Predicting aboveground biomass in Arctic landscapes using very high spatial resolution satellite imagery and field sampling

Aleksi Räsänen, corresponding author

Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, and Helsinki Institute of Sustainability Science (HELSUS), P.O Box 65, FI-00014 University of Helsinki, Finland, aleksi.rasanen@helsinki.fi

Department of Geography, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

Sari Juutinen

Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, and Helsinki Institute of Sustainability Science (HELSUS), P.O Box 65, FI-00014 University of Helsinki, Finland, sari.juutinen@helsinki.fi

Mika Aurela

Finnish Meteorological Institute, PO Box 503, FI-00101 Helsinki, Finland, mika.aurela@fmi.fi

Tarmo Virtanen

Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, and Helsinki Institute of Sustainability Science (HELSUS), P.O Box 65, FI-00014 University of Helsinki, Finland, tarmo.virtanen@helsinki.fi

word count: 9383

Predicting aboveground biomass in Arctic landscapes using very high spatial resolution satellite imagery and field sampling

Remote sensing based biomass estimates in Arctic areas are usually produced using coarse spatial resolution satellite imagery, which is incapable of capturing the fragmented nature of tundra vegetation communities. We mapped aboveground biomass using field sampling and very high spatial resolution (VHSR) satellite images (QuickBird, WorldView-2 and WorldView-3) in four different Arctic tundra or peatland sites with low vegetation located in Russia, Canada, and Finland. We compared site-specific and cross-site empirical regressions. First, we classified species into plant functional types and estimated biomass using easy, non-destructive field measurements (cover, height). Second, we used the cover/height-based biomass as the response variable and used combinations of single bands and vegetation indices in predicting total biomass. We found that plant functional type biomass could be predicted reasonably well in most cases using cover and height as the explanatory variables (adjusted R^2 0.21–0.92), and there was considerable variation in the model fit when the total biomass was predicted with satellite spectra (adjusted R^2 0.33–0.75). There were dissimilarities between cross-site and site-specific regression estimates in satellite spectra based regressions suggesting that the same regression should be used only in areas with similar kinds of vegetation. We discuss the considerable variation in biomass and plant functional type composition within and between different Arctic landscapes and how well this variation can be reproduced using VHSR satellite images. Overall, the usage of VHSR images creates new possibilities but to utilize them to full potential requires similarly more detailed in-situ data related to biomass inventories and other ecosystem change studies and modelling.

1. Introduction

Biomass is a key parameter for tracking plant productivity, which is central to the flow of energy and nutrients in an ecosystem (Epstein et al. 2012; van der Wal and Stien 2014). In Arctic tundra and other northern landscapes with low-growth vegetation, knowledge of the biomass distribution is a prerequisite for understanding and mapping changes in key ecosystem parameters such as the carbon cycle and permafrost dynamics

(Chen, Li, et al. 2009; Epstein et al. 2012). Biomass patterns have been estimated for decades using satellite images which allow the mapping of vast areas with little field work (Laidler and Treitz 2003; Raynolds, Walker, and Maier 2006; Epstein et al. 2012; Buchhorn, Raynolds, and Walker 2016).

In Arctic environments, satellite based estimates of biomass distribution have mostly been carried out using rather coarse spatial resolution images (Walker et al. 2003; Heiskanen 2006; Raynolds, Walker, and Maier 2006; Epstein et al. 2012; Raynolds et al. 2012; Buchhorn et al. 2013; Doiron et al. 2013; Johansen and Tommervik 2014; Berner et al. 2018), such as Landsat (30 m pixel size) (Heiskanen 2006; Johansen and Tommervik 2014; Berner et al. 2018), MODIS (250 m pixel size) (Westergaard-Nielsen et al. 2015), and AVHRR (>1 km pixel size) (Walker et al. 2003; Raynolds, Walker, and Maier 2006; Epstein et al. 2012; Raynolds et al. 2012; Buchhorn et al. 2013; Doiron et al. 2013). Although the images with coarse spatial resolution have high temporal resolution and they have proved to be suitable for circumpolar studies and detecting coarse-scale biomass patterns (coefficient of determination (R^2) up to 0.89) (Walker et al. 2003), they are incapable of representing the fragmented nature of tundra environment and fine-scale changes in vegetation and carbon dynamics (Laidler and Treitz 2003; Virtanen and Ek 2014; Siewert et al. 2015; Beamish et al. 2017).

Very high spatial resolution (VHSR, spatial resolution 0.5–2.5 m) satellite images could offer a potential method to map landscape-scale biomass distribution in an ecologically sound pixel size (Laidler and Treitz 2003; Virtanen and Ek 2014). In addition, field data is usually collected in small plots, which are more comparable to the pixel size of VHSR than coarser resolution imagery. However, the use of VHSR images has been modest in biomass prediction (Fuchs et al. 2009; Atkinson and Treitz 2013; Collingwood et al. 2014; Greaves et al. 2016). In Canadian tundra landscapes, reasonably high prediction capabilities (R^2 of 0.55 to 0.79) have been obtained with VHSR images (Atkinson and Treitz 2013; Collingwood et al. 2014). When reflectance data are combined with other types of data, such as radar and topographical data (Chen, Blain, et al. 2009; Collingwood et al. 2014) or LiDAR (Greaves et al. 2016), higher explanatory power can be obtained. The use of VHSR satellite imagery, radar and LiDAR data is hampered by the low availability of such data in VHSR at the global scale (Sinha et al. 2015; Steele-Dunne et al. 2017), and logistical and practical issues limit data collection possibilities with remotely piloted aircraft systems in remote Arctic locations. Nevertheless, the use and availability of LiDAR, radar, and VHSR optical images is increasing and they present an interesting research frontier in Arctic vegetation studies. So far, to the best of our knowledge, there are no studies in which biomass has been estimated using VHSR data and compared in various tundra environments across the circumpolar Arctic, although there have been calls for biome-wide observation methodologies (Walker et al. 2016). Therefore, there is a need to test whether cross-site biomass models that include data from divergent Arctic landscapes can be developed.

To get field validation data for remote sensing studies, harvested biomass samples at a plot scale are needed (Hope, Kimball, and Stow 1993; Walker et al. 2003; Raynolds, Walker, and Maier 2006; Kushida et al. 2015; Greaves et al. 2016). Previously, it has been suggested that non-destructive methods, such as estimations of height, %-cover and volume of plant species or plant functional types (PFTs, which are groups of plants with functional similarity and similar growth form), are sufficient for estimating plot-scale biomass in landscapes with low-growth vegetation and allow the collection of larger validation sets (Chen, Li, et al. 2009; Axmanova et al. 2012; Suvanto, Le Roux, and Luoto 2014). Also in this case, models have been developed and applied only in one specific location and studies that test whether one model can be applied in various tundra environments are lacking.

Our objective was to predict biomass distribution by using VHSR satellite imagery in different Arctic tundra and peatland communities and evaluate whether the same predictive regressions can be applied across circumpolar Arctic sites. Therefore, in this study, we first estimated PFT-specific biomass using harvested biomass as the response variable and field-measured height and %-cover as predictors. Second, we estimated total biomass using modelled cover/height-based biomass as the response variable and single bands and vegetation indices of VHSR satellite images as predictors. At both steps, we compared different predictor combinations and transformations as well as site-specific and cross-site regressions. We concluded our study by discussing how general the regressions are for biomass prediction and how different PFTs contribute to biomass across circumpolar northern landscapes with low-growth vegetation.

2. Materials and methods

2.1.Study sites

We included four different study sites which present a continuum from northern boreal to sub-Arctic and to Arctic landscapes: Sodankylä in Finland, northwestern (NW) Russia, Herschel in Canada and Tiksi in Russia (Figures 1 and 2, Table 1). All the study sites are characterized by low-growth vegetation, but having some variation in landscape patterns and vegetation communities, which make them a good combination for a circumpolar comparison and testing whether simple and general approach could be used for spatial modelling.

[FIGURE 1 approximately here]

[FIGURE 2 approximately here]

[TABLE 1 approximately here]

The Sodankylä study site is an open north-boreal fen in northern Finland (Figure 1). The vegetation pattern consists of strings with shrubs and birch trees, and lawns and flarks with *Sphagnum* and brown mosses and sedges (Figure 2a, see Appendix 1 in the supplemental material for the dominant species). For further site description see (Dinsmore et al. 2017).

The study sites Khosedayu, Rogovaya 1 and 2, and Seida located in NW Russia within 150 km to each other near the Ural Mountains (Figure 1). Generally, the vegetation, geomorphology, climate and soil characteristics were similar in these four sites; therefore, they are analyzed together. The vegetation belongs to the ecotone of forest-tundra – tundra zones, and is a mosaic of peatlands, heaths with different kind of shrubs and willow thickets and meadows along streams (Figure 2b, Appendix 1 in the supplemental material, see also Hugelius et al. (2011) and Virtanen and Ek (2014)).

The Canadian study site is located on Herschel Island, which has an area of approximately 100 km^2 and is located a few kilometres off the Yukon Coast in the southern Beaufort Sea, Canada (Figure 1). Prominent geomorphic features are smooth hills, river channels, and numerous retrogressive thaw slumps. In the lowland tundra, there are several different herb rich plant community types but also other type of vegetation (Figure 2c, Appendix 1 in the supplemental material, see also (Myers-Smith et al. 2011) and (Obu et al. 2017)).

The Tiksi study site is located near the coast of the Laptev Sea about 120 km southeast of the Lena River delta in Siberia, Russia (Figure 1). The site consists of relatively flat lowlands and gently sloping hillslopes with elevations at 200–300 m a.s.l. Sedge and moss dominated peatlands, tundra heaths with low shrubs, and rocky and lichen covered surfaces alternate in the site (Figure 2d, Appendix 1 in the supplemental material, see also Grosswald et al. (1992), Juutinen et al. (2017) and Mikola et al. (2018)).

2.2.Biomass field data

We measured the biomass, %-cover and height of the following PFTs: (1) dwarf shrubs, (2) herbs, (3) graminoids, (4) dwarf birch (*Betula nana)*, (5) *Salix spp.* and other tall shrubs (height ≤ 1.5 m), and (6) mosses (height not measured). Examples of the common species or genera included in each PFT for each study site are listed in Appendix 1 in the supplemental material. The PFT classification we used was a slight modification of the one presented by Chapin III et al. (1996), and was used earlier at one of the study sites (Hugelius et al. 2011).

In each study site, we sampled 48 to 182 circular plots either randomly or using transects (Table 1, Appendix 3 in the supplemental material). When sampling, plots of each major vegetation type were included, and plots w were representative of the overall landscape in each study site. Sampling set-up differed between study sites, because data were collected during different field campaigns and projects over several years. Field plots were classified in the field into vegetation types which were defined in previous studies in each site (Smith et al. 1989; Virtanen and Ek 2014; Juutinen et al. 2017; Obu et al. 2017). Each plot had a radius of 5 meters and contained 3 or 4 rectangular subplots of sides 30–50 cm in length (depending on the study site). These subplots were located 1.0–2.5 m from the plot centroid at right angles to each other. We visually estimated the %-cover of each PFT and measured the mean height of each PFT using a ruler in each subplot. One of the subplots was harvested during the peak growing season to measure aboveground biomass. All vascular plant material was collected. For mosses, we collected a subsample of 5 cm x 5 cm with variable depths, determined by the height of photosynthesizing part of the mosses. The harvested biomass was sorted by PFT, oven dried at 60°C for 24 hours, and weighed. In Sodankylä and NW Russia, i.e. sites with scattered trees (height > 1.5 m), tree biomass in 5 m radius circular plots was calculated on the basis of tree height, mean stem diameter at breast height and basal area by using allometric equations (Nyyssönen 1955; Varmola and Vuokila 1986; Alexeyev et al. 1995; Kauppi, Tomppo, and Ferm 1995; Shepashenko, Shvidenko, and Nilsson 1998; Starr, Hartman, and Kinnunen 1998; Korpela 2001) as specified in Appendix 2 in the supplemental material.

2.3.Satellite imagery and its preprocessing

From each study site, we used one cloud free QuickBird (2.4 m pixel size in multispectral bands), WorldView-2 (2 m pixel size in multispectral bands), or WorldView-3 (1.6 m pixel size in multispectral bands) VHSR satellite image (Digital Globe, Westminster, CO, USA) acquired at approximately peak biomass and at approximately the same time as the field work in the respective study sites (Table 1).

Generally, the temporal availability of VHSR satellite images is quite low in the Arctic (Stow et al. 2004; Westergaard-Nielsen et al. 2013). Some of the images were taken some days earlier in the summer than the field work, but these were the best matching images at the time of image acquisition. In Tiksi and Seida, images were taken some years before the field work but phenologically in a relatively similar phase as the field work. This should not affect image interpretation due to small or lacking disturbance and slow vegetation growth in these study sites. This interpretation is based on our field observations, a MODIS trend analysis (Appendix 4 in the supplemental material) and global comparisons in which only small changes have been observed in these sites (Epstein et al. 2012; Myers-Smith et al. 2015). Nevertheless, the timing of satellite images may produce some uncertainties in our analysis, and there might be small-scale dynamics which cannot be observed in the MODIS images.

Satellite images were first atmospherically corrected and transformed to ground reflectance values using the dark object subtraction method (Chavez 1988; Song et al. 2001). After atmospheric correction, images were orthorectified to match the field work data. The geometric error based on our visual interpretations was a maximum of a couple of meters. Due to lack of high precision GPS data or precisely georeferenced maps or images, it was not possible to calculate the exact accuracy of the images. However, the accuracy should be adequate, because we used 5 m radius circular plots (plot area 78.5 m^2) when predicting biomass with satellite images. In each plot, there were between 13.6 (Quickbird) and 30.7 (WorldView-3) pixels. Satellite data values were averaged to obtain a mean value per circular plot.

2.4.Data analysis overview

We first estimated PFT-specific biomass using PFT %-cover and height measured in the field as explanatory variables in the regression (referred as biomass-cover/height regressions). Second, we used the predicted total biomass for each 5 m radius plot as the response variable when developing regressions to estimate total aboveground biomass distribution based on VHSR satellite images (referred as biomass-satellite spectra regressions). For all sites and both regression steps, we tested both site-specific regressions with data from one study site only and cross-site regressions in which data from all study sites were used. As tree biomass was calculated using existing allometric equations, biomass-cover/height regressions were not built for them, but they were included in the biomass-satellite spectra regressions. We carried out all biomass-cover/height and biomass-satellite spectra estimations with ordinary least squares linear regressions. We acknowledge that there are also more sophisticated modelling frameworks for biomass (Collingwood et al. 2014; Greaves et al. 2016) but our goal was to test simple regression equations that can be easily interpreted and applied at different sites. We performed data analyses in R 3.2.2. (R Core Team 2015), using the car (Fox and Weisberg 2011), caret (Kuhn et al. 2016) and MASS (Venables and Ripley 2003) packages.

2.5. Plant functional group specific regressions for biomass based on vegetation height and %-cover

We predicted area-normalized PFT biomass for the subplots using the field measured %-cover and average height of the respective groups in the subplot as explaining factors. The data from harvested subplots were used to build the regressions. For all variables, we tested the transformations suited to our data distribution in order to achieve better normality for data and to find the best fitting regressions (McDonald 2014). For biomass and height, we used the following transformations (1) no transformation, (2) square root, (3) natural logarithm + 1. For %-cover, we tried (1) no transformation and (2) arcsine transformation (asin(sqrt(%-cover/100))), as %-cover distribution varies between 0 and 1 (McDonald 2014).

For each functional group, we tried all the possible parameter combinations with different transformations. We tested regressions with either one or both explanatory variables but did not include two explanatory variables in the same regression if their Pearson correlation was >0.7. We formed the empirical relationships separately for each study site and also carried out cross-site regressions. We evaluated the regressions based on their root mean square error (RMSE) and chose the regressions with the lowest RMSE value. Once the best regression was determined for each PFT, it was applied to all subplots. Some of the regression equations had a negative intercept and predicted negative biomass values for a small minority of the subplots. In these cases, the biomass was set to 0 for the respective PFT in the subplot. Finally, we added up the biomass values of every PFT to calculate the total biomass per area of each subplot. For some functional groups at some sites (*Salix* spp. at Sodankylä and *Betula nana* on Herschel), species were present only in one or two harvested subplots. In other situations, we did not harvest biomass but measured PFT %-cover (mosses on Herschel). In these cases, we used the cross-site biomass-cover/height regression estimations for the respective functional groups when we summed up site-specific total biomass.

2.6.Predicting total biomass using VHSR satellite images

We built biomass-satellite spectra regressions to predict biomass using estimated cover/height-based total biomass as the response variable and individual spectral bands and spectral indices of VHSR satellite images as predictors (Table 2). We carried out three different types of regressions: site-specific cover/height predictions combined with satellite image data from one site, cross-site cover/height predictions combined with satellite image data from one site, and cross-site cover/height predictions with satellite image data from all study sites.

[TABLE 2 APPROXIMATELY HERE]

Estimated cover/height-based biomass was calculated as the mean of the predicted subplot biomass values in the respective 5 m radius plot. To evaluate the uncertainty in biomass-cover/height regressions, we also carried out alternative biomass-satellite spectra regressions, in which we used harvested subplot-scale biomass data as response variable. In these alternative calculations, used site-specific cover/height-based moss biomass estimate for Tiksi and Seida 2016 data and cross-site cover/height-based moss biomass estimate for Herschel as mosses were not systematically harvested in these datasets. For Sodankylä and NW Russia, we included tree biomass in all biomass-satellite spectra regressions.

Individual spectral bands consisted of blue, green, red, and near infrared (NIR) for the Quickbird images. For the WorldView images, the following bands were also included: coastal, yellow, red-edge, and NIR2. We calculated the mean value per band or index for the 5 m radius circle corresponding to each plot location.

We transformed biomass values with a natural logarithm as this transformation has been used usually in tundra biomass studies, and it has been found in several studies that there is a logarithmic relationship between biomass and satellite spectra (Walker et al. 2003; Raynolds et al. 2012; Atkinson and Treitz 2013; Berner et al. 2018). Satellite image values were not transformed. In each regression, those predictors whose correlations were < 0.7 were chosen in the same model. All combinations were tested, and also regressions with only one vegetation index. Variables were selected on the basis of Akaike's Information Criteria and 10-fold cross-validation. Finally, we compared the regressions built using different explanatory variable sets by comparing the RMSE values.

3. Results

3.1.Aboveground biomass of different tundra and peatland vegetation types

The highest total biomass values were found at the NW Russian sites consisting of both mineral tundra and peatland. The southernmost site Sodankylä, a treeless fen, had lower total biomass values than NW Russia study sites. The most Arctic site, Tiksi had the lowest total biomass. The proportion of different PFTs varied among the study sites. At Sodankylä, a major proportion of the biomass consisted of mosses; at NW Russia, *Betula nana* and other shrubs had high biomass values; herbaceous biomass was higher on Herschel than at other study sites, whereas at Tiksi, graminoids contributed most to total biomass (when mosses were excluded) (Table 3). Similar trends could also be seen in the average %-cover and height of PFTs, but there were variation in habitat type specific biomass at each study site (Appendix 5 in the supplemental material). At Sodankylä, there were trees in 32% (n = 16) of the 5 m radius plots with an average biomass of 117.2 g m^{-2}, while at NW Russia trees were present in 6.8% (n = 25) of the plots with an average biomass of 903.0 g m^{-2}.

[TABLE 3 APPROXIMATELY HERE]

3.2.Predicting biomass using easily measurable plant height and %-cover

PFT-specific empirical regressions to predict biomass by %-cover and height of plants performed well in most of the cases, with the adjusted coefficient of determination (R^2_{adj}) values varying between 0.21 and 0.92 (Table 4). Overall, the lowest R^2_{ajd} values were obtained for mosses (R^2_{adj} 0.21–0.38), but also in some vascular plant regressions RMSE values were relatively high. There was variation across study sites which PFT regressions had the lowest RMSE and highest R^2_{adj} values.

[TABLE 4 APPROXIMATELY HERE]

Overall, the total predicted cover/height-based biomass values ranged between 0 and 2000 g m^{-2} in the study plots (Figure 3). On average, the biomass was greatest and had the largest variation at NW Russia. Based on site-specific estimate and excluding trees, average total biomass was 423 g m^{-2} and standard deviation 321 g m^{-2}. Sodankylä had relatively high average biomass (282 g m^{-2}) and low standard deviation (115 g m^{-2}), while Herschel (average 196 g m^{-2}, standard deviation 102 g m^{-2}) and Tiksi (163 g m^{-2}, standard deviation 73 g m^{-2}) had low average biomass values and low variation.

[FIGURE 3 APPROXIMATELY HERE]

Cross-site regressions performed quite differently between the study sites (Figures 3 and 4), by underestimating total biomass on Herschel (21% difference) and at NW Russia (2%), and overestimating at Sodankylä (10%) and at Tiksi (3%). At NW

Russia, PFT-specific average patterns between observed and predicted values were close to 1:1 line, whereas at other study sites, there were more evident underestimation or overestimation (Figure 4). NW Russia had the highest number of observations, which may have an undue influence on the regression. In individual subplots and in PFTs, disparities between cross-site and site-specific estimations were often significantly higher than differences between average total site biomass.

[FIGURE 4 APPROXIMATELY HERE]

3.3.Using VHSR imagery to estimate total aboveground biomass distribution

In biomass-satellite spectra regressions, R^2_{adj} values ranged between 0.33 and 0.75 (Table 5). The best fits were obtained on Herschel and at Tiksi, whereas at Sodankylä the R^2_{adj} values were seemingly low and at NW Russia RMSE values high. RMSE values in cross-site biomass-satellite spectra regression were larger than in site-specific biomass-satellite spectra regressions, with the RMSE value being especially high on Herschel (Table 6). Cross-site biomass-satellite spectra regression overestimated biomass values for Herschel, and underestimated for NW Russia and Tiksi (Figure 5, Table 6). At Sodankylä, there was overestimation in plot-specific predicted values and underestimation in the landscape (Table 6). Alternative biomass estimations having harvested data as the response variable had higher RMSE values than biomass estimations using cover/height-modelled biomass as the response variable. The differences in average biomass values between regressions using cover/height-based biomass estimate and harvested biomass were small at Sodankylä and Tiksi and a little higher at NW Russia and on Herschel (Table 6). Finally, there was fine-scale spatial variation in biomass distribution across the landscapes, and spatial pattern of biomass was divergent in different study sites (Figure 6).

[TABLE 5 APPROXIMATELY HERE]

[TABLE 6 APPROXIMATELY HERE]

[FIGURE 5 APPROXIMATELY HERE]

[FIGURE 6 APPROXIMATELY HERE]

4. Discussion

Aboveground plant biomass in tundra environments can be predicted reasonably well at the plot scale with easily measurable field data (height, %-cover) (Table 4, Figure 4). This is also supported by the fact that biomass-satellite spectra regressions using cover/height modelled biomass as the response variable had lower RMSE and relatively similar average biomass estimate than biomass-satellite spectra regressions using harvested biomass as the response variable (Table 6). The finding suggests that it is more recommendable to measure plant cover and height in a larger area and estimate biomass based on these measurements than to use only small harvested biomass samples when carrying out biomass-satellite spectra models.

Previously, it has been shown that both %-cover and height information are needed for the most accurate biomass predictions at the plot scale (Chen, Li, et al. 2009; Axmanova et al. 2012; Suvanto, Le Roux, and Luoto 2014). Our results show instead that in some sites and in some PFTs, the lowest RMSE values were obtained with %-cover measurements only, but in most regressions for vascular plants, height measurements were needed for the best predictions (Table 4). Biomass-cover/height regressions had higher R^2_{adj} values for vascular plants than for mosses. The poor regression performance of mosses compared to vascular plants could be related to small size of harvested moss samples, to moisture content of the mosses as changing moisture changes the volume, colour and productivity of mosses and to the heterogeneity of the growth forms of moss genera. Possibly separate regressions for different types of mosses, like liverworts, peat-mosses, and other mosses (possibly further divided into sub-groups), should be used. In future studies, more samples in a more systematic way from different types of moss growth forms should be collected to allow better model development for moss biomass.

There are differences in the explanatory potential of satellite image regressions across tundra or other northern landscapes with low-growth vegetation (Table 5). In a comparison between two sites at Nunavut, Canada, Atkinson and Treitz (2013) got higher R^2 values for their southern site which had lower average biomass and lower biomass variation across plots. Also in our study, the sites with low average biomass and low variation (Herschel and Tiksi) had low RMSE and high R^2_{adj} values, whereas sites with high average biomass (Sodankylä and sites in NW Russia) had higher RMSE and lower R^2_{adj} values. These differences could be related to within-site characteristics and variation in vegetation. For instance, the relationship between NDVI (or other vegetation indices) and biomass has been strong in VHSR evaluations in the Canadian Arctic associated with clear NDVI gradients from non-vegetated surfaces with low NDVI values to vegetated areas with high NDVI values (Atkinson and Treitz 2013; Collingwood et al. 2014). Of our sites, Herschel and Tiksi had large areas with no vegetation and, on the other hand, high herbaceous biomass in other places. In turn, especially at the north boreal fen Sodankylä, the biomass and NDVI gradients were short, and the landscape was dominated by an almost continuous moss cover. Moreover, it has been shown that variation of biomass in wetter sites, such as Sodankylä in our case, is not always evident in reflectance patterns as soil moisture suppresses NIR reflectance (Buchhorn et al. 2013).

It also is noteworthy that the PFT composition differs between the study sites, and this may affect the relationships between spectral reflectance and biomass. Vegetation indices such as NDVI and RATIO are connected to greenness as well as cellular and volume scattering by vegetation (Birth and McVey 1968; Rouse et al. 1973). It may be that in moss and shrub vegetation, which dominate in Sodankylä and among the sites in NW Russia, biomass and reflectance variables used in our analyses are not as tightly connected as in herbaceous vegetation, which has higher relative biomass in Herschel and Tiksi. In shrubs, the woody part has a large contribution to biomass, but they do not have as high reflectance values as leaves and other green parts found in herbaceous plants. In addition, the canopy structure is different in herbs and shrubs, which also affects reflectance. Furthermore, in the previous research, there have been problems in estimating moss biomass with the help of spectral reflectance (Bratsch et al. 2017). Especially *Sphagnum* have narrow absorption peaks in red and NIR, which hampers the value of vegetation indices in biomass estimation (Bubier, Rock, and Crill 1997). The variation in PFTs confuse the universal relationships but are worth examining in future research.

Examination of previous studies in sub-Arctic or Arctic environments suggests that coarse patterns in vegetation and biomass distribution are easier to detect than fine-scale variations. Usually moderate to high R^2_{ajd} values (>0.4) have been obtained in studies from plot to circumpolar scale (Hope, Kimball, and Stow 1993; Walker et al. 2003; Riedel, Epstein, and Walker 2005; Heiskanen 2006; Fuchs et al. 2009; Kushida et al. 2009; Raynolds et al. 2012; Atkinson and Treitz 2013; Buchhorn et al. 2013; Doiron et al. 2013; Collingwood et al. 2014; Johansen and Tommervik 2014; Kushida et al. 2015). The highest R^2_{ajd} values (>0.7) between NDVI and biomass have been obtained in studies that use moderate to coarse resolution satellite datasets (Walker et al. 2003; Heiskanen 2006; Raynolds et al. 2012; Buchhorn et al. 2013; Johansen and Tommervik 2014; Berner et al. 2018). In studies that use plot-scale NDVI measurements or VHSR imagery, the R^2_{ajd} values have often been near 0.5 or even below it (Hope, Kimball, and

Stow 1993; Riedel, Epstein, and Walker 2005; Fuchs et al. 2009; Kushida et al. 2009; Atkinson and Treitz 2013; Kushida et al. 2015) but there are also some exceptions (Atkinson and Treitz 2013; Buchhorn et al. 2013; Collingwood et al. 2014). One reason behind this disparity might be that the variation in spectral reflectance patterns is more evident at coarser scales, which usually also include areas with no or little biomass and low NDVI values. Nevertheless, more research is needed to analyze how the biomass distribution varies from fine to coarse scale across the different land cover and vegetation types.

We showed that cross-site regressions functioned relatively well in biomass-cover/height regression, with the underestimation and overestimation being relatively small (Figures 3 and 4). Nevertheless, there were large potential biases and high RMSE values in cross-site biomass-satellite spectra regression predictions (Figure 5, Table 6). This was evident on Herschel, where there were 2–3-fold differences in the landscape-scale average biomass when different regression combinations were used (Table 6). This finding is in line with the study by Atkinson and Treitz (2013), who, however, had only two sites at Nunavut, Canada for their comparison. The differences between sites suggest that satellite image based shrub tundra models work well in different shrub tundra landscapes such as NW Russian sites, but their value is limited in herbaceous environments such as Herschel, and it is tedious to find suitable cross-site models. Nevertheless, on Herschel, the combination of cross-site biomass-cover/height regression and site-specific biomass-satellite spectra regression yielded lower RMSE values than the combination of two site-specific regressions. This might be due to the fact that cross-site biomass-cover/height regressions were more realistic as they had a bigger sample size. Another possibility is, that although cross-site biomass-cover/height regressions slightly underestimated biomass values, modelled values were such that they could be modelled with satellite spectra. Nevertheless, it might be that cross-site models are more robust to outliers due to larger sample size, and they can give better fit, if there is no large differences in the environmental characteristics of the study sites.

The biomass-satellite spectra regressions that included multiple explanatory variables had better prediction capability than regressions with only one index as explanatory variable. Furthermore, in previous research, good explanatory power has been obtained using models that combine optical imagery and other types of remote sensing data both in the Arctic areas (Chen, Blain, et al. 2009; Collingwood et al. 2014; Greaves et al. 2016) and other landscapes with low-growth vegetation (Glenn et al. 2016). In particular, features related to vegetation height, topography, moisture gradients, soil properties and geomorphology could improve biomass models (Chen, Blain, et al. 2009; Axmanova et al. 2012; Collingwood et al. 2014; Suvanto, Le Roux, and Luoto 2014; Glenn et al. 2016; Greaves et al. 2016). In a similar manner, in peatland landscapes such as Sodankylä, carbon exchange and other ecosystem properties are linked to microtopographical variation (Lees et al. 2018), which could be captured with VHSR digital elevation models. It may be that the benefit of other datatypes is greater in areas where the relationship between reflectance and biomass is weak. Therefore, future research should combine VHSR imagery with other VHSR data and test what kind of models and predictor variables should be used in each kind of landscape and if some predictor sets are locally optimal but not as useful in a larger area.

Finally, our analysis did not include phenological dynamics nor did we analyze the optimal timing for satellite imagery in mapping biomass distribution. Biomass and other vegetation parameters change during the growing season; furthermore, the relative importance of different PFTs change as the growing season proceeds (Anderson et al.

2016; Wang et al. 2016; Juutinen et al. 2017). It has thus been shown that the seasonal phase of VHSR images affect the interpretation of vegetation parameters such as leaf-area index (Juutinen et al. 2017). However, more work is needed on evaluating how the timing of satellite images affects success in mapping biomass distribution.

5. Conclusions

We estimated aboveground biomass in four different Arctic landscapes using field sampling based biomass-cover/height regressions and biomass-satellite spectra regressions. We tested both site-specific regressions and cross-site regressions across all the study sites, and showed that biomass-cover/height regressions perform well in most cases (R^2_{adj} 0.21–0.92), and their performance varies in biomass-satellite spectra regressions (R^2_{adj} 0.33–0.75). The cross-site regressions should be used with care in biomass-satellite spectra regressions, as they underestimated biomass in some study sites and overestimated them in other sites. However, in biomass-cover/height regressions there was no large differences in predicted biomass values when site-specific regressions were compared with cross-site regressions. Moreover, due to larger sample size, cross-site regressions are more robust to outliers, and may yield better fit than site-specific regressions when they combine data from study sites which have similar vegetation and landscape characteristics. We showed that there is considerable variation in biomass distribution both within and between different Arctic landscapes, and the biomass and proportion of different PFTs vary between Arctic landscapes. Nevertheless, there is further need for model building and validation across different tundra environments, including landscape types which were not included in our study or still have limited field datasets. To summarize, the usage of VHSR images creates new possibilities to map the fine-scale spatial variability in biomass in landscapes with patchy vegetation cover for different kind of ecosystem and modelling purposes, but some caution is needed when trying to develop models performing well in different environments.

Acknowledgments: We thank Malin Ek, Hanna Hyvönen, Maria Kröger, Maiju Linkosalmi, Johanna Nyman, Tiina Ronkainen, Lauri Rosenius, Sanna Susiluoto and Emmi Vähä for field and laboratory assistance. We would also like to thank the following individuals from various institutes and organizations for help with different aspects and phases of field campaigns and data collection: Komi Biological Institute in Syktyvkar, Tiksi Observatory and Yakutian Service for Hydrometeorology, The Aurora Research Institute in Inuvik, Arctic Research Centre of Finnish Meteorological Institute in Sodankylä, University of Stockholm, Alfred Wegener Institute in Potsdam, University of Eastern Finland, and Natural Resources Institute Finland. Data used in this study was collected in EU 6th Framework CARBONorth project [contract 036993], and in the following projects funded by the Academy of Finland: COUP [project 291736] and Greenhouse gas, aerosol and albedo variations in the changing Arctic [project 269095]. The paper was finalized with financial support from CAPTURE [project 296423 funded by the Academy of Finland].

References

AARI. 2017. "Electronic archive AARI term meteorological and upper-air observations Hydrometeorological Observatory AARI 2016 (station) Tiksi for 19322014." Accessed September 7. www.aari.ru/main.php?sub=2&id=3.

———. 2017. "Electronic archive AARI term meteorological and upper-air observations Hydrometeorological Observatory AARI 2016 (station) Tiksi for 1932 2014." Accessed September 7. www.aari.ru/main.php?sub=2&id=3.

Alexeyev, V., R. Birdsey, V. Stakanov, and I. Korotkov. 1995. "Carbon in vegetation of Russian forests: Methods to estimate storage and geographical distribution." *Water, Air, and Soil Pollution* 82 (1-2):271-82. doi: 10.1007/BF01182840.

Anderson, H. B., L. Nilsen, H. Tommervik, S. R. Karlsen, S. Nagai, and E. J. Cooper. 2016. "Using ordinary digital cameras in place of near-infrared sensors to derive vegetation indices for phenology studies of high arctic vegetation." *Remote Sensing* 8 (10). doi: 10.3390/rs8100847.

Atkinson, D. M., and P. Treitz. 2013. "Modeling biophysical variables across an Arctic latitudinal gradient using high spatial resolution remote sensing data." *Arctic Antarctic and Alpine Research* 45 (2):161-78. doi: 10.1657/1938-4246-45.2.161.

Axmanova, I., L. Tichy, Z. Fajmonova, P. Hajkova, E. Hettenbergerova, C. F. Li, K. Merunkova, et al. 2012. "Estimation of herbaceous biomass from species composition and cover." *Applied Vegetation Science* 15 (4):580-9. doi: 10.1111/j.1654-109X.2012.01191.x.

Beamish, A. L., N. C. Coops, S. Chabrillat, and B. Heim. 2017. "A Phenological Approach to Spectral Differentiation of Low-Arctic Tundra Vegetation Communities, North Slope, Alaska." *Remote Sensing* 9 (11):1200. doi: doi:10.3390/rs9111200.

Berner, L. T., P. Jantz, K. D. Tape, and S. J. Goetz. 2018. "Tundra plant aboveground biomass and shrub dominance mapped across the North Slope of Alaska." *Environmental Research Letters* In press. doi: https://doi.org/10.1088/1748-9326/aaaa9a.

Birth, G. S., and G. R. McVey. 1968. "Measuring the color of growing turf with a reflectance spectrophotometer." *Agronomy Journal* 60 (6):640-3. doi: doi:10.2134/agronj1968.00021962006000060016x.

Bratsch, S., H. Epstein, M. Buchhorn, D. Walker, and H. Landes. 2017. "Relationships between hyperspectral data and components of vegetation biomass in Low Arctic tundra communities at Ivotuk, Alaska." *Environmental Research Letters* 12 (2). doi: 10.1088/1748-9326/aa572e.

Bubier, J. L., B. N. Rock, and P. M. Crill. 1997. "Spectral reflectance measurements of boreal wetland and forest mosses." *Journal of Geophysical Research Atmospheres* 102 (24):29483-94.

Buchhorn, M., M. K. Raynolds, and D. A. Walker. 2016. "Influence of BRDF on NDVI and biomass estimations of Alaska Arctic tundra." *Environmental Research Letters* 11 (12):125002. doi: 10.1088/1748-9326/11/12/125002.

Buchhorn, M., D. A. Walker, B. Heim, M. K. Raynolds, H. E. Epstein, and M. Schwieder. 2013. "Ground-based hyperspectral characterization of Alaska tundra vegetation along environmental gradients." *Remote Sensing* 5 (8):3971-4005. doi: 10.3390/rs5083971.

Burn, C. R. 2012. "Climate." In *Herschel Island—Qikiqtaryuk: A Natural and Cultural History.*, edited by C. R. Burn, 48-53. Calgary, Canada: University of Calgary Press.

Chapin III, F. S., M. S. Bret-Harte, S. E. Hobbie, and H. Zhong. 1996. "Plant functional types as predictors of transient responses of arctic vegetation to global change." *Journal of Vegetation Science* 7 (3):347-58.

Chavez, P. S. 1988. "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data." *Remote Sensing of Environment* 24 (3):459-79. doi: 10.1016/0034-4257(88)90019-3.

Chen, W., D. Blain, J. Li, K. Keohler, R. Fraser, Y. Zhang, S. Leblanc, I. Olthof, J. Wang, and M. McGovern. 2009. "Biomass measurements and relationships with Landsat-7/ETM + and JERS-1/SAR data over Canada's western sub-arctic and low arctic." *International journal of remote sensing* 30 (9):2355-76. doi: 10.1080/01431160802549401.

Chen, W. J., J. H. Li, Y. Zhang, F. Q. Zhou, K. Koehler, S. Leblanc, R. Fraser, I. Olthof, Y. S. Zhang, and J. X. Wang. 2009. "Relating biomass and leaf area index to non-destructive measurements in order to monitor changes in arctic vegetation." *Arctic* 62 (3):281-94.

Collingwood, A., P. Treitz, F. Charbonneau, and D. M. Atkinson. 2014. "Artificial neural network modeling of high arctic phytomass using synthetic aperture radar and multispectral data." *Remote Sensing* 6 (3):2134-53. doi: 10.3390/rs6032134.

Coops, N. C., M. Johnson, M. A. Wulder, and J. C. White. 2006. "Assessment of QuickBird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation." *Remote Sensing of Environment* 103 (1):67-80. doi: 10.1016/j.rse.2006.03.012.

Dinsmore, K. J., J. Drewer, P. E. Levy, C. George, A. Lohila, M. Aurela, and U. M. Skiba. 2017. "Growing season CH_4 and N_2O fluxes from a subarctic landscape in northern Finland; From chamber to landscape scale." *Biogeosciences* 14 (4):799-815. doi: 10.5194/bg-14-799-2017.

Doiron, M., P. Legagneux, G. Gauthier, and E. Levesque. 2013. "Broad-scale satellite Normalized Difference Vegetation Index data predict plant biomass and peak date of nitrogen concentration in Arctic tundra vegetation." *Applied Vegetation Science* 16 (2):343-51. doi: 10.1111/j.1654-109X.2012.01219.x.

Eckert, S. 2012. "Improved forest biomass and carbon estimations using texture measures from worldView-2 satellite data." *Remote Sensing* 4 (4):810-29. doi: 10.3390/rs4040810.

Epstein, H. E., M. K. Raynolds, D. A. Walker, U. S. Bhatt, C. J. Tucker, and J. E. Pinzon. 2012. "Dynamics of aboveground phytomass of the circumpolar Arctic tundra during the past three decades." *Environmental Research Letters* 7 (1). doi: 10.1088/1748-9326/7/1/015506.

Finnish Meteorological Institute. 2017. "July statistics." Accessed September 9. http://ilmatieteenlaitos.fi/heinakuu.

Fox, J., and S. Weisberg. 2011. *An {R} Companion to Applied Regression, Second Edition.* Thousand Oaks, CA, USA: Sage.

Fuchs, H., P. Magdon, C. Kleinn, and H. Flessa. 2009. "Estimating aboveground carbon in a catchment of the Siberian forest tundra: Combining satellite imagery and field inventory." *Remote Sensing of Environment* 113 (3):518-31. doi: 10.1016/j.rse.2008.07.017.

Glenn, N. F., A. Neuenschwander, L. A. Vierling, L. Spaete, A. Li, D. J. Shinneman, D. S. Pilliod, R. S. Arkle, and S. K. McIlroy. 2016. "Landsat 8 and ICESat-2: Performance and potential synergies for quantifying dryland ecosystem vegetation cover and biomass." *Remote Sensing of Environment* 185:233-42. doi: 10.1016/j.rse.2016.02.039.

Greaves, H. E., L. A. Vierling, J. U. H. Eitel, N. T. Boelman, T. S. Magney, C. M. Prager, and K. L. Griffin. 2016. "High-resolution mapping of aboveground shrub biomass in Arctic tundra using airborne lidar and imagery." *Remote Sensing of Environment* 184:361-73. doi: 10.1016/j.rse.2016.07.026.

Grosswald, M. G., W. Karlen, Z. Shishorina, and A. Bodin. 1992. "Glacial landforms and the age of deglaciation in the Tiksi area, east Siberia." *Geografiska Annaler, Series A* 74 A (4):295-304.

Heiskanen, J. 2006. "Estimating aboveground tree biomass and leaf area index in a mountain birch forest using ASTER satellite data." *International journal of remote sensing* 27 (5-6):1135-58. doi: 10.1080/01431160500353858.

Hope, A. S., J. S. Kimball, and D. A. Stow. 1993. "The relationship between tussock tundra spectral reflectance properties and biomass and vegetation composition." *International journal of remote sensing* 14 (10):1861-74.

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

Hugelius, G., T. Virtanen, D. Kaverin, A. Pastukhov, F. Rivkin, S. Marchenko, V. Romanovsky, and P. Kuhry. 2011. "High-resolution mapping of ecosystem carbon storage and potential effects of permafrost thaw in periglacial terrain, European Russian Arctic." *Journal of Geophysical Research: Biogeosciences* 116 (3). doi: 10.1029/2010JG001606.

Jiang, Z. Y., A. R. Huete, K. Didan, and T. Miura. 2008. "Development of a two-band enhanced vegetation index without a blue band." *Remote Sensing of Environment* 112 (10):3833-45. doi: 10.1016/j.rse.2008.06.006.

Johansen, B., and H. Tommervik. 2014. "The relationship between phytomass, NDVI and vegetation communities Svalbard." *International Journal of Applied Earth Observation and Geoinformation* 27:20-30. doi: 10.1016/j.jag.2013.07.001.

Juutinen, S., T. Virtanen, V. Kondratyev, T. Laurila, M. Linkosalmi, J. Mikola, J. Nyman, A. Räsänen, J-P. Tuovinen, and M. Aurela. 2017. "Spatial variation and seasonal dynamics of leaf-area index in the arctic tundra-implications for linking ground observations and satellite images." *Environmental Research Letters* 12:095002. doi: https://doi.org/10.1088/1748-9326/aa7f85.

Kauppi, P. E., E. Tomppo, and A. Ferm. 1995. "C and N storage in living trees within Finland since 1950s." *Plant and Soil* 168-169 (1):633-8. doi: 10.1007/BF00029377.

Korpela, I. 2001. "Metsänarviointi. Luentomoniste teknillisen korkeakoulun metsänarvioinnin kurssille." In. Helsinki, Finland: University of Helsinki.

Kuhn, M., J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper, et al. "caret: classification and regression training. R package version 6.0-71.". http://CRAN.R-project.org/package=caret.

Kushida, K., S. Hobara, S. Tsuyuzaki, Y. Kim, M. Watanabe, Y. Setiawan, K. Harada, G. R. Shaver, and M. Fukuda. 2015. "Spectral indices for remote sensing of phytomass, deciduous shrubs, and productivity in Alaskan Arctic tundra." *International journal of remote sensing* 36 (17):4344-62. doi: 10.1080/01431161.2015.1080878.

Kushida, K., Y. Kim, S. Tsuyuzaki, and M. Fukuda. 2009. "Spectral vegetation indices for estimating shrub cover, green phytomass and leaf turnover in a sedge-shrub tundra." *International journal of remote sensing* 30 (6):1651-8. doi: 10.1080/01431160802502632.

Laidler, G. J., and P. Treitz. 2003. "Biophysical remote sensing of arctic environments." *Progress in Physical Geography* 27 (1):44-68. doi: 10.1191/0309133303pp358ra.

Lees, K. J., T. Quaife, R. R. E. Artz, M. Khomik, and J. M. Clark. 2018. "Potential for using remote sensing to estimate carbon fluxes across northern peatlands – A review." *Science of the Total Environment* 615:857-74. doi: 10.1016/j.scitotenv.2017.09.103.

Liu, H. Q., and A. Huete. 1995. "A feedback based modification of the NDVI to minimize canopy background and atmospheric noise." *Ieee Transactions on Geoscience and Remote Sensing* 33 (2):457-65.

Marushchak, M. E., I. Kiepe, C. Biasi, V. Elsakov, T. Friborg, T. Johansson, H. Soegaard, T. Virtanen, and P. J. Martikainen. 2013. "Carbon dioxide balance of subarctic tundra from plot to regional scales." *Biogeosciences* 10 (1):437-52. doi: 10.5194/bg-10-437-2013.

McDonald, J.H. 2014. *Handbook of Biological Statistics*. 3rd ed. ed. Baltimore, Maryland, USA: Sparky House Publishing.

Mikola, J., T. Virtanen, M. Linkosalmi, E. Vähä, J. Nyman, O. Postanogova, A. Räsänen, et al. 2018. "Spatial variation and linkages of soil and vegetation in the Siberian Arctic tundra - Coupling field observations with remote sensing data." *Biogeosciences* 15 (9):2781-801. doi: 10.5194/bg-15-2781-2018.

Myers-Smith, I. H., S. C. Elmendorf, P. S. A. Beck, M. Wilmking, M. Hallinger, D. Blok, K. D. Tape, et al. 2015. "Climate sensitivity of shrub growth across the tundra biome." *Nature Climate Change* 5 (9):887-+. doi: 10.1038/nclimate2697.

Myers-Smith, I. H., D. S. Hik, C. Kennedy, D. Cooley, J. F. Johnstone, A. J. Kenney, and C. J. Krebs. 2011. "Expansion of canopy-forming willows over the

twentieth century on Herschel Island, Yukon Territory, Canada." *Ambio* 40 (6):610-23. doi: 10.1007/s13280-011-0168-y.
Nyyssönen, A. 1955. "Metsikön kuutiomäärän arvioiminen relaskoopin avulla." *Metsäntutkimuslaitoksen julkaisuja* 44 (6):1-31.
Obu, J., H. Lantuit, I. Myers-Smith, B. Heim, J. Wolter, and M. Fritz. 2017. "Effect of terrain characteristics on soil organic carbon and total nitrogen stocks in soils of Herschel Island, western Canadian Arctic." *Permafrost and Periglacial Processes* 28 (1):92-107. doi: 10.1002/ppp.1881.
Qi, J., A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian. 1994. "A modified soil adjusted vegetation index." *Remote Sensing of Environment* 48 (2):119-26. doi: 10.1016/0034-4257(94)90134-1.
R Core Team. "R: A language and environment for statistical computing." R Foundation for Statistical Computing. https://www.R-project.org/.
Raynolds, M. K., D. A. Walker, H. E. Epstein, J. E. Pinzon, and C. J. Tucker. 2012. "A new estimate of tundra-biome phytomass from trans-Arctic field data and AVHRR NDVI." *Remote Sensing Letters* 3 (5):403-11. doi: 10.1080/01431161.2011.609188.
Raynolds, M. K., D. A. Walker, and H. A. Maier. 2006. "NDVI patterns and phytornass distribution in the circumpolar Arctic." *Remote Sensing of Environment* 102 (3-4):271-81. doi: 10.1016/j.rse.2006.02.016.
Riedel, S. M., H. E. Epstein, and D. A. Walker. 2005. "Biotic controls over spectral reflectance of arctic tundra vegetation." *International journal of remote sensing* 26 (11):2391-405. doi: 10.1080/01431160512331337754.
Rouse, J. W. Jr., R. H. Haas, J. A. Schell, and D. W. Deering. 1973. "Monitoring vegetation systems in the Great Plains with ERTS." In *Third Earth Resources Technology Satellite-1 Symposium*, 309-17. Washington, DC: NASA.
Shepashenko, D., A. Shvidenko, and S. Nilsson. 1998. "Phytomass (live biomass) and carbon of Siberian forests." *Biomass & Bioenergy* 14 (1):21-31. doi: 10.1016/s0961-9534(97)10006-x.
Siewert, M. B., J. Hanisch, N. Weiss, P. Kuhry, T. C. Maximov, and G. Hugelius. 2015. "Comparing carbon storage of Siberian tundra and taiga permafrost ecosystems at very high spatial resolution." *Journal of Geophysical Research-Biogeosciences* 120 (10):1973-94. doi: 10.1002/2015jg002999.
Sinha, S., C. Jeganathan, L. K. Sharma, and M. S. Nathawat. 2015. "A review of radar remote sensing for biomass estimation." *International Journal of Environmental Science and Technology* 12 (5):1779-92. doi: 10.1007/s13762-015-0750-0.
Smith, C. A. S., C. E. Kennedy, A. E. Hargrave, and McKenna K. M. 1989. *Soil and vegetation survey of Herschel Island, Yukon Territory., Yukon Soil Survey Report, No. 1. LRRC Contribution No. 88-26*. Whitehorse Yukon: Agriculture Canada.
Song, C., C. E. Woodcock, K. C. Seto, M. P. Lenney, and S. A. Macomber. 2001. "Classification and change detection using Landsat TM data: When and how to correct atmospheric effects?" *Remote Sensing of Environment* 75 (2):230-44. doi: 10.1016/s0034-4257(00)00169-3.
Starr, M., M. Hartman, and T. Kinnunen. 1998. "Biomass functions for mountain birch in the Vuoskojärvi Integrated Monitoring area." *Boreal Environment Research* 3 (3):297-303.
Steele-Dunne, S. C., H. McNairn, A. Monsivais-Huertero, J. Judge, P. W. Liu, and K. Papathanassiou. 2017. "Radar remote sensing of agricultural canopies: A

Review." *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10 (5):2249-73. doi: 10.1109/JSTARS.2016.2639043.

Stow, D. A., A. Hope, D. McGuire, D. Verbyla, J. Gamon, F. Huemmrich, S. Houston, et al. 2004. "Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems." *Remote Sensing of Environment* 89 (3):281-308. doi: 10.1016/j.rse.2003.10.018.

Suvanto, S., P. C. Le Roux, and M. Luoto. 2014. "Arctic-alpine vegetation biomass is driven by fine-scale abiotic heterogeneity." *Geografiska Annaler Series a-Physical Geography* 96 (4):549-60. doi: 10.1111/geoa.12050.

Walker, D. A., F. J. A. Daniëls, I. Alsos, U. S. Bhatt, A. L. Breen, M. Buchhorn, H. Bültmann, et al. 2016. "Circumpolar Arctic vegetation: A hierarchic review and roadmap toward an internationally consistent approach to survey, archive and classify tundra plot data." *Environmental Research Letters* 11 (5). doi: 10.1088/1748-9326/11/5/055005.

Walker, D. A., H. E. Epstein, G. J. Jia, A. Balser, C. Copass, E. J. Edwards, W. A. Gould, et al. 2003. "Phytomass, LAI, and NDVI in northern Alaska: Relationships to summer warmth, soil pH, plant functional types, and extrapolation to the circumpolar Arctic." *Journal of Geophysical Research-Atmospheres* 108 (D2). doi: 10.1029/2001jd000986.

van der Wal, R., and A. Stien. 2014. "High-arctic plants like it hot: a long-term investigation of between-year variability in plant biomass." *Ecology* 95 (12):3414-27.

Wang, P., L. Mommer, J. van Ruijven, F. Berendse, T. C. Maximov, and Mmpd Heijmans. 2016. "Seasonal changes and vertical distribution of root standing biomass of graminoids and shrubs at a Siberian tundra site." *Plant and Soil* 407 (1-2):55-65. doi: 10.1007/s11104-016-2858-5.

Varmola, M., and E. Vuokila. 1986. "Pienten mäntyjen tilavuusyhtälöt ja -taulukot. Tree volume functions and tables for small-sized pines." In *Folia Forestalia 652*, 24. Helsinki, Finland: The Finnish Forest Research Institute.

Venables, W. N., and B. D. Ripley. 2003. *Modern Applied Statistics with S. Fourth Edition*. New York, NY, USA: Springer.

Westergaard-Nielsen, A., A. B. Bjørnsson, M. R. Jepsen, M. Stendel, B. U. Hansen, and B. Elberling. 2015. "Greenlandic sheep farming controlled by vegetation response today and at the end of the 21st Century." *Science of the Total Environment* 512-513:672-81. doi: 10.1016/j.scitotenv.2015.01.039.

Westergaard-Nielsen, A., M. Lund, B. U. Hansen, and M. P. Tamstorf. 2013. "Camera derived vegetation greenness index as proxy for gross primary production in a low Arctic wetland area." *ISPRS Journal of Photogrammetry and Remote Sensing* 86:89-99. doi: 10.1016/j.isprsjprs.2013.09.006.

Virtanen, T., and M. Ek. 2014. "The fragmented nature of tundra landscape." *International Journal of Applied Earth Observation and Geoinformation* 27:4-12. doi: 10.1016/j.jag.2013.05.010.

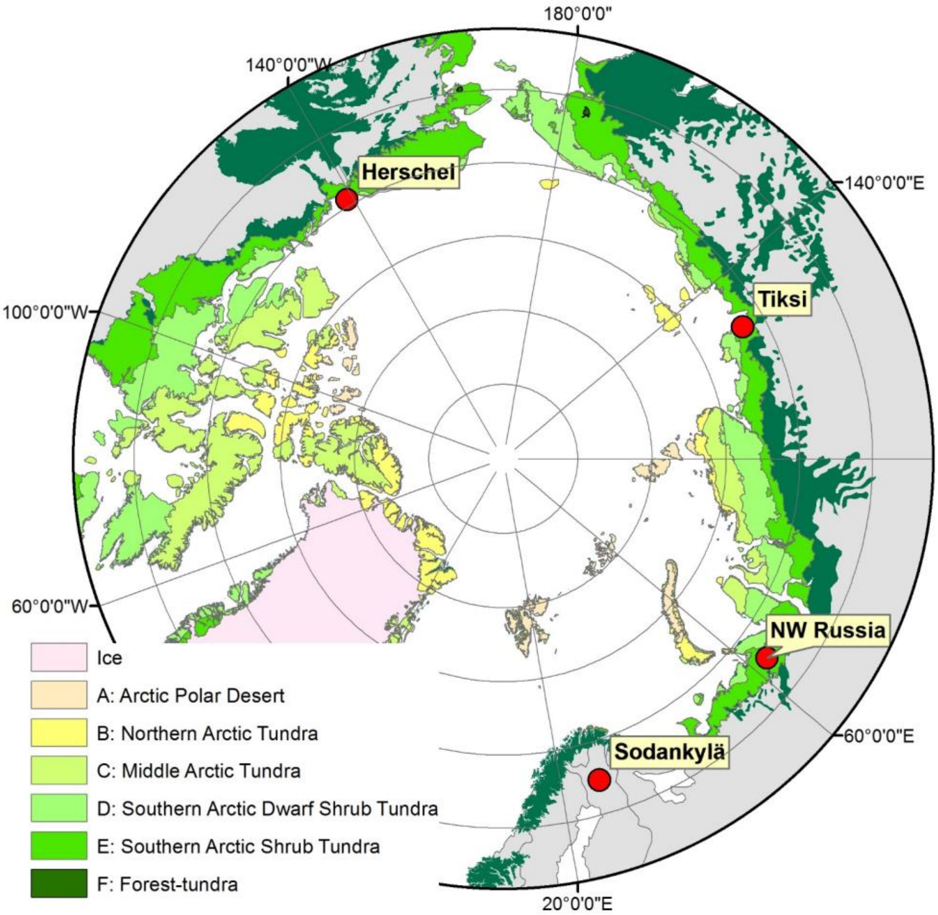

Figure 1. Location of the study sites. Vegetation zones A–E are based on Walker et al. (2005) and forest tundra (F) on Olson et al. (2001). Sodankylä belongs to the northern boreal vegetation zone.

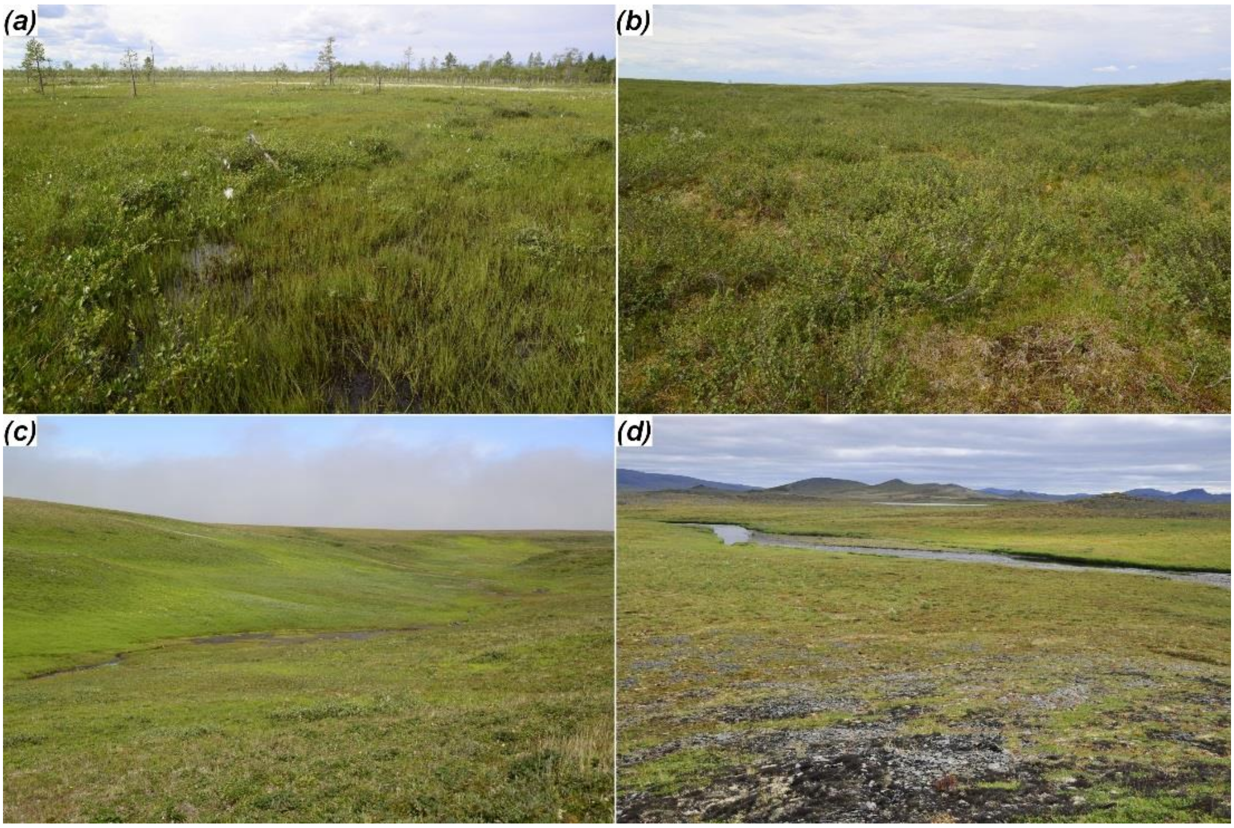

Figure 2. Study sites (a) Sodankylä, (b) Seida in Northwestern Russia, (c) Herschel Island, and (d) Tiksi.

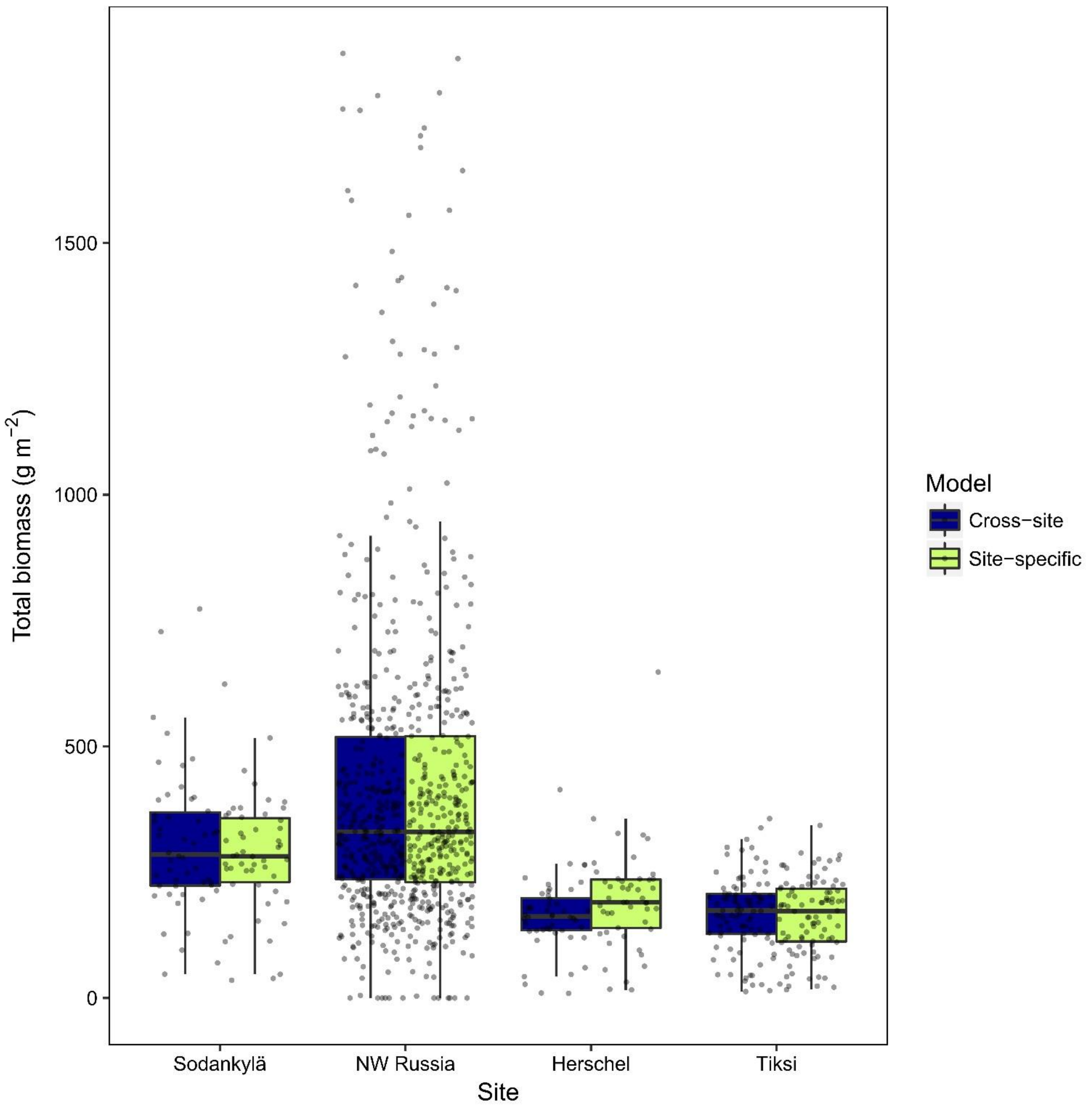

Figure 3. Distribution of total aboveground biomass estimated for 5 m radius plots at each study site using cross-site and site-specific regressions. Note: Sodankylä and NW Russia estimates do not include tree biomass. Lines in the middle of the boxes show the median value and the lower and upper hinges are the first and third quartiles. The lower and upper whiskers extend to the smallest and largest values which are no further than 1.5 times the distance between the first and third quartiles.

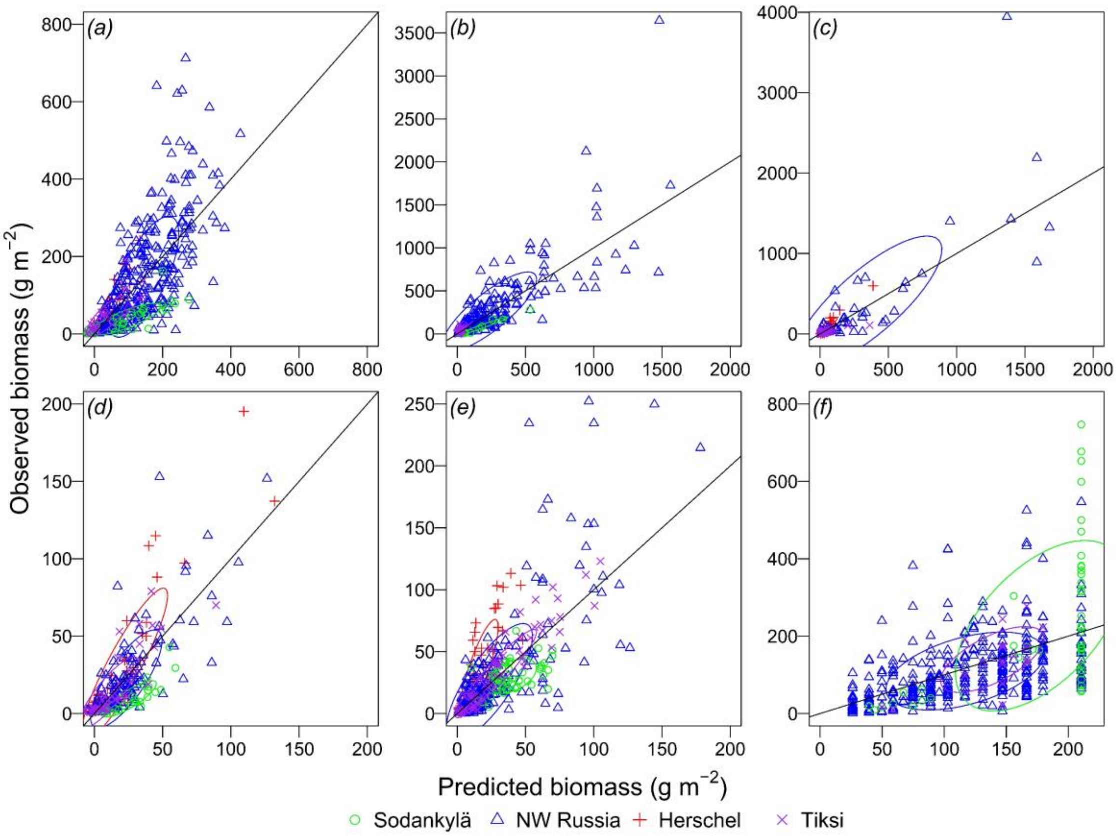

Figure 4. Observed (y-axis) vs. predicted i.e. harvested (x-axis) biomass for harvested plots (symbols) for (a) dwarf shrubs, (b) *Betula nana*, (c) *Salix* spp., (d) herbs, (e) graminoids and (f) mosses. Cross-site regressions for different plant functional groups were applied and the 1:1 line is shown. For each study site, individual plots are marked with symbols and the 50% confidence interval with an ellipsoid.

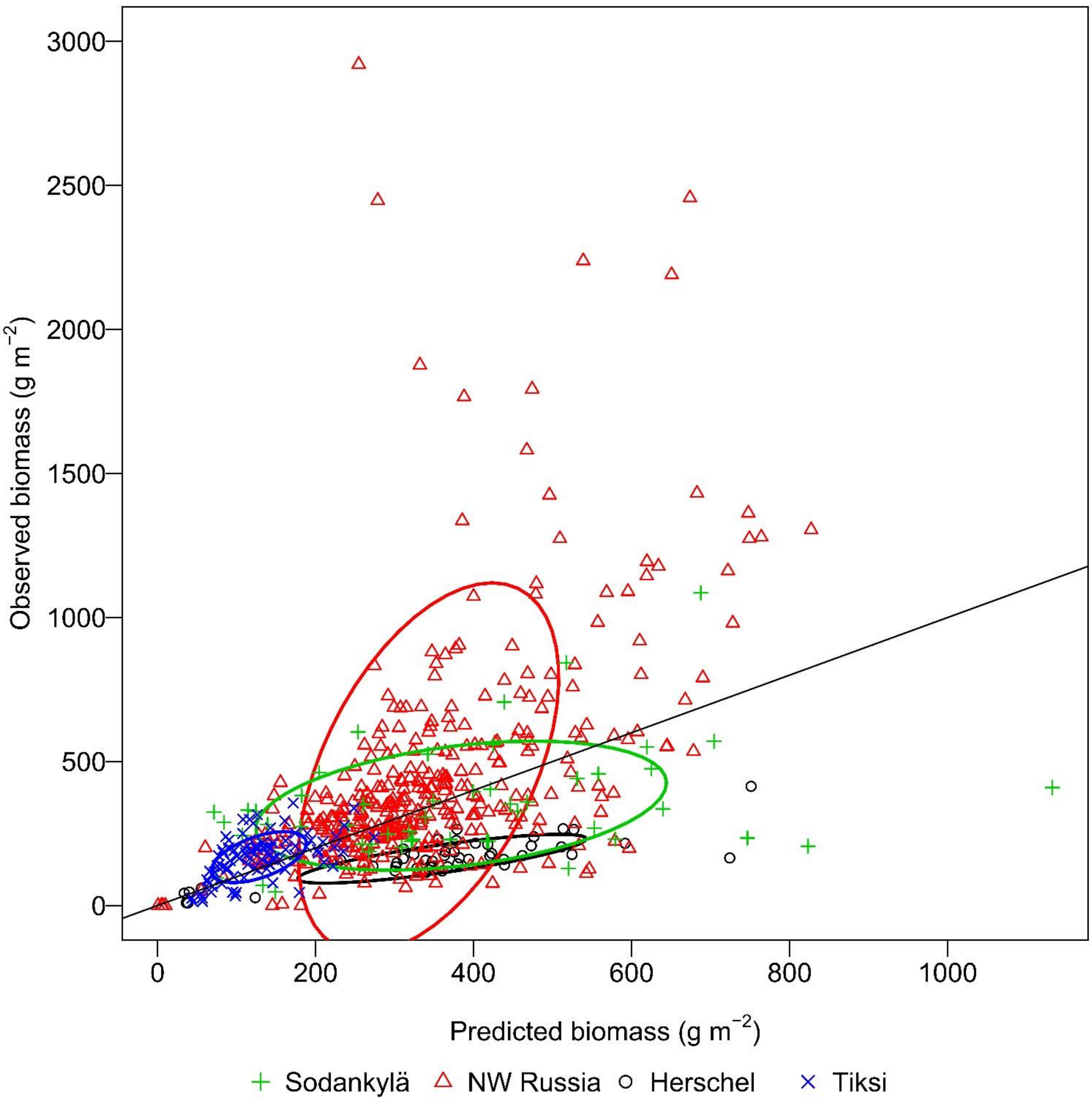

Figure 5. Predicted biomass (x-axis) against observed biomass in the best fitting cross-site biomass- satellite spectra regression. Observed biomass is the sum of the predicted values of the best fitting cross-site biomass-cover/height regressions for each PFT and tree biomass. For each study site, individual plots are marked with asterisks and a 50% confidence interval with an ellipsoid. Line represents a 1:1 line. To increase readability of the plot, three observations with > 3000 g m^{-2} observed biomass were removed from the plot.

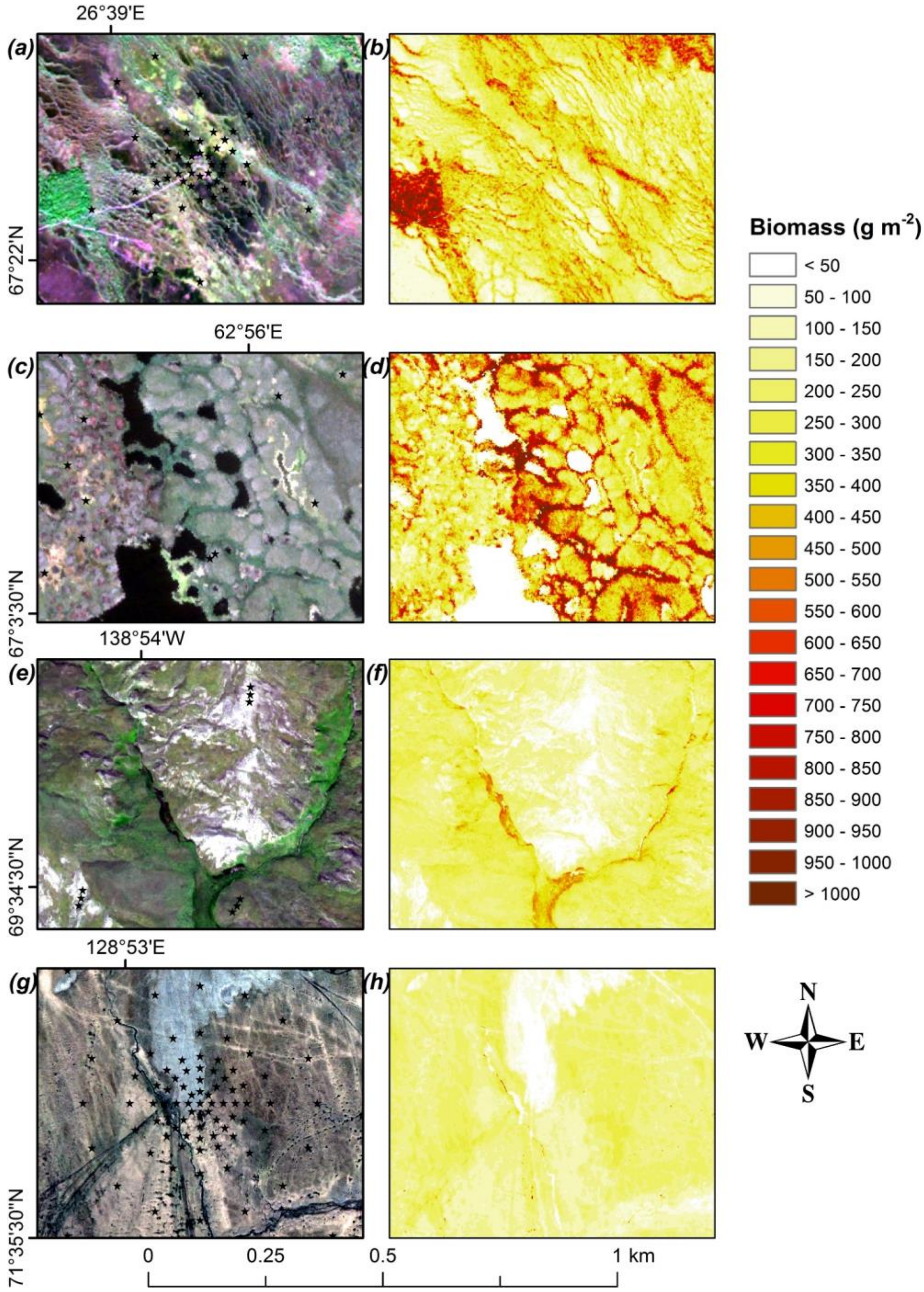

Figure 6. RGB satellite images and biomass maps for Sodankylä (a, b), Seida in NW Russia (c, d), Herschel (e, f), and Tiksi (g, h). Biomass maps were produced with the best fitting site-specific regressions (see Table 5). Spatial resolution of the images is shown in Table 1 and biomass maps have same pixel size as the images. In the satellite images, the location of the field sampling plots are shown with star symbols. For Seida/NW Russia and Herschel, only part of the field sampling plots are shown, because the plots were collected from a larger area. Satellite images ©Digital Globe.

Table 1. Coordinates of the study sites, mean July temperature sensor and imagery information, and date of the field data collection. In the column “Sensor”, WV refers to WorldView and QB to QuickBird.

Study site	Location	Mean July temperature (°C)	Sensor	Pixel size (m)	Imagery date	Fieldwork date	Number of plots
Sodankylä	67° 22' N 26° 39' E	14.5[1]	WV-2	2	4 July 2015	15–20 July 2014	50
Khosedayu[2]	67° 3' N 59° 25' E	13[3]	QB	2.4	30 June 2008	19–25 July 2007	60
Seida[2]	67° 4' N 62° 56' E	13[3]	QB	2.4	6 July 2007	5 July–6 Aug 2007 24–27 July 2016	150[4] 32
Rogovaya 1[2]	67° 22' N 62° 15' E	13[3]	QB	2.4	4 July 2007	8–11 July 2007	62
Rogovaya 2[2]	67° 17' N 62° 6' E	13[3]	QB	2.4	4 July 2007	13–17 July 2007	62
Herschel	69° 35' N 138° 55' W	9[5]	WV-3	1.6	8 Aug 2015	23 July–3 Aug 2015	48
Tiksi	71° 35' N 128° 53' E	7[6]	QB	0.6[7]	15 July 2005	23–27 July 2014	91

[1](Finnish Meteorological Institute 2017)
[2]Analyzed together with other northwestern (NW) Russia study sites (Khosedayu, Seida, Rogovaya 1 and 2).
[3](Marushchak et al. 2013)
[4]In addition, we used data from 34 extra subplots in biomass-cover/height regressions.
[5](Burn 2012)
[6](AARI 2017b)
[7]Image was delivered as a pan-sharpened product, i.e. all multispectral bands had 0.6 m resolution.

Table 2. Calculated vegetation indices, their references, and equations. Variables in the table refer to the respective spectral bands of the images and NIR to near infrared. Normalized difference vegetation index 2 was calculated only for WorldView images as it includes the near infrared 2 (NIR2) band.

Index	Equation
Normalized difference vegetation index (NDVI) (Rouse et al. 1973)	$\mathrm{NDVI} = \frac{((\mathrm{NIR}) - (\mathrm{red}))}{((\mathrm{NIR}) + (\mathrm{red}))}$
Normalized difference vegetation index 2 (NDVI2) (Eckert 2012)	$\mathrm{NDVI2} = \frac{((\mathrm{NIR2}) - (\mathrm{red}))}{((\mathrm{NIR2}) + (\mathrm{red}))}$
Red-green index (RGI) (Coops et al. 2006)	$\mathrm{RGI} = \frac{((\mathrm{green}) - (\mathrm{red}))}{((\mathrm{green} + (\mathrm{red}))}$
Simple ratio (RATIO) (Birth and McVey 1968)	$\mathrm{RATIO} = \frac{(\mathrm{red})}{(\mathrm{NIR})}$
Enhanced vegetation index (EVI) (Liu and Huete 1995)	$\mathrm{EVI} = 2.5 \times \frac{((\mathrm{NIR}) - (\mathrm{red}))}{((\mathrm{NIR}) + 6 \times (\mathrm{red}) - 7.5 \times (\mathrm{blue}) + 1)}$
Enhanced vegetation index 2 (EVI2) (Jiang et al. 2008)	$\mathrm{EVI2} = 2.5 \times \frac{((\mathrm{NIR}) - (\mathrm{red}))}{((\mathrm{NIR}) + 2.4 \times (\mathrm{red}) + 1)}$
Soil-adjusted vegetation index (SAVI) (Huete 1988)	$\mathrm{SAVI} = \frac{((\mathrm{NIR}) - (\mathrm{red}))}{((\mathrm{NIR}) + (\mathrm{red}) + 1)} \times 1.5$
Modified SAVI (MSAVI2) (Qi et al. 1994)	$\mathrm{MSAVI2} = \frac{(2 \times (\mathrm{NIR}) + 1 - \sqrt{(2 \times (\mathrm{NIR}) + 1)^2 - 8 \times ((\mathrm{NIR}) - (\mathrm{red}))}}{2}$

Table 3. Average ± standard deviation of biomass values for each plant functional type and study site based on field samples. The plant functional type with highest biomass at each site is shown in bold. Note: moss biomass for Herschel and Tiksi and tree biomass are not included in this table.

Study site	Biomass (g m^{-2})						
	Dwarf shrubs	*B. nana*	*Salix* spp.	Herbs	Graminoids	Mosses	Total
Sodankylä	27.9±33	16.5±51.9	0.5±8.4	5.4±8.7	30.3±11.8	**217.9±187.6**	298.4±210
NW Russia	114.7±132.0	**157.3±329.3**	46.4±270.5	7.9±19.1	18.7±36.8	102.9±87.0	439.3±431.6
Herschel	37.2±54.2	12.2±61.2	**65.1±98.1**	28.8±40.0	35.5±32.7	-	178.8±116.1
Tiksi	15.1±32.6	10.4±22.3	16.9±24.2	8.8±16.7	**28.2±30**	-	79.4±46.9

Table 4. Equations for different biomass-cover/height regressions for predicting plant functional type biomass (bm). Adjusted coefficient of determination (R^2_{adj}) values, average biomass in the training data (harvested data), root mean square error (RMSE) values, and *p*-values are also given. In the table, *c* refers to %-cover and *h* to height.

Plant functional group	Site	Equation	$R^2_{adj.}$	Average biomass ($g\ m^{-2}$)	RMSE ($g\ m^{-2}$)	*p*-value
Dwarf shrubs	Cross-site	$\mathrm{bm} = -31.85+2.98\mathrm{x}c+8.01\mathrm{x}h$	0.52	121.6	86.9	< 0.0001
	Sodankylä	$\mathrm{bm} = -15.12+90.46\mathrm{xasin}(c)+1.58\mathrm{x}h$	0.65	35.7	19.3	< 0.0001
	NW Russia	$\mathrm{bm} = -186.89+265.36\mathrm{xasin}(c)+61.06\mathrm{x}\sqrt{h}$	0.50	146.6	92.9	< 0.0001
	Herschel	$\mathrm{bm} = -0.86+19.75\mathrm{xasin}(c)$	0.87	61.2	18.9	< 0.0001
	Tiksi	$\mathrm{bm} = -46.83+136.72\mathrm{xasin}(c)+26.56\mathrm{x}\sqrt{h}$	0.61	40.4	26.2	< 0.0001
Betula nana	Cross-site	$\sqrt{\mathrm{bm}} = 1.05+0.22\mathrm{x}c+ 0.22\mathrm{x}h$	0.79	226.7	212.4	< 0.0001
	Sodankylä	$\sqrt{\mathrm{bm}} = -0.38+0.22\mathrm{x}c+0.11\mathrm{x}h$	0.92	74.8	16.4	< 0.0001
	NW Russia	$\sqrt{\mathrm{bm}} = -5.07+0.22\mathrm{x}c+ 2.56\mathrm{x}\sqrt{h}$	0.79	256.9	224.7	< 0.0001
	Herschel	n.a.	-	-	-	-
	Tiksi	$\mathrm{bm} = 1.37+3.30\mathrm{x}c$	0.62	30.6	17.4	< 0.0001
Salix spp.	Cross-site	$\sqrt{\mathrm{bm}} = 0.88+ 0.28\mathrm{x}c+0.13\mathrm{x}h$	0.78	154.8	252.4	< 0.0001
	Sodankylä	n.a.	-	-	-	-
	NW Russia	$\sqrt{\mathrm{bm}} = -0.36+0.30\mathrm{x}c+0.14\mathrm{x}h$	0.72	403.2	424.4	< 0.0001
	Herschel	$\sqrt{\mathrm{bm}} = -0.09\mathrm{x}0.40\mathrm{x}c+0.15\mathrm{x}h$	0.83	77.3	30.2	< 0.0001
	Tiksi	$\mathrm{bm} = 0.95+1.97\mathrm{x}c$	0.84	25.2	10.3	< 0.0001
Herbs	Cross-site	$\mathrm{bm} = -8.46+94.21\mathrm{xasin}(c)-3.24\mathrm{xln}(h)$	0.63	20.2	17.1	< 0.0001
	Sodankylä	$\sqrt{\mathrm{bm}} = 0.41+0.10\mathrm{x}c+0.05\mathrm{x}h$	0.83	9.3	4.1	< 0.0001
	NW Russia	$\mathrm{bm} = -20.75+80.38\mathrm{xasin}(c)+3.95\mathrm{xln}(h)$	0.63	22.2	16.2	< 0.0001
	Herschel	$\mathrm{bm} = -0.02+1.78\mathrm{x}c+0.85\mathrm{xln}(h)$	0.84	31.3	16.0	< 0.0001
	Tiksi	$\mathrm{bm} = 1.15+86.58\mathrm{xasin}(c)-6.67\mathrm{xln}(h)$	0.73	12.6	9.9	< 0.0001
Graminoids	Cross-site	$\sqrt{\mathrm{bm}} = -2.06+6.53\mathrm{xasin}(c)+1.08\sqrt{h}$	0.64	27.9	23.8	< 0.0001
	Sodankylä	$\mathrm{bm} = 20.62+0.46\mathrm{x}c-0.29\mathrm{x}h$	0.24	30.3	9.9	0.0005
	NW Russia	$\sqrt{\mathrm{bm}} = -2.48+7.04\mathrm{xasin}(c)+1.07\mathrm{x}\sqrt{h}$	0.66	25.3	25.5	< 0.0001
	Herschel	$\sqrt{\mathrm{bm}} = -4.58+12.48\mathrm{xasin}(c)+2.36\mathrm{xln}(h)$	0.77	36.9	16.8	< 0.0001
	Tiksi	$\mathrm{bm} = -5.71+1.23\mathrm{x}c+0.89\mathrm{x}h$	0.89	31.6	9.8	< 0.0001
Mosses	Cross-site	$\mathrm{bm} = -4.71+136.97\mathrm{xasin}(c)$	0.28	124.9	91.7	< 0.0001
	Sodankylä	$\mathrm{bm} = -96.01+244.10\mathrm{xasin}(c)$	0.28	227.0	154.2	< 0.0001
	NW Russia	$\mathrm{bm} = 20.72+100.37\mathrm{xasin}(c)$	0.21	110.1	75.5	< 0.0001
	Herschel	n.a.	-	-	-	-
	Tiksi	$\mathrm{bm} = -49.06+180.46\mathrm{xasin}(c)$	0.38	139.7	52.9	0.0022

Table 5. Regression equations, adjusted coefficient of determination (R^2_{adj}) values, and *p*-values for biomass-satellite spectra regressions. For all sites, the results are shown for regressions with the lowest root mean square error value using both site-specific and cross-site height-cover based biomass estimations. In the table, bm refers to aboveground biomass.

Site	Cover/height-based estimation	Best regression	R^2_{adj}	*p*-value
Cross-site	Cross-site	ln(bm) = 3.84+31.19xSAVI+13.02xRED-23.54xNIR	0.47	< 0.0001
Sodankylä	Cross-site	ln(bm = 5.73-12.03xRATIO+135.53xCOASTAL-13.39xGREEN	0.33	< 0.0001
	Site-specific	ln(bm) = 6.01-11.17xRATIO+77.93xCOASTAL	0.33	< 0.0001
NW Russia	Cross-site	ln(bm) = 8.32-8.35xRATIO+2.75xRGI-1.86xNIR	0.51	< 0.0001
	Site-specific	ln(bm) = 8.39-8.37xRATIO+2.91xRGI-2.01xNIR	0.51	< 0.0001
Herschel	Cross-site	ln(bm) = 7.26-6.07xRATIO-5.03xRED-EDGE	0.75	< 0.0001
	Site-specific	ln(bm) = 2.89+4.69xNDVI-4.43xRED-EDGE	0.68	< 0.0001
Tiksi	Cross-site	ln(bm) = 9.50-17.0xRATIO+35.8xRED-8.71xNIR	0.63	< 0.0001
	Site-specific	ln(bm) = 8.57-14.24xRATIO+25.76xRED-5.55xNIR	0.66	< 0.0001

Table 6. Average satellite spectra based biomass estimate in the study plots (Plot), root mean square error of the biomass-satellite spectra model (RMSE), and average biomass estimate in the overall landscape for each study site. In the "Regression combination" column, ss refers to site-specific and cs to cross-site model, with first acronym pointing to biomass-cover/height regression and second acronym to biomass-satellite spectra regression. Combinations subplot-ss and subplot-cs refer to alternative biomass-satellite spectra regressions that use harvested biomass instead of cover/height modelled biomass as the response variable. In cs-cs and subplot-cs combinations, regressions were carried out with data from all study sites, but RMSE is calculated based on study site specific training and fitted data. Water bodies were masked out of the images before calculating the average landscape biomass.

	Sodankylä			NW Russia			Herschel			Tiksi		
Regression combination	Plot biomass (g m $^{-2}$)	RMSE (g m $^{-2}$)	Landscape biomass (g m $^{-2}$)	Plot biomass (g m $^{-2}$)	RMSE (g m $^{-2}$)	Landscape biomass (g m $^{-2}$)	Plot biomass (g m $^{-2}$)	RMSE (g m $^{-2}$)	Landscape biomass (g m $^{-2}$)	Plot biomass (g m $^{-2}$)	RMSE (g m $^{-2}$)	Landscape biomass (g m $^{-2}$)
ss-ss	291.9	150.0	315.2	392.7	485.9	477.5	185.1	71.9	222.8	154.6	48.5	151.4
cs-ss	317.9	171.8	341.4	388.3	480.8	469.3	156.2	43.9	191.4	158.8	54.3	154.1
cs-cs	384.2	234.1	290.5	342.9	509.8	367.5	359.8	222.6	428.9	128.2	75.1	132.1
subplot-ss	287.4	231.1	360.4	505.1	535.2	587.5	213.5	94.9	262.2	144.8	70.6	141.6
subplot-cs	335.9	237.5	342.6	617.8	584.4	442.5	230.6	258.9	718.1	158.9	87.1	183.0