Estimating above-ground biomass of the pasture using sentinel-2 data: a case study in Bayandelger soum, Tuv province

Boldbaatar Natsagdorj[1], Sainbayar Dalantai[1], Bayartungalag Batsaikhan[1], Tsogtjagal Batsaikhan[1], Altantuya Dorjsuren[1]

*[1] Institute of Geography and Geoecology, Mongolian Academy of Sciences, Ulaanbaatar 15170, Mongolia
Emails: boldbaatarn@mas.ac.mn; sainbayard@mas.ac.mn; bayartungalag_b@mas.ac.mn; tsogtjargal_b@mas.ac.mn; altantuyad@mas.ac.mn

EST 2023

INTRODUCTION

The pasture ecosystem covers 109645.6 thousand hectares of Mongolian territory, and it is an important source of livestock and livelihoods of herders in Mongolia. Pasture ecosystems play a key role in providing food, goods, and services for humans, and are crucial to livestock grazing (Boval and Dixon, 2012). Pastoralism plays the main role in the economy and livelihood of herders of Mongolia. Pasture biomass is one of the pasture health indicators in animal husbandry (Otgonbayar et. 2018). Therefore, it is necessary to accurately investigate the biomass of pasture that comprise more than 70% of the territory of Mongolia. Moreover, determining the pasture yield in different geographical regions by pattern, type, and season is the main part of forming a pasture monitoring system (Gendaram et al. 2012). Accurate estimation of the spatial distribution of the AGBs under different grazing pressures is indispensable for identifying the livestock carrying capacity and planning their appropriate management. (Primi et al. 2016). The biomass production, defined as the amount of dry matter content (DMC) per unit area produced (kg ha−1), is a key focus for monitoring grassland production for livestock (Quan et al. 2017). In addition, the AGB influences environmental processes, including carbon balance, soil nutrient dynamics, soil erosion, and water cycling (Anaya et al. 2009; Askar et al. 2018; Wen et al. 2013). As it is known from the existing literature studies, pasture management has been widely used across the world since the 1980's. At the regional level with a lack of field data and large spatial areas are under consideration, the best way to estimate biomass is remote sensing (Kumar et al. 2015; Anaya et al. 2009). The most precise way to determine the grassland AGB is the field surveys, but they are too time-consuming and expensive for vast areas (Quan et al. 2017). The main principle of remote sensing is to make decisions based on spectral reflectance characteristics of any object. Based on spectral reflectance characteristics, the remote sensing images are analyzed and the outputs are used for mapping and biomass estimation (Baloloy et al. 2018; Wang et al. 2021; Huete, 1988). The objectives of this study are to evaluate newly proposed or not commonly used algorithms (RF, SVR) for AGB estimation and explore how they perform in comparison with established models, and to investigate how the performances of these machine learning regression algorithms for AGB estimation from multiple satellite data products vary with the AGB ranges.

METHODS

According to our research goal, field samplings of botanical composition were carried out at a total of 68 plots for 5 days between June 16-21, 2020 in Bayandelger soum. Our field study was conducted in preparatory and field measurement research. Transect and mowing methods were used to measure biomass from each plot. Plots are square, 50x50cm in size, and vegetation is harvested and sealed in plastic bags. In the laboratory, field-measured biomass was dried in 80ºC for 24 hours. Two scenes of Sentinel-2 data acquired on 13 June in 2020 were downloaded from USGS site (https://earthexplorer.usgs.gov/). Sentinel-2 images were processed by the ENVI 5.1 and mapping of spatial distribution was performed in ArcGIS 10.8 software.

Sentinel-2 level 1-C processing includes radiometric correction. Atmospheric corrections were conducted using the SEN2COR plugin tool in Sentinel Application Platform (SNAP) software. Also, bands for classification algorithms and vegetation indices were conducted using SNAP software. To develop a robust methodology to estimate pastureland biomass from Sentinel-2 data, field-measured biomass samples were analysed together with spectral information derived from Sentinel-2. In order to validate estimated biomass, a total of 68 field-measured biomass samples were available in Bayandelger soum (Figure 1). In addition, a combination of a machine learning regression model and remote sensing techniques was used to develop the biomass estimation model of vegetation and its mapping. A flowchart of the processes to obtain the spatial distribution map is shown in Figure 2.

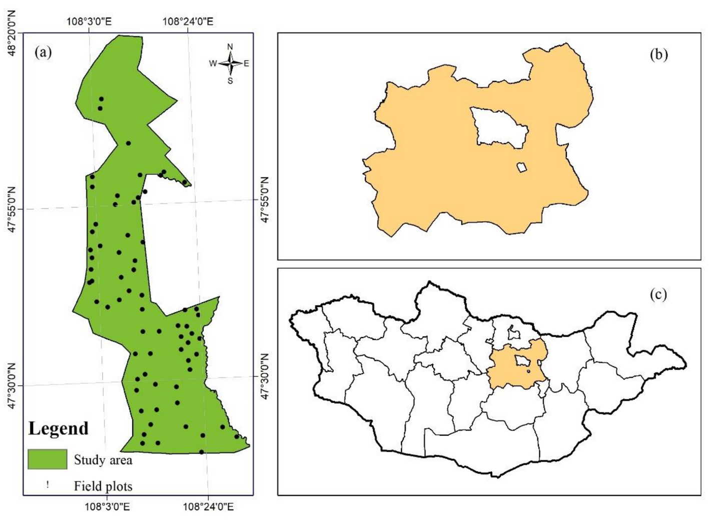

Figure 1. *Location map of study area: (a) Bayandelger soum, (b) Tuv aimag, (c) Map of Mongolia*

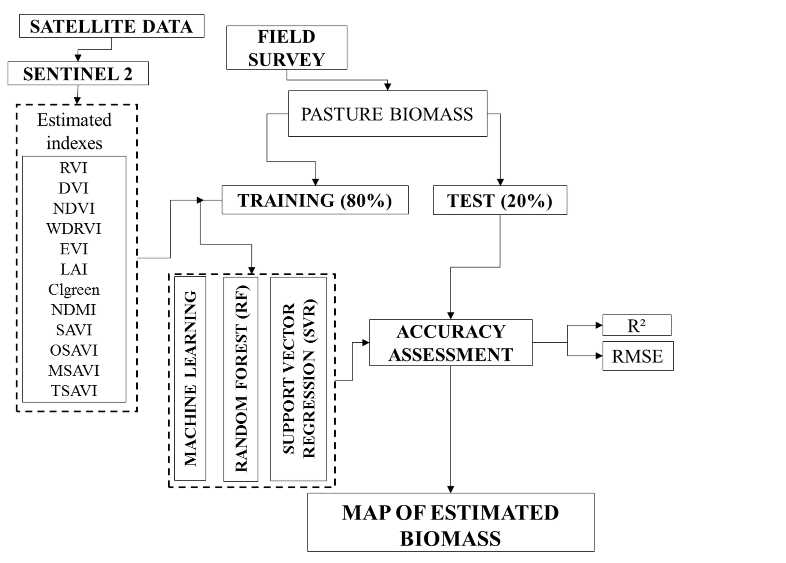

Figure 2. *Flowchart for estimating pasture biomass using Sentinel-2 and machine learning for generating a biomass map.*

These variables from Sentinel-2 data are used in machine learning models for estimating pasture biomass: We used high-resolution Sentinel 2 data (10m), which was based on various vegetation indices including NDVI, DVI, RVI, OSAVI, TSAVI, SAVI, MSAVI, NDMI, Clgreen, WDRVI, LAI and EVI, and field biomass samples collected from 68 sites, to estimate the pasture biomass in Bayandelger soum of Tuv province, Mongolia.. We modeled the biomass estimation using two machine-learning models, Random Forest (RF) and Support Vector Regression (SVR), and compared the outcomes.

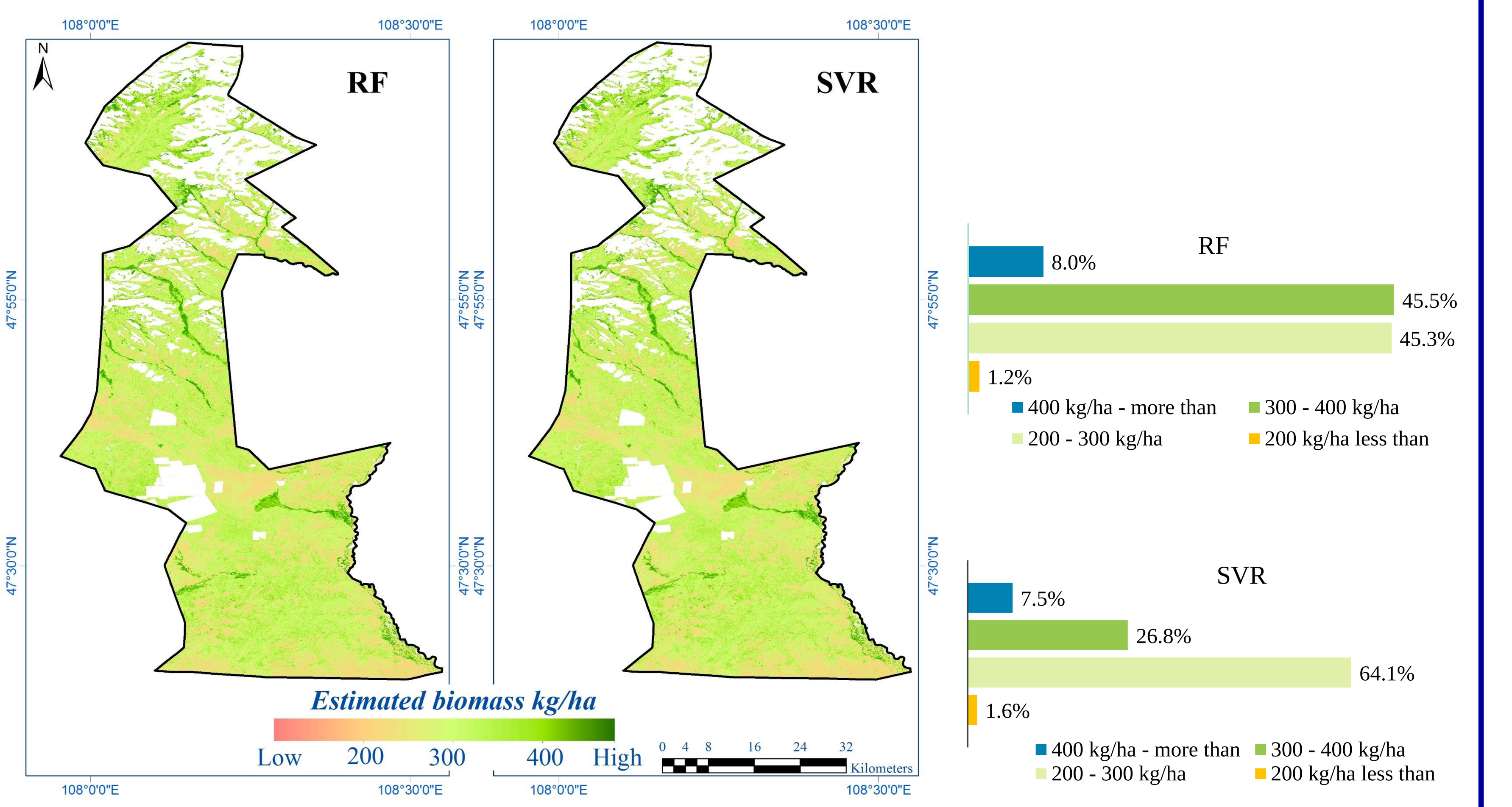

Figure 7. *Estimated biomass using Sentinel 2 data vegetation indices and machine-learning model.*

Machine learning models estimated that the areas with less than 200 kg/ha biomass covered the smallest areas and accounted for 1.2-1.6% of the total area, whereas, areas with 200 – 300 kg/ha biomass accounted for the largest and ranged between 45.3-64.1%. For areas with 300 – 400 kg/ha and more than 400 kg/ha biomass accounted for 26.8-45.5% and 7.5-8%, respectively. Areas with less than 200 kg/ha biomass were found in the southern part of the soum and areas with more than 400 kg/ha occurred in the meadows, wetlands, river valleys, and forest of the central and northern parts of the Bayandelger soum of Tuv province.

RESULTS

Modeling results, assessment, and comparison. The result showed that the prediction accuracy of the SVR was relatively higher than SVR and R^2=0.70 and RMSE=67.12 kg/ha for RF and R^2=0.65 and RMSE=72.32 kg/ha for the RF Table. 1. and Fig. 4. Variable importance. Because RF and SVR can identify the importance of characteristic variables, this study put all 12 variables into two models to rank their importance to AGB. The importance degree was expressed by a score. The results are shown in Fig. 2. According to the result, importance of EVI, TSAVI, RVI, DVI, NDMI and OSAVI was significantly higher in the RF and SVR, respectively. However, the importance of LAI and WDRVI in both methods was significantly lower than in the other variable.

Table. 1. Performance metrics of the two models in estimating AGB

Machine learning models	RMSE (kg/ha)	R^2
Random forest regression (RF)	72.32	0.65
Support Vector Regression (SVR)	67.12	0.70

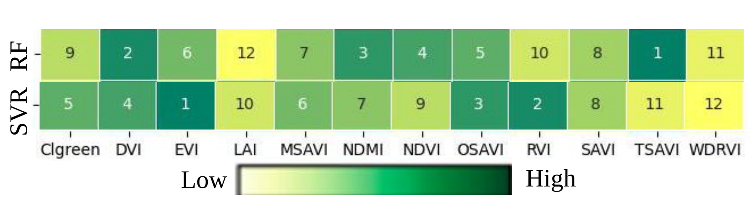

Figure 4. *Feature importance values in the two models.*

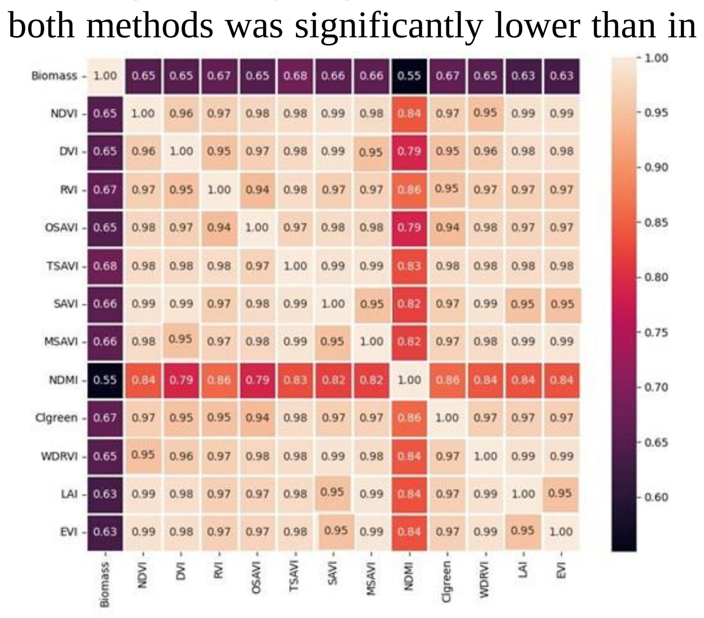

Figure 5. *Correlation matrix between biomass and selected vegetation indices*

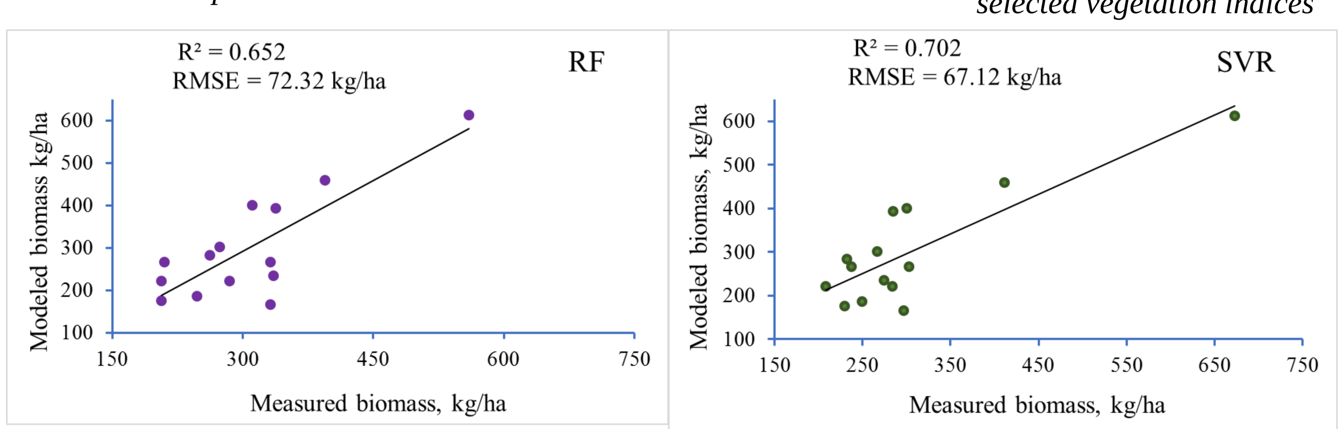

Figure 6. *Scatter plots of the two models in estimating AGB in the testing datasets*

DISCUSSION

A spatial map of pasture biomass covering Mongolia at a spatial solution of 8x8 km across Mongolia was generated using remote sensing data, based on NOAA satellite NDVI (Erdenetuya, 2004). In addition, data from the MODIS satellite was combined with ground data to map the vegetation condition and summer situation in the country in 2002 by Information and Research Institute of Meteorology, Hydrology, and Environment. Otgonbayar et al. (2018) calculated pasture biomass with a spatial resolution of 30 meters across Mongolia based on a total of 17 indices using Landsat 8 multi-temporal satellite imagery. Our research offers advantages of better accuracy, specifically 10 meters spatial ratio for machine learning modeling of pasture biomass. The findings of this research will further serve as the grounds for the environmental baseline study.

CONCLUSIONS

In this study, we wanted to evaluate biomass in Bayandelger soum of Tuv province, Mongolia using vegetation indices derived from Sentinel-2 satellite data, along with machine learning. Research indicated that including NDVI (R=0.65, p<0.01), DVI (R=0.65, p<0.01), RVI (R=0.67, p<0.01), OSAVI (R=0.65, p<0.01), TSAVI (R=0.68, p<0.01), SAVI (R=0.66, p<0.01), MSAVI (R=0.66, p<0.01), NDMI (R=0.55, p<0.01), Clgreen (R=0.67, p<0.01), WDRVI (R=0.655, p<0.01), LAI (R=0.63, p<0.01) and EVI (R=0.63, p<0.01),

NDVI (R=0.71, p<0.01) , and NDI45 (R =0.67, p<0.01) showed a good correlation with field-measured biomass. We evaluated two machine learning models for estimating forest AGB from Sentinel data products and ancillary information. The results showed that pasture AGB estimated with RF and SVR, had the mean R^2 for runs ranging from 0.65 to 0.70, RMSE ranging from 67.12 to 72.32 kg/ha, and SVR were more accurate than RF.

References:[1].Anaya, J.A., Chuvieco, E. and Palacios-Orueta, A., 2009. Aboveground biomass assessment in Colombia: A remote sensing approach. Forest Ecology and Management, 257(4), pp.1237-1246.Askar, Nuthammachot, N., Phairuang, W., Wicaksono, P. and Sayektiningsih, T., 2018. Estimating aboveground biomass on private forest using Sentinel-2 imagery. Journal of Sensors, pp. 1-11. [2] Baloloy, A.B., Blanco, A.C., Candido, C.G., Argamosa, R.J.L., Dumalag, J.B.L.C., Dimapilis, L.L.C. and Paringit, E.C., 2018. Estimation of mangrove forest aboveground biomass using multispectral bands, vegetation indices and biophysical variables derived from optical satellite imageries: rapideye, planetscope and Sentinel-2. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 4(3). [3] Boval, M. and Dixon, R.M., 2012. The importance of grasslands for animal production and other functions: a review on management and methodological progress in the tropics. animal, 6(5), pp.748-762. [4] Clevers, J.G.P.W., Van der Heijden, G.W.A.M., Verzakov, S. and Schaepman, M.E., 2007. Estimating grassland biomass using SVM band shaving of hyperspectral data. Photogrammetric Engineering & Remote Sensing, 73(10), pp.1141-1148. [5] Delegido, J., Alonso, L., Gonzalez, G. and Moreno, J., 2010. Estimating chlorophyll content of crops from hyperspectral data using a normalized area over reflectance curve (NAOC). International Journal of Applied Earth Observation and Geoinformation, 12(3), pp.165-174. [6] Mori, N., Debeljak, B., Škerjanec, M., Simčič, T., Kanduč, T., & Brancelj, A., 2019. Modelling the effects of multiple stressors on respiration and microbial biomass in the hyporheic zone using decision trees. Water Research, 149, 9–20. https://doi.org/10.1016/j.watres.2018.10.093. [7] Mountrakis, G., Im, J., & Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 66(3), 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001. [8] Otgonbayar, M., Atzberger, C., Chambers, J., & Damdinsuren, A. (2019). Mapping pasture biomass in Mongolia using Partial Least Squares, Random Forest regression and Landsat 8 imagery. International Journal of Remote Sensing, 40(8), 3204–3226. https://doi.org/10.1080/01431161.2018.1541110. [9] Rouse, J., W., Haas, R., H., Schell, J., A., & Deering, D., W. (1974). Monitoring Vegetation Systems in the Great Plains with Erts. NASA. 1974NASSP.351..309R.

Acknowledgement: We thank all colleagues of the Division of Environmental and Natural Resource Management, Institute of Geography and Geoecology, Mongolian Academy of Sciences who helped in this study.