Delineation of the forest-tundra ecotone using texture-based classification of satellite imagery

Wenkai Guo[a]*, Gareth Rees[a], and Annika Hofgaard[b]

[a] *Scott Polar Research Institute, University of Cambridge, Cambridge, United Kingdom;*
[b] *Norwegian Institute for Nature Research, Trondheim, Norway*
**Corresponding author. Email: wg241@cam.ac.uk*

Delineation of the forest-tundra ecotone using texture-based classification of satellite imagery

The transition zone between the boreal forest and Arctic tundra, the forest-tundra ecotone (FTE), is an area of high ecological and climatological significance. Despite its importance, a globally consistent high spatial resolution mapping is lacking. Accurate mapping of the FTE requires the use of satellite remote sensing data. Here we use the Landsat Vegetation Continuous Fields (VCF) product and reference point data to derive the location and characteristics of the FTE. An image texture-based supervised classification scheme is developed based on a study area in Central Eurasia to statistically exploit the spatial patterns of the transition zone. Texture statistics for the VCF image are derived from the grey-level co-occurrence matrix (GLCM) based on which the study area is classified into forest, tundra, and FTEs. Adaptive parameterisation is implemented to achieve optimal classification performance in the study area. This method is further applied to six additional study areas around the circumarctic region to test its adaptability. In all study areas, this method achieves better FTE delineation results than previously reported methods, showing better classification accuracies (average of 0.826) and more realistic and complete representation of the FTE as shown by visual examination. This shows the universal applicability of the method and its potential to be used to achieve more detailed and accurate circumarctic mapping of the FTE, which could serve as the basis of time series analysis of FTE positions, eventually contributing to a better understanding of the inter-relations between climate change and shifts in sub-arctic vegetation.

Keywords: forest-tundra ecotone; Landsat VCF; sub-arctic vegetation, texture analysis; image classification

1. Introduction

The forest-tundra ecotone (FTE), also interchangeably termed the taiga-tundra ecotone (TTE) or the ‘arctic treeline,’ is the transition zone from closed canopy forest to treeless tundra, featuring changes in tree cover and density, tree size and shape (Sveinbjörnsson, Hofgaard and Lloyd, 2002), with a variety of spatial patterns which challenge globally consistent high resolution mapping (ref.). As a circumarctic phenomenon, the FTE is the world’s largest vegetation transition zone (Ranson, Montesano and Nelson, 2011), spanning more than 13400km in length and up to several hundred kilometres in width (Callaghan *et al.*, 2002). The configuration, composition and dynamics of the FTE vary greatly through time and across the circumarctic region according to local to regional abiotic and biotic drivers including disturbance history (Hofgaard, Harper and Golubeva, 2012; Timoney *et al.*, 2018).

Most current global vegetation models predict encroachment of boreal forest onto the treeless tundra in response to global warming (Larsen *et al.*, 2014). Recent Earth System Models (Settele *et al.*, 2014) show a general northward migration trend, with forest areas displacing between 11% and 50% of the tundra within 100 years (Larsen *et al.*, 2014). However, circumarctic and worldwide, forest advance in FTE areas since 1900 has been observed in only about half of sites studied (Harsch *et al.*, 2009) despite considerable climate change (Larsen *et al.*, 2014). Regional influences complicate the actual patterns of FTE movement (Callaghan, Werkman and Crawford, 2002; Rees *et al.*, 2002; Harsch *et al.*, 2009; Van Bogaert *et al.*, 2011; Hofgaard *et al.*, 2013; Larsen *et al.*, 2014). The FTE in different subarctic regions has been found to remain static or in dynamic equilibrium (Masek, 2001), show an increase in biomass or crown closure without moving (Payette, Fortin and Gamache, 2001), or be within a northward (Esper and Schweingruber, 2004; Gamache and Payette, 2005; Hofgaard *et al.*, 2013) or southward (Vlassova, 2002; Crawford, Jeffree and Rees, 2003; Montesano *et al.*, 2009) movement stage, with different displacement rates found for different species and different structures (Payette, Fortin and Gamache, 2001; Crawford, 2008; Hofgaard *et al.*, 2013). This diversity of FTE change modes and discrepancy between model predictions and empirical findings (Van Bogaert *et al.*, 2011; Hofgaard *et al.*, 2013; Timoney *et al.*, 2018; Rees et al. in prep.) emphasise the need for precise spatial mapping of the current circumarctic FTE.

A clear characterisation of the circumarctic FTE and its temporal progression through observation is lacking despite its high ecological significance (Harsch and Bader, 2011; Montesano *et al.*, 2016). Because of the vastness and predominant remoteness of the transition zone, remote sensing is the only feasible approach to retrieve its configuration (Rees *et al.*, 2002; Ranson, Montesano and Nelson, 2011). Various remote sensing methods can be utilised for this purpose, including spectral based methods using products from multispectral imagery, e.g. NDVI (Normalised Difference Vegetation Index) and VCF (Vegetation Continuous Fields) products, Synthetic Aperture Radar (SAR) data, or hyperspectral imagery through which spectral profiles for different vegetation types can be established and monitored (Govender, Chetty and Bulcock, 2007; Hu and Li, 2007; Darvishzadeh, 2008; Im and Jensen, 2008; Li, Chen and Chen, 2010; Wu and Peng, 2012, Walther et al, 2019). Traditionally, studies have defined the FTE according to vegetation metrics including tree density and cover, tree height, biomass, tree growth form and proportions of different vegetation types e.g. tree-to-tundra area ratio (Timoney *et al.*, 1992; Callaghan *et al.*, 2002; Montesano *et al.*, 2014, 2016).

The spatial structure of the FTE can potentially be exploited as a tool through which the FTE can be statistically separated from other landcover classes (forest and tundra). In remote sensing imagery, the spatial arrangement of surface feature can potentially be recognised using surface texture analysis. The inclusion of texture information into image classification and land cover mapping have improved classification accuracies (Blom and Daily, 1982; Greenspan and Goodman, 1993; Haack and Bechdol, 2000; Ferro and Warner, 2002; Coburn and Roberts, 2004; Herold, Haack and Solomon, 2004; Otukei, Blaschke and Collins, 2012), and proven helpful in vegetation analysis and mapping (Coburn and Roberts, 2004; Wood *et al.*, 2012). However, no study has yet incorporated texture information into the derivation of FTE areas.

The aim of this study is to provide an FTE delineation algorithm which incorporates texture measures into a supervised classification scheme using Landsat VCF and reference point data. We further aim to make the algorithm adaptable to be used in different regions through variable parameterisation adjusted for optimal performance locally. For this purpose, we developed the algorithm in a study area in Central Eurasia, and tested the method in additional study areas around the circumarctic region.

2. Materials and methods

2.1. Study areas

This study splits the circumarctic region into seven sub-regions following the scheme of Montesano et al. (2009). Longitudinal limits of each region are: Eastern Canada (ECA): 55°W–80°W; Central/Western Canada (CWCA): 80°W–130°W; Alaska (ALA): 130°W–170°W; Eastern Eurasia (EEU): 180°E–110°E; Central Eurasia (CEU): 110°E–60°E; Western Eurasia (WEU): 60°E–40°E; Scandinavia (SCA): 40°E–4°E (Figure 1). The algorithm development part of this study focuses on a region in Central Eurasia (128.52 by 150.72 km, centring on 61.928E, 66.953N) which straddles the transition from forest to tundra, thus encompassing the regional FTE (Figure 2). The region was chosen where a recent circumarctic FTE characterisation (Ranson, Montesano and Nelson, 2011) overlap with the location of the northern limit of boreal forest as shown by the Circumpolar Arctic Vegetation Map (CAVM), which is a circumarctic-scale vegetation map based on Advanced Very High Resolution Radiometer (AVHRR) data (Walker *et al.*, 2005). Six additional study areas of similar sizes to the Central Eurasia study area are chosen to test the applicability of the developed FTE delineation method, one in each of the other six sub-regions (Figure 1).

Please put Figure 1 here.

2.2. Data

2.2.1. Vegetation Continuous Fields

The VCF is an estimate of the proportion of a pixel occupied by tree cover derived from multi-spectral satellite remote sensing images (Hansen & DeFries 2004). Formally, the pixel value is an estimate of the amount of skylight obstructed by tree canopies of at least 5m in height (Montesano et al. 2009). Thus, a VCF image is a continuous (per-pixel) representation of vegetation cover across space which depicts areas of heterogeneous landcover, such as the FTE, better than traditional discrete classification schemes (Montesano *et al.*, 2009; DiMiceli *et al.*, 2011; Townsend *et al.*, 2011). The first VCF product is generated from Moderate Resolution Imaging Spectroradiometer (MODIS) data at a spatial resolution of 250m, with yearly coverage from 2000 to present (DiMiceli et al 2011). The 250m spatial resolution and relatively long temporal coverage make such products potentially suitable for large-scale study of ecotone dynamics (Stow *et al.*, 2004; Montesano *et al.*, 2009), and have been used by numerous studies to map tree cover (Cross and Settle, 1991; Zhu and Evans, 1994; Mayaux and Lambin, 1997; Tottrup *et al.*, 2007; Heiskanen and Kivinen, 2008). A global FTE product already exists at MODIS resolution, i.e. the Ranson et al. (2011) FTE, which is based on image segmentation on MODIS VCF data adjusted using Quickbird-derived tree cover estimates.

This study uses the Landsat VCF product as the primary data source, which is the MODIS VCF product densified to 30 m resolution using Landsat images. It thus having improved discriminatory power for small forest patches and increased ability to identify vegetation transitions more accurately. It is currently the highest-resolution multi-temporal global dataset of tree cover, and has been shown to have similar accuracies to MODIS VCF (Sexton *et al.*, 2013). The most recent version of the dataset ,version 3 (Sexton *et al.*, 2013), is used in this study, which covers four nominal epochs: 2000, 2005, 2010 and 2015, derived from MODIS VCF data in the corresponding years. However, visual examination of Landsat VCF in our study area shows that the product suffers from artefacts that are cloud and shadow contamination and inconsistencies among VCF values from different scenes, and the severity of these defects varies greatly between epochs. This study uses the Landsat VCF dataset having the fewest apparent artefacts and thus the best quality among the available epochs (the 2000 epoch for the Central Eurasia study area), which is ensured through visual inspection (Figure 2(c)).

Please put Figure 2 here.

2.2.2. Reference point data

Reference data points where the type of vegetation cover can be identified are needed for training and validation purposes. For this, 100 randomly distributed points are generated for the selected study region through the 'Create Random Points' function in the ArcMap 10.4 software. The landcover class of the data points are determined through visual interpretation of vegetation distribution in the area surrounding each point, thus taking into consideration local context. This is achieved through the examination of high-resolution Google Earth coverage of the study area (Sentinel-2 data, 10m spatial resolution). Thus, the points are divided into four landcover classes (forest, two types of FTE and tundra, Figure 2(c)).

It is necessary to separate FTEs into two small-scale and large-scale ones (hereafter referred to as FTE1 and FTE2, respectively) as they both represent a transition from forest to tundra, but at considerably different spatial scales, thus having vastly different spatial texture features. Therefore, they can confuse the classification scheme if regarded as a single class. This distinction between two FTE classes is different than in the Ranson et al. (2011) study where the FTE is also separated into class 1, which are image segments with mean VCF values between 5 and 20, and class 2 which are those with mean VCF values of less than 5 but with standard deviation values of larger than 5. The examination of spatial texture relies on focal analysis on a small area around each VCF pixel, and the transition in FTE1 occurs in similar spatial scale to these focal areas. Thus, FTE1 are mostly altitudinal FTEs, but also small-scale FTEs without significant elevational change. On the other hand, FTE2 represents transition zones much larger in scale, and thus appear to be pixels surrounded by windows composed of relatively uniform pixels having 'intermediate' VCF values. These pixels correspond to large-scale latitudinal FTEs. Additional data points are manually added for the FTE1 class which have very few data points randomly generated, while still ensuring relatively even distribution of all data points. The final numbers of forest, FTE1, FTE2 and tundra data points are 20, 22, 39, and 19, respectively. Additionally, three altitudinal FTE data points (therefore FTE1 points) available from the published literature (Wilmking *et al.*, 2012) are included in the study (Figure 2(c)). Landsat VCF data and the reference point data are then processed in subsequent steps (Figure 3) into an FTE map of the study area:

Please put Figure 3 here.

2.3. VCF thresholding

To provide a baseline for performance assessment of the texture-based classifiers, we first perform simple thresholding of the VCF data. The Ranson et al. (2011) study identified the FTE as image segments with mean VCF percentages between 5 and 20, or those with mean VCF percentages of less than 5 and standard deviation values larger than 5. This threshold pair envelops the 'intermediate VCF values' that are considered to represent the core of the FTE (Ranson, Montesano and Nelson, 2011). However, forests in various parts of the circumarctic region may have different ranges of VCF values because of differences in structure and composition. Therefore, an experiment is conducted to find the pair of VCF thresholds with which reference data points in different landcover classes could be best distinguished. Thus, the intermediate VCF value envelope is derived programmatically to best fit the study area. According to this model, a pixel is classified as forest if its VCF value is above some upper threshold and as tundra if it is below some lower threshold. Pixels with VCF values in between are classified as FTE. All possible combinations of two VCF thresholds from 1 to 100 are investigated, and the classification accuracies and kappa coefficients (Cohen, 1960) are recorded. The threshold pair that gives the best accuracies as measured by these two metrics is selected as the optimal threshold pair for FTE characterisation in the study area. Preference is given to the threshold pair that gives the highest kappa coefficients when the result judging from the two metrics differ, since the kappa coefficient takes into account the possibility of agreement occurring by chance and is considered more robust statistic than simple accuracy. This new threshold pair (hereafter referred to as the adaptive threshold pair) is compared with the Ranson et al. (2011) 5-20 threshold pair to test their abilities to correctly separate different landcover classes.

2.4. FTE delineation based on supervised classification utilising texture measures

2.4.1. Texture measures used in this study

Common measures of texture include first-order statistics such as variance, and second-order statistics calculated on the basis of the grey-level co-occurrence matrix (GLCM) (Ferro and Warner, 2002). The calculation of GLCMs, as proposed by Haralick et al. (1973), has proved to be one of the most powerful tools to extract information of spatial structure from remote sensing images (Weszka, Dyer and Rosenfeld, 1976; Conners and Harlow, 1980). It is a tabulation of how often different combinations of grey levels co-occur in an image or image section (Yang *et al.*, 2009), based on which numerous texture features can be derived to represent local spatial variations at pixels of interest. In this study, a total of 11 GLCM-based texture measures (termed primary texture measures, Table 1, Equations (1)-(11)) are analysed for their ability to distinguish between different landcover classes. In addition, eight texture measures derived from the primary GLCM-based textures are included in the analysis (termed secondary texture measures, Table 1, Equations (12)-(19)). Thus, a total of 19 texture measures are used in this study.

Please put Table 1 here.

Where:

$P_{i,j}$ is the (i, j)th entry in the GLCM; μ_x, μ_y,σ_x and σ_y are the means and standard deviations of p_x and p_y; N_g is the number of distinct grey levels in the quantised image;

$\sum_i$ is $\sum_{i=1}^{N_g}$; $\sum_j$ is $\sum_{j=1}^{N_g}$;

$P_{x+y}(k) = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P_{i,j}$, ($i$+$j$=$k$; k=2,3,…,2N_g); $P_{x-y}(k) = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P_{i,j}$, (|$i$ –j|=k; k=0,1,…,N_g-1);

$P_x(i) = \sum_{j=1}^{N_g} P_{i,j}$; $P_y(j) = \sum_{i=1}^{N_g} P_{i,j}$;

$\mathrm{HXY} = -\sum_i\sum_j P_{i,j}\log P_{i,j}$; $\mathrm{HXY1} = -\sum_i\sum_j P_{i,j}\log p_x(i)p_y(j)$;

$\mathrm{HXY2} = -\sum_i\sum_j p_x(i)p_y(j)\log p_x(i)p_y(j)$; and HX and HY are entropies of p_x and p_y.

2.4.2. Derivation of optimal window size

Our FTE characterisation method relies on texture analysis, which considers not only VCF values of the selected points, but also the spatial configuration of the landscape within the surrounding windows. The actual implementation of texture analysis needs to be adapted for different regions because of the difference in the spatial configuration of FTE areas, requiring different parameterisation in the texture analysis algorithm. An appropriate window size and a suitable set of texture measures are key parameters in the texture analysis, and need to be determined first. In this study, an optimal window size is determined before the derivation of optimal textures. This is because the optimal window size is a distance at which textures from different landcover classes can be properly separated. It is therefore a geographic phenomenon independent of texture selection, and is only dependent on the scale at which the unique textures of the FTE are identifiable. More importantly, differences in window size can directly influence the separating power of the texture measures, i.e. texture measures perform differently when applied with different window sizes (Ge *et al.*, 2006).

The determination of an appropriate window size for texture analysis is crucial for two reasons. Firstly, texture measures are calculated within a window around each point, and the window size must be appropriate so that it is smaller than the object, in our case the FTE, but big enough to include the characteristic variability of the object (Hall-Beyer, 2017). Secondly, past studies have shown increased class separability with the incorporation of texture in addition to spectral information in image classification, and this benefit generally increases with larger window sizes which reduce random error and thus produces more stable textures. However, larger window sizes also lead to larger edge effects and introduce systematic errors. More importantly, the window size needs to be compatible with the scale of texture resolvable by the remote sensing product used. Instead of using arbitrary and fixed geometric windows regardless of study area, this study produces data-driven geographic windows in a window size with which texture analysis is able to produce maximum separability between different landcover classes.

Spatial statistical methods like the semivariogram can potentially be used to determine the scales of spatial variability in the VCF image, and thus to estimate optimal window sizes in texture analysis. However, in this study we utilise the information from the data points to specifically find the scale at which the FTE classes can be optimally separated from other classes, thus yielding more focused and meaningful spatial scale outcome. Specifically, the separability between data points of different landcover classes is calculated for the Central Eurasia region using all the 19 GLCM-based texture measures. This process is repeated for window sizes from 3 to 91 pixels to encompass the range of window sizes in which different landcover classes can be identified through visual inspection. We adopt the Transformed Divergence as a statistical measure to assess the separability between landcover classes. Transformed Divergence and the Jeffries Matusita Distance are both commonly used for this purpose (Davis *et al.*, 1978), and while they have been found to have similar performances in assessing class separability (Gong, Marceau and Howarth, 1992), the Jeffries Matusita Distance is computationally less efficient (Jensen and Lulla, 1987). Transformed Divergence (TD) is defined as follows: (Otukei, Blaschke and Collins, 2012),

$$\mathrm{TD}_{ij} = 2\left(1 - e^{\frac{-D_{ij}}{8}}\right) \tag{20}$$

where:

$$D_{ij} = \frac{1}{2}\mathrm{trace}\left(\left(\mathbf{C}_i - \mathbf{C}_j\right)\left(\mathbf{C}_i^{-1} - \mathbf{C}_j^{-1}\right)\right) + \frac{1}{2}\mathrm{trace}\left(\left(\mathbf{C}_i^{-1} - \mathbf{C}_j^{-1}\right)\left(\mu_i - \mu_j\right)\left(\mu_i - \mu_j\right)^T\right) \tag{21}$$

The subscripts i, j represent signatures of the selected classes; $\mathbf{C}_i$ and $\mathbf{C}_j$ are covariance matrices of i and j; μ_i and μ_j are mean vectors of i and j.

Transformed Divergence has a range of 0 to $2\sqrt{2}$, with higher values showing higher separability. Usually, Transformed Divergence values of higher than 1.9 are deemed to represent separable classes, while those between 1.7 and 1.9 represent good separation and those below 1.7 shows poor separation (Jensen 1996). Since the purpose of our study is to isolate FTE from other landcover classes, transformed divergence values are only calculated between the FTE category (both FTE1 and FTE2) and forest and tundra, thus resulting in calculated values for four landcover class pairs (FTE1 – forest, FTE1 – tundra, FTE2 – tundra, and FTE2 – forest). The window size at which maximum total separability is achieved in all the class pairs is chosen as the optimal window size to be used in subsequent steps.

2.4.3. Derivation of suitable texture measures

All 19 GLCM-based texture measures (Table 1) are calculated for all data points (Figure 2(c)) using the determined optimal window size (cf. above). The next task is the determination of an optimal set of texture measures which can separate the FTE from other landcover classes in a statistically robust way. This is conducted in a two-step process. In the first step, for each landcover class, mean values of all 19 texture measures for all the data points are calculated. T-tests are then performed to assess the separability between average texture values from data points in each landcover class pair. Since the variance of the VCF values of the four landcover classes and therefore that of the resulted texture measures may not be equal, two-sample F-tests are conducted to determine the equality of variance, and subsequent t-tests are altered in accordance to the F test results. If variances are determined to be unequal, Satterthwaite's approximation of the effective degrees of freedom is used (Satterthwaite, 1946). A texture measure is retained only if it shows the ability to separate either or both the FTE classes from other classes, i.e. reporting with statistical significance that the texture measure averages of points in FTE classes are different from those in both forest and tundra classes.

In the second step, the remaining texture measures filtered by the t-tests go through the Spearman rank correlation test to determine their collinearity, and texture measures which correlate strongly with others and hence provide minimal additional discriminating power are excluded. This test is used because of its nonparametric properties and tolerance of extreme values, and its ability to test for monotonic relationships that are not necessarily linear. Specifically, an 8-pixel neighbourhood area (3 by 3) is constructed centred on each data point, and each remaining texture measure is calculated for every pixel within this neighbourhood. Summary mean and standard deviation values are calculated for the texture measures in these neighbourhoods.

Spearman rank correlation is then calculated for each pair of texture measures based on the mean and standard deviation summaries for all data points, assessing their collinearity. The result are Spearman rank matrices for the mean and standard deviation summaries of every texture pair.

For each texture measure, its Spearman rank correlation coefficients with all other texture measures are averaged, and the five texture measures with the lowest averages of the mean summaries are kept for further analysis. Then, the rest of the texture measures with at least one mean summary that shows no significant correlation (p value > 0.01) with others are kept. Texture measures filtered out by these two steps have relatively strong correlation with others and should be eliminated from further analysis. However, exceptions can be made when the standard deviation summaries are not strongly correlated, suggesting their ability to capture unique textural heterogeneity (Wood *et al.*, 2012). Thus, five texture measures among those filtered out by the two-step process having the smallest averaged standard deviation summaries are kept for further analysis.

2.4.4. Supervised classification

The selected texture measures from the previous step are calculated for the entire image using the optimal window size. The resulting texture measures are then fed into the maximum likelihood classification algorithm. Water body pixels (identified in the VCF product by the mask value of 200) are ignored in the classification process. In order to convey the distinction between VCF values for different landcover classes to the classification algorithm, thresholded VCF images are created both using the 5-20 and adaptive threshold pairs, and then also fed into the classification process. The classification is executed with randomly selected half of the data points as training data, and the other half for validation, while ensuring that half of each landcover class are kept for both training and validation. Due to the fragmented nature of the classification result, a generalisation process involving image segmentation is performed to filter out FTE segments too small in size in order to achieve a more desirable transition zone feature. After the classification process, the FTE1 and FTE2 classes are merged into a single FTE landcover class. The final FTE derivation results are compared to previously delineated FTEs qualitatively through visual inspection and quantitatively through classification accuracy and kappa coefficient.

2.5. Application to additional study areas

The above FTE delineation method developed in Central Eurasia is used on FTEs in the other six study areas to test its applicability. To streamline the data retrieval process and enhance the adaptability of our method, we explore the feasibility of vegetation data retrieval and processing from the Google Earth Engine platform, hereafter referred to as GEE (Gorelick *et al.*, 2017). Landsat VCF data intersecting with the study areas of the best quality are downloaded, and the derivation of optimal window sizes and texture measures are performed locally in MATLAB. Texture image calculation using the derived parameters are performed in GEE, and the resulting texture images are downloaded to be used in supervised classification in ArcMap. GEE currently has 18 GLCM textures available, two of which are duplicates (inertia and contrast), thus making a total of 17 usable texture measures. Autocorrelation (AUT) and inverse difference (IND) are not available in GEE, and we replace them by similar-performing measures, i.e. correlation (COR) and homogeneity (HOM), respectively (Haralick, Shanmugan and Dinstein, 1973). If either or both of the latter two are also among the selected list of texture measures, no replacement of the former two is given. The same supervised classification method as used in the Central Eurasia study area is implemented to separate FTE with other landcover classes, in which training and validation data also come from randomly generated reference points in the study areas (Figure 10). This workflow offloads the most time-consuming tasks (VCF data retrieval and texture image calculation) to GEE's cloud-computation platform which saves a considerable amount of processing time. It also ensures that the detailed statistical procedures developed in this study are followed through local processing, which are much less time-consuming and not available in GEE.

3. Results

3.1. VCF thresholding

The adaptive VCF threshold pair enveloping FTE pixels is determined to be 5 and 10 for the Central Eurasia region using the method described above, as this threshold pair yields the highest overall accuracy and kappa coefficient in separating forest, FTE and tundra data points (Figure 4).

Please put Figure 4 here.

The thresholding approach results in a pixelated thresholded image not desirable for the delineation of a transition zone (Figure 5), but serves as a reference of the distribution of VCF values within the study area. A more detailed look at a subset of the image (Figure 5 c-e) shows that the thresholded image produced from the adaptive threshold pair gives a more realistic representation of the forest areas corresponding to Google Earth visualization (Figure 5 d&e), and the forest areas in the 5-20 thresholded image show heavy encroachment from FTE points which produces very fragmented forest patches (Figure 5 c&e). This is presumably attributable to the adaptive threshold pair being derived directly using the VCF values of the data points, thus reflecting a better distinction between different landcover classes. Therefore, the adaptive threshold pair will be used to threshold the VCF image to be used in the supervised classification process.

Please put Figure 5 here.

3.2. FTE delineation based on supervised classification utilising texture measurements

3.2.1. Selection of optimal window size and optimal textures measures

Calculated Transformed Divergence of the class pairs (Figure 6) shows that the FTE1-Forest and FTE2-Tundra class pairs have generally higher separability. Both the FTE1-Tundra and FTE2-Forest class pairs reach a local maximum at the 15×15 window size, where the FTE1-Forest and FTE2-Tundra class pairs are also maintaining high levels of separability. Thus, a window size of 15×15 is deemed to be the optimal window size based on which the GLCMs and texture measures will be calculated and incorporated into the classification.

Please put Figure 6 here.

The selected window size is used to derive the optimal set of texture measures. The t-test keeps 17 texture measures based on which a Spearman rank matrix is established. Mean summaries of most texture measures are highly correlated, while the Spearman rank correlation coefficients calculated from standard deviation summaries have a wider spread. Based on the selection criteria described above, seven texture measures are kept: cluster shade, correlation, difference variance, homogeneity, information measure of correlation 2, inverse difference and maximum probability. They have low collinearity with other texture measures or higher collinearity with others but relatively low coefficients calculated from the standard deviation summaries.

3.2.2. Classification and segmentation results

Texture images are constructed based on the final list of texture measures using the optimal window size of 15×15, and are then fed into the classification process along with the thresholded VCF image. The classification based on the 5-20 and 5-10 threshold pairs after merging the two FTE classes (Figure 7) loses some of the fine details on the surface, which is expected from the nature of the windowing approach in texture analysis.

Please put Figure 7 here.

A more detailed qualitative comparison between and FTE areas derived using different methods in a subset of the study area (Figure 8) shows that classification based on the adaptive threshold pair yields a more realistic picture of FTE distribution comparatively when compared to the Google Earth coverage of the study area.

Please put Figure 8 here.

Quantitative evaluation of the classification result is conducted using half of the data points as validation (section 2.4.4). Quantitative assessment of the results (Table 2) shows that classification based on the adaptive threshold pair yields higher accuracies than that based on the 5-20 threshold pair. Simple thresholding produces similar and higher accuracies than classification based on the 5-20 threshold pair, but is outperformed by that based on the adaptive one, which produces higher accuracy and kappa coefficient than all other methods. It is therefore the optimal

FTE delineation approach for our study area. Classification accuracy and kappa coefficient calculated for the Ranson et al. (2011) FTE are based on two categories: FTE vs. non-FTE.

Please put Table 2 here.

The supervised classification method based on calculated texture measures and the adaptively thresholded VCF image, which produces the highest accuracies, is used to create the final output of this study (Figure 9) – a map of FTE pixels (in green) in the study area.

Please put Figure 9 here.

3.3. Application to additional study areas

The application of the above FTE delineation method to the additional six study areas is conducted mainly locally in MATLAB and ArcMap, with GLCM texture images calculated in GEE. On average, the application of GEE-based texture image calculation reduces the processing time from approximately 2.5 hours to approximately 3 minutes per texture measure, greatly expediting the analysis. FTEs derived using our classification method with these adaptive VCF threshold pairs consistently produce the highest classification accuracy compared to other methods, as can be seen from the comparison between Google Earth coverages of the study areas, Landsat

VCF dataset and reference point data, the Ranson et al., (2011) FTE, and the FTE derived using our method shown in Figure 10 and Table 3. FTE delineated using VCF thresholding also shows higher accuracies when using the adaptive threshold pairs. Supervised classification using the 5-20 threshold yields generally lower accuracies than VCF thresholding, except for the ALA and CWCA study areas where they show similar or higher accuracies.

Please put Table 3 here.

Please put Figure 10 here.

4. Discussion

The results suggest that the Landsat VCF product is a useful data source for FTE delineation which provides reasonable spatial resolution, and a texture-based classification method based on VCF values is able to reliably extract FTE information. For the Central Eurasia study area, the Landsat VCF product produces a more detailed depiction of the FTE area than the previous global FTE product (Ranson, Montesano and Nelson, 2011) derived from MODIS VCF, which is based on segmentation before thresholding with arbitrary limits of segment sizes. The MODIS-based FTE product creates large FTE patches that often include tundra areas that have been recognised as being within the same segments as the FTE pixels, see for example the FTE segment designated by letter ‘A’ in Figure 8.

Selecting the correct threshold pair is crucial for satisfactory performance of the texture-based classification method. FTE derived from classification based on the adaptive threshold pair produces smaller FTE patches than that based on the 5-20 one (Figure 7), which is also true for simple VCF thresholding (Figure 5), as expected. This corresponds to the forest areas in the study area producing VCF values of mostly around 10 to 20 due to relatively small biomass, thus making the 5-20 threshold pair unreliable. This hypothesis was partially validated by the typical tree heights of around 3-5 m in forests calculated from shadow length and capture time from the Google Earth coverage (e.g. Mathisen *et al.*, 2013). The inclusion of texture images into classification has resulted in improved classification accuracies, consistent with previous findings (e.g. Coburn and Roberts, 2004; Ferro and Warner, 2002; Otukei et al., 2012). Compared to other methods, classification based on the adaptive threshold pair yields a more realistic representation of FTE distribution when compared to the Google Earth coverage of the study area (Figure 8). This method also produces higher accuracy and kappa coefficient than all other methods. Simple thresholding produces similar and higher accuracies than classification based on the 5-20 threshold pair, further confirming the importance of adaptive thresholding in the classification algorithm.

For the additional study areas, the optimal VCF threshold pairs, window sizes and texture measures derived for different study areas vary considerably. The results show that FTE delineation using texture-based classification based on adaptive VCF thresholding produces consistently highest accuracies (Table 3), and again emphasises the need for adaptive parameterisation in achieving optimal FTE delineation results. Qualitatively, our method produces FTEs corresponding well with transition areas from forest to tundra shown in the Google Earth coverages (Figure 10). Our method largely produces FTEs with similar placements to the Ranson et al. (2011) FTE product, but with additional representation of small-scale FTEs and with more spatial details for large-scale FTEs. They have more similar FTE placements for study areas where large-scale FTEs are more spatially concentrated (WEU, EEU, ALA and CWCA). In other study areas with more spread-out FTEs (as verified by visual examination of the Google Earth coverage and also placement of VCF pixels with 'intermediate' VCF values), our method produces a more complete representation of the transition zone. The MODIS-based FTE product misses part of the FTE due to the limit in segment sizes and thus incomplete derivation of transition zones with spread-out FTE pixels. Thus, our study provides a viable approach to delineating both large and small-scale FTE areas across the circumarctic region.

The FTE delineation problem is highly scale-dependent. The MODIS VCF product provides good spatial and temporal coverage for circumarctic FTE delineation, but FTE recognition based on this product is limited by its 250m spatial resolution whereby local transitional details can be overlooked. The Landsat VCF product also provides global coverage but with finer spatial resolution, and our study proves that it can be used to derive large-scale FTE areas with the use of texture analysis. The Landsat resolution also enables the recognition of small-scale FTEs not resolvable by the MODIS VCF product. It is therefore a more versatile tool for the purpose of FTE delineation. With the even higher spatial resolution of the other satellite image products, e.g. Sentinel-2 data (K. Fletcher, 2012), more spatial characteristics of FTEs can be revealed, but the limited availability of usable cloud-free imagery limits its use in the effort at deriving a universally adaptable method for circumarctic FTE delineation.

One important source of error in this study is the high dependence on the selection of reference data points, which is based on inspection of high-resolution Google Earth coverage of the study areas in addition to point data derived from previous work. The adaptive selection of threshold pairs for dividing landcover classes based on VCF values, the calculation of optimal window size and optimal set of texture measures are all dependent on correct classification of reference data points. Data sources apart from locally generated random points are desirable to improve confidence in the ground truth. Such data are available, for example, through the PPS Arctic long-term monitoring network (http://ppsarctic.nina.no). Also, this study is built upon the VCF products and thus affected by inaccuracies in these datasets including systematic errors as well as the prevalent image artefacts (White, Shaw and Ramsey, 2005; Sexton *et al.*, 2013), which is a major consideration in the selection of the epoch of the VCF dataset. Future application of our method is likely limited by the availability of quality data in the areas of interest, which can potentially be remedied by future improvements in Landsat VCF data quality, local image fusion of Landsat scenes and MODIS VCF data, or the incorporation of higher-resolution datasets.

In this study, the maximum probability classifier is chosen in consideration of processing time given the number and sizes of the study areas, an also because the emphasis of this study is on the incorporation of image texture into the classification workflow. In future application of this method, more advanced classification techniques can be used to further improve on the performance of the classification process. Finally, this study only looks into the horizontal spatial arrangement of the landscape and does not include an analysis of the vertical dimension of the FTE. FTE delineation can benefit from elevation information since the occurrence and placement of altitudinal FTEs are associated with local topographic variation. Also, at a very high spatial resolution, FTE delineation can benefit from tree height information e.g. from satellite LiDAR products (Montesano et al., 2016b), as tree height variation is also an important component of vegetation structural change through the FTE.

4.1. Future tasks

This study provides an adaptable method for FTE delineation based on Landsat VCF which can potentially be used in different parts of the circumarctic region. A future task would be to create a circumarctic FTE map based on our method, a prerequisite of which is a reasonable division scheme of the circumarctic region which recognises the ranges of VCF values of different landcover classes in different regions. For example, the Montesano et al. (2009) division of the circumarctic region can be used as a starting point, based on which sensitivity analyses can be conducted to achieve geographically and ecologically meaningful sub-regions. Adaptive thresholds can then be established for each sub-region. This circumarctic FTE product based on different epochs of the Landsat VCF product (currently 2000, 2005, 2010 and 2015) can be used to construct a times series of FTE change through the past two decades. The derivation of a circumarctic product demands sufficient reference data points to be established whose landcover classes can be determined and verified, either through field work or visual recognition based on satellite imagery.

As previously noted, the location and spatial pattern of both latitudinal and altitudinal FTEs vary greatly across the circumarctic region. These regional differences represent the effect of a wide range of local influencing factors, the relative importance of which has great implications on the ecotone's vulnerability to shift with climatic change. Therefore, it is necessary to move beyond the task of FTE delineation and explore more detailed spatial patterns within the FTE areas. In this study, texture information is only used to separate FTE areas from tundra and forest. However, texture analysis is also potentially useful in the examination of the spatial configuration of FTEs in different regions. Through observation, recent studies have confirmed a close link between different FTE spatial patterns (FTE 'forms') and FTE movement in response to climate change (Holtmeier, 2010; Harsch and Bader, 2011). Each FTE form is unique in the spatial arrangement of vegetation which will be represented in their varying textures in remotely sensed images, which can be exploited to identify and map different FTE forms, thus facilitating the identification of FTEs that are the most vulnerable to shift with climate change. The analysis of these local variations will rely on higher resolution datasets such as Sentinel-2 data, and the correspondence between FTE forms and vulnerability can be validated by the incorporation of study sites where historical records of FTE movements are available.

5. Conclusion

This study introduces a texture-based classification approach to the FTE delineation problem. The incorporation of texture measures is theoretically relevant in FTE delineation because the FTE is a unique transition zone in which the mosaic distribution of forest and tundra creates unique spatial patterns inexistent in either side of the ecotone. Compared to other vegetation products, the reliable global coverage and reasonable spatial resolution provided by the

Landsat VCF is considered to be optimal for the purpose of FTE delineation. This study provides a versatile delineation approach of multi-scale FTEs based on the Landsat VCF dataset, and provide objective and adaptable approaches to every component of the texture-based FTE delineation process through statistical determination of analysis parameters. It is based on reference data points derived from expert knowledge and thus takes the specificities of the study area into consideration, and also considers the spatial patterns surrounding the data points. Compared to pixel-based thresholding and segmentation, our method provides a relatively natural representation of a transitional area, utilising the information of VCF gradient while preserving reasonable continuity of the interface, and is robust in handling small-scale variations. Quantitative assessment also suggests that our method is able to provide more accurate FTE delineation results than others. Our method can be potentially used to create a circumarctic map of the FTE based on which a time series of circumarctic FTE change can be derived. This can potentially serve as a more accurate baseline for future studies seeking to understand the interactions between arctic vegetation and climatic change, and help models to explain and predict vegetation response to global warming.

Acknowledgement

W.G. acknowledges financial support by the Cambridge Trust, and Trinity and Fitzwilliam colleges, University of Cambridge. W.G. expresses his immense gratitude towards the handling editor and the anonymous reviewers for their time and attention which helped the authors achieve a clear and relevant presentation of the research materials.

References

Blom, R. and Daily, M. (1982) 'Radar image processing for rock-type discrimination', *IEEE Transactions on Geoscience and Remote Sensing*, (3), pp. 343–351. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4157309.

Van Bogaert, R. *et al.* (2011) 'A century of tree line changes in sub-Arctic Sweden shows local and regional variability and only a minor influence of 20th century climate warming', *Journal of Biogeography*, 38(5), pp. 907–921. doi: 10.1111/j.1365-2699.2010.02453.x.

Callaghan, T. V *et al.* (2002) 'The Dynamics of the Tundra-Taiga Boundary: An Overview and Suggested Coordinated and Integrated Approach to Research', *Ambio Special Report*, (12), pp. 3–5.

Callaghan, T. V, Werkman, B. R. and Crawford, R. M. M. (2002) 'The Tundra-Taiga Interface and Its Dynamics: Concepts and Applications', *Ambio Special Report*, (12), pp. 6–14. doi: 10.2307/25094570.

Coburn, C. a. and Roberts, a. C. B. (2004) 'A multiscale texture analysis procedure for improved forest stand classification', *International Journal of Remote Sensing*, 25(20), pp. 4287–4308. doi: 10.1080/0143116042000192367.

Cohen, J. (1960) 'A coefficient of agreement for nominal scales', *Educational and psychological measurement*. Sage Publications Sage CA: Thousand Oaks, CA, 20(1), pp. 37–46.

Conners, R. W. and Harlow, C. A. (1980) 'A theoretical comparison of texture algorithms', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(3), pp. 204–222. doi: 10.1109/TPAMI.1980.4767008.

Crawford, R. M. M. (2008) *Plants at the margin: ecological limits and climate change*. Cambridge University Press.

Crawford, R. M. M., Jeffree, C. E. and Rees, W. G. (2003) 'Paludification and Forest Retreat in Northern Oceanic Environments', *Annals of Botany*, 91(2), pp. 213–226. doi: 10.1093/aob/mcf185.

Cross, A. and Settle, J. (1991) 'Subpixel measurement of tropical forest cover using AVHRR data', *International Journal of Remote Sensing*, 12(5), pp. 1119–1129. doi: 10.1080/01431169108929715.

Darvishzadeh, R. (2008) *Hyperspectral remote sensing of vegetation parameters using statistical and physical models, GeoInformation Science*. Available at: http://www.itc.nl/library/papers_2008/phd/roshanak.pdf.

Davis, S. M. *et al.* (1978) 'Remote sensing: the quantitative approach', *New York, McGraw-Hill International Book Co., 1978. 405 p.*

DiMiceli, C. M. *et al.* (2011) *MODIS Vegetation Continuous Fields*, *Global Land Cover Facility*. Available at: http://glcf.umd.edu/data/vcf/description.shtml.

Esper, J. and Schweingruber, F. H. (2004) 'Large-scale treeline changes recorded in Siberia', *Geophysical Research Letters*, 31(6), p. L06202. doi: 10.1029/2003GL019178.

Ferro, C. J. S. and Warner, T. A. (2002) 'Scale and texture in digital image classification', *Photogrammetric Engineering & Remote Sensing*, 68(1), pp. 51–63.

Gamache, I. and Payette, S. (2005) 'Latitudinal response of subarctic tree lines to recent climate change in eastern Canada', *Journal of Biogeography*, 32, pp. 849–862. doi: 10.1111/j.1365-2699.2004.01182.x.

Ge, S. *et al.* (2006) 'Texture analysis for mapping Tamarix parviflora using aerial photographs along the Cache Creek, California', *Environmental Monitoring and Assessment*, 114(1–3), pp. 65–83. doi: 10.1007/s10661-006-1071-z.

Gong, P., Marceau, D. J. and Howarth, P. J. (1992) 'A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data', *Remote Sensing of Environment*, 40(2), pp. 137–151. doi: 10.1016/0034-4257(92)90011-8.

Gorelick, N. *et al.* (2017) 'Google Earth Engine: Planetary-scale geospatial analysis for everyone', *Remote Sensing of Environment*. The Author(s), 202, pp. 18–27. doi: 10.1016/j.rse.2017.06.031.

Govender, M., Chetty, K. and Bulcock, H. (2007) 'A review of hyperspectral remote sensing and its application in vegetation and water resource studies', *Water SA*, 33(2), pp. 145–151. doi:

10.4314/wsa.v33i2.49049.

Greenspan, H. and Goodman, R. M. (1993) 'Remote Sensing Image Analysis via a Texture Classification Neural Network', *Advances in Neural Information Processing Systems 5*, pp. 425–432. Available at: http://papers.nips.cc/paper/636-remote-sensing-image-analysis-via-a-texture-classification-neural-network.pdf%5Cnfiles/3248/Greenspan ? Goodman - 1993 - Remote Sensing Image Analysis via a Texture Classi.pdf%5Cnfiles/3249/636-remote-sensing-image-analysis.

Haack, B. and Bechdol, M. (2000) 'Integrating multisensor data and RADAR texture measures for land cover mapping', *Computers and Geosciences*, 26(4), pp. 411–421. doi: 10.1016/S0098-3004(99)00121-1.

Hall-Beyer, M. (2017) 'Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales', *International Journal of Remote Sensing*, 38(5), pp. 1312-1338. doi: 10.1080/01431161.2016.1278314.

Haralick, R., Shanmugan, K. and Dinstein, I. (1973) 'Textural features for image classification', *IEEE Transactions on Systems, Man and Cybernetics*, pp. 610–621. doi: 10.1109/TSMC.1973.4309314.

Harsch, M. A. *et al.* (2009) 'Are treelines advancing? A global meta-analysis of treeline response to climate warming', *Ecology Letters*, 12(10), pp. 1040–1049. doi: 10.1111/j.1461-0248.2009.01355.x.

Harsch, M. A. and Bader, M. Y. (2011) 'Treeline form - a potential key to understanding treeline dynamics', *Global Ecology and Biogeography*, 20(4), pp. 582–596. doi: 10.1111/j.1466-8238.2010.00622.x.

Heiskanen, J. and Kivinen, S. (2008) 'Assessment of multispectral, -temporal and -angular MODIS data for tree cover mapping in the tundra-taiga transition zone', *Remote Sensing of Environment*, 112(5), pp. 2367–2380. doi: 10.1016/j.rse.2007.11.002.

Herold, N. D., Haack, B. N. and Solomon, E. (2004) 'An evaluation of radar texture for land use/cover extraction in varied landscapes', *International Journal of Applied Earth Observation and Geoinformation*, 5(2), pp. 113–128. doi: 10.1016/j.jag.2004.01.005.

Hofgaard, A. *et al*. (2013) 'Latitudinal forest advance in northernmost Norway since the early 20th century', *Journal of Biogeography*, 40(5), pp. 938–949. doi: 10.1111/jbi.12053.

Hofgaard, A., Harper, K. A. and Golubeva, E. (2012) 'The role of the circumarctic forest-tundra ecotone for Arctic biodiversity', *Biodiversity*, 13(3–4), pp. 174–181. doi: 10.1080/14888386.2012.700560.

Holtmeier, K. (2010) 'Altitudinal and polar treelines in the northern hemisphere Causes and response to climate change', *Polarforschung*, 79(September 2009), pp. 139–153.

Hu, B. and Li, Q. (2007) 'Vegetation classification using hyperspectral remote sensing and singular spectrum analysis', in *Proc. SPIE*, p. 66960N. Available at: http://dx.doi.org/10.1117/12.735278.

Im, J. and Jensen, J. R. (2008) 'Hyperspectral remote sensing of vegetation', *Geography Compass*, 2(6), pp. 1943–1961. doi: 10.1111/j.1749-8198.2008.00182.x.

Jensen, J. R. and Lulla, K. (1987) 'Introductory digital image processing: a remote sensing perspective'. Taylor & Francis.

K. Fletcher (2012) *Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services*. Available at: https://sentinel.esa.int/web/sentinel/home.

Larsen, J. N. *et al*. (2014) 'Polar regions', in Barros, V. R. et al. (eds) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 1567–1612.

Li, D., Chen, S. and Chen, X. (2010) 'Research on method for extracting vegetation information based on hyperspectral remote sensing data', *Transactions of the Chinese Society of Agricultural Engineering*. Editorial Office of Transactions of the Chinese Society of Agricultural Engineering, 26(7), pp. 181–185.

Masek, J. (2001) 'Stability of boreal forest stands during recent climate change: evidence from Landsat satellite imagery', *Journal of biogeography*, 28(8), pp. 967–976. doi: 10.1046/j.1365-2699.2001.00612.x.

Mathisen, I. E. *et al.* (2013) 'Fifty years of tree line change in the Khibiny Mountains, Russia: advantages of combined remote sensing and dendroecological approaches', *Applied Vegetation Science*, 17(1), pp. 6–16. doi: 10.1111/avsc.12038.

Mayaux, P. and Lambin, E. F. (1997) 'Tropical forest area measured from global land-cover classifications: Inverse calibration models based on spatial textures', *Remote Sensing of Environment*, 59(96), pp. 29–43. doi: 10.1016/S0034-4257(96)00077-6.

Montesano, P. M. *et al.* (2009) 'MODIS tree cover validation for the circumpolar taiga–tundra transition zone', *Remote Sensing of Environment*, 113(10), pp. 2130–2141. doi: 10.1016/j.rse.2009.05.021.

Montesano, P. M. *et al.* (2014) 'The uncertainty of plot-scale forest height estimates from complementary spaceborne observations in the taiga-tundra ecotone', *Remote Sensing*, 6(10), pp. 10070–10088. doi: 10.3390/rs61010070.

Montesano, P. M. *et al.* (2016) 'Spaceborne potential for examining taiga-tundra ecotone form and vulnerability', *Biogeosciences Discussions*, (January), pp. 1–36. doi: 10.5194/bg-2015-575.

Otukei, J. R., Blaschke, T. and Collins, M. (2012) 'A Decision Tree Approach for Identifying the Optimum Window Size for Extracting Texture Features from TerraSAR-X Data', *GI Forum 2012: Geovizualisation, Society and Learning*, pp. 466–474.

Payette, S., Fortin, M.-J. and Gamache, I. (2001) 'The Subarctic Forest–Tundra: The Structure of a Biome in a Changing Climate', *BioScience*. Oxford University Press, 51(9), pp. 709–718.

Ranson, K. J., Montesano, P. M. and Nelson, R. (2011) 'Object-based mapping of the circumpolar taiga–tundra ecotone with MODIS tree cover', *Remote Sensing of Environment*. Elsevier B.V., 115(12), pp. 3670–3680. doi: 10.1016/j.rse.2011.09.006.

Rees, G. *et al.* (2002) 'How Can the Dynamics of the Tundra-Taiga Boundary Be Remotely Monitored?', *Ambio Special Report*. Springer on behalf of Royal Swedish Academy of Sciences, (12), pp. 56–62. doi: 10.2307/25094576.

Satterthwaite, F. E. (1946) 'An approximate distribution of estimates of variance components', *Biometrics bulletin*. JSTOR, 2(6), pp. 110–114.

Settele, J. *et al.* (2014) ‘Terrestrial and inland water systems’, in Field, C. B. et al. (eds) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 271–359.

Sexton, J. O. *et al.* (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), pp. 427–448. doi: 10.1080/17538947.2013.786146.

Stow, D. A. *et al.* (2004) ‘Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems’, *Remote Sensing of Environment*, 89, pp. 281–308. doi: 10.1016/j.rse.2003.10.018.

Sveinbjörnsson, B., Hofgaard, A. and Lloyd, A. (2002) ‘Natural causes of the tundra-taiga boundary.’, *Ambio Special Report*, (12), pp. 23–29. doi: 10.2307/25094572.

Timoney, K. P. *et al.* (1992) ‘The high subarctic forest-tundra of northwestern Canada: position, width, and vegetation gradients in relation to climate’, *Arctic*, 45(1), pp. 1–9.

Timoney, K. P. *et al.* (2018) ‘Tree cover response to climate change in the forest-tundra of north-central Canada : fire-driven decline , not northward advance’, *Écoscience*. Taylor & Francis, 26(2), pp. 133–148. doi: 10.1080/11956860.2018.1532868.

Tottrup, C. *et al.* (2007) ‘Mapping fractional forest cover across the highlands of mainland Southeast Asia using MODIS data and regression tree modelling’, *International Journal of Remote Sensing*. Taylor & Francis, 28(1), pp. 23–46.

Townsend, J. *et al.* (2011) *User Guide for the MODIS Vegetation Continuous Fields product Collection 5 version 1*.

Vlassova, T. K. (2002) ‘Human impacts on the tundra-taiga zone dynamics: the case of the Russian lesotundra’, *Ambio Special Report*, (12), pp. 30–36. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12374057.

Walker, D. a *et al.* (2005) ‘The Circumpolar Arctic vegetation map’, *Journal of Vegetation Science*, (16), pp. 267–282. doi: 10.1658/1100-9233(2005)016[0267:TCAVM]2.0.CO;2.

Walther, C. et al. (2019) ‘Modelling the Arctic taiga-tundra ecotone using ALOS PALSAR and optical earth observation data’, *Int J Appl Earth Obs Geoinformation*, (81), pp. 196-206. doi: 10.1016/j.jag.2019.05.008.

Weszka, J. S., Dyer, C. R. and Rosenfeld, A. (1976) ‘Comparative Study of Texture Measures for Terrain Classification.’, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-6(4), pp. 269–285. doi: 10.1109/TSMC.1976.5408777.

White, M. a., Shaw, J. D. and Ramsey, R. D. (2005) ‘Accuracy assessment of the vegetation continuous field tree cover product using 3954 ground plots in the south-western USA’, *International Journal of Remote Sensing*, 26(July 2014), pp. 2699–2704. doi: 10.1080/01431160500080626.

Wilmking, M. *et al.* (2012) ‘Continuously missing outer rings in woody plants at their distributional margins’, *Dendrochronologia*. Elsevier GmbH., 30(3), pp. 213–222. doi: 10.1016/j.dendro.2011.10.001.

Wood, E. M. *et al.* (2012) ‘Image texture as a remotely sensed measure of vegetation structure’, *Remote Sensing of Environment*. Elsevier Inc., 121, pp. 516–526. doi: 10.1016/j.rse.2012.01.003.

Wu, J. and Peng, D. (2012) ‘Vegetation classification technology of hyperspectral remote sensing based on spatial information’, *Transactions of the Chinese Society of Agricultural Engineering*. Editorial Office of Transactions of the Chinese Society of Agricultural Engineering, 28(5), pp. 150–153.

Yang, W. *et al.* (2009) ‘Supervised land-cover classification of TerraSAR-X imagery over urban areas using extremely randomized clustering forests’, *2009 Joint Urban Remote Sensing Event*, pp. 1–6. doi: 10.1109/URS.2009.5137603.

Zhu, Z. and Evans, D. L. (1994) ‘US forest types and predicted percent forest cover from AVHRR data’, *Photogrammetric engineering and remote sensing*, 60(5).

List of figure captions

Figure 1. Locations of additional study areas in different sub-regions (red).

Figure 2. (a) Location of the study area; (b) Google Earth coverage of the study area; (c) Landsat VCF data and reference data points used in this study. The data points from Wilmking et al. (2012) study are of the FTE1 class.

Figure 3. Processing steps for FTE delineation.

Figure 4. Optimal threshold pairs derived from the Central Eurasia study area. Corresponding maximum accuracy and kappa coefficient also displayed (T1acc: first threshold based on classification accuracy; T2acc: second threshold based on classification accuracy; T1kappa: first threshold based on kappa coefficient; T2kappa: first threshold based on kappa coefficient). Bars representing the numbers of points are placed on every 5 bin number from 0 to 100 in the order of forest (black), tundra (white) and FTE (grey), from left to right.

Figure 5. FTE delineated from VCF thresholding in the study area (a) using the 5-20 threshold pair; (b) using the derived 5-10 threshold pair. FTE delineated from VCF thresholding in a subset (red rectangle) of the study area: (c) using the 5-20 threshold pair; (d) using the derived 5-10 threshold pair. (e) Google Earth coverage over the subset of the study area.

Figure 6. Transformed divergence between class pairs in different window sizes.

Figure 7. FTE delineated using supervised classification based on texture analysis: (a) using the 5-20 threshold pair; (b) using the 5-10 threshold pair.

Figure 8. Comparison between (a) FTE areas derived from texture-based classification using 5-20 threshold pair and (b) 5-10 threshold pair, (c) FTE derived by Ranson et al., (the FTE segment

designated by letter ‘A’ is an example of tundra areas being recognised as FTE in this product), and (d) Google Earth coverage over a subset of the study area whose location is shown by the red rectangle in the left-hand panel.

Figure 9. Final derived FTE area in the Central Eurasia study area.

Figure 10. Application of the FTE delineation on additional study sites. (From left to right) Google Earth coverage; Landsat VCF and reference data points; Landsat VCF and the Ranson et al. (2011) FTE; classified image using our texture-based classification.

List of table captions

Table 1. GLCM-based texture measures used in this study.

Table 2. Classification accuracy and kappa coefficient of FTE delineation using VCF thresholding and texture-based classification based on the 5-20 and 5-10 threshold pairs, and the Ranson et al. (2011) FTE in the study area.

Table 3. Parameterisation and classification accuracies of FTE delineation for the additional study areas. Those for the Central Eurasia study area are also listed as reference.