A comparison of two models with Landsat data for estimating above ground grassland biomass in Inner Mongolia, China

Yichun Xie\textsuperscript{a,b}, Zongyang Sha\textsuperscript{b,c}, Mei Yu\textsuperscript{a,d,}\textsuperscript{*}, Yongfei Bai\textsuperscript{c}, Lei Zhang\textsuperscript{e}

\textsuperscript{a} State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, PR China
\textsuperscript{b} Department of Geography and Geology, Eastern Michigan University, Ypsilanti, MI 48197, United States
\textsuperscript{c} International Software School, Wuhan University, Wuhan 430079, PR China
\textsuperscript{d} Institute for Tropical Ecosystem Studies, University of Puerto Rico-Río Piedras, San Juan, PR 00931-1910, United States
\textsuperscript{e} Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, PR China

\textbf{Abstract}

Two models, artificial neural network (ANN) and multiple linear regression (MLR), were developed to estimate typical grassland aboveground dry biomass in Xilingol River Basin, Inner Mongolia, China. The normalized difference vegetation index (NDVI) and topographic variables (elevation, aspect, and slope) were combined with atmospherically corrected reflectance from the Landsat ETM+ reflective bands as the candidate input variables for building both models. Seven variables (NDVI, aspect, and bands 1, 3, 4, 5 and 7) were selected by the ANN model (implemented in Statistica 6.0 neural network module), while six (elevation, NDVI, and bands 1, 3, 5 and 7) were picked to fit the MLR function after a stepwise analysis was executed between the candidate input variables and the above ground dry biomass. Both models achieved reasonable results with RMSEs ranging from 39.88% to 50.08%. The ANN model provided a more accurate estimation (RMSE\textsubscript{r} = 39.88\% for the training set, and RMSE\textsubscript{r} = 42.36\% for the testing set) than MLR (RMSE\textsubscript{r} = 49.51\% for the training, and RMSE\textsubscript{r} = 53.20\% for the testing). The final above ground dry biomass maps of the research area were produced based on the ANN and MLR models, generating the estimated mean values of 121 and 147 g/m\textsuperscript{2}, respectively.

\section{Introduction}

Natural grasslands and savannas cover some 40\% of the terrestrial globe (Moore, 1966; Chapin \textit{et al.}, 2001) and provide important resources to modern societies. Most of the world's food crops and draft animals were domesticated in grasslands and today's urban cultures trace their roots to the ancient cities of the Asian plains and river valleys. Grasslands are highly complex biomes that vary greatly in species composition, density, and biomass over space and over time. Grasslands are sensitive to variable edaphic conditions, local management, and climate and weather conditions (Lauenroth \textit{et al.}, 1999). Grassland degradation is a serious concern in China due to climate variability and growing human pressures (He \textit{et al.}, 2005). Therefore, accurate spatial monitoring and assessment of grassland resources are increasingly critical to their conservation and, in some cases, their restoration (He \textit{et al.}, 2005; Martin \textit{et al.}, 2005). However, obtaining reliable estimates of grassland biomass and its spatial and seasonal variability poses major challenges (Schino \textit{et al.}, 2003).

Field surveys provide the most accurate method for obtaining vegetative data, but are far too time-consuming and costly to cover large expanses. However, the emergence of new techniques for geographical data collection and processing raises the possibility of quantifying grassland properties remotely (Gao, 2006). Satellite platforms in particular, offer an effective means of collecting contemporary data from vast areas and in short periods of time (Moreau \textit{et al.}, 2003; Nordberg and Evertson, 2003). Historically the US Landsat series of satellites (starting in 1972) has been the workhorse for vegetation assessments of grasslands. In particular, medium resolution (15–30 m) Landsat multispectral thematic mapper (TM) data has been widely used in monitoring and assessing the vegetative resources of grassland areas (Sha \textit{et al.}, 2008). By coupling ground observations with Landsat data, acceptable and affordable estimates of biomass can be made (Li and Liu, 2001; Zheng \textit{et al.}, 2004).

Various analysis methods have been developed for estimating grassland biomass with remotely sensed data. Moreau \textit{et al.} (2003) categorized these into three groups: (1) the empirical relationships of spectral vegetation indices (VIs), (2) Monteith’s efficiency model (Moreau \textit{et al.}, 2003), and (3) the canopy process-based models (Cayrol \textit{et al.}, 1999; van der Werf \textit{et al.}, 2007). The VI-based methods...
estimate above ground biomass by establishing empirical relationships between recorded biomass and the transformations of two or more remote sensed spectral bands (Tueller, 1989; He et al., 2005). NDVI is the most widely known VI for estimating green vegetation biomass by enhancing chlorophyll reflectance differences in the red and near-infrared spectral regions (Boelman et al., 2003; Calvao and Palmeirim, 2004; Wessels et al., 2006). Despite of the usefulness of NDVI, Mutanga and Skidmore (2004) pointed out that NDVI is limited by its non-linear responses at green vegetation extremes, making it insensitive to differences at low and high densities. Accordingly, other VIs have been developed, among which are the modified normalized difference vegetation index (MNDVI), enhanced vegetation index (EVI), simple ratio (SR), and transformed vegetation index (TVI) (Nagler et al., 2005; Mutanga and Skidmore, 2004).

The objective of this study is to compare two different methods for estimating the above-ground dry biomass (hereafter simply “biomass”) in the Xilingol River Basin of the Inner Mongolia Autonomous Region, China, by combining conventionally processed NDVI data from the Landsat ETM+ (enhanced thematic mapper) and coupled with such topographic variables as terrain elevation, slope and aspect. An introduction of the study area is provided in Section 2.1. The research design is discussed in Section 2.2. The procedures for data processing and analysis are explained in Section 2.3. Section 3 presents the model results, error assessment of the biomass mapping. Sections 4 and 5 conclude with observations on the strengths, limitations, and possible improvements of these two methodologies.

2. Study area, research design, and data processing and analysis

2.1. Study area

Located on a rolling semiarid high plain between 43°26’ and 44°29’ N and 115°32’ and 117°12’ E, the Xilingol River Basin, is one of the most representative short grass steppes in northern China (Li et al., 1988). More than 90% of the land in the region is covered by grasses that include associations of Leymus chinensis (Trin.) Tzvel., Stipa grandis P. Smirn., Stipa krylovii Roshev., Artemisia frigida Willd., and Filipolium sibiricum (L.) Kitam. (He et al., 2005). This area has an annual average temperature of 1.71°C, an annual mean precipitation of 29.5 cm, 60–80% of which falls in June, July and August, coinciding with the highest temperatures and the growing season (May–September) (Kang, 2002; He et al., 2005; Kawamura et al., 2005). The elevation decreases from southeast to northwest, with the highest elevation at 1608 m and the lowest at 902 m above sea level (Fig. 1).

Scientific research in Xilingol River Basin started in the early-1950s, when the Xilingol Breeding Stock Rangeland was established (The Inner Mongolia and Ninxia Survey Team of CAS, TIMNSTCAS, 1985). Xilingol River Basin was also designated as a biological experiment site by University of Inner Mongolia in 1957. A large area scientific survey was conducted here in 1964–1965. A permanent ecosystem observation station was established by the Chinese Academy of Sciences in 1979 (Li et al., 1988). Systematic collection of climate, soil, vegetation, and ecosystem data have been conducted since then, which provides the basis for continued grassland research in this region and this manuscript development.

2.2. Research design

Grassland degradation is a major concern in Xilingol River Basin because of growing population pressures, an influx of new settlers, and intensified grazing (Tong et al., 2004; He et al., 2005; Li et al., 2005). Half of Xilingol River Basin has been identified as having medium to severe degrees of desertification, and significant soil and wind erosion (Kang, 2002). Growing human interference, in addition to increasing land use complexities, compounds the difficulty of applying the Landsat ETM+ data for estimating the grassland biomass of this area. Ancillary topographic data and local knowl-
edge were considered important for estimating biomass. Shi et al. (1993) tried a purely geographical model (incorporating elevation, latitude and longitude) to estimate vegetation biomass over a larger region in Xilingol Grassland (containing the current study area). They concluded that this geographic model could produce small-scale biomass maps, but suggested that it might be more accurate if remote sensing data were added to the geographical model. This suggestion was confirmed by follow-up research using both geographical data and remote sensing data (NOAA-AVHRR) (Shi et al., 1994). Moreover, the density of vegetation cover was found to be closely related to soil moisture conditions (He et al., 2005). Since reliable soil moisture data were not available in the study area, topographical features were used as surrogate variables (Sulebak et al., 2000).

The correlations of NDVI remote sensing data with topographic features have been shown to increase at moderate to coarse resolutions (Walsh et al., 1997, 1999). Large local variations between NDVI values and topography occur at scales smaller than the average farm size. The noticeable spatial variations of correlations at fine scales make the assumption of their linear relationships meaningless (Nelson et al., 2007). Therefore, the application of artificial neural network (ANN) appeared reasonable because of their non-parametric nature, making them suitable for the analysis of nearly non-dimensional (Nelson et al., 2007). Moreover, the density of vegetation cover was found to be closely related to soil moisture conditions (He et al., 2005). Since reliable soil moisture data were not available in the study area, topographical features were used as surrogate variables (Sulebak et al., 2000).

The correlations of NDVI remote sensing data with topographic features have been shown to increase at moderate to coarse resolutions (Walsh et al., 1997, 1999). Large local variations between NDVI values and topography occur at scales smaller than the average farm size. The noticeable spatial variations of correlations at fine scales make the assumption of their linear relationships meaningless (Nelson et al., 2007). Therefore, the application of artificial neural network (ANN) appeared reasonable because of their non-parametric nature, making them suitable for the analysis of nearly non-dimensional (Nelson et al., 2007). Moreover, the density of vegetation cover was found to be closely related to soil moisture conditions (He et al., 2005). Since reliable soil moisture data were not available in the study area, topographical features were used as surrogate variables (Sulebak et al., 2000).

The correlations of NDVI remote sensing data with topographic features have been shown to increase at moderate to coarse resolutions (Walsh et al., 1997, 1999). Large local variations between NDVI values and topography occur at scales smaller than the average farm size. The noticeable spatial variations of correlations at fine scales make the assumption of their linear relationships meaningless (Nelson et al., 2007). Therefore, the application of artificial neural network (ANN) appeared reasonable because of their non-parametric nature, making them suitable for the analysis of nearly non-dimensional (Nelson et al., 2007). Moreover, the density of vegetation cover was found to be closely related to soil moisture conditions (He et al., 2005). Since reliable soil moisture data were not available in the study area, topographical features were used as surrogate variables (Sulebak et al., 2000).

The correlations of NDVI remote sensing data with topographic features have been shown to increase at moderate to coarse resolutions (Walsh et al., 1997, 1999). Large local variations between NDVI values and topography occur at scales smaller than the average farm size. The noticeable spatial variations of correlations at fine scales make the assumption of their linear relationships meaningless (Nelson et al., 2007). Therefore, the application of artificial neural network (ANN) appeared reasonable because of their non-parametric nature, making them suitable for the analysis of nearly non-dimensional (Nelson et al., 2007). Moreover, the density of vegetation cover was found to be closely related to soil moisture conditions (He et al., 2005). Since reliable soil moisture data were not available in the study area, topographical features were used as surrogate variables (Sulebak et al., 2000).

The correlations of NDVI remote sensing data with topographic features have been shown to increase at moderate to coarse resolutions (Walsh et al., 1997, 1999). Large local variations between NDVI values and topography occur at scales smaller than the average farm size. The noticeable spatial variations of correlations at fine scales make the assumption of their linear relationships meaningless (Nelson et al., 2007). Therefore, the application of artificial neural network (ANN) appeared reasonable because of their non-parametric nature, making them suitable for the analysis of nearly non-dimensional (Nelson et al., 2007). Moreover, the density of vegetation cover was found to be closely related to soil moisture conditions (He et al., 2005). Since reliable soil moisture data were not available in the study area, topographical features were used as surrogate variables (Sulebak et al., 2000).


dand MLR models. The DEM was reprocessed to fit the ETM+ images using ERDAS Imagine (v. 8.5).

A 1st order polynomial rectification using the GPS data provided an RMSE accuracy of less than half a pixel. Both scenes fit well with the topographic and ancillary data. An improved C-correction method for TM related images (Riano et al., 2003) was applied to reduce any topographic distortion. Digital numbers (DNs) from equivalent dark objects in both scenes (e.g. Xilin River, deserted areas) were found to vary by four or less DN’s in the infrared bands and two or less DN’s in the visible bands. The two scenes were then mosaicked to create a single image with no apparent irregularities. Although there were no observable clouds in the images, atmospheric haze could not be ruled out. Since we had no in situ atmospheric measurements, an image-based atmospheric correction, COST (see Chavez, 1996), was implemented on the image.

Data from the above image was clipped using a boundary polygon for the research area. The topographic data aided in the editing out roads and urban areas. Prior to carrying out our biomass assessment of natural grasslands, farmed and fenced lands were also delineated and edited from the image. TM Band 6 (the thermal band) was eliminated because of its coarse spatial resolution. The spectral statistics for the preprocessed image (Img I) are listed in Table 1.

All field samples with GPS coordinates were registered on Img I after a conversion of coordinates from the latitude/longitude form to the UTM form was made. Several samples were discarded to avoid possible noise: (1) two or more survey quadrats at a sampling site had statistically abnormal values based on 3σ; (2) the sampling sites were close to main roads (<5 pixels); (3) the sampling sites were located within man-fenced farms or cultivated fields. As a result, 461 samples of the 568 remained as validated samples. The mean value of the above ground biomass over the spots satisfying 3σ at each sample site provided the plant biomass for that site. The biomass frequency distribution of all validated samples is shown in Fig. 3.

Two new images were generated to facilitate the biomass assessment. The first one was the NDVI image, which was produced by ratioing Band 3 and Band 4 as follows:

\[
\text{Band 4} - \text{Band 3}
\]

\[
\text{Band 4} + \text{Band 3}
\]

where Band 4 represents the spectral reflection of Near Infrared (NIR) band and Band 3 is the spectral reflection of red band. This NDVI image was then combined with three DEM-derived topographic variables (slope, aspect and elevation) to form a 10-layers stacked image. The data layers of slope, aspect and elevation were produced from the DEM using “ERDAS Modeler”. Statistical information of the input variables and output variable biomass was summarized in Table 2. Taken the fact that the variables with higher values may suppress the influence of the variables of smaller values when fitting models, the pixel values of all layers were normalized

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.156</td>
<td>0.957</td>
<td>0.267</td>
<td>0.057</td>
</tr>
<tr>
<td>2</td>
<td>0.059</td>
<td>0.764</td>
<td>0.144</td>
<td>0.041</td>
</tr>
<tr>
<td>3</td>
<td>0.051</td>
<td>0.987</td>
<td>0.209</td>
<td>0.093</td>
</tr>
<tr>
<td>4</td>
<td>0.008</td>
<td>0.850</td>
<td>0.316</td>
<td>0.056</td>
</tr>
<tr>
<td>5</td>
<td>0.005</td>
<td>0.987</td>
<td>0.480</td>
<td>0.121</td>
</tr>
<tr>
<td>7</td>
<td>0.005</td>
<td>0.987</td>
<td>0.227</td>
<td>0.103</td>
</tr>
</tbody>
</table>

1 Values greater than 3σ were regarded as abnormal, where σ is the standard deviation for all 5 spots.
based on the following function:

\[ B_i = \frac{B - B_{\text{Min}}}{B_{\text{Max}} - B_{\text{Min}}} \]  

where \( B_i \) is the transformed value of an original pixel value \( B \) on layer \( i \) that has maximum and minimum pixel values given by \( B_{\text{Max}} \) and \( B_{\text{Min}} \), respectively. The same operation was also applied to the output variable (biomass) for all field samples. In the latter case, \( B_i \) represents a normalized biomass, \( B \) is the actual biomass, \( B_{\text{Max}} \) and \( B_{\text{Min}} \) are the maximum and minimum biomass values in the sample set. Conversely, the following function is used to derive the actual biomass value \( (B) \):

\[ B = (B_{\text{Max}} - B_{\text{Min}}) \times B_i + B_{\text{Min}} \]

After the normalization, the pixel values of all the layers, including both the input variables and the output variable, were scaled between 0 and 1, which helps minimize the effects of magnitude

**Table 2**

<table>
<thead>
<tr>
<th>Candidate variable</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 1 (N1)</td>
<td>0.201</td>
<td>0.513</td>
<td>0.278</td>
<td>0.049</td>
</tr>
<tr>
<td>Band 2 (N2)</td>
<td>0.086</td>
<td>0.305</td>
<td>0.152</td>
<td>0.036</td>
</tr>
<tr>
<td>Band 3 (N3)</td>
<td>0.098</td>
<td>0.505</td>
<td>0.227</td>
<td>0.083</td>
</tr>
<tr>
<td>Band 4 (N4)</td>
<td>0.129</td>
<td>0.537</td>
<td>0.305</td>
<td>0.048</td>
</tr>
<tr>
<td>Band 5 (N5)</td>
<td>0.105</td>
<td>0.796</td>
<td>0.498</td>
<td>0.111</td>
</tr>
<tr>
<td>Band 7 (N7)</td>
<td>0.035</td>
<td>0.490</td>
<td>0.247</td>
<td>0.094</td>
</tr>
<tr>
<td>Elevator (m)</td>
<td>905.00</td>
<td>1503.00</td>
<td>1148.03</td>
<td>133.27</td>
</tr>
<tr>
<td>Aspect</td>
<td>0.00</td>
<td>361.00</td>
<td>213.30</td>
<td>125.44</td>
</tr>
<tr>
<td>Slope (degree)</td>
<td>0.00</td>
<td>32.83</td>
<td>3.96</td>
<td>4.70</td>
</tr>
<tr>
<td>NDVI</td>
<td>–0.13</td>
<td>0.63</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td>Biomass (g/m²)</td>
<td>2.13</td>
<td>585.23</td>
<td>149.20</td>
<td>90.44</td>
</tr>
</tbody>
</table>

Note: Biomass is the dependent (predicted) variable while all the others are predictive variables.
selected cases from the sample set was input to train the ANN model, while biomass was set as the output variable. Totally 310 randomly selected samples were used for training and testing the ANN and MLR models. Finally, the 461 validated samples were randomly divided into a training set and a testing set with a ratio of about 2 to 1. As a result 310 samples were used for training and 151 samples provided the test set. This unequal proportion favoring the training set data was dictated by our relatively small number of samples.

The model performance, for both ANN and MLR, was assessed based on the agreements between the modeled values and the observed values. The agreements were quantified using RMSE (total root-mean-square errors) and \( \text{RMSE}_r \) (relative RMSE) as the indicators, which were defined as follows (Mäkeä and Pekkarinen, 2004):

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

\[
\text{RMSE}_r = \frac{\sqrt{1/n \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}}{\bar{y}} \times 100
\]

where \( \hat{y}_i \) is the modeled value, \( y_i \) is the observed value, \( \bar{y} \) is the mean of the observed values, and \( n \) is the number of samples (cases) in the testing set.

### 2.4. The ANN model

The ANN model for biomass assessment was built using Statistica 6.0 neural network module, by intelligent problem solver (IPS) and by customizing the number of neurons (from 5 to 15) with a single hidden layer. The results of IPS recommended ANN models and the customized ANN models were compared. To build an ANN model, the general tasks include training ANN, testing ANN and applying ANN. IPS is a toolbox capable of doing the first two tasks, creating and testing neural networks for data analysis and prediction tasks. Training task is of the most fundamental importance to build ANN models in which the observed values of the output variable is compared to the network output, and then the error is minimized by adjusting the weights and biases. IPS can automatically design a number of neural networks and recommend a list of ANN architectures. During the training process in this study, IPS used a back propagation algorithm at the beginning and then depended on a conjugate gradient descent (CGD) algorithm to converge and optimize the ANN models so that the minimum error (the observed values versus the modeled values) was achieved. Additional ANN models, with neurons increased from 5 to 15 with a single hidden layer were tried and compared to the IPS derived models. The ultimate ANN architecture was decided by balancing performance against complexity because a complicated ANN requires significant processing time and data storage space.

All ten variables (see Table 3) were selected as candidate inputs while biomass was set as the output variable. Totally 310 randomly selected cases from the sample set was input to train the ANN models, respectively. Dozens of ANN architectures, including the ones from IPS and the customized, were examined. Finally an multi-layer perception (MLP) network from IPS (with a regression ratio of 0.558 and a correlation 0.830) provided a reasonable complexity as well as accuracy for predicting biomass. This network included 3 layers, one input layer, one hidden layer and one output layer (Fig. 4). The input layer used a linear function and corresponded to seven input data items, NDVI, aspect, N1, N3, N4, N5 and N7. The hidden layer had 7 neurons and used a basic logistic sigmoid transfer function for each neuron. The one-neuron output layer was the modeled biomass corresponding to the specific inputs. Furthermore, the training was realized by setting the learning rate as 0.75 and the momentum as 0.45, respectively. At the same time, the epoch size was set equal to the training set size (i.e., 300). During the training period, an error trend indicated by RMSE was noticed to decrease with the increasing number of iteration. After 30 training cycles, no significant effect on error reduction has been detected. In addition, to verify the possible influence of the selected training samples contained in the training dataset on the selected ANN models, two more tries of building ANN models were tested using other two sets of randomly selected training cases, respectively, and a similar architecture of the ultimate ANN model was obtained. Little variation (within 5.0%) was noticed among all of the three trials in terms of values of both RMSE and \( \text{RMSE}_r \). Thus, this ANN architecture was used to create the final biomass map.

### 2.5. The MLR model

The construction of MLR model for biomass assessment was executed through a step-wise regression analysis using the ten predictive variables and the biomass variable. Six predictive variables (elevation, NDVI, and TM bands 1, 3, 5 and 7) that statistically had significant contributions to the prediction of biomass were selected to create the final biomass map (Table 3).

### 3. The results, error assessment, and biomass mapping

Analysis of the ANN model is shown in Table 4. Seven input variables were chosen to predict the biomass when the pruning ratio...
threshold for the inputs was set at 1.05 (Table 4). Band 7 (N7), mid-IR, (with the ratio value of 3.19) was ranked the first and assumed to be the most sensitive indicator for biomass prediction. This was somewhat confirmed with the correlation analysis, as N7 had the highest correlation coefficient among all input variables. The next one was Band 1 (N1), blue, (with the ratio value 2.95), which was considerably higher than the ratio values of the remaining variables. The other three bands (N3, N5 and N4) were all significantly related to the modeled value of biomass. NDVI and aspect were the two least sensitive, but still statistically significant, input variables to the ANN model (the ratio values were 1.36 and 1.15, respectively).

MLR model picked up six variables, four spectral bands (1, 3, 5 and 7), NDVI, and elevation. All of the variables were mapped against biomass through a simple linear regression analysis fed with the selected samples after a robust estimator analysis (a trimmed mean estimator to delete 5% samples from both sides). Result showed that all 4 selected spectral bands indicated negative relationships with the biomass (the correlation coefficients ranged from $-0.475$ to $-0.414$). NDVI, as a frequently used indicator to study the intensity of vegetation cover, represented a positive correlation with the biomass. The correlation between elevation and biomass ($r=0.233$) was the lowest among the six predictive variables. The bi-variate linear regression plots of the six predictive variables versus the biomass were depicted in Fig. 5. When the six variables were jointly used to fit a multi-variate regression model, the explanation and prediction power was much improved compared with any single one (Fig. 6).

**Table 4**

Significance analysis of ANN (for the training dataset, $n = 310$)\(^a\).

<table>
<thead>
<tr>
<th></th>
<th>N7</th>
<th>N1</th>
<th>N3</th>
<th>N5</th>
<th>N4</th>
<th>NDVI</th>
<th>Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Missing error</td>
<td>213.41</td>
<td>197.33</td>
<td>127.30</td>
<td>108.56</td>
<td>95.34</td>
<td>91.14</td>
<td>77.20</td>
</tr>
<tr>
<td>Ratio</td>
<td>3.19</td>
<td>2.95</td>
<td>1.90</td>
<td>1.62</td>
<td>1.43</td>
<td>1.36</td>
<td>1.15</td>
</tr>
</tbody>
</table>

Baseline error = 66.89, ratio = missing error/baseline error, ratio-threshold = 1.05.

\(^a\) To define the significance of a particular variable, $v$, ANN was firstly run on a set of training cases, and the accumulated network error was called baseline error. Then ANN was run again on the same training cases by replacing the observed values of $v$ with the estimated values of $v$. The accumulated network error was called missing error.

Both ANN and MLR used the same training and testing sets to ensure the comparability of the results. RMSEs were calculated for the training dataset, the testing dataset, and the entire dataset, respectively (Table 5). First, the overall values of the RMSE of both ANN and MLR models were relatively small, which indicated the predictions of these models were relatively accurate. Second, ANN had lower RMSE and RMSE\(_r\) values for the training, testing and entire datasets than MLR, suggesting that ANN was more accurate for estimating the biomass than MLR. The result also showed that the RMSE and RMSE\(_r\) values, for both ANN and MLR, were lowest

**Table 5**

Accuracy assessment for the training set ($n = 310$) and the testing set ($n = 151$).\(^a\)

<table>
<thead>
<tr>
<th>Model</th>
<th>Training set (%)</th>
<th>Testing set (%)</th>
<th>Entire set (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>RMSE(_r)</td>
<td>RMSE</td>
</tr>
<tr>
<td>ANN</td>
<td>59.60</td>
<td>39.88</td>
<td>63.20</td>
</tr>
<tr>
<td>MLR</td>
<td>78.36</td>
<td>49.51</td>
<td>79.36</td>
</tr>
<tr>
<td>ANN(^a)</td>
<td>74.41</td>
<td>45.84</td>
<td>76.65</td>
</tr>
</tbody>
</table>

\(^a\) Indicates the results using the same set of variables used by MLR.
Fig. 7. Biomass maps modeled by MLR and ANN with the histograms.

for the training set, medium for the entire dataset, and highest for the testing set.

Fig. 6 displayed the scatter plots and the fitted lines between the observed biomass against the values modeled by ANN and MLR based on Eq. (3). ANN and MLR made closer prediction values for biomass when the observed biomass located around the range of 160–170 (g/m²) where the fitted lines of ANN and MLR met. Both the modeled values from ANN and MLR showed larger deviations from the observed biomasses towards the two ends of the fitted lines. Compared with MLR ($R^2 = 0.591$ and $\beta = 0.542$), ANN ($R^2 = 0.817$ and $\beta = 0.748$) had higher accuracies. Moreover, ANN and MLR behaved significantly different. The scatter plot of MLR was more clustered in the lower range and the fitted line passed over a relatively high position, overestimating biomasses for a good portion of the training cases. The scatter plot of ANN was also concentrated in the lower range, but the fitted line was across along the middle line.

The selected ANN and MLR models were applied over Img II to produce the final biomass maps to quantify the biomass over the research area (Fig. 7). The average biomass modeled by MLR and ANN was around 147 and 121 (g/m²), respectively. As illustrated by Fig. 6, the higher estimate of biomass modeled by MLR than ANN was largely caused by the overestimate of biomasses by MLR in the range of 75–150 (g/m²). Furthermore, when comparing the biomass maps modeled by both ANN and MLR (Fig. 7) and the elevation map (Fig. 1), it is found that the biomass distribution correlates with the elevation. As the elevation declines from southeast to northwest, the vegetation biomass shows a decreasing trend. The density of vegetation cover is related to the landforms and decreases in the following order: low mountains → lava tablelands → hilly plains → plains (Tong et al., 2004; He et al., 2005).

4. Discussions

4.1. Field sample validation

A concern arises from the removal of nearly 20% of the field samples. This is not an advisable practice and was necessitated by the limited number of acceptable samples. This was attributed to three factors: (1) the personnel cost for field sampling, (2) the limited accessibility to certain areas, especially in the southeastern mountainous region, and (3) the extensive impact of human land uses. For example, some of the initial sample sites were located close to the roads and some were within cultivated fields. Moreover, natural plant communities were often disturbed to different degrees by human land use or animal grazing. (The disturbed sites were removed from consideration.)

4.2. Impact of vegetation community variations

Stratified sampling on the basis of different grassland communities would likely improve biomass estimates. However, some of the vegetation communities had few samples due to their small areas or scattered distributions, or were inaccessible for sampling. Other
research projects conducted at the Grassland Ecosystem Observation Station of Chinese Academy of Sciences show that limited field samples often constrain establishing reliable characterizations of the vegetation communities. The number of field samples used in our research is the highest for studies of Xilingol River Basin so far reported.

4.3. Impact of topographic effect on biomass and NDVI

We suspected that topography (elevation, aspect and slope) affects the nature and distribution of plant growth (Matthes-Sears et al., 1988; Perelman et al., 2001; López et al., 2003; Bennie et al., 2006). This notion is confirmed by both the ANN model (aspect as a contributing variable) and MLR model (elevation as a contributing variable). However, the relationship between topography and biomass is clearly not linear, and different patches are localized. For instance, although five variables (N1, N3, N5, N7 and NDVI) were picked up by both ANN and MLR models, other selected variables were different. N4 and aspect were selected as the input variables for the ANN model. On the contrary, elevation, while adopted by MLR, was not listed among the input variables for ANN. ANN is basically a non-linear method. Therefore, ANN is able to find non-linear relationships between variables (Gross et al., 1999).

As expected, NDVI was correlated with biomass ($r = 0.467$, Table 3) and was selected by both ANN and MLR models. However, NDVI had only limited success of predicting the actual biomass (Boelman et al., 2003; Niu and Ni, 2003; Calvao and Palmierim, 2004; Mutanga and Skidmore, 2004; Wessels et al., 2006). NDVI’s explanation power may be compromised in the areas with either very dense or sparse vegetation (Shupe and Marsh, 2004). Moreover, it is difficult to capture local interactions between topography and NVDI. Their relationships, as well as the geographically weighted correlations (GWCs), increased significantly and consistently when the scales of analysis increased (Nelson et al., 2007). Thus, the goal of applying a GWC test in many cases is to ensure that interesting local patterns are identified rather than to establishing a form of statistical inference (Fotheringham et al., 2002, 165–167; Nelson et al., 2007). As discussed above, there are obvious local patterns of plant community compositions, topographic landscapes, climate conditions, and human influences in Xilingol River Basin. Spatial explorations based on stratified vegetation communities would likely generate more meaningful biomass estimates, which remains as a goal of future research.

4.4. Discussions on the accuracy

The ANN model performed well, even in grasslands where diagnostic variables for biomass were poor and the data were noisy. Our study showed the biomass values modeled by ANN ($R^2 = 0.817$, RMSE$_r = 40.61\%$) were closer to the observed values than MLR ($R^2 = 0.591$, RMSE$_r = 50.08\%$). Moreover, the overall accuracy levels achieved by the ANN and MLR models in our study were improved over other studies conducted in the same area using different remote sensing platforms, e.g. $R^2 = 0.414$ by Kawamura et al. (2003), and $R^2 = 0.447$ by Kawamura et al. (2005). The levels of accuracy tests from our research are comparable to vegetation studies of other biomes. For example, RMSE$_s$ of the total aboveground biomass of the forest stands in the study of biomass in boreal forests using ASTER satellite, were 44.7% and 41.0% (Muukkonen and Heiskanen, 2005). Zheng et al. (2004) developed 3 regression functions to estimate aboveground biomass using Landsat 7 ETM+ derived spectral indexes with $R^2$ (the regression fitness between the aboveground biomass and Landsat 7 ETM+ derived spectral indexes) $= 0.82$, 0.86 and 0.95, respectively, which was also comparable to our ANN model ($R^2 = 0.82$).

5. Conclusion

Xilingol River Basin is a typical grassland area in Inner Mongolia, China which is one of the largest remaining grassland ecosystems in the world (Kawamura et al., 2005). Modeling the biomass distribution at a regional scale is of great importance to support the study of grassland ecology and social–economic environment. Unfortunately, no systematic studies have been yet implemented to map biomass for the region. Therefore, we proposed in this study two models, ANN and MLR, to do this work. The ANN model was developed to provide a quantitative analysis of the natural grassland biomass for, while the standard MRL model provided a methodological control. The general accuracies of both models show that both models are applicable to map aboveground biomass in the study area. For both models, topographic elements, such as elevation, aspect and slope, contributed to the remotely sensed input variables. In comparing the two models, ANN produced more accurate estimates than MLR in terms of the fitness between the modeled and the observed biomass values. Considering the fact that 2004 was considered a normal year in terms of the climate conditions (rainfall, temperature, solar radiation, etc.), the biomass map derived from ANN may be representative and can be used as a benchmark to assist in assessing current and future impacts of land-use development, grassland grazing capacities, and animal production potentials. However, since ANN is usually built without knowing underlying relationships between different variables (Sofouglu, 2008), it is difficult to interpret the resulting ANN structure for information concerning why the selected variables explained the above ground biomass. It is therefore more appropriate to apply ANN to applications aimed at predicting biomass rather than understanding its phenomenology. Moreover, a time series of biomass mapping for selected years spanning a period should be carried out in order to enlighten how the grassland in Xilingol River Basin has been degraded.

Acknowledgements

This project was partially supported by The One Hundred Scholars—Distinguished Overseas Scholar Funds from The Institute of Botany at Chinese Academy of Sciences (CAS), and three grants from Natural Science Foundation of China (30590384 and 30521002), and Chinese Academy of Sciences (KZZX2-XB2-01). The authors are also grateful to the research staff and graduate assistants at CAS—Inner Mongolia Grassland Research Station (IMGERS) who assisted collecting the field samples for this research, and the insightful comments from the anonymous reviewers.

References


