BioOne COMPLETE

Reconstruction of a Cloud-free Vegetation Index Time Series for the Tibetan Plateau

Authors: Jun, Wen, Zhongbo, Su, and Yaoming, Ma

Source: Mountain Research and Development, 24(4) : 348-353

Published By: International Mountain Society

URL: https://doi.org/10.1659/0276-4741(2004)024[0348:ROACVI]2.0.CO;2

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://complete.bioone.org/journals/Mountain-Research-and-Development on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

Mountain Research and Development Vol 24 No 4 Nov 2004: 348–353

Wen Jun, Su Zhongbo, and Ma Yaoming

Reconstruction of a Cloud-free Vegetation Index Time Series for the Tibetan Plateau

In this investigation, an advanced Harmonic Analysis of Time Series (HANTS) technique has been introduced in order to remove cloud contamination from the satellite-based Normalized Difference Vegetation Index (NDVI) over the Tibetan Plateau. Due to high elevation, Asian monsoon, and the prevailing local convective weather, the probability of cloud occurrence in the daytime is very high over the Tibetan Plateau. It is therefore very difficult to directly obtain cloud-free vegetation index images using satellite optical remote sensing data over the entire Tibetan Plateau. Applying the HANTS algorithm made it possible to successfully replace the cloud contamination in NDVI images by the HANTS synthesized results, not only in temporal variation but also in regional distribution. The reconstructed NDVI images are cloud-free and more reliable; the HANTS-generated parameters are also potentially useful in understanding vegetation evolution and phenological characteristics over the mountainous regions of the Tibetan Plateau.

Keywords: *Vegetation; remote sensing; HANTS; time series; Tibetan Plateau; China.*

Peer reviewed: *May 2004* ***Accepted:*** *July 2004*

Introduction

Satellite remote sensing data have been successfully applied to monitoring and studies of land surface vegetation and land cover change (Zomer et al 2001; Bewket 2002). However, cloud contamination in satellite images seriously hampers their applicability. The Normalized Difference Vegetation Index (NDVI) is a preliminary parameter directly derived from satellite optical remote sensing data that has been widely used to characterize land surface vegetation and relevant applications (Valor and Caselles 1996; Maselli and Rembold 2001). Many vegetation variables can also be parameterized using NDVI; for instance, vegetation fractional coverage and leaf area index (Su 2000). Even certain physical parameters have to be parameterized using the vegetation index, by up-scaling ground point observation to regional distribution using remote sensing data. For example, soil heat flux can be parameterized with the vegetation index in land surface energy flux estimates (Su et al 1999). The vegetation index has also been widely and successfully used in regional and temporal vegetation classification and drought monitoring (Peters et al 2002). Since the NDVI is derived from satellite optical remote sensing data, cloud contamination not only significantly influences the precision of its application but also makes it unfeasible to obtain completely cloud-free NDVI images for a limited period over a large area or in mountainous regions.

The Tibetan Plateau is characterized by high above sea level elevation and local convective weather. Vegetation thrives mainly in summer. It is almost impossible to directly derive regional cloud-free vegetation index images using optical remote sensing data owing to weather characteristics. Even if NDVI values are rearranged using a maximum value-composing algorithm within every 10 days, cloud contamination still occurs in the NDVI images over the Tibetan mountainous regions. However, cloud-free images are essential for understanding the regional vegetation distribution and temporal evolution over the entire Tibetan Plateau. Therefore, an algorithm for removal of cloud contamination from NDVI images is necessary for vegetation study and associated applications.

A Harmonic Analysis of Time Series (HANTS) algorithm for time series dataset analysis has been developed (Verhoef et al 1996) and successfully applied in vegetation index analyses on the European continent (Roerink et al 2000). The object of the present investigation is to explore the HANTS's potential for cloud removal in NDVI images over the mountainous regions of the Tibetan Plateau; the cloud-free NDVI images will be reconstructed. The characteristics of vegetation evolution and phenological characteristics will also be outlined using the HANTS-generated parameters over the Tibetan Plateau.

Algorithm

In practical cases, any natural variation can be separated into many simple harmonic waves defined by different amplitudes, frequencies and phases. These amplitudes, frequencies and phases can be obtained by a so-called Fourier Transformation, by fitting the oscillation curve using observational samples in a finite discrete time series. The first step is to transfer observational samples from time domain to frequency domain; then, a new time series can be reconstructed using a reversed Fourier Transformation or synthesis, which can generate values of the variable at any desired time. A Fast Fourier Transformation (FFT) is often used in the Fourier Transformation process in order to save computational time.

However, the observational samples may sometimes incorrectly represent the true value of land surface geophysical parameters. For example, a satellite-

Downloaded From: https://complete.bioone.org/journals/Mountain-Research-and-Development on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

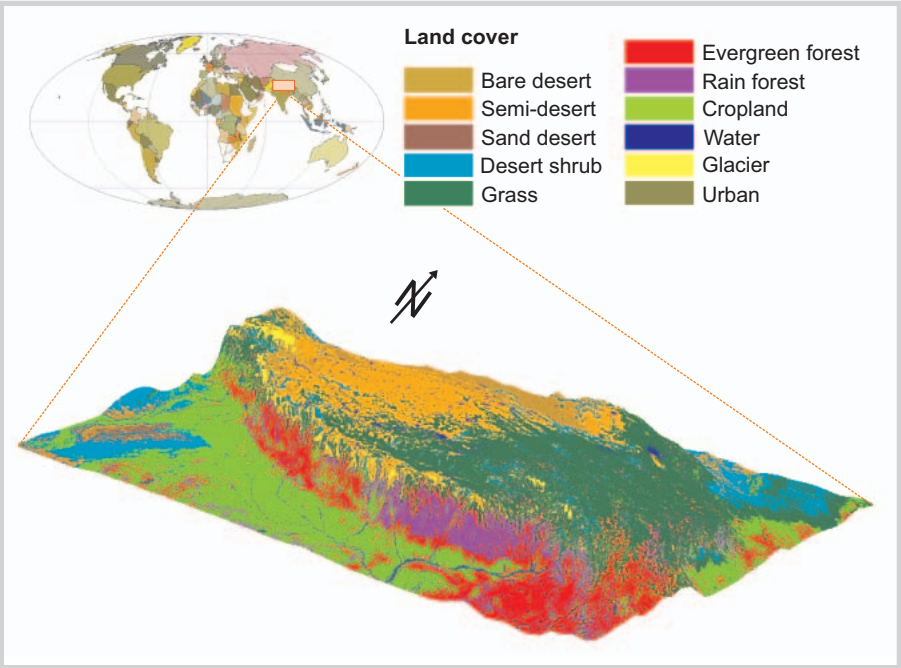


FIGURE 1 Map of the land cover types combined with elevation over the Tibetan Plateau. The western plateau is the arid Gobi Desert region, the southeastern plateau regions are pasturelands, and the temperate forests extend along the southern mountainous regions. Snow cover and glaciers are mainly confined to the Himalayan mountain regions.

based vegetation index might be incorrect if cloud contamination occurs in the satellite viewing field. If the new time series is reconstructed using the FFT algorithm based upon all observational samples, the reconstructed NDVI time series will misinterpret the temporal characteristics of land surface vegetation. Thus, the invalid observational samples should be eliminated or rejected in the curve fitting process. The Harmonic Analysis of Time Series (HANTS) algorithm was designed to satisfy the above requirements (Verhoef et al 1996). The invalid data can be removed by associating zero weights to them; the original curve-fitting problem thereby becomes a weighted least squares curve fitting process. A constant and a number of harmonic sines and cosines are chosen as the basic functions for the weighted least squares fitting.

In the HANTS algorithm, zero weights are assigned to outliers; the curve fitting is only based on the valid observational samples, and the samples no longer need to be equidistant in time. In comparing the observational samples to the curve values, if a value in the observational sample is obviously lower than that of the curve values, which are qualified by the pre-assigned errors, it should be eliminated because this sample is not the true value of the sample. In this case, a new curve is computed on the remaining samples, and the procedure is repeated. This iteration eventually leads to a smooth curve that approaches the upper envelope over the data points. By this procedure, the invalid observational samples (outliers) are removed. The computed amplitudes and phases are much more reliable than those based directly on a standard FFT algorithm.

In order to obtain reliable amplitudes and phases, 5 controlling parameters need to be specified in advance to run the HANTS algorithm:

Number of frequencies (NOF): The number of frequencies determines how many frequencies need to be used in reconstructing the new time series. A small number of frequencies will lead to a smooth fitting curve but give a less detailed description of the time series, while a large number of frequencies gives a more accurately fitting curve. The frequencies consist of the zero frequency, which determines the average value of the valid observational samples, the base frequency (period equal to the length of the time series), and a few multiples of the base frequency (overtones). Annual (12 months), semi-annual (6 months) and seasonal (3 months) frequencies are used in this investigation.

HiLo flag: The direction of the values that should be rejected as invalid values in the curve fitting process. For instance, low NDVI values should be rejected in processing time series datasets, because cloud contamination always corresponds to low or negative NDVI values. However, high values should be obviated in reconstructing surface reflectance time series, because cloud contamination leads to very high (apparent) surface reflectance.

Invalid Data Reject Threshold (IDRT): In the current software, a Valid Data Range can be specified in order to flag data as being invalid beforehand. If, for instance, the values 0 and 255 mark sea pixels and

Downloaded From: https://complete.bioone.org/journals/Mountain-Research-and-Development on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

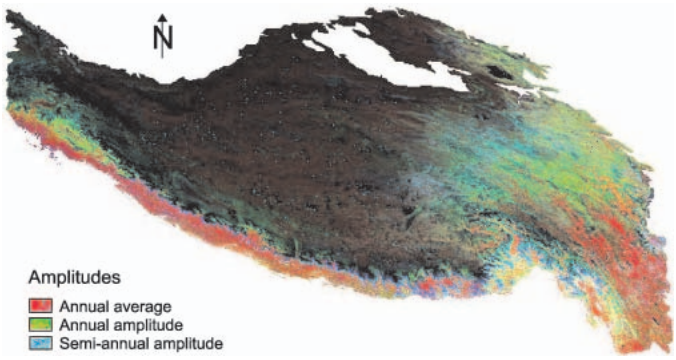


FIGURE 2 Color composite of the HANTS-generated NDVI annual average (Red), annual amplitude (Green) and semi-annual amplitude (Blue) over the Tibetan Plateau. The annual and semi-annual amplitudes are negligible in the western area, but the annual average is low in this area, which causes it to appear dark red. The annual amplitude is dominant in the southwest area; this area appears green. The annual average is dominant in the easternmost region and on southern slopes in Himalayan mountain regions; these regions are shown in red.

clouds detected by some algorithm, one could specify a Valid Data Range of 1-254 in order to indicate potentially valid samples.

Fitting Error Tolerance (FET): The absolute deviation in the chosen direction that is still acceptable in the curve fitting process. It determines when the curve fitting will stop. A large FET will lead to only a few invalid values being removed and the iteration soon coming to a halt, but the fitting might be based on many samples, including some invalid samples, which would yield unreliable results. However, small FET will lead to the rejection of too many observational samples, since there will be fewer observations used in the fitting, and this result will not be reliable either. In NDVI units, 0.05 is assigned as the FET in this investigation.

Degree of Overdeterminedness (DOD): HANTS describes the new time series with several amplitudes and phases. Since zero frequency has only one amplitude, the minimum number of remaining samples that describe the reconstructed curve is equal to $2 \times NFO-1$. The number of valid observational samples must always be equal to or greater than this minimum number. DOD defines how many observational samples are still available in addition to the minimum number of observational samples in the curve fitting process. DOD is set to 13 in this investigation.

As the HANTS attaches zero weight to invalid observational samples in the curve fitting process, it will involve several matrix multiplications and a matrix inversion; it will require more computational time than the FFT algorithm in processing time series datasets. There are also several parameters for which suitable values need to be found experimentally. However, the specified parameters are easy to assign, the invalid samples are rejected in the least squares fitting process, and the cloud-contaminated NDVI values are removed from the original NDVI image, which makes the reconstructed time series more reliable and reasonable; nor is it necessary that the observational samples be equidistant. Therefore, it is worthwhile adopting the HANTS algorithm in time series analysis. The HANTS software is downloadable free of charge from the NLR's web site (http://remotesensing@nlr.nl).

Cloud removal and its applications

The dominant land cover types over the Tibetan Plateau are grassland and semi desert; evergreen forest and rain forest are distributed along the southern mountain slopes. A map of the geographical location and land cover types on the Tibetan Plateau, combined with elevation, is presented in Figure 1. Because of the geographical location and local convective weather characteristics, it is difficult to obtain regional cloud-free vegetation index images from optical remote sensing over the mountainous regions of the Plateau. A so-called maximum value compositing technique (MVC) applied to NOAA/AVHRR NDVI data can provide fairly reliable NDVI images within every 10 days, but cloud contamination still occurs in the composite NDVI images, especially over mountainous regions. In order to explore the potential of the HANTS in cloud contamination removal over the Tibetan Plateau, 10 days of periodic NDVI images from February 1995 to February 1996 were used in this

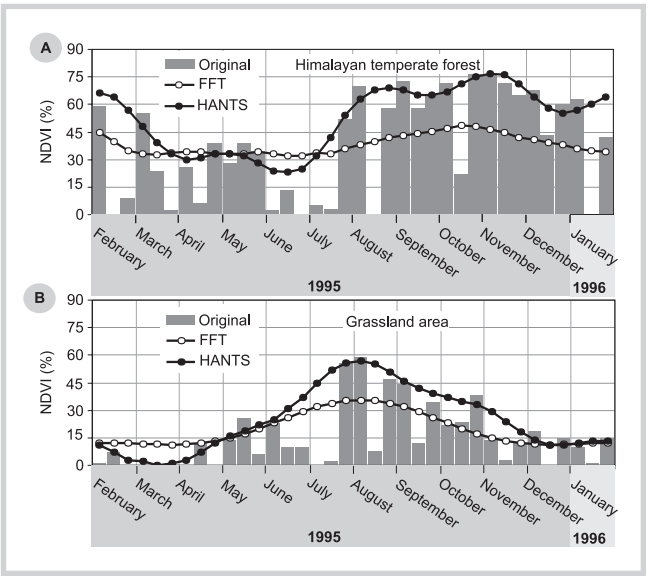


FIGURE 3 Comparison of the original and reconstructed NDVI time series. (A) Himalayan temperate forest area; (B) Grassland area. This illustration demonstrates the reliability of the reconstructed NDVI time series.

Downloaded From: https://complete.bioone.org/journals/Mountain-Research-and-Development on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

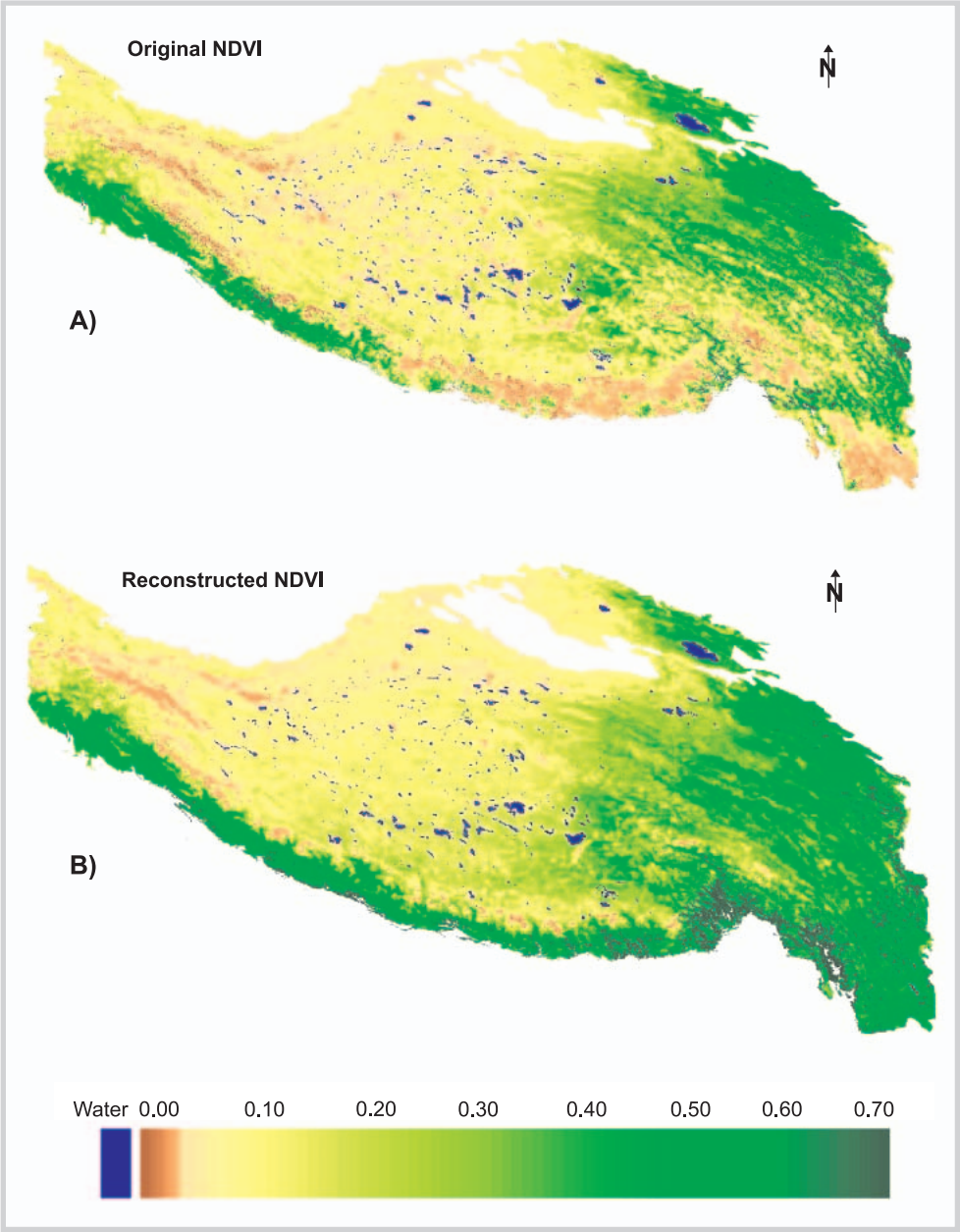


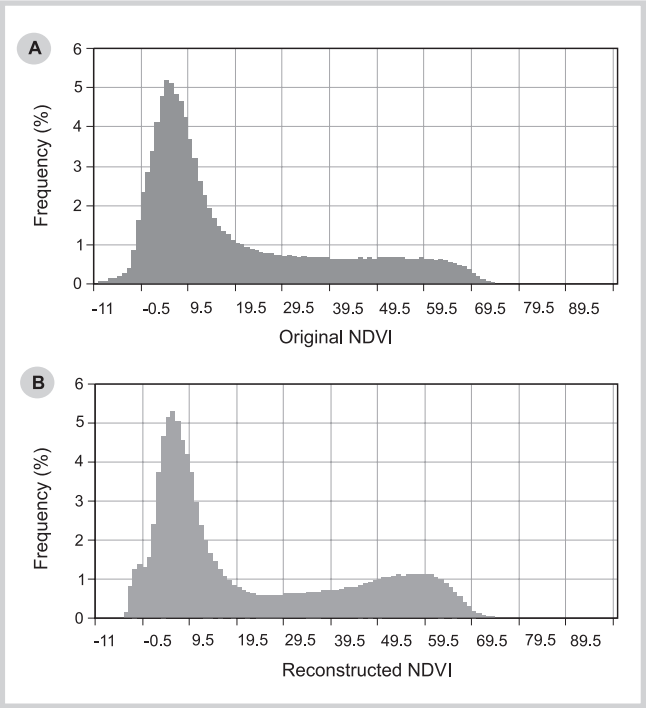
FIGURE 4 NDVI images over the Tibetan Plateau in the 3rd 10-day period of August 1995. (A) the original image; (B) the HANTS reconstructed image. The success of cloud removal is evident on the southern Himalayan mountain slopes.

investigation. The HANTS algorithm can generate amplitudes and phases of different periodic vegetation cycles. A color composite image of the annual average NDVI in red, and annual and semi-annual amplitudes in green and blue, are mapped in Figure 2 in order to compare their magnitudes. In order to focus on the Tibetan Plateau area and test the success of the HANTS in cloud contamination removal over the Tibetan Plateau and Himalayan southern mountainous regions, only areas above 2000 m above sea level and the Himalayan southern mountainous regions are shown on this and subsequent plates.

Due to the different vegetation types, climate, and phenology or vegetation cycles, the HANTS-generated NDVI amplitudes composite image displays different colors over the Tibetan Plateau area. The western area is in dark red, which means the average NDVI is dominant among 3 amplitudes, while the annual and semi-annual amplitudes are negligible. However, the vegetation is very poor over this area, the average annual NDVI value is still very small, and it causes this area to be dark red. The southeastern plateau and the southern slopes of Himalayan mountain regions are magenta or red, which means the vegetation, which is evergreen, grows very well

Downloaded From: https://complete.bioone.org/journals/Mountain-Research-and-Development on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

FIGURE 5 Histogram (0.01 bin-box) of the NDVI image over the Tibetan Plateau in the 2nd 10-day period of August 1995. (A) the original NDVI; (B) the reconstructed NDVI. The histograms show that some of the cloud-contaminated NDVI values in the original dataset appear as large values in the histogram of the reconstructed NDVI image.



there throughout the year. The southern regions, among others, show green, which corresponds to a strong annual but weak semi-annual vegetation cycle; this is a typical characteristic of the plateau grasslands. The southwestern and eastern plateau areas are yellow or sallow, which means that the cycle of vegetation has only annual variation but weak semi-annual variation in this area. Some areas are in cyan, meaning that the vegetation is characterized not only by annual but also by semi-annual variation. The northern slopes of the Himalayan mountain regions are light blue; this corresponds to a strong semi-annual vegetation cycle. It associates with the precipitation caused by the yearly arrival and withdrawal of the Asian monsoon. The above analysis demonstrates that the HANTS-generated amplitudes can characterize vegetation evolution and its phenological variations.

With the HANTS-generated amplitudes and phases, the cloud-free NDVI images over the entire Tibetan Plateau can be reconstructed in time series. Two reconstructed NDVI time series extracted over 2 typical plateau land surfaces to test the success of the HANTS algorithm in cloud contamination removal and results are presented in Figure 3. One site is located in the Himalayan mountain regions with a high fractional vegetation coverage during the entire year; the other site is located in the southeastern plateau with a clear annual vegetation cycle. The original NDVI values are also plotted in the figures for purposes of comparison. This shows that although the maximum values are picked up every 10 days, the original NDVI time series is still chaotic because of cloud contamination. It is not feasible to characterize temporal vegetation variation with these kinds of dataset, but the HANTS-reconstructed time series clearly shows yearly variation, since the cloud-contaminated values are replaced by the reconstructed values. In the original time series, there are extremely low NDVI values in summer, which is impossible since land surface vegetation cannot change at such a rapid rate within such a short period. If these values are used to describe vegetation character, temporal variation will be misinterpreted, whereas cloud-contaminated values are removed in the reconstructed time series and give reasonable and reliable results. Therefore, the HANTS has a powerful potential in time series analyses. For comparison, the FFT results computed using the same frequencies are also plotted in the same figures. The FFT algorithm uses all observational samples; the reconstructed time series underestimates the true vegetation index. Although the computational speed of the FFT is faster than that of the HANTS, it is not worthwhile adopting FFT in the time series analysis. More plots demonstrate that the HANTS-generated NDVI time series are reliable in most cases (figures are not presented). Only a few cases cannot deliver the expected results.

The original NDVI and the reconstructed NDVI images for the 3rd 10-day period in August 1995 are mapped in Figure 4. In the original NDVI image, there were very low values on the southern slope of the Himalayan mountain regions, in the southeastern area of the Tsangpo River basin, and in some other mountainous regions on the southeastern part of the Plateau. As the temperate forest extends throughout the southern slope of the Himalayan mountains, the Tsangpo River basin is also well vegetated in summer, while the original NDVI image does not correctly display the true vegetation situation. Since cloud occurrence probability is high in the mountainous regions, the original NDVI values take cloud contamination into account in time series analysis; some of the original NDVI values are measured at the top of the clouds. In the reconstructed time series, the cloud-contaminated NDVI values are replaced, and the Himalayan mountain regions and their southern slopes reveal large vegetation index values. However, extremely low NDVI values still occur in certain places in the reconstructed images along the Himalayan mountain regions and other places. These locations are demarcated as bare soil surface, permanent snow cover, or water surface, with reference to the land cover map.

The success of cloud removal in the reconstructed NDVI cloud-free image can also be testified by histogram illustrations, as shown in Figure 5. In the original NDVI images, clouds occur over the vegetated land surfaces in some locations. The cloud-contaminated NDVI only corresponds to the low NDVI values in the histogram, especially in the Himalayan mountain regions, although the true vegetation index might be

Downloaded From: https://complete.bioone.org/journals/Mountain-Research-and-Development on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use

very large. In the HANTS-reconstructed NDVI image, the cloud-contaminated NDVIs are replaced by the reconstructed cloud-free values, and this area will appear as large NDVI values in the histogram. There are still some low NDVI values in the histogram of the reconstructed image; these are contributions from snow cover, water surface, and the Gobi Desert.

The HANTS algorithm can also be applied to other parameters that have similar characteristics to those of NDVI in time series. For example, the cloud-free land surface evaporation fraction was reconstructed at spatial and temporal scales (Su et al 2001). In ecological assessments, Gross Primary Production (GPP) and Net Primary Production (NPP) can be estimated with the aid of satellite remote sensing data at temporal and spatial scales (Hicke et al 2002), but cannot be correctly evaluated over cloud-covered land surfaces. In this case, the HANTS algorithm has the potential to reconstruct cloud-free temporal GPP and NPP images and time series. This also applies to satellite-derived land surface reflectance, because the values of cloud-contaminated surface reflectance are much higher than those of the true surface reflectance.

Conclusion

In this study, the spatial and temporal vegetation index has been reconstructed with the HANTS technique over the entire Tibetan Plateau. This has improved understanding of spatial and temporal vegetation characteristics over the Tibetan Plateau. The results demonstrate that the HANTS algorithm can be successfully used for cloud contamination removal in NDVI images, not only in temporal variation but also in spatial distribution over the Tibetan Plateau. Only a few amplitudes and phases are required to reconstruct the temporal vegetation index in each location. The HANTS algorithm does not require equidistant temporal sampling. This makes it more convenient for processing of practical observational samples.

The color composite of the HANTS-generated NDVI periodic amplitudes gives the growing characteristics of land surface vegetation on the Tibetan Plateau, which is useful for understanding vegetation change and phenological variations. The applications can also potentially be extended to other geophysical parameters with periodic variations and similar characteristics to NDVI, especially in mountainous regions.

ACKNOWLEDGMENTS

This study was carried out under the auspices of the innovation projects of the Chinese Academy of Sciences (KZCX3-SW-329 and KZCX3-SW-339) and the Centurial Program sponsored by the Chinese Academy of Sciences. Preparation of this publication was supported, in part, by the Dutch Remote Sensing Board, and the Royal Netherlands Academy of Arts and Sciences, USDA ARS Hydrology, and Remote Sensing Lab. We wish to thank Dr W Verhoef and Mr GJ Roerink for technical support, and Mr Seth Passell for correcting the English. The authors are grateful to 2 anonymous reviewers for their constructive comments.

AUTHORS

Wen Jun and Ma Yaoming
Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu 730000, People's Republic of China.
wentiexin@yahoo.com (W.J.); ymma@ns.lzb.ac.cn (M.Y.)

Su Zhongbo
International Institute for Geo-Information Science and Earth Observation (ITC) Hengelosestraat 99, PO Box 6 7500, AA Enschede, The Netherlands.
b_su@itc.nl

REFERENCES

Bewket W. 2002. Land cover dynamics since the 1950s in Chemoga Watershed, Blue Nile Basin, Ethiopia. *Mountain Research and Development* 22:263–269.

Hicke JA, Asner GP, Randerson JT, Tucker C, Los S, Birdsey R, Jenkins JC, Field C, Holland E. 2002. Satellite-derived increases in net primary productivity across North America, 1982–1998. *Geophysical Research Letters* 29(10). Available at http://www.agu.org/pubs/crossref/2002.../2001GL013578.shtml; accessed on 7 July 2004.

Maselli F, Rembold F. 2001. Analysis of GAC NDVI data for cropland identification and yield forecasting in Mediterranean African Countries. *Photogrammetric Engineering & Remote Sensing* 67:593–602.

Peters AJ, Walter-Shea EA, Ji L, Vina A, Hayes M, Svoboda MD. 2002. Drought monitoring with NDVI-based standardized vegetation index. *Photogrammetric Engineering & Remote Sensing* 68:71–75.

Roerink JG, Menenti M, Verhoef W. 2000. Reconstructing cloud-free NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing* 21:1911–1917.

Su ZB. 2000. Remote sensing of land use and vegetation for mesoscale hydrological studies. *International Journal of Remote Sensing* 21:213–233.

Su ZB, Menenti M, Van Den Hurk BJJM. 1999. Large area case study with SEBAL: Europe with AVHRR data. *In*: Su ZB, Menenti M, editors. *Mesoscale Climate Hydrology: The Contribution of the New Observation.* Report USP-2. BCRS [Netherlands National Remote Sensing Board]:, pp 63–77.

Su ZB, Roerink GJ, Jacobs C. 2001. Development of a procedure to produce time series of evaporation fractions. *In*: Su Z and Jacobs C, editors. *Advanced Earth Observation – Land Surface Climate.* Publications of the Netherlands National Remote Sensing Board (BCRS):REPORT USP-2, 01–02, pp 184–186.

Valor E, Caselles V. 1996. Mapping land surface emissivity from NDVI: Application to European, Africa and South American area. *Remote Sensing of Environment* 57:167–184.

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

Zomer RJ, Ustin SL, Carpenter CC. 2001. Land cover change along tropical and subtropical riparian corridors within the Makalu Barun National Park and Conservation Area, Nepal. *Mountain Research and Development* 21:175–183.

Downloaded From: https://complete.bioone.org/journals/Mountain-Research-and-Development on 28 Mar 2024
Terms of Use: https://complete.bioone.org/terms-of-use