

SCIENTIFIC REPORTS



OPEN

Forest cover mapping in post-Soviet Central Asia using multi-resolution remote sensing imagery

He Yin^{1,2}, Asia Khamzina³, Dirk Pflugmacher⁴ & Christopher Martius⁵

Despite rapid advances and large-scale initiatives in forest mapping, reliable cross-border information about the status of forest resources in Central Asian countries is lacking. We produced consistent Central Asia forest cover (CAFC) maps based on a cost-efficient approach using multi-resolution satellite imagery from Landsat and MODIS during 2009–2011. The spectral-temporal metrics derived from 2009–2011 Landsat imagery (overall accuracy of 0.83) was used to predict sub-pixel forest cover on the MODIS scale for 2010. Accuracy assessment confirmed the validity of MODIS-based forest cover map with a normalized root-mean-square error of 0.63. A general paucity of forest resources in post-Soviet Central Asia was indicated, with 1.24% of the region covered by forest. In comparison to the CAFC map, a regional map derived from MODIS Vegetation Continuous Fields tended to underestimate forest cover, while the Global Forest Change product matched well. The Global Forest Resources Assessments, based on individual country reports, overestimated forest cover by 1.5 to 147 times, particularly in the more arid countries of Turkmenistan and Uzbekistan. Multi-resolution imagery contributes to regionalized assessment of forest cover in the world's drylands while developed CAFC maps (available at <https://data.zef.de>) aim to facilitate decisions on biodiversity conservation and reforestation programs in Central Asia.

Central Asia comprises five predominantly agricultural countries of the former Soviet Union (Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan) that, although largely covered by drylands, are home to diverse and important forest ecosystems, ranging from continuous forests in the higher-rainfall upstream areas to patchy riparian forests in the more arid downstream areas^{1–3}. Notwithstanding their relatively small coverage area, forests in Central Asia play a key role in the regional hydrology by helping maintain a steady river discharge from the high mountains to the irrigated lowlands⁴. Mountainous forests are particularly rich in terms of the genetic diversity of wild fruit and nut tree species and serve as carbon sinks^{5–8}. Given their great ecological and environmental importance, most of the forests in Central Asia have been placed in Group I of the Soviet Union forestry system, which is designated for conservation and protection^{4,9–12}. However, following the collapse of the Soviet Union, regular and standardized forest inventories have stopped, with the effect that the status and condition of these forests is currently uncertain.

Forest statistics provided by these Central Asian countries to the Food and Agriculture Organization of the United Nations (FAO) are mostly out-dated or missing. For example, the country report by Uzbekistan states that the last forest inventory was conducted in 1987; no national statistics are provided by Turkmenistan; Kazakhstan reports results based on a desk study (<http://www.fao.org/countryprofiles/en/>); and in general, the capacities of these countries for forest monitoring is unknown^{13–15}. There are substantial discrepancies between datasets generated from different sources, with the effect that there is no commonly agreed upon estimate for the total forested area in Central Asia⁴. Based on individual country reports, the Global Forest Resources Assessments (FRA)¹⁶ estimates that 11.7 million ha, or less than 5% of the land area, are covered by forest in Central Asia. However, official statistics often overestimate the total forested area and do not reflect the actual forest condition¹⁷. The use of different definitions for forest, some of which deviate from international definitions, introduces further

¹Center for Development Research (ZEF), University of Bonn, Walter-Flex-Str. 3, 53113, Bonn, Germany. ²SILVIS Lab, Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI, 53706, USA. ³Division of Environmental Science and Ecological Engineering, Korea University, 145 Anam-Ro, Seongbuk-Gu, Seoul, 02841, Korea. ⁴Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099, Berlin, Germany. ⁵Center for International Forestry Research (CIFOR), P.O. Box 0113 BOCBD, Bogor, 16000, Indonesia. Correspondence and requests for materials should be addressed to A.K. (email: asia_khamzina@korea.ac.kr)

ambiguities between estimates from Central Asian countries^{18–20}. Spatially explicit, comparable methodologies for obtaining forest information are much needed in Central Asian countries, so that high quality datasets can be established that will enable adequate assessment of ecosystem services (e.g., biodiversity, climate, and water cycle regulation) and development of sustainable management options²¹.

Remote sensing provides seamless and periodic observations of the Earth, and for countries with few or no inventories, it is the only viable source of forest area estimates e.g.²². The independence of remote sensing from national forest agencies is an additional advantage for forest monitoring across political boundaries²³. Several continental and global land cover mapping efforts based on coarse resolution satellite data (i.e., spatial resolutions greater than 250 m) provide a number of land cover products, such as the GLC2000²⁴, GlobCover²⁵, and MODIS Land Cover Type^{26,27} products. These datasets map the distribution of forest types with acceptable accuracies on a global scale but, due to their coarse resolution, are insufficient for national forest area reporting²⁸. Products that estimate the sub-pixel proportion of forest area, such as MODIS Vegetation Continuous Fields (VCF)²⁹, seek to overcome the coarse resolution bias^{30,31}. However, the accuracy of global maps can vary substantially between regions^{32–39}. It is difficult and not always feasible to capture regionally important ecological gradients with global products that are based on globally selected reference data^{26,40,41}. Large discrepancies between global land cover products have been found, especially at biome borders and in sparsely vegetated environments^{42,43}.

With recent developments in data access and distribution policies for Landsat and Sentinel-2, improved forest monitoring using medium-resolution satellite image time series data has become feasible^{44,45}. Medium-resolution sensors capture more spatially detailed information, which is important for forest monitoring in heterogeneous drylands^{46,47}. In addition, using multi-temporal instead of single-date imagery allows for more accurate land cover classification^{48–50}. Recently, the concomitant use of all available Landsat imagery time series data has provided opportunities for large-area forest cover mapping^{23,51,52}. Taking advantage of all available observations per pixel, it is possible to use Landsat time series data for improved classification^{53–55}.

Combining medium- and coarse-resolution imagery can support cost-effective forest mapping over large areas^{56–58}. In spite of the rapid developments in computational capability, processing the large volumes of data associated with higher resolution images is still challenging. In addition, data scarcity due to low observation density and persistent cloud and snow coverage make it difficult to use Landsat to produce wall-to-wall forest cover maps^{59,60}. Therefore, coarse resolution archive imagery, such as from AVHRR and MODIS, has been combined with Landsat data for the land cover mapping of large areas^{56,57,61–63}. The higher resolution of Landsat makes it feasible to estimate the percentage of different land cover types at the MODIS level of coverage and predict wall-to-wall sub-pixel forest cover through regression or other approaches^{64–67}.

The goal of this study was to develop a consistent and improved forest cover map for Central Asia, and to assess the current national and global forest cover estimates for this region. To achieve this, we first produced land cover maps from a sample of Landsat footprints corresponding to the World Reference System 2. We then used the Landsat-based map as a reference to train a fractional forest cover model based on MODIS time series data at a resolution of 250 m. The objectives of this study were thus to: 1) assess the usefulness of dense Landsat time series data for forest cover mapping in drylands; 2) produce a Central Asia Forest Cover (CAFC) map using MODIS time series data for 2010 that distinguishes the prevalent forest types; and 3) compare this regionalized forest cover map with the FRA and other available global coverage products to evaluate the relative strengths and limits of all available tools for forest monitoring in Central Asia.

Results

Classification of Landsat imagery. The use of spectral-temporal metrics generated from three years of imagery yielded the highest classification accuracy (overall accuracy OA: $OA = 0.83 \pm 0.04$) while using single-date composite produced the lowest ($OA: 0.56 \pm 0.04$) (see Supplementary Table S1, and Fig. S1). Using temporal metrics improved the separation of deciduous forest and herbaceous vegetation (see Supplementary Fig. S2). In addition, the confusion between mixed and evergreen forest decreased, resulting in improved mapping accuracy for these two classes. Extending the observation period from one to three years reduced the classification uncertainty further. Most of the land cover classes, particularly forest, exhibited fewer classification errors. Fewer cases of confusion occurred between forest and wetland (see Supplementary Fig. S2). Adding observations from 2009 and 2011 to the 2010 dataset led to more accurate distinctions between different forest classes, especially for deciduous forest. The data scarcity effect, e.g. the strip artefact of ETM + Scan Line Corrector (SLC)-off, was largely eliminated when including temporal metrics and the extended observation period (see Supplementary Fig. S2).

Forest cover percentage prediction using MODIS. The Central Asia Forest Cover (CAFC) map produced from the MODIS data indicates that forest covers 4.96 million ha (1.24% of the land area in Central Asia) with a distinct variation in forest types between the countries (Figs 1 and 2). Kyrgyzstan was found to have the largest proportion of forest cover (3.3% of its land area), while Turkmenistan carried the least (0.06% of the land area). Due to its large territory, Kazakhstan is home to most of the forest (79.4% of the total forest in Central Asia), with forests relatively evenly distributed between deciduous, evergreen, and mixed forest types. In the mountainous countries Tajikistan and Kyrgyzstan, evergreen forests were more prominent, occupying 67.9% and 60.2% of the total forest area, respectively. Deciduous forests prevailed in Turkmenistan (68.3% of the country's forest area), a much more arid country. Prediction accuracy varied by forest type (Fig. 3). Overall forest cover was predicted with the highest accuracy. The normalized root-mean-square error (NRMSE) of the predicted evergreen forest cover ($NRMSE = 1.19$) was lower than the errors for deciduous (1.35) and mixed forests (1.84). The overall forest cover map had the lowest NRMSE (0.63). The linear relationship between predicted forest cover

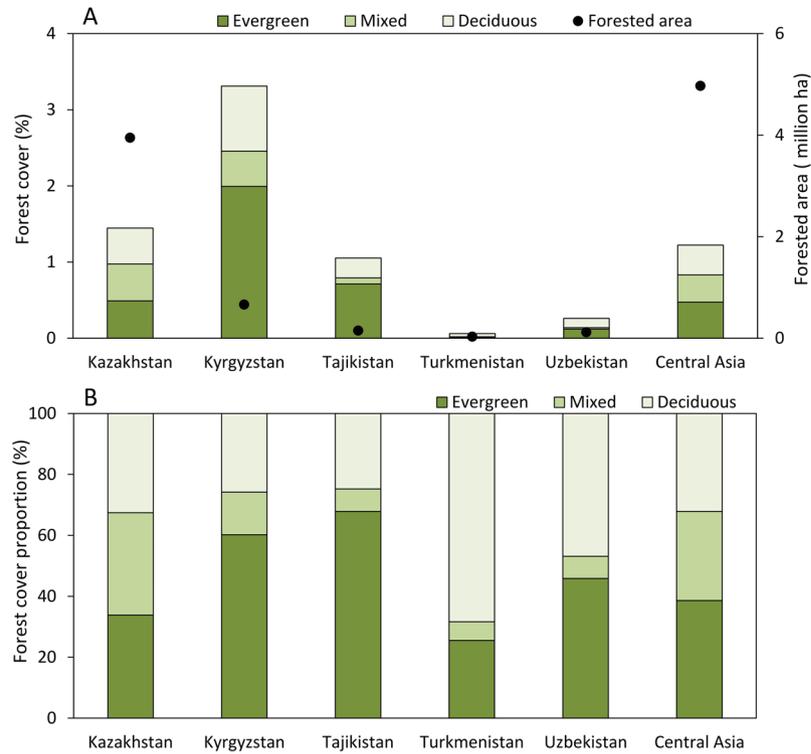


Figure 1. Land area covered by forests (A) and the composition of forest types (B) in Central Asia.

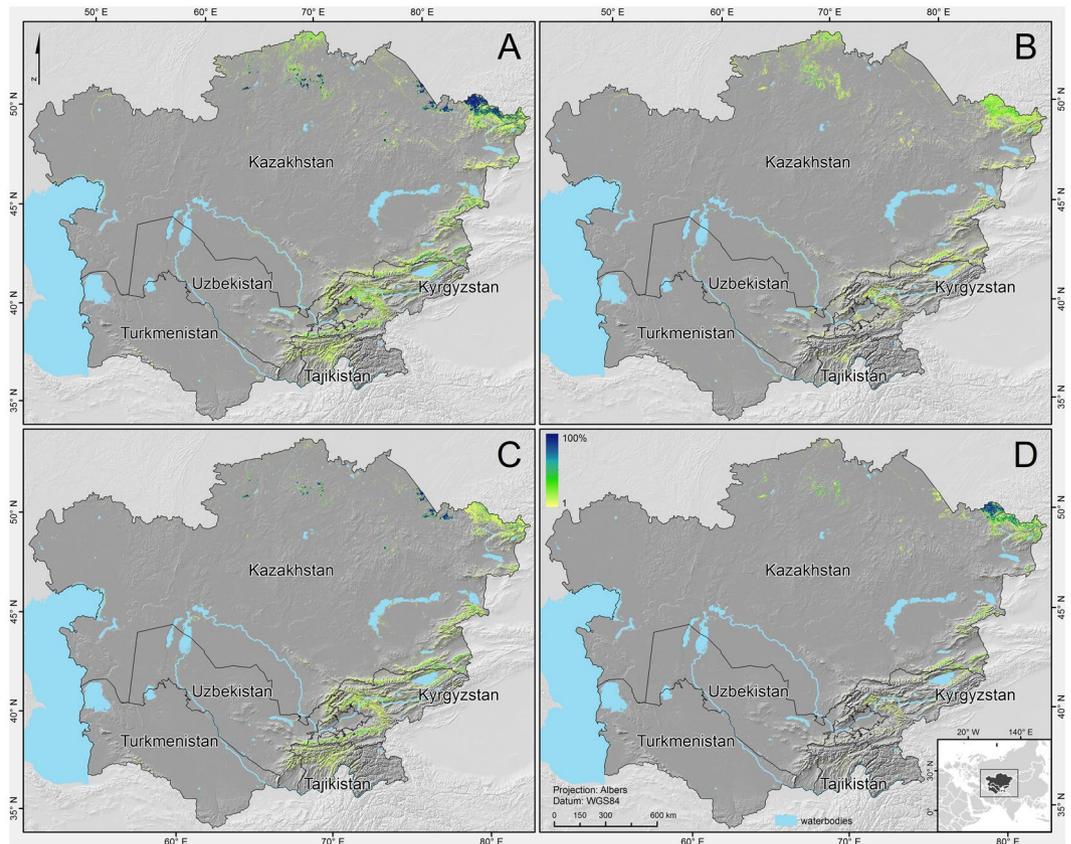


Figure 2. Forest cover percentage with (A) all forest types, (B) deciduous, (C) evergreen, and (D) mixed forest. This figure was produced using ArcGIS 10.3.

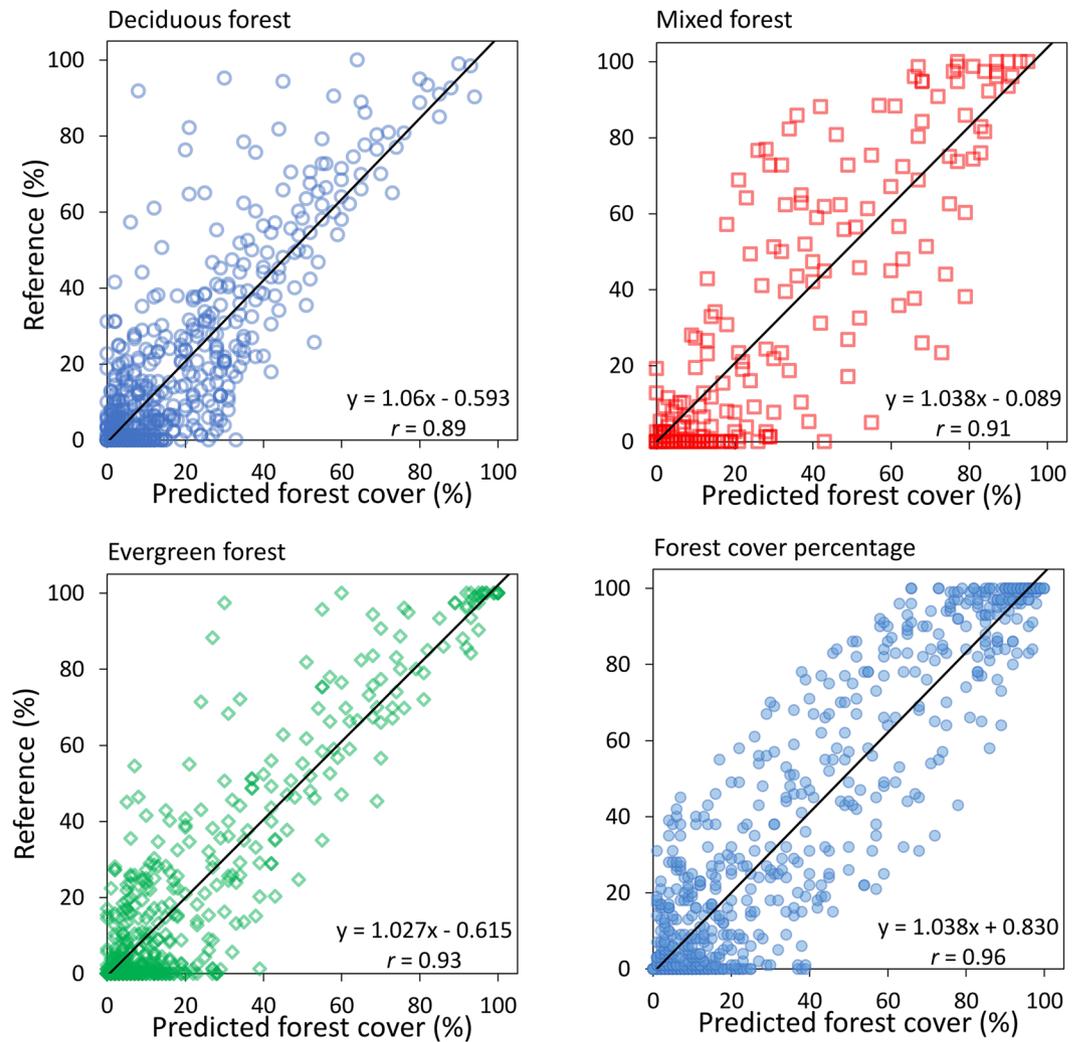


Figure 3. Scatterplots of reference forest cover derived from Landsat land cover maps and the estimated forest cover from MODIS data as deciduous, mixed, evergreen, and overall forest.

and reference forest cover was stronger (Pearson's $r = 0.96$) than that between the cover estimates of any single forest type and the reference.

Comparison of forest cover estimates. Comparison of the forest cover estimates for Central Asia showed both agreement and discrepancies between the CAFC and the global products. In general, the CAFC produced patterns similar to GFC (Fig. 4; also see Supplementary Fig. S3). Although the GFC is more detailed because of its higher spatial resolution, we found localized but substantial over- and underestimations in this map. For example, GFC overestimated forest cover in non-woody wetlands (Figure S3, Ili River Delta) and tended to underestimate forest cover in riparian forests (Figure S3, Amu Darya) and evergreen forests (Figure S3, Eastern Kazakhstan Region). When comparing GFC and MODIS VCF, we found that MODIS VCF overestimated forest cover in regions with low forest cover and underestimated the dense forest areas in most of Central Asia (Fig. 4, see also Supplementary Fig. S3). Overall, the strong agreement between the CAFC and GFC further improved when we aggregated the maps to coarser spatial resolutions (Fig. 4). The correlation (Pearson's r) between the CAFC and GFC increased from 0.77 at a pixel size of 250×250 m to 0.86 at 5×5 km. At a pixel size of 5×5 km, the regression slope was close to 1 (0.95). MODIS VCF and GFC had a strong positive relationship at pixel sizes of 250×250 m ($r = 0.7$) to 5×5 km ($r = 0.89$); however, MODIS VCF made lower tree cover predictions with regression slopes ranging between 1.13 and 1.37. Forest cover estimates from the CAFC and GFC tended to agree closely, with the exception of mountainous Tajikistan. MODIS VCF typically provided a much lower estimate, totalling 0.3% of Central Asia as a whole (Table 1). The FAO FRA 2010⁶⁸ and FRA 2015¹⁶, relying on individual country reports, largely overestimated forest cover, with the exception of Kazakhstan and Kyrgyzstan in most recent assessment. For the other countries, the FRA estimates were between 1.5 and 147 times higher than the CAFC estimates, with particular extremes for the more arid, downstream countries of Turkmenistan and Uzbekistan.

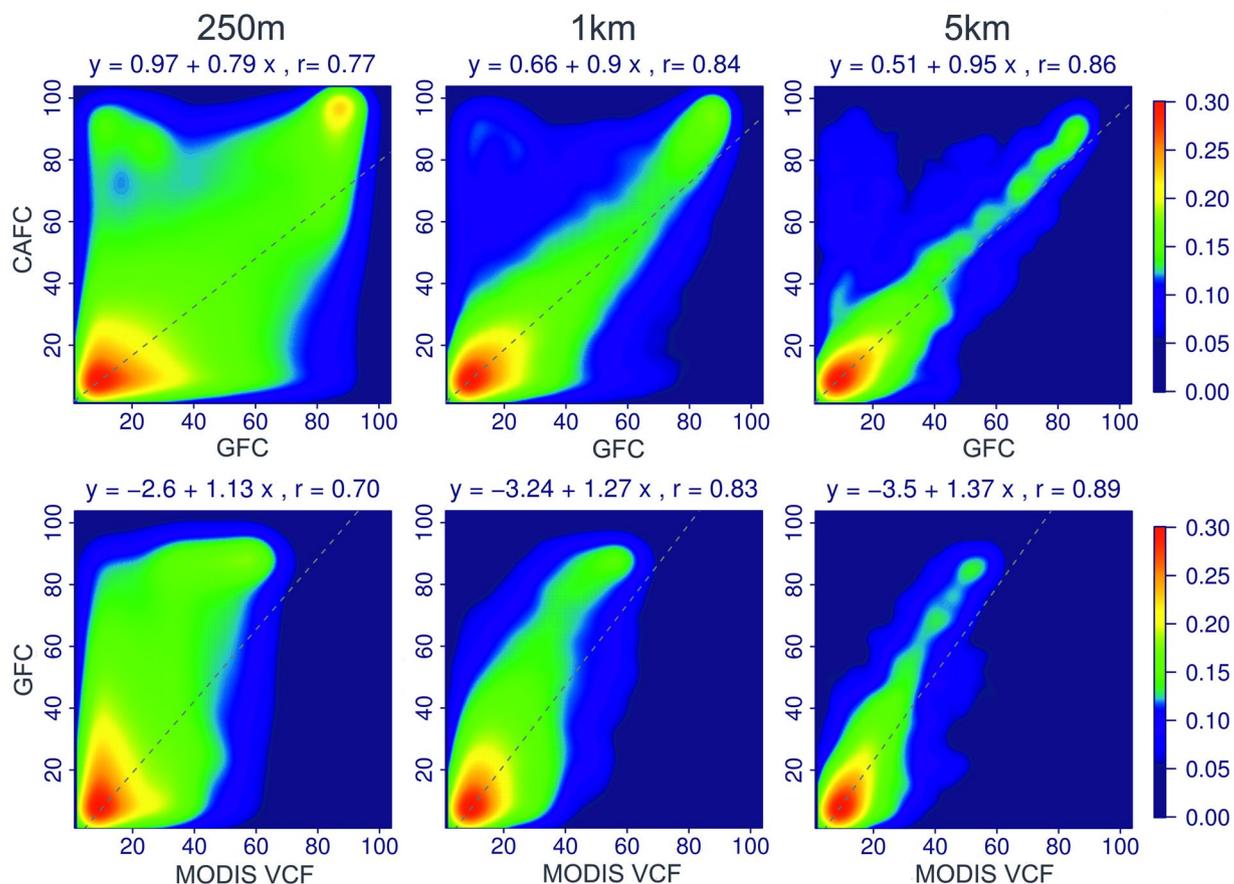


Figure 4. Joint distribution of forest cover percentage from CAFC vs. GFC (upper row) and tree cover percentage from MODIS VCF vs. GFC (lower row) at the sampling resolutions of 250 m, 1 km, and 5 km, with the colour scheme representing local densities at each point of the scatterplot (obtained through a kernel density estimate) and the dotted regression line demonstrating the coherence between different products.

Data source	Kazakhstan	Kyrgyzstan	Tajikistan	Turkmenistan	Uzbekistan
CAFC	1.45	3.3	1.05	0.06	0.29
GFC	1.59	3.42	0.43	0.02	0.24
MODIS VCF	0.41	0.54	0.02	0.00	0.01
FRA (2010)	1.2	5	2.9	8.8	7.7
FRA (2015)	1.2	3.3	3.0	8.8	7.3

Table 1. Forest cover percentage (%) in Central Asian countries estimated from the Central Asia Forest Cover map (CAFC), Global Forest Change (GFC), MODIS VCF, and Forest Resources Assessments (FRA).

Discussion

Using all available Landsat imagery rather than single-date spectral reflectance, we were able to adequately distinguish between forests and other land cover classes with an overall accuracy of 0.83 ± 0.04 , increased from 0.56 ± 0.04 when only one date was used. Metrics derived from time series data provide information on phenology patterns and therefore are less influenced by inter-annual variability in the imaging dates and overall image availability⁵⁵. This is critical for land cover mapping in drylands characterized by high diversity in plant phenology. Moreover, metrics calculated from full time series are more resistant to the noise from geometric errors in single images⁶⁹.

Extending the observation period from one year to three years (called epoch imagery) for land cover classification with the Landsat data further reduced mapping errors (OA: 0.83 ± 0.04 compared to 0.71 ± 0.04). Using MODIS time series data, Hüttich *et al.*⁷⁰ arrived at similar accuracy levels with increasing length of observation period in the drylands of Africa. In contrast to the frequent MODIS observations, Landsat imagery has lower temporal resolution and the time series data has many gaps^{45,71}. Our study suggests that data scarcity due to sensor errors or frequent cloud/snow cover could be reduced by including imagery from neighbouring years. Furthermore, classification errors due to considerable intra-annual variations in dryland vegetation can be largely avoided by including multi-annual time series data from remote sensing observations⁷².

Extreme land cover conversions could produce large variations in a time series, thus undermining its suitability for use in epoch imagery for classification. To the best of our knowledge, there were no dramatic changes for the examined land cover categories in Central Asia during the observation period of 2009–2011. The mapping confusion between deciduous forest and herbaceous land was likely due to general landscape heterogeneity in Central Asia. Mixed land cover produces mixed pixels which are prone to classification errors⁷³. Hindsight examination on the confused validation samples revealed that 53% of all deciduous forest/herbaceous land misclassifications were mixed pixels dominated by herbaceous vegetation and scattered trees. Another challenge is the separation of land cover classes with similar spectral and temporal responses. Mixed forest has a wide range of spectral reflectance and is thus often confused with deciduous or evergreen forests⁷⁴. These uncertainties are a problem inherent to land cover mapping based on remote sensing in heterogeneous landscapes and can be largely reduced through the use of temporal metrics from multi-annual Landsat time series data. Compared to other forest cover maps our mapping approach using Landsat time series captures forest cover in drylands reasonably well (Supplementary Table S1). For example, the National Land Cover Database (NLCD) of the Conterminous United States, also using Landsat data, reported the forest user's accuracy of 47% and the producer's accuracy of 99% in the dryland region, covering the great plains and deserts⁷⁵.

We found strong agreement between MODIS- and Landsat-based forest cover estimates indicating that MODIS can provide reliable data for forest cover monitoring in dryland forests. Although uncertainty was relatively high at the pixel level (e.g., an NRMSE of 0.63 for overall forest cover), the precision of the estimates improved when aggregated over larger areas. Using MODIS to estimate the proportional cover of different forest types in Central Asia revealed uncertainty due to contamination of the Landsat training dataset. Mapping errors also originated from training data imperfections⁷⁶. Using a land cover map derived from high resolution imagery as a reference in the estimation of fractional land cover at coarse resolutions could lead to inter-resolution error propagation⁴⁶. The slight overestimation of forest cover in the wetlands of the Ili River delta (see Supplementary Fig. S3) could be traced back to the residual error in the Landsat map (see Supplementary Fig. S2). The higher mapping error of deciduous forest in the Landsat map may decrease a mapping accuracy of the CAFC deciduous forest map (see Supplementary Fig. S1). The challenge of distinguishing between spectrally similar classes persisted at the MODIS scale, as shown by the relatively moderate mapping accuracy for mixed forest (Fig. 3). In the same vein, mapping accuracy was highest for the overall forest fractional cover, eliminating confusion errors between forest types (Fig. 3).

We detected considerable differences between the forest maps generated from continuous tree cover products for Central Asia. Saturation of spectral reflectance, phenological variation, and confusion with dense herbaceous vegetation are possible reasons for commission errors in sparsely covered areas and omission errors in densely covered regions in the MODIS VCF product^{77,78}. In contrast, GFC generally provided a reliable forest estimate in Central Asia. However, calibration, especially to correct mapping errors in riparian and evergreen forests, would be needed for regional applications of the GFC product in Central Asia. In our study model-based inferential framework was used to obtain forest area estimates^{79,80}. In contrast to design-based inference, the validity of the model-based inference was supported by the strong linear relationship ($r = 0.96$) and the coefficient close to 1. Both suggest the unbiasedness of our model-based area estimate. To overcome uncertainty in the map data, there is a need to integrate various resources such as ground-based measurements and satellite-derived maps to generate systematic forest cover estimates at a national level^{28,81}.

In general, the differences between the FRA estimates and all satellite-derived data were substantial. The greater forest cover reported by Turkmenistan and Uzbekistan in the FRA is at least partially attributed to these countries reporting sparse desert woodland as part of the national forest areas^{18,82}. Desert woodlands are characterized by low primary production, but are highly valuable in the regional ecology¹⁸, which explains the high importance given to them in the national forest reports.

The minimum area used in the forest definition could be another factor causing differences between the remote sensing estimates and the country reports. Applying a minimum mapping unit can influence forest area estimates from map data²⁰, particularly in heterogeneous and fragmented landscapes. Forest inventories in Central Asian countries traditionally have not followed a fixed minimum area criterion hence various values were used, which might be above or below the uniform FAO requirement of >0.5 ha^{16,82}. In this study, we did not apply a minimum area requirement for the definition of a forest. Because the majority of the forests in Central Asia are contiguous areas, the effect is likely negligible or could even increase the difference between the remote sensing estimates and the country reports.

While countries might choose their own forest definition when reporting to the FRA, the large discrepancy in forest area estimates between the remote sensing assessment and the national statistics⁶⁸ calls for a more coherent approach to forest monitoring across Central Asia. To this end, our proposed cost-effective approach using time series data from multi-resolution imagery and coherent maps of forest cover with 250 m resolution across Central Asia can complement and facilitate national forest inventories still largely dependent on resource demanding field surveys.

In particular, land cover classification using the Landsat imagery revealed the value of dense time series for land cover mapping in the drylands. With the release of more advanced medium resolution imagery such as the upcoming Sentinel-2 archive, uncertainty due to data scarcity might be further reduced. In contrast to the available global coverage products, the CAFC map provides information on prevalent forest types while considering the specific environmental settings of Central Asia. It could serve as a useful basis for standardized forest monitoring across country borders in Central Asia and, for example, carbon reporting to the United Nations Framework Convention on Climate Change (UNFCCC) and Reducing Emissions from Deforestation and Forest Degradation (REDD+). Moreover, our approach to the forest cover map for 2010 might be applicable to historic imagery, thus aiding in the continuous monitoring of forest cover changes in the Central Asian drylands.

Methods

Study area. Landlocked deep inside the Eurasian continent, Central Asia is dominated by a distinctive continental arid to semi-humid climate^{83–85}. The sparsely vegetated deserts (Karakum and Kyzylkum) and the Eurasian Steppe belt stretching from the Caspian Sea to the Tian Shan Mountains cover most of the region, but a strong gradient in temperature and precipitation exists from north to south and from the lowlands to the mountains (see Supplementary Fig. S4). Associated with the diverse climates and terrains, forests in Central Asia consist of various types, with a rich species composition³. Dense needle-leaved evergreen forests are found in the mountains of Kazakhstan, Kyrgyzstan, and Tajikistan. The prevalent tree species include Scots pine (*Pinus sylvestris*), Siberian fir (*Abies sibirica*), Asian spruce (*Picea schrenkiana*), and junipers (*Juniperus* spp.). Broadleaved trees such as Tianshan birch (*Betula tianschanica*) and English oak (*Quercus robur*), and wild fruit and nut bearing species (e.g., walnut [*Juglans regia*], pistachio [*Pistacia vera*], and wild apple [*Malus sieversii*]) are more abundant in the lowlands and the lower mountain zone. Along the large river systems of the Amu Darya and Syr Darya, riparian forests dominated by Euphrates poplar (*Populus euphratica*) provide important corridors for wildlife and harbor a rich biodiversity⁸⁶. Sparse stands of saxaul trees (*Haloxylon* spp.) and various xeric shrubs form “desert forests” in the sandy deserts of Turkmenistan and Uzbekistan¹⁸.

Land cover classification from Landsat imagery. We randomly selected sixteen Landsat footprints covering the principal forest biomes of Central Asia based on the terrestrial ecoregions used by the World Wildlife Fund (WWF)⁸⁷. Two extra footprints (152/31 and 154/31) for which ground data were available were added manually. We downloaded all available precision terrain corrected Landsat TM and ETM + L1T products with less than 100% cloud coverage for the period 2009–2011 for the eighteen Landsat footprints from the U.S. Geological Survey (USGS) (see Supplementary Fig. S4). Detailed information about the Landsat imagery used in this study can be found in Supplementary Table S2.

For each Landsat image, we performed atmospheric correction and radiometric calibration using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) to ensure consistency between different sensors, dates, and footprints⁸⁸. We used an object-based algorithm FMASK to generate cloud/shadow/snow masks for each image⁸⁹. All imagery was then reprojected to a regional Albers equal-area conic projection.

To develop cloud- and gap-free forest cover maps with Landsat data for 2010, we built temporal composites using all available imagery. We tested three different composite types yielding three different sets of predictors (see Supplementary Table S3). The first set of predictors was based on a single date compositing strategy. For each reflectance band, a target day of the year (DOY) in mid-summer (DOY 210) was set and a clear observation closest to the target day was selected for compositing a gap-free clear image. The second set was comprised of all clear observations from 2010 and was used to calculate five metrics for each reflectance band: the mean, median, standard deviation, 25th percentile, and 75th percentile. The third set used the same statistical metrics as the second set, but for an expanded observation period of 2009 to 2011.

To map the forest composition in Central Asia, we classified forest types into deciduous forest, evergreen forest, and mixed forest (in which deciduous or evergreen tree species have a canopy cover of less than 60%). In addition to forests, we also mapped four other land cover classes, including non-woody wetland, herbaceous land (including grassland and cropland), barren land, and water bodies, to gain a deeper understanding of error sources when mapping forests in the drylands. The definition of forest and other land cover types was adapted from the IGBP DISCover dataset⁹⁰. Because tree stands are relatively sparse in the drylands, we defined forest as lands dominated by tree cover greater than 30%.

We used the random forest (RF) classification algorithm⁹¹ to predict land cover for each set of predictors. The number of variables randomly sampled as candidates at each split was set to the square root of the number of input variables, and the minimum sizes of the terminal nodes and the number of the trees were set to 10 and 1000, respectively. Per-pixel class was predicted based on the majority of tree votes for a given class. The package ‘randomForest’⁹² implemented in the statistical software CRAN R⁹³ was employed to conduct the analysis.

The training and validation samples for Landsat land cover mapping were selected independently from references collected in the field, high-resolution images from Google Earth and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), for assessing Landsat imagery and MODIS time series acquired for the period 2009–2011. First, 4389 training polygons were labelled with the aid of high resolution imagery from Google Earth. To guarantee geometric accuracy, we cross-checked the geolocation of the polygons on ASTER imagery to ensure there were no major terrain displacements. Multi-date imagery, such as winter acquisitions together with the spring leaf-onset and autumn leaf-shading imagery (either from Google Earth, ASTER or Landsat), were used to better distinguish deciduous, evergreen, and mixed forests. Second, Landsat pixels within the training polygons were selected as the final training samples. In all, we created 24,238 training pixel samples, with 8,661 labelled as forest.

To avoid smaller samples for smaller classes such as forest or wetland, we used disproportionate stratified sampling at the Landsat pixel level for accuracy assessment⁹⁴. First, we randomly selected 300 Landsat pixels as reference samples within each land cover class from the map produced from the temporal metric 2009–2011. Second, each sample was manually labelled by an expert interpreter, without knowledge of the mapped class label, using high resolution imagery from Google Earth, ASTER, and Landsat images acquired around 2010. In addition to the Landsat images, the interpreter also used temporal profiles of the MODIS NDVI time series (MOD13Q1) to aid in the labelling of samples with data gaps in the Google Earth, ASTER, and Landsat images. For each set of classifications, we constructed an error matrix from which we calculated overall accuracy, commission and omission errors. Finally, we selected the map with the best classification accuracy for use as a reference when predicting forest cover percentage using the MODIS time series.

MODIS preprocessing and fractional forest cover mapping. Two MODIS Vegetation Index (VI) Collection 5 products from Terra (MOD13Q1) and Aqua (MYD13Q1) were used as the main sources for the production of the fractional forest cover map for Central Asia. The two products have a nominal spatial resolution of 250 m and contain 16-day composites of two vegetation indices and four spectral reflectance bands: the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), surface reflectance in the blue, red, near-infrared (NIR), and shortwave-infrared (SWIR) wavelengths. We downloaded all imagery for eight MODIS tiles (h21v03, h22v03, h23v03, h21v04, h22v04, h23v04, h22v05, and h23v05) covering the entire Central Asia land area (see Supplementary Fig. S4) between January and December 2010 from the United States Geological Survey (USGS) Land Processes Distributed Active Archive Center (LPDAAC, available: <https://mrtweb.cr.usgs.gov/>). All imagery was reprojected to an Albers equal-area conic projection.

To increase the observation density, we combined the indices and spectral reflectance acquired from Terra and Aqua to achieve a temporal resolution of 8 days. The combined time series was reconstructed based on the order of the DOY information provided along with the MOD13Q1/MYD13Q1 spectral reflectance. To reduce residual noise in the time series caused by clouds, ozone, dust, off-nadir viewing, and low sun zenith angles, we smoothed the data using a Savitzky-Golay filter in the TIMESAT software⁹⁵. Pixels flagged as no data, snow/ice, or cloud in the MOD13Q1/MYD13Q1 pixel reliability layer were excluded prior to the filtering. Outliers that were not identified as such in the MODIS reliability layer were removed with a seasonal-trend decomposition based on loess (STL)⁹⁶, which is global in character and not dependent on ancillary data. We calculated several seasonal statistics from the MODIS time series to serve as predictor variables. For each smoothed MODIS time series (NDVI, EVI, and the four reflectance bands), we computed the mean, minimum, maximum, range, and standard deviation for spring (March-May), summer (June-August), and autumn (September-November).

The Landsat land cover map served as a source for the generation of training data. First, we randomly selected MODIS pixel samples, during which we enforced a minimum distance of 1000 m between samples to avoid spatial autocorrelation and along-scan triangular point spread function (PSF) effects⁹⁷. Second, we calculated the proportion of each forest type in each sample using the Landsat land cover map. As a result, we obtained 16,173 training samples comprising the percentage cover of each forest type to train the MODIS forest cover model.

We applied Random Forest (RF) regression using the R-package randomForest⁹² implemented in R⁹³ to predict per-pixel forest cover percentage. While RF classification estimates class membership of categorical data by applying a majority vote to individual tree predictions, RF regression predicts a continuous response by averaging individual tree predictions per-pixel. The model was run separately for deciduous, evergreen, and mixed forests. We also summed the forest cover fractions of all forest types to estimate the percentage of overall forest cover for each pixel.

The accuracy of the final forest cover map was assessed based on a stratified random sample using the mapped class proportions to define two strata: pixels with an overall forest proportion >0% (forest) and pixels with an overall forest proportion of 0% (non-forest). Within each stratum, we randomly selected 1000 MODIS pixel samples. Again, we used the Landsat maps to estimate the reference forest cover proportion for each pixel. We then assessed the agreement between the MODIS forest proportions and the Landsat-based proportions of deciduous, evergreen, and mixed forests by estimating the normalized root-mean-square error (NRMSE) between predictions and reference data:

$$NRMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - R_i)^2}{n}} / \bar{R}, \quad (1)$$

where P_i and R_i are the model prediction and reference for forest cover, respectively, for pixel i with a sample size of n , and \bar{R} is the mean value of the reference.

Comparison of tree cover maps in Central Asia. Two global tree cover products were compared with our MODIS forest cover map (CAFC) to determine the strengths and limits of all available maps for Central Asia. We selected the Global Forest Change (GFC) dataset and MODIS Vegetation Continuous Fields (VCF) for these comparisons. GFC is a state-of-the-art global forest change product by Hansen *et al.*⁵² with a 30-meter resolution. It estimates per-pixel tree cover percentage for 2000 and annual forest cover change between 2000 and 2013 using a rigorous data processing chain. To derive tree cover in 2010, we summarized forest changes during 2000 and 2010 and used the results to update the tree cover layer produced for the year 2000. Pixels labelled as forest loss during 2000–2010 were assigned 0% tree cover in 2010. Inversely, pixels mapped as forest gain before 2010 were labelled as 100% tree cover. In all, a number of pixels equal to 2.1% of the area with >25% tree cover in 2000 were re-labelled. The MODIS VCF product was developed by the University of Maryland at a spatial resolution of 250 m²⁹. The inputs for MODIS VCF include the 16-day surface reflectance composites of ten MODIS spectral bands. MODIS VCF Collection 5 (MOD44B) for the year 2010 was downloaded and used for the comparison. Several other global land cover maps provide categorical land cover information at a coarse resolution. However, we only selected continuous tree cover data because sub-pixel proportions are likely to capture heterogeneous and sparse forest cover better than discrete maps at that spatial resolution, and these data are more comparable to our map.

In the absence of information from GFC and MODIS VCF on specific forest types, we only selected the overall forest cover layer from the CAFC for the comparison. GFC was translated into a forest cover map for comparison with the CAFC. MODIS VCF was then compared with GFC without translation into forest cover because of the coarse resolution of MODIS VCF product. We followed this procedure to identify similarities and discrepancies between the forest cover products without losing detail. Comparisons were performed at the same spatial scale. First, all maps were re-projected to an Albers equal-area conic projection. Second, the GFC product was translated to a forest/non-forest map using a tree cover threshold of 30%. Third, we standardized all of the maps originating from the products to a resolution of 250-m by averaging all pixel values within a 250-m sampling unit. Because geolocation errors in remote sensing data and maps can influence per-pixel comparisons, we also

included a comparison of larger pixel blocks by gradually decreasing the spatial resolution of the maps to 1 km and 5 km. Further, we only selected pixels with a tree cover or forest cover greater than 0% according to GFC, MODIS VCF, or CAFC. We also compared per-country forest cover statistics according to CAFC, GFC, MODIS VCF, and FAO FRA⁶⁸. MODIS VCF was converted to a forest/non-forest map by applying a 30% tree-cover threshold before using it for per-country forest coverage calculations.

To improve information on worldwide forest cover, FRA 1990 and FRA 2000 included a remote sensing survey in tropical areas to complement national reporting. The FRA 2010 expanded the previous surveys to a global scale based on a systematic sampling approach with the aim of providing statistics on forest area change and land use dynamics at the regional, biome, and global scales⁹⁸. We took forest data from the FRA 2010 and 2015 reports without adjustment since, according to the individual country reports, the national data had been adjusted and re-classified to comply with the FRA format, as well as because of a lack of details provided for the (re-) classification procedures⁹⁹.

References

1. Eastwood, A., Lazkov, G. & Newton, A. C. *The red list of trees of Central Asia*. (Fauna & Flora International Cambridge, 2009).
2. Turdieva, M. *et al.* In *Climate change and terrestrial carbon sequestration in Central Asia* (eds R. Lal *et al.*) 25–32 (CRC Press 2007).
3. FAO. *People, forests and trees in West and Central Asia: Outlook for 2020*. (Rome, 2007).
4. Akhmadov, K. *Forest and forest products country profile: Tajikistan*. (United Nations Economic Commission for Europe/Food and Agriculture Organization of the United Nations 2008).
5. USAID. *Biodiversity assessment for Central Asia: Regional overview*. (USAID Central Asian Republics Mission, Almaty, 2001).
6. Blaser, J., Carter, E. J. & Gilmour, D. A. Biodiversity and sustainable use of Kyrgyzstan's walnut-fruit forests: Proceedings of the seminar, Arslanbob, Dzalal-abab Oblast, Kyrgyzstan, 4–8 September 1995. (IUCN, 1998).
7. Czudek, R. *Wildlife issues and development prospects in West and Central Asia*. (FAO, Rome, 2005).
8. Li, C. *et al.* Carbon stock and its responses to climate change in Central Asia. *Global Change Biol*, **21**, 1951–1967, doi:10.1111/gcb.12846 (2015).
9. Kushlin, A. V., Schillhorn van Veen, T. W. & Sutton, W. Kazakhstan: forest sector in transition: the resource, the users and sustainable use. (World Bank, Washington, D.C., 2003).
10. Oromzbekov, A. *et al.* Forest rehabilitation in Kyrgyzstan. 131–182 (IUFRO, Vienna, 2009).
11. Botman, E. Forest rehabilitation in the Republic of Uzbekistan. 253–299 (IUFRO, Vienna, 2009).
12. Popov, K. In *Biogeography and Ecology of Turkmenistan Vol. 72 Monographiae Biologicae* (eds V. Fet & K., Atamuradov) Ch. 10, 173–186 (Springer Netherlands, 1994).
13. Romijn, E. *et al.* Assessing change in national forest monitoring capacities of 99 tropical countries. *Forest Ecol Manag* **352**, 109–123, doi:10.1016/j.foreco.2015.06.003 (2015).
14. FAO. *Forests and climate change in Eastern Europe and Central Asia*. 188 (Rome, 2010).
15. De Pauw, E. In *Climate change and terrestrial carbon sequestration in Central Asia* (eds R., Lal *et al.*) 3–24 (CRC Press 2007).
16. FAO. *Global Forest Resources Assessment 2015*. (Food and agriculture organization of the United nations (FAO), 2015).
17. Kirchhoff, J. F. & Fabian, A. *Forestry sector analysis of the republic of Tajikistan*. (Eschborn: Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), 2010).
18. Worbes, M. *et al.* Scope and constraints for tree planting in the irrigated landscapes of the Aral Sea Basin: Case studies in Khorezm Region, Uzbekistan., (ZEF Discussion Papers on Development Policy 112 ZEF (Center for Development Research), the University of Bonn, 2006).
19. Verchot, L. V., Zomer, R., Van Straaten, O. & Muys, B. Implications of country-level decisions on the specification of crown cover in the definition of forests for land area eligible for afforestation and reforestation activities in the CDM. *Climatic Change* **81**, 415–430, doi:10.1007/s10584-006-9111-9 (2007).
20. Sexton, J. O. *et al.* Conservation policy and the measurement of forests. *Nature Clim. Change* **6**, 192–196, doi:10.1038/nclimate2816 (2015).
21. Safriel, U. *et al.* In *Ecosystems and human well-being: current state and trends* (The Millennium Ecosystem Assessment series) (eds R., Hassan, R., Scholes & N., Ash) 623–662 (Island Press, Washington, DC, 2005).
22. Sannier, C., McRoberts, R. E., Fichet, L. V. & Makaga, E. M. K. Using the regression estimator with Landsat data to estimate proportion forest cover and net proportion deforestation in Gabon. *Remote Sens Environ* **151**, 138–148, doi:10.1016/j.rse.2013.09.015 (2014).
23. Potapov, P., Turubanova, S. & Hansen, M. C. Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia. *Remote Sens Environ* **115**, 548–561, doi:10.1016/j.rse.2010.10.001 (2011).
24. Bartholome, E. & Belward, A. S. GLC2000: a new approach to global land cover mapping from Earth observation data. *Int J Remote Sens* **26**, 1959–1977, doi:10.1080/01431160412331291297 (2005).
25. Defourny, P. *et al.* In *Proceedings of the ISPRS Commission VII mid-term symposium, Remote sensing: from pixels to processes*. 8–11 (Enschede, the Netherlands).
26. Friedl, M. A. *et al.* MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens Environ* **114**, 168–182, doi:10.1016/j.rse.2009.08.016 (2010).
27. Friedl, M. A. *et al.* Global land cover mapping from MODIS: algorithms and early results. *Remote Sens Environ* **83**, 287–302, doi:10.1016/S0034-4257(02)00078-0 (2002).
28. GFOI. *Integrating remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: Methods and Guidance from the Global Forest Observations Initiative*. (Geneva, Switzerland, 2014, 2013).
29. DiMiceli, C. M. *et al.* Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000–2010, Collection 5 Percent Tree Cover, University of Maryland, College Park, MD, USA. (2011).
30. Boschetti, L., Flasse, S. P. & Brivio, P. A. Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sens Environ* **91**, 280–292, doi:10.1016/j.rse.2004.02.015 (2004).
31. Leinenkugel, P., Wolters, M. L., Oppelt, N. & Kuenzer, C. Tree cover and forest cover dynamics in the Mekong Basin from 2001 to 2011. *Remote Sens Environ* **158**, 376–392, doi:10.1016/j.rse.2014.10.021 (2015).
32. Fritz, S. & See, L. Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biol* **14**, 1057–1075, doi:10.1111/j.1365-2486.2007.01519.x (2008).
33. Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A. & Schmullius, C. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens Environ* **112**, 2538–2556, doi:10.1016/j.rse.2007.11.013 (2008).
34. Montesano, P. M. *et al.* MODIS tree cover validation for the circumpolar taiga-tundra transition zone. *Remote Sens Environ* **113**, 2130–2141, doi:10.1016/j.rse.2009.05.021 (2009).
35. Tropek, R. *et al.* Comment on “High-resolution global maps of 21st-century forest cover change”. *Science* **344**, 981–981, doi:10.1126/science.1248753 (2014).
36. White, M. A., Shaw, J. D. & Ramsey, R. D. 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**, 2699–2704, doi:10.1080/01431160500080626 (2005).

37. Giri, C., Zhu, Z. L. & Reed, B. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sens Environ* **94**, 123–132, doi:10.1016/j.rse.2004.09.005 (2005).
38. Kuenzer, C., Leinenkugel, P., Vollmuth, M. & Dech, S. Comparing global land-cover products – implications for geoscience applications: an investigation for the trans-boundary Mekong Basin. *International Journal of Remote Sensing* **35**, 2752–2779, doi:10.1080/01431161.2014.890305 (2014).
39. Qin, Y. *et al.* Mapping forests in monsoon Asia with ALOS PALSAR 50-m mosaic images and MODIS imagery in 2010. *Scientific Reports* **6**, 20880, doi:10.1038/srep20880 (2016).
40. Yin, H., Pflugmacher, D., Kennedy, R. E., Sulla-Menashe, D. & Hostert, P. Mapping annual land use and land cover changes using MODIS time series. *Ieee J-Stars* **7**, 3421–3427, doi:10.1109/Jstars.2014.2348411 (2014).
41. Verburg, P. H., Neumann, K. & Nol, L. Challenges in using land use and land cover data for global change studies. *Global Change Biol* **17**, 974–989, doi:10.1111/j.1365-2486.2010.02307.x (2011).
42. Pflugmacher, D. *et al.* Comparison and assessment of coarse resolution land cover maps for Northern Eurasia. *Remote Sensing of Environment* **115**, 3539–3553, doi:10.1016/j.rse.2011.08.016 (2011).
43. Kobayashi, T., Tsend-Ayush, J. & Tateishi, R. A new tree cover percentage map in Eurasia at 500 m resolution using MODIS data. *Remote Sensing* **6**, 209–232, doi:10.3390/Rs6010209 (2014).
44. Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R. & Woodcock, C. E. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sens Environ* **122**, 2–10, doi:10.1016/j.rse.2012.01.010 (2012).
45. Kennedy, R. E. *et al.* Bringing an ecological view of change to Landsat-based remote sensing. *Front Ecol Environ* **12**, 339–346, doi:10.1890/130066 (2014).
46. Gessner, U., Machwitz, M., Conrad, C. & Dech, S. Estimating the fractional cover of growth forms and bare surface in savannas. A multi-resolution approach based on regression tree ensembles. *Remote Sens Environ* **129**, 90–102, doi:10.1016/j.rse.2012.10.026 (2013).
47. Cohen, W. B. & Goward, S. N. Landsat's role in ecological applications of remote sensing. *Bioscience* **54**, 535–545, doi:10.1641/0006-3568(2004)054[0535:Lrieao]2.0.Co;2 (2004).
48. Wolter, P. T., Mladenoff, D. J., Host, G. E. & Crow, T. R. Improved Forest Classification in the Northern Lake-States Using Multitemporal Landsat Imagery. *Photogramm Eng Rem S* **61**, 1129–1143 (1995).
49. Oetter, D. R., Cohen, W. B., Berterretche, M., Maier-Sperger, T. K. & Kennedy, R. E. Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data. *Remote Sens Environ* **76**, 139–155, doi:10.1016/S0034-4257(00)00202-9 (2001).
50. Helmer, E. H. *et al.* Detailed maps of tropical forest types are within reach: Forest tree communities for Trinidad and Tobago mapped with multiseason Landsat and multiseason fine-resolution imagery. *Forest Ecol Manag* **279**, 147–166, doi:10.1016/j.foreco.2012.05.016 (2012).
51. Gong, P. *et al.* Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing* **34**, 2607–2654, doi:10.1080/01431161.2012.748992 (2012).
52. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853, doi:10.1126/science.1244693 (2013).
53. Griffiths, P., van der Linden, S., Kuemmerle, T. & Hostert, P. A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **6**, 2088–2101, doi:10.1109/JSTARS.2012.2228167 (2013).
54. Mueller, H., Rufin, P., Griffiths, P., Siqueira, A. J. B. & Hostert, P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *Remote Sens Environ* **156**, 490–499, doi:10.1016/j.rse.2014.10.014 (2015).
55. Zhong, L. H., Gong, P. & Biging, G. S. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sens Environ* **140**, 1–13, doi:10.1016/j.rse.2013.08.023 (2014).
56. Potapov, P., Hansen, M. C., Stehman, S. V., Loveland, T. R. & Pittman, K. Combining MODIS and Landsat imagery to estimate and map boreal forest cover loss. *Remote Sens Environ* **112**, 3708–3719, doi:10.1016/j.rse.2008.05.006 (2008).
57. Kennedy, P. & Bertolo, F. Mapping sub-pixel forest cover in Europe using AVHRR data and national and regional statistics. *Can J Remote Sens* **28**, 302–321, doi:10.5589/m02-015 (2002).
58. Hansen, M. C. *et al.* A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. *Remote Sens Environ* **112**, 2495–2513, doi:10.1016/j.rse.2007.11.012 (2008).
59. Conrad, C., Lamers, J. P. A., Ibragimov, N., Löw, F. & Martius, C. Analysing irrigated crop rotation patterns in arid Uzbekistan by the means of remote sensing: A case study on post-Soviet agricultural land use. *J Arid Environ* **124**, 150–159, doi:10.1016/j.jaridenv.2015.08.008 (2016).
60. Prishchepov, A. V., Radeloff, V. C., Dubinin, M. & Alcantara, C. The effect of Landsat ETM/ETM plus image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. *Remote Sens Environ* **126**, 195–209, doi:10.1016/j.rse.2012.08.017 (2012).
61. Stehman, S. V., Sohl, T. L. & Loveland, T. R. An evaluation of sampling strategies to improve precision of estimates of gross change in land use and land cover. *International Journal of Remote Sensing* **26**, 4941–4957, doi:10.1080/01431160500222632 (2005).
62. Stehman, S. V., Sohl, T. L. & Loveland, T. R. Statistical sampling to characterize recent United States land-cover change. *Remote Sens Environ* **86**, 517–529, doi:10.1016/S0034-4257(03)00129-9 (2003).
63. DeFries, R. *et al.* Subpixel forest cover in central Africa from multisensor, multitemporal data. *Remote Sens Environ* **60**, 228–246, doi:10.1016/S0034-4257(96)00119-8 (1997).
64. Huang, C. & Townshend, J. R. G. A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover. *International Journal of Remote Sensing* **24**, 75–90, doi:10.1080/01431160110115032 (2003).
65. Foody, G. M. & Cox, D. P. Sub-Pixel Land-Cover Composition Estimation Using a Linear Mixture Model and Fuzzy Membership Functions. *International Journal of Remote Sensing* **15**, 619–631, doi:10.1080/01431169408954100 (1994).
66. Xu, M., Watanachaturaporn, P., Varshney, P. K. & Arora, M. K. Decision tree regression for soft classification of remote sensing data. *Remote Sens Environ* **97**, 322–336, doi:10.1016/j.rse.2005.05.008 (2005).
67. Atkinson, P. M., Cutler, M. E. J. & Lewis, H. Mapping sub-pixel proportional land cover with AVHRR imagery. *International Journal of Remote Sensing* **18**, 917–935, doi:10.1080/014311697218836 (1997).
68. FAO. *Global Forest Resources Assessment 2010: main report*. (Food and agriculture organization of the United nations (FAO), 2010).
69. Banskota, A. *et al.* Forest monitoring using Landsat time series data: A review. *Can J Remote Sens* **40**, 362–384, doi:10.1080/07038992.2014.987376 (2014).
70. Huttich, C. *et al.* Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring. *Remote Sens Environ* **115**, 2445–2459, doi:10.1016/j.rse.2011.05.005 (2011).
71. Morton, D. C. *et al.* Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data. *Remote Sens Environ* **115**, 1706–1720, doi:10.1016/j.rse.2011.03.002 (2011).
72. Tchuente, A. T. K., De Jong, S. M., Roujean, J. L., Favier, C. & Mering, C. Ecosystem mapping at the African continent scale using a hybrid clustering approach based on 1-km resolution multi-annual data from SPOT/VEGETATION. *Remote Sens Environ* **115**, 452–464, doi:10.1016/j.rse.2010.09.015 (2011).
73. Foschi, P. G. A geometric approach to a mixed pixel problem - Detecting subpixel woody vegetation. *Remote Sens Environ* **50**, 317–327, doi:10.1016/0034-4257(94)90081-7 (1994).
74. Griffiths, P. *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 Sens Environ* **151**, 72–88, doi:10.1016/j.rse.2013.04.022 (2014).

75. Wickham, J. D., Stehman, S. V., Fry, J. A., Smith, J. H. & Homer, C. G. Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. *Remote Sens Environ* **114**, 1286–1296, doi:10.1016/j.rse.2010.01.018 (2010).
76. Foody, G. M. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sens Environ* **114**, 2271–2285, doi:10.1016/j.rse.2010.05.003 (2010).
77. Sexton, J. O. *et al.* Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *Int J Digit Earth* **6**, 427–448, doi:10.1080/17538947.2013.786146 (2013).
78. Heiskanen, J. Evaluation of global land cover data sets over the tundra–taiga transition zone in northernmost Finland. *International Journal of Remote Sensing* **29**, 3727–3751, doi:10.1080/01431160701871104 (2008).
79. McRoberts, R. E. Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. *Remote Sens Environ* **114**, 1017–1025, doi:10.1016/j.rse.2009.12.013 (2010).
80. Ståhl, G. *et al.* Use of models in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation. *Forest Ecosystems* **3**, 5, doi:10.1186/s40663-016-0064-9 (2016).
81. Sannier, C., McRoberts, R. E. & Fichtel, L.-V. Suitability of Global Forest Change data to report forest cover estimates at national level in Gabon. *Remote Sens Environ* **173**, 326–338, doi:10.1016/j.rse.2015.10.032 (2016).
82. Atamuradov, A. & Karryeva, S. Global Forest Resources Assessment: Turkmenistan country report. (Rome, 2005).
83. Klein, I., Gessner, U. & Kuenzer, C. Regional land cover mapping and change detection in Central Asia using MODIS time-series. *Appl Geogr* **35**, 219–234, doi:10.1016/j.apgeog.2012.06.016 (2012).
84. de Beurs, K. M., Henebry, G. M., Owsley, B. C. & Sokolik, I. Using multiple remote sensing perspectives to identify and attribute land surface dynamics in Central Asia 2001–2013. *Remote Sens Environ* **170**, 48–61, doi:10.1016/j.rse.2015.08.018 (2015).
85. Propastin, P. & Kappas, M. Modeling Net Ecosystem Exchange for Grassland in Central Kazakhstan by Combining Remote Sensing and Field Data. *Remote Sensing* **1**, 159–183, doi:10.3390/rs1030159 (2009).
86. Thevs, N. *et al.* Structure and wood biomass of near-natural floodplain forests along the Central Asian rivers Tarim and Amu Darya. *Forestry*, 1–10, doi:10.1093/forestry/cpr056 (2011).
87. Olson, D. M. *et al.* Terrestrial Ecoregions of the World: A New Map of Life on Earth A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *Bioscience* **51**, 933–938, doi:10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2 (2001).
88. Masek, J. G. *et al.* A Landsat surface reflectance dataset for North America, 1990–2000. *Geoscience and Remote Sensing Letters, IEEE* **3**, 68–72, doi:10.1109/LGRS.2005.857030 (2006).
89. Zhu, Z. & Woodcock, C. E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens Environ* **118**, 83–94, doi:10.1016/j.rse.2011.10.028 (2012).
90. Belward, A. S., Estes, J. E. & Kline, K. D. The IGBP-DIS global 1-km land-cover data set DISCover: A project overview. *Photogramm Eng Rem S* **65**, 1013–1020 (1999).
91. Breiman, L. Random forests. *Machine learning* **45**, 5–32, doi:10.1023/A:1010933404324 (2001).
92. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R news* **2**, 18–22 (2002).
93. R Development Core Team. R: A language and environment for statistical computing. (Vienna, Austria, 2011).
94. Olofsson, P., Foody, G. M., Stehman, S. V. & Woodcock, C. E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens Environ* **129**, 122–131, doi:10.1016/j.rse.2012.10.031 (2013).
95. Jonsson, P. & Eklundh, L. TIMESAT - a program for analyzing time-series of satellite sensor data. *Comput Geosci-Uk* **30**, 833–845, doi:10.1016/j.cageo.2004.05.006 (2004).
96. Cleveland, R. B., Cleveland, W. S., McRae, J. E. & Terpenning, I. STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics* **6**, 3–73 (1990).
97. Tan, B. *et al.* The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. *Remote Sens Environ* **105**, 98–114, doi:10.1016/j.rse.2006.06.008 (2006).
98. FAO & JRC. Global forest land-use change 1990–2005. *FAO forestry paper* **169** (2012).
99. FAO. Global Forest Resources Assessment 2010 country report: Kazakhstan. (Food and agriculture organization of the United Nations (FAO), 2010).

Acknowledgements

This study was conducted under the framework of the joint ZEF/CIFOR project “Mapping of Forest Cover in Central Asia using Satellite Remote Sensing” funded under the CGIAR Research Program “Forests, Trees, and Agroforestry.” Additional support from the O-Jeong Eco-Resilience Institute (South Korea) and the Robert Bosch Foundation (Germany) is gratefully acknowledged. We thank Mr. Alfredo Jakob, Ms. Susanne Haas, Mr. Alexander Lee, and Mr. Murodjon Sultanov for the training data collection, and Mr. Salikhjan Khamzin, Mr. Nikoljai Lyutsian, Dr. Venera Surappaeva, and Mr. Nursultan Kubatbekov for their support during field validation.

Author Contributions

Y.H. and A.K. designed the research and conducted field work. Y.H. performed the data analysis. All authors co-wrote and co-edited the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-01582-x

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017