Suitability Analysis and Evaluation of GIMMS NDVI3g Product in Plateau Region

Jianyun Zhao¹, *, Yuanyuan Ding¹, and Jing Yang²

¹Department of Geologic Engineering, Qinghai University, Xining 810016, China
²Meteorological Station of Gangcha County, Qinghai Haibei 812300, China

* Corresponding author: zhaojianyun1981@163.com

Abstract. Quantitative Remote Sensing product have many challenges and uncertainties used in Tibet plateau, GIMMS NDVI3g is the longest NDVI time series products reflected the vegetation growing and changing at present. This paper takes Qinghai plateau as example, and using multi-source remote sensing data including SPOT, MODIS and Landsat to analyze its suitability and evaluate its accuracy. The results indicate that the value of GIMMS NDVI3g is bigger than others dataset, and there has no significant difference between the datasets according to the root mean square error, and the correlation coefficients are greater than 0.95. The vegetation phenology and changing characteristic reflected by the datasets is in line with each other. There have strong correlation between NDVI3g and Landsat NDVI, and the root mean square error is 0.055, and the deviation is in accord with the Gaussian normal distribution. So NDVI3g dataset has favorable suitability and reliable accuracy used in vegetation change detection of plateau region.

1. Introduction
Vegetation is an important component of ecosystem which plays a vital role in the processes of energy exchange and material cycle. It bridges the connection between soil, atmosphere and moisture. Vegetation coverage is the result of multiple actions of regional climate and environment, and its change is the important indicator of global environmental change, and has great significance to regional ecosystem evaluation[1]. Vegetation index realizes vegetation classification, growth status assessment and parameter inversion using the vegetation chlorophyll, moisture and dry matter’s spectrum characteristic and bands calculation of vegetation, which is a common method of remote sensing geological application. Normalized difference vegetation index (NDVI) is the most common used index which reflects the green degree, photosynthesis intensity, seasonal and annual variation of vegetation, and compensates the image change induced by illumination, topography or observation direction [2,3].

At present, there have AVHRR GIMMS NDVIg, AVHRR GIMMS NDVI3g, SPOT NDVI, MODIS NDVI, etc. NDVI product and the NDVI3g is the latest and longest product, which is composted by 15 day’s AVHRR data between 1982 and 2015. NDVI3g weather can reflect the vegetation change in plateau region actually and in line with other NDVI product should be test and validation [4]. Some scholars have studied the consistency of the NDVI product, and used them in the field of regional vegetation coverage change [5]. Some results indicated that the time-space variation tendency of vegetation phenology reflected by the GIMMS NDVI and MODIS NDVI is the same. Although some results also indicated that the vegetation coverage acquired by GIMMS NDVI had no seasonal features, still MODIS NDVI has higher precision in detecting the change of vegetation than GIMMS NDVI, and the vegetation phenology features reflected by GIMMS NDVI were not compatible with others because of its serious
quality problems, and GIMMS NDVI3g has the higher precision and robustness in the vegetation activity
detecting [6].

This paper aims at those uncertainty problems, taking the Qinghai plateau of China as example, using
the AVHRR GIMMS NDVI, SPOT NDVI, MODIS NDVI and Landsat data to analyze and evaluate the
suitability of the NDVI3g, and testing the precision of vegetation coverage change, and providing the
decision basis to the using of long time series NDVI product in plateau area.

2. Experimental Area And Data Source
Qinghai is located in the northeast of the Qinghai-Tibet plateau. The east longitude is 89°35′~103°04′,
the north latitude is 31°39′~39°19′. It is about 1200km in east and west, and 800km in north and south,
the area is about 72.12 thousand km². The fluctuation of west terrain is smaller than the east, and the
gemorphologic types are multitudinous and have big difference. Qinghai plateau is the Chinese water
tower, and the three river sources. The land resource type is abundant, and the main vegetation type is
coniferous forest, shrub, grassland, meadow and alpine vegetation (www.resdc.cn).

AVHRR GIMMS NDVI datasets include the first and third product of NOAA (National Oceanic and
Atmospheric Administration) AVHRR (Advanced Very High Resolution Radiometer) GIMMS (Global
Inventory Modeling and Mapping Studies) sensor. NDVI3g combines the AVHRR/2 and AVHRR/3
sensor of NOAA-7/9/11/14/16/17/18/19, and corrects the orbit displacement, atmospheric radiation and
geometric distortion, and its spatial resolution is 8km, temporal resolution is 15 days. MODIS (Moderate
Resolution Imaging Spectroradiometer) is the important sensor of American EOS (Earth Observation
System), and the spectral region is 0.4μm~14.4μm, and has 36 bands. The sensor covers the global in
1~2 days, and the spatial resolution of MOD13A2 NDVI V6 is 250m, 500m and 1km, and the projection
of the product is Sinusoidal Grid. Its temporal resolution is 16 days, and the time duration is between
2000 and 2015.

SPOT NDVI dataset is generated by SPOT VEGETATION sensor, and corrects the deviation of
atmospheric, radiation and geometric. It is a 10-day maximum synthetic data with a spatial resolution
of 1 km, and the time duration is between 1998 and 2007.

Landsat is an Earth observation satellite launched and operated by the United States in 1972. Its main
sensors include RBV (Return Beam Vidicon, Landsat1-3), MSS (Multispectral Scanner System Landsat
1-5), and ETM (Enhanced Thematic Mapper, Landsat6), ETM+, and OLI (Operational Land Imager).
The spatial resolution has been increased from the original 120m to 30m, where the resolution of the
ETM+ and OLI panchromatic bands is 15m.

Qinghai province administrative division data is from the Chinese Academy of Sciences Resources
and Environmental Science Data Center (http://resdc.cn). The above NDVI data sets were cropped and
projected using administrative division data to obtain the NDVI data set of the multi-temporal and spatial
scales in the study area. Albers Conical Equal Area projection is used for each data (Southern standard
parallel: 25°N; North standard: 47°N; Central meridian: 105°E).

3. Research Methods
The consistency of NDVI datasets is generally analyzed by the methods of static numerical values and
dynamic trends on the scales of pixels, regions, and vegetation types, which through the correlation
coefficient, regression equation, mean deviation, root mean square deviation, and histogram. because
the spatial resolution of GIMMS NDVI is 8km, and the spatial resolutions of SPOT NDVI and MODIS
NDVI are all 1km, we first resample the last two datasets to 8km resolution. At the same time,
considering that the complete overlap of the four NDVI datasets is only 6 years, this study only uses the
static numerical method for consistency analysis at the regional and pixel scales.

Calculate the average NDVI value of each pixel on the annual scale of the region, as follows:

$$\bar{NDVI}_i = \frac{\sum NDVI_j}{n}$$

(1)
In the formula: \( NDVI^T_i \) is the Annual mean value of each pixel; \( NDVI_j \) is the NDVI value for each pixel during the year; \( n \) is the number of NDVI values greater than zero during the year. Using \( NDVI^T_i \) obtained above calculates the maximum, minimum, and mean value for each year as follows:

\[
\begin{align*}
NDVI^T_{i, \text{Max}} &= \text{Max}NDVI^T_i \\
NDVI^T_{i, \text{Min}} &= \text{Min}NDVI^T_i \\
NDVI^T_{i, \text{Mean}} &= \text{Mean}NDVI^T_i
\end{align*}
\]  

Considering that the minimum value of the data set is generally small (close to zero), the paper just calculates the deviation and root mean square error of the maximum value and the average value of each data set. The method is as follows:

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (NDVI_i - \overline{NDVI})^2}{n}}
\]

In the formula: \( (NDVI_i - \overline{NDVI}) \) is the deviation value; \( \overline{NDVI} \) is the mean value. On the monthly scale of the region, calculate the average monthly NDVI for each pixel. The method is as follows:

\[
\overline{NDVI}^M_i = \frac{\sum_{j=1}^{j_{\text{year}}} NDVI_j}{k}
\]

In the formula: \( \overline{NDVI}^M_i \) is the Monthly mean value of each pixel. \( NDVI_j \) is the NDVI value for each year. \( k \) is the Statistics year.

Using the \( \overline{NDVI}^M_i \) obtained above, and considering that GIMMS NDVI is a 15-day composite data, SPOT NDVI is a 10-day composite data, and MODIS is a 16-day composite data, the maximum, minimum, and mean values on the monthly scale are calculated as follows:

\[
\begin{align*}
\overline{NDVI}^M_{i, \text{Max}} &= \text{Max}\overline{NDVI}^M_i \\
\overline{NDVI}^M_{i, \text{Min}} &= \text{Min}\overline{NDVI}^M_i \\
\overline{NDVI}^M_{i, \text{Mean}} &= \text{Mean}\overline{NDVI}^M_i
\end{align*}
\]  

In order to obtain the correlation characteristics of each data set, regression analysis is performed using the formula (6), based on the mean value in the monthly scale, the determination coefficient is calculated as formula (7):

\[
\overline{NDVI}^T_i = \beta_0 + \beta_1 \overline{NDVI}^M_i + \varepsilon_i
\]

\[

r^2 = \frac{\sum_{i=1}^{n} (\overline{NDVI}^T_i - \overline{NDVI}^T) (\overline{NDVI}^M_i - \overline{NDVI}^M))^2}{\left(\sum_{i=1}^{n} (\overline{NDVI}^T_i - \overline{NDVI}^T)^2 \sum_{i=1}^{n} (\overline{NDVI}^M_i - \overline{NDVI}^M)^2\right)}
\]

In the formula: \( \beta_1 \) is the regression coefficient; \( \varepsilon_i \) is the residual; \( \overline{NDVI}^T_i \) and \( \overline{NDVI}^M_i \) are the corresponding NDVI datasets.

4. Analysis Results

Using equations (1) and (2) to calculate the maximum, minimum, and mean values of GIMMS NDVI_3g, NDVI_3g, SPOT NDVI and MODIS NDVI datasets in the regional scale for 2001-2006, and using formula (3) to calculate the deviation of the maximum value and the average value, the root mean square error is shown in Table 1.
Table 1  Bias and RMSE of Regional Yearly Dimension of the NDVI Data Set

| Dataset  | Deviation | 2001  | 2002  | 2003  | 2004  | 2005  | 2006  | RMSE  |
|----------|-----------|-------|-------|-------|-------|-------|-------|-------|
| NDVI₃g   | Max       | 0.034 | 0.032 | 0.015 | -0.002| 0.008 | 0.010 | 0.015 |
|          | Mean      | 0.022 | 0.021 | 0.017 | 0.016 | 0.017 | 0.015 | 0.014 |
| NDVI₉    | Max       | -0.050| -0.057| -0.045| -0.052| -0.037| -0.041| 0.037 |
|          | Mean      | -0.004| -0.003| -0.011| -0.013| -0.009| -0.009| 0.008 |
| NDVIₛ    | Max       | -0.042| -0.036| -0.022| -0.006| -0.019| -0.026| 0.022 |
|          | Mean      | -0.022| -0.023| -0.010| -0.011| -0.016| -0.014| 0.013 |
| NDVIₘ    | Max       | 0.059 | 0.061 | 0.052 | 0.059 | 0.048 | 0.057 | 0.046 |
|          | Mean      | 0.004 | 0.006 | 0.004 | 0.008 | 0.008 | 0.009 | 0.006 |

Using formulas (4) and (5) to calculate the maximum, minimum and mean values of the monthly scale of each data set, and deviations and root mean square errors of maximum and average values are shown in Table 2. Draw each data set maximum, mean curve distribution shown in Figure 1.

Table 2  Bias and RMSE of Regional Monthly Dimension of the NDVI Data Set

| Month | NDVI₃g | NDVI₉ | NDVIₛ | NDVIₘ |
|-------|--------|-------|-------|-------|
|       | Max    | Mean  | Max    | Mean  |
|       | Max    | Mean  | Max    | Mean  |
|       | Max    | Mean  | Max    | Mean  |
|       | Max    | Mean  | Max    | Mean  |
| 1     | 0.023  | 0.016 | -0.053 | -0.002 |
| 2     | 0.000  | 0.014 | -0.034 | -0.005 |
| 3     | 0.026  | 0.016 | -0.036 | -0.004 |
| 4     | 0.001  | 0.024 | -0.051 | -0.012 |
| 5     | 0.007  | 0.050 | -0.096 | -0.025 |
| 6     | 0.026  | 0.032 | -0.061 | -0.016 |
| 7     | -0.003 | 0.005 | -0.005 | -0.006 |
| 8     | -0.007 | 0.002 | -0.003 | -0.008 |
| 9     | 0.016  | 0.014 | -0.040 | -0.014 |
| 10    | 0.059  | -0.005| -0.088 | -0.022 |
| 11    | 0.005  | 0.004 | -0.064 | -0.010 |
| 12    | 0.019  | 0.015 | -0.039 | -0.003 |
| RMSE  | 0.023  | 0.021 | 0.055  | 0.013 |

Figure 1  Max and Average Value Distribution of the NDVI Data Set
Using equations (6) and (7), regression analysis was performed on the GIMMS NDVI3g and GIMMS NDVIg, SPOT NDVI, and MODIS NDVI data sets, and correlation coefficients were calculated. The results are shown in Table 3. At the pixel scale, using the monthly mean value of each pixel obtained from equation (4), the mean value of each pixel of each data set from 2001 to 2006 is calculated.

Table 3: Regression Analysis between NDVI3g and other NDVI Data Sets

| Response variable | Explanatory variables | Slope   | Intercept | Decisive factor |
|-------------------|-----------------------|---------|-----------|-----------------|
| NDVI3g            | NDVIg                 | 0.9528  | 0.0343    | 0.9520          |
|                   | NDVIg                 | 0.9212  | 0.0426    | 0.9675          |
|                   | NDVIm                 | 0.9436  | 0.0238    | 0.9626          |

The above diagram indicated that the GIMMS NDVI3g value is larger than other datasets in the regional inter-annual, monthly, and pixel scales. From the root mean square error of its maximum value and mean value, the numerical difference between the data sets is not significant. From the monthly scale, the vegetation phenology characteristics reflected by GIMMS NDVI3g, NDVIg, SPOT NDVI, and MODIS NDVI are also basically in the same. From the regression analysis of the regional monthly scale, the coefficient of determination between the datasets was above 0.95, indicating a strong correlation between datasets. From the distribution of the histograms of the mean of the four datasets, the consistency is also obviously. This shows that the vegetation changes reflected by the four datasets are consistent. The use of GIMMS NDVI3g data to monitor vegetation cover changes in the Qinghai Plateau is reliable.

5. Accuracy Verification

Due to the airborne calibration of the Landsat series of data, it has a higher spatial resolution than the NDVI dataset, which is best evaluation and verification data for low-resolution remote sensing product. In the study area, 12 sample areas of 24km×24km were selected. The cloudless Landsat data in July, August of the sample areas in 1996, 2000, 2004 and 2011 are collected and selected. After radiometric calibration and atmospheric correction of the collected Landsat data, calculate the NDVI of 432 samples in 12 sample areas and resample to 8km resolution.

According to the time obtained from each sampling area Landsat NDVI (NDVI_L), the NDVI value is obtained on the corresponding GIMMS NDVI3g nearest image, and the NDVI of two data sets is statistical analysis to calculate average NDVI for each sample. The result indicated that there was systematic bias in the 4th NDVI_L area in 1996, which may be covered by clouds; Mean deviations and RMSE were further calculated for mean values of the remaining areas. The RMSE values are all around 0.055 except for 2000, which indicates that the NDVI3g and NDVI_L values are close with each other. Statistical analysis was performed on each sample point in each sample area, and the statistical histograms of the data deviations of each year were plotted in Figure 2.

(a) Statistical Histogram of 1996
(b) Statistical Histogram of 2000
In order to further clarify whether the characteristics of the vegetation coverage changes expressed by the two data are consistent, we use the equations (6) and (7) to perform regression analysis on the obtained NDVI_{3g} and NDVI_L mean values, the result show as figure 3.

It can be seen from Figure 2 that the deviation distribution of NDVI_{3g} data and NDVI_L data conforms to the normal distribution rule, indicating that both data come from the same parent. From Figure 3 and the regression analysis result, it can be seen that NDVI_{3g} and NDVI_L are strongly related, and the overall trend of vegetation changes reflected by NDVI_{3g} and NDVI_L is consistent. Therefore, the GIMMS NDVI_{3g} data accuracy in the Qinghai Plateau is reliable.

6. Conclusion
The authenticity inspection of remote sensing products is one of the main scientific goals of quantitative remote sensing research and is the basis for safeguarding theoretical models, scientific application and accuracy of production applications. The Qinghai-Tibet Plateau at the third pole of the world, due to its unique climatic environment and complex landforms, there are large uncertainties and controversies in the application of various quantitative remote sensing products. The third-generation GIMMS NDVI is the longest released NDVI product in the time series. This paper makes use of the current NDVI product data commonly used in the world such as AVHRR GIMMS, SPOT, MODIS, etc. to conduct a comparatively comprehensive analysis, and using Landsat data to verify its accuracy. To a certain extent, the study promotes the product's application and research in Plateau Areas.
Acknowledgments
This study was supported by the natural science foundation of Qinghai science and technology agency of China (No. 2021-ZJ-743).

References
[1] Verma R, Dutta S. (2013) Vegetation Dynamics from Denoised NDVI Using Empirical Mode Decomposition. Journal of the Indian Society of Remote Sensing, 41(3):555-566.
[2] Verbyla D. (2015) Comparison of eMODIS and MOD/MYD13A2 NDVI products during 2012–2014 spring green-up periods in Alaska and northwest Canada. International Journal of Applied Earth Observations & Geoinformation, 36:83-86.
[3] Li Z, Guo X. (2013) A suitable NDVI product for monitoring spatiotemporal variations of LAI in semiarid mixed grassland. Canadian Journal of Remote Sensing, 38(6):683-694.
[4] Lunetta RS, Knight JF, Ediriwickrema J, et al. (2009) Land-cover change detection using multi-temporal MODIS NDVI data. Remote Sensing of Environment, 105(2):142-154.
[5] Walker JJ, De Beurs KM, Wynne RH. (2014) Dryland vegetation phenology across an elevation gradient in Arizona, USA, investigated with fused MODIS and Landsat data. Remote Sensing of Environment, 144:85-97.
[6] Zhang Y, Ye A. (2020) Spatial and temporal variations in vegetation coverage observed using AVHRR GIMMS and Terra MODIS data in the mainland of China [J]. International Journal of Remote Sensing, 41(11):4238-4268.