

Understanding the Spatially Explicit Distribution of Regional Tree Species Using Multi-Seasonal Sentinel-1&2 Imagery within Google Earth Engine

Bo Xie

AIR: Chinese Academy of Sciences Aerospace Information Research Institute

Chunxiang Cao (✉ caocx@aircas.ac.cn)

AIR: Chinese Academy of Sciences Aerospace Information Research Institute

Min Xu

AIR: Chinese Academy of Sciences Aerospace Information Research Institute

Barjeece Bashir

AIR: Chinese Academy of Sciences Aerospace Information Research Institute

Yiyu Chen

AIR: Chinese Academy of Sciences Aerospace Information Research Institute

Robert S. Duerler

AIR: Chinese Academy of Sciences Aerospace Information Research Institute

Kaimin Wang

AIR: Chinese Academy of Sciences Aerospace Information Research Institute

Research

Keywords: Multi-sensor, Tree species mapping, Large areas, Cloud-computing, Machine learning

Posted Date: February 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-245409/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Accurate information on tree species is much in demand for forestry management and further investigations on biodiversity and forest ecosystem services. Over regional or large areas, discriminating tree species at high resolution is deemed challenging by lack of representative features and computational power.

Methods A novel methodology to delineate the explicit spatial distribution of dominated six tree species (Pinus, Quercus, Betula, Populus, Larch, and Apricot) and one residual class using the analysis-ready large volume multi-sensor imagery within Google Earth Engine (GEE) platform is demonstrated and used to map a 10 m classification with detail analysis of spatial pattern for an area covering over 90,000 km² between 41° N and 45° N. Random Forest (RF) algorithm built into GEE was used for tree species mapping, together with the multi-temporal features extracted from Sentinel-1/2 and topographic imagery data. The composition of tree species in natural forests and plantations in city and county-level were performed in detail afterwards.

Results The proposed model achieved a reliable overall agreement (77.5%, 0.71 kappa), and the detailed analysis on the spatial distributing of targeted species indicated that the plantations (Pinus, Populus, Larch, and Apricot) outnumber natural forests (Quercus and Betula) by 6%, and they were mainly grown in the northern and southern regions, respectively. Moreover, Arhorchin had the largest total forest area of over 4,500 km², while Hexingten and Aohan ranked first in natural forest and plantation area, and the class proportion of the number of tree species in Karqin and Ningcheng was more balanced.

Conclusions It is our belief that combined multi-source information of the machine learning algorithm within cloud platforms is beneficial to map a reliable spatial tree species over large areas on a fine scale. High-resolution tree species information based on online tools could be more easily considered for practical forestry management and further studies on forest ecosystems.

Introduction

A clear understanding of spatial distribution of tree species is crucial for afforestation decision-making, carbon cycle estimation, biodiversity assessment (Fassnacht et al., 2016; Grabska et al., 2019), as well as further interest in analyzing tree-environment interactions (Blaschke, 2010; Turner et al., 2003).

Conventional forestry inventories, though time-consuming and inefficient, have been a long history. Remote sensing technology has greatly improved efficiency because it is able to capture forest type composition and forest structure information in a larger and inaccessible areas through multi-band and multi-mode sensors in comparison to conventional field works (Grabska et al., 2019), which brings possible solutions to the challenging but promising topic that tree species identification based on remotely sensed data.

Remotely sensed imagery data with very high spatial resolution (better than 5m or even sub-meter) have been used for tree species discrimination (Adelabu et al., 2013; Karlson et al., 2016; Wangda et al., 2019;

Yang et al., 2014), because they can assist in reducing the impact of the occurrence of mixed pixels on tree species classification, which is an inherent characteristic determined by remote sensing imaging mechanism, especially in heterogeneous forests (Griffiths et al., 2014; Xie et al., 2008). However, the operational applications of high-resolution data are limited due to the high cost. The high similarity of features captured by sensors among trees is another challenge that has been tried to address in many studies using hyperspectral sensors. They provide narrow and contiguous spectral curves capable of characterizing small differences in the biochemical components of vegetation that cannot be captured by multispectral sensors (Ghosh et al., 2014; Van Coillie et al., 2015), which has been demonstrated in many studies (Ballanti et al., 2016; Dudley et al., 2015). However, the processing process of hyperspectral image is time-consuming due to its large volume (Shukla and Kot, 2016), and it is very delicate for optimal bands filtering from many high-correlation bands of hyperspectral image (George et al., 2014). With the rapid development of airborne laser scanning (ALS) and unmanned aerial vehicle (UAV), their light detection and ranging (LiDAR) data are also used in conjunction with high-resolution multi-spectral and hyper-spectral images for tree species mapping and individual trees identification studies, leading to a high accuracy achievement (Burai et al., 2019; Dalponte et al., 2019; Sasaki et al., 2012; Shi et al., 2018; Wang et al., 2019; Wu and Zhang, 2020).

Although these data mentioned above possess good potential for species identification, they are practically restricted owing to limited availability. Instead, the cost-free multispectral Landsat and Sentinel-2 images appear to be the best solution to vegetation studies, especially to large regional areas, in which distribution range of tree species is the preferred not individual trees extraction (Grabska et al., 2019). Landsat data have been useful in many typical studies relevant to vegetation mapping (Adams and Matthews, 2018; Chowdhury et al., 2021; Hobbs et al., 1989; Permana et al., 1998). Increasingly, scholars begin to focus on forest type classification without detailed in tree species composition in collaboration with Landsat and other data (Immitzer et al., 2018; Townshend et al., 2012). Moreover, the single-specie (e.g., mangrove, bamboo, and eucalyptus) studies were also undertaken based on time-series Landsat imagery (Khairuddin et al., 2016; Long and Giri, 2011; Soares and Hoffer, 1995; Zhao et al., 2018). Since the launch of Sentinel-2 mission in 2015, as another freely available data, it has brought new opportunities for fine monitoring of vegetation owing to its unique red-edge band and excellent spatial and temporal resolution (Grabska et al., 2019). The high potential of red-edge and shortwave infrared (SWIR) bands of Sentinel-2 data for vegetation mapping was confirmed by Immitzer et al. in 2016 when they assessed the capabilities of pre-operational (August 2015) Sentinel-2 data for mapping tree species in Austria (Immitzer et al., 2016). In addition, adequate studies have shown that, consistent with Landsat data (Mickelson et al., 1998), the time-series metrics of multispectral Sentinel-2 are crucial for tree species classification (Misra et al., 2020). A case study realized the identification of complex tree species composition in mountains areas, and proved that using time-series Sentinel-2 features instead of single-date images can improve accuracy by 5-10% (Grabska et al., 2019). Active imaging radar, because of its all-weather and all-day working advantages, has become one of the important data for forest monitoring. However, most former studies focused on the discrimination of broad forest types using (synthetic aperture radar) SAR omitting the species level (Fabian et al., 2016).

Regarding the methodology used for tree species identification, the object-based method is generally used for research that uses only extremely high spatial resolution images (Agarwal et al., 2013; Wang et al., 2018) or collaborates with other remotely sensed data (Sasaki et al., 2012). Furthermore, the multi-temporal approaches are indispensable to delineate tree species with multispectral images (Hoscilo and Lewandowska, 2019; Mickelson et al., 1998). Machine learning algorithms have been commonly applied for tree species classification because of its convenience in coordinating multi-source features, and relevant studies have given evidence that among machine learning models, random forest (RF) and support vector machines (SVM) outperformed others (Wessel et al., 2018). Therefore, the two models have been widely used to map tree species together with imagery data of multi-sensor (Ballanti et al., 2016; Lv and Ma, 2016; Raczko and Zagajewski, 2017). The accuracy of the machine learning model is very dependent on feature engineering affected by prior knowledge, while the deep learning model can directly implement end-to-end image classification based on the original image (Xi et al., 2020). Recently deep learning algorithms have been applied in the studies on tree species classification based on high spatial resolution and hyperspectral images due to their strong capability of feature mining (dos Santos et al., 2019; Fricker et al., 2019; Hartling et al., 2019; Nezami et al., 2020; Schiefer et al., 2020; Zhang et al., 2020). However, deep learning models can only be driven by a large number of labeled samples and great computational power, which is the major obstacle to its wide usage. (Ball et al., 2018). The advent of a wide variety of tools facilitated the geospatial data processing of large-scale with the development of high-performance computing systems (Gorelick et al., 2017), among which GEE has been widely used in vegetation monitoring by remote sensing in large and even global regions with its easy-to-use advantages (Mutanga and Kumar, 2019; Srinet et al., 2020; Venkatappa et al., 2020; Wu, 2020; Yang et al., 2019).

All the previous studies provided different solutions to tree species mapping, but they focused on the small-scale without detailed analysis on the spatial pattern of tree species composition. In present study, focusing on the remaining problem of how to reliably delineate and analyze spatial distribution for complex tree species composition at regional scale, we built on previous studies using RF algorithm, but also proposed a promising methodology for tree species mapping within Google Earth Engine (GEE) cloud-computing platform that is as simple as possible and can be scaled for processing larger datasets. The objectives were: (i) to map the seven tree species with high resolution (10 m); and (ii) to figure out the tree species composition of large regional area on the basis of at different scales within GEE.

Materials And Methods

Study area

Chifeng city, one prefecture-level cities, centered at 119°22'58.38"E, 60°35'07.2"N underlies over 90,000 km² of ten counties in the southeast of Inner Mongolia in northeastern China (Fig. 1). The topography is complex and diverse with only small mountain flats and alluvial plains along the river, the abundant forest resources and diversity tree species are deciduous-dominated by *Quercus*, *Betula*, *Populus*, *Larch*, and *Apricot* trees, and the evergreen trees are mainly *Pinus*.

Field measurements

The forest resources inventory sample plots were used as the ground-truth of the seven tree species in this study, and a total of 342 square plots (60*10 m) surveyed in 2018 were distributed throughout the area (Fig. 1). These plots have a high positioning accuracy because the reset accuracy of their position once every five year is required to be above 98%. Therefore, they are enabled to function as a reliable modeling and validation data. Additional measurements of natural forests and plantations polygons from forestry inventory were also used as mask layers.

Satellite data in Google Earth Engine

Google Earth Engine (GEE) is a big data cloud platform with high performance computing where a petabyte analysis-ready dataset is freely available and the programming interface is also quite access-friendly (Gorelick et al., 2017; Kumar and Mutanga, 2018; Mutanga and Kumar, 2019). The study was conducted based on the Sentinel-1A (S1) Ground Range Detected (GRD) scenes, Sentinel-2 (S2) surface reflectance (SR) images, and the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) dataset of GEE platform. The S1 GRD data was processed by thermal noise removal, radiation correction, and terrain correction, and the 20 m dual band VV+VH and HH+HV of Interferometric Wide Swath (IW) mode was selected for further processing. For the S2 SR image, geometric, radiation, and atmospheric correction were performed, and the final SR image composites consisted of 10 bands with two spatial resolutions, including 10 m visible and 20 m infrared and red-edge bands. Additionally, cloud mask (QA60) band was also used to help mask cloud in image scenes, leaving only pixels cloudless with good quality. To reflect the temporal characteristics of different tree species in the four seasons of summer, autumn and winter, the S1 and S2 images of four months (March, June, September, and December) of 2019 were selected. The 30 m DEM data was used to depict the varied topographic feature, generating aspect and slope variables. Table 1 provided details regarding the parameter of the satellite data in this study.

Table 1 Details of adopted satellite dataset in Google Earth Engine.

Satellite image	Year	Month	Bands	Spatial resolution (m)
Sentinel-1A GRD Images	2019	3, 6, 9, 12	VV+VH	20
			HH+HV	
Sentinel-2 SR Images		3, 6, 9, 12	Blue	10
			Green	
			Red	
			Red Edge 1	20
			Red Edge 2	
			Red Edge 3	
			NIR	
			Red Edge 4	
SWIR 1	30			
SWIR 2				
SRTM DEM				30

Methods

Tree species classification overview

Our goal was to leverage the powerful computing ability of GEE platform for producing a high-resolution typical tree species distribution map across Chifeng and figure out the spatial pattern of each tree species. We proposed and implemented the novel methodology within GEE cloud-computing platform, which was designed as four processes including field sample plots addressing, multi-temporal feature mining, RF model optimizing, and classification and analyzing. Fig. 2 was an overview of our workflow, which will be described in detail in subsequent sections.

Addressing field plots

These plots were not built-in data of the GEE platform, and they were made up of 307 modeling and independent 35 validation plots, of which the spatial distribution was provided in Fig. 1. Taking the differentiation of regional scales into account, modeling and verification points are reasonably distributed throughout the study area to better represent the local characteristics of each tree species and reduce the systematic errors in modeling and accuracy assessment. All pixels covered by polygon, not only the geometric center pixels, were used in subsequent analysis.

Mining multi-temporal features from Sentinel-1/2 imagery

The multi-temporal features were all derived from the S1 and S2 products accessed to GEE on line. We imported the S1 and S2 datasets from the data catalog of GEE and filtered out all the images covering the entire study area in March, June, September, and December according to the image acquisition date. Fig. 3 provided the visual characteristic changes reflected by S2 surface reflectance images of four seasons, from which the serious influence of cloud occlusion on the application of optical satellite images was also apparently told. Therefore, the cloud-mask operation was performed on each S2 scene afterwards, while GRD images of S1 were further screened, leaving only image scenes of IW mode.

When all S1 and S2 images of four months were analysis-ready for further procedures, we made full use of the advantages of the GEE platform to integrate the three-dimensional (time, space, and spectrum) features mined from these multitemporal images. Specifically, 23 metrics derived from S1 and S2 dataset were divided into seven categories, of which 16 were from SR images, and the others from S1 GRD images (Fig. 4). For the S2 spectral index, we calculated eight commonly used indices incorporating the visible, near-infrared, and red-edge bands, including Infrared Percentage Vegetation Index (IPVI) (Crippen, 1990), Transformed Normalized Difference Vegetation Index (TNDVI) (Senseman et al., 1996), Green Normalized Difference Vegetation Index (GNDVI) (Gitelson et al., 1996), the second Brightness Index (BI2) (Escadafal et al., 1989), Meris Terrestrial Chlorophyll Index (MTCI) (Dash and Curran, 2004), Red-Edge Inflection Point Index (REIP) (Guyot et al., 1988), Inverted Red-Edge Chlorophyll Index (IRECI) (Clevers et al., 2000; Guyot et al., 1988), normalized difference vegetation index (NDVI), and enhanced vegetation index (EVI) (Huete et al., 2002). To take full use of the high spatial resolution, Grey-Level Co-occurrence Matrix (GLCM) was performed on the NIR bands with highest resolution (10 m) and sensitivity to vegetation to generate four texture features (the second moment, contrast, homogeneity, and entropy) of

S1 scenes. Furthermore, we addressed the like-polarization (VV/HH) and cross-polarization (VH/HV), yielding four radar indices (division, difference, amplitude, and normalization). Finally, we applied linear regression on the EVI and VH variables to capture the gradient of spectral and radar back scatter over time in one month as well.

Table 2 summarized all the monthly variables, of which except for the four characteristics of slope, aspect, EVI_scale, and VH_scale, we used the 20th and 80th percentiles of the remaining monthly characteristics instead for subsequent analysis to reduce sensitivity to noise such as residual cloud and shadows as well as taking full use of all the images with same scene in one month (Zhang and Roy, 2017), yielding 191 features of March, June, September, and December.

Table 2 Detailed description of all the features generated from the satellite images in GEE and used for Random forest classification.

Feature	Short name	Formula	Source
First shortwave infrared band	SWIR1		Sentinel-2
Second shortwave infrared band	SWIR2		Sentinel-2
Normalized difference vegetation index	NDVI	$(nir - red) / (nir + red)$	Sentinel-2
Enhanced vegetation index	EVI	$2.5 * ((nir - red) / (nir + 6 * red - 7.5 * blue + 1))$	Sentinel-2
Infrared Percentage Vegetation Index	IPVI	$0.5 * ((nir - red) / (nir + red) + 1)$	Sentinel-2
Transformed Normalized Difference Vegetation Index	TNDVI	$\sqrt{((nir - red) / (nir + red) + 0.5)}$	Sentinel-2
Green Normalized Difference Vegetation Index	GNDVI	$(nir - green) / (nir + green)$	Sentinel-2
Second Brightness Index	BI2	$\sqrt{((red * red) + (green * green) + (nir * nir))} / 3$	Sentinel-2
Meris Terrestrial Chlorophyll Index	MTCI	$(red_{oe2} - red_{oe1}) / (red_{oe1} - red)$	Sentinel-2
Red-Edge Inflection Point Index	REIP	$705 + 35 * ((red + red_{ge3}) / 2 - red_{ge1}) / (red_{ge2} - red_{ge1})$	Sentinel-2
Inverted Red-Edge Chlorophyll Index	IRECI		Sentinel-2
NIR: Angular Second Moment	asm		Sentinel-2
NIR: Contrast	contrast		Sentinel-2
NIR: Inverse Difference Moment	idm		Sentinel-2
NIR: Entropy	ent		Sentinel-2
Gradient of EVI	EVI_grad	$(EVI - b) / t$	Sentinel-2
Cross polarization band	VH		Sentine-1
Like polarization band	VV		Sentine-1
Back scatter division	div	VH / VV	Sentine-1
Back scatter difference	diff	$VH - VV$	Sentine-1
Back scatter amplitude	amp	$\sqrt{(VH * VH) + (VV * VV)}$	Sentine-1
Back scatter normalization	norm	$(VH - VV) / (VH + VV)$	Sentine-1
Gradient of VH	VH_grad	$(VH - b) / t$	Sentine-1
Terrain Slope	Slope		SRTM DEM
Terrain Aspect	Aspect		SRTM DEM

(t and b indicated the intercept of linear regression and acquire time of each image scene)

Additional ancillary features

Topographic factors (slope and aspect) closely related to the vegetation distribution were derived from SRTM DEM data, which was widely used in forestry research (Xie et al., 2020). We used the built-in terrain algorithm of GEE to calculate slope and aspect, which were then reclassified to convert the two continuous variables to categorical variables according to the technical regulations for continuous inventory of forest resources. Table 3 listed the criteria of slope and aspect reclassification. The result of one-hot encoding of the two reclassified variables was used as the final topographic feature.

Table 3 Summary of slope & aspect reclassification criteria.

Slope		Aspect	
Value (°)	Class	Value (°)	Class
5	I		Non-directional
5-14	II	338 - 22	North
15-24	III	292.5 - 337.5	Northwest
25-34	IV	23 - 67	Northeast
35-44	V	68 - 112	East
45	VI	113- 157	Southeast
		158 - 202	South
		203- 247	Southwest
		248- 292	West
		293-337	Northwest

Optimizing Random Forest classifier

Machine learning algorithm tuning is of great importance of obtaining a stable and high-performance classifier. Here, the Random Forest algorithm built-in GEE platform named “smileRandomForest” (RF) was leveraged for capturing regional tree species distribution. Tree-based RF model was one of the most commonly used typical bagging learning algorithms (Lv and Ma, 2016; Raczko and Zagajewski, 2017). RF builds multiple decision trees and merges them to get a more accurate and stable model. The prediction result for the sample is the prediction result of each tree in the forest for the sample, and then the final result is selected from these prediction results by voting. The critical hyper-parameter “numberOfTrees” of RF classifier was supposed to be optimized by balancing model complexity and model generalization accuracy. The learning curves was used to characterize the generalization capability of RF model with “numberOfTrees” increasing from 1 to 100 (Fig. 5). Moreover, RF built-in attribute of out-of-bag score (oob) was used in the model tuning process, which takes advantage of the unused samples during the random decision tree generation process to evaluate the accuracy of each tree, yielding the quantified performance of RF algorithms by taking the average accuracy value of all trees. The smaller the difference between kappa detailed in Eq. (1) and oob learning curves, the better the robustness of the model given a specific parameter value.

Classification, accuracy assessment, and zonal statistics

To map the spatial distribution of the six targeted tree species (Pinus, Quercus, Betula, Populus, Larch, and Apricot) and one residual class, a bagging learning model was carried out using the optimal RF classifier based on multi-temporal features within GEE cloud-computing platform. Quality evaluation of the classification result is always of indispensability in remote sensing since it proves how well the used classifier is capable of identifying the desired objects from certain image. Here, we applied two accuracy measures on the classification validation and additional one metric on model optimization. As a first evaluation, overall accuracy (OA), expressed as Eq. (1), gives quantified evidence of the proportion of all correctly classified reference samples, that is, the class assignment of the classification agrees with the reference classification. Eq. (2) represented another accuracy measure that is Cohen’s kappa coefficient (kappa), which measures the consistency of assigned and reference classification assuming they are

equally reliable and independent. To quantitatively clarify the tree species composition in whole areas and local districts, on the basis of the spatial distribution area, the zonal statistics of each tree species of the whole area and of local regions was respectively carried out, leading to assessment on the natural forest and plantation forest areas.

$$OA = \sum_{i=1}^n x_{ii} / N \quad (1)$$

$$\text{kappa} = \frac{N \sum_{i=1}^n x_{ii} - \sum_{i=1}^n (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^n (x_{i+} x_{+i})} \quad (2)$$

where n is the number of reference class, x_{ii} is the number of the correctly classified pixels of the i -th category, N is the total number of the reference pixels, and x_{i+} and x_{+i} represents the total number of the i -th category of reference classification and assignment of classification respectively.

Results

Multi-source feature composition

The feature dataset composed of 176 multi-temporal features (e.g., original image bands, spectral/radar indices, textures and gradient) and additional two one-hot encoded topographic features (slope and aspect). The multi-temporal features were listed as Fig. 6, of which the ordinate and abscissa individually represented the feature name and feature score, which was computed by the importance attribute built into RF classifier within GEE and was proportional to the contribution of the corresponding feature to RF model. Additional two terrain features were represented as Fig. 7 and Fig. 8 where the grayscale pictures on the left were the unprocessed continuous numeric features extracted from terrain data, while the classified color images were their corresponding reclassification results according to the criteria in Table 3.

Optimization of Random Forest model

Fig. 5 implied the accuracy of RF classifier generally increased with the increase of parameter number of trees. The growth rate increased firstly and then decreased, and finally tends to be stable with small fluctuations. The two learning curves (kappa and out_bag_score) had the minimum difference (0.06), that was, RF model had the best generalization capability and robustness when the key hyper-

parameter numberOfTrees reached 71. Therefore, we built a random forest model with 71 trees as the optimal model for tree species classification.

Tree species classification and accuracy assessment

The spatial distribution mapping of objective tree species across Chifeng city was presented as Fig. 9, of which the spatial pattern clearly depicted that Pinus was mainly distributed in the southern four districts, while Quercus was concentrated in the northernmost part of Chifeng, additionally, Apricot and Populus trees were found in the central region as well as in the western and eastern regions, respectively, as for Betula, it was along the western and northwestern forest margins, and Larch was mainly distributed in the western and southwestern regions.

There were significant differences in the distribution range and tree species composition between natural forests and plantations (Fig. 10). The planted forests were distributed in a wider area compared to natural forests, and the former was distributed almost throughout the area but mostly in the south, while the latter was dominated in the north. In addition, the planted trees in this region were composed of Pinus, Populus and Larch, while the natural tree species mainly consisted of Quercus, Betula, and Apricot.

The classification accuracy assessment was carried out to yield the confusion matrix as Fig. 11, from which the quantified accuracy of the classification results was calculated according to the accuracy evaluation metrics of OA and kappa that detailed in Section 2.4.6. The overall accuracy (OA) = 77.5% and kappa = 0.71 for the seven classes in Chifeng based on the RF classifier with multi-temporal features within GEE.

Quantitative analysis on the tree species distribution

The tree species area results (Fig. 12) of the entire region revealed that Natural forests and plantations separately accounted for 47% and 53% of the total forest area. Moreover, Apricot was roughly equal in proportion to natural and cultivated trees, and it had the largest distribution area of more than 10,000 km² followed by the Populus trees covering an area over 8,000 km². Almost all Pinus (84%) and Populus (80%) were planted trees, while Quercus (91%) and Betula (94%) were natural forest species.

The statistical results of tree species composition and spatial area at county-level was summarized as Fig. 13. It quantitatively revealed the regional differences of forest resources among districts in Chifeng. In terms of total forest resources, Arhorchin and Aohan ranked first and second with total forest area of more than 4800 and 4100 km², respectively, and the proportion of natural forest and artificial forest in the former was relatively balanced, while the latter was mainly artificial forest. In addition, Linxi and Ningcheng dominated by plantations, having the smallest total forest area, both were less than 2,000 km². From the perspective of diversity of tree species, the distribution area of these typical tree species in Kalaqin Banner and Ningcheng was more balanced, both with only slightly more Pinus, of which area were separately about 280 and 480 km². Pinus, Populus, and Apricot trees were dominant in multiple

regions, and the regions of Pinus were Karqin and Ningcheng, of Populus were Aohan, Arhorchin, and Ongniud, and of Apricot were Bahrain left & right, Hexingten, Linxi, and the municipal district.

Discussion

Tree species discrimination by use of remote sensing technique is a quite challenging topic due to the mixed pixels and low separability among trees (Fabian et al., 2016). There are numerous studies have been dedicated to tree species mapping using remotely sensed data, but they mainly focused on a relatively small areas, of which the experiences are not representative enough to be applied to assist in forestry inventories, environmental monitoring, and carbon cycle estimation, which require works in a large area (Fassnacht et al., 2016). While greater computing capacity to handle a large volume satellite images is a prerequisite for large geographic regions, while it is generally not available locally. This study explored the use of non-parametric RF classifier built into GEE cloud computing platform to classify the dominant tree species over a regional area of more than 90,000 km² to assess the potential of GEE in identification of forest fine categories over large areas.

The satellite image features were extracted from the analysis-ready Sentinel-1/2 and SRTM images within GEE by multi-temporal characteristics integration, and the multi-dimensional indicators, such as radar back scatter, spectral index, gradient, and topography were considered for single-phase images. The 20th and 80th percentiles used to extracted from these indicators of same month and scene can avoid the noise effect of the maximum and minimum value of the images (Zhang and Roy, 2017). Additionally, the labeled samples, however, are also of importance to regional forest type discrimination, especially using non-parametric models. The 307 ground-based polygons provided 2,272 training samples by collecting all the pixels within their respective coverage.

The results detailed in Figs.9 and 10 are representative of the spatial pattern of forest across the study area, in which most of the natural forests and plantations had their own concentrated areas, and together with the composition of different dominant tree species, indicating the obvious the heterogeneity of forest site conditions between the northern and southern regions. Clearly, however, the quantitative comparative analysis outcomes illustrated in Figs. 12 and 13 further proved that although it varies in the distribution of tree species in each county, on the whole, the composition of tree species in neighboring counties is more similar, that is, there is homogeneity of class proportion and forest site among local districts.

The classification results of the dominant species was validated by independent samples, which is a recommended way to assess accuracy (Olofsson et al., 2014), and the distribution range of seven targeted forest classes was delineated with an acceptable accuracy ($\kappa=0.71$, OA=77.5%) across the study area, which was comparable to the existing related studies, where the object-based methods was used together with UAV or multi-temporal multi-spectral images, and an overall classification accuracy was achieved from 73%-82% (Apostol et al., 2020; Franklin and Ahmed, 2018; Liu et al., 2018; Michez et al., 2016). Furthermore, some of previous studies obtained better accuracy, but the number of targeted

category was relatively fewer, and they were very demanding in terms of spatial resolution of the image, which is not applicable for studies on large areas (Fabian et al., 2016; Grabska et al., 2019).

Further work was suggested for two aspects: (i) monitoring the growth status of different tree species, and (ii) particularly analyzing the driving factors of their spatial distribution patterns to figure out the interactions of trees and environment.

Conclusion

A novel methodology for regional tree species identification using Sentinel-1/2 images in four months (March, June, September, and December) within GEE platform was demonstrated in an area over 90,000 km², yielding a 10-m spatial map of dominated six trees, including Pinus, Quercus, Betula, Populus, Larch, and Apricot. The reliable I agreement (OA=77.5%, kappa=0.71) also facilitated the subsequent analysis on global and local spatial patterns of different tree species.

There was a clear discordant distribution range between plantations and natural forests, the former was mainly distributed in the southern region, while the latter is mainly distributed in the northern region. The area of artificial forest was 6% more than that of natural forests. Moreover, Pinus and Populus were mainly artificially planted, while Quercus and Betula are the typical natural forest species in this region. Overall, Populus and Apricot trees had the most area in the study area, with an area over 8,657 km² and 11,099 km², respectively.

Through the comparative analysis of tree species distribution in different districts and counties, it was found that Karqin and Ningcheng county were the areas where Pinus and Larch trees were concentrated, and their category proportions were more balancing than any other regions. Besides, Populus were mainly planted in Aohan, Arhorchin, Bahrain right, Nicheng, Ongniud, and the municipal district, and Quercus trees mainly grown in Bahrain left and Karqin. Betula concentrated in Hexingten, while Apricot trees were relatively evenly distributed across the whole region.

The findings demonstrate the proposed cloud-computing workflow is capable of classifying forest type and analyzing spatial pattern over a regional-scale area and yielding a sufficiently satisfactory accuracy when using only freely-accessible Sentinel-1/2 imagery instead of more expensive high-resolution multispectral or hyperspectral data. We conclude that the novel design is well-suited to be applied on larger geographic areas to assist to helping forestry inventories.

Abbreviations

GEE: Google Earth Engine; S1: Sentinel-1A; GRD: Ground Range Detected; S2: Sentinel-2; SR: surface reflectance; SRTM: Shuttle Radar Topography Mission; DEM: digital elevation model; IW: Interferometric Wide; RF: Random Forest; GLCM: Grey-Level Co-occurrence Matrix; OA: overall accuracy.

Declarations

Authors' contributions

Bo Xie: Conceptualization, Investigation, Methodology, Software, Visualization, Writing - Original draft preparation, Validation. **Chunxiang Cao:** Conceptualization, Methodology, Data curation, Supervision, Funding acquisition, Writing - Reviewing and Editing. **Min Xu:** Conceptualization, Funding acquisition, Data curation. **Barjeece Bashir:** Investigation, Writing - Reviewing and Editing, Visualization. **Yiyu Chen:** Investigation, Software, Validation. **Robert S. Duerler:** Investigation, Writing - Reviewing and Editing. **Kaimin Wang:** Investigation, Software.

Acknowledgments

The authors are grateful to the National Forestry and Grassland Administration for sharing their forest data. Also, the authors acknowledge the contribution of Inner Mongolia Agricultural University for providing the necessary information about the study area. Special thanks go to Google and Copernicus for making available GEE and Sentinel imagery freely.

Funding

This study was jointly funded by the National Key Research and Development of China (NO. 2017YFD0600903) and the National Natural Science Foundation of China (No. 41971394).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing Interest

The authors declare that they have no conflict of interest in this study.

Author details

¹ State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100101, China;

² University of Chinese Academy of Sciences, Beijing, 100094, China.

References

- Adams BT, Matthews SN. Enhancing Forest and Shrubland Mapping in a Managed Forest Landscape with Landsat-LiDAR Data Fusion. *Natural Areas Journal* 2018; 38: 402-418.
- Adelabu S, Mutanga O, Adam E, Cho MA. Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. *Journal of Applied Remote Sensing* 2013; 7.
- Agarwal S, Vailshery LS, Jaganmohan M, Nagendra H. Mapping Urban Tree Species Using Very High Resolution Satellite Imagery: Comparing Pixel-Based and Object-Based Approaches. *Isprs International Journal of Geo-Information* 2013; 2: 220-236.
- Apostol B, Petrila M, Lorent A, Ciceu A, Gancz V, Badea O. Species discrimination and individual tree detection for predicting main dendrometric characteristics in mixed temperate forests by use of airborne laser scanning and ultra-high-resolution imagery. *Science of the Total Environment* 2020; 698.
- Ball JE, Anderson DT, Chan CS. Feature and Deep Learning in Remote Sensing Applications. *Journal of Applied Remote Sensing* 2018; 11.
- Ballanti L, Blesius L, Hines E, Kruse B. Tree Species Classification Using Hyperspectral Imagery: A Comparison of Two Classifiers. *Remote Sensing* 2016; 8.
- Blaschke T. Object based image analysis for remote sensing. *Isprs Journal of Photogrammetry and Remote Sensing* 2010; 65: 2-16.
- Burai P, Beko L, Lenart C, Tomor T, Kovacs Z, Ieee. INDIVIDUAL TREE SPECIES CLASSIFICATION USING AIRBORNE HYPERSPECTRAL IMAGERY AND LIDAR DATA. 2019 10th Workshop on Hyperspectral Imaging and Signal Processing - Evolution in Remote Sensing, 2019.
- Chowdhury S, Peddle DR, Wulder MA, Heckbert S, Shipman TC, Chao DK. Estimation of land-use/land-cover changes associated with energy footprints and other disturbance agents in the Upper Peace Region of Alberta Canada from 1985 to 2015 using Landsat data. *International Journal of Applied Earth Observation and Geoinformation* 2021; 94.
- Clevers JGPW, Jong H, Epema GF, Meer, Bakker W, Skidmore A, et al. The use of the MERIS standard band setting for deriving the red edge index. In: Proc. of the ISSSR International Symposium "Sensors and Systems for the new Millennium. Las Vegas, Oct 31 - Nov, 5, 1999. - [S.l.] : [s.n.], 2000 2000.
- Crippen RE. CALCULATING THE VEGETATION INDEX FASTER. *Remote Sensing of Environment* 1990; 34: 71-73.

Dalponte M, Frizzera L, Gianelle D. Individual tree crown delineation and tree species classification with hyperspectral and LiDAR data. *Peerj* 2019; 6.

Dash J, Curran PJ. The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing* 2004; 25: 5403-5413.

dos Santos AA, Marcato J, Araujo MS, Di Martini DR, Tetila EC, Siqueira HL, et al. Assessment of CNN-Based Methods for Individual Tree Detection on Images Captured by RGB Cameras Attached to UAVs. *Sensors* 2019; 19.

Dudley KL, Dennison PE, Roth KL, Roberts DA, Coates AR. A multi-temporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. *Remote Sensing of Environment* 2015; 167: 121-134.

Escadafal R, Girard MC, Courault D. MUNSELL SOIL COLOR AND SOIL REFLECTANCE IN THE VISIBLE SPECTRAL BANDS OF LANDSAT MSS AND TM DATA. *Remote Sensing of Environment* 1989; 27: 37-46.

Fabian, Ewald, Fassnacht, Hooman, Latifi, Krzysztof, et al. Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment* 2016.

Fassnacht FE, Latifi H, Stereńczak K, Modzelewska A, Lefsky M, Waser LT, et al. Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment* 2016; 186: 64-87.

Franklin SE, Ahmed OS. Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *International Journal of Remote Sensing* 2018; 39: 5236-5245.

Fricker GA, Ventura JD, Wolf JA, North MP, Davis FW, Franklin J. A Convolutional Neural Network Classifier Identifies Tree Species in Mixed-Conifer Forest from Hyperspectral Imagery. *Remote Sensing* 2019; 11.

George R, Padalia H, Kushwaha SPS. Forest tree species discrimination in western Himalaya using EO-1 Hyperion. *International Journal of Applied Earth Observation and Geoinformation* 2014; 28: 140-149.

Ghosh A, Fassnacht FE, Joshi PK, Koch B. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *International Journal of Applied Earth Observation and Geoinformation* 2014; 26: 49-63.

Gitelson AA, Kaufman YJ, Merzlyak MN. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment* 1996; 58: 289-298.

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 2017; 202: 18-27.

Grabska E, Hostert P, Pflugmacher D, Ostapowicz K. Forest Stand Species Mapping Using the Sentinel-2 Time Series. *Remote Sensing* 2019; 11: 1197.

Griffiths P, Kuemmerle T, Baumann M, Radeloff VC, Abrudan IV, Lieskovsky J, et al. Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment* 2014; 151: 72-88.

Guyot G, Frederic B, Major D. High spectral resolution: Determination of spectral shifts between the red and the near infrared. *International Archives of Photogrammetry and Remote Sensing* 1988; 11: 750-760.

Hartling S, Sagan V, Sidike P, Maimaitijiang M, Carron J. Urban Tree Species Classification Using a WorldView-2/3 and LiDAR Data Fusion Approach and Deep Learning. *Sensors* 2019; 19.

Hobbs RJ, Wallace JF, Campbell NA. CLASSIFICATION OF VEGETATION IN THE WESTERN AUSTRALIAN WHEATBELT USING LANDSAT MSS DATA. *Vegetatio* 1989; 80: 91-105.

Hoscilo A, Lewandowska A. Mapping Forest Type and Tree Species on a Regional Scale Using Multi-Temporal Sentinel-2 Data. *Remote Sensing* 2019; 11: 23.

Huete A, Didan K, Miura T, Rodriguez EP, Gao X, Ferreira LG. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 2002; 83: 195-213.

Immitzer M, Bock S, Einzmann K, Vuolo F, Pinnel N, Wallner A, et al. Fractional cover mapping of spruce and pine at 1 ha resolution combining very high and medium spatial resolution satellite imagery. *Remote Sensing of Environment* 2018; 204: 690-703.

Immitzer M, Vuolo F, Atzberger C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing* 2016; 8.

Karlson M, Ostwald M, Reese H, Bazie HR, Tankoano B. Assessing the potential of multi-seasonal WorldView-2 imagery for mapping West African agroforestry tree species. *International Journal of Applied Earth Observation and Geoinformation* 2016; 50: 80-88.

Khairuddin B, Yulianda F, Kusmana C, Yonvitner. Degradation mangrove by using Landsat 5 TM and Landsat 8 OLI image in Mempawah Regency, West Kalimantan Province year 1989-2014. In: Setiawan Y, Prasetyo LB, Siregar IZ, Effendi H, editors. 2nd International Symposium on Lapan-lpb Satellite. 33, 2016, pp. 460-464.

Kumar L, Mutanga O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sensing* 2018; 10.

Liu YA, Gong WS, Hu XY, Gong JY. Forest Type Identification with Random Forest Using Sentinel-1A, Sentinel-2A, Multi-Temporal Landsat-8 and DEM Data. *Remote Sensing* 2018; 10: 25.

- Long JB, Giri C. Mapping the Philippines' Mangrove Forests Using Landsat Imagery. *Sensors* 2011; 11: 2972-2981.
- Lv J, Ma T. Discrimination of tree species using random forests from the Chinese high-resolution remote sensing satellite GF-1. In: Liu W, Wang J, editors. *Hyperspectral Remote Sensing Applications and Environmental Monitoring and Safety Testing Technology*. 10156, 2016.
- Michez A, Piegay H, Lisein J, Claessens H, Lejeune P. Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environmental Monitoring and Assessment* 2016; 188.
- Mickelson JG, Civco DL, Silander JA. Delineating forest canopy species in the northeastern United States using multi-temporal TM imagery. *Photogrammetric Engineering and Remote Sensing* 1998; 64: 891-904.
- Misra G, Cawkwell F, Wingler A. Status of Phenological Research Using Sentinel-2 Data: A Review. *Remote Sensing* 2020; 12.
- Mutanga O, Kumar L. Google Earth Engine Applications. *Remote Sensing* 2019; 11.
- Nezami S, Khoramshahi E, Nevalainen O, Polonen I, Honkavaara E. Tree Species Classification of Drone Hyperspectral and RGB Imagery with Deep Learning Convolutional Neural Networks. *Remote Sensing* 2020; 12.
- Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment* 2014; 148: 42-57.
- Permana DS, Nakajima T, Yuasa T, Akatsuka T. A vegetation classification method with a spectral, spatial and temporal variability for Landsat/TM imagery. In: Tescher AG, editor. *Applications of Digital Image Processing Xxi*. 3460, 1998, pp. 834-843.
- Raczko E, Zagajewski B. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing* 2017; 50: 144-154.
- Sasaki T, Imanishi J, Ioki K, Morimoto Y, Kitada K. Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data. *Landscape and Ecological Engineering* 2012; 8: 157-171.
- Schiefer F, Kattenborn T, Frick A, Frey J, Schall P, Koch B, et al. Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks. *Isprs Journal of Photogrammetry and Remote Sensing* 2020; 170: 205-215.
- Senseman G, Tweddale S, Anderson A, Bagley C. Correlation of Land Condition Trend Analysis (LCTA) Rangeland Cover Measures to Satellite-Imagery-Derived Vegetation Indices. 1996: 33.

- Shi YF, Wang TJ, Skidmore AK, Heurich M. Important LiDAR metrics for discriminating forest tree species in Central Europe. *Isprs Journal of Photogrammetry and Remote Sensing* 2018; 137: 163-174.
- Shukla A, Kot R. An Overview of Hyperspectral Remote Sensing and its applications in various Disciplines. *IRA-International Journal of Applied Sciences (ISSN 2455-4499)* 2016; 5: 85.
- Soares VP, Hoffer RM. EUCALYPTUS FOREST CHANGE CLASSIFICATION USING MULTI-DATE LANDSAT TM DATA. Vol 2314, 1995.
- Srinet R, Nandy S, Padalia H, Ghosh S, Watham T, Patel NR, et al. Mapping plant functional types in Northwest Himalayan foothills of India using random forest algorithm in Google Earth Engine. *International Journal of Remote Sensing* 2020; 41: 1-14.
- Townshend JR, Masek JG, Huang CQ, Vermote EF, Gao F, Channan S, et al. Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. *International Journal of Digital Earth* 2012; 5: 373-397.
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M. Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution* 2003; 18: 306-314.
- Van Coillie FMB, Liao W, Kempeneers W, Vandekerkhove K, Gautama S, Philips W, et al. OPTIMIZED FEATURE FUSION OF LIDAR AND HYPERSPECTRAL DATA FOR TREE SPECIES MAPPING IN CLOSED FOREST CANOPIES. 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2015.
- Venkatappa M, Sasaki N, Anantsuksomsri S, Smith B. Applications of the Google Earth Engine and Phenology-Based Threshold Classification Method for Mapping Forest Cover and Carbon Stock Changes in Siem Reap Province, Cambodia. *Remote Sensing* 2020; 12.
- Wang DZ, Wan B, Qiu PH, Su YJ, Guo QH, Wu XC. Artificial Mangrove Species Mapping Using Pleiades-1: An Evaluation of Pixel-Based and Object-Based Classifications with Selected Machine Learning Algorithms. *Remote Sensing* 2018; 10.
- Wang KP, Wang TJ, Liu XH. A Review: Individual Tree Species Classification Using Integrated Airborne LiDAR and Optical Imagery with a Focus on the Urban Environment. *Forests* 2019; 10.
- Wangda P, Hussin YA, Bronsveld MC, Karna YK. Species stratification and upscaling of forest carbon estimates to landscape scale using GeoEye-1 image and lidar data in sub-tropical forests of Nepal. *International Journal of Remote Sensing* 2019; 40: 7941-7965.
- Wessel M, Brandmeier M, Tiede D. Evaluation of Different Machine Learning Algorithms for Scalable Classification of Tree Types and Tree Species Based on Sentinel-2 Data. *Remote Sensing* 2018; 10: 21.

Wu Q. geemap: A Python package for interactive mapping with Google Earth Engine. *Journal of Open Source Software* 2020; 5: 2305.

Wu YS, Zhang XL. Object-Based Tree Species Classification Using Airborne Hyperspectral Images and LiDAR Data. *Forests* 2020; 11.

Xi ZX, Hopkinson C, Rood SB, Peddle DR. See the forest and the trees: Effective machine and deep learning algorithms for wood filtering and tree species classification from terrestrial laser scanning. *Isprs Journal of Photogrammetry and Remote Sensing* 2020; 168: 1-16.

Xie B, Cao C, Xu M, Bashir B, Singh RP, Huang Z, et al. Regional Forest Volume Estimation by Expanding LiDAR Samples Using Multi-Sensor Satellite Data. *Remote Sensing* 2020; 12: 360.

Xie YC, Sha ZY, Yu M. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology* 2008; 1: 9-23.

Yang AX, Zhong B, Wu JH, lee. Monitoring winter wheat in ShanDong province using Sentinel data and Google Earth Engine platform, 2019.

Yang XH, Rochdi N, Zhang JK, Banting J, Rolfson D, King C, et al. MAPPING TREE SPECIES IN A BOREAL FOREST AREA USING RAPIDEYE AND LIDAR DATA. 2014 IEEE International Geoscience and Remote Sensing Symposium, 2014.

Zhang B, Zhao L, Zhang XL. Three-dimensional convolutional neural network model for tree species classification using airborne hyperspectral images. *Remote Sensing of Environment* 2020; 247.

Zhang HK, Roy DP. Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sensing of Environment* 2017; 197: 15-34.

Zhao YY, Feng DL, Jayaraman D, Belay D, Sebrala H, Ngugi J, et al. Bamboo mapping of Ethiopia, Kenya and Uganda for the year 2016 using multi-temporal Landsat imagery. *International Journal of Applied Earth Observation and Geoinformation* 2018; 66: 116-125.

Figures

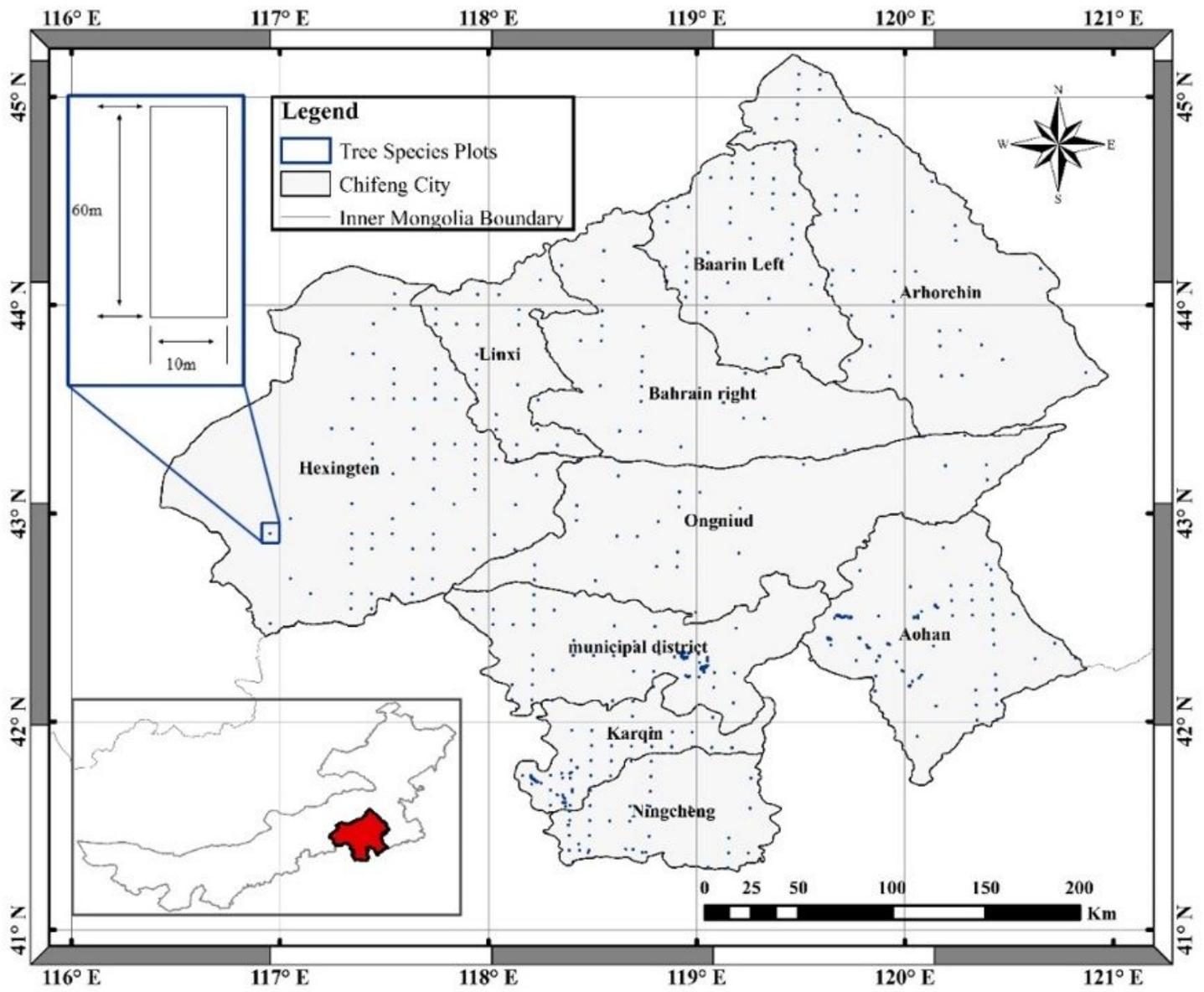


Figure 1

The study area (the blue plots represent the field measurements and the enlarged one is the shape of the square plots). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

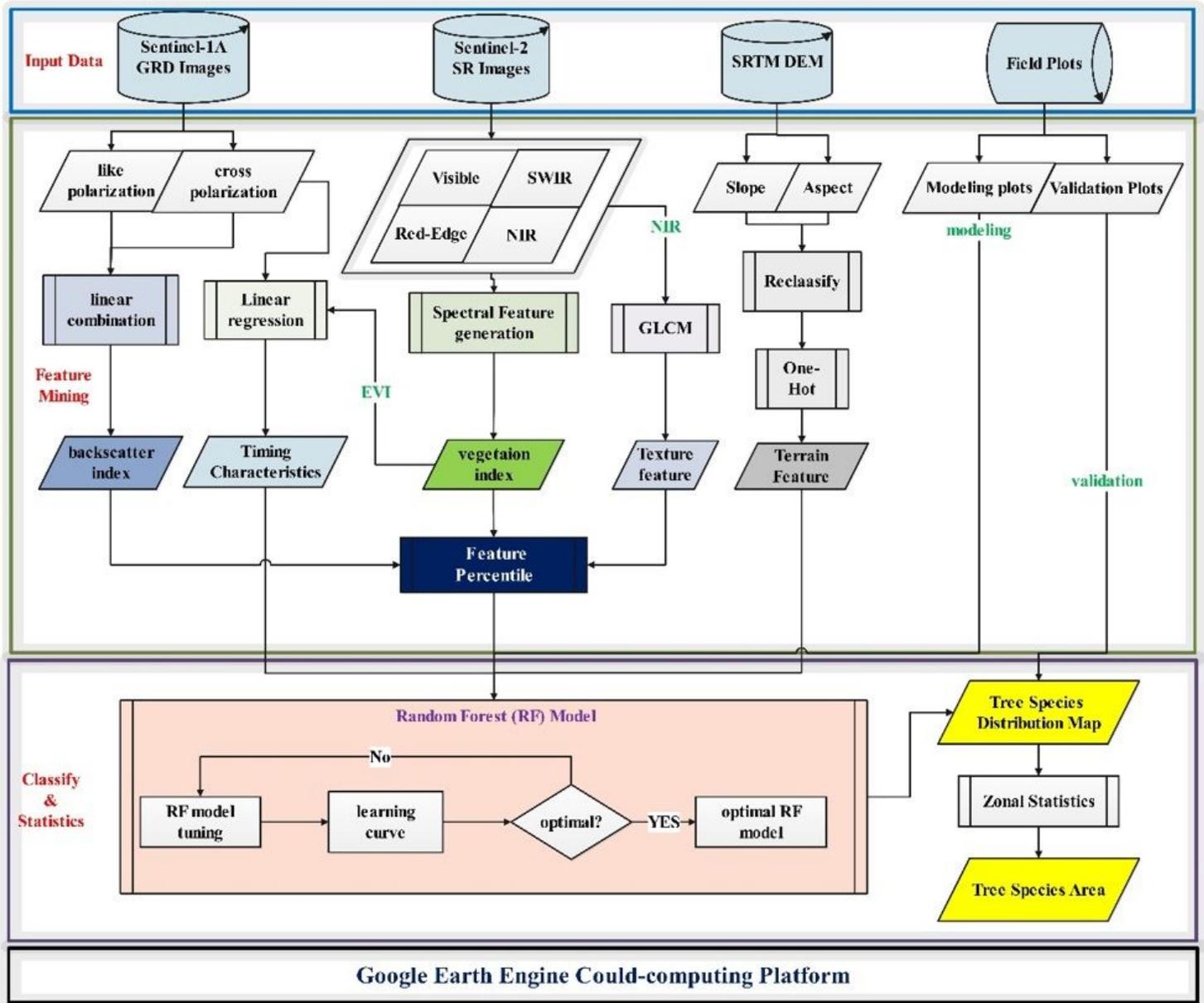


Figure 2

Workflow overview (GLCM means Grey-Level Co-occurrence Matrix).

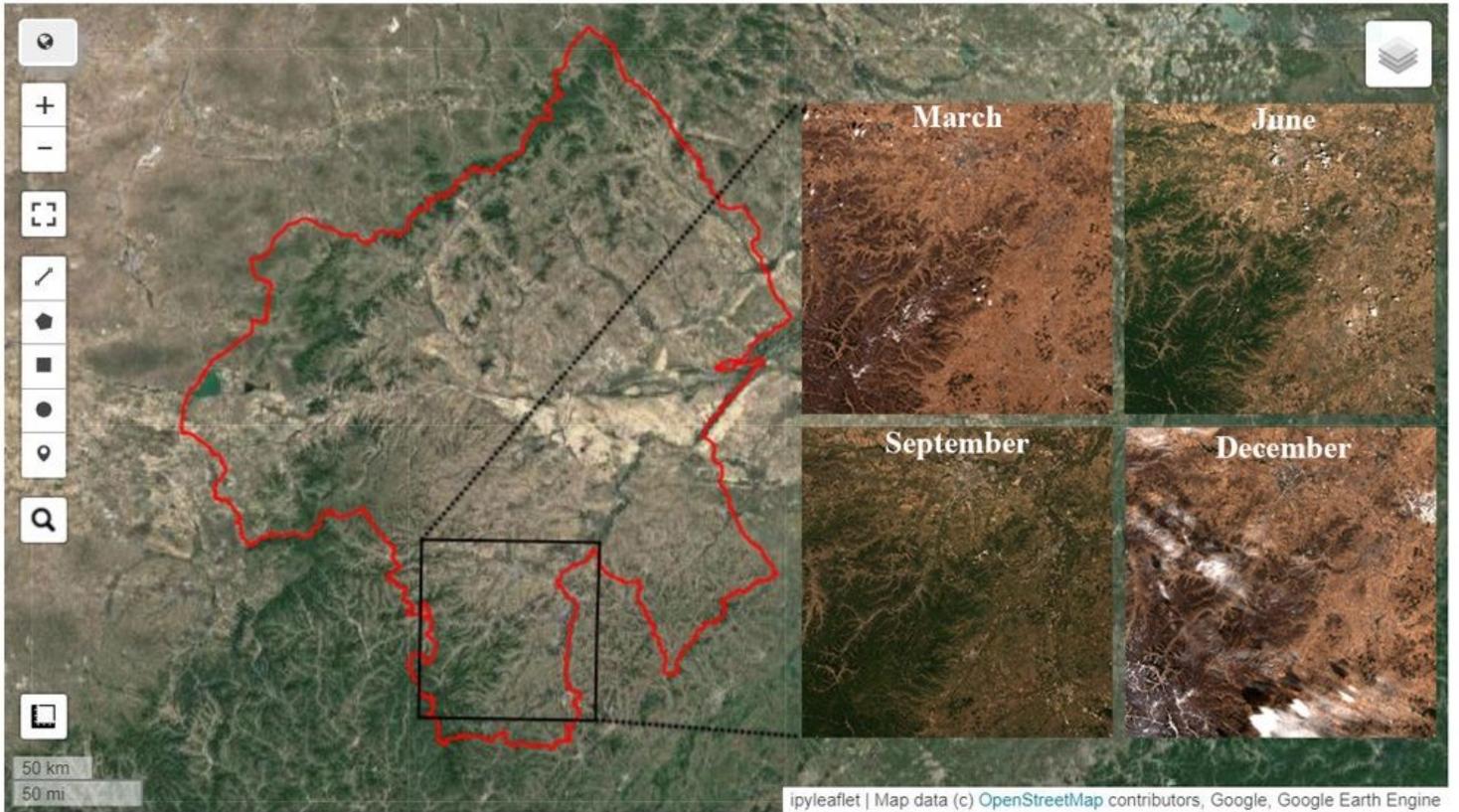


Figure 3

The enlarged Sentinel-2 images of four months in south Chifeng city. (Each scene was the image with the least cloudiness in the same area that month and was displayed in true color). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

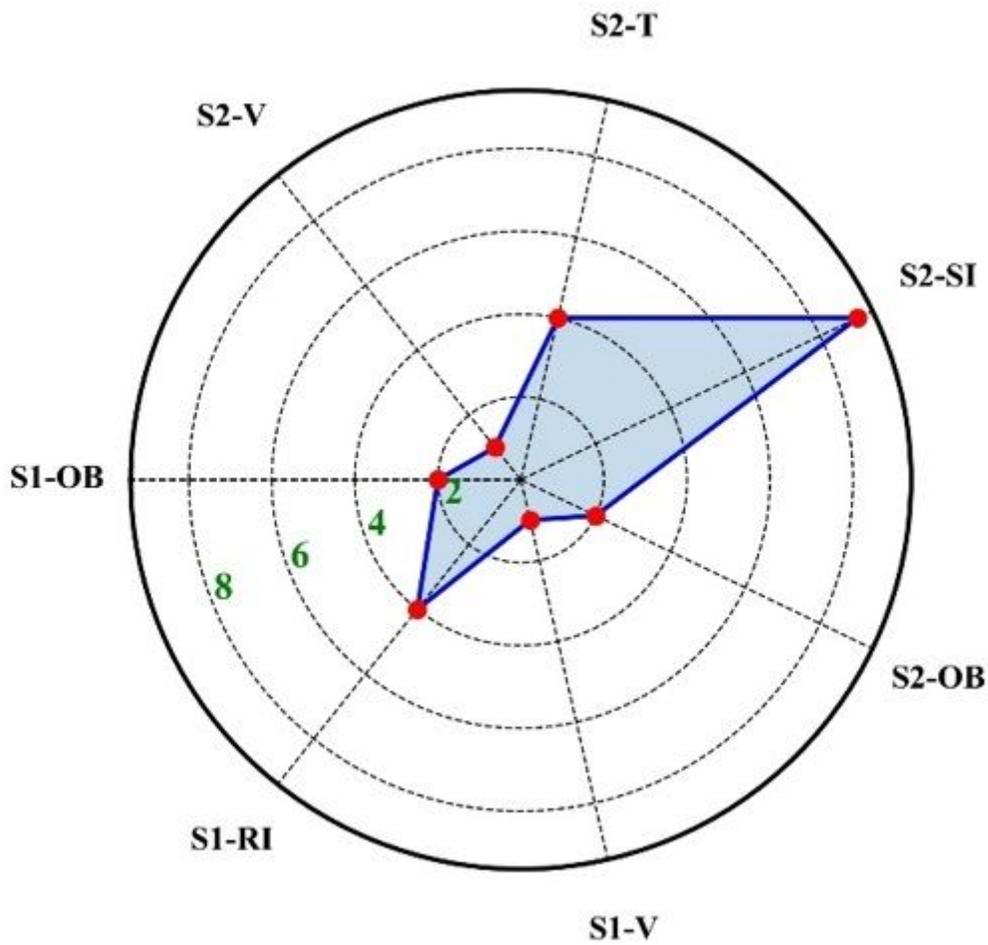


Figure 4

Monthly features composition of extracted from S1/2 data in GEE. (S1-OB=S1-Original bands, S1-RI=S1-Radar index, S1-V=S1-Variation, S2-OB=S2-Original bands, S2-SI=S2-Spectral index, S2-T=S2-Textures, S2-V=S2-Variation).

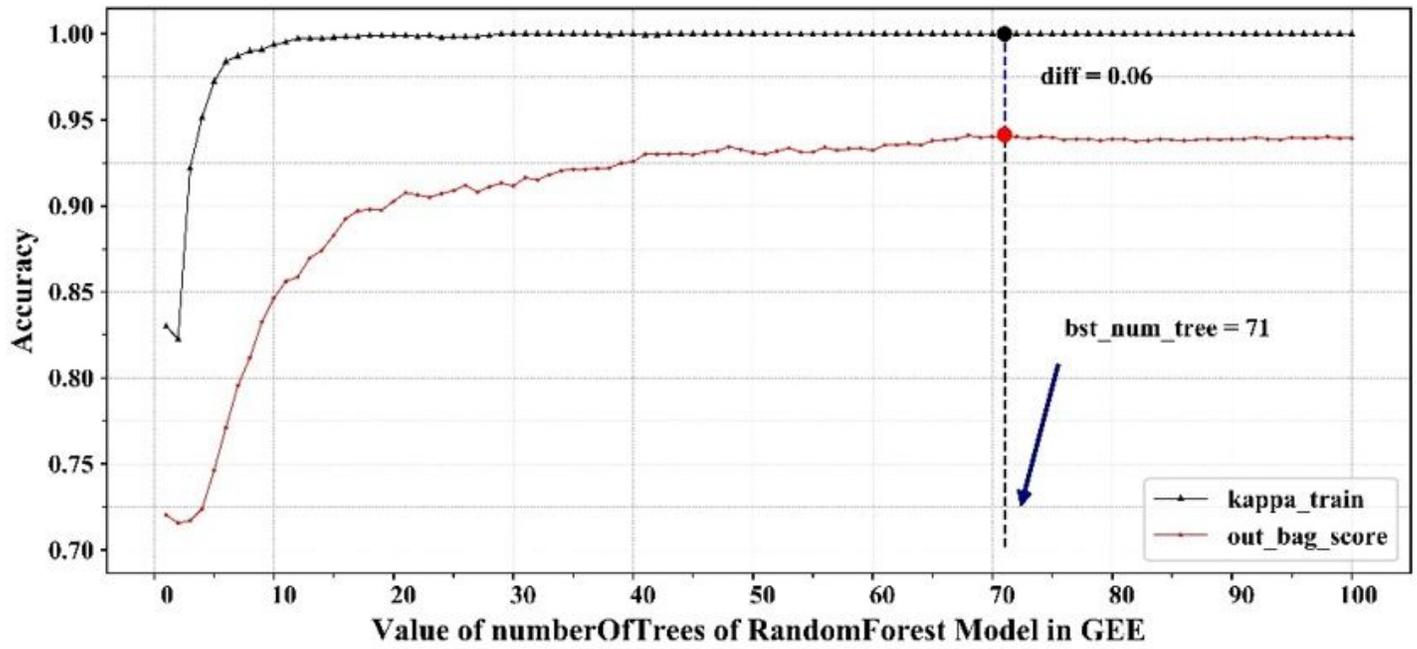


Figure 5

Learning curve of random forest classifier tuning process.

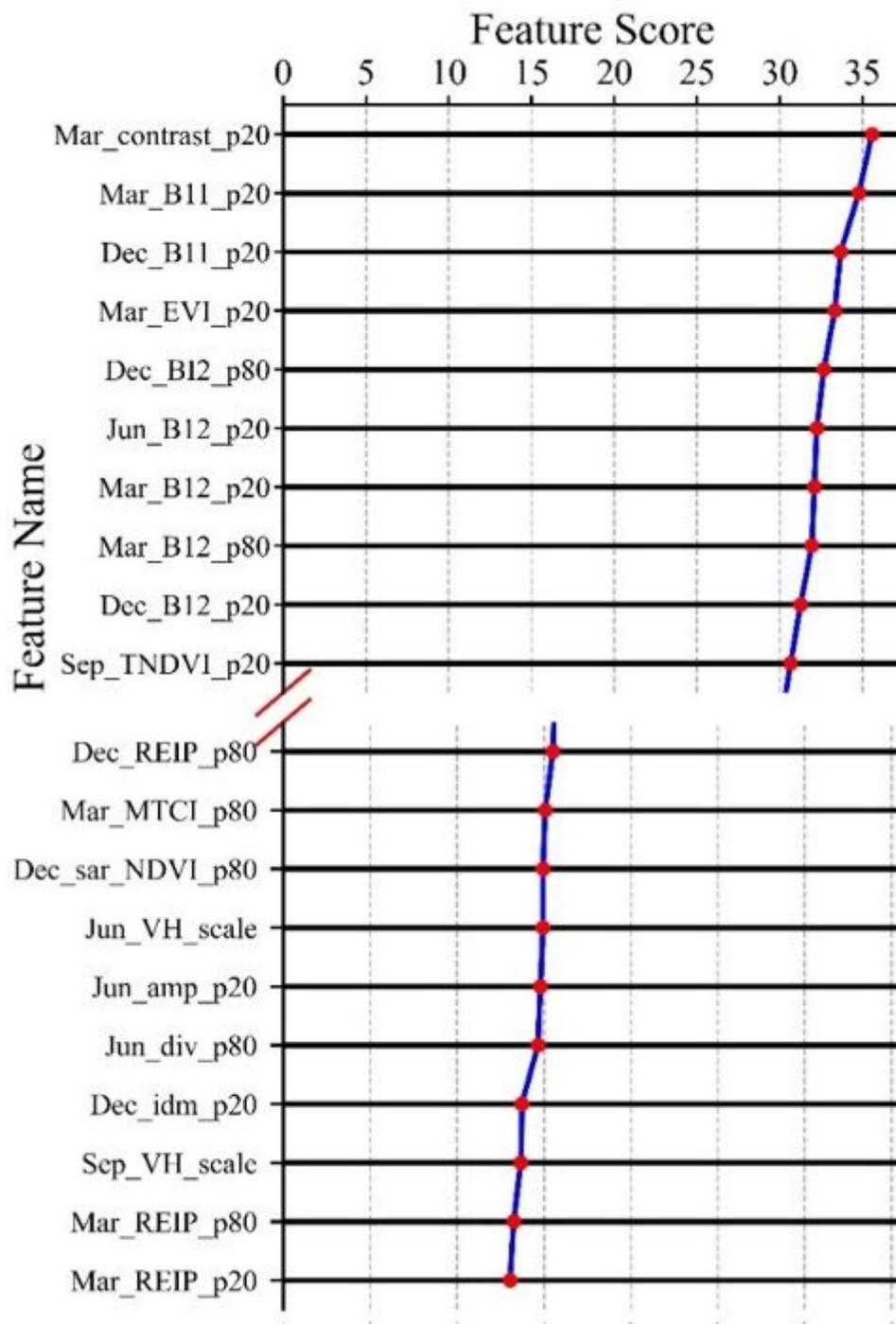


Figure 6

Multi-temporal feature collection (The suffixes p20 and p80 of characteristic variables individually denote 20th and 80th percentiles, and the prefixes Mar June Sep and Dec are the abbreviations of March, June, September and December, respectively).

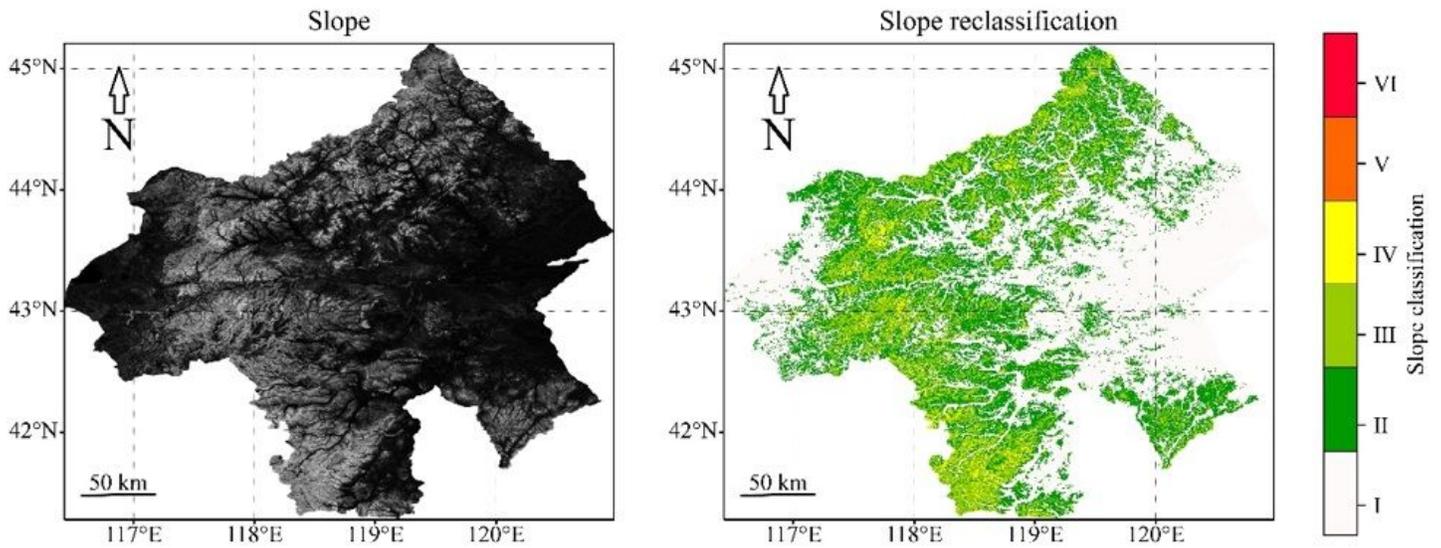


Figure 7

The slope feature extracted from topographic data (the slope: VI>V>IV>III>II>I). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

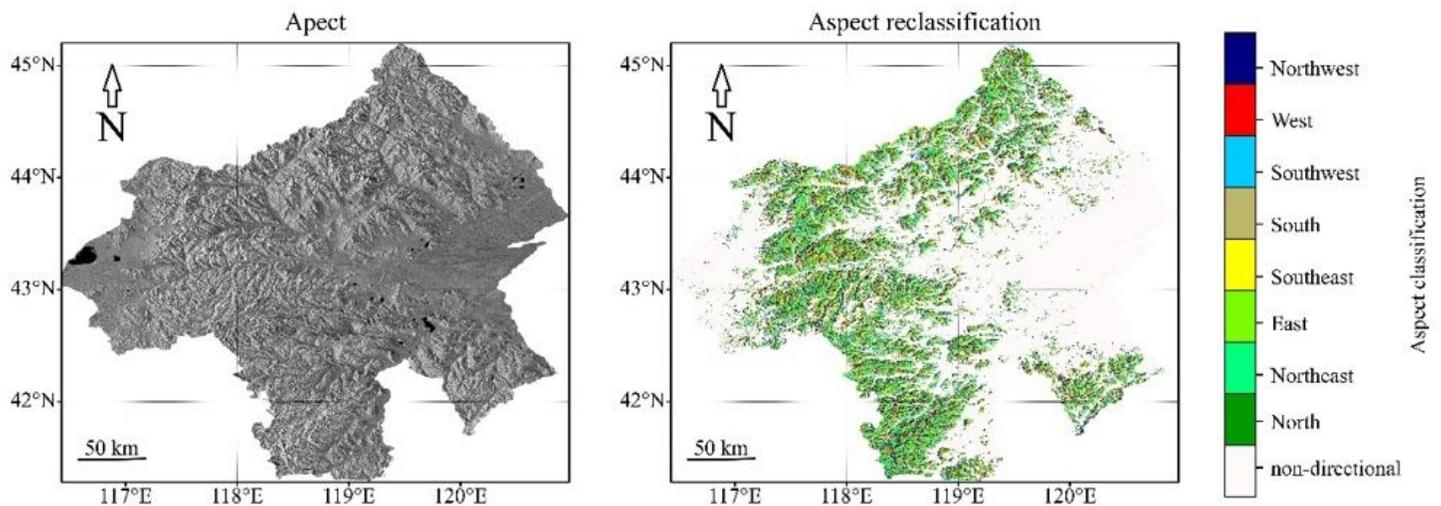


Figure 8

The aspect feature extracted from topographic data. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

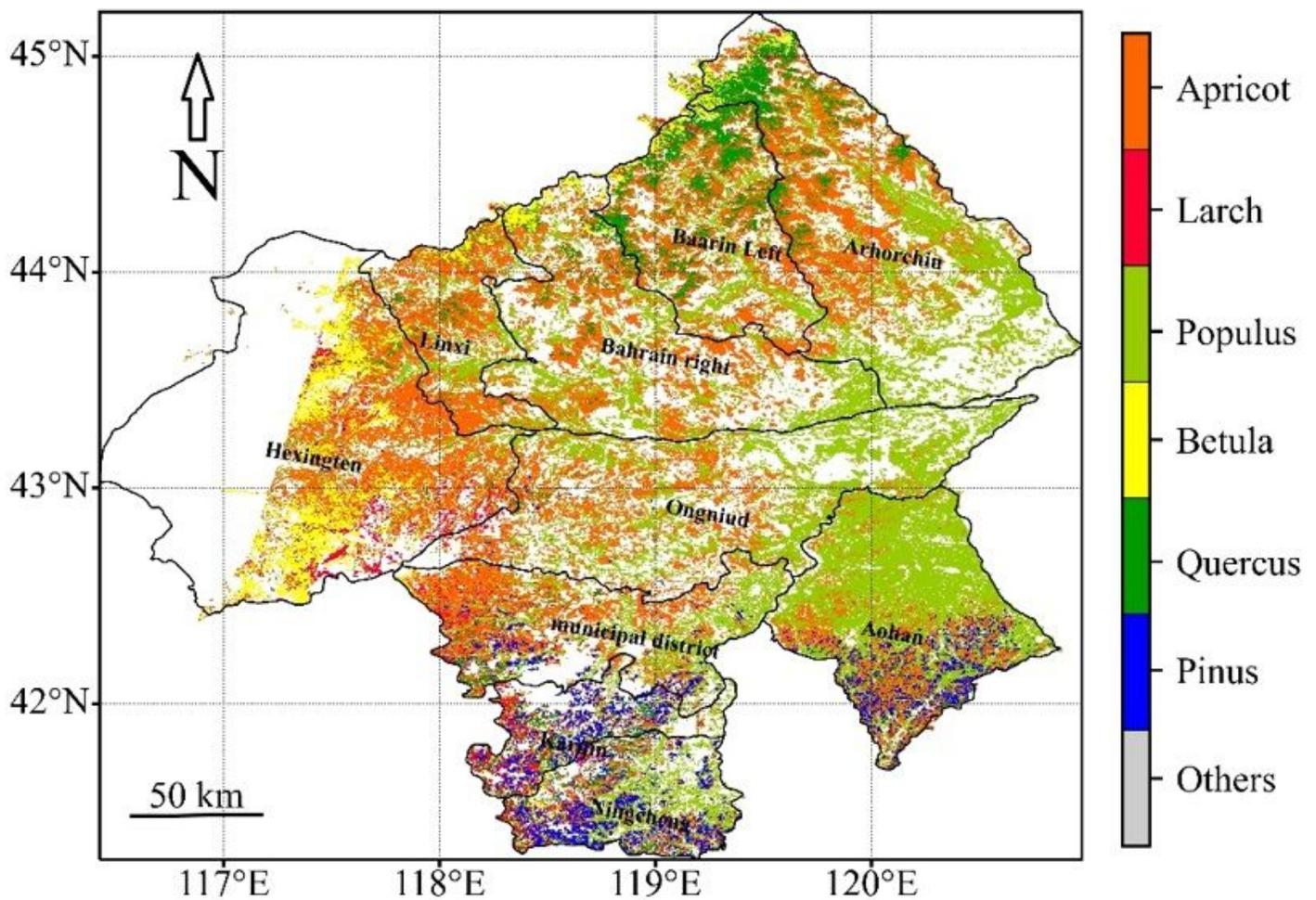


Figure 9

Spatial distribution of the objective six tree species in Chifeng city. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

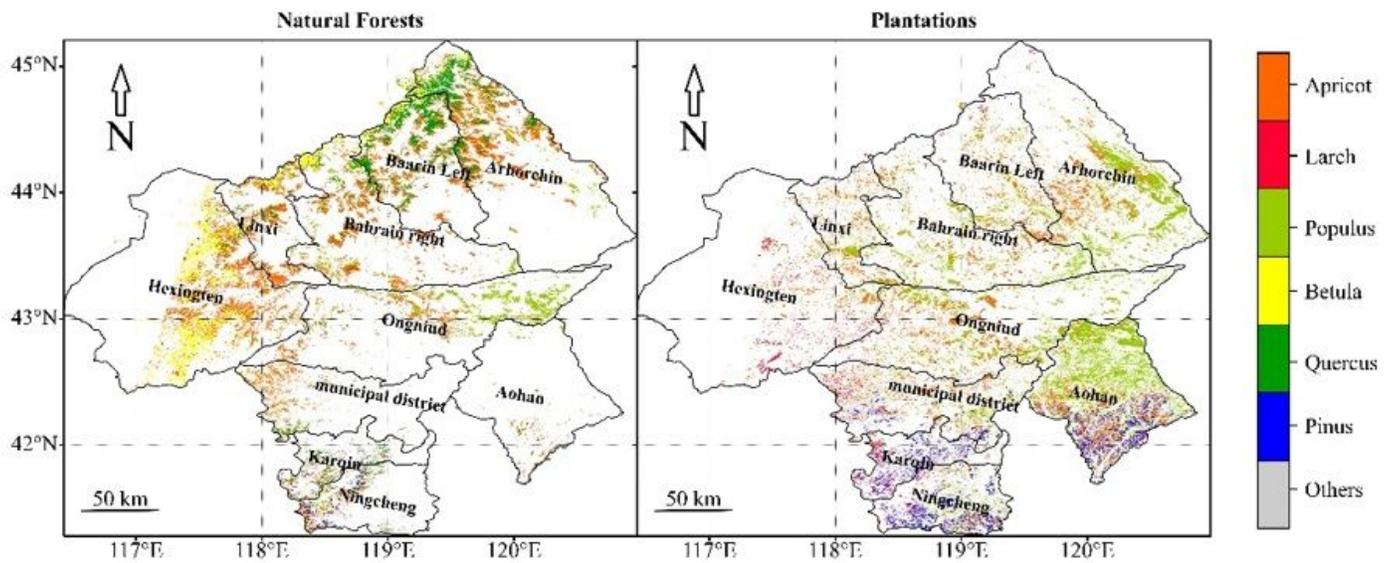


Figure 10

Spatial distribution of the objective six tree species in natural forest and plantation forest, respectively in Chifeng city. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

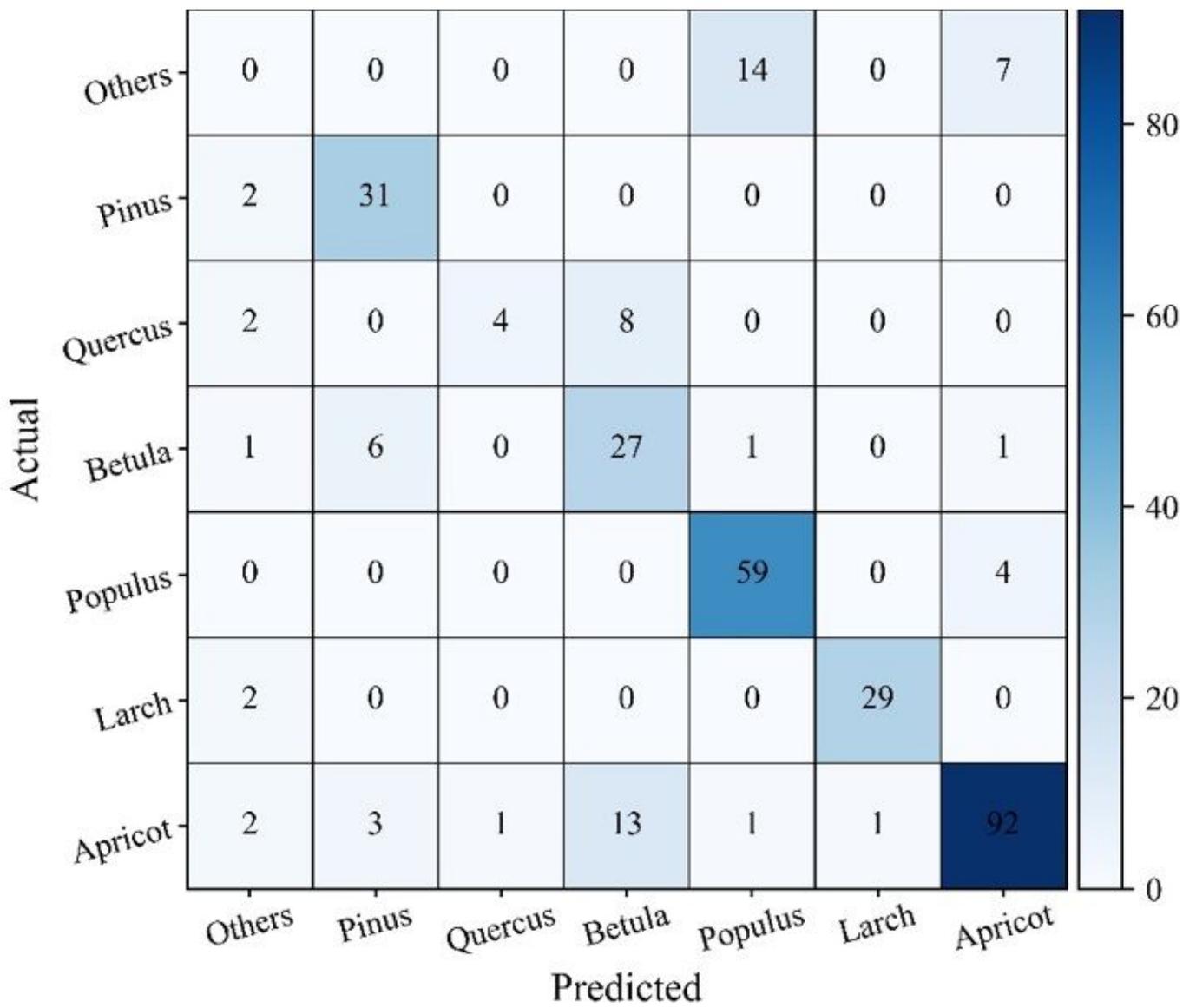


Figure 11

The confusion matrix of objective seven tree species categories.

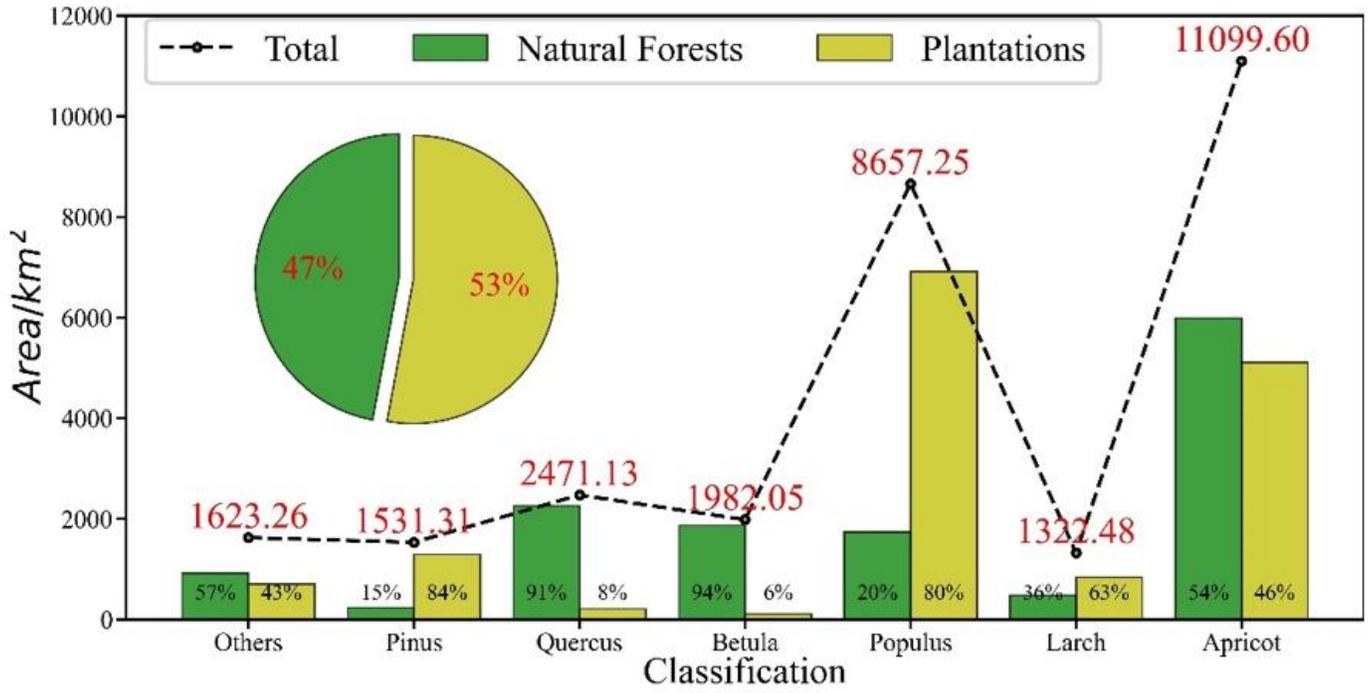


Figure 12

Area statistics of tree species according to planting mode in Chifeng city.

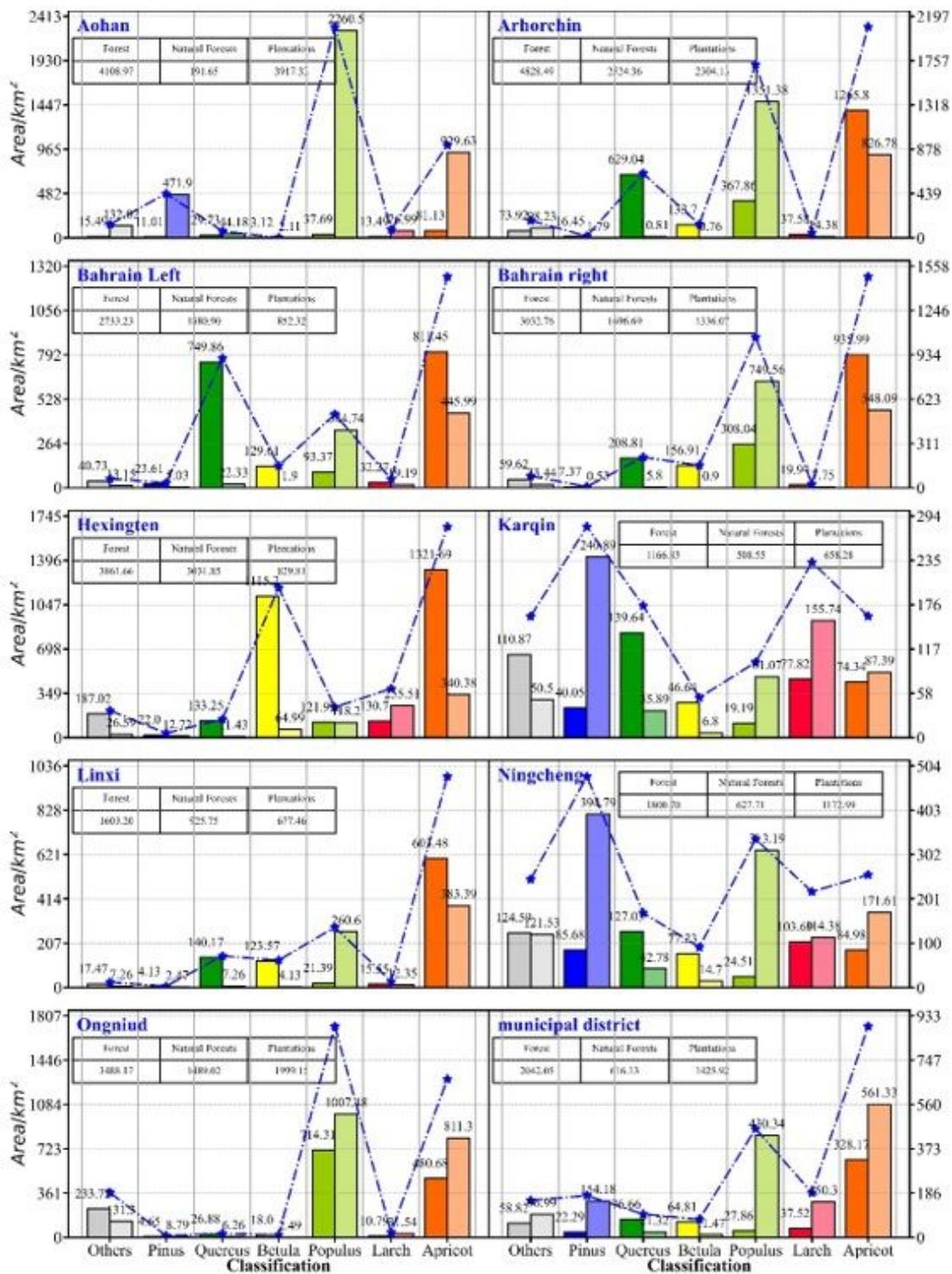


Figure 13

Quantitative description of spatial distribution of typical tree species at county level (The high saturation in the same color was the natural forests, the light color represented the plantations, and the blue polyline denoted the total area of the same tree species).