

Evaluation of Above Ground Biomass Estimation Accuracy for Alpine Meadow Based on MODIS Vegetation Indices

Bao-Ping MENG, Tian-Gang LIANG*, Jing GE, Jin-Long GAO, Jian-Peng YIN

State Key Laboratory of Grassland Agro-ecosystems, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China

mengbp09@lzu.edu.cn, Correspondence: tgliang@lzu.edu.cn, gej12@lzu.edu.cn, gaojl16@lzu.edu.cn, yinjp2013@lzu.edu.cn

Abstract—Animal husbandry is the main agricultural type over the Tibetan Plateau, above ground biomass (AGB) is very important to monitor the productivity for administration of grassland resources and grazing balance. The MODIS vegetation indices have been successfully used in numerous studies on grassland AGB estimation in the Tibetan Plateau area. However, there are considerable differences of AGB estimation models both in the form of the models and the accuracy of estimation. In this study, field measurements of AGB data at Sangke Town, Gansu Province, China in four years (2013–2016) and MODIS indices (NDVI and EVI) are combined to construct AGB estimation models of alpine meadow grassland. The field measured AGB are also used to evaluate feasibility of models developed for large scale in applying to small area. The results show that (1) the differences in biomass were relatively large among the 5 sample areas of alpine meadow grassland in the study area during 2013–2016, with the maximum and minimum biomass values of 3,963 kg DW/ha and 745.5 kg DW/ha, respectively, and mean value of 1,907.7 kg DW/ha; the mean of EVI value range (0.42-0.60) are slightly smaller than the NDVI's (0.59-0.75); (2) the optimum estimation model of grassland AGB in the study area is the exponential model based on MODIS EVI, with root mean square error of 656.6 kg DW/ha and relative estimation errors (REE) of 36.3%; (3) the estimation errors of grassland AGB models previously constructed at different spatial scales (the Tibetan Plateau, the Gannan Prefecture, and Xiahe County) are higher than those directly constructed based on the small area of this study by 9.5%–31.7%, with the increase of the modeling study area scales, the REE increasing as well. This study presents an improved monitoring algorithm of alpine natural grassland AGB estimation and provides a clear direction for future improvement of the grassland AGB estimation and grassland productivity from remote sensing technology.

1 Introduction

Grassland biome is supported by the basic principles of primary productivity, especially for various animals of the grassland [1]. The above-ground biomass (AGB) of grassland provides the basis for estimating the net primary productivity of grassland [2]. Therefore, it is very important to monitor the productivity for administration of grassland resources and grazing balance, i.e. the number of livestock units that grassland can feed for specific length of time, and the quality and quantity of AGB [3]. In particular, as the main grazing area, the AGB estimation is more prominent in the Tibet plateau of China [4-5].

Numerous studies about grassland AGB have indicated that there is an increasing prominent role for grassland ecosystems [1]. However, grasslands in Tibetan Plateau (TP) are facing soil erosion, salinization, desertification and degradation caused by human activities and climate change. Accurate measurements of grassland biomass and its

temporal and spatial variation are important for the utilization and protection of grassland resources. Grassland AGB can be estimated by using both traditional (or ground based) and remote sensing methods. With the development of science and technology, monitoring using remote sensing data has gradually replaced the traditional methods, for its large scale, frequency, low cost, and massive information [6]. The MODIS vegetation indices have been successfully used in numerous studies on grassland AGB in the TP. However, there are considerable differences of AGB estimation models both in the form of the models and accuracy of estimation. Previous studies stated a relatively large uncertainty in the evaluation of AGB and a lack of quantitative accuracy analysis on relatively large spatial scales. For alpine meadow grassland, the most suitable type of remote sensing estimation model is currently unknown and the applicability of these models still needs to be proved.

In this study, considering the factors discussed above, a region of alpine meadow grassland in the TP area is used to

conduct the following investigations: (1) to construct and analyze the AGB estimation models and their accuracies based on the MODIS NDVI and EVI in study area; (2) the data from the sample plots observed inside the study area are used to validate the applicability of previous grassland models based on MODIS data and to investigate the reasons for errors from different models, and (3) based on the above research results, we propose a method to improve the accuracy of grassland AGB estimation.

2 Data And Methods

2.1 Study Area

The study area (102°23'-102°26'E, 35°5'-35°7'N) is located in the Yangji Community of Sangke Town in Xiahe County, Gansu Province, with the size of approximately 3.86 km (N-S) × 2.77 km (E-W) and a mean elevation of 3,050 m. The natural grassland type in the study area is alpine meadow. The study area is cold and wet throughout the year, and it belongs to the continental monsoon climate of the temperate plateau. The annual average temperature is 2.1 °C; the annual average rainfall is 580 mm.

The study area consists of five fenced experimental sample areas (A–E) of natural alpine meadow grassland (Fig. 1), with a total area of 161.36 ha. Specifically, area A (19.38 ha) is used to conduct natural grassland reseeding tests, area B (16.06 ha) is used for grazing utilization experiments; area C (7.52 ha) acts as a non-treatment testing control area; area D (19.30 ha) is an experimental area for grassland enclosure; and area E (99.10 ha, accounting for 61.55% of the total experimental area) was used for artificial fertilization experiments. The grassland growth in these five experimental areas differed because of the above different usages.

2.2 Sampling Strategy And Data Collection

A total of 13 sample plots of 30 m×30 m are set up inside the 5 experimental sample areas. In each sample plot, five quadrats (1.5 m×1.5 m each) are set up as shown in figure 1, the central point and the four corner points are used to represent the entire plot. To reduce artificial sampling error of biomass measurements, a strategy of 9 sub-quadrats are also used for avoiding repeat sampling in each year, The sub-quadrat identification number (1 to 9) is the order that we used to sample grass each time, e.g., sub-quadrat 1 was used in the first time of any year, 2 was used in the second time of the same year, etc. (Fig 1). The locations of the 13 sample plots are selected based on the following factors: 1) the growth status of the grassland is relatively uniform and is spatially representative, and 2) each sample area, in order to corresponding MODIS pixel, the distance between adjacent plots is more than 250m. During the grassland growth seasons from 2013–2016, observations were performed

approximately once a month and total of 9 field investigations were conducted over the four years.

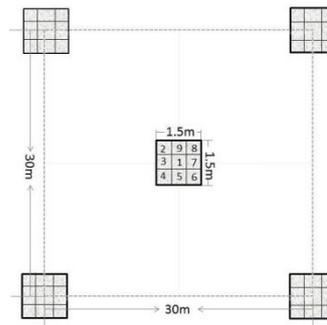


Fig. 1. Distributions of plot

2.3 Previously Established Biomass Inversion Models Based on the MODIS Vegetation Index Over the Tibetan Plateau

In this study, in order to validate the applicability of previous grassland models, we select five previous constructed AGB inversion models based on MODIS data in Tibetan plateau (table 1). The model I is an exponential function model, with coefficient of determination (R^2) of 0.75 and overall estimation accuracy up to 80%, which was based on the 10-d maximum composite MODIS NDVI (with a 1 km resolution) and field measurement data from the end of July to the end of September in 2007 in the TP area [1]. Model II is an exponential function model, with R^2 of 0.49, and RMSE of 671.8 kg DW/ha, which was constructed by measurement data from the TP during July–August of 2005 and 2006 with the monthly maximum composition 1 km resolution MODIS NDVI [5]. Model III and IV are power function models, with the R^2 values of 0.63 and 0.62, calculated by the MODIS EVI at 500 m and 1 km resolution in the Gannan prefecture [7- 8]. Model V is a power function model, with R^2 values of 0.63, which constructed by 250m resolution MODIS EVI in Xiahe County [9].

Table 1 The existing AGB estimation models for alpine grassland in the Tibetan Plateau region

Model	MODIS	Formula	R^2	Literature
I	NDVI	$y=225.42 \times e^{4.4368x}$	0.75	Xu et al, 2007
II	NDVI	$y=268.810 \times e^{2.398x}$	0.49	Feng et al, 2011
III	EVI	$y=3738.073x^{1.553}$	0.63	Cui et al, 2012
IV	EVI	$y=5320.7x^{1.9776}$	0.62	Wang et al, 2010
V	EVI	$y=1719.1x^{2.2588}$	0.63	Bao et al, 2010

2.4 Preprocessing of MODIS Vegetation Index Data

The MODIS vegetation index data are selected from the MODIS 16-d maximum composite NDVI vegetation index product (MOD13Q1) of the United States National Aeronautics and Space Administration. In total, 9 images spanning the years 2013–2016 are selected with a spatial resolution of 250 m from orbit number h26v05 (table2). During the period of field sampling (2013-2016), the NDVI and EVI are used to construct the grassland AGB models and performance of the models using the field measurements assessment.

Table 2 Date between MODIS images and field measurements in the study area

Date of MODIS	Measurement time
2013.07.29-08.13	2013.08.06-09
2014.06.27-07.12	2014.06.28-29
2014.07.13-28	2014.07.27-31
2014.08.14-29	2014.08.14-15
2014.09.15-30	2014.09.26-28
2015.07.13-28	2015.07.11-17
2015.08.14-29	2015.08.10-11
2015.08.30-09.14	2015.09.14-18
2016.07.29-08.13	2016.07.29-30

2.5 Construction of Grassland Biomass Monitoring Model And Accuracy Evaluation

The average dry weight of AGB in all the sub-quadrats of each plot is estimated as the biomass for that sample plot. The grassland biomass of different sample plots are taken as the dependent variable and NDVI, EVI as the independent variables. The Leave-One-Out Cross Validation (LOOCV) method, RMSE(Equation 1), and relative estimate error (REE) (Equation 2) are used to evaluate the accuracy and estimation errors of the linear, exponential, logarithmic, and power regression models for the 2 vegetation indices.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (1)$$

$$REE = \sqrt{\frac{\sum_{i=1}^n [(y_i - \hat{y}_i) / y_i]^2}{N}} \quad (2)$$

where y_i represents the ground-measured value, \hat{y}_i is the estimated value, i represents for a sampling plot, n stands for the number of sample plots.

In addition, in order to study the relative estimation error of the alpine meadow grassland AGB monitoring model on the regional scale (e.g., the experimental area of this study; the Xiahe County; the Gannan prefecture; and the whole Tibetan Plateau), the REE are used to assess the performance of each previous model.

3 Results And Analysis

3.1 Statistical Analysis of Ground Observation Agb And the Corresponding MODIS NDVI

Table 3 shows the statistical results of AGB in the surveyed sample plots and the NDVI, EVI calculated using the corresponding remote sensing data from 2013–2016. There are considerable differences in the biomass during the grass growing season in the 5 sample area of study area. The average values range from 1,480.3–2,097.8 kg DW/ha over the past four years, while the coefficient of variation range from 0.40–0.64. The largest biomass from the 5 sample areas is 3,963 kg DW/ha, and the minimum is 745.5 kg DW/ha. The magnitude of the variation in the EVI values and dispersion degree are larger than NDVI (standard deviation and coefficient of variation range from 0.07–0.11 and 0.09–0.17), with standard deviation and coefficient of variation range of 0.05–0.12 and 0.15–0.24, respectively. In contrast to the mean of EVI value range are slightly smaller than the NDVIs (0.59-0.75) with mean value range of 0.42-0.60.

Table 3. Descriptive statistics of grassland above ground biomass (AGB) and corresponding MODIS indices during July-September of 2013-2015 in the study area, Xiahe County.

Index	Statistic	Sample area					
		A	B	C	D	E	all
Biomass (kg DW/ha)	Max	3963	2415	3348	3669	3788	3963
		.0	.5	.9	.3	.0	.0
	Min	745.	928.	954.	944.	786.	745.
		5	1	6	1	5	5
	Ave	2097	1480	1775	2182	2002	1907
		.8	.3	.3	.6	.4	.7
NDVI	Std	1336	586.	877.	924.	810.	916.
		.3	3	5	5	2	6
	CV	0.64	0.40	0.49	0.42	0.40	0.48
	Max	0.79	0.71	0.75	0.84	0.84	0.84
	Min	0.49	0.53	0.50	0.66	0.60	0.49
	Ave	0.65	0.59	0.64	0.75	0.74	0.67
EVI	Std	0.11	0.07	0.10	0.07	0.07	0.09
		0.17	0.12	0.16	0.09	0.09	0.14
	CV	0.17	0.12	0.16	0.09	0.09	0.14
	Max	0.52	0.44	0.75	0.62	0.61	0.75
	Min	0.26	0.30	0.46	0.37	0.32	0.26
	Ave	0.42	0.36	0.60	0.50	0.51	0.48
EVI	Std	0.10	0.05	0.12	0.08	0.08	0.09
		0.24	0.15	0.20	0.15	0.16	0.15
	CV	0.24	0.15	0.20	0.15	0.16	0.15

3.2 Grassland Biomass Monitoring Model in the Study Area And Evaluation of Its Accuracy

The results of the accuracy evaluation validated by LOOCV for the grassland biomass models constructed on MODIS

NDVI and EVI are listed in table 4. It is found that in the four types of grassland AGB estimation models (linear, logarithmic, power, and exponential), the exponential model performs the best, with the smallest RMSE and REE, while the power model performs the second. Among the two vegetation indices, the exponential model based on EVI is the best, with the smallest RMSE of 656.6 kg DW/ha and REE of 36.3%, and the highest estimation accuracy of the AGB (Fig. 3). Table 5 shows parameters of the best fit model for MODIS EVI. As shown, the model passes the T test and F test at a significance level of 0.01.

Table 4 The validation results by Leave-One-Out Cross Validation for the grassland biomass models based on MODIS NDVI and EVI

Vegetation index	Model	Accuracy evaluation	
		RMSE(kg/ha)	REE (%)
NDVI	Linear	686.8	40.0
	Exponential	676.3	36.5
	Logarithm	720.7	61.5
	Power	680.1	36.8
EVI	Linear	686.9	38.0
	Exponential	656.6	36.3
	Logarithm	715.8	58.6
	Power	662.8	36.4

Table 5 The results of model fits with the optimum inversion models based on MDOIS EVI

Vegetation index	Parameter estimation and T test			Regression significance test	
	Parameter	Estimated value	T	R ²	F
EVI	b	3.862	9.235**	0.55	85.291**
	a	286.599	4.963**		

Note: **represents P<0.01; a and b represent the constant and exponential term of the model, respectively; T and F are the significant values according to T and F tests.

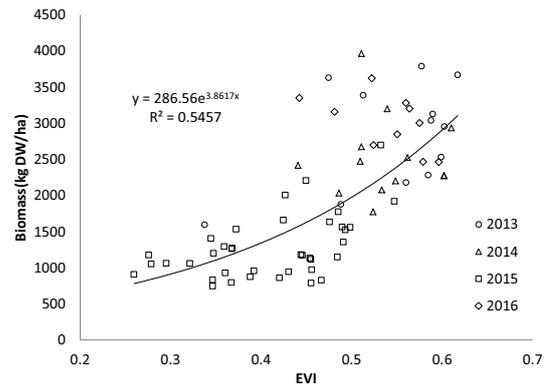


Fig. 3. The best fit models constructed based on MODIS EVI

3.3 Assessment of Previously Established Biomass Inversion Models Based on the MODIS Vegetation Index Over the Tibetan Plateau

Table 6 shows the statistical analysis using the biomass data measured on the ground surface during 2013–2016 in our study area to examine the REE of the aforementioned five models. As shown, the model V error is the smallest (45.8%), followed by IV, III, II, and I. Among the five models, the models I and II based on the entire Tibetan Plateau perform the worse. Although model I achieves the lowest error in the Tibetan Plateau (Table 8), its estimation error of grassland AGB in the study area is highest (68.0%). Models III, V, and IV are built on the regional area (Xiahe and Gannan) of the Tibetan Plateau, their estimation errors are much smaller (45.8%-48.9%).

Compared with this study, the estimation error of grassland AGB is the smallest in the study area scale (namely, the study area of Xiahe County), with REE of 36.3%. Although the previous inversion Models I–V can reflect the overall variation trend of grassland biomass, their estimation errors are higher when applied to our study area.

Table 6 The estimation error for the grassland AGB based on existing inversion models

Model	Study area	REE (%)	Literature
I	Tibetan Plateau	68.0	Xu et al., 2007
II	Tibetan Plateau	55.8	Feng et al, 2011
III	The northeast of Tibetan Plateau (Gannan Prefecture)	48.9	Cui et al., 2012
IV	The northeast of Tibetan Plateau (Gannan Prefecture)	46.5	Wang et al,2010
V	The northwest of Gannan Prefecture (Xiahe County)	45.8	Bao et al., 2010

4 Conclusions

Feasibility study on previous models developed at broad scales based on MODIS indices applied to our small study area suggests that the REE of these models is higher than that of the model constructed in this study by 9.5%–31.7%. Models V, IV, and III based on Xiahe County and Gannan Prefecture do not show considerably differences on the estimation error of AGB, ranging from 45.8%–48.9%, however, the models I and II based on the Tibetan Plateau scale show much larger estimation error, up to 68.0% and 55.8%, respectively.

Discussions

In this study, although the optimum models for the remote sensing monitoring of grassland biomass in the study area are determined, due to the limitations of factors such as the duration of sampling on the ground (only the growing seasons of the grassland from 2013–2016 were sampled), there is still some error and uncertainty for these inversion models. Compared with MSS, TM, HJ-CCD and SPOT, MODIS has relatively low spatial resolution and large estimation error of grassland AGB. Nevertheless, it is still particularly suitable for monitoring grassland AGB of widely distributed scale, due to its high temporal resolution (daily) and large spatial coverage (2,330 km). The result of this study shows that with the increase of the modeling study area, the REE of estimation model are increased gradually. The reason is that in large areas (the Tibetan Plateau area or an entire prefecture or county), the types of natural grassland pasture are complicated, and the geographical distribution is wide, the spatial heterogeneity is strong, and the vegetation index value is subjected to influence from many features and factors (e.g., the ecological environment of the grassland). Therefore, the models built based on large scale, when applied to small local areas (such as this experimental area), would result large errors and low accuracy, poor stability, and large spatial variations. This conclusion is similar to other studies in other areas [10-12].

Therefore, it is important to explore new research approaches to grassland AGB monitoring based on MODIS data to improve the accuracy of grassland remote sensing inversion over large regions in the future. These approaches mainly involve the following aspects: (1) enhance the spatial representativeness of ground sampling sites, for example, by increasing their number and area and improving the range observed by the ground sampling site and the corresponding spatial matching problem given the size of satellite image pixels; (2) improve the temporal matching between ground sampling sites and remote sensing data, for example, by better scheduling the times of field investigations to reduce the time differences between ground surveys and satellite image acquisitions; (3) incorporate new remote sensing

observation techniques (e.g., a hyperspectral imagery and the UAV remote sensing technique) and strengthen research on the spectral characteristics of the grassland vegetation community and the applications of the narrow band remote sensing vegetation index in monitoring grassland AGB [11, 13]; and 4) construct multi-factor grassland AGB estimation models based on statistical analysis and machine learning techniques. These multi factors include climatic factors (e.g., sunlight, temperature, and rainfall), soil factors (e.g., soil nutrients, soil structure and fertility), biological factors (e.g., grassland type, species richness, and distribution of malignant weeds), and management factors (e.g., pasture, fencing enclosures, and rotational grazing) [14-16].

Acknowledgments

This study was supported by the National Natural Science Foundation of China (31672484, 31372367, 41101337 and 41401472), the Project of the Ministry of Agriculture (201203006)

References

- [1] Xu B.; Yang X.; Tao W.; Qin Z.H.; Liu H.Q.; Miao J.M. Remote sensing monitoring upon the grass production in China. *Acta Ecologica Sinica*, 2007, 27(2):405-413, DOI:10.1016/S1872-2032(07) 60012-2.
- [2] Lauenroth W.K.; Hunt H.W.; Swift D.M.; Singh J.S. Estimating aboveground net primary production in grasslands- a simulation approach. *Ecological Modelling*, 1986, 33:297–314, DOI: 10.1016/0304-3800(86)90045-1.
- [3] Ali I, Cawkwell F, Green S, et al. Application of statistical and machine learning models for grassland yield estimation based on a hypertemporal satellite remote sensing time series. *Geoscience and Remote Sensing Symposium. IEEE*, 2014, 5060-5063, DOI: 10.1109/IGARSS.2014.6947634
- [4] Ma W.H.; Fang J.Y.; Yang Y.H.; Mohammat A. Biomass carbon stocks and their changes in northern China's grasslands during 1982-2006. *Science China Life sciences*, 2010, 53(7): 841-50, DOI: 10.1007/s11427-010-4020-6.
- [5] Feng Q.S.; Gao X.H.; Huang X.D.; Yu H.; Liang T.G. Remote sensing dynamic monitoring of grass growth in Qinghai-Tibet plateau from 2001 to 2010. *Journal of Lanzhou University*, 2011, 47(4):75-74.
- [6] Ali I.; Cawkwell F.; Dwyer E.; Barrett B.; Green S. Satellite remote sensing of grasslands: from observation to management—a review. *Journal of Plant Ecology*, 2016, DOI: 10.1093/jpe/rtw005.
- [7] Cui X.; Guo Z.G.; Liang T.G.; Shen Y.Y.; Liu X.Y.; Liu y. Classification management for grassland using MODIS data: a case study in the Gannan region, China. *International Journal of Remote Sensing*, 2012,

- 33(10):3156-3175,DOI:10.1080/01431161.2011.634861.
- [8] Wang Y, Xia W T, Liang T G, Wang C. Spatial and temporal dynamic changes of net primary product based on MODIS vegetation index in Gannan grassland. *Acta Prataculturae Sinica*, 2010, 19(1): 201-210.
- [9] Bao H M. Dynamic Monitoring and Prediction of Aboveground Biomass of Natural Grassland—A case study in Xiahe County of Gansu Province. master thesis, Gansu agricultural university, Lanzhou, 2010.
- [10] Ullah S.; Si Y.; Schlerf M.; Skidmore A.K.; Shafique M.; Iqbal I A. Estimation of grassland biomass and nitrogen using MERIS data. *International Journal of Applied Earth Observation and Geoinformation*, 2012, 19: 196–204, DOI: 10.1016/j.jag.2012.05.008.
- [11] Gao T.; Yang X.Y.; Jin Y.X.; Ma H.L.; Li J.Y.; Yu Q.Y.; Zheng X. Spatio-Temporal Variation in Vegetation Biomass and Its Relationships with Climate Factors in the Xilingol Grasslands, Northern China. *Plos One*, 2013,8(12):e83824-e83824,DOI:10.1080/01431161.2013.823000.
- [12] Ahamed T.; Tian L.; Zhang Y.; Ting K.C. A review of remote sensing methods for biomass feedstock production. *Biomass and Bioenergy*, 2011, 35:2455-2469, DOI: 10.1016/j.biombioe.2011.02.028.
- [13] Liu M.; Liu G.H.; Gong L.; Wang D.B.; Sun J. Relationships of biomass with environmental factors in the grassland area of Hulunbuir, China. *Plos One*, 2014, 9(9):e102344-e102344, DOI: 10.1371/journal.pone.0102344.
- [14] Li F.; Jiang L.; Wang X.F.; Zhang X.Q.; Zheng J.J.; Zhao Q.J. Estimating grassland aboveground biomass using multitemporal MODIS data in the West Songnen Plain, China. *Journal of Applied Remote Sensing*, 2013, 7(1):124-131, DOI: 10.1117/1.JRS.7.073546.
- [15] Liang T.G.; Yang S.X.; Feng Q.S.; Liu B.K.; Zhang R.P.; Huang X.D.; Xie H.J. Multi-factor modeling of above-ground biomass in alpine grassland: A case study in the Three-River Headwaters Region, China. *Remote Sensing of Environment*, 2016, 186:164-172, DOI: 10.1016/j.rse.2016.08.014.
- [16] Diouf A.A.; Hiernaux P.; Brandt M.; Faye G.; Djaby B.; Diop M.B.; Ndione J.A.; Tvchon B. Do Agrometeorological Data Improve Optical Satellite-Based Estimations of the Herbaceous Yield in Sahelian Semi-Arid Ecosystems?. 2016, 8(8). DOI: 10.3390/rs8080668