05 can be set as a threshold for CV(RMSD) to indicate whether the rebuilt series are acceptable. The larger CV(RMSD) is, the worse the accuracy of the reconstruction.

Considering the overall performance response to MGS, as one can expect, the CV(RMSD) shows an increasing trend as MGS increases for each pattern. The M-S pattern has the largest rate of increase followed by the profile of M-B, L-L and H-L patterns. When the MGS is larger than 8, 9 and 12 for M-S, M-B and L-L patterns respectively, CV(RMSD) begins to exceed 0.05. CV(RMSD) of the H-L pattern, otherwise it has a low value (less than 0.05). For even larger MGS, the CV(RMSD) increases dramatically and rises to 0.4 for M-S pattern and 0.26 for M-B pattern. NL shows less sensitivity to CV(RMSD) compared to MGS. CV(RMSD) exceeds 0.05 for M-S and M-B pattern when MGS is larger than 13 and 20 respectively. CV(RMSD) typically is not larger than 0.06 under all NL conditions for H-L and L-L phenology. CV(RMSE) has an obvious downtrend as the NG increases (see Figure 3 (right panel)), since the larger NG indicates a more uniformed distribution of gaps in the time series.

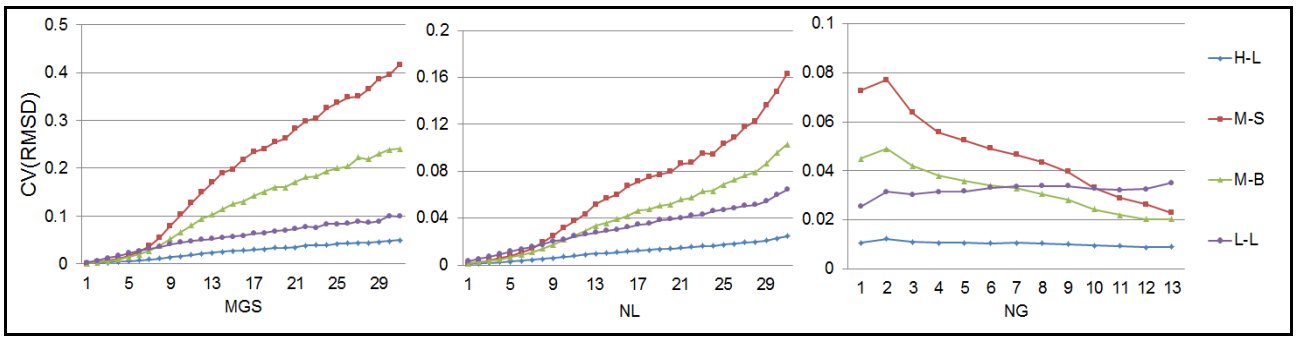

Figure 3: Mean of CV(RMSD) as a function of MGS (left panel), NL (middle panel) and NG (right panel) level and time series pattern. The total number of samples in each time series is 46.

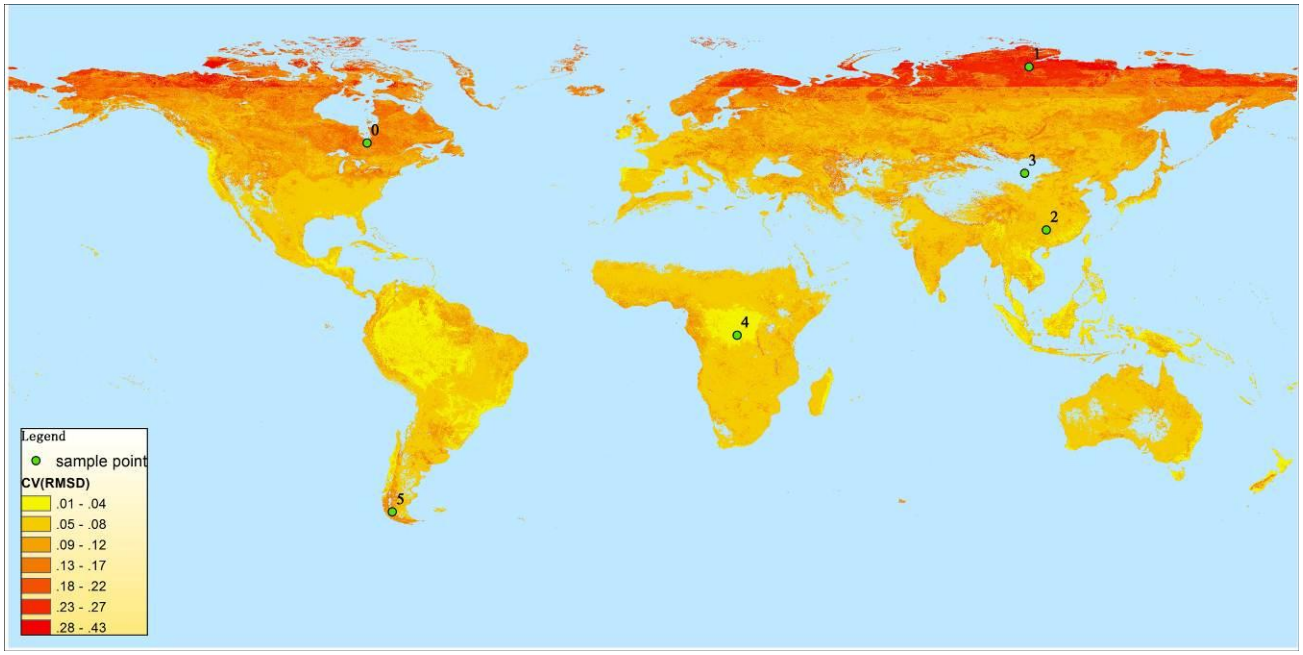

Figure 4: Quantified CV(RMSD) of 2009 global NDVI time series processed by HANTS.

Quantify the global accuracy of HANTS

The evaluation results derived above provide a Look Up Table (LUT) of accuracy in fact. This LUT is applied to assess the accuracy of the HANTS reconstruction of the Global MOD13C1 (Vegetation Indices 16-Day L3 Global 0.05Deg CMG) data in 2009. In detail, after the first sub-process of HANTS, the outliers have been removed and the MGS, NL and NG in the series can be calculated. Then the accuracy of a specific reconstruction result can be determined using the LUT for the corresponding deviation under same gap conditions.

Figure 4 gives the quantified results, which indicates the estimated CV(RMSD) of reconstruction results for individual pixels. At the same time, the reconstructed series of six random selected sample points has been given in Figure 5. As one can see, the vegetation area northern than 60^{o} (point0 in Figure 5) and mountainous areas of earth (point5 in Figure 5) have unrealistic reconstruction performance due to persistent snow cover in these regions at the beginning and end of the yearly series and a snow free period in the central segment. The CV(RMSD) values show a large spatial variation in the region between 40^{o}N and 60^{o}N (point-0 in Figure 5), which may be attributed to intermittent snow at two terminals of yearly series. Due to the cloud cover over the

Amazon, Central African (point4 in Figure 5) and Southeast Asian rainforest are distributed randomly all over the year other than only at the two terminals, the reconstruction results show a good performance. Although another cloud prone area, e.g. in India and Southwest China (point2 in Figure 5), is covered by cloud all over the summer, they get a better reconstruction performance than snow cover area because of the gaps mostly appear in the middle of series. Further analysis of impact of gap position on the reconstructed time series by HANTS will be addressed in the future study. The Gobi, desert, water snow/ice regions (point3 in Figure 5) all have no significant vegetation signal (NDVI less than 0.2 all over the year), thus no analysis were given in this study.

Considering that the main cause of outliers in NDVI time series are clouds and snow cover, and these two factors both appear in relatively persistent period in a year for a specific location on the earth, so it is reasonable to accept that the quantified accuracy in Figure 7 depicts the outline of the global normal accuracy condition although the results are just derived for a single year of NDVI time series data.

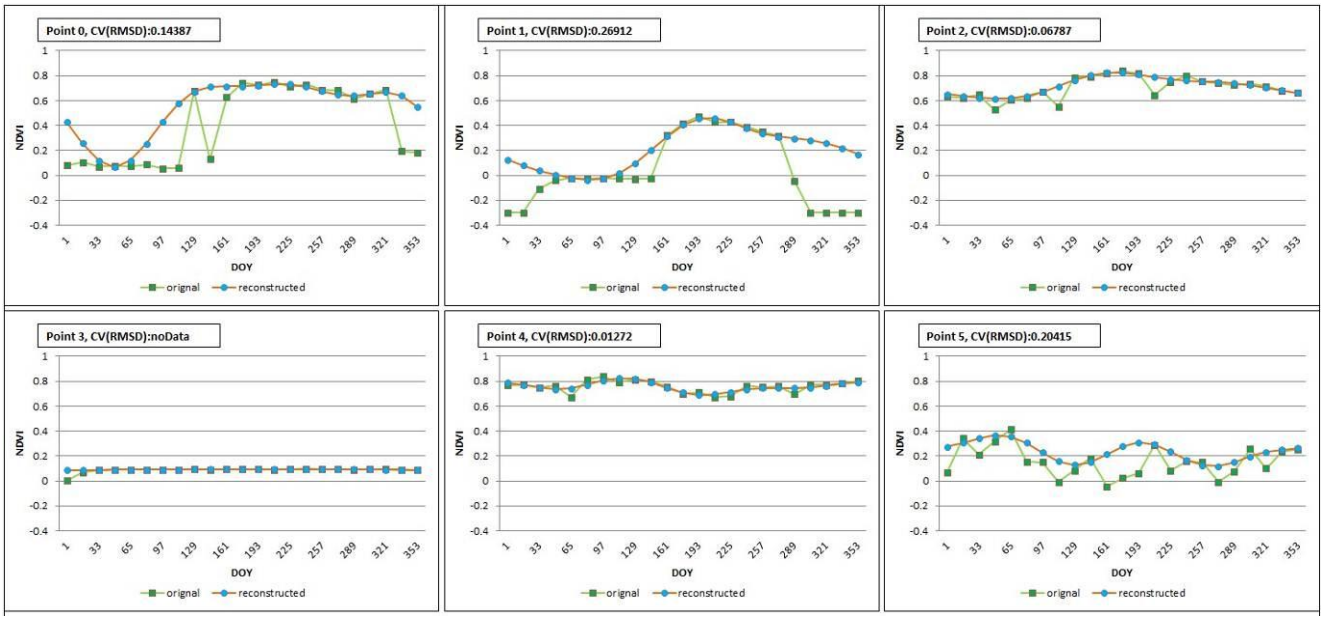

Figure 5: The original and reconstructed time series of sample points in Figure 4.

DISCUSSION AND CONCLUSIONS

This study proposed a novel method to quantify the accuracy of harmonic analysis based time series reconstruction by parameterizing the gaps in time series using the MGS, NL and NG parameters. The same gap configurations can exert distinct effects on time series with different patterns. M-S and M-B patterns are more sensitive to gaps, which imply that the NDVI time series variation can be a potential controller of reconstruction performance and more attention should be paid to areas covered by vegetation with high annual variations.

Following the methods and results in this paper, an accuracy LUT has been constructed and used to quantify the accuracy of a global NDVI time series processed by HANTS. The results advise that both periodical cloud and snow cover are main cause of bad HANTS performance, and the snow cover factor is more significant. HANTS algorithm may be more suitable for low-latitude area and new improvement should be developed to fulfill the time series reconstruction required over high latitude areas. Fortunately, the real gap conditions of global NDVI dataset are limited, if we can improve HANTS algorithm to reach an acceptable accuracy under gap conditions that give a bad reconstruction accuracy currently, a global gap-free NDVI dataset can be produced..

This evaluation method can also be extended to assess other reconstruction algorithms than HANTS. For the purpose of global representativeness, we simulated a NDVI time series dataset

with distinct patterns. We also highly recommend, however, to use a synthetic dataset (other than a simulated one in this study) fulfilling the global representativeness to evaluate the accuracy of reconstruction algorithms (12, 13).

ACKNOWLEDGEMENTS

This work was jointly supported by the "Chinese Academy of Sciences Visiting Professorships for Senior International Scientists", the EU-FP7 project CEOP-AEGIS (grant number 212921), the "CAS 100-Talent Project" and the project supported by KB VI program under framework of Dutch Ministry of Economic Affairs, Agriculture and Innovation.

REFERENCES

1. Sellers, P.J., et al., *Modeling the exchanges of energy, water, and carbon between continents and the atmosphere.* Science, 1997. **275**(5299): p. 502-509.
2. Myneni, R.B., et al., *Increased plant growth in the northern high latitudes from 1981 to 1991.* Nature, 1997. **386**(6626): p. 698-702.
3. Zhao, M.S. and S.W. Running, *Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 Through 2009.* Science, 2010. **329**(5994): p. 940-943.
4. Tucker, C.J., *Red and Photographic Infrared Linear Combinations for Monitoring Vegetation.* Remote Sensing of Environment, 1979. **8**(2): p. 127-150.
5. Holben, B.N., *Characteristics of Maximum-Value Composite Images from Temporal Avhrr Data.* International Journal of Remote Sensing, 1986. **7**(11): p. 1417-1434.
6. Viovy, N., O. Arino, and A.S. Belward, *The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series.* International Journal of Remote Sensing, 1992. **13**(8): p. 1585-1590.
7. Menenti, M., et al., *Mapping agroecological zones and time lag in vegetation growth by means of fourier analysis of time series of NDVI images.* Advances in Space Research, 1993. **13**(5): p. 233-237.
8. Verhoef, W., M. Menenti, and S. Azzali, *A colour composite of NOAA-AVHRR-NDVI based on time series analysis (1981-1992).* International Journal of Remote Sensing, 1996. **17**(2): p. 231-235.
9. Chen, J., et al., *A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter.* Remote Sensing of Environment, 2004. **91**(3–4): p. 332-344.
10. Roerink, G.J., M. Menenti, and W. Verhoef, *Reconstructing cloudfree NDVI composites using Fourier analysis of time series.* International Journal of Remote Sensing, 2000. **21**(9): p. 1911-1917.
11. M. Menenti, L.J., *Analysis of vegetation response to climate variability using extended time series of multispectral satellite images* in *Remote Sensing Optical Observations of Vegetation Properties*, F. Maselli, Menenti, M., Brivio, P.A. , Editor 2010. p. 131-163.
12. Atzberger, C. and P.H.C. Eilers, *Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements.* International Journal of Remote Sensing, 2011. **32**(13): p. 3689-3709.
13. Hird, J.N. and G.J. McDermid, *Noise reduction of NDVI time series: An empirical comparison of selected techniques.* Remote Sensing of Environment, 2009. **113**(1): p. 248-258.