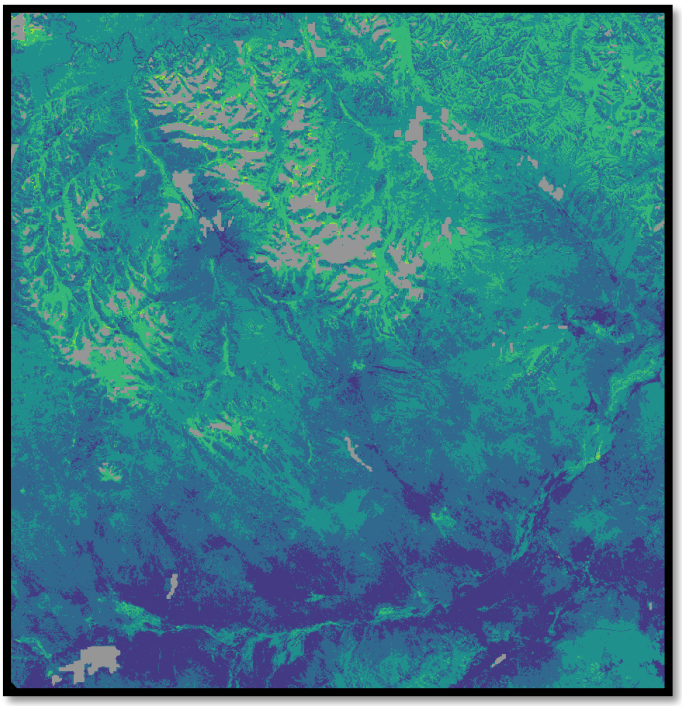
Geo-information Science and Remote Sensing

Thesis Report GIRS-2020-32

Assessing Utilization Intensity of Rangelands Based on Remote Sensing-Derived Carrying Capacity Models and Stocking Density: A case study in Bulgan, Mongolia

Sietse van der Woude, sietsepier.vanderwoude@wur.nl

June 23, 2020



WAGENINGEN
UNIVERSITY & RESEARCH

Assessing Utilization Intensity of Rangelands Based on Remote Sensing-Derived Carrying Capacity Models and Stocking Density: a case study in Bulgan, Mongolia

Sietse van der Woude, sietsepier.vanderwoude@wur.nl

Registration number: 95 09 12 973 010

Supervisors:

dr. NE Tsendbazar

(Geo-Information Science & Remote Sensing, WUR)

dr.ir. JGPW Clevers

(Geo-Information Science & Remote Sensing, WUR)

A thesis submitted in partial fulfilment of the degree of Master of Science

at Wageningen University and Research Centre,

The Netherlands.

23-06-2020

Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2020-32
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Abstract

Rangelands cover a large proportion of the earth's surface, upon which a great deal of biodiversity, but also livelihoods of people depend. Rangeland management on varying scales has been able to benefit from the implementation of remote-sensing based information. However, the interests of different scales of rangeland management, from local to international, have differing requirements for remote-sensing based data in terms of cost, accuracy and speed in order to asses rangeland utilization intensity. In this research, the accuracy of a MODIS Net Primary Productivity-based model (process-based model) developed for Azerbaijan for aboveground carrying capacity in rangelands was compared to that of locally-trained a Landsat 8 surface reflectance-based empirical (Cubist) model. These models were coupled with stocking density estimates derived from summer herder locations to identify areas in which overgrazing had occurred. This was performed for the study area of Bulgan province, Mongolia. It was shown that the performance of the process-based model when applied to Bulgan resulted in low accuracy, however the model performed comparably to the model developed in a previous study for Azerbeijan. Despite this, the mean of the biomass predictions differed significantly from the mean of the observed values ($p < 0.05$). The empirical model delivered better performance than the process-based model and did not result in significant differences between the means of predictions and observed biomass values ($p > 0.05$). Due to overprediction of biomass by the process-based model, overgrazing was not estimated to be as extensive by the process-based model as by the empirical model. However, the process-based model and validation strategies could be further fine-tuned until the accuracy of an NPP-based model could be comparable to that of a strictly empirical model and applicable for use in inter-regional or international comparison of overgrazing levels.

Keywords: rangeland, biomass, overgrazing, forage, degradation, Landsat, Modis

Acknowledgements

I thank my parents, Cees and Christel, for putting up with me moving back home during difficult times, keeping my spirits up, and all the lovely meals and express-zo's without which I would have not even been close to finishing my thesis. I would like to thank Nandika Tsendbazar and Jan Clevers for their guidance, tips and flexibility, especially in transitioning from working in the office and working at home. Finally, I would like to thank all my friends for always providing a listening ear and encouragement when I needed it.

Contents

List of Figures

List of Tables

1. Introduction

1.1. Context & Background

Despite the superficial simplicity of the term 'rangeland' and an immediate association to a singular, idyllic image that might be conjured in one's mind, there is no universally accepted definition of the term (Reeves et al., 2015). According to the definition maintained by the Society for Range Management, land can be considered rangeland when it is predominantly covered by "[...] grasses, grass-like plants, forbs, or shrubs and is managed as a natural ecosystem" (Reeves et al., 2015, p.239). Naturally, this entails that a large proportion of the earth's land surface can be classified as rangeland, although exact estimates range between 18% to 80% (Lund, 2007). Since no single international organization is responsible for reporting on and assessing global rangeland cover such as there exists for global forest cover, like the United Nation's FAO, estimates of rangeland cover vary greatly (Lund, 2007). Despite this lack of monitoring, many people are dependent on rangelands for their livelihoods, providing food and income. Pressures on rangelands are increasing in the form of overuse, land conversion and climate change: estimates of degraded rangeland globally vary from 20% to 73% (Lund, 2007). This conflict between extensive reliance on rangelands and degradation can be seen in the context of the Sustainable Development Goals (SDG's) as compiled by the United Nations. Specifically, the goals 15.3 ("End desertification and restore degraded land") and 2 ("Zero Hunger") are of concern (United Nations, 2016). As populations increase, the intensification of the rangeland use likely increases to maintain a sufficient level of food production, thereby contributing towards the "Zero-Hunger" SDG. However, intensification of land use or the change of land use altogether can be in conflict with reaching the SDG 15.3, as desertification and intensified rangeland use, in combination with other factors, have been shown to be linked (Bedunah & Angerer, 2012).

The interests of monitoring rangelands are many; on a global scale, cross-border comparable and accurate data is required on the state of rangelands in order to identify areas which require extra attention in terms of reaching related Sustainable Development Goals such as SDG 2 and 15.3, and subsequently, in order to direct international efforts accordingly. Unlike the global Forest Resource Assessment (FRA)'s surveys of forest cover, such globally comparable data is

hardly available for rangelands (Lund, 2007). Understanding and balancing the fine and complex relationship between maintaining rangeland-dependent livelihoods and maintaining rangeland quality is necessary and can be aided by defining the 'carrying capacity' of the land using remote sensing and other spatially explicit data (de Leeuw et al., 2019). In the context of rangeland management, carrying capacity is defined as: "[...] the number of livestock per unit area that can be sustained given the amount of available forage and is expressed as the number of animals and days that an area of land may be grazed." (de Leeuw et al., 2019, p.66). In practice, defining the carrying capacity would allow policy makers to more effectively steer grazing pressure, both spatially and temporally, in such a manner that livestock productivity and reduction of land degradation are in balance, by avoiding or more closely monitoring areas potentially at risk of overgrazing.

However, the concept of carrying capacity, and thus overgrazing, may not be universally applicable to all climatic zones and land-cover types, which has been a topic of debate in the field of rangeland ecology for many years, especially since the definition of rangelands can span a variety of land-cover types (Addison, Friedel, Brown, Davies, & Waldron, 2012; de Leeuw et al., 2019; Wehrden, Hanspach, Kaczensky, Fischer, & Wesche, 2012). Many argue that it is fundamentally only appropriate to employ carrying capacity as a conceptual tool when an ecosystem is considered to be an 'equilibrium' system (de Leeuw et al., 2019; Ellis & Swift, 1988). In this system it is assumed there is a limited amount of forage available, and that overgrazing of the forage, and thus over-exploitation of the land, leads to rangeland degradation. The system is therefore mainly influenced and bound by grazing pressure (Ellis & Swift, 1988). Conversely, non-equilibrium systems have been found to be bound by other factors, but mainly the variability of rainfall (de Leeuw et al., 2019; Vetter, 2005; Wehrden et al., 2012). In such systems, if the variability of rainfall is above a certain threshold (often 33% is maintained), the productivity of the land can no longer be considered to be bound by grazing, as periods of droughts largely influence the stocking density, rather than the overgrazing of the livestock itself (Ellis & Swift, 1988). In other words, periods of drought in non-equilibrium systems reduce the amount of available forage, therefore reducing livestock numbers and stocking density, before the stocking density can become high enough to have a significant effect on the availability of forage. This difference between in equilibrium and non-

equilibrium systems is critical in rangeland monitoring, and consequently rangeland management as well, because the carrying capacity approach must be applied only to areas for which it is valid.

Simultaneously, differing interests in rangeland monitoring mean that the requirements of the monitoring activity can differ according to the varying needs of the parties involved. As determined before, global-scale monitoring is required for international comparisons in relation to the monitoring of the state of the Sustainable Development Goals. However, in local-scale management, there is a greater need for a high level of accuracy of the data in order to effectively manage stocking densities locally. In addition, local-scale management also necessitates insight into local systems and factors which can affect the risk of overgrazing (Bedunah & Angerer, 2012). For example, overgrazing and land degradation is often found close to urban areas and water sources, but these factors are not necessarily globally valid (John et al., 2018). Also, data that could prove useful on this scale, such as local stocking densities, is not consistently available globally (Lund, 2007). Therefore, there seems to be an inherent difficulty in fulfilling the requirements of rangeland (overgrazing) monitoring at different scale levels, while employing similar methods to achieve results.

A country which has a large interest in rangeland monitoring is Mongolia. An estimated 70% of Mongolia's land can be classified as rangeland (Bruegger, Jigjsuren, & Fernández-Giménez, 2014). Arid and semi-arid regions such as Mongolia are especially at risk of desertification and land degradation, as there generally exists a high degree of variability of precipitation (Eisfelder, Kuenzer, & Dech, 2012; Huang, Yu, Guan, Wang, & Guo, 2016). Although in 2014 28% of the employment was dependent on rangelands, and approximately 13.5% of the Gross Domestic Product (GDP) stemmed from livestock (National Statistics Office of Mongolia, 2019), its rangeland vegetation productivity has been estimated to have decreased by 20-30% in the last 40 years, and the vegetation composition of an expansive area is projected to change greatly (Angerer, Han, Fujisaki, & Havstad, 2008; Bruegger et al., 2014). Despite a decreasing productivity, the amount of livestock has almost doubled (from 33.6 million to 66.7 million) between 1999 and 2017 despite severe droughts and extreme cold winters in 2001 and 2009, of which the former killed up to 35% of the livestock at the time (Angerer, 2012; Erdenesan, 2018). Currently, a large degree of rangeland management is performed on a highly localized

level, based on traditional communal knowledge (Bruegger et al., 2014). While this can be effective on local scales, cross-boundary policymaking and management capabilities are currently weak (Bruegger et al., 2014). Furthermore, the differing climatic zones which exist in Mongolia complicate the issue of regional and national rangeland management, as ecological systems can display conflicting characteristics of differing ecological models, such as the equilibrium and non-equilibrium paradigms, which require different management methods (Ellis & Swift, 1988; Fernandez-Gimenez & Allen-Diaz, 1999). Therefore, developing an accurate and interpretable method for the estimation of carrying capacity and, potentially, overgrazing risk in Mongolia is key in coping with both anthropogenic and non-anthropogenic pressure sources, which endanger not only the sustainability of rangeland ecosystems, but also the livelihoods of many pastoralists. At the same time, this method should ideally maintain cross-border comparability in the context of the Sustainable Development Goals and be applicable in other regions as well, thus not be dependent on highly cost- and labor-intensive data.

1.2. Problem Definition

It is perhaps an inescapable fact: the narrower the scope for which a prediction model is developed, the better its performance. For example, if one develops a model to predict apple harvest yields in the field of Farmer A, this model will likely not perform as well on harvest yields in the field of Farmer B. By expanding the scope of the model to include the fields of Farmer B in the development and training of the model, the resulting model will perform better for those fields, but due to its wider scope, worse for Farmer A than if only the fields of Farmer A had been involved in the training of the model. The development of models for rangeland monitoring faces this problem, as the interests and requirements of local rangeland policy makers and users can differ greatly from the requirements of global rangeland monitoring. Especially for the estimation of (aboveground) biomass, a wide range of models has been developed. Two main approaches using input from remote sensing images can be distinguished according to Reeves et al.: “[…] (1) empirical approaches that estimate the forage biomass or quality based on a statistical relationship between the spectral bands (or some combination of bands) in the imagery and field-collected vegetation data and (2) process models that use remote-sensing data as inputs for predicting vegetation biomass or quality”

(Reeves et al., 2015). Both approaches are subject to the issue mentioned before (Eisfelder et al., 2012).

Empirical, statistical models derive relations between data points without previous knowledge, thus being entirely dependent on the input data itself, and likely inaccurate for regions outside of the extent from which the training data originated. Approaches using linear and logarithmic regression have been employed in many regions, such as estimating dry matter production in Senegal (Tucker, Vanpraet, Boerwinkel, & Gaston, 1983), grass yield in Golog prefecture, China (Yu, Zhou, Liu, & Zhou, 2010), and aboveground biomass on the Qinghai-Tibet Plateau (Liu et al., 2015). While these are often adequate for limited extents, they struggle with up-scaling (John et al., 2018). Non-parametric approaches using machine learning techniques such as Cubist and Random Forest attempt to overcome problems with linear (parametric) techniques related to limitations in characterizing complex relations in multiple climatic zones and land cover types simultaneously (John et al., 2018). In addition, they allow for an expansion of the indices used for prediction – as frequently only the Normalized Difference Vegetation Index (NDVI) is employed in linear models (John et al., 2018). The drawback of this method is the comparatively extensive field sampling required in order to obtain enough in-situ training and validation points.

Process-based models also often require calibration according to the local conditions in which they are applied (Reeves et al., 2015). Calibrating a process-based model for a wider extent thus likely implies lower accuracy than calibrating for a smaller extent. A process-based model was implemented in Mongolia in order to predict forage availability, in which biomass estimates (derived from MODIS NDVI) were fed into the Phytomass Growth Simulation model (Angerer, 2012). In order to be applicable for the entirety of Mongolia, the model included 65 variables for various categories of soil surface, plant species, livestock and climate data, all of which had to be set according to data acquired in the field (Angerer, 2012). Another, similar model was developed for Kazakhstan using net primary productivity (NPP) estimates derived from remote sensing fed into the BETHY/DLR model; however, this also required local parameterization (Eisfelder, Klein, Niklaus, & Kuenzer, 2014). In order to overcome the issue of the necessity of a large degree of local parameterization, a model for carrying capacity in

Azerbaijan has been developed with promising results by de Leeuw et al., which uses the MODIS global NPP product (de Leeuw et al., 2019). Importantly, this model uses only a very limited range of variables and parameters, for which no field data is required apart from for validation purposes. However, it is unknown how this model performs for regions outside Azerbaijan, and how it compares to empirical models developed from in-situ data and other process-based models with a higher degree of parameterization according to local conditions (de Leeuw et al., 2019).

Apart from the estimation of aboveground biomass and usable forage, and thus carrying capacity, the estimation of spatial distribution of stocking densities is another element in the estimation of overgrazing risk. This can often only be characterized on a rough scale within certain administrative boundaries (National Statistics Office of Mongolia, 2019). However, with the incorporation of known herder locations, the spatial distribution of stocking density within local administrative boundaries can be more accurately estimated.

Finally, the high degree of local parameterization complicates the use of developed models for other regions worldwide, hindering cross-border compatibility of results. In the specific context of Mongolian rangeland monitoring and management, there is a need for a model such as developed by de Leeuw et al. (2019), which combined with stocking densities can provide insight into overgrazing risk on country-wide or regional scale without the need for costly and time-intensive widespread sampling of aboveground biomass, and thus also locally help steer policy. Simultaneously, the model should ideally be easily applicable in other similar regions on the globe without the necessity for extensive local-scale parameterization and potentially aid in the global comparison and monitoring of land degradation in the context of the United Nation’s Sustainable Development Goals.

1.3. Objectives

The general research objective of this thesis is to assess the accuracy of a process-based model for aboveground carrying capacity in rangelands, compared to that of an empirical model, and couple these models with stocking density estimates to identify areas in which overgrazing has occurred, as well as investigate the differences between models in predicted overgrazing levels.

Research questions:

- How suitable is an existing process-based model for rangeland biomass developed for Azerbaijan, for use in Bulgan province, Mongolia, based on validation with in-situ biomass samples?

- Does an empirical model for rangeland biomass increase prediction accuracy compared to the process-based model?

- What is the spatial distribution of forage intake by herder livestock in their summer locations in Bulgan, Mongolia?

- What are the differences in overgrazing predictions between the empirical and process-based models, and what factors influence these differences?

4. Methodology

4.1. Study Area

The area this thesis has used for the analysis is Bulgan province, Mongolia. It is located in the north of Mongolia and covers an area of 48,733 km^2, which is approximately the same size as Slovakia. It is home to 62,214 people, of which approximately 75% live in a rural environment (National Statistics Office of Mongolia, 2018b). On average, the province thus has a population density of only +-1.3 per square kilometer, slightly over half of the country-wide population density of Mongolia. The average annual temperature is -2.4 °C, and temperatures fluctuate between -49 °C in winter and 38 °C in summer (Yu et al., 2013). The vegetation types vary, generally along the north/south axis: while the north is characterized by alpine forests in the Sayan montane conifer forests ecoregion, the landscape gradually transforms to arid steppe plains in the Mongolian Manchurian grassland ecoregion further southwards (Yu et al., 2013). Most of the province lies in the Selenge-Orkhon forest steppe ecoregion of which the vegetation cover is a mixture of forest and grassland (Figure 1).

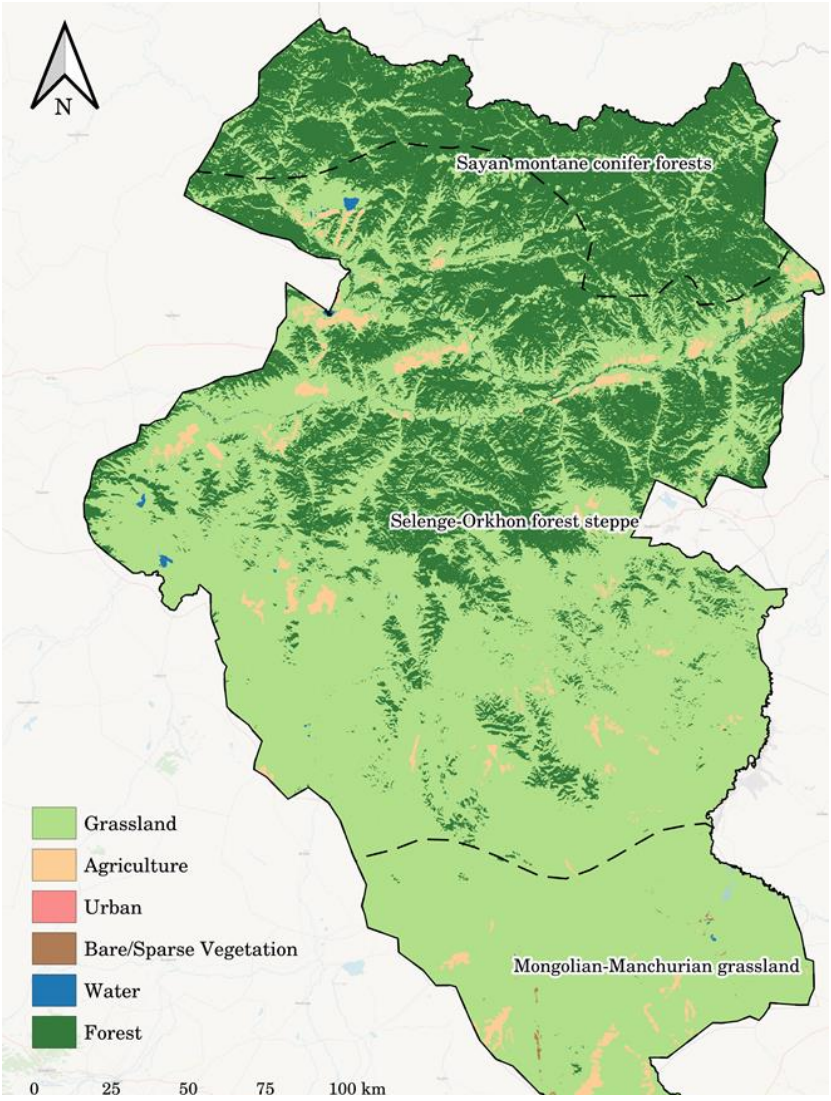


Figure 1: Map of Bulgan Province, Mongolia. The land cover in Bulgan consists of mostly grassy vegetation in the south and forest in the north. The province can be categorized into three ecoregions, as divided by the dashed lines: Sayan montane conifer forests, Selenge-Orkhon forest steppe and Mongolian-Manchurian grassland .

Urban areas are very sparsely distributed throughout the province, and agriculture mostly occurs in the Selenge- and Eg river basins in the north of the Selenge-Orkhon forest steppe ecoregion (Figure 1). The average yearly precipitation ranges from 200 to 350 mm (Yu et al., 2013), and the coefficient of variation of precipitation is below the threshold of 33%, which generally allows for the determination of carrying capacity as outlined in the introduction of this thesis (Figure 2). Herding livestock is an important source of income in Bulgan: the province counts 10,683 households which depend on herding as their primary source of income, which is more than half of the total number of households (19,072) (National Statistics Office of Mongolia, 2018a). Herders generally move in a rotational system according to the seasons, from summer to winter pastures (Al-Jaloudy et al., 2005).

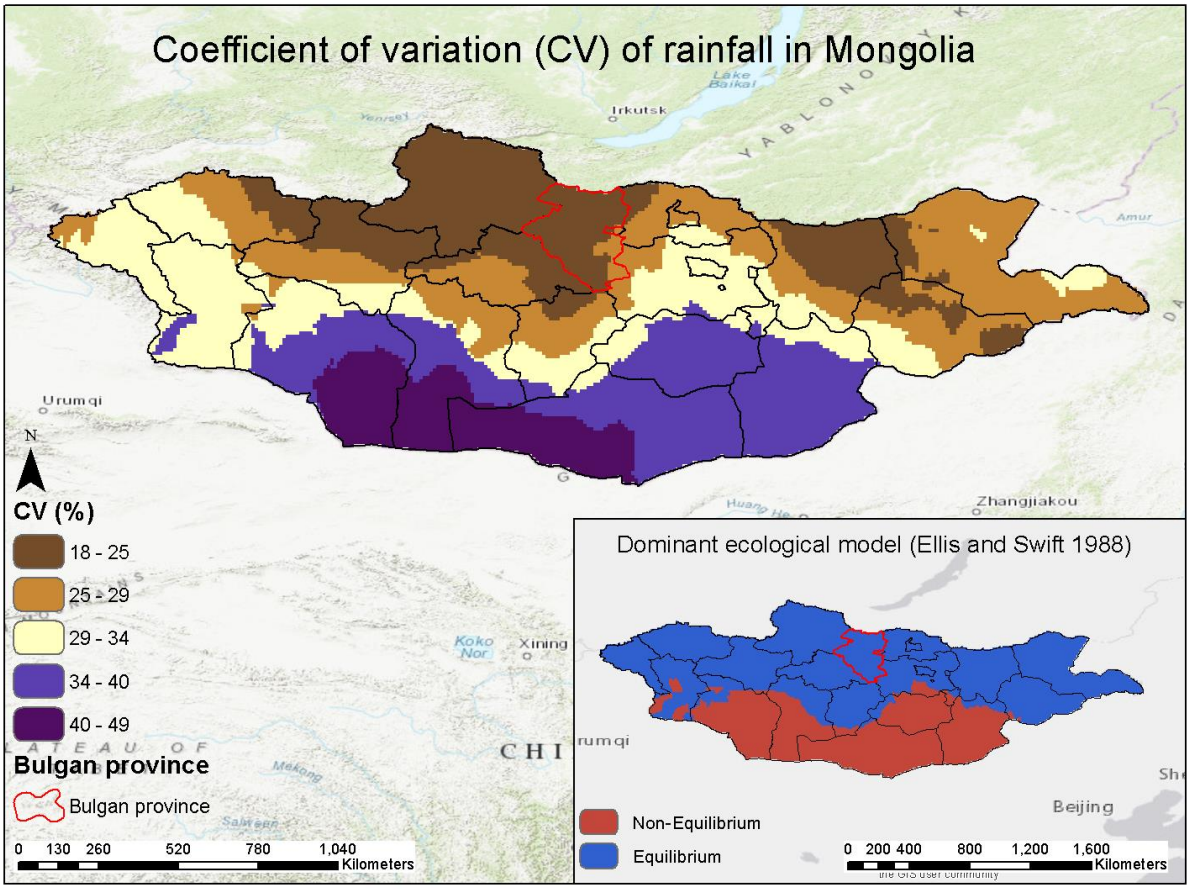


Figure 2: Map of the spatial distribution of coefficient of variation (CV) of precipitation in Mongolia. Inset map shows the resulting divide between equilibrium and non-equilibrium systems according to the 33% threshold as proposed by Ellis and Swift (Ellis & Swift, 1988). Bulgan province is outlined in red. Rainfall variability data obtained from (Wehrden et al., 2012)

4.2. Data

The data used can be categorized as: data and parameters required as input for the process-based biomass prediction model (Table 1), vegetation indices used as input for the empirical biomass prediction model (Table 2), other data required as input for the empirical biomass prediction model (Table 3), parameters required for the adjustment of biomass estimates to sustainably usable values (Table 4) and data and parameters required for forage intake estimation (Table 5). The analysis has been performed for the year 2014, as this was the year for which there were georeferenced biomass sample points and herder locations available.

4.2.1. Process-based model

The remote sensing data used to generate biomass predictions consists of two composite MODIS (Moderate Resolution Imaging Spectroradiometer) products, the MOD17A3HGF v006 Net Primary Production Yearly L4 Global 500 m SIN Grid and the MOD17A2HGF v006 Net Photosynthesis 8-day composite 500 m, and a Landsat 8 derived yearly maximum top-of-atmosphere NDVI composite. Net Primary Production is defined as the rate at which all flora in an ecosystem are producing useful energy, and is thus the difference between net photosynthesis (PsnNet) and woody maintenance plus growth respiration (Zhao, Heinsch, Nemani, & Running, 2005). It is valid to use the sum of PsnNet in this use case, since the target of the analysis and the majority of vegetation in the study area that is of interest for biomass prediction are herbaceous species, which do not grow woody biomass. The MODIS NPP and GPP data were retrieved from the USGS's AppEEARS (Application for Extracting and Exploring Analysis Ready Samples) platform. Other data and parameters used in the process-based model are a mean annual temperature raster, and parameters for the conversion of carbon to biomass and the calculation of fraction of aboveground biomass. The interpolated mean average temperature from 1970 – 2000 was retrieved at 30 arcsec resolution from WorldClim Version 2 (Fick & Hijmans, 2017), and the parameters from IPCC and Hui & Jackson (Hui & Jackson, 2006; IPCC, 2006).

Table 1: Data and parameters required as input for the process-based biomass prediction model.

Process-based biomass estimation			
Data/Parameter	**Value, Unit**	**Type, Resolution**	**Source**
Annual Net Primary Production (NPP)	kg C ha^{-1} yr^{-1}	Raster, 500 m	MOD17A3HGF v006
Eight-day net photosynthesis (PSnNet)	kg C ha^{-1} eight $days^{-1}$	Raster, 500 m	MOD17A2HGF v006
Annual maximum NDVI composite		Raster, 30 m	USGS/Google
Mean Annual Temperature (MAT)	°C	Raster, 30 arcsec	Fick & Hijmans, 2017
Factor for conversion biomass to carbon	0.47	Parameter	IPCC 2006
FANPP and MAT relation	fANPP=0.171+0.0129 MAT	Parameter	Hui & Jackson, 2006

4.2.2. Empirical model

It was chosen to use the 30m resolution Landsat 8 tier 1 surface reflectance product instead of a surface reflectance product with comparable resolution as the MODIS NPP product (250- or 500m resolution MODIS) as input for the empirical model. This was done in order to minimize the effect of within-pixel variance on accuracy assessment of the empirical model, since the georeferenced biomass sample plots were small (1 m^2) compared to the surface area of a 500m resolution MODIS pixel (250,000 m^2), or even a 250m resolution pixel (62,500 m^2).

Table 2: Vegetation indices used as input for the empirical model. Blue (B), Green (G), Red (R), NIR, SWIR1 and SWIR2 indicate surface reflectance SR at wavelengths (µm): B: $SR_{0.452-0.512}$, G: $SR_{0.533-0.590}$, R: $SR_{0.636-0.673}$, NIR: $SR_{0.851-0.879}$, SWIR1: $SR_{1.566-1.651}$, SWIR2 $SR_{2.107-2.294}$. The parameter L denotes the canopy background adjustment term, C_1 and C_2 are coefficients correcting for aerosol influences, Gf is a gain factor. Terms b_{1-6}, g_{1-6} and w_{1-6} are Tasseled Cap indices coefficients. See Devries, Pratihast, Verbesselt, Kooistra, & Herold (2016) for all values used.

Name	Index	Equation	Source
Normalized Difference Vegetation Index	NDVI	$\frac{(\text{NIR}-\text{R})}{(\text{NIR}+\text{R})}$	(Rouse, Schell, & Deering, 1973)
Modified Soil-Adjusted Vegetation Index (v2)	MSAVI2	$\frac{\left(2*\text{NIR}+1-\sqrt{(2*\text{NIR}+1)^2-8*(\text{NIR}-\text{R})}\right)}{2}$	(Qi, Chehbouni, Huete, Kerr, & Sorooshian, 1994)
Soil-Adjusted Vegetation Index	SAVI	$\left(\frac{(\text{NIR}-\text{R})}{(\text{NIR}+\text{R}+\text{L})}\right)*(1+\text{L})$	(Huete, 1988)
Enhanced Vegetation Index (v2)	EVI2	$\text{Gf}*\left(\frac{(\text{NIR}-\text{R})}{(\text{NIR}+C_1*\text{R}-C_2*\text{B}+\text{L})}\right)$	(Huete et al., 2002)
Normalized Difference Moisture Index	NDMI	$\frac{(\text{NIR}-\text{SWIR1})}{(\text{NIR}+\text{SWIR1})}$	(Wilson & Sader, 2002)
Tasseled Cap Brightness	TCB	$b_1\text{B}+b_2\text{R}+b_3\text{G}+b_4\text{NIR}+b_5\text{SWIR1}+b_6\text{SWIR2}$	(Devries et al., 2016; Kauth, 1976)
Tasseled Cap Greenness	TCG	$g_1\text{B}+g_2\text{R}+g_3\text{G}+g_4\text{NIR}+g_5\text{SWIR1}+g_6\text{SWIR2}$	(Devries et al., 2016; Kauth, 1976)
Tasseled Cap Wetness	TCW	$w_1\text{B}+w_2\text{R}+w_3\text{G}+w_4\text{NIR}+w_5\text{SWIR1}+w_6\text{SWIR2}$	(Devries et al., 2016; Kauth, 1976)

Bands 2 to 7 were used from the Landsat 8 monthly composite, which were retrieved from Google Earth Engine. A modified form of the script provided by the Centre for Development and Environment, Bern University, was used to create the monthly composite, providing cloud masking, best-pixel selection based on distance from clouds and 'greenness' levels, as well as correcting for differences in topographic illumination (Hurni, Heinimann, & Würsch, 2017). From the Landsat 8 monthly composite, eight spectral indices were calculated to use as input (Table 2).

In addition, 'time-integrated' values of each of these indices, as well as precipitation, until the approximate date of sampling were calculated and used as input for the empirical models:

$$VIint = \sum_{i=1}^{d} VI$$

where VIint = Integrated Vegetation Index, d= month of biomass sampling and VI = Vegetation Index. It was chosen to use a simple summation of the monthly VI values, rather than compute an integral, as fitting a curve to the monthly VI values per pixel would have been too time- and computation-intensive.

Other datasets used for the estimation of carrying capacity are in-situ georeferenced biomass samples, a DEM raster, a monthly precipitation raster, an ecoregion vector dataset, a land cover vector dataset, and a tree cover percentage raster (Table 3). The tree cover product is at 30m resolution for the year 2010 and available for download from the Global Land Cover Facility (GLCF) (Sexton et al., 2013). A DEM of Bulgan province was retrieved from the Shuttle Radar Topography Mission (SRTM) 90m Digital Elevation Database v4.1 (Reuter, Nelson, & Jarvis, 2007). In-situ biomass measurements were collected by the Institute of Geography and GeoEcology, Mongolian Academy of Sciences, and are available for the year 2014 (120 samples) as point data. The dataset contains attributes for the date of collection, administrative area, biomass, grazing pressure and landcover type. The sampling technique used was purposive sampling, in which only grassland samples were taken from relatively homogeneous areas. Care was taken to include areas with varying levels of grazing pressure. The sample plots were 1 m^2 in area, from which the aboveground biomass was harvested, dried, and weighed. Monthly precipitation data was retrieved using Google Earth Engine from the TerraClimate Monthy Climate and Climatic Water Balance for Global Terrestrial Surfaces dataset at 2.5 arcmin resolution (Abatzoglou, Dobrowski, Parks, & Hegewisch, 2018). The ecoregion data was retrieved using Google Earth Engine from the RESOLVE Ecoregions dataset (Dinerstein et al., 2017). Although the ecoregion dataset originally included three ecoregions intersecting Bulgan province, it was chosen to merge the two most northern ecoregions (Sayan Montane Conifer forests and the Selenge-Orkhon forest steppe) into one ecoregion, 'Forest-

steppe'. This was done because there were too few biomass samples located in the Sayan Montane Conifer forest ecoregion. The landcover data was retrieved from Copernicus, filtered and converted to vector format via Google Earth Engine before download (ESA, 2018).

Table 3: Data required as input for the empirical biomass prediction model.

Empirical biomass estimation			
Data/Parameter	**Value, Unit**	**Type, Resolution**	**Source**
Monthly median composite Landsat 8, Bands 2-7		Raster, 30m	USGS/Google
DEM	m	Raster, 90m	SRTM
Ecoregion		Vector (polygon)	RESOLVE Ecoregions 2017
Tree cover	Tree cover %	Raster, 30m	GLCF/Google
Land cover		Raster, 100m	CGLS-LC100 collection 2/Google
Monthly precipitation	mm	Raster, 2.5 arcmin	TerraClimate/Google
Georeferenced biomass samples	dry weight, g m^{-2}	Vector (point)	Mongolian Academy of Sciences

4.2.3. Sustainably usable biomass

For the conversion of predicted biomass values to sustainably usable values, several parameters were used from de Leeuw et al., (2019) such as a 'proper use' factor, and 'proper use' factor reduction based on slope percentage (Table 4).

Table 4: Parameters required for the adjustment of biomass estimates to sustainably usable biomass values.

Biomass proper use estimation			
Data/Parameter	**Value, Unit**	**Type, Resolution**	**Source**
Proper use factor	0.65	Parameter	(de Leeuw et al., 2019)
Reduction proper use factor for slopes 10-30%	30%	Parameter	(de Leeuw et al., 2019)
Reduction proper use factor for slopes 30-60%	60%	Parameter	(de Leeuw et al., 2019)
Reduction proper use factor for slopes > 60%	100%	Parameter	(de Leeuw et al., 2019)

4.2.4. Forage Intake

For the estimation of forage intake, generally two types of data were required: herder locations, and sub-district statistics on livestock numbers (Table 5). A dataset containing the seasonal locations of herders as points for the 2014 year was collected by the Institute of Geography and GeoEcology, Mongolian Academy of Sciences. Attributes included are province, Soum (district administrative level), Bag (sub-district administrative level), season and herder names. Sub-district statistics on livestock numbers were also provided for the year 2014 by the Institute of Geography and GeoEcology, and contained attributes on season, district, sub-district, herder names, livestock numbers by animal type, and total livestock numbers in sheep-equivalent units (SEU). Due to missing and/or incomplete data in some areas, census data on livestock numbers by subdistrict and livestock type was retrieved from the Mongolian Statistical Information Service portal in order to supplement the seasonal data where possible (National Statistics Office of Mongolia, 2019). Finally, for the final calculation of forage intake, several parameters are required such as a SEU conversion matrix, with which different livestock types can be converted and represented in SEU (Appendix A), a grazing range matrix, representing the distance typically grazed per season per livestock type (Appendix A), a forage intake parameter, denoting the daily forage intake requirement per SEU, and duration of stay parameters, which represent the duration that herders generally stay per location, per season.

Table 5: Data and parameters required for forage intake estimation

Forage intake estimation			
Data/Parameter	**Value, Unit**	**Type, Resolution**	**Source**
Administrative boundaries		Vector (polygon)	Mongolian Academy of Sciences
Georeferenced seasonal herder locations		Vector (point)	Mongolian Academy of Sciences
Seasonal livestock per herder group and subdistrict		Table	Mongolian Academy of Sciences
Livestock per subdistrict		Table	National Statistics Office of Mongolia
Conversion to SEU matrix		Table/Parameter	(Sainbuyan, 2016)
Grazing range matrix		Table/Parameter	(Sainbuyan, 2016)
Forage intake	1.5 Kg SEU^{-1} day^{-1}	Parameter	(Sainbuyan, 2016)
Duration of stay summer pastures	90 days	Parameter	Mongolian Academy of Sciences
Duration of stay summer + autumn pastures	150 days	Parameter	Mongolian Academy of Sciences

4.3. Models

Two types of models will be discussed in this section, which for the sake of consistency will be identified henceforth as the 'Process-based' models and the 'Empirical' models. It is important to note that, while the process-based model is identified as such, it does contain empirical elements, such as the parameters in the fraction of aboveground biomass formula, which were derived through empirical methods.

4.3.1. Process-based model

The process-based model for carrying capacity estimation is based on the model developed by de Leeuw et al. (de Leeuw et al., 2019). The main inputs necessary are the MODIS NPP product, and the WorldClim mean annual temperature (Figure 3). NPP values can be converted to total biomass using a conversion factor of 0.47 (IPCC, 2006). Then, the aboveground biomass is derived from the product of the fraction of aboveground biomass (fANPP) and the converted NPP values.

$$AB = NPP \times fANPP$$

where AB = aboveground biomass, NPP = Net Primary Productivity and fANPP = fraction of Aboveground Biomass. fANPP is related to mean annual temperature through the following equation (Hui & Jackson, 2006).

$$fANPP = 0.171 + 0.0129MAT$$

where MAT = Mean Annual Temperature. In addition, in order to minimize the effect of biomass growth that had occurred past the sampling date, the model can potentially be adjusted by using the sum of PsnNet up to the sampling date instead of yearly NPP (Figure 4):

$$AB = \sum_{i=1}^{d} PsnNet \times fANPP$$

where d = number of eight-week periods up to the sampling date and PsnNet = Net Photosynthesis.

Due to the relatively low resolution of the MODIS NPP product compared to the in-situ biomass sampling plots (500 m vs. 1 m), it is suggested by de Leeuw et al. to employ high-resolution Landsat images to correct for possible landscape heterogeneity within a MODIS pixel when assessing model accuracy (de Leeuw et al., 2019). This will be elaborated upon in the section on validation.

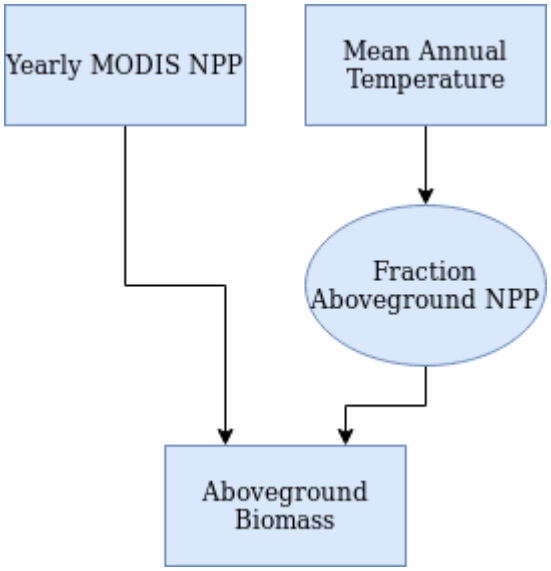


Figure 3: Schematic representing the process-based model for aboveground biomass calculation. NPP = Net Primary Production

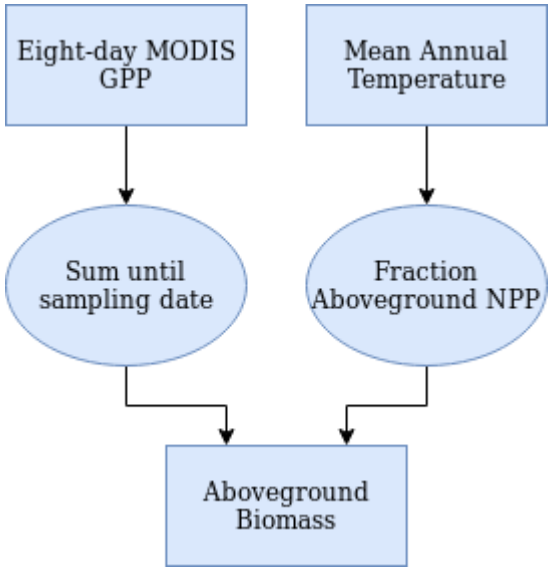


Figure 4: Schematic representing the process-based model for aboveground biomass calculation using the sum of eight-day GPP. NPP = Net Primary Production

4.3.2. Empirical model

An empirical model was chosen for biomass estimation after evaluating the performances of two machine learning methods, Random Forests and Cubist. The input data for the models were the same for each model (Tables 2 & 3).

Random Forest

Random Forest (RF) is a machine learning method based on the concept of decision trees (Breiman, 2001). Decision trees themselves often perform poorly because they do not extend well to prediction using unseen data, thus beyond the data for which they were trained. However, RF addresses this issue through several methods. First, RF 'bootstraps' datasets, by sampling from the training dataset with replacement. Secondly, RF randomly selects variables at each node, until each decision tree is complete. Thirdly, this process is repeated for n number of trees. The predictions, in the use case of regression, are the means of the predications of each tree. The advantages of RF is that it is relatively efficient, it can handle large numbers of input variables, it is a flexible method that generally offers a good trade-off between bias and variance (Segal, 2004). In addition, RF has been used with success for the estimation of a wide range of biomass types using remote sensing data. For example, it has

been employed for the estimation of wheat biomass with remote sensing data as input by Wang et al. (Wang et al., 2016), the estimation of woody biomass using Landsat data (Avitabile, Baccini, Friedl, & Schmullius, 2012), and high density biomass estimation in wetland areas using WorldView2 data (Mutanga, Adam, & Cho, 2012). However, it must be noted that RF does occasionally suffer from overfitting when faced with noisy tasks, especially when the dataset is relatively small (Segal, 2004).

Cubist

Cubist is similar to tree-based models in that a tree is grown with linear regression models at the terminal nodes. Each intermediate step also contains a linear model. Then, all paths through a decision tree are 'flattened' into rules, which are pruned and combined for simplification. Cubist is different from other tree-based models in several ways: most importantly, a boosting-like technique called 'committees' can be used, and predictions can be adjusted using instance-based corrections (Kuhn et al., 2013). The advantages of using Cubist for predictions are that it is robust when faced with noisy data (as is often the case with remote sensing) due to the use of committees and instance-based corrections, and that it can handle non-linear patterns in the data, while remaining relatively interpretable (John et al., 2018). In addition, it can handle different forms of input data besides numerical, such as factors. Cubist has been successfully used for rangeland productivity biomass estimation in several studies, such as the development of a 30m grassland productivity estimation for central Nebraska by Gu & Wylie (Gu & Wylie, 2015) and for a study estimating grassland biomass and canopy cover in Mongolia (John et al., 2018).

Model tuning and selection

The steps taken to create the models was as follows: First in order to assess the general effectiveness of each model compared to one another, the models were compared using a so-called 'benchmark'. In a 'benchmark' test, the models can be compared using the same resampling instances, thus eliminating variance otherwise introduced from differences in cross-validation sets. Using pre-defined hyperparameter search spaces, both models were allocated 200 instances for which random combinations of hyperparameters were selected to use in model training: this is dubbed a 'random search algorithm'. These have been shown to offer a good trade-off between finding optimal hyperparameter combinations and computational load (Bergstra & Bengio, 2012). For each 200 iterations of random hyperparameters, the best combination was selected using spatial cross-validation; this is called the inner resampling loop. This entire process was then repeated in an outer resampling loop using repeated spatial cross-validation, comparing the best models between the two learning methods, RF and Cubist. Performance metrics were then calculated using the mean performances of the outer resampling loop. If one model performs much better than the others, it could be decided to use only this one for further training. However, in this case there did not seem to be a significant difference at that point between learner performances. Therefore, it was chosen to continue with both methods.

Again, hyperparameter search spaces were chosen for each learning method. However, the hyperparameter tuning was performed with repeated spatial cross-validation and with a budget of 5000 random iterations, rather than the single spatial-cross validation and the 200 random iterations used in the benchmark test. After hyperparameter tuning, the best hyperparameter combination was extracted and used to set the hyperparameters of a model. This model was then resampled using repeated spatial cross-validation in order to reduce bias in the performance estimation that would have been introduced if the performance was solely estimated based on tuning performance results. Finally, the model performance was estimated and compared between learning methods. In addition, scatterplots were created of predicted vs ground-truth values using the mean predictions from the outer resampling loop in order to visually compare the overall performances between the RF and Cubist methods.

4.4. Validation

4.4.1. Cross-validation

The most important method to prevent overfitting of the empirical models, as well as provide indications of model prediction accuracy was repeated spatial cross-validation. This method is based on the well-known method k-fold cross-validation, in which data is randomly split into *k* folds. Each model is trained using the *k* – 1 folds, and then validated using the leftover *k*th fold. This is then repeated until every fold has served as the validation set. The mean performance of the test scores is then taken as the overall performance metric for the model. However, this can lead to issues when applied to spatial data, as spatial data is likely affected by spatial autocorrelation. This means that datapoints located closer to one another are more likely to be similar; therefore, not accounting for spatial autocorrelation often leads to overoptimistic model performance (Brenning, 2005). By using spatial cross-validation, it is avoided that training and test folds are located close to one another, splitting the data into spatially disjoint subsets (Brenning, 2005). Repeating this process further reduces variance caused by partitioning: this is known as repeated spatial cross-validation.

4.4.2. NDVI-adjustment of process-based model

Due to the relatively large difference between the area of a MODIS NPP pixel (250,000m^2) and the area of a biomass sampling plot (1m^2), it was suggested by de Leeuw et al. (2019) to perform an adjustment of the observed biomass, so that these would be a better reflection of the true biomass which one MODIS pixel encompassed. Although the sampling in this study was performed purposefully, with sampling plots being chosen in relatively homogeneous areas, it is possible that this did not occur everywhere. Potential prediction errors could be attributed to the biomass sampling plot not correctly reflecting the mean biomass of the pixel in which it is located, rather than unexplained variation in the true phenomena linked to biomass levels which the model could not capture. For example, if the MODIS pixel in which the sampling plot was located also included a permanent body of water, deciduous forest, or rocky outcrops, this would of course not be reflected in the biomass sampling plot, resulting in measurement error. Therefore, it could be worthwhile to perform an adjustment to account for this. This can be done by incorporating a relatively high-resolution 30m resolution

(compared to the 500m resolution MODIS pixels) yearly NDVI composite. Assuming that there exists a relationship between NDVI and biomass, and that the interfering, non-grassland areas within a MODIS pixel have a significantly different mean biomass value, and thus NDVI signature, the observed biomass value could be adjusted by the ratio between the NDVI of the smaller, high-resolution pixel and the mean NDVI of the large, low-resolution MODIS pixel. If the values of the error metrics of model predictions are then lower, it could be a sign that there is some effect of the mismatch between spatial resolutions of sampling versus prediction areas.

4.4.3. Performance Metrics

The performance metrics that were chosen to evaluate the accuracy of model biomass predictions were Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and comparison of means using the paired Wilcoxon signed-rank test. MAE is a measure for overall magnitude of prediction errors, without considering their direction. It is calculated by taking the average of the absolute differences between observed values and predicted values (Willmott & Matsuura, 2005). RMSE is also a measure of magnitude of error; it is the square root of the mean of squared differences between the observed and predicted values (Wilmott & Matsuura, 2005). In practice, this means that RMSE is proportionally more sensitive to larger errors than it is too smaller errors, while errors are weighted equally by MAE (Wilmott & Matsuura, 2005). This difference between MAE and RMSE is why these metrics were chosen to estimate and compare the performance of the empirical and process-based models, as the relation between the two metrics allows for a better interpretation of model prediction behavior, than if only one metric was used (Wilmott & Matsuura, 2005). Next, a statistical test, namely the Wilcoxon signed-rank test, was performed to determine if the predicted values of the models differed significantly from the observed values. This test is often regarded as the non-parametric equivalent of the paired sample t-test. However, the assumptions about the distributions of the samples are different; while the paired sample t-test requires the dependent variable to be normally distributed, the Wilcoxon test only requires the differences between the independent and dependent variables to be symmetrically distributed (Wilcoxon, 1945). De Leeuw et al. (2019) used a paired sample t-test to compare observed and predicted biomass values; however, after checking the normality of the predicted values in our study

using the Shapiro-Wilk test of normality, it could not be concluded that the predicted values followed a normal distribution, therefore the paired sample t-test could not be used in this case.

4.5. Workflow

The workflow of all data processing can be divided into roughly four parts (Figure 5). Part 1 deals with the workflow of the prediction of aboveground biomass using the process-based model. Part 2 describes the workflow of aboveground biomass prediction using the empirical model. In part 3, it is described how the data processing for the estimation of forage intake was done. Finally, part 4 describes the steps taken to overlay the results of the previous parts.

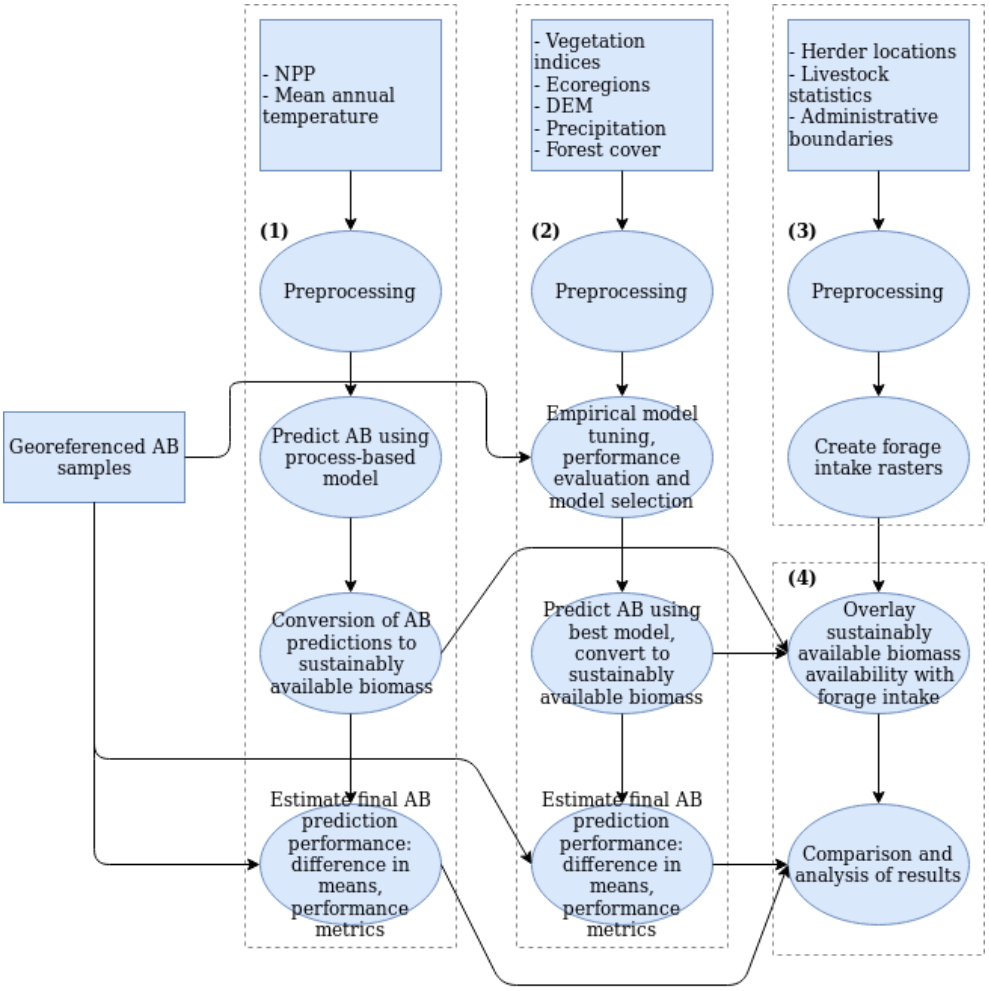


Figure 5: Workflow schematic showing the steps that were taken in data processing. NPP = Net primary production, DEM = digital elevation model, AB = aboveground biomass. More detailed schematics including all steps taken can be found in Appendix B.

4.5.1. Aboveground biomass prediction using empirical model

To start, the no data-values in the Landsat 8 terrain-corrected median monthly composite were set to NA, as they were originally set to 0. Then, the images were cropped to the study area extent and reprojected to UTM Zone 48N projection. These were then split into 1500x1500 pixel tiles to allow for the implementation of parallel processing to reduce overall computation length. For each tile, the vegetation indices (Table 2) were calculated, and capped to a -1, 1 range if applicable. The resulting vegetation indices tiles were then mosaiced, and in QGIS a manual check was carried out using the eight-day NPP composites as an overlay to approximate when the growing season started and ended (Assumed was an NPP value above 0 meant the start of the growing season, and the inverse for the end of the growing season). In the first and last month of the pre-mosaiced tiles corresponding with growing season months, NA-values were reclassified to the mean value of each vegetation index. This was done so that a linear time-series interpolation could be carried out in the remaining months if values were missing, important for the calculation of time-integrated vegetation indices. After the interpolation of NA-values in the vegetation index tile raster stacks, the tiles were mosaiced. From these stacks, the time-integrated vegetation index rasters using values up to the sampling date were generated. After all other input data for the process-based model had been cropped, reprojected and resampled to the same resolution as the vegetation indices rasters, all rasters were stacked. Next, all values at the locations of the georeferenced biomass sample points were extracted. All variables except for the ecoregions were rescaled to a 0,1 range. Before the prediction could be performed, the raster stack with all predictor variables were split into 1000x1000 pixel tiles to increase processing speed. After applying the model to each tile, all tiles were mosaiced to create one predicted aboveground biomass raster at 30m resolution covering the entire study area. Then, using the Copernicus land cover product, which had been filtered to only include areas that were not categorized as grassland/shrubland, the biomass raster was inversely masked to include only grassland areas. Finally, performance metrics were calculated over all predictions.

4.5.2. Aboveground biomass prediction using process-based model

First, the MODIS NPP yearly composite and WorldClim mean yearly temperature rasters were cropped to the study area and reprojected to UTM zone 48N projection at 500m resolution.

These were then stacked, and an aboveground biomass raster was generated using the process-based model as described in section 4.3.1.

To gain information of the variability of NPP within the 500m resolution pixels compared to the smaller $1m^2$ biomass plot size, several additional steps were taken, as described in 4.4.2: the georeferenced biomass sample points were used to mask the yearly NPP composite raster. The remaining pixels were then converted to polygons. These polygons were then used to extract the mean NDVI values from a 30m resolution, yearly maximum NDVI composite, as well as the standard deviation of the Landsat NDVI values per 500m resolution pixel. The mean NDVI value of the entire 500m pixel, the NDVI value of the 30m resolution Landsat raster, the standard deviation of Landsat NDVI values within each 500m pixel, as well as the predicted biomass were then extracted at the location of each biomass sample point location. The ratio between the mean NDVI value at 500m resolution, and the NDVI value at 30m resolution was then calculated and used to create an adjusted ground-truth aboveground biomass column.

Besides investigating spatial variability within each pixel, it was also investigated if it was worth adapting the process-based model to use only NPP values up until the date of sampling, as described in 4.3.1. For this, the MODIS eight-day NPP composite was used to extract NPP values at the biomass sample point locations. The sum of NPP until the approximate date of sampling was then calculated. The resulting NPP values were then used in conjunction with the mean annual temperature to calculate aboveground biomass, which was masked using the Copernicus land cover to include only grassland areas. Finally, performance metrics were calculated for the adjusted and unadjusted biomass values, yearly and eight-day composite biomass prediction values, and the best performing results were used for further processing and analysis.

4.5.3. Forage intake prediction

For the creation of the forage intake rasters, the livestock statistics obtained from the Institute of Geosciences, Mongolia, first had to be cleaned, and subdistrict identification codes added. Then, only livestock numbers recorded in the summer season were selected. This was done because reliable estimates of biomass could not be made in other seasons due to the biomass sampling occurring only in the summer of 2014. Using a conversion matrix, livestock figures were converted to sheep equivalent units (SEU), before being summed per subdistrict. These

sums per subdistrict were then spatially joined to herder location points which had been filtered to only contain locations of herders in the summer, and summer through autumn locations (many herders do not change locations for autumn). Some subdistricts contained information on livestock but did not contain herder locations: these were mostly district centers. For these district centers, 'herder locations' were generated at their centroids. Next, the SEU per subdistrict were equally divided among the herder location points within that subdistrict. Then, livestock was divided into 'large' and 'small' livestock, and buffers were created around the herder location points with a radius according to livestock size using the grazing range matrix (Appendix A). The buffers were then exported to QGIS and a difference operation was carried out, using the vectorized 100m resolution Copernicus land cover product-based grassland/shrubland areas. The resulting masked buffers were re-imported to R. From the masked buffers, the SEU per cell, for both 30m and 500m resolution were calculated by dividing the SEU total per buffer by the approximate number of pixels each buffer would occupy (= buffer area divided by pixel size). Then the buffers were rasterized, using the previously generated 30m and 500m aboveground biomass rasters as templates, and the SEU per cell value raster values. Finally, the small- and large-radius buffer rasters were summed and converted to SEU ha values.

4.5.4. Further processing and overlay of outputs

Since the aboveground biomass prediction rasters did not account for the proportion of biomass that could be used sustainably, they had to be adjusted by a 'proper use' factor according to de Leeuw et al. (2019). This proper use factor takes into account the proportion of biomass which gets trampled by livestock, as well the proportion of biomass which must remain after grazing in order to prevent land degradation from occurring due to erosion or replacement of palatable forage species (de Leeuw et al., 2019). In addition, it also takes slope steepness into account. Therefore, the proper use factor of 0.65 was adjusted by a factor based on the slope steepness as described in 4.2.3. The resulting 'proper use' adjustment factor raster was then multiplied with the predicted aboveground biomass rasters in order to obtain sustainably usable biomass availability rasters, at 30m and 500m resolution. For the absolute shortage/oversupply of biomass, defined as the difference between sustainably usable biomass and forage intake, the forage intake rasters were subtracted from the biomass

availability rasters. For the relative intake of biomass, the forage intake rasters were divided by the biomass availability rasters. Finally, summary statistics were calculated for the entire area, as well as per district and ecoregion, and the overall results were analyzed and visualized.

4.6. Software

All data processing, except if otherwise mentioned in the workflow section was done in R (v3.6.3 The model training and resampling was performed using *mlr* (v2.17.1) (Bischl et al., 2016), *caret* (v6.0-86) (Kuhn et al., 2013), *ranger* (v0.12.1) (Wright & Ziegler, 2017), and *Cubist* (v0.2.3) (Kuhn et al., 2013).

All bar charts and graphs were produced in R using base R, as well as the *ggplot2* package (v3.3.0). All maps were made in Qgis.

5. Results

5.1. Process-based model

5.1.1. Performance

According to the calculated performance metrics, the performance of the process-based model was not significantly better than that of the Null model, which used the average of all observed biomass values as biomass predictions (Table 11). Note that the observed mean is slightly different here than for the empirical model; this is due to the masking of the input data with non-grassland areas covering a pixel in which a georeferenced biomass sample was located, therefore removing it from the training dataset. The performance of the process-based model using the sum of eight-day NPP composites did not increase significantly over the process-based model which used the yearly NPP composite according to the RMSE (Table 6). However, the MAE, or the average magnitude of error decreased greatly, from 39.48 to 20.20. Scatterplots were also compiled of the predicted versus observed biomass values of both the model using the yearly composite and the model using the sum of eight-day composites (Figure 6 & 7). While the yearly composite model generally overpredicts, the eight-day

composite model both over- and underpredicts. The lower MAE of the eight-day composite model translates to a narrower range of predicted values overall, without necessarily having a higher predictive power, as the predictions do not deviate greatly from the prediction mean of 62.55 g m^{-2}.

Table 6: Performance of the process-based model compared with the performance of a Null model. MAE = Mean Absolute Error, RMSE = Root Mean Square Error. Mean values shown with a 95% confidence interval.

Model	MAE	RMSE	Mean	Observed Mean
Process-Based, yearly composite	39.48	49.49	97.58 ± 4.55 g m^{-2}	82.15 ± 9.63 g m^{-2}
Process-Based, eight-day composites	20.20	51.57	62.55 ± 2.76 g m^{-2}	
Null	36.94	50.3	82.15 g m^{-2}	

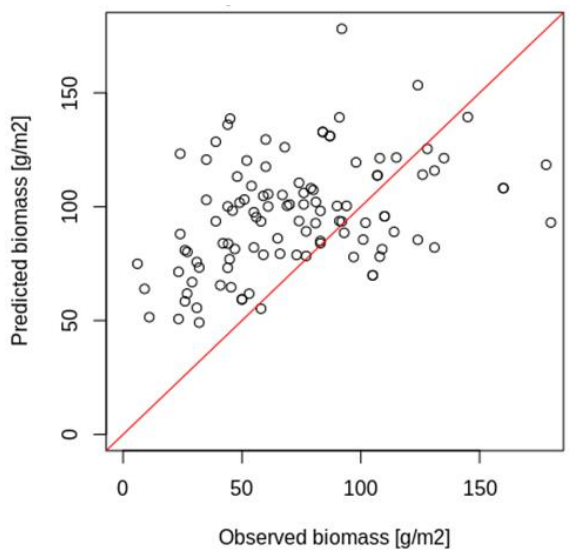


Figure 6: Scatterplot of observed biomass vs predicted biomass using the yearly NPP composite.

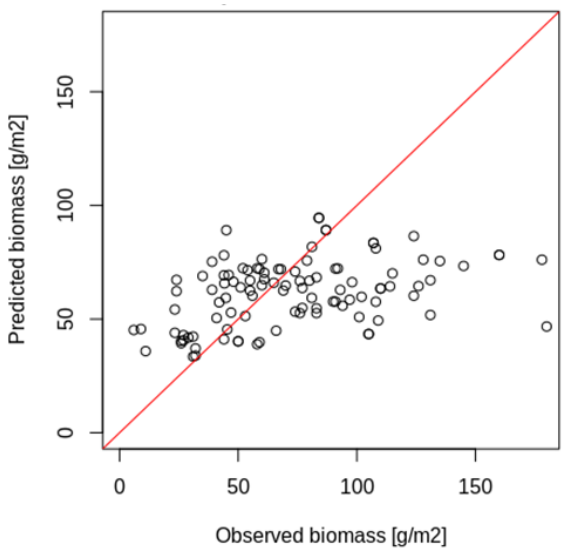


Figure 7: Scatterplot of observed biomass vs predicted biomass using sum of eight-day NPP composites.

Due to the relatively high difference in area between the sampling plots and the pixels they were located in, an adjustment of the observed biomass values was carried out as proposed in section 4.4.2. Inspecting the pixel with the highest internal variance, as well as the pixel with the lowest internal variance shows the potential for adjusting the ground truth biomass plot values for spatial variability within the pixels they are located in (Figure 8).

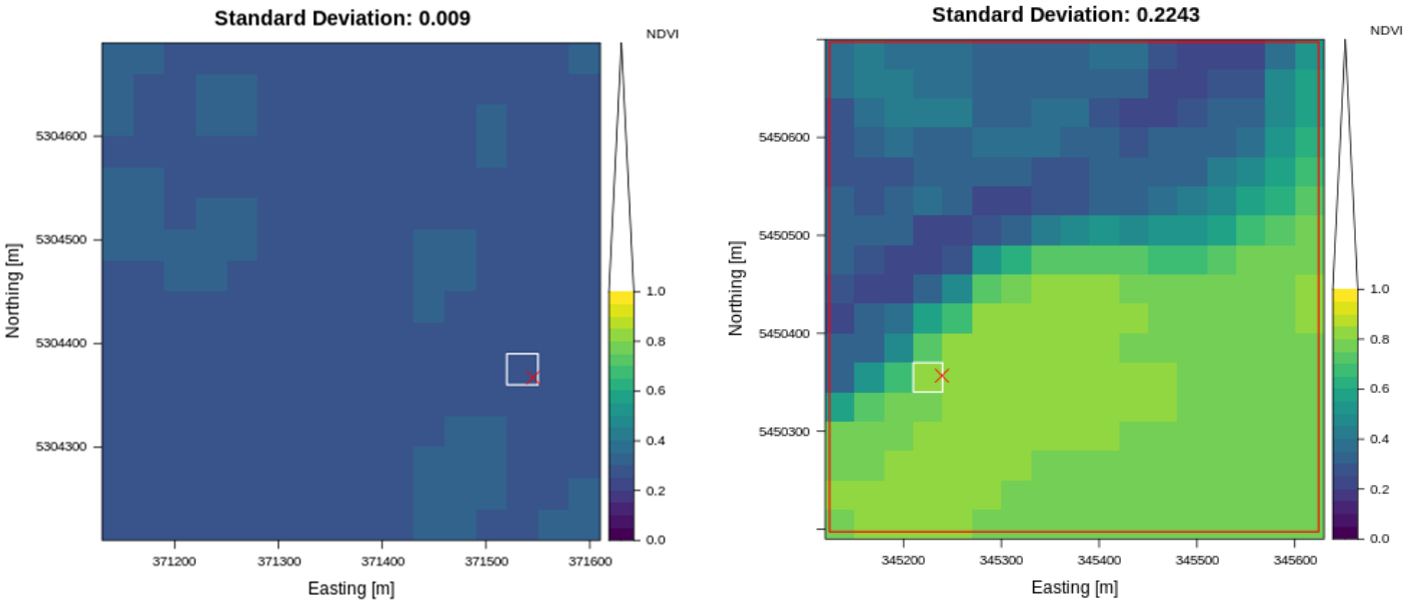


Figure 8: 500m resolution Modis pixel with the lowest (left) and highest (right) within-pixel variation of 30m resolution Landsat NDVI pixels. The red cross indicates the location of the biomass sampling location. The white box indicates the Landsat pixel the biomass sample falls within.

With this adjustment, the mean of the observed biomass increased slightly, from 82.03 g m^{-2} to 84.06 g m^{-2}, therefore reducing the overprediction from 15.8% to 13.9%. However, this did not significantly affect the accuracy of the model predictions, as the RMSE was increased from 49.49 to 50.76 (yearly composite), and 51.57 to 54.53 (sum of eight-day composites) (Table 7).

Table 7: Performance of the process-based model with NDVI-adjusted observed biomass values.

Model	MAE	RMSE	Mean	Observed Mean
Process-Based, yearly composites (NDVI adjusted)	39.07	50.76	97.58 ± 4.55 g m^{-2}	84.06 g m^{-2}
Process-Based, eight-day composites (NDVI-adjusted)	21.99	54.53	62.55 ± 2.76 g m^{-2}	

The Wilcoxon signed-rank tests were performed for the Process-based models, to test whether the mean of the predictions was significantly different from the mean of the observed values. In all cases, including the use of NDVI-adjusted sample biomass values, a significant difference in means could be found (Table 8). Since the use of the sum of eight-day NPP composites did

not significantly improve the predictions of the Process-based model, biomass predictions with the Process-based model were carried out using the yearly NPP composite.

Table 8: Results of the Wilcoxon sign rank test for differences of means. The '' indicates that the P-value is lower than 0.05, thus disproving the null-hypothesis that there is no significant difference in means.*

Input data	NDVI-Adjusted	P-value
Yearly NPP composite	No	$1.202e^{-05}$*
Eight-day NPP composite	No	$2.202e^{-05}$*
Yearly NPP composite	Yes	$5.196\ e^{-4}$*
Eight-day NPP composite	Yes	$2.312\ e^{-4}$*

5.1.2. Biomass predictions: Process-Based model

The sustainably usable biomass prediction values by the process-based model were highest in the northern forest steppes, with most values exceeding 700 kg ha in these areas (Figure 9). The lowest values were found in the south, the Mongolian-Manchurian grassland ecoregion (Figure 9).

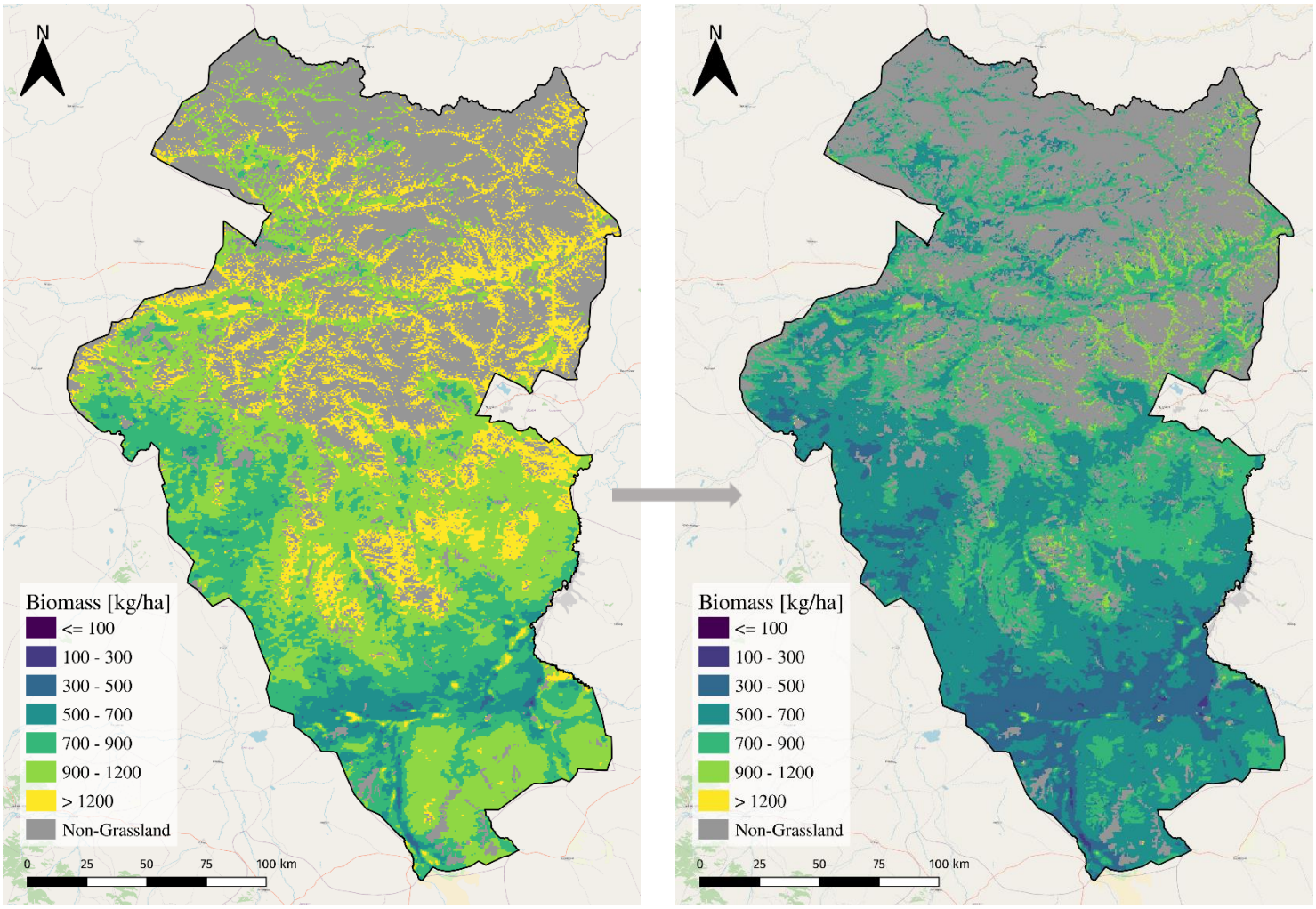


Figure 9: Prediction of aboveground biomass [kg/ha] (left), and biomass adjusted for sustainable use [kg/ha] (right) for Bulgan province using the process-based model.

The three districts with the highest mean predicted sustainable biomass values were Selenge, Xangal, and Bugat, with 886, 862, and 777 kg ha biomass prediction means. The three lowest mean predictions values were found in Gurvanbulag, Saixan, and Dashinchilen, with 509, 566, and 575 kg ha sustainably usable biomass respectively (Figure 11). The confidence interval of

the observed values, visible as black bars in figure 10, show that there were often only a low number of observed values per district, which resulted in relatively wide confidence intervals. Therefore, it is difficult to make conclusions on the performance of the model within certain districts, such as Bayannuur (wide confidence interval), Bulgan (only one sampling point) or Rashaant (no sampling points) (Figure 10). With the resulting overall mean sustainably usable biomass figures, the overall theoretical carrying capacity can be calculated. According to the process-based model, the total carrying capacity, or the number of sheep-equivalent units that can sustainably graze in a 150-day period in Bulgan province is 9.3 million SEU.

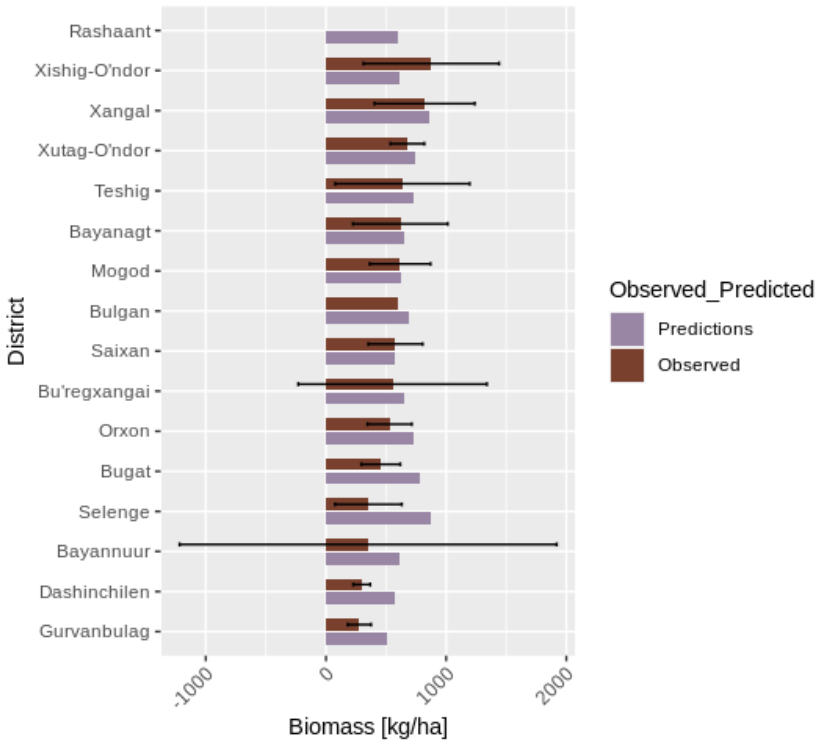


Figure 10: Sustainably usable biomass predictions compared to (converted to sustainable useable) observed biomass values, per district using the process-based model. The black bars represent the 95% confidence interval of the observed values.

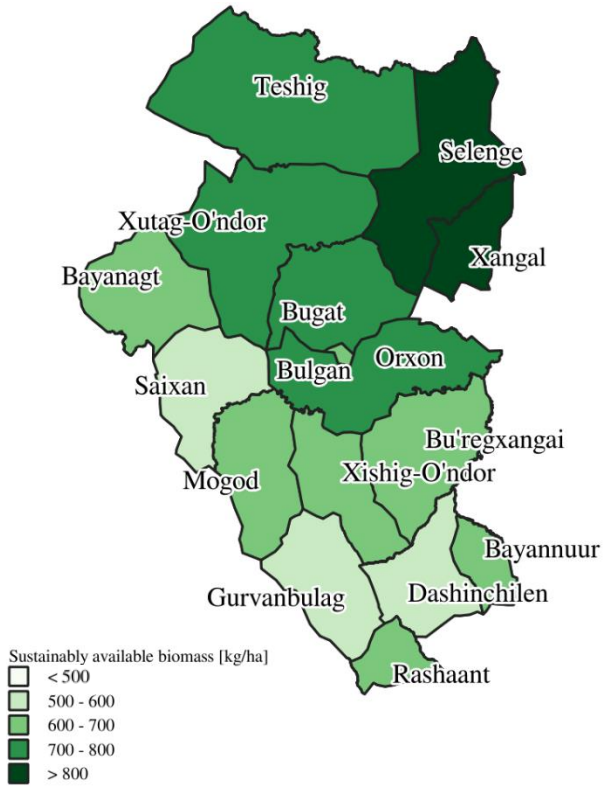


Figure 11: Sustainably usable biomass predictions per district; process-based model.

5.2. Empirical model

5.2.1. Exploratory analysis

Before the training of models and model selection, an exploratory analysis was carried out of the chosen vegetation indices and other input variables. The correlations between input variables were calculated using the Spearman method, since the normality of these could not be confirmed using the Shapiro-Wilk test. The three input variables with the highest correlation with biomass were NDVI (0.48), TCW (0.45) and EVI (0.44). Generally, the time-integrated vegetation indices were not significantly correlated with biomass (Figure 12). The highest correlation non-vegetation index input variables were precipitation and time-integrated precipitation (0.42 each). From the scatterplots of the highest correlated vegetation index (NDVI) and non-vegetation index input variables (precipitation), it is visible that at high (> 150 g/m^2) biomass values, the correlation between NDVI and biomass, as well as precipitation and biomass is not strong (Figure 13). In addition, NDVI and precipitation values seem to 'taper off' with high corresponding biomass values, indicating non-linear relationships (Figure 13).

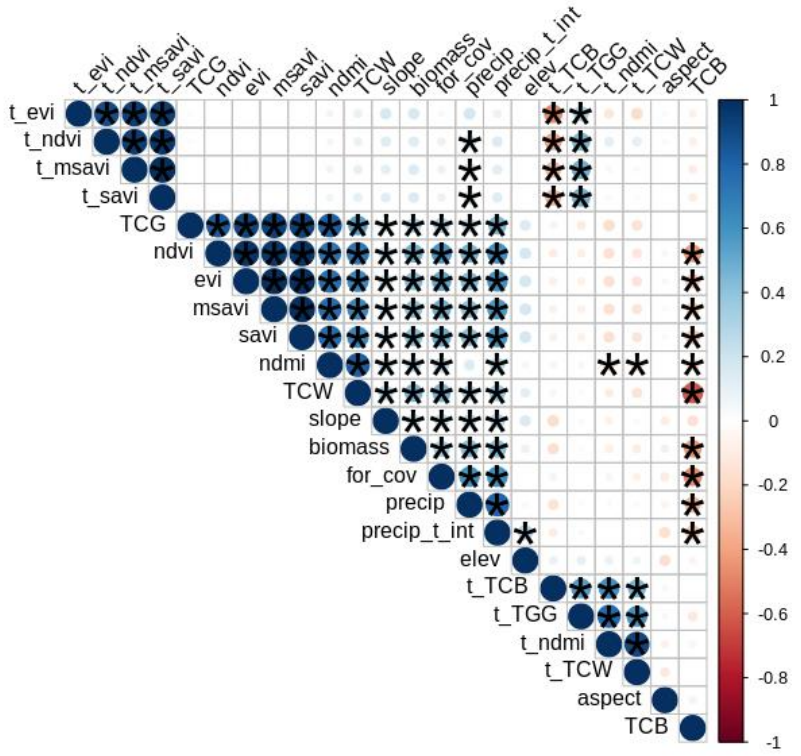


Figure 12: Correlation between all input variables except ecoregion type, and between these input variables and biomass. Red indicates a negative correlation, while blue indicates a positive correlation. The size of the circles indicates the magnitude of the correlation. The asterisks are placed on statistically significant correlations (p<0.05).

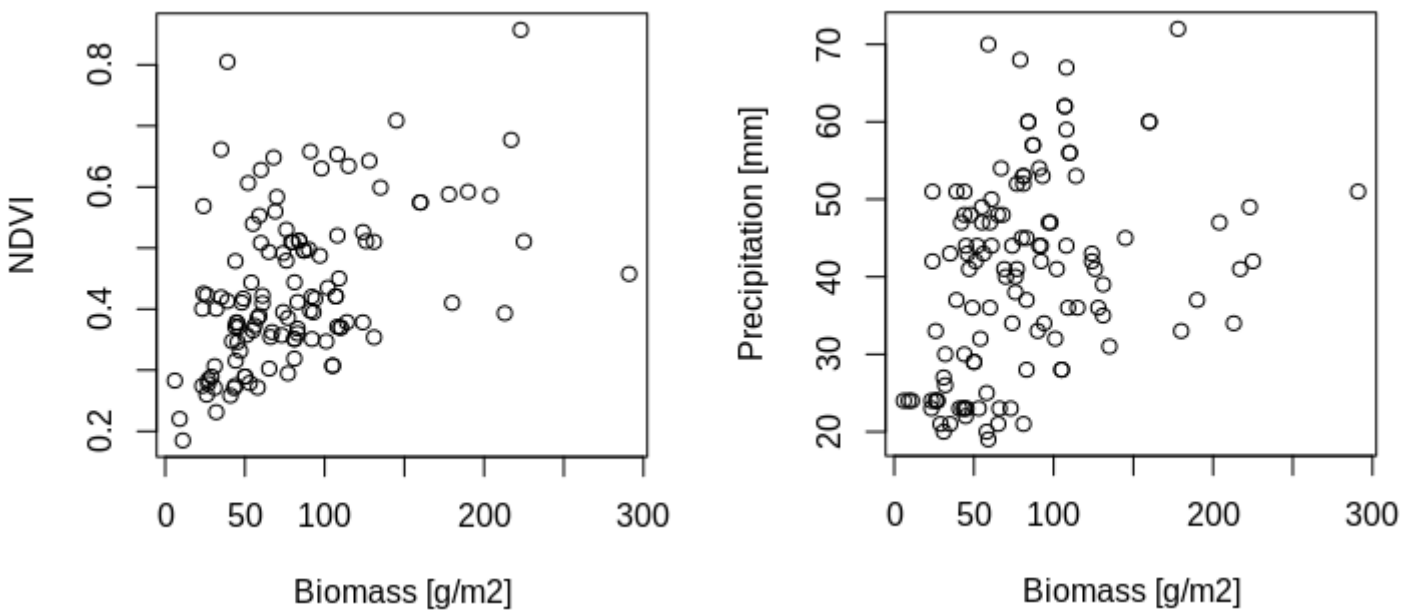


Figure 13: Scatterplots of biomass versus NDVI (left) and biomass versus precipitation (right).

5.2.2. Model selection and performance

From the initial 'benchmark' comparison, no significant difference in predictive performance could be observed between the Random Forest and Cubist methods (Table 9). In addition, the models both did not perform much better than the 'Null' model, which simply takes the mean of the training data as prediction values. Both scaled and non-scaled predictors were tested. For the comparison, a 5-fold spatial cross-validation was used for the hyperparameter tuning, while a 5-fold, 25-repetition repeated spatial cross-validation was used for the performance estimation. The hyperparameter search algorithm was allocated 200 iterations per fold for tuning. The hyperparameters and their respective search spaces used can be seen in Table 10.

Table 9: Results of 'benchmark' comparison between the proposed machine learning methods and a null model, which takes the mean of all observed biomass values as predictions.

Scaled Predictors	Model	MAE	RMSE
Yes	Random Forest	34.55	48.81
Yes	Cubist	33.98	49.68
No	Random Forest	34.60	48.54
No	Cubist	33.69	49.08
-	Null	36.94	50.3

Table 10: Hyperparameters and search space used in the 'benchmark' comparison.

Model	Parameter	Search space
Random Forest	mtry	1-23
	min.node.size	1-10
	num trees	1-500
Cubist	committees	1-100
	rules	1-100
	extrapolation	0-100

Since one method did not perform significantly better than the other, both methods were further tuned and resampled separately. However, predictor variables were not scaled.

In tuning the separate models, 5-fold 10-repetition repeated spatial cross-validation was used for the hyperparameter tuning, and 5-fold 100-repetition repeated spatial cross-validation for the performance estimation. 5000 iterations were allocated per fold for hyperparameter tuning. The same search spaces as in the method comparison was used for Random Forest (Table 11). The *rules* parameter was slightly adjusted for the tuning of the Cubist model (Table 11). While further tuning seemed to have improved results of both models, Cubist seemed to have attained higher performance compared to the Random Forest model (MAE: 31.81 vs 33.77) (Table 12).

Table 11: Parameters and search spaces used in model tuning, as well as optimal hyperparameter values.

Model	Parameter	Search space	Optimal value
Random Forest	mtry	1-23	1
	min.node.size	1-10	6
	num trees	1-500	315
Cubist	committees	1-100	83
	rules	2-300	289
	extrapolation	0-100	16.1

Table 12: Performance of the empirical models compared with the performance of a Null model. MAE = Mean Absolute Error, RMSE = Root Mean Square Error. Mean values shown with a 95% confidence interval.

Model	MAE	RMSE	Mean predictions	Observed Mean
Random Forest	33.77	47.89	82.84 ± 3.58 g m^{-2}	81.76 ± 9.12 g m^{-2}
Cubist	31.81	47.60	75.21 ± 4.48 g m^{-2}	
Null	36.94	50.3	81.76 g m^{-2}	

Besides performance metrics, scatterplots were also made of the mean predictions obtained during cross-validation (Figures 14 & 15). The mean predictions of the Random Forest (Figure 14) seem to be concentrated around the mean of all predictions, especially at mid- to high-observed biomass values. In almost all cases, Random Forest overpredicted biomass in the range of 0-100 g m^{-2}, while underpredicting at observed biomass values higher than 100 g m^{-2}, demonstrating a possible upward bias in this range of values. The mean predictions of the Cubist model (Figure 15) exhibit a similar pattern to those of the Random Forest. However, Cubist generally overpredicted to a lesser extent in the 0-100 g m^{-2} range, with the exception of several outliers. This is in spite of the fact that the mean of the Cubist predictions overall is further from the mean of the observed biomass values than that of the Random Forest predictions.

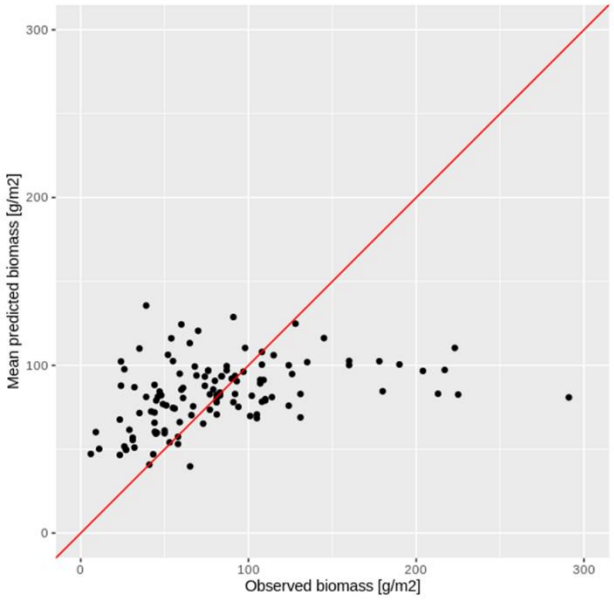


Figure 14: Mean aboveground biomass predictions during cross-validation: Random Forest. The red line represents a 1:1 relationship between observed and predicted biomass.

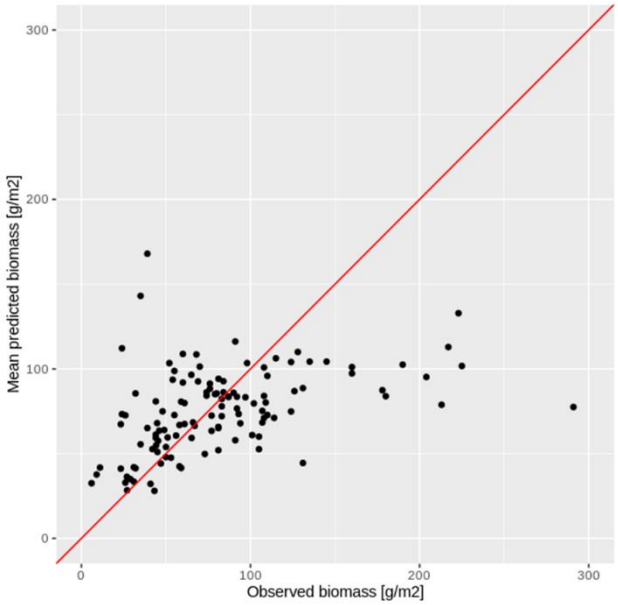


Figure 15: Mean aboveground biomass predictions during cross-validation: Cubist. The red line represents a 1:1 relationship between observed and predicted biomass.

In order to properly investigate overall prediction bias, the mean predictions were also compared to the observed biomass values using the Wilcoxon signed-rank test as outlined in section 4.4.2. In the case of the Cubist model, there was no significant difference of means between the predicted and the observed biomass values, both for the entire dataset as for the ecoregions separately (Table 13). The predictions of Random Forest model were not significantly different from the observed values, however the predictions which were done for the Mongolian-Manchurian Grassland Ecoregion were significantly different (Table 13).

Table 13: Results of the Wilcoxon sign rank test for differences of means. The '' indicates that the P-value is lower than 0.05, thus disproving the null-hypothesis that there is no significant difference in means.*

Model	Data (sub-)set	P-value
Random Forest	All points	0.06
	Mongolian-Manchurian Grassland Ecoregion	< 0.01*
	Forest-steppe Ecoregion	0.63
Cubist	All points	0.44
	Mongolian-Manchurian Grassland Ecoregion	0.10
	Forest-steppe Ecoregion	0.10

The importance of variables which were used in the training of the Cubist and Random Forest models were also extracted (Figures 16 & 17). In the case of the Random Forest model, the values were generated based on permutation importance, or the decrease in model performance when a variable is randomly shuffled. The importance values of the Cubist model were calculated as a combination of the frequency of variable use in the linear models in the nodes and the frequency of variable use in model rules. Note that the importance values cannot be compared between models.

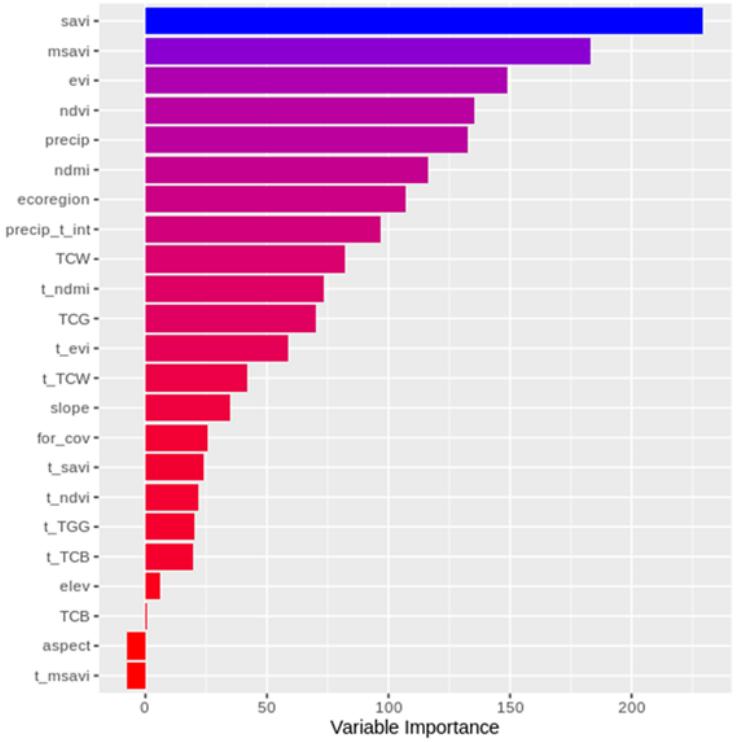


Figure 16: Variable importance of the Random Forest model.

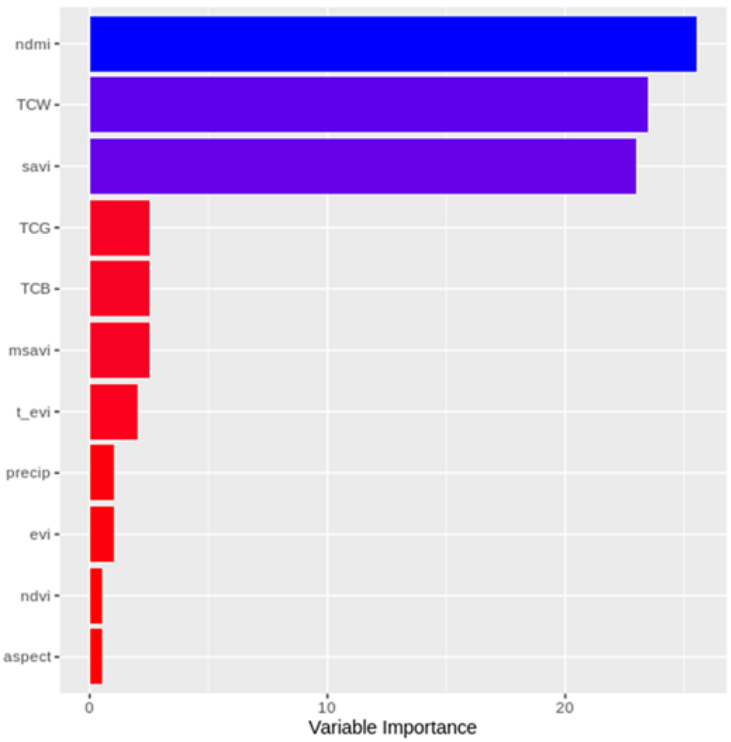


Figure 17: Variable importance of the Cubist model.

From the variable importance graph of the Random Forest model, the vegetation indices SAVI, MSAVI, EVI and NDVI seem to be the most important for the model (Figure 16). However, there was not a great difference in importance value between each variable and the next higher- or lower-ranked variables. The variable importance graph of the Cubist model shows three variables which had significantly higher importance than the others: NDMI, TCW and SAVI (Figure 17).

Since the performance of the Cubist model was shown to be better than the Random Forest model according to the calculated performance metrics, visual inspection of the scatterplots and Wilcoxon test, the Cubist model was chosen to be used as the 'Empirical model' for the remainder of the analysis.

5.2.3. Biomass predictions: Empirical model

The predictions for aboveground biomass and sustainably usable biomass, in which non-grassland areas such as agriculture, urban settlements and bodies of water were masked out can be seen in Figure 18. Overall, the highest amounts of sustainably usable biomass per hectare according to the empirical model were found in the northern forest steppes while lower biomass values were generally found in the southern, Mongolian-Manchurian Grassland ecoregion. Since the slopes in Bulgan are generally not very steep, the reduction of biomass from total to proper use does not deviate greatly from the proper use fraction of 0.65.

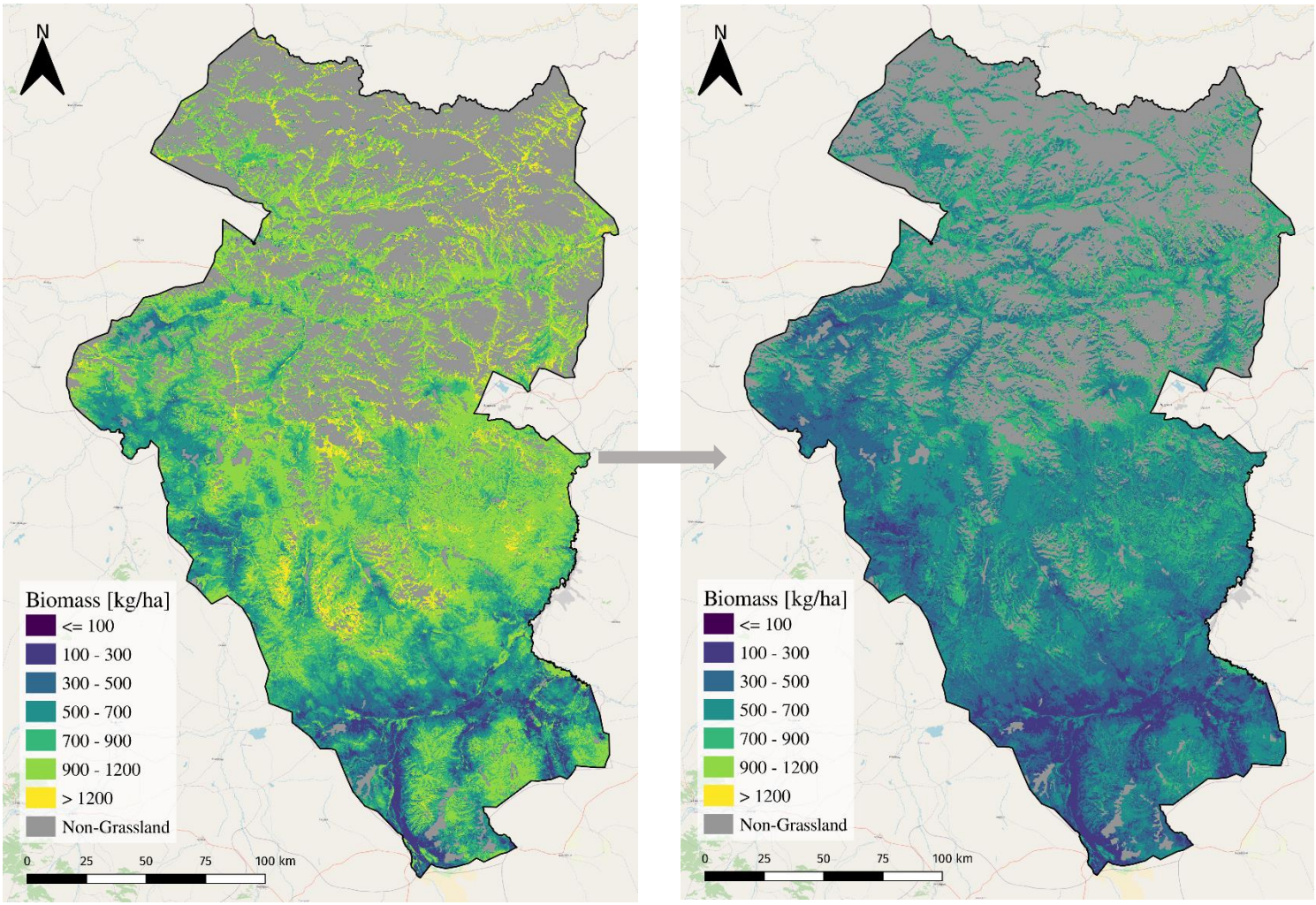


Figure 18: Prediction of aboveground biomass [kg/ha] (left), and biomass adjusted for sustainable use [kg/ha] (right) for Bulgan province using the empirical model.

The three districts with the highest mean predicted biomass per hectare values were Selenge, Teshig and Bugat with 643, 636, and 623 kg/ha biomass, respectively (Figure 20). The three districts with the lowest mean predicted biomass per hectare values were Gurvanbulag, Rashaant and Bayannuur, with 368, 424, and 443 kg ha biomass, respectively (Figure 20). The confidence interval of the observed values, visible as black bars in Figure 19, show that there were often only a low number of observed values per district, which resulted in relatively wide confidence intervals. Therefore, it is difficult to make conclusions on the performance of the model within certain districts, such as Bayannuur (wide confidence interval), Bulgan (only one sampling point) or Rashaant (no sampling points). However, on a district-by-district basis, the process-based model prediction means (Figure 10) fell outside of the confidence intervals of

the observed biomass samples more often than the empirical model predictions did (Figure 19).

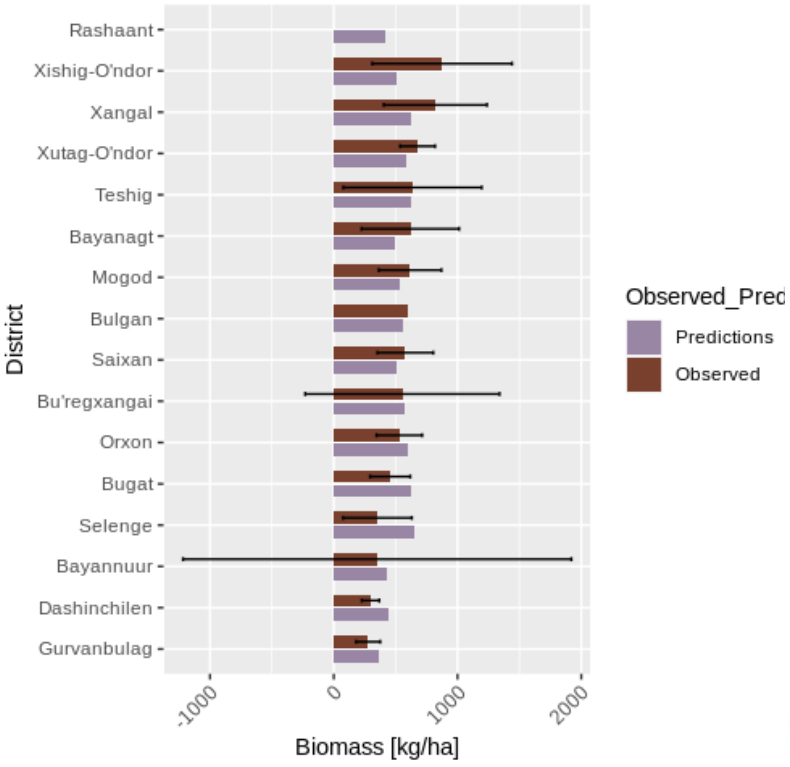


Figure 19: Sustainably usable biomass predictions compared to converted observed biomass values, per district using the empirical model. The black bars represent the 95% confidence interval the observed values.

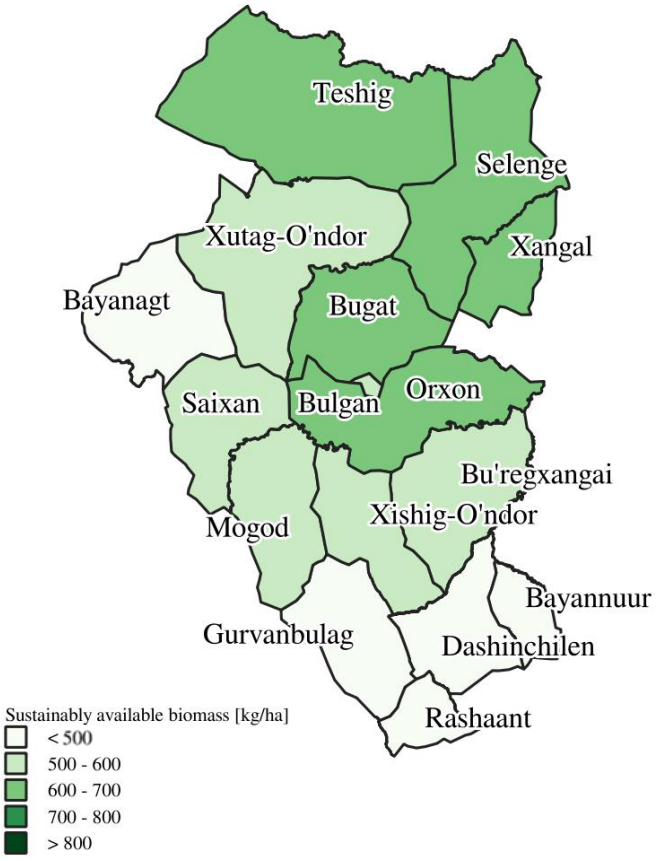


Figure 20: Mean sustainably usable biomass predictions per district; empirical model.

With the resulting overall mean proper use biomass figures, the overall theoretical carrying capacity can be calculated. According to the empirical model, the carrying capacity, or the total amount of sheep-equivalent units that can sustainably graze in a 150-day period in Bulgan province is 7.5 million SEU, compared to the 9.3 million as estimated using the process-based model.

5.3. Forage Intake

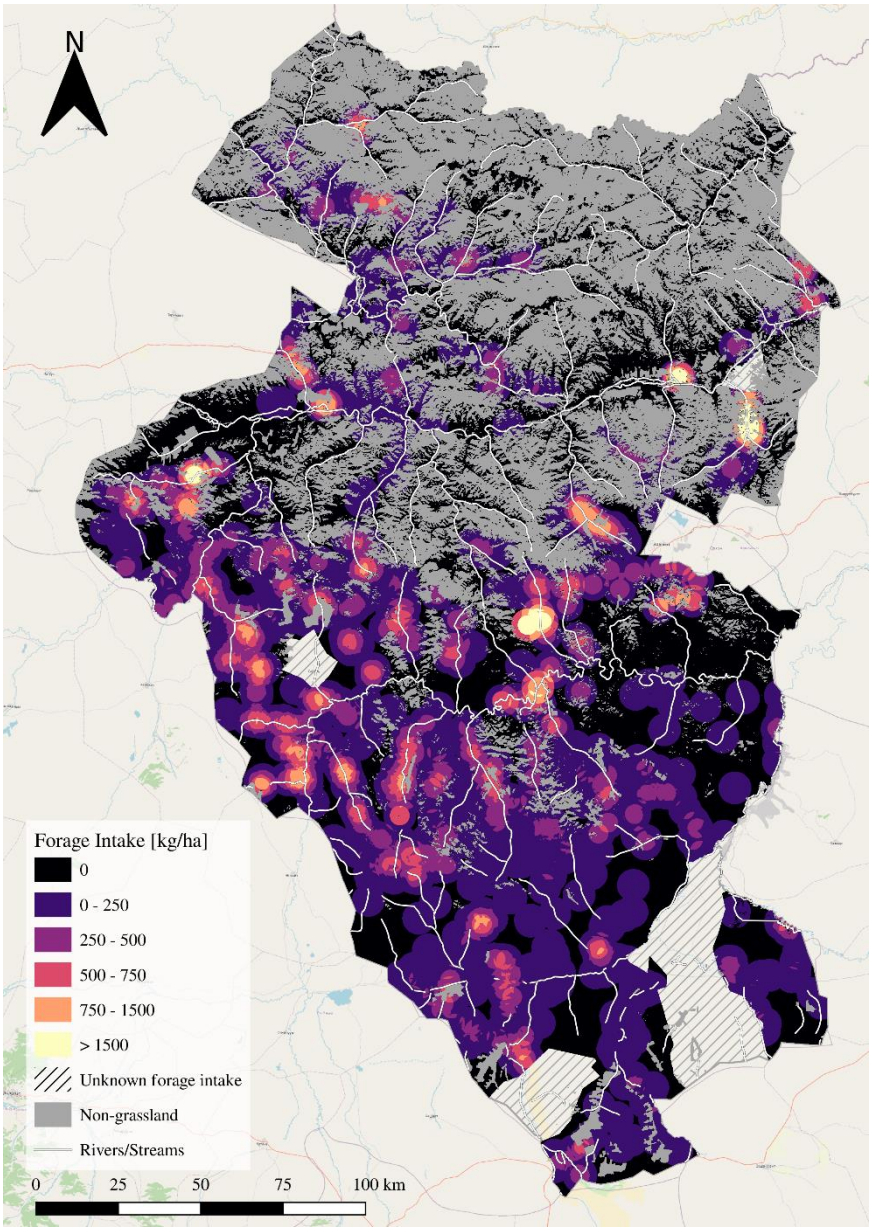


Figure 21: Estimated forage intake [kg/ha] for herders in their summer locations (90-150 days). Hashed areas are regions for which there was no/incomplete data. Grey areas represent non-grassland areas, which include mainly forest and agricultural land. White lines are rivers and streams.

The majority of forage intake from the summer locations of herders, which encompasses a 90-150-day period, occurred in the southern half of Bulgan province (Figure 21). Generally, forage intake 'hotspots' were concentrated around rivers and streams (Figure 21). Areas characterized by high degrees of forest cover such as in the north of the province did not experience high forage intake levels (Figure 21).

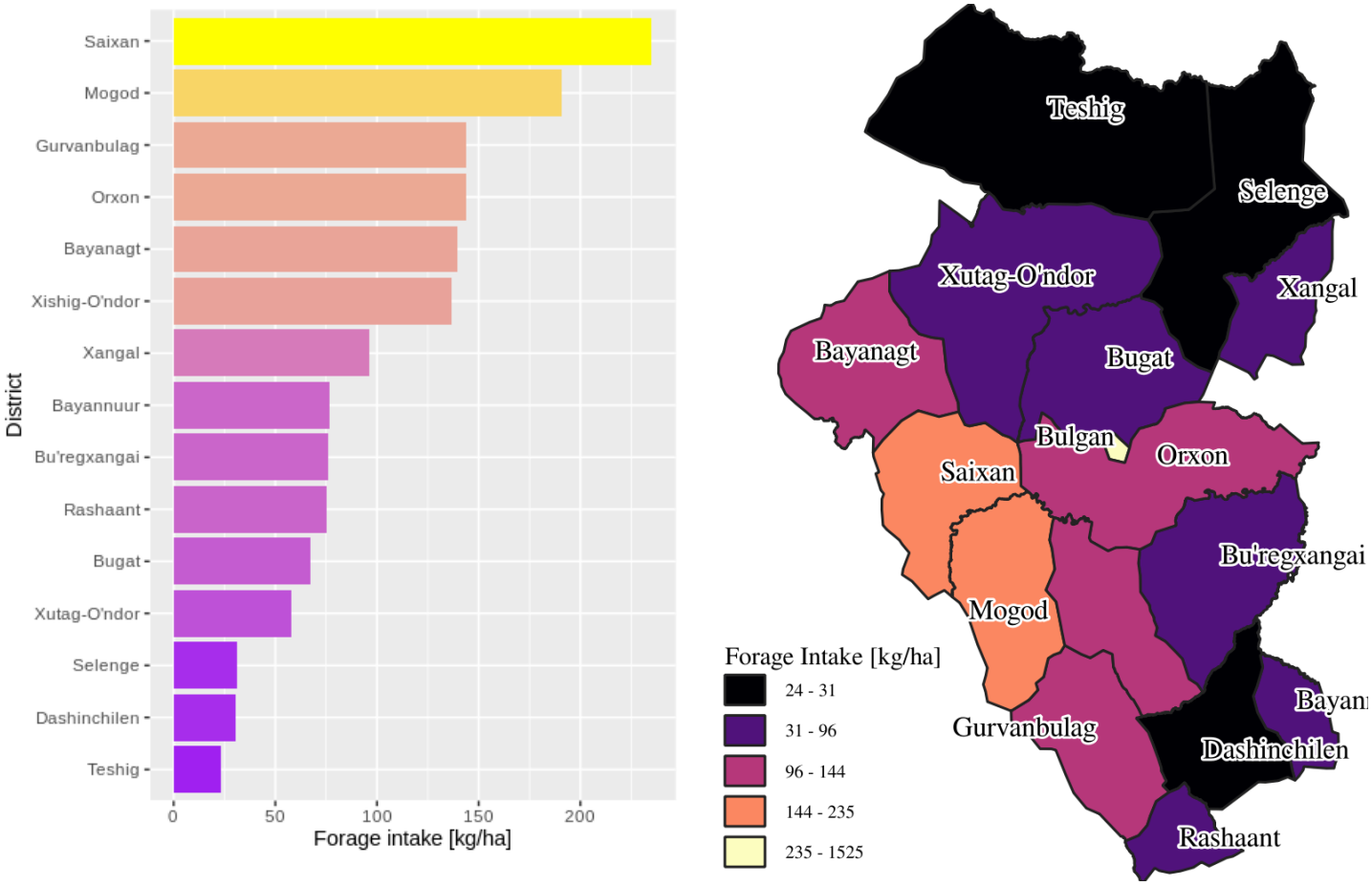


Figure 22: Mean forage intake in grassland areas by district [kg/ha]. Note that Bulgan district was not included in the bar chart, as the mean forage intake of this district exceeds 1500 kg/ha, which would have hindered the legibility of this chart.

The top three districts with highest mean forage intake levels in grassland areas were Bulgan , Saixan and Mogod, with 1524, 235, and 190 kg/ha mean biomass intake respectively (Figure 23). The three districts with the lowest forage intake levels were Selenge, Dashinchilen, and Teshig, with 31, 31 and 24 kg/ha mean biomass intake levels, respectively.

5.4. Overgrazing

5.4.1. Relative grazing pressure: Overview

Dividing the previously produced sustainably usable biomass rasters by the forage intake rasters produced with the process-based and empirical models resulted in maps showcasing the relative use of sustainably available biomass according to each model (Figure 23). Since the mean biomass predictions produced by the empirical model were lower than those produced by the process-based model, the overall mean relative use of sustainably available biomass was predicted to be higher by the empirical model than by the process-based model. In addition, the relative usage data produced by the empirical model is a higher resolution than the process-based model data, allowing for a higher degree of spatial detail.

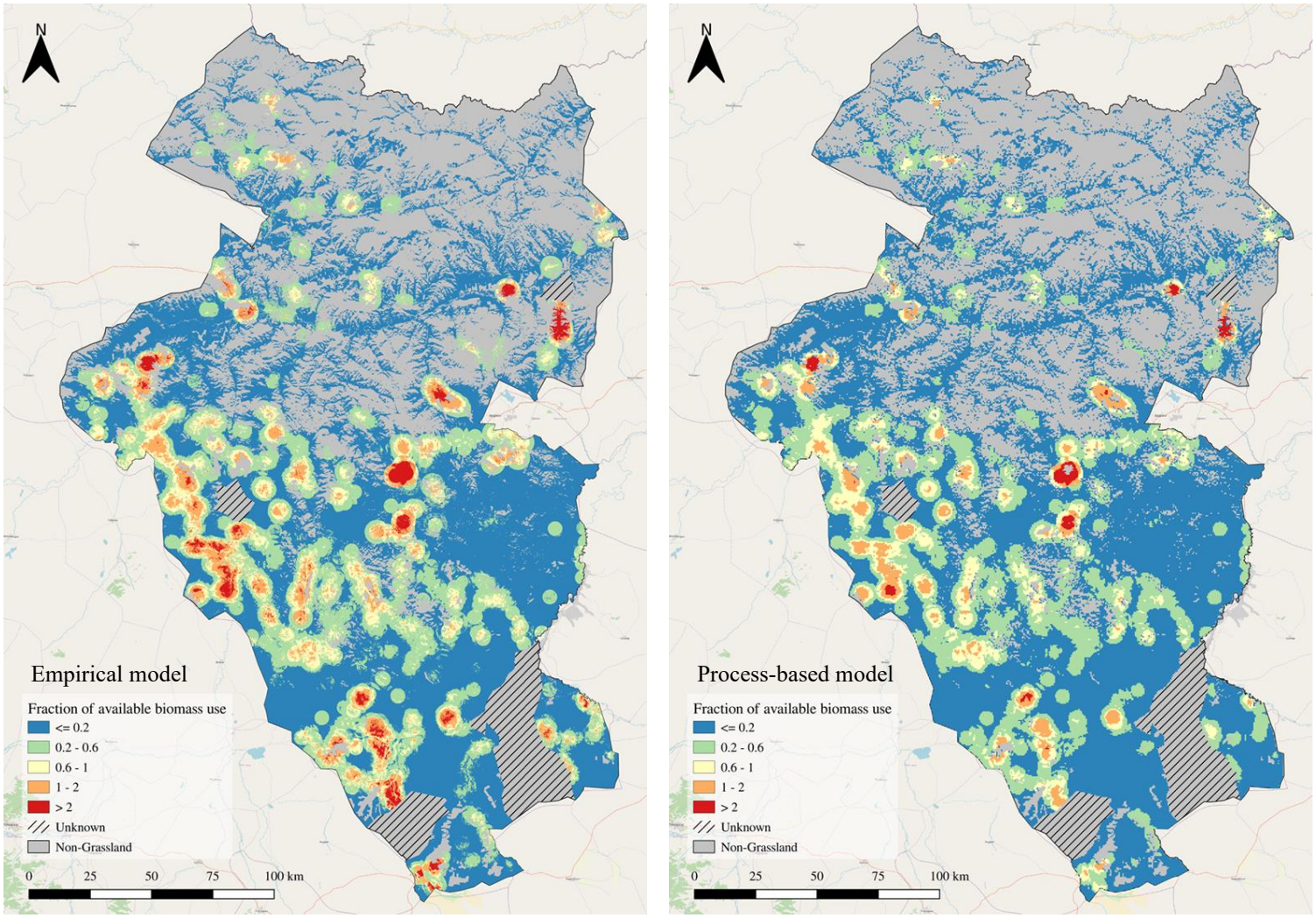


Figure 23: Fraction of available sustainably usable biomass used by herders in their summer locations, as predicted employing the empirical model (left) and process-based model (right). A fraction of use below 1 indicates that more forage could potentially be used sustainably, while a fraction of use over 1 indicates there was potentially overgrazing occurring.

Besides the pixel-scale forage use factor, the overall overgrazed grassland area and overgrazed grassland area per district was also calculated (Figures 24 and 25). ‘Overgrazed’ was defined as the percentage of area with a forage intake higher than sustainably available biomass (fraction of sustainably available biomass used > 1), while ‘Severely Overgrazed’ is overgrazed by a factor of at least 2. The three districts with the largest relative overgrazed area coverage according to the empirical model were: Bulgan (73.29%), Saixan (20.78%), and Gurvanbulag (16.99%). Using the process-based model, the three most-overgrazed districts were: Bulgan (67.33%), Saixan (13.59%), and Xangal (9.62%). The districts with the lowest percentage of overgrazed area according to the empirical model were: Dashinchilen (0.15%), Buregxangai (0.49%), and Teshig (1.93%). Using the process-based model, the districts with the lowest overgrazing percentage were Dashinchilen (0%), Buregxangai (0.13%), and Bayannuur (0.39%). The largest differences between the empirical and the process-based models occurred in Gurvanbulag (9.13%), Saixan (7.22%) and Rashaant (6.67%) (Figure 25).

In general, the most overgrazed areas were located in the central and western parts of the province (Figures 26).

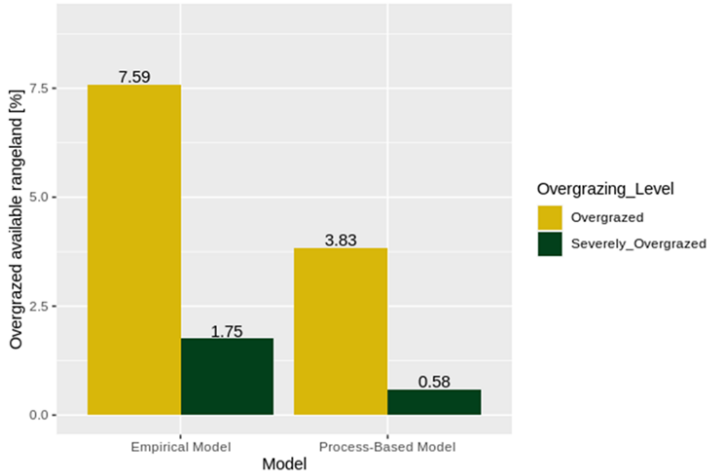


Figure 24: Overall overgrazed area in Bulgan province as a percentage of available rangeland area. 'Overgrazed' is defined as the percentage of land in which more forage is used than available, while 'Severely Overgrazed' is the percentage of land in which there is more than twice as much forage intake as there is available sustainably.

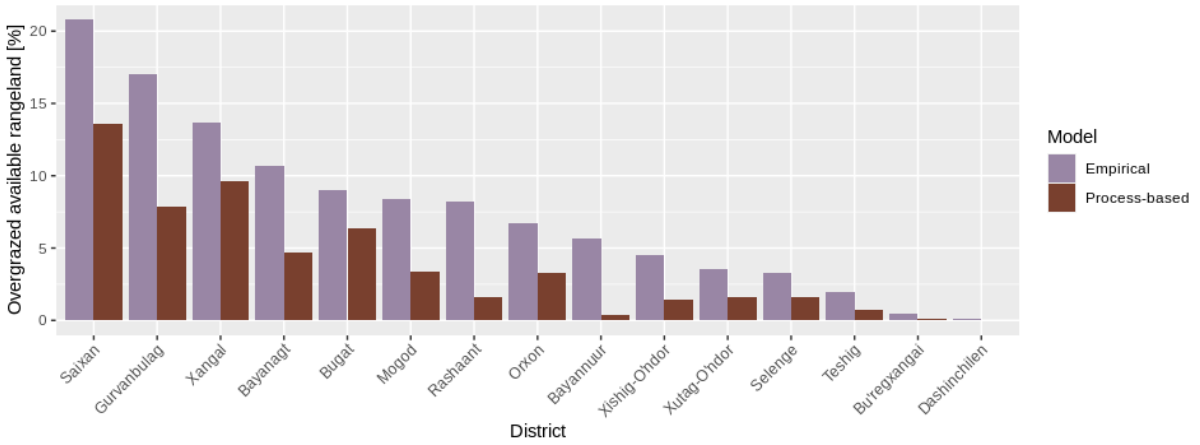


Figure 25: Overgrazing per district as a percentage of available rangeland. Bulgan district was not included, since the overgrazing percentage was so high that it would have impeded the legibility of this chart (73% and 67% for empirical and process-based models, respectively).

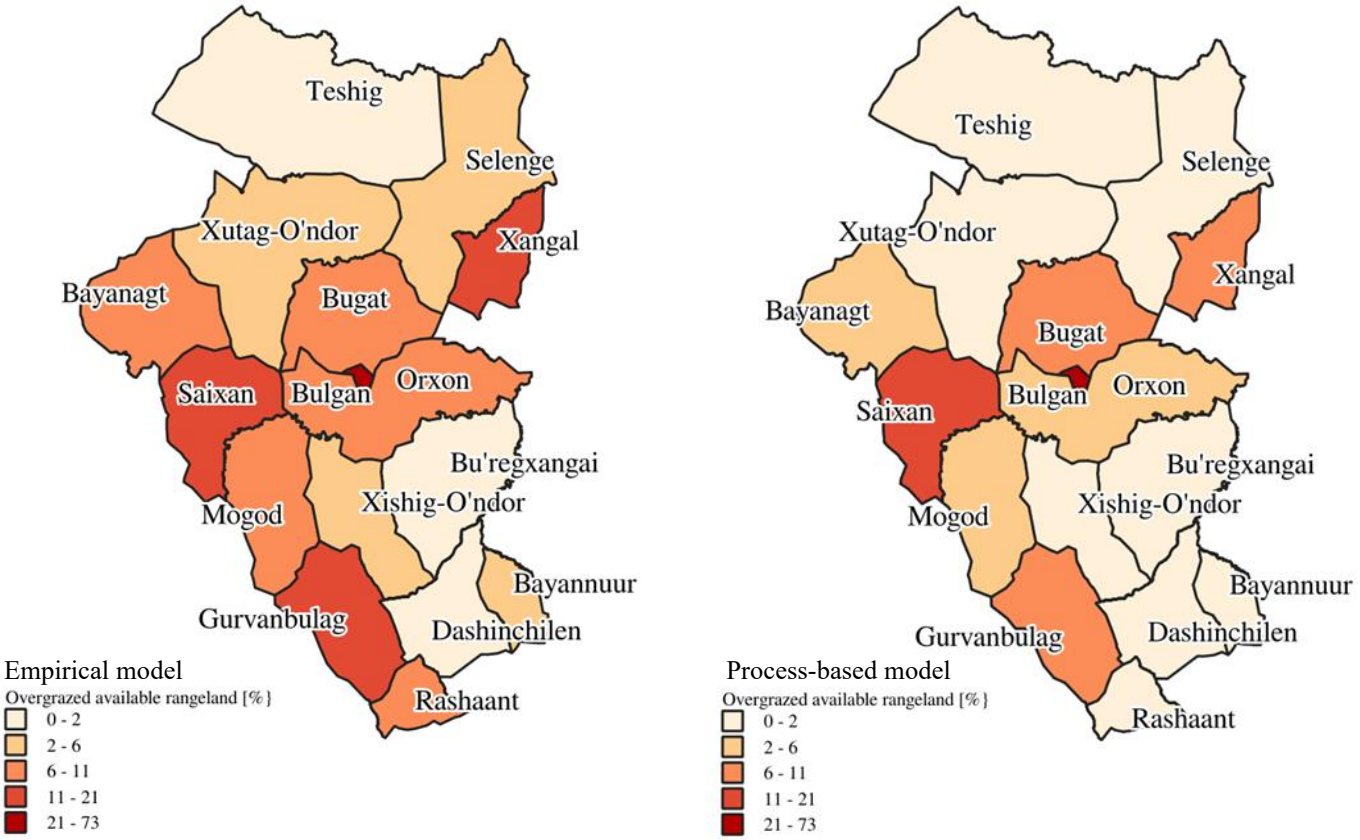


Figure 26: Overgrazed available rangeland area (fraction of biomass used > 1) as a percentage of total available rangeland area, using the empirical model (left) and process-based model (right).

5.4.2. Ecoregion and local-scale differences in overgrazing

Beyond differences in overgrazing between districts, it was also investigated if there were differences in overall overgrazing between the ecoregions, Forest-steppe and Mongolian Manchurian grassland, and between the model overgrazing predictions of the ecoregions. According to the empirical model, the ecoregion with the largest overgrazed area as a percentage of its total area was the Mongolian-Manchurian grassland in the south of the province, with 8.34% of the total available rangeland overgrazed, of which 2.28% was severely overgrazed, compared to the Forest-steppe ecoregion with 7.39% and 1.6% (Figure 27). However, according to the process-based model, the ecoregion with the largest relative overgrazed area was the Forest-Steppe ecoregion, in the central and northern part of the province, with 3.97% overgrazed area, compared to only 3.3% in the Mongolian-Manchurian grassland (Figure 27).

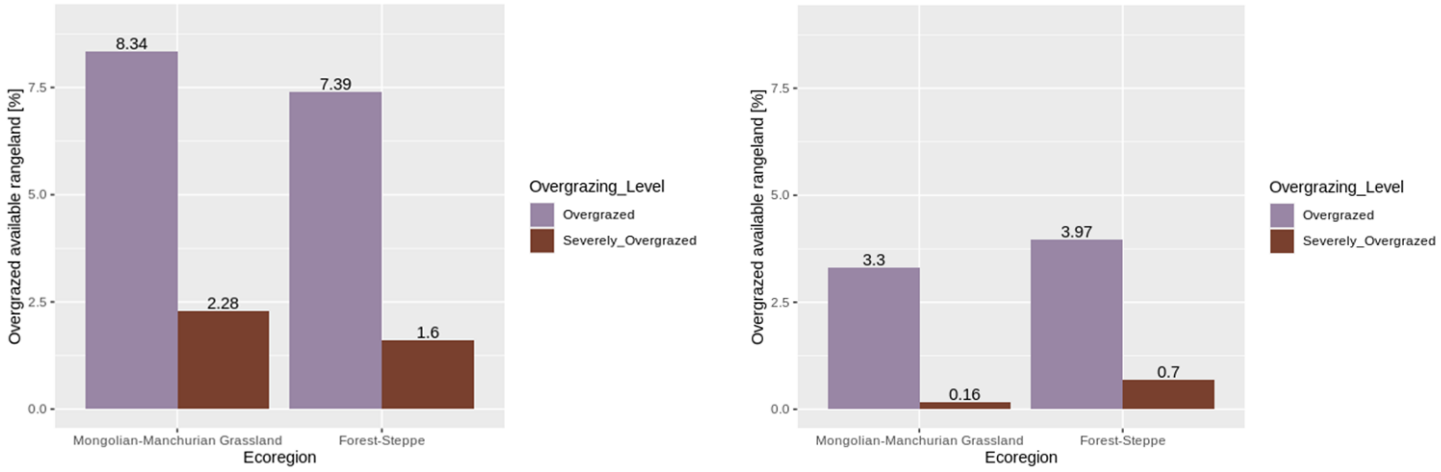


Figure 27: Percentage of overgrazed and severely overgrazed area per ecoregion, using the empirical model (left) and the process-based model (right)

In order to illustrate this difference between empirical and process-based models in the ecoregions, two districts were taken as examples, one from each ecoregion: Bayanagt from the Forest-steppe ecoregion (Figure 28), and Gurvanbulag, from the Mongolian-Manchurian grassland ecoregion (Figure 29). Although the fraction of sustainably available biomass use is generally higher for the empirical model, regardless of the ecoregion/district, the difference is larger in Gurvanbulag, located in the generally low-biomass Mongolian-Manchurian grassland, than in Bayanagt, located further north in the Forest-steppe ecoregion, where biomass was

predicted to be higher overall. This mirrors the results of the biomass predictions, as the process-based model overpredicted strongly for low observed biomass values, while the empirical model was able to predict lower biomass values with a higher degree of accuracy. In addition, the higher resolution of the input data used for the empirical model allows for varying overgrazing predictions in highly spatially varied biomass values at a local level, as is especially the case in mountainous terrain such as in Gurvanbulag (Figure 29). In contrast, the overgrazing predictions made using the process-based model capture much less of potentially important local variation.

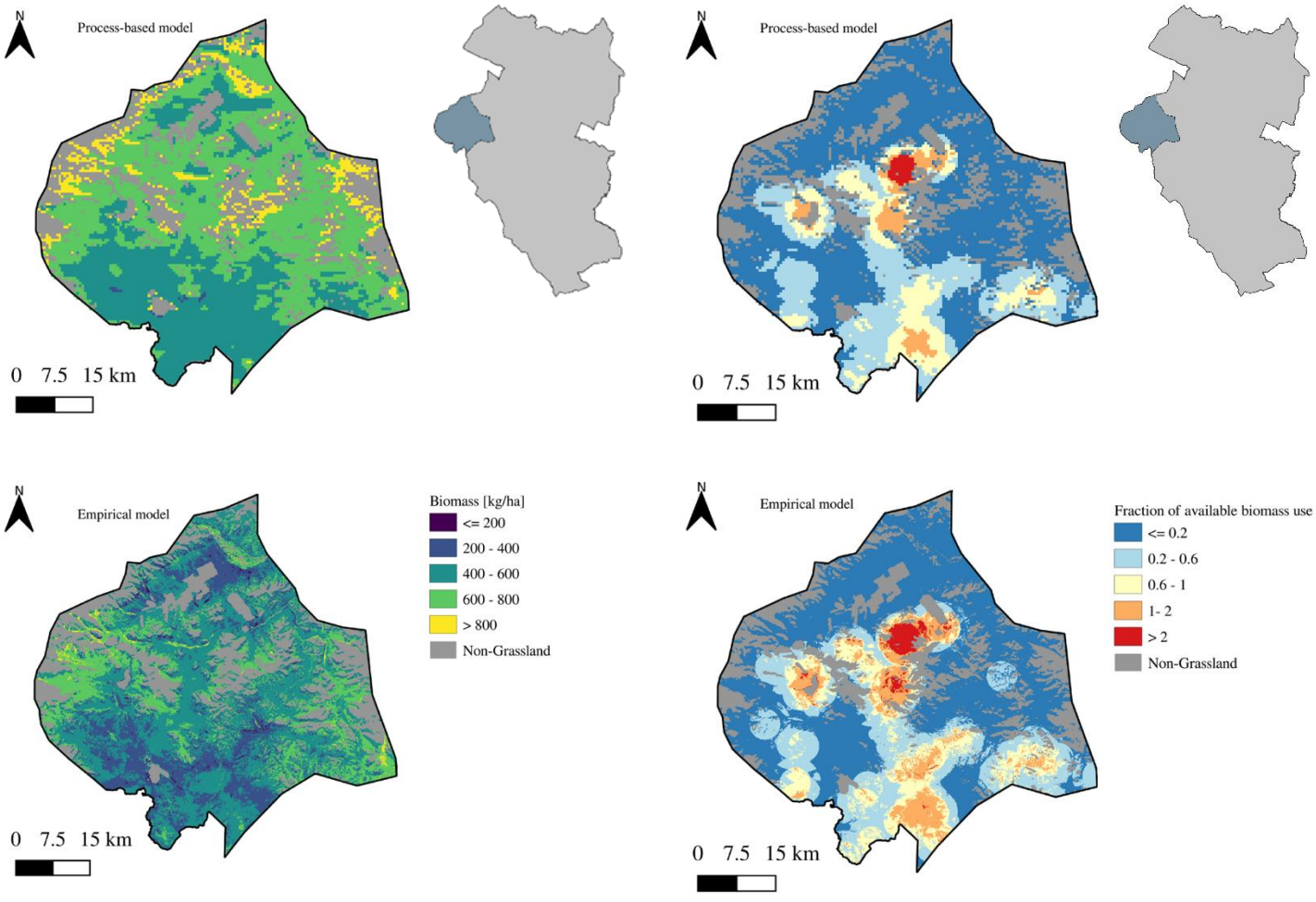


Figure 28: Comparison between overgrazing levels predicted by the process-based and empirical models in Bayanagt, located in the Forest-Steppe ecoregion. Also included on the left side are the respective sustainably available biomass predictions, to illustrate the cause of the difference in fraction of biomass use between the models.

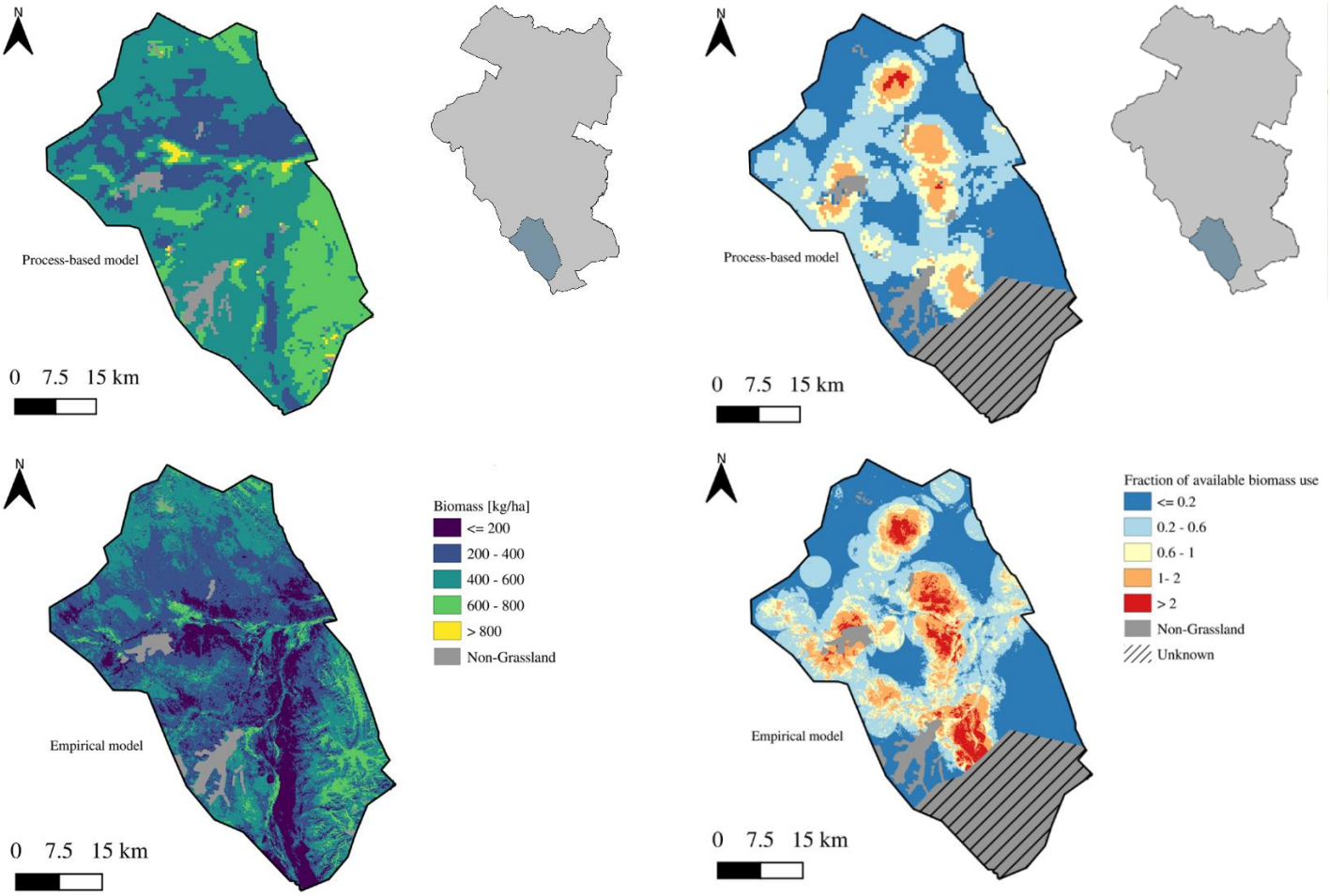


Figure 29: Comparison between overgrazing levels predicted by the process-based and empirical models in Gurvanbulag, located in the Forest-Steppe ecoregion. Also included on the left side are the respective sustainably available biomass predictions, to illustrate the cause of the difference in fraction of biomass use between the models.

6. Discussion

6.1. Model performance

Overall, it is clear that the empirical model outperformed the process-based model. The empirical model predictions resulted in better results in terms of performance metrics; a MAE of 31.81 and an RMSE of 47.60 were obtained using the Cubist method, compared to 39.48 and 49.39 respectively obtained from the process-based model using the yearly composite NPP as input data. Furthermore, the Wilcoxon tests also confirmed that there was no significant different between the observed values of aboveground biomass and the predicted values obtained using the empirical model, while the process-based model showed significant differences between observed and predicted values (Figure 30). In addition, when dividing the data into ecoregions, the mean empirical model biomass predictions were still closer to the mean of the observed biomass values than the predictions of the process-based model (Figure 30).

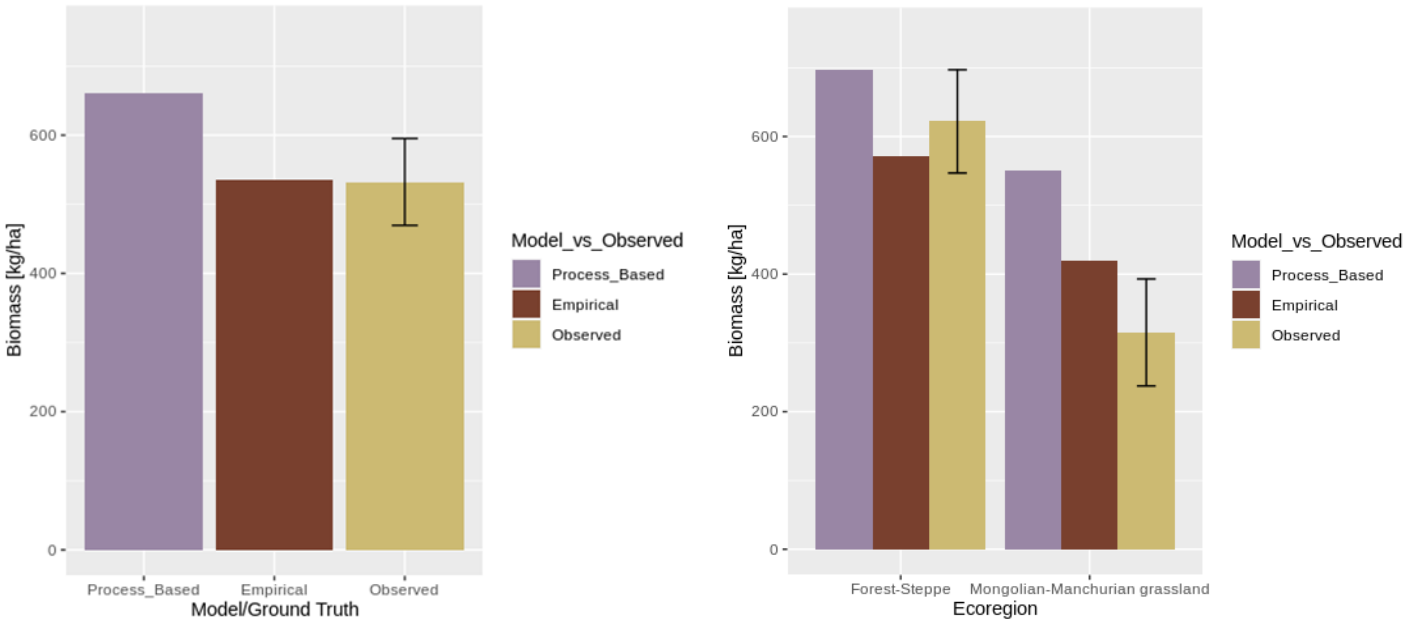


*Figure 30: Comparison between mean process-based and empirical sustainably available biomass values and the mean observed sustainably available biomass values **overall** (left), and per **ecoregion** (right). The black bar represents the 95% confidence interval of the observed mean.*

It must be noted that this is not surprising, as the process-based model was not parameterized for this area and was simply transplanted from one region to another, from Azerbaijan to Bulgan, Mongolia (de Leeuw et al., 2019). Although the biomass prediction portion of the model was not parametrized specifically for the region of study in the analysis of de Leeuw et al. (2019), the results of that study were much more optimistic about the potential of the process-based method of biomass prediction. There are several potential explanations.

First, the method of adjustment using Landsat imagery to adjust the mean field biomass values significantly increased accuracy in Azerbaijan (38% mean underprediction to 16% mean underprediction), while this same method applied in Bulgan did not result in sufficiently higher accuracies (15.8% to 13.9% overprediction), so that there was no longer a significant difference in means between observed and predicted data. It is possible that the landscape in Azerbaijan is more heterogenous on the scale of a MODIS pixel compared to the landscape of Bulgan. It must also be noted that the predicted biomass values were manually adjusted upwards in the study by de Leeuw et al., to '[…] be as close as possible to the field-based measurements' (2019, p.72). For the proposed use of this model as indicated at the start of this thesis, namely, to develop a model which avoids the necessity of extensive gathering of field data, manually adjusting all predictions up- or downwards would defeat the purpose of being relatively independent from field samples.

Furthermore, the fANPP (fraction of aboveground net primary production) term used in the process-based model could be ill-suited for the rangeland type of Bulgan province. For example, the mean of field sample biomass that was gathered in Bulgan (820 kg/ha) was significantly lower than in Azerbaijan (2974 kg/ha) (de Leeuw et al., 2019). This, combined with the fact that the process-based model was heavily upwards-biased especially in the low biomass ranges (0-1000 kg/ha), of which there are very few areas in Azerbaijan, indicates (in accordance with the advice of de Leeuw et al.) that further work should be done on the relation between fANPP and aboveground biomass for different rangeland types.

Thirdly, it is possible that due to the sampling design, which was not conducted with this research in mind, there is a significant bias introduced into the results which affect the validation of model results. Although a correction was carried out using Landsat NDVI as previously discussed, this was rather rudimentary, and cannot compensate for a sampling design specifically tailored to this research. For example, instead of many dispersed sampling

plots, clustered sampling could be performed, so that the spatial variability within one 500m resolution pixel could be better characterized. In addition, instead of purposeful sampling, sampling could potentially be performed in a stratified manner, where the study area is divided into strata and then samples are taken weighted by strata size. However, this is a costly exercise, especially in difficult-to-access remote areas.

Finally, since it was chosen to use the yearly NPP composite as input for the process-based method for biomass estimation, the overprediction of biomass values could be partly due to the biomass sampling at approximately 'peak greenness' not capturing the total production of biomass in one year. In essence, the process-based model estimated the biomass production for a year, while the empirical model estimated the biomass at one point in time. If there was no further growth after the point of sampling, than one might expect the models to return similar biomass values. However, if there is heavy removal and re-growth of biomass within one year, and after the biomass sampling date, the results of the process-based model might not be able to be validated through a set of field sample of biomass taken at one point in time only and appear to be heavily overestimating aboveground biomass.

Therefore, it cannot be concluded that the notion of an NPP- process-based model is not suitable for aboveground biomass prediction. However, some local parameterization is necessary, as the model proposed by de Leeuw et al. did not perform well in the study area of this thesis, Bulgan province. In addition, sampling must be done with the spatial resolution of model predictions in mind, as well as the temporal resolution in order to ensure a reliable validation strategy.

Although it was not the focus of this thesis, the biomass prediction results from the empirical method were not as accurate as results from previous studies. For example, in a study by John et al., in which aboveground biomass was estimated for varying rangeland types across Mongolia using Cubist models, the best model resulted in an MAE of 59.83 g/m (9% of the observed values mean), while the Cubist model in this thesis resulted in an MAE of 31.8 g/m (38.9% of the observed values mean) (John et al., 2018). However, due to the sampling design which utilized much larger sampling plots, clustered and stratified sampling, as well as a much larger number of samples, John et al. (2018) were able to use 500m MODIS as input data, which potentially greatly reduced issues related to cloud cover due to shorter revisiting intervals than Landsat 8, as well as reducing potential issues with geometric accuracy related

to the higher resolution that Landsat 8 offers. Interestingly, the best models predicting aboveground biomass in the study by John et al. (2018) found that the most important input variables were often spectral indices related to moisture content, such as Land Surface Water Index (LSWI) and the Normalized Difference Water Index (NDWI). This is similar to the best Cubist model found in this analysis, as these were NDMI, Tasseled Cap Wetness and SAVI (Figure 17). In addition, the aboveground biomass prediction models of John et al., also underpredicted at higher biomass levels, however to a lesser extent, as there were likely more high-biomass sample points in their training dataset (John et al., 2018).

6.2 Carrying capacity and forage intake

As mentioned in the results section, the overall maximum sustainable carrying capacity in Bulgan in the summer is estimated to be 7.6 million SEU by the empirical model, and 9.3 million SEU by the process-based model. However, this assumes that livestock is optimally spread out, that areas are not overgrazed, and that there were no previously overgrazed areas as a result of grazing in the spring (using more than the sustainably available forage). This is of course not the reality, as we have seen from the model predictions that sustainably available biomass can range greatly between ecoregions and districts, and even on a pixel-by-pixel basis. The model predictions for carrying capacity were highest overall in the Forest-steppe ecoregion (572 kg/ha, or 2.54 SEU/ha, empirical model), in the central and northern parts of the province, while the dryer Mongolian-Manchurian grassland in the south and south-west, such as in the districts of Gurvanbulag, Rashaant, Dashinchilen and Bayanuur were estimated to have lower carrying capacities (419 kg/ha, or 1.86 SEU/ha, empirical model) (Figure 19). Even within districts themselves, carrying capacity can vary as well, with biomass values in Gurvanbulag ranging from below 200 to the 600-800 kg/ha range (Figure 29).

In many studies on carrying capacity and overgrazing of rangeland by (nomadic) livestock, stocking density estimates can only be made on the basis of district or, at best, sub-district wide levels, since the exact locations of herders are not known (Gao, Angerer, Fernandez-Gimenez, & Reid, 2015). Using the known locations of herders, it was found that the locations of herders are often highly concentrated in clusters. By using livestock statistics aggregated to administrative boundaries, there is thus a risk of greatly underestimating forage intake in areas where herders are clustered locally, as is visible when comparing Figures 21 & 22. According

to the Mongolian National Statistics Office, the total amount of livestock in Bulgan in 2014 was 5.8 million SEU (conversion to SEU done using livestock conversion matrix provided by Mongolian Institute of Science) (National Statistics Office of Mongolia, 2019; Sainbuyan, 2016). Spreading this number over the available land for grazing would result in a stocking density of 0.54 SEU/ha, or a mean forage intake of 122.4 kg/ha over a 150-day period. Judging from the results of the forage intake analysis (Figure 22), in some areas, especially the district centers where non-nomadic herders are located, forage intake was estimated to reach up to 1500 kg/ha in a 150-day period. Furthermore, even within districts it was estimated that there was a wide range of forage intake values possible. For example, in Xangal almost all summer forage intake was concentrated in one river valley, resulting in extremely high estimated forage intake values in this area, while the rest of the district was not grazed at all. However, by averaging forage intake over the total area of the district, Xangal appears to be ranked halfway among all districts in terms of mean forage intake, while it contains some of the heaviest-grazed land in the province.

It must be noted that there could be some inaccuracies arising from the methods used to estimate forage intake. First, the herder location points were not the locations of single herder families, but usually groups of herders. These groups do not remain consistent however throughout the seasons. The data containing livestock statistics, contained herder names along with the sub-district in which they were located, but these could not be linked to the exact herder locations. Therefore, it was chosen to divide the total livestock numbers within each sub-district equally among the herder locations, regardless of the number of herder families referenced by each point. Since it is unknown how consistently herders were grouped into single point locations, this could result in over- or underestimations of forage intake. In order to potentially increase accuracy, the allocation of livestock among herder locations could be done according to number of herders per point, with points representing more herders being allocated a higher number of livestock. However, since not all herder location points contained this data, this was not done in the analysis. Secondly, for some sub-districts, there were livestock statistics, but these did not contain herder location points. As these were almost exclusively district centers (permanent settlements), it was chosen to assign these livestock figures to a simulated central ‘herder location’ point. This is also the explanation for the extremely high forage intake in Bulgan district. Finally, several subdistricts contained herder

locations, but the livestock data was either missing, corrupted, or unusable. Livestock statistics on a sub-district level from the Mongolian National Statistics office were then used if the per-point density of livestock was not unreasonably larger than the district average. This was the case for Huremt (Saixan district), Dorgont & Haraat (Dashinchilen district), and Hyalganat (Xangal district). These sub-districts were then marked as having an 'unknown' forage intake density, and not considered when calculating overall forage intake and overgrazing levels.

6.3. Overgrazing

Combining the summer carrying capacity, or maximum sustainably usable forage, with the summer forage intake resulted in the fraction of use maps (Figure 23). For both the process-based and the empirical model, the overgrazing results showed that there was a highly localized pattern of overgrazing. It is difficult to compare these results to previous findings of other authors, since there is no agreed upon definition for proper use in Mongolia (Gao et al., 2015). In addition, overgrazing is often estimated on a yearly basis, rather than on a seasonal basis, which does not take into account potential regrowth of biomass during the year. Regardless, in a study by Gao et al. (2015), 2014 was found to be the year with the highest level of overgrazing in a 10 year period starting in 2003; the percentage of land overgrazed as a whole in Mongolia, defined as a percentage of use (of total aboveground biomass) over 70%[1], or a usage factor of sustainably available biomass larger than 1, was found to be 39.4 (calculated as an average of percentages over different land cover classes). If only the landcover classes were taken which correspond to those in Bulgan, Forest-Steppe and Steppe/Dry-Steppe, the mean overgrazing percentage according to Gao et al., was 46.5 (Gao et al., 2015). The overall percentage of overgrazed land area in Bulgan, defined as land on which there was a usage factor of sustainably available biomass larger than 1, during the summer season in 2014 according to this thesis, was 7.59% according to the empirical model, and 3.83% according to the process-based model (Figure 27). It must be noted that using the districts, subdistricts or even the province of Bulgan as 'potential' grazing area is problematic and might portray an overly optimistic view of the land available for grazing in one season,

[1] The proportion of sustainably usable biomass as defined by Gao et al. (70%) is similar to the 'proper use' factor used in this study (65%, with further reductions according to slope steepness).

since the theoretical boundaries of seasonal grazing areas are not fixed according to administrative boundaries. It would make more sense to determine the overgrazed area in the summer as a percentage of the theoretical total summer grazing area. Unfortunately, this data is unavailable. Nevertheless, the analysis provided by this thesis thus implies a generally lower percentage of land area as overgrazed, using either the process-based model or the empirical model, than the study by Gao et al. This difference can be attributed to several factors: First, it must be noted that the scales of comparison are different, and that figures representative for all Forest-steppe and Steppe/Dry-Steppe (as defined by Gao et al.) grassland in Mongolia are not representative for only Bulgan province. Secondly, the forage intake calculations performed in the study of Gao et al. used stocking densities averaged over administrative boundaries, rather than a point-and-buffer based estimate such as performed in this thesis. As previously discussed, this could mean that total area of estimated overgrazed land has the potential to be overestimated, while underestimating the overgrazing on a very local scale. For example, fractions of use of sustainably available biomass in excess of 500% could be found on a highly local level in Gurvanbulag, Bayanagt, Saixan and Bulgan districts, while overgrazing simultaneously did not occur in large areas in these districts (Figure 23). Finally, the analysis performed for this thesis did not include the forage intake for the whole year, rather, only the summer locations of herders were taken into account. Therefore, if the forage intake over the entire year would have been estimated, the resulting percentage of overgrazed land would have been higher as well. The grazing pressure as estimated by this study should therefore not be seen as a representation of the overall condition of rangeland at a certain point in time, but rather an estimation of areas which were over- or under-grazed as a result of the temporary settling of a herder(group) for one season only. In other words, one could only determine the overall condition of the rangeland using long-term data describing all movements of the herders in combination with biomass (re-growth) data spanning the year, rather than a single season.

The question if Bulgan as a whole was overgrazed in the summer season of 2014 is difficult to answer. While a large portion of land was not overgrazed as a result of the grazing of livestock in their summer locations, areas previously (over)grazed could not be taken into account. However, areas that were overgrazed were often very overgrazed to a higher level than previously assumed.

7. Conclusion

The main objective of this thesis was to assess the accuracy of a process-based model for aboveground carrying capacity in rangelands, compared to that of an empirical model, and couple these models with stocking density estimates to identify areas in which overgrazing has occurred in Bulgan, Mongolia. Since existing models of biomass estimation require high degrees of local parameterization, it was investigated if an existing process-based model with minimal local parameterization developed for Azerbaijan could be applied to Bulgan, Mongolia in order to estimate biomass, and ultimately estimate overgrazing levels. The performance of the process-based model when applied to Bulgan resulted in low accuracy when validated with in-situ biomass samples, which was not better than predicting with the mean of observed values. However, the model performed comparably to the model in Azerbeijan (without manual adjustment). Despite this, the mean of the biomass predictions differed significantly from the mean of the observed values ($p < 0.05$). The empirical model, developed using Cubist, delivered better performance than the process-based model, and did not result in significant differences between the means of predictions and observed biomass values ($p > 0.05$). Upon conversion of aboveground biomass to sustainably usable biomass, the overall theoretical summer carrying capacity in Bulgan province in 2014 was estimated to be 7.6 million sheep equivalent units (SEU), with higher carrying capacity in the forest steppes in the north compared to the southern grasslands. Forage intake was estimated to be generally highest in the southern half of the province. However, forage intake was often highly localized to clusters in river valleys. As a result, the total area of overgrazed land returned by this study by both models was lower than previous studies, but overgrazed land was estimated to be more intensively overgrazed. The heaviest overgrazing was estimated to have occurred in roughly the same areas using both the process-based and empirical model, but due to overprediction of biomass by the process-based model, overgrazing was not estimated to be as extensive by the process-based model as by the empirical model. Overall, the process-based model should be considered as unsuitable in its current form for rangeland biomass prediction in Bulgan province until it can be tested with better-suited validation strategies, as well as with the implementation of slight local parameterization. However, this does not definitively disqualify a process-based approach to rangeland biomass, and thus carrying capacity prediction, as the

model and validation strategies could be further fine-tuned until an NPP-based model could be applicable for use in inter-regional or international comparison of overgrazing levels.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. *Scientific Data*, *5*, 1–12. https://doi.org/10.1038/sdata.2017.191

Addison, J., Friedel, M., Brown, C., Davies, J., & Waldron, S. (2012). A critical review of degradation assumptions applied to Mongolia's Gobi Desert. *Rangeland Journal*, *34*(2), 125–137. https://doi.org/10.1071/RJ11013

Al-Jaloudy, M. A., Berkat, O., Tazi, M., Coulibally, A., Dost, M., Fitzherbert, A. R., … Wangdil, K. (2005). *Grasslands of the World*. (J. M. Suttie, S. G. Reynolds, & C. Batello, Eds.). Rome: FAO. Retrieved from http://www.fao.org/3/y8344e00.htm

Angerer, J. (2012). Gobi forage livestock early warning system. *Conducting National Feed Assesments. FAO Animal Production and Health Manual No. 15*, 115–130.

Angerer, J., Han, G., Fujisaki, I., & Havstad, K. (2008). Climate change and ecosystems of Asia with emphasis on inner Mongolia and Mongolia. *Rangelands*, *30*(3), 46–51. https://doi.org/10.2111/1551-501X(2008)30[46:CCAEOA]2.0.CO;2

Avitabile, V., Baccini, A., Friedl, M. A., & Schmullius, C. (2012). Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. *Remote Sensing of Environment*, *117*, 366–380. https://doi.org/10.1016/j.rse.2011.10.012

Bedunah, D. J., & Angerer, J. P. (2012). Rangeland degradation, poverty, and conflict: How can rangeland scientists contribute to effective responses and solutions? *Rangeland Ecology and Management*, *65*(6), 606–612. https://doi.org/10.2111/REM-D-11-00155.1

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*, 281–305.

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., … Jones, Z. M. (2016). Mlr: Machine learning in R. *Journal of Machine Learning Research*, *17*(September).

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1201/9780367816377-11

Brenning, A. (2005). Spatial prediction models for landslide hazards: Review, comparison and evaluation. *Natural Hazards and Earth System Science*, *5*(6), 853–862.

https://doi.org/10.5194/nhess-5-853-2005

Bruegger, R. A., Jigjsuren, O., & Fernández-Giménez, M. E. (2014). Herder observations of rangeland change in Mongolia: Indicators, causes, and application to community-based management. *Rangeland Ecology and Management*, *67*(2), 119–131. https://doi.org/10.2111/REM-D-13-00124.1

de Leeuw, J., Rizayeva, A., Namazov, E., Bayramov, E., Marshall, M. T., Etzold, J., & Neudert, R. (2019). Application of the MODIS MOD 17 Net Primary Production product in grassland carrying capacity assessment. *International Journal of Applied Earth Observation and Geoinformation*, *78*(September 2018), 66–76. https://doi.org/10.1016/j.jag.2018.09.014

Devries, B., Pratihast, A. K., Verbesselt, J., Kooistra, L., & Herold, M. (2016). Characterizing forest change using community-based monitoring data and landsat time series. *PLoS ONE*, *11*(3), 1–26. https://doi.org/10.1371/journal.pone.0147121

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., … Saleem, M. (2017). An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience*, *67*(6), 534–545. https://doi.org/10.1093/biosci/bix014

Eisfelder, C., Klein, I., Niklaus, M., & Kuenzer, C. (2014). Net primary productivity in Kazakhstan, its spatio-temporal patterns and relation to meteorological variables. *Journal of Arid Environments*, *103*, 17–30. https://doi.org/10.1016/j.jaridenv.2013.12.005

Eisfelder, C., Kuenzer, C., & Dech, S. (2012). Derivation of biomass information for semi-arid areas using remote-sensing data. *International Journal of Remote Sensing*, *33*(9), 2937–2984. https://doi.org/10.1080/01431161.2011.620034

Ellis, J., & Swift, D. (1988). Stability of African pastoral ecosystems: alternate paradigms and implications for development. *Journal of Range Management*, *41*(6), 450–459. https://doi.org/10.2307/3899515

Erdenesan, E. (2018). Mongolia experience on generating livestock statistics. Retrieved October 17, 2019, from http://www.fao.org/3/bu567en/bu567en.pdf

ESA. (2018). Copernicus Global Land cover 100 m, (1), 1–44.

Fernandez-Gimenez, M. E., & Allen-Diaz, B. (1999). Testing a non-equilibrium model of rangeland vegetation dynamics in Mongolia. *Journal of Applied Ecology*, *36*(6), 871–885.

https://doi.org/10.1046/j.1365-2664.1999.00447.x

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Gao, W., Angerer, J. P., Fernandez-Gimenez, M. E., & Reid, R. S. (2015). Is Overgrazing A Pervasive Problem Across Mongolia? An Examination of Livestock Forage Demand and Forage Availability from 2000 to 2014. *Proceedings of the Trans-Disciplinary Research Conference: Building Resilience of Mongolian Rangelands,* (January), 35–41.

Gu, Y., & Wylie, B. K. (2015). Developing a 30-m grassland productivity estimation map for central Nebraska using 250-m MODIS and 30-m Landsat-8 observations. *Remote Sensing of Environment*, *171*, 291–298. https://doi.org/10.1016/j.rse.2015.10.018

Huang, J., Yu, H., Guan, X., Wang, G., & Guo, R. (2016). Accelerated dryland expansion under climate change. *Nature Climate Change*, *6*(2), 166–171. https://doi.org/10.1038/nclimate2837

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, (83), 195–213. https://doi.org/10.1080/0965156x.2013.836857

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Hui, D., & Jackson, R. B. (2006). Geographical and interannual variability in biomass partitioning in grassland ecosystems: A synthesis of field data. *New Phytologist*, *169*(1), 85–93. https://doi.org/10.1111/j.1469-8137.2005.01569.x

Hurni, K., Heinimann, A., & Würsch, L. (2017). Google Earth Engine Image Pre-processing Tool : Examples.

IPCC. (2006). *2006 IPCC Guidelines for National Greenhouse Gas Inventories* (Vol. 4).

John, R., Chen, J., Giannico, V., Park, H., Xiao, J., Shirkey, G., … Qi, J. (2018). Grassland canopy cover and aboveground biomass in Mongolia and Inner Mongolia: Spatiotemporal estimates and controlling factors. *Remote Sensing of Environment*, *213*(July 2017), 34–48. https://doi.org/10.1016/j.rse.2018.05.002

Kauth, R. J. (1976). Tasselled Cap - a Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen By Landsat., 41–51.

Kuhn, A. M., Weston, S., Keefer, C., Coulter, N., Quinlan, R., & Kuhn, M. M. (2013). *Package ‘ Cubist .’*

Kuhn, M., Splitting, D., Performance, E., Pre-processing, D., Models, T. T., Support, T. A., … Processing, P. (2013). useR ! 2013 Modeling Conventions in R.

Liu, S., Su, X., Dong, S., Cheng, F., Zhao, H., Wu, X., … Li, J. (2015). Modeling aboveground biomass of an alpine desert grassland with SPOT-VGT NDVI. *GIScience and Remote Sensing*, *52*(6), 680–699. https://doi.org/10.1080/15481603.2015.1080143

Lund, G. (2007). Accounting for the World’s Rangelands. *Rangelands*, *29*(1), 3–10. https://doi.org/https://doi.org/10.2111/1551-501X(2007)29[3:AFTWR]2.0.CO;2

Mutanga, O., Adam, E., & Cho, M. A. (2012). High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation*, *18*(1), 399–406. https://doi.org/10.1016/j.jag.2012.03.012

National Statistics Office of Mongolia. (2018a). NUMBER OF HERDER HOUSEHOLDS, by region, aimag, the capital, soum and bag. Retrieved from http://www.1212.mn/tables.aspx?TBL_ID=DT_NSO_1001_027V1

National Statistics Office of Mongolia. (2018b). POPULATION OF MONGOLIA, by region, aimag and the capital, urban and rural. Retrieved October 22, 2019, from http://www.1212.mn/Maps.aspx?TBL_ID=DT_NSO_0300_004V1&year=2018&CHAR_ITM_ID=T1&OV_L1_ID=11

National Statistics Office of Mongolia. (2019). LIVESTOCK STATISTICS IN MONGOLIA. Retrieved October 17, 2019, from http://www.fao.org/fileadmin/templates/ess/documents/apcas26/presentations/APCAS-16-6.3.2_-_Bangladesh_-_Fisheries_Statistics_in_Bangladesh.pdf

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modify soil adjust vegetation index. *Remote Sensing of Environment*, *126*, 119–126.

Reeves, M. C., Washington-Allen, R. A., Angerer, J., Hunt, E. R., Kulawardhana, R. W., Kumar, L., … Ramsey, R. D. (2015). Global view of remote sensing of rangelands: Evolution, applications, future pathways. *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing*, 237–275. https://doi.org/10.1201/b19322

Reuter, H. I., Nelson, A., & Jarvis, A. (2007). An evaluation of void-filling interpolation

methods for SRTM data. *International Journal of Geographical Information Science*, *21*(9), 983–1008. https://doi.org/10.1080/13658810601169899

Rouse, J., RH, H., Schell, J. A., & Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. In *Third ERTS Symposium, NASA SP-351* (pp. 309–317).

Sainbuyan, B. (2016). *Methodologies for assessing pasture utilization and its changes*.

Segal, M. R. (2004). Machine learning benchmarks and random forest regression, (May).

Sexton, J. O., Song, X. P., Feng, M., Noojipady, P., Anand, A., Huang, C., … Townshend, J. R. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth*, *6*(5), 427–448. https://doi.org/10.1080/17538947.2013.786146

Tucker, C. J., Vanpraet, C., Boerwinkel, E., & Gaston, A. (1983). Satellite remote sensing of total dry matter production in the Senegalese Sahel. *Remote Sensing of Environment*, *13*(6), 461–474. https://doi.org/10.1016/0034-4257(83)90053-6

United Nations. (2016). Transforming our World: The 2030 Agenda for Sustainable Development. Retrieved September 24, 2019, from https://sustainabledevelopment.un.org/content/documents/21252030 Agenda for Sustainable Development web.pdf

Vetter, S. (2005). Rangelands at equilibrium and non-equilibrium: Recent developments in the debate. *Journal of Arid Environments*, *62*(2), 321–341. https://doi.org/10.1016/j.jaridenv.2004.11.015

Wang, X., Zhang, J., Shahid, S., Guan, E., Wu, Y., Gao, J., & He, R. (2016). Adaptation to climate change impacts on water demand. *Mitigation and Adaptation Strategies for Global Change*, *21*(1), 81–99. https://doi.org/10.1007/s11027-014-9571-6

Wehrden, H. Von, Hanspach, J., Kaczensky, P., Fischer, J., & Wesche, K. (2012). Global assessment of the non-equilibrium concept in rangelands. *Ecological Applications*, *22*(2), 393–399. https://doi.org/10.1890/11-0802.1

WILCOXON, F. (1945). Individual comparisons of grouped data by ranking methods. *Journal of Economic Entomology*, *39*(6), 269. https://doi.org/10.1093/jee/39.2.269

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079

Wilson, E. H., & Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, *80*(3), 385–396. https://doi.org/10.1016/S0034-4257(01)00318-2

Wright, M. N., & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), 1–17. https://doi.org/10.18637/jss.v077.i01

Yu, H., Lee, J.-Y., Lee, W.-K., Lamchin, M., Tserendorj, D., Choi, S., … Kang, H. D. (2013). Feasibility of Vegetation Temperature Condition Index for monitoring desertification in Bulgan, Mongolia. *Korean Journal of Remote Sensing*, *29*(6), 621–629. https://doi.org/10.7780/kjrs.2013.29.6.5

Yu, L., Zhou, L., Liu, W., & Zhou, H. K. (2010). Using Remote Sensing and GIS Technologies to Estimate Grass Yield and Livestock Carrying Capacity of Alpine Grasslands in Golog Prefecture, China. *Pedosphere*, *20*(3), 342–351. https://doi.org/10.1016/S1002-0160(10)60023-9

Zhao, M., Heinsch, F. A., Nemani, R. R., & Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, *95*(2), 164–176. https://doi.org/10.1016/j.rse.2004.12.011

Appendices

A Livestock conversion and grazing range

Table 14: Livestock conversion matrix. SEU = Sheep equivalent unit.

Livestock	Livestock in SEU
1 Horse	7 SEU
1 Cow	6 SEU
1 Camel	5 SEU
1 Sheep	1 SEU
1 Goat	0.9 SEU

Table 15: Livestock grazing range mantrix. 'Large' livestock are: Horse, Cattle, Camel. 'Small' livestock are Sheep and Goat. Summer-Autumn distances were used.

Season	Livestock Size	
	Large	Small
Summer-Autumn	5 km	3 km
Winter-Spring	4 km	2 km

B Workflow

B.1. Process-based model workflow chart

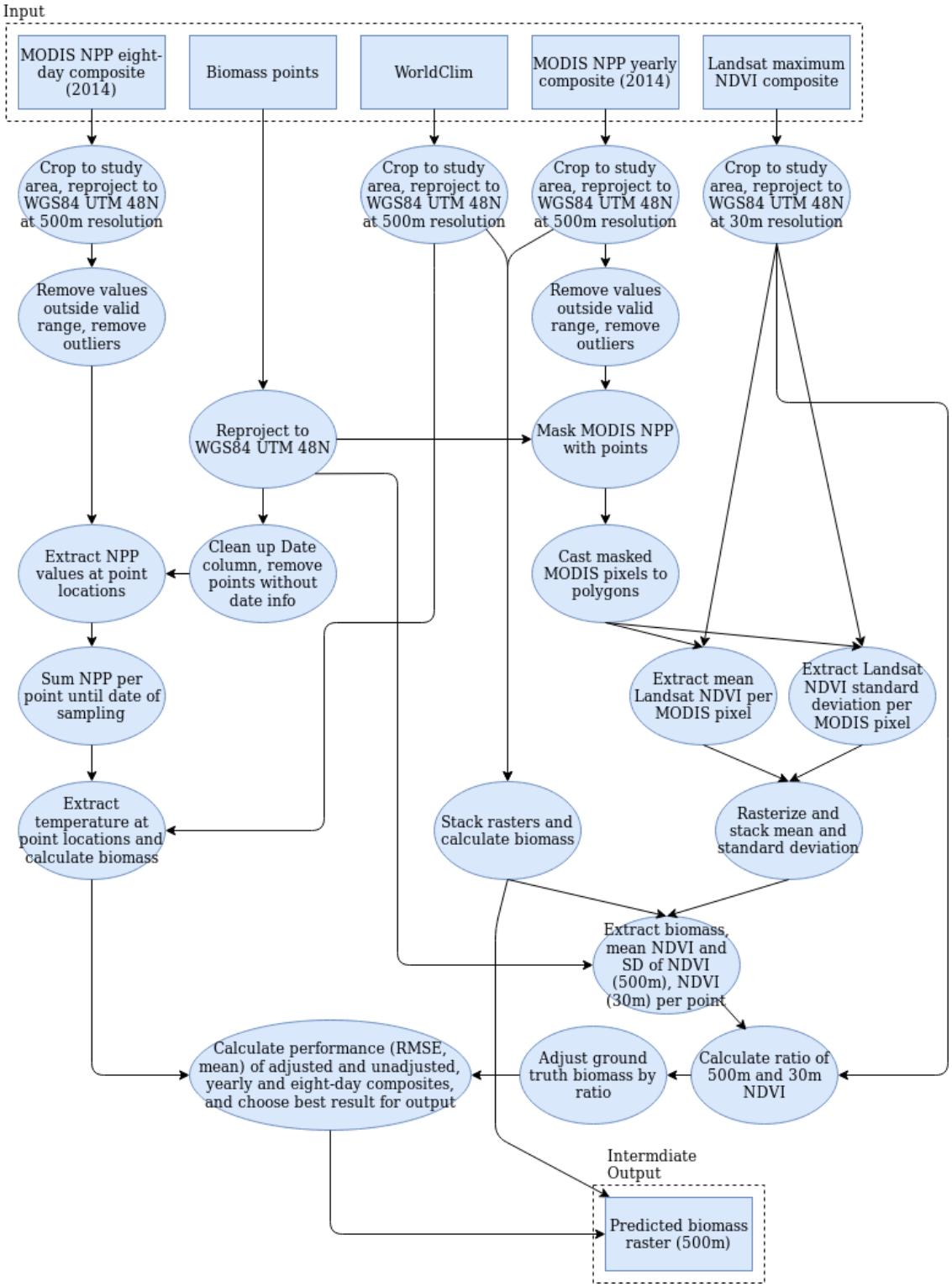


Figure 31: Workflow chart showing the steps taken to output the aboveground biomass prediction raster using the process-based model.

B.2. Empirical model workflow chart

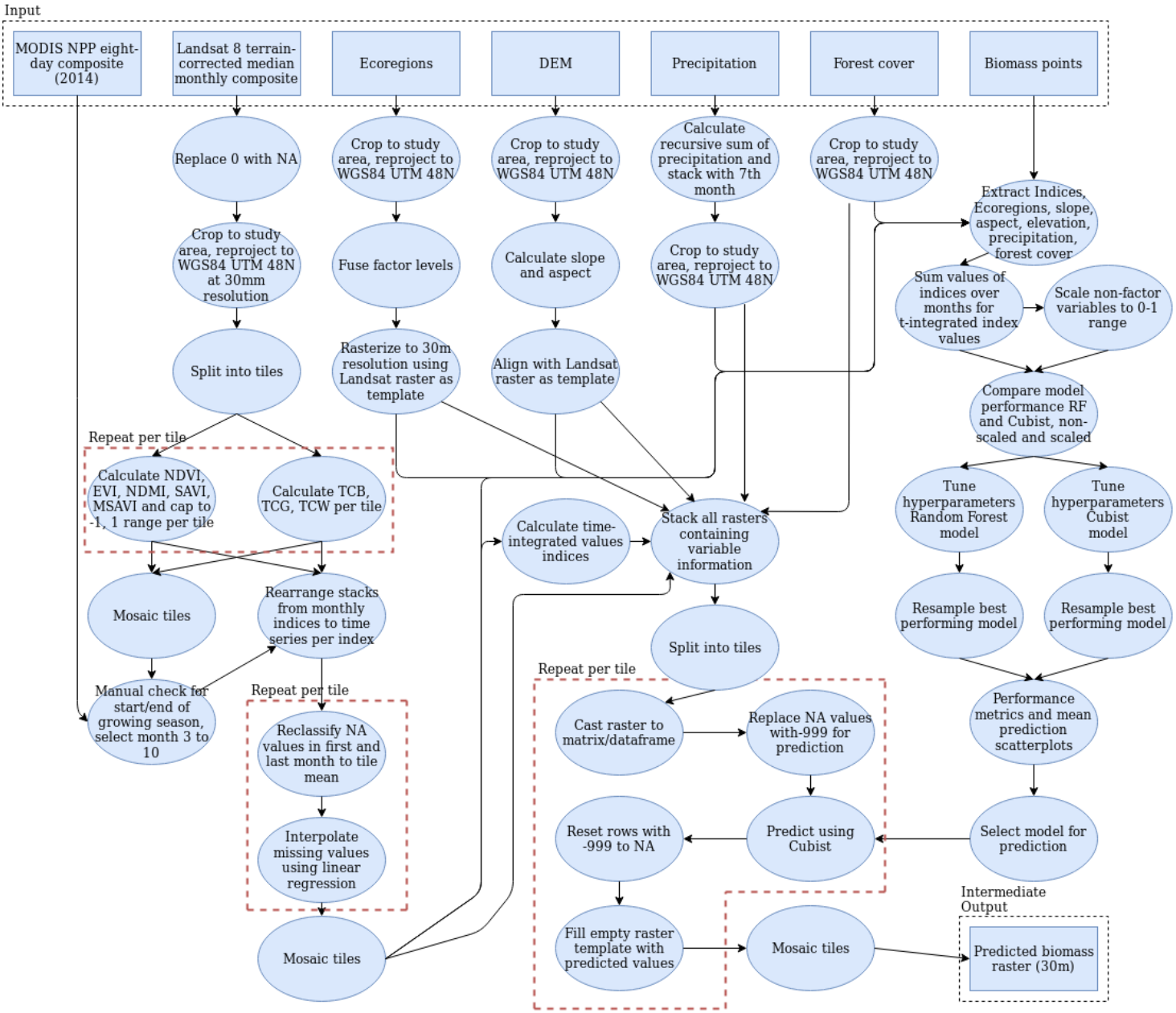


Figure 32: Workflow chart showing the steps taken to output the aboveground biomass prediction raster using the empirical model.

B.3. Stocking density and forage intake workflow chart

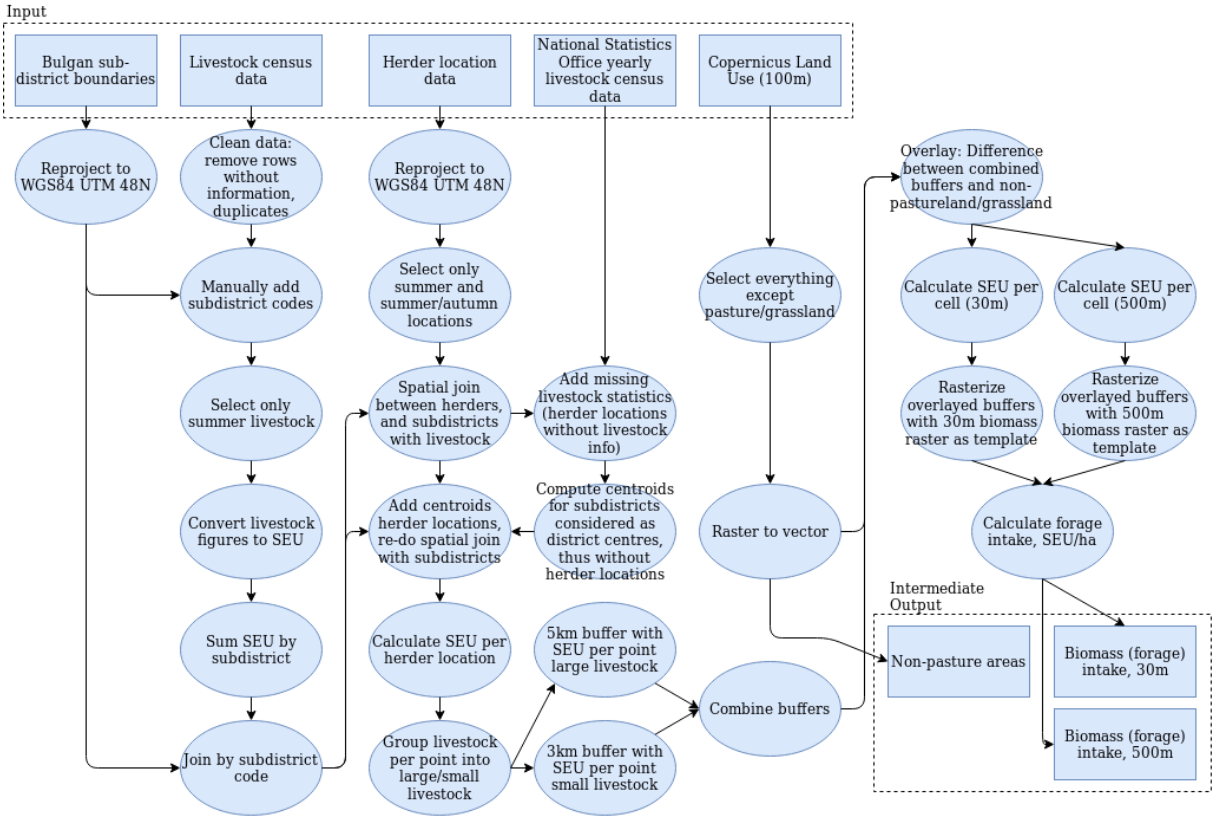


Figure 33: Workflow chart depicting the steps taken in order to output the forage intake rasters.

B.4. Over- and undercapacity in biomass availability workflow chart

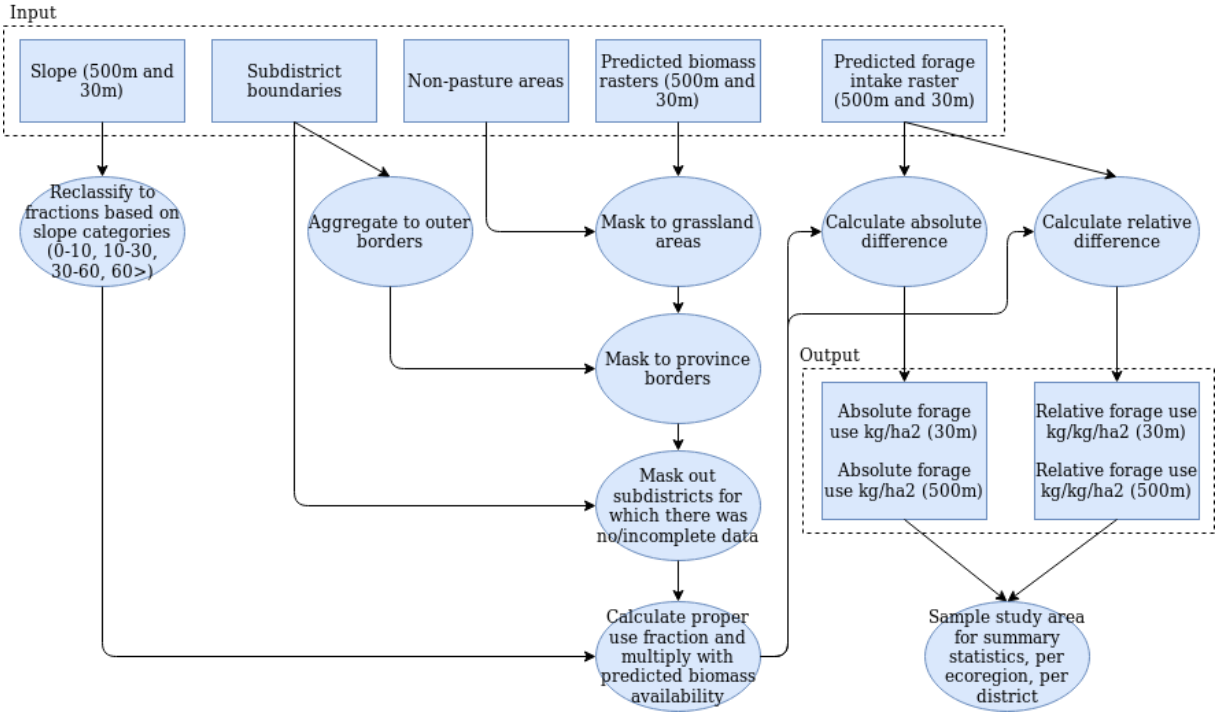


Figure 34: Workflow chart depicting the steps taken to output relative and absolute forage use rasters.