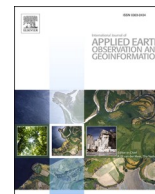


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Mapping fractional woody cover in an extensive semi-arid woodland area at different spatial grains with Sentinel-2 and very high-resolution data

Elham Shafeian^{a,*}, Fabian Ewald Fassnacht^a, Hooman Latifi^{b,c}^a Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany^b Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran 15433-19967, Iran^c Department of Remote Sensing, University of Würzburg, Oswald Kühle Weg 86, 97074 Würzburg, Germany

ARTICLE INFO

Keywords:

Semi-arid woodlands
Zagros
Woody cover
Sentinel-2
VHR imagery
Land cover

ABSTRACT

Woody canopy cover is an essential variable to characterize and monitor vegetation health, carbon accumulation and land-atmosphere exchange processes. Remote sensing-based global woody and forest cover maps are available, yet with varying qualities. In arid and semi-arid areas, existing global products often underestimate the presence of woody cover due to the sparse woody cover and bright soil background. Case studies on smaller regions have shown that a combination of collected field data and medium-to-high resolution free satellite data (e.g., Landsat / Sentinel-2) can provide woody cover estimates with practically-sufficient accuracies. However, most earlier studies focused on comparably small regions and relied on costly field data. Here, we present a fully remote sensing-based work-flow to derive woody cover estimates over an area covering more than 0.5 million km². The work-flow is showcased over the Zagros Mountains, a semi-arid mountain range covering western Iran, the northeast of Iraq and some smaller fraction of southeast Turkey. We use the Google Earth Engine to create homogeneous Sentinel-2 mosaics of the region using data from several years. These data are combined with reference woody cover values derived by a semi-automatic procedure from Google® and Bing® very high resolution (VHR) imagery. Several random forest (RF) models at different spatial grains were trained and at each grain validated with iterative splits of the reference data into training and validation sets (100 repetitions). Best results (considering the trade-off between model performance and spatial detail) were obtained for the model with 40 m spatial grain which showed stable relationships between the VHR-derived reference data and the Sentinel-2 based estimates of woody cover density. The model resulted in median values of coefficient of determination (R²) and RMSE of 0.67 and 0.11, respectively. Our work-flow is potentially also applicable to other arid and semi-arid regions and can contribute to improve currently available global woody cover products, which often perform poorly in semi-arid and arid regions. Comparisons between our woody cover products with common global woody or forest-cover products indicate a clear superiority of our approach. In future studies, these results may be further improved by taking into account regional differences in the drivers of woody-cover patterns along the environmental gradient of the Zagros area.

1. Introduction

Woody vegetation canopy cover (hereafter termed woody cover) is one of the most commonly used structural parameters to describe forest and woodland ecosystems and is an essential biodiversity variable under the class of ecosystem structure (Jongman et al., 2017). Woody cover is defined as the fraction of land that is covered by the vertical projection of trees' and shrubs' canopies relative to the entire area (Gonsamo et al., 2013, Yang et al., 2017). Woody cover is also an important input parameter to describe biosphere ~ atmosphere exchange processes such

as evapotranspiration (Villegas et al., 2015). Further, woody cover can be used to describe the current state of an ecosystem: reductions in woody cover may indicate decline or overexploitation processes in forest and woodland areas (e.g., Wang, Cochrane, 2005) with corresponding losses in biomass and, hence, the sequestered carbon. Increases in woody cover may indicate reduced anthropogenic pressure, for example due to land-use changes (e.g., Baumann et al., 2012) or climate-change induced encroachment of woody species in temperature- or in some cases precipitation-limited ecosystems (e.g., García Criado et al., 2020). In dryland ecosystems, an increase in woody cover may indicate

* Corresponding author at: Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany.
E-mail address: Elham.shafeian@partner.kit.edu (E. Shafeian).

<https://doi.org/10.1016/j.jag.2021.102621>

Received 4 July 2021; Received in revised form 21 October 2021; Accepted 10 November 2021

Available online 23 November 2021

0303-2434/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

unfavorable shrub encroachment processes, which may fragment habitats and endanger biodiversity (e.g., Yang, Crews, 2019). Regular monitoring of woody cover at ecosystem level over large geographical extents is hardly feasible with terrestrial surveys alone. Hence, numerous studies have combined field data with remote sensing data (e.g., Wagenseil, Samimi, 2007) or directly used remote sensing approaches to map forest or woody cover (e.g., Yang, Crews, 2019; Zhang et al., 2019; Nagelkirk and Dahlin, 2020). Woody cover over local to regional extents has for example been estimated from airborne remote sensing data including aerial photographs (e.g., Fadaei et al., 2010; López et al., 2016) and laser scanning data (e.g., Andersen et al., 2005, Lee and Lucas, 2007). While these approaches typically provide realistic woody cover estimates, these data are often restricted in their spatial coverage due to quite high costs and thus may not be available for some parts of the Earth. Therefore, other studies examined freely-available satellite data to estimate forest and woody cover using data from sensors, such as Landsat TM (Rikimaru et al., 2002, Nandy et al., 2003; Deka et al., 2012), Landsat ETM+ (Deka et al., 2012; Joshi et al., 2006), Landsat 8-OLI (Korhonen, 2017) and Sentinel-2 (S2) (e.g., Korhonen et al., 2017; Zhang et al., 2019). As summarized by Yang et al. (2017) most of these studies applied one of three methodical approaches to estimate woody cover. These include pixel-wise/spectral unmixing, physically-based and empirical modeling approaches.

In contrast to empirical methods, the spectral unmixing method does

not necessarily involve field-based measurements, as all required information can principally be derived directly from the image being unmixed (Nagelkirk and Dahlin, 2020). The basic principle of pixel-wise spectral unmixing follows the assumption that the spectrum of each pixel is a (linear) mixture of the spectra of a limited number of so-called endmembers. Each endmember represents the spectral properties of one of the “pure” landcover classes typically occurring in the region (e.g., soil, woody vegetation, rocks). Comparing the spectrum measured for an individual pixel against (mixed) spectra of the endmembers then allows to estimate the percentage cover of each of the corresponding landcover classes in the pixel. In a recent study in presence of field data, spectral unmixing outperformed linear regression for mapping woody cover in a savannah environment, while in turn random forest (RF) outperformed the unmixing approach (Nagelkirk and Dahlin, 2020). The selection of endmembers was stated to be challenging in semi-arid regions due to the varying spectral properties of vegetation and soil over time and space (Nagelkirk and Dahlin, 2020).

The inversion of physically-based models, which simulate the physical relationships between vegetation canopies’ spectral reflectance and a set of parameters including leaf and canopy traits, soil properties and view-geometry parameters has also been used to estimate fractional vegetation cover (FVC), a variable related to woody cover (Baret et al., 2007, Yang et al., 2017). Direct inversion of these models is usually difficult due to the complexity of the physical models that often require

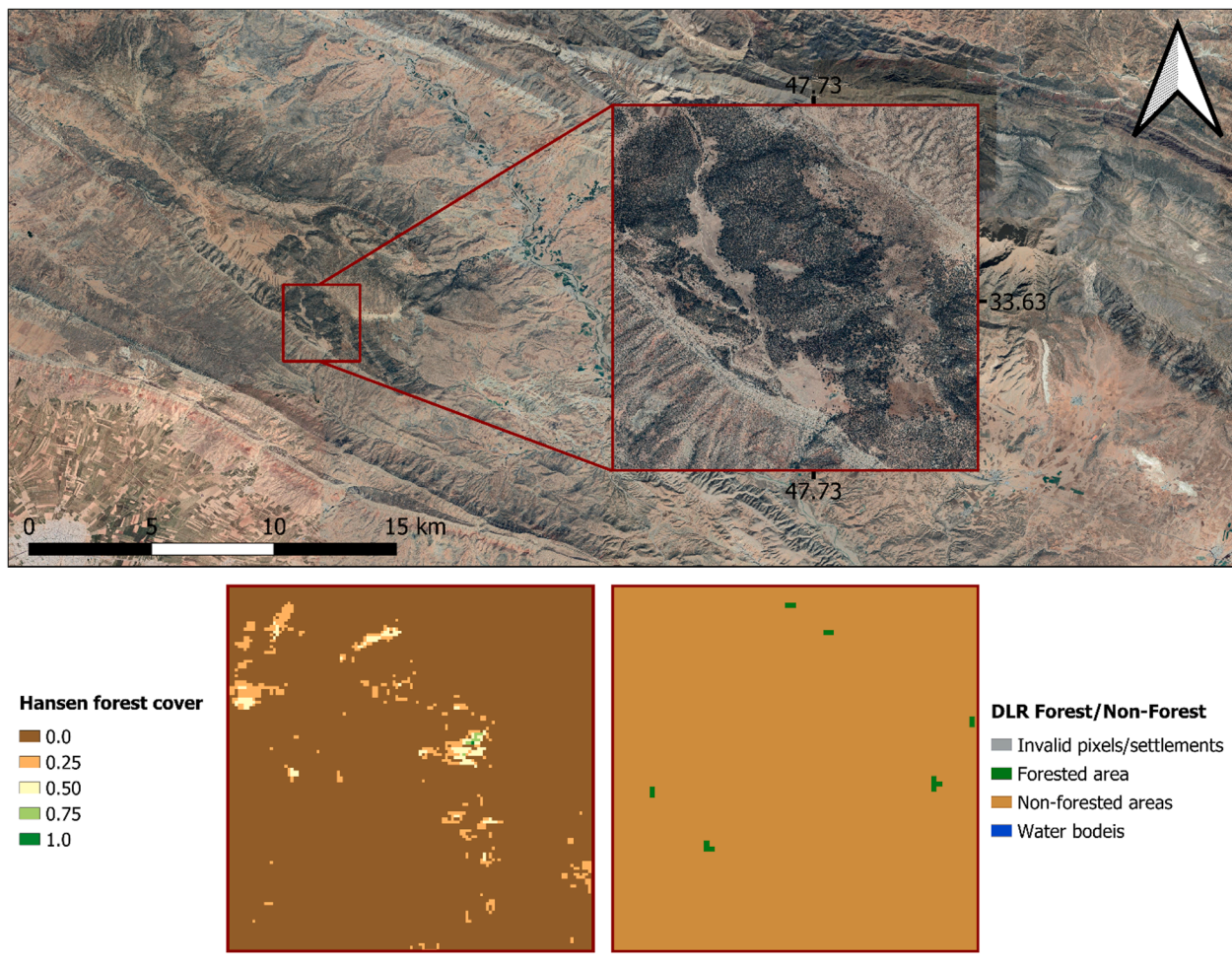


Fig. 1. The top panel shows a high-resolution Google satellite image of a small area in Zagros. The bottom left panel shows the Hansen treecover2000 layer of the global forest cover product. In this area of Zagros, the Hansen Forest cover is very low in most parts even though the high-resolution images show a very high woody cover; In the bottom right panel, the DLR TanDEM-X Forest/Non-Forest (FNF) product for the same area is depicted. Almost the whole area is classified as non-forest as well. (The comparison of this area with the results of our woody cover map, is also shown in the Supplementary Material 1). All images are projected in the geographic coordinate system (EPSG 4326).

the definition of numerous parameters, while corresponding data are limited. The comparably heterogeneous horizontal vegetation structure of semi-arid woodlands may also not match key assumptions of some radiative transfer models. For example, the SAIL model assumes a horizontally homogeneous canopy (Verhoef 1984) which cannot be assumed for most semi-arid woodlands at the spatial scale of Landsat or Sentinel-2. Hence, most studies and data products that follow a physical modelling approach to estimate variables related to vegetation cover were based on coarse resolution data. Using coarse resolution data, some generalizations can be made concerning the parameters required for the inversion procedure and the assumption of a homogeneous canopy cover may become acceptable. For example, Yang et al. (2017) mention products based on MODIS and MERIS data and Baret et al. (2007) used the PROSAIL model to estimate FVC from VEGETATION data at approximately 1.15 km spatial resolution. Contrarily, studies applying physically-based models to estimate vegetation cover at finer resolution are lacking, most likely due to the restrictions outlined above.

Finally, the presumably most common and straightforward approach to estimate woody cover is via empirical models, in which a statistical relationship between vegetation cover and remotely sensed spectral information is established (Yang et al., 2017). Many earlier studies combined woody cover reference values obtained from field data or very high spatial resolution imagery (VHR) with moderate to high spatial resolution satellite data (e.g., Wagenseil, Samimi, 2007; Bucini et al., 2009), via statistical or machine learning models (e.g., Wingate et al., 2019; Liao et al., 2020; Anchang et al., 2020). Machine learning models are nowadays often applied because of their computational efficiency and robustness against noisy data, as well as their ability to estimate multivariate nonlinear relationships (Yang et al., 2017). A general key limitation of empirical models is their commonly-observed restriction to the environmental conditions under which they were established (Schlerf and Atzberger, 2006). As a result, empirical models often cannot be extrapolated to new conditions that were not covered in their calibration data (Tewry and Weiskittel, 2013). Calibration data, in turn, are typically constrained by the availability of field or other reference datasets. On the other hand, empirical models need only basic inputs, are simple to implement and often perform better than physical models and spectral unmixing approaches within a limited dataspace or study area. Hence, they are a viable option if the corresponding reference data demands can be matched. This high data demand of empirical models has become less problematic over the last years. Numerous new satellite sensors deliver continuous streams of satellite data with some of the data being available free of cost (Turner et al., 2015). Also, corresponding reference data become increasingly available through data sharing and direct derivation of reference data from very high-resolution imagery

collected by aerial surveys, unmanned aerial systems or satellites (e.g., Ludwig et al., 2019; Kattenborn et al., 2019; Fassnacht et al., 2021). In summary, empirical models have shown to work well if sufficient reference data are available, i.e. the advantages of empirical models over physical models and unmixing approaches may prevail if large amounts of reference data can be obtained at low cost.

Independent from the applied methodology, there is a tendency of earlier studies examining work-flows to estimate woody cover to focus on temperate (e.g., Heckel et al., 2020), boreal (e.g., Korhonen et al., 2017) and tropical forests (e.g., Waśniewski et al., 2020). Contrarily, the estimation of woody cover from satellite imagery has been less frequently examined in arid and semi-arid regions (Soleimannejad et al., 2018; Yang et al., 2012). In the context of remote sensing, such regions differ from tropical, temperate and boreal forests in both their structural and spectral properties. Dryland ecosystems like the woodlands found in the semi-Mediterranean Zagros area on which we focus in this study are associated with patchy occurrences of trees and a corresponding discontinuous crown cover. In Zagros, these patterns are a result of the joint effects of a longstanding land-use history and the climatic conditions. These patchy and often sparse canopies typically build a strong contrast to the bright soil background (Abdollahnejad, et al., 2019), which notably affects the relation between woody cover and the observed satellite signal.

Depending on the exact location and season in which remote sensing data are acquired, herbaceous vegetation may additionally have a notable influence on the spectral signal (Rautiainen and Heiskanen, 2013; Rautiainen et al., 2011). In arid and semi-arid regions, vegetation growth is mostly limited by water availability, thus vegetation phenology displays pronounced seasonal trends following the precipitation patterns (wet and dry seasons) of the region (Wagenseil, Samimi, 2007). Hence, previous studies suggested that applying multi-temporal intra-annual data can be useful for mapping woody cover in arid and semi-arid regions (Zandler et al., 2015; Wingate et al., 2019; Ludwig et al., 2019). It has also been observed that spectral bands in the visual and near-infrared regions often show a greater contrast between soil and woody vegetation during the dry season due to the lack of herbaceous vegetation (Ludwig et al., 2019). These differences in the spectral properties of arid and semi-arid woodlands forests as compared to boreal, temperate and tropical forests may be one important reason why global forest cover products often perform poorly in arid and semi-arid regions (Cunningham et al., 2019). In particular, underestimations of woody cover are common in these ecosystems (Friedl et al., 2002; Bastin et al., 2017; Cunningham et al., 2019). This can be problematic as these global forest cover products are often used as input to environmental modelling studies and for the planning of forest restoration and

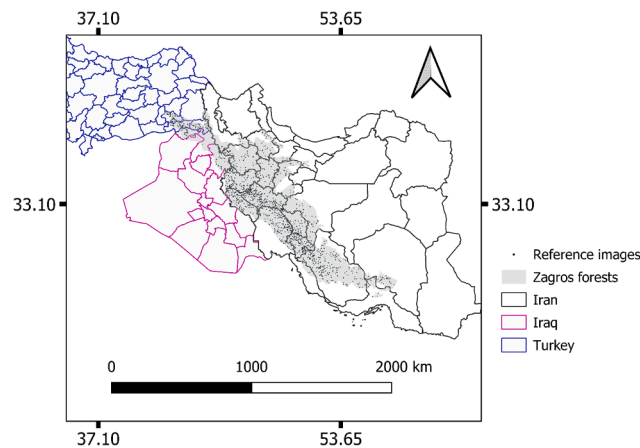


Fig. 2. Location of the Zagros study area. The small black dots indicate the location of the reference images. The colored polygons show the provinces of the three countries comprising parts of the Zagros area.

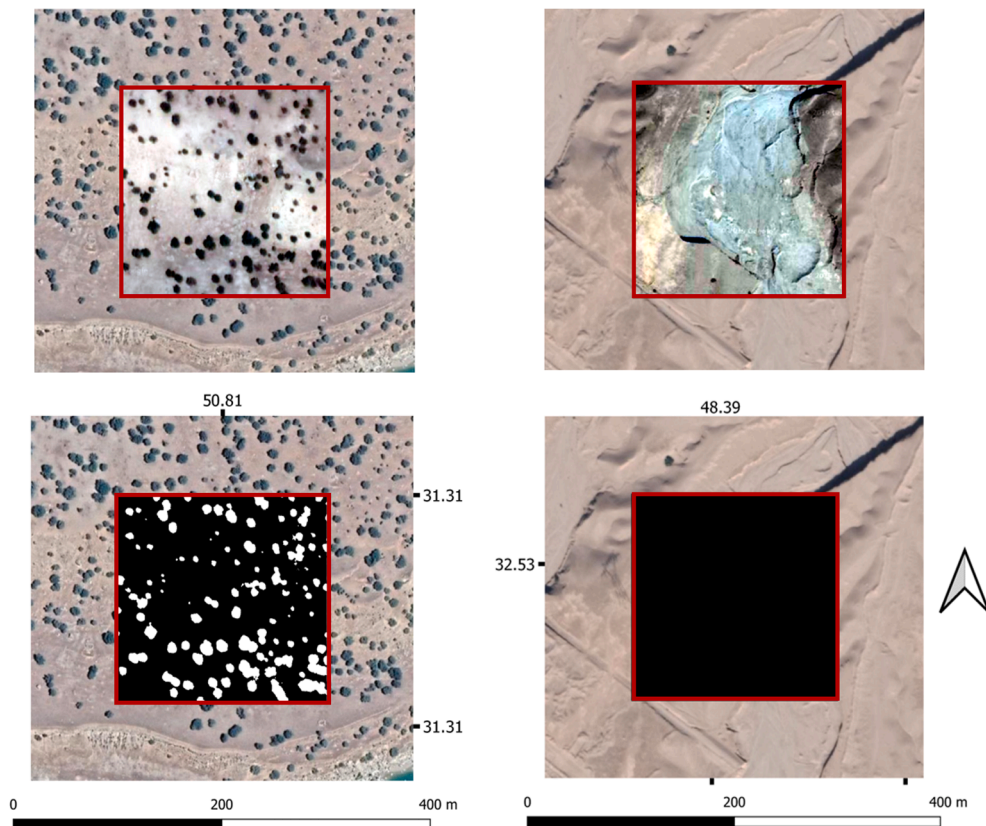


Fig. 3. Preparation of the binary reference images for a wooded area (left panels) and a non-wooded area/bare soil (right panels) based on a simple threshold-based procedure.

afforestation measures, but are rarely scientifically evaluated in these studies (Fagan 2020). Inaccuracies in these products may also affect future decision-making, especially in regions that are associated with data scarcity (Cunningham et al., 2019). This local quality issue of global products is shown for example in Fig. 1 where areas densely stocked with trees are depicted as mostly unforested in two global forest cover products (i.e., the Hansen Global Forest Change map 2000–2014, and the Global TanDEM-X Forest/Non-Forest Map by DLR). These two products show less than 5 percent woody vegetation/forests in most parts of the Zagros area, although in reality there are some areas in which these woodlands have cover-values of up to 80–90 percent (Sagheb-Talebi et al., 2014). It has to be mentioned though, that for example the Hansen global forest cover product is estimated for trees with heights greater than 5 m. These tree heights are not achieved in all areas of Zagros, however, a comparably large fraction of the oak forests of the region have notably higher trees. This is only one example which shows that the accurate quantification of woody cover from satellite imagery over dryland ecosystems with sparse vegetation cover (Soleimannejad et al., 2018, and Yang et al., 2012) and often limited amounts of field reference data (Bai 2010) still remains a challenge.

This challenge has been widely discussed in several recent studies. For example, Fagan (2020) shows that relying on current global forest cover product may lead to a wide overestimation of the afforestation potential in dryland regions as the actual forest cover is notably underestimated. Bastin et al. (2017) show in a study based on an extensive dataset of photo-interpreted forest and woodland plots in drylands (using very high-resolution imagery available in Google Earth) that the actual tree cover of global drylands may be 40–47% higher than reported by global tree cover products. A more general discussion of the challenges of remote sensing approaches in dryland ecosystem can be found in the recent review of Smith et al. (2019).

In this study, we focus on the “Zagros Forests” which spread over a

semi-arid mountain range covering western Iran, the northeast of Iraq and some smaller fraction of southeast Turkey. Baseline information is generally sparse in this enormous area (about 5.5 million ha) and obtaining a woody cover map at fine spatial grain is an important contribution to characterize the forest structure of this vast area. Furthermore, information on woody cover is an important prerequisite for subsequent forest decline analyses from satellite time-series as the comparably large influence of the soil background has to be accounted for (Soleimannejad et al., 2018; Yang et al., 2012). To our knowledge, no study so far has estimated the woody cover of the entire Zagros Forests. Previous studies only examined smaller subset of the Zagros area (e.g., Darvishsefat and Saroei, 2003; Ahmadi Sani et al., 2007; Abdolahi and Shataee Joybari, 2012; Shahvali Kouhshour et al., 2012; Mirzaeizadeh et al., 2015).

The objective of this study is to develop a work-flow to derive woody cover estimates across a very large semi-arid area based on freely available Sentinel-2 satellite data. As a sub-objective, we additionally examine how the spatial grain of the analysis affects the model quality. The latter is interesting for at least two reasons: 1. we expect that at the spatial grain of Sentinel-2, the patchy occurrence of trees and shrubs may complicate the accurate estimation of woody cover and switching to slightly coarser spatial grains may improve the model performance. 2. A slightly coarser grain may additionally reduce variability introduced by sub-pixel spatial shifts between reference data and Sentinel-2 data. The present work-flow is potentially also applicable to other arid and semi-arid forest and woodland areas and can contribute to improve currently available global forest cover products, which often perform poorly across areas dominated by sparse woody vegetation on a bright background soil.

2. Materials and methods

2.1. Study area

Our study area includes the entire Zagros Forests covering more than 5.5 million ha across a length of approximately 1300 km (Khabazi, 2020; Rahimi, et al., 2020). A third of the total population of Iran lives in this region (Pourmoghadam et al., 2013). The Zagros area encompasses more than 42 percent of Iran's forests. The area spreads from the southern part of West Azerbaijan Province to the Fars Province (Mahdavi, et al., 2014) and expands to neighboring countries (northern Iraq and southeastern Turkey) (Fig. 2). The climate in the Zagros area is semi-arid with a mean annual precipitation from 250 to 800 mm and a mean annual temperature from 9 to 25° C (Attarod et al., 2016). The woodlands in Zagros are associated with patchy occurrences of trees and a corresponding discontinuous crown cover. The most dominant and frequent tree species include: Brant's Oak (*Quercus brantii* var. *persica*) that are partly mixed with *Quercus infectoria* G.Olivier, *Quercus libani* G.Olivier, wild Pistachio (*Pistacia atlantica* Desf.) and *Acer monspessulanum* L., *Crataegus* spp., *Amygdalus* spp. and *Pyrus* spp. (Erfanifard et al., 2014). The prevailing mixtures depend on multiple site factors along the latitudinal gradient. The Zagros Forests play an important role in maintaining endemic vegetation and wildlife habitats, providing forage and shelter for nomadic and rural populations and their cattle, as well as non-wood forest products like natural gum, oak seeds and oak gall (Jazirei, Ebrahimi-Rostaghi, 2003; Sagheb-Talebi et al., 2014). Over the last decades, the Zagros area has been deeply affected by environmental and political processes such as the occurrence of weather extremes, the 8-year Iran-Iraq war (1980–1988), and the still ongoing oak dieback (Alibakhshi et al., 2019; Arsalani et al., 2018) which have all contributed to a decline of the ecosystems in Zagros.

2.2. Remote sensing data

This research is based on two main data sources: multispectral S2 satellite data; Google Earth and Bing very high-resolution satellite images.

Reference data preparation for woody cover estimation

A large woody cover reference data set was collected using very high-resolution Google Earth and Bing imagery. The most recent acquisition dates for VHR imagery were 2015 and 2016 and the spatial resolution of Google Earth and Bing imagery ranges from 15 m to 15 cm. In this study we did not use images with less than 1.5 m spatial resolution to reduce uncertainty in the reference dataset. To compose our reference dataset, a total of 8000 points were randomly sampled inside the Zagros area and a square buffer with 200 × 200 m area was generated around each point. To create these buffers, we used the `gBuffer` function available in the `rgeos` library (Bivand et al., 2020) in R (R Development Core Team, 2021). Then, the corresponding Google Earth and Bing images were visually screened in QGIS to allocate the Google and Bing images at each location to one of the three classes ("Wooded area = 1", "No wooded area = 0" and "Non-interpretable area" = 99) according to the area within each buffer. The Non-interpretable areas mostly referred to image scenes of low quality (e.g., coarse spatial resolution, blurry images, image dominated by dark shadows) in which the woody cover could not clearly be identified. Non-interpretable areas were excluded and the remaining images (in total 1270 images) composed the basis of our reference dataset. The reference samples were well distributed over the entire Zagros area (Fig. 2) and had an average distance of approximately 8.8 km from each other.

All the selected reference images were saved as geo-located screenshots in QGIS and cropped to the size of the squared buffers of 200 ×

200 m. Then, a semi-automatic threshold-based approach was applied in R using raster and `Rgdal` libraries to prepare a binary dataset indicating the presence and absence of woody cover in each pixel of the high-resolution reference images. For this, we systematically applied several thresholds (typically ranging from 50 to 150 while the radiometric resolution of the screenshots was 8bit with values from 0 to 255) to the blue channel of the RGB images until the resulting binary image agreed well with the visual impression of the woody cover presence. Due to the very dark appearance of the woody vegetation in front of the bright soil background, this procedure was straightforward and reliable. This procedure resulted in high-resolution binary images indicating areas of woody cover and bare soil (Fig. 3). From these images, the fractional woody cover in percent could be easily obtained for any subset of the high-resolution reference image by dividing the number of woody cover pixels by the total number of pixels. For each RF model (see Fig. 3) we automatically produced the woody cover reference values from the high-resolution binary reference images by overlapping them with a polygon/vector file corresponding to the pixel size and selected grain of S2, as suggested by previous studies (e.g., Higginbottom et al., 2018). All reference images can be accessed via a link provided in the Supplementary Material 5.

Preparation of S2 mosaics and predictor variables

We applied S2 Surface Reflectance imagery available in the Google Earth Engine (GEE) platform (<https://earthengine.google.org/>) data catalogue – detailed description of this data product and the applied radiometric, geometric and atmospheric corrections can be found at the GEE webpage (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR, last access 11th of October 2021). These data have a 5-day temporal resolution. For creating cloud-free mosaics, we first defined a filter to only consider S2 images with a maximum cloud cover of 30%. Then, from all remaining images acquired in the time-period between 2018 and 2020 and during the summer months (June - September) we calculated a composite mosaic image directly in the GEE. We did this by first using the quality assessment band (QA60) to eliminate cloud and cloud shadow affected pixels for each of the S2 scenes (Gao et al., 2017; Zhang et al., 2020). Then we calculated the median-reflectance value for each pixel and band from the remaining cloud-free image stack, which resulted in a cloud-free high-quality mosaic. A visual screening of the mosaic showed that no recognizable data artefacts were present. In preliminary analyses, we also examined the potential benefits of adding S2 mosaics of other seasons, but these did not improve the results. For the sake of parsimony, we hence focused on the summer-mosaics only.

In addition to the 10 S2 bands with 10- and 20-meter resolutions, we calculated the normalized difference vegetation index (NDVI) and the reduced simple ratio (RSR) index. We included the RSR as it was found to decrease background effects in areas with open canopy cover in earlier studies (e.g.; Brown et al., 2000; Zhu et al., 2010). Furthermore, a binary vegetation layer predictor variable was produced by applying a threshold of 0.17 to the S2 NDVI image which resulted in a binary image of vegetated and non-vegetated areas. Then, the fraction of vegetation pixels at 10 m pixel size within the reference areas was determined and used as predictor variable. The relatively low threshold was selected due to generally lower NDVI values in semi-arid regions such as the Zagros Forests (Eskandari et al., 2020). To account for potential influences of the large environmental gradient covered by the Zagros area on the relationship between remote sensing data and woody cover, we also included Bioclim variables (BIO12 = Annual Precipitation, BIO15 = Precipitation Seasonality (Coefficient of Variation), BIO16 = Precipitation of Wettest Quarter, BIO17 = Precipitation of Driest Quarter) (www.worldclim.org) from the GEE data catalogue in our woody cover models. Furthermore, we calculated grey-level co-occurrence texture metrics

1- Reference data preparation from VHR images

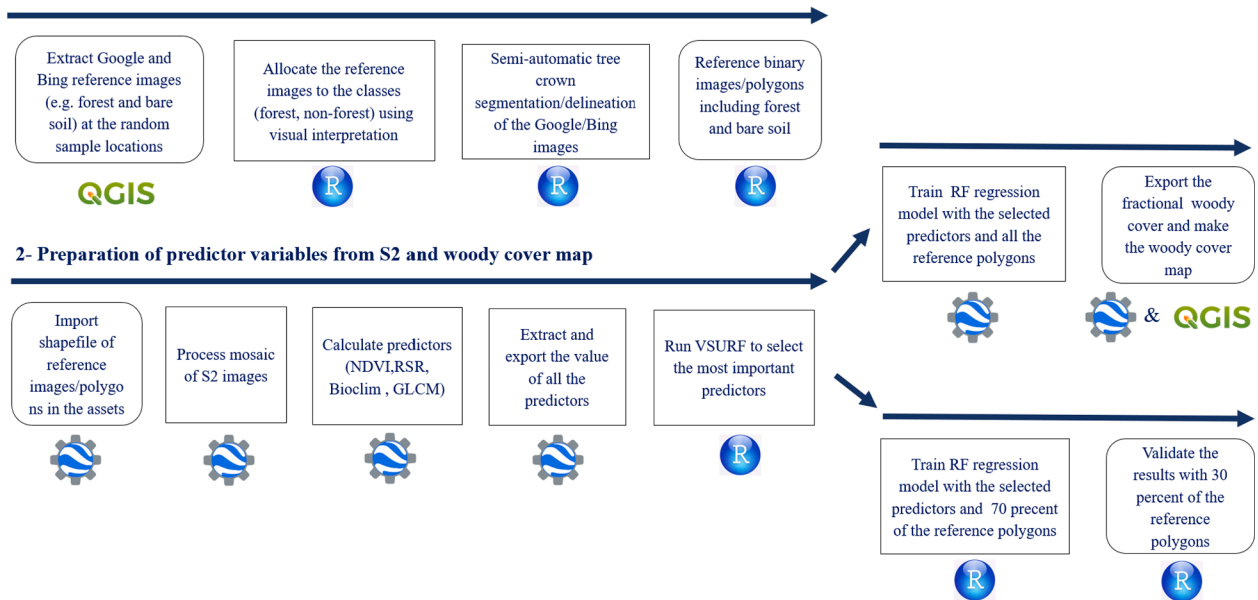


Fig. 4. The applied work-flow for woody cover mapping, the arrows represent the direction of the process.

(GLCM) using the Near Infrared band - NIR (Band 8) of S2 and the NDVI to also consider spatial context (Wood et al., 2012). GLCM texture variables were previously stated to be a strong proxy for vegetation structure (Wood et al., 2012), and were found to be beneficial for vegetation-related remote sensing analyses in many earlier studies (e.g., Coburn and Roberts, 2004; Dobrowski et al., 2008). We calculated texture metrics in the GEE for neighborhoods of 3×3 and 5×5 pixels using the “gcmTexture” function and selected the metrics entropy, contrast, asymilarity/dissimilarity, correlation and variance. The selection of metrics was based on the recommendations of Deur et al. (2020).

2.3. Methods

The methodical work-flow for the study is summarized in Fig. 4 and explained with more details below.

The Google Earth Engine codes applied in this study can be found in the Supplementary Material 5.

Rf algorithm

We used the decision-tree ensemble method RF (Breiman, 2001) to map the canopy cover and land cover of the study area. For a summary of some earlier applications of RF in the context of remote sensing we refer to Belgiu and Drăguț (2016). The benefits of RF over traditional classification and regression methods include its ability to manage

Table 1

The VSURF selected predictors for woody cover estimation selected from the complete set of 40 predictors. The numbers represent the rank at which the corresponding variable was selected in the model of the corresponding grain size (columns).

Grain [m]	10	20	30	40	50	60	70	80	90	100	110	120
Predictor												
B2	6	5	6	6	6	5	6	6	6	6	6	6
B3	5	4	4	5	5	4	5	5	5	5	5	4
B4	2	2	3	3	3	3	2	2	3	2	2	3
B5	4	6		4	4		4	4	4	4	4	
B6			12									12
B7		11	11	13	13	12			13			11
B8	11			11	11		11	11	11	11	11	
B8a						10						
B12						11						
NDVI	1	1	1	1	1		1	1	1	1	1	1
thres	3	3	2	2	2	2	3	3	2	3	3	2
RSR	7	7	5	7	7		7	7	7	7	7	5
NDVI_contrast	9	8	7	9	9	7	9	9	9	9	9	7
NDVI_corr_1			13									13
NDVI_asm_1	15			14	14		15	15	13	15	14	
NDVI_texture55						14						
NDVI_entropy												
NDVI_var						13						
NDVI_ent_1						16						
B8_contrast			14									14
B8_ent_1	14	12					14	14		14	13	
B8_var	16						16	16		16	15	
bioclim 12	10	9	8	10	10	15	10	10	10	10	10	8
bioclim 15	8	10	10	8	8	9	8	8	8	8	8	10
bioclim 16	12		9	12	12	8	12	12	12	12	11	9
bioclim 17	13		13	15	15	9	13	13	14	13	12	13

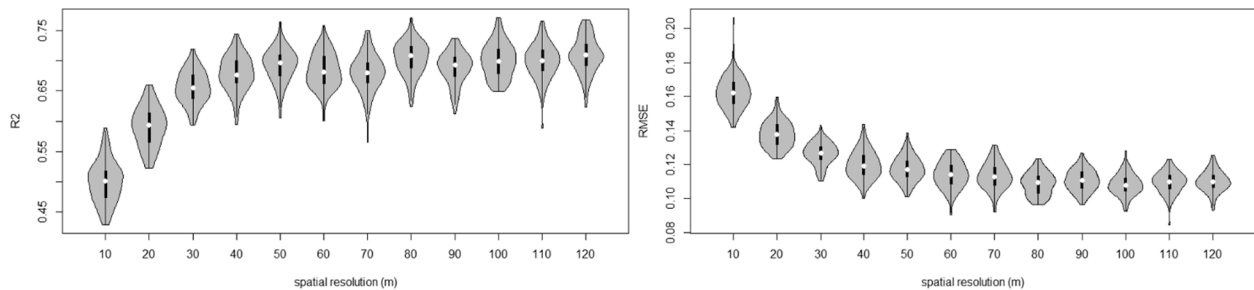


Fig. 5. From left to right, the development of coefficient of determination (R2) and RMSE with decreasing spatial grain for the RF models estimating woody cover are illustrated. The grey-shaded area in the violin plots show the distribution of values, the white point in each violin plot shows the median value.

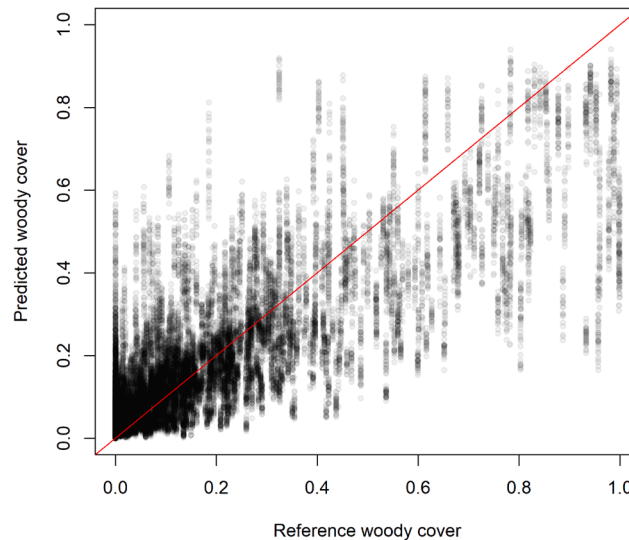


Fig. 6. Scatter plot between observed and predicted woody cover estimates using the RF model at 40 m spatial grain. Results of all 100 model runs conducted in R are depicted.

numerous input variables, to run effectively on large datasets, to be less susceptible to noisy input data and outliers and to be hardly prone to overfitting (Mellor et al., 2015; Gislason et al., 2006). RF is directly available in the GEE (Gaffarian et al., 2020). The availability of RF in GEE is particularly valuable as image processing on local computers would require a lot of memory, a long processing time, and a lot of storage space for our extensive study area.

Variable selection, modeling and multi-scale analysis

RF models were trained to predict the woody cover reference values using the 40 predictor variables described above and listed in [Supplementary Material 2](#). We ran multiple RF models trained at different spatial grains. We started with the original S2 resolution of 10 m and then reduced the resolution with 10 m steps. In case of the model with 10 m resolution, we resampled the 20 m bands of S2 to 10 m in the GEE. This resulted in datasets with a pixel size of 10, 20, 30, 40, ... up to 120 m. We ran RF models separately for each spatial grain. For each model, we automatically produced the woody cover reference values from the high-resolution woody cover- delineated RGB reference images by overlapping them with a polygon/vector file corresponding to the S2 pixel size of the current model, and then counting the number of pixels in the high-resolution binary reference images that were assigned to the woody cover class. The corresponding fractional woody cover values were attached to the reference polygons. This step was conducted in R using the raster (Hijmans and Etten, 2012) and rgdal (Bivand et al., 2021) packages. After obtaining the reference values, the polygon vector file was imported to GEE. Then, mean values of all predictor variables

for the pixels covered by the polygons were obtained and exported as csv file using GEE. The extracted values were then imported to R to run a feature selection using the “three-step variable selection using random forests” (VSURF) package (Genuer et al., 2015). The VSURF method is based on the RF model and is designed to manage high-dimensional data. The VSURF package is highly scalable and can be used to select features in both, regression and supervised classification problems. To perform VSURF, the first step is to exclude all non-essential variables from the dataset. The second step is to pick all variables that are correlated to the response and contribute to a good model performance. For prediction purposes, the third step refines the collection by removing redundancy in the range of variables chosen in the second step (Genuer et al., 2015). The selected predictor variables for each spatial grain are reported in [Table 1](#).

Using the VSURF-selected predictors, we then trained a RF model (ee.Classifier.smileRandomForest function with 500 trees) to predict woody cover across the entire Zagros area. This part of the analysis was again conducted in GEE. For model training in GEE, we used all available reference data to maximize the information content in the model used to create the prediction map of woody cover density. To further validate our models, we ran additional RF models in R by iteratively splitting the whole reference data sets into 70% training and 30% validation samples. Optimal mtry parameter for the models in R were selected using the tuneRF-function of the randomForest-package. The tuneRF function was run with settings: ntreeTry = 50, stepFactor = 2, improve = 0.05, trace = TRUE, doBest = TRUE. For the tuned models,

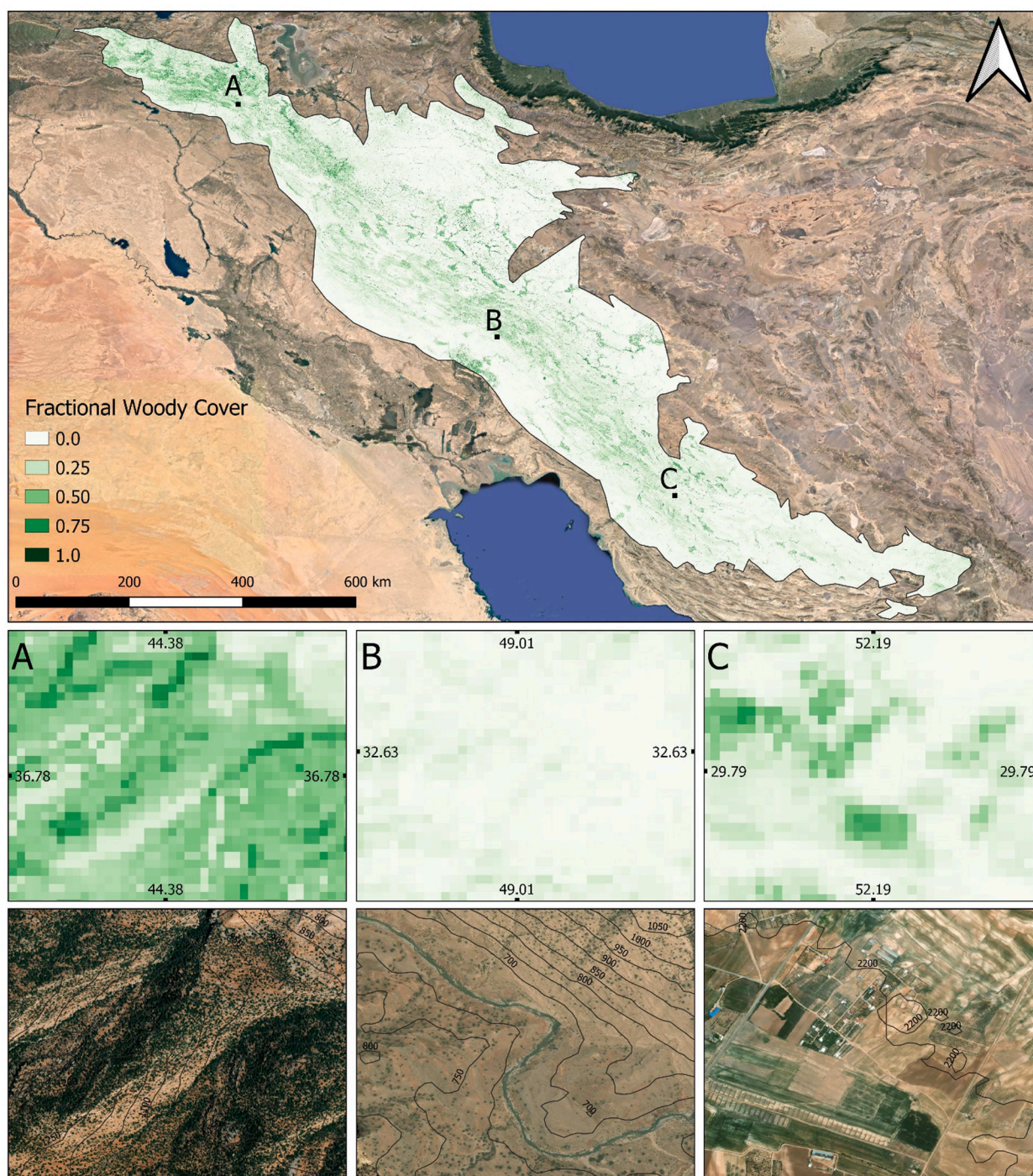


Fig. 7. Map of woody cover estimates for the entire Zagros area. The three subsets A, B and C show two focus areas in topographically complex regions and one focus area with buildings and agricultural fields. The upper row shows the S2-based woody cover map and the bottom row the corresponding areas of high-resolution RGB imagery. Contour lines showing the elevation above sea level are shown on top of the RGB images.

we calculated RMSE and coefficient of determination (R^2) values based on the validation samples for each of the 100 model runs.

2.3.1. Land cover

Our woody cover predictions were created for the entire Zagros area, which also includes non-forested areas. Hence, a land cover map was required to exclude other land cover classes such as built-up, water, and agricultural areas. A recent country-wide land cover map of Iran is available (Ghorbanian et al., 2020) and some other local products were provided by the provincial branches of the Research Institute of Forests

and Rangelands, yet based on varied datasets and unclear methodology. Initial visual interpretation of those maps showed that they do not appropriately depict the wooded areas in our study region. Almost the entire woodland areas in Zagros are classified as rangeland in the corresponding products (e.g., Ghorbanian et al., 2020) but also other land cover types were mis-classified in the available maps. Hence, a more suitable land cover product was produced within this study. For this, we first prepared S2 mosaics following the work-flow described above but containing all seasons. Using the images of different seasons may improve the land cover classification results, especially when classifying

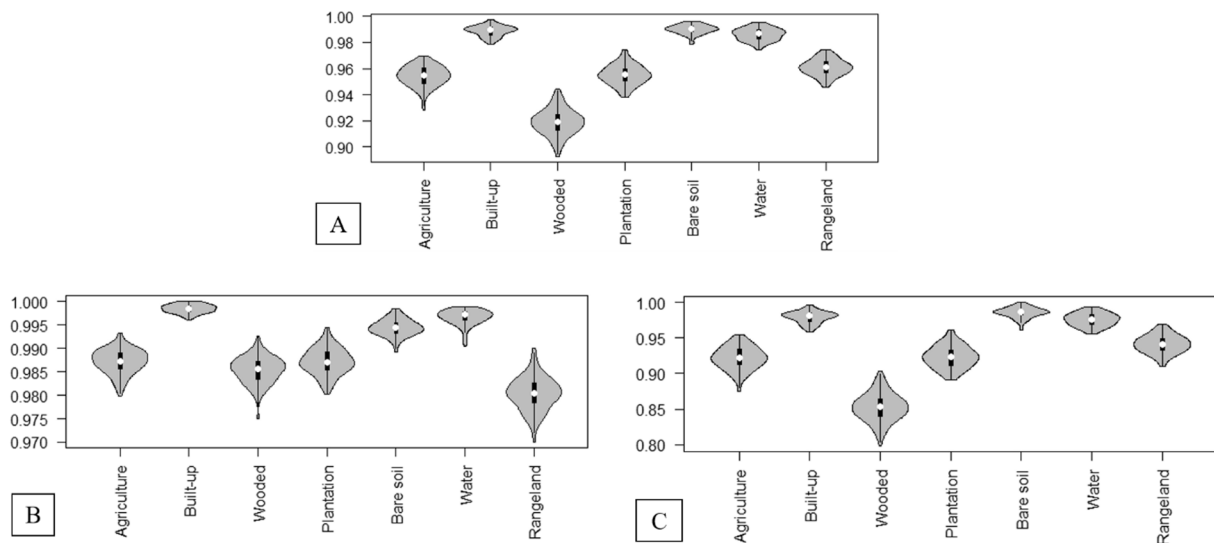


Fig. 8. (A) Balanced accuracy, (B) specificity and (C) sensitivity of the land cover classes in the supervised classification.

agricultural areas (Wingate et al., 2019; Zandler et al., 2015). Thus, we used mosaics based on the S2 images of the last three years including spring (March, April, May), summer (June, July, August) and fall (September, October, November) seasons. We excluded winter as the corresponding mosaics were more frequently affected by data artefacts from clouds and snow. For each season we calculated a grey-scale image from the S2 image using equation (1). Texture variables for each of these greyscale images were calculated using the `gcmTexture` function in GEE. We used a neighborhood size of 2 and a kernel size of 3×3 and all available 21 texture metrics were used as predictors (e.g., entropy, contrast, asimilarity, correlation, etc.).

$$\text{GreyScaleImage} = (0.3 * \text{S2Band8}) + (0.59 * \text{S2Band4}) + (0.11 * \text{S2Band3}) \quad (1)$$

We further used additional vegetation indices (e.g., Soil Adjusted Vegetation Index (SAVI), Normalized Difference Tillage Index (NDTI) introduced by Van Deventer et al (1997), Enhanced Vegetation Index (EVI)), SRTM digital elevation data (Farr et al., 2007) and predictor variables from Sentinel-1 images. We used Sentinel-1C-band Ground Range Detected SAR data from GEE (ImageCollection ID: COPERNICUS/S1 GRD) at 10 m pixel size for spring (March, April, May), summer (June, July, August) and fall (September, October, November) seasons between 2018 and 2020. To construct a homogeneous subset of Sentinel-1 data, metadata properties were used to filter the S1 data collection with the following setting: Polarization of the transmitter and receiver: ['VV', 'VH'], instrument Mode of "IW" (Interferometric Wide Swath and orbit Properties pass: "DESCENDING"). Then we calculated a composite mosaic image directly in the GEE obtaining median-backscatter values for each pixel.

We defined seven land cover classes for the supervised classification including built-up, plantations (orchards), agriculture, water, wooded area, rangeland and bare soil. We collected approximately 1000 reference data points per class via visual interpretation of high-resolution Google and Bing imagery in GEE. 70 percent of these points were used to train a RF model in the classification mode to classify the stacked image for the entire Zagros area. The point-based training for RF classifier was stated to result in more accurate classification compared to polygon-based training (Corcoran et al., 2015). We used an equal number of training points for each class, since RF tends to classify more pixels to the classes with a larger number of training points (e.g., Latifi et al., 2015). Analogously to the woody cover models, we used the VSURF package to select the most important predictors to improve the model. In the final model, 25 important variables were selected by

VSURF (see Supplementary Material 3). Then after adding the important variables, we ran the model by splitting the whole reference data set to 70 % training and 30% validation samples to create the prediction map. We ran additional RF models to validate our model by iteratively splitting the whole reference data set to 70 % training and 30% validation samples following the same approach outlined above for the woody cover models. We calculated confusion matrices from the 100 iterations and report the accuracy of the classification via overall accuracy, kappa and class-specific accuracies (sensitivity, specificity, balanced accuracy).

2.3.2. Woody cover histogram

As final step, we created 25,000 random points across the entire Zagros area and at the location of each random point, we extracted the estimated woody cover as well as the landcover class assigned to the point. These data were then used to depict a histogram of the woody cover percentages across the entire Zagros area.

3. Results

3.1. Woody cover map

Fig. 5 shows violin plots for the model performance metrics obtained during the iterative evaluations of the RF models for estimating woody cover. The results notably improve from the highest resolution at 10×10 m until 40×40 m but then improvements slowly level off and remain approximately constant until they reach the coarsest spatial grain of 120×120 m. As the results hardly improved at resolutions coarser than 40 m, we prepared the predictions maps at 40 m spatial resolution as a finer spatial grain was considered beneficial.

For the selected model with 40 m resolution, coefficient of determination (R^2) values range between 0.56 and 0.74 with a median value of 0.67. The RMSE values range between approximately 0.10 and 0.14 with a median of 0.12.

The corresponding scatter plot of the RF model with 40 m pixel size indicates that the model works reasonably well but also that a quite notable amount of variation in the reference data cannot be explained. Particularly for higher cover values the model only works moderately well (Fig. 6). The tendency of RF to underestimate high and overestimate low values is also apparent.

The final woody cover map for the entire Zagros area is depicted in Fig. 7. High woody cover values are particularly visible for the Northern, more humid parts of Zagros and for the central mountain regions with

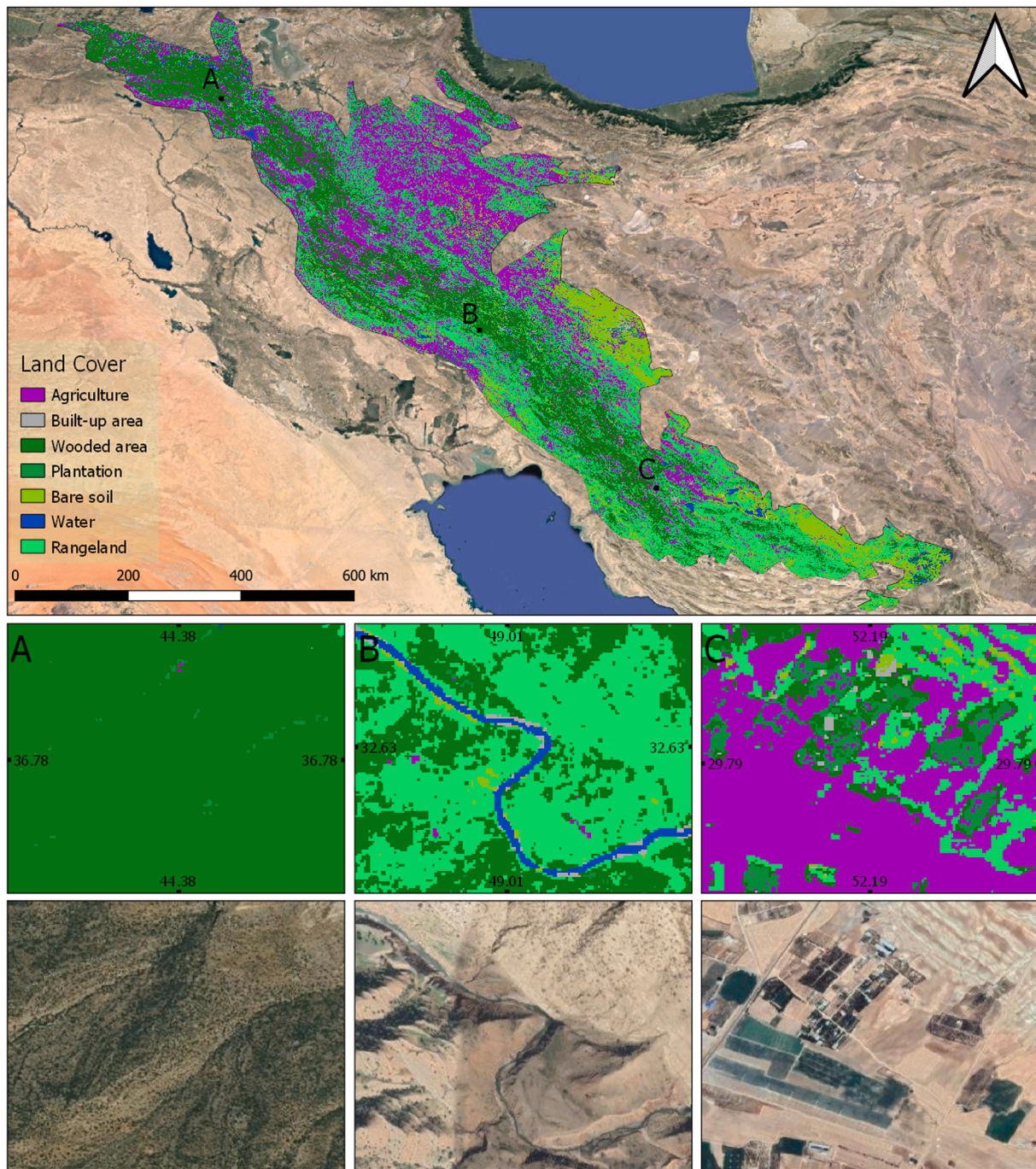


Fig. 9. Map of land cover estimates for the entire Zagros area. The panels A, B and C show focus areas of the S2-based land cover map and below them, the corresponding areas of high-resolution RGB imagery (classes with non-green colors were masked out in the final woody cover product). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

higher elevations. This agrees very well with the existing knowledge on the cover distribution in the region.

The three focal areas of the woody cover map (in Fig. 7) show that the developed woody cover product is able to depict varied cover-situations quite well, even though areas with over- and underestimated cover values are also clearly apparent. Judging from the displayed contour-maps, no obvious influence of topography on the woody cover estimates can be observed in the depicted focal area maps.

The selected predictors that were applied in the final RF models at each spatial grain are summarized in Table.1. The selected variables are overall very stable across the examined spatial grains. In almost all models, most S2-bands were selected except for the SWIR bands B10-

B12 and the red-edge band B6. The B8a band is also only selected in one model, while for all other models B8 (which includes the spectral region of B8a) was selected. Besides the original bands, the two examined indices NDVI and RSR as well as the threshold variable based on NDVI were chosen in almost all models, whereby NDVI and the threshold variable were always amongst the first ranked predictors. Additionally, two NDVI-based texture metrics (NDVI contrast and NDVI asimilarity) were frequently selected. All four bioclim variables were selected in almost all models, even though mostly at higher ranks.

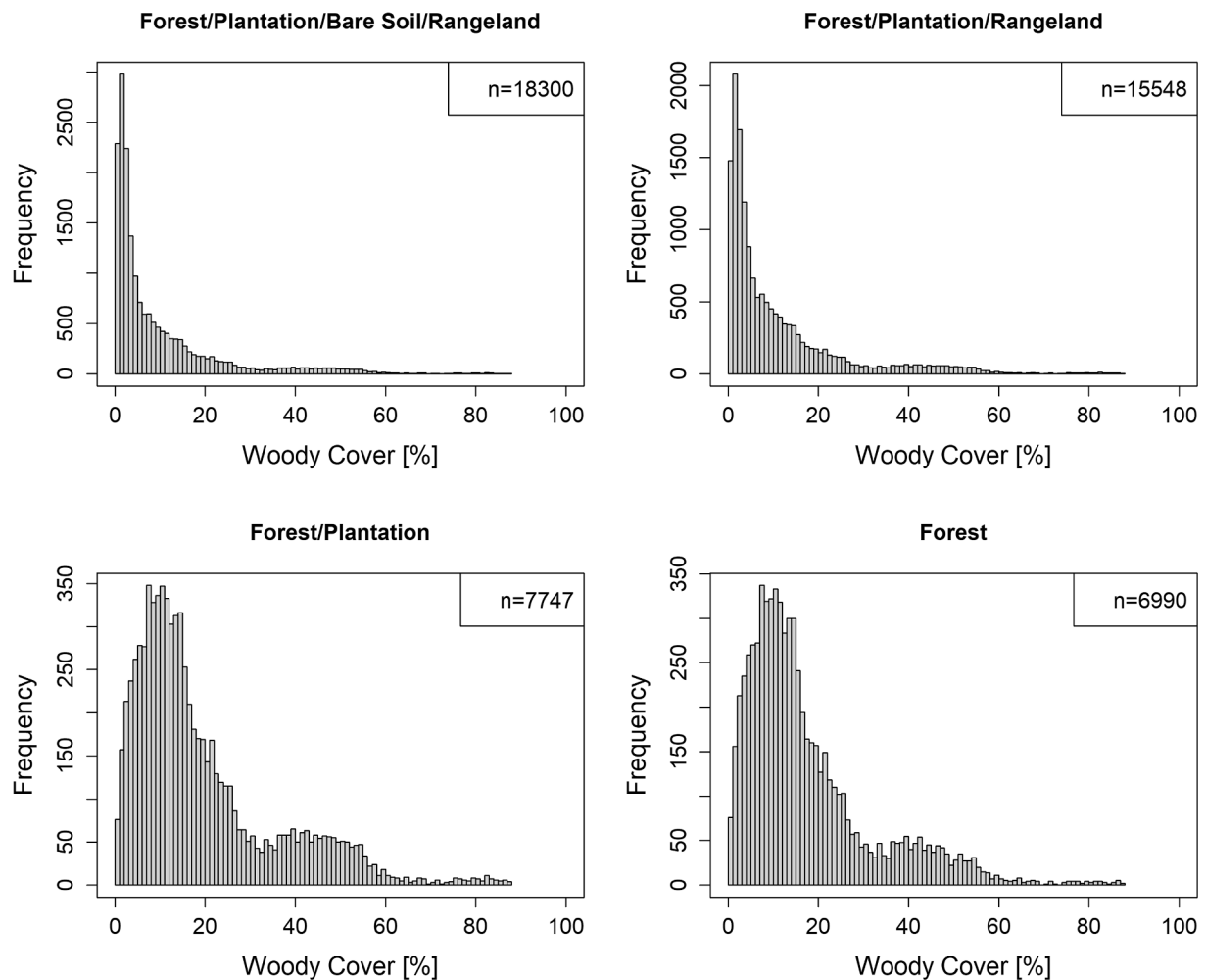


Fig. 10. Histograms of woody cover estimates for n random points across the complete Zagros area and for different combinations of land-cover types as obtained from the land cover classification.

3.2. Land cover map

Fig. 8 displays a summary of the validation metrics of the land cover classification. The highest sensitivity was gained by bare soil and built-up classes with median values of both > 0.98 . Lowest sensitivity values (0.85 and 0.92) were observed for the wooded area and agriculture classes, respectively. The highest specificity values were obtained for water, built-up classes and bare soil (all three > 0.99) and the lowest specificity values for rangeland, wooded area, agriculture and plantation, respectively (approximately 0.98).

The corresponding trends are also mirrored in the balanced accuracies which were again highest for the bare soil with median value of 0.99, built-up and water with median values of more than 0.98 and lowest for the wooded area and agriculture with approximately 0.91 and 0.95, respectively. All metrics were comparably stable throughout the 100 iterations which indicates that the number of training samples were sufficient. The land cover classification model reached an overall accuracy and a kappa value of more than 0.94 and 0.93, respectively.

The final S2-based land cover map for the entire study area and some focal area maps with corresponding high-resolution RGB imagery are shown in Fig. 9. The dark-green belt of “wooded areas” reaching from the North-West to the South-East of the study area agrees very well with the fractional woody cover map (Fig. 7). The huge agricultural areas in the North of Zagros are also clearly visible, while agricultural areas are more sparsely distributed in the Southern parts of Zagros. The other land-cover classes mostly occur with smaller proportions, except for

rangeland which becomes more continuous in the Southern parts of Zagros, where tree growth is more and more limited due to the climatic conditions. Overall, the land-cover map agrees well with the known distribution of land-cover in the Zagros area.

As shown in the focus area maps, discriminating the plantation/orchards class from the wooded areas was not always possible (see Fig. 9 – panel C and additional graphs in the Supplementary Material 3) and especially in areas with dense natural woody cover some confusions occurred. Some further confusion occurred between built-up areas and very bright bare soil or pebble areas (see Fig. 9 – panel B, along the river). Finally, the classes bare soil, rangeland and wooded area expectedly showed some confusions. This is no surprise as these three classes more or less form a continuum from bare soil to woodlands with rangeland having higher fractions of herbaceous vegetation. A clear separation of the three classes is hence challenging, even in the field.

3.3. Woody cover histogram

According to the final woody cover map presented in this study, most areas classified as forests, have a woody cover below approximately 60% even though higher cover values exist in some areas (bottom right panel of Fig. 10). The highest frequencies can be observed for areas with a cover of approximately 10% which is FAO’s minimum cover to fulfill the definition of a forest. If woody cover estimates for bare soil, rangeland and plantations areas are additionally considered in the histogram, the proportion of very low cover values increases notably (top panels in Fig. 10).

4. Discussion

In this study, we suggest a fully remote sensing-based work-flow to estimate fractional woody cover across the entire Zagros mountains. Our results demonstrate the capability of S2 images to estimate woody cover in the semi-arid Zagros Forest area using reference data from VHR satellite imagery of Google Earth and Bing instead of field measurements. Our methodology is practical when focusing on vast areas where field sampling can be very time-consuming and expensive. Recently published studies proved that high-resolution imagery can be used as complementary to field sampling (e.g., Anchang et al., 2020; Eskandari et al., 2020), whereas some studies solely relied on VHR imagery as surrogate of field data (e.g., Kattenborn et al., 2019; Fassnacht et al., 2021; Ludwig et al., 2019). The latter studies argued that for certain objectives, the VHR imagery might even be more reliable than field data. For example, visual estimates of woody cover on the ground may be less accurate as the field observer is not sharing the bird eye's perspective of the satellite data (Kattenborn et al., 2019). Visual comparisons of our resulting woody cover map showed very good agreements with the VHR data and seem to confirm the validity of this approach.

One challenge when working with extensive study areas is the environmental variability. The Zagros area is considered a geographic entity, but the environmental and socio-political settings within Zagros vary widely. The Zagros area has varying rainfall patterns and crosses several climatic zones from very hot and dry (in southern Zagros) to semi-arid and even cold climates (in northern Zagros) (Ashraf Vaghefi et al., 2019). Different dominant woody species may occur in each of these climatic zones, exemplified by 1) the high mixture of *Q. libani* and *Q. infectoria* with *Q. brantii* in the northern part and their absence in central / southern parts and 2) the dominance of *Amygdalus* spp. in the very southern part, occasionally in absence of *Q. brantii*. Another significant source of variation is the traditional land-use practices outside the main urban centers, which range from a dominance of settled villages in the northern parts to different types of nomadic life in central/southern parts, resulting in variable forest tenure and thus densities due to different silvo-pastoral systems. This might be one reason why our model was able to predict woody cover reasonably well but still a notable amount of unexplained variation remained. Given, the comparably low woody cover in most parts of Zagros (Figs. 7 and 10), the obtained median RMSE value of 0.12 (12%) is comparably high and indicates that there is still room for improving the suggested work-flow. Concerning the variability introduced by differing climatic conditions, we tried to at least partly account for the environmental gradients of the Zagros area by adding bioclim variables to our work-flow. These turned out to be amongst the selected predictors in almost all RF models, which confirms the importance of considering information on environmental gradients in estimating woody cover fraction. Some earlier studies focusing on woody cover estimation classified their study area into various bioclimatic zones and found a significant relationship between bioclimatic zones and woody cover density (e.g., Brandt et al., 2018; Sankaran et al., 2005). Thus, an alternative way to our suggested approach could be to classify the entire study area to different climatic zones and then train separate zone-specific models.

Another source or variation in our work-flow relates to the VHR Bing and Google images. Most Google Satellite images for the Zagros area are more recent than the Bing Maps images (Lesiv et al., 2018). This was also reflected in the data used for our analysis (7000 Google Earth vs. 1000 Bing images). For most sub-regions, particularly in the central part of Zagros, no VHR imagery with an acquisition date after 2015 / 2016 was available. Hence, these older reference data that might have not been fully representative for the current conditions mirrored in the S2-imagery and may have had a negative effect on our model performance. This issue could be addressed by collecting UAV data in areas with outdated VHR images, however, that would also lead to notably increased costs.

4.1. Effects of spatial grain on model performance

One objective of our study was to examine how the spatial grain of the analysis affects the model quality. We initially used S2 bands at 10 m pixel size, and then additionally ran the model with down-sampled S2 data with pixel sizes from 20 to 120 m spatial resolution. Our results show a decrease in RMSE and increase in R^2 from 10 to 60 m pixel size, followed by a saturation and comparably stable results with only marginal fluctuations afterwards. This model improvement at coarser spatial grain matches our expectations and is in line with the findings of Korhonen et al. (2017), who reported better model performances when using nine S2 image pixels instead of one. Higginbottom et al. (2018) also estimated woody cover density using Landsat TM/ ETM + imagery at 30, 60, 90, 120 m pixel sizes, and suggested that the most accurate model was the model with the coarsest resolution ($R^2 = 0.8$, RMSE = 8.9, at the 120 m pixel scale) when solely using dry season images. They discussed that while the maps at 120 m were of the highest model accuracy, they were less suitable to detect smaller clumps and canopies of dryland vegetation. Accordingly, and as a compromise between model performance and the ability to depict finer spatial patterns, we used the 40 m model to produce our final map. The tendency of the models to perform better at coarser spatial grain may relate to at least two aspects: First, potential co-location errors between S2 and VHR imagery at the sub-pixel scale have less influence at coarser grain. Second, by increasing the spatial grain of the analysis, the number of extreme values in the reference data (very high and very low woody cover estimates) will decrease (see Supplementary Material 4). The latter point may lead to reductions in RMSE, particularly if a model like RF is applied that tends to predict towards intermediate values. This phenomenon has for example been described for biomass estimation models trained with field plots of differing sizes and using several remote sensing data types (e.g., Fassnacht et al., 2018; Hosseini et al., 2020), and is also related to the inability of RF to predict out of the range of its reference values. It was consequently reported in many studies that RF tends to under-predict high and over-predict low values, which was also confirmed in our study (Fig. 8).

4.2. Selected predictor variables

The spatial scale only had a minor influence on the selected predictor variables. According to the results of VSURF-based variable selection, the frequency and order in which predictors variables were selected remained fairly constant across all examined scales. One exception was band 8 (spectrally centered at 842 nm), which was repeatedly selected in models based on coarser grain data (>70 m). It was remarkably also the case for two texture metrics based on the same band (B8_ent_1 and B8_var), which indicates that NIR-based texture information may become more relevant at coarser spatial grains. This seems logical as spatial variation in reflectance in the NIR is likely to relate to varied woody cover situations in our study area. This variation may play less of a role at finer spatial grains where the information in the individual pixel may be sufficient to describe the reference woody cover values extracted from the VHR images.

The most frequently selected predictor variables that were present in all RF models included NDVI, B4/ Red band, thres, B3/Green band, bio12/Annual Precipitation, B2/Blue band. Hence, particularly bands from the visual part of the spectrum were constantly selected. This partly agrees with Higginbottom et al. (2018) who reported B3 as the second-best individual S2 band (with $R^2 = 0.66-0.70$) for monitoring canopy properties in patchy discontinuous woody vegetation in South Africa. Heckel et al. (2020) and Waśniewski et al. (2020) also reported the bands from the visual domain to notably contribute to estimating forest cover. In contrast to Higginbottom et al. (2018) and some studies in other forest ecosystems (e.g., Heckel et al., 2020; Waśniewski et al., 2020), we found the red edge bands (B5-B7) and also the bands located in the near-infrared (B8, B8a) to be of only intermediate importance for

estimating woody cover. While this has also been reported by earlier studies focusing on related variables such as LAI (Meyer et al., 2019), it is very likely that this is also a consequence of NDVI being the most important predictor variables in all models. As NDVI captures the most important information contained in the red-edge, the additional, finer information from the red-edge bands may simply not be relevant when focusing on a straightforward target variable such as woody cover. The high importance of NDVI comes at no surprise as it has been previously reported to be a highly significant predictor variable for estimating tree/woody canopy cover fraction in the Zagros region (Eskandari et al., 2020). It was interesting to see that also our NDVI threshold predictor (named thres), a binary mask derived by applying a threshold of 0.17 to the NDVI was amongst the most frequently selected predictor variables. The idea of this predictor was to imitate the approach with which the reference data was created and we assumed that it will be particularly useful at coarser spatial resolutions. However, the predictor was relevant even at the finest examined spatial resolutions.

The two SWIR bands of S2 played almost no role in our study and only B12 was selected once in the model at 60 m spatial grain. This is not surprising, as the differentiation of the dark vegetation in front of the soil background is unlikely to depend on additional information on canopy structure or water content that is mostly contained in the SWIR bands. Furthermore, similar to the red-edge band, the most important information related to the SWIR bands (or at least Band 11) might have already been captured by the RSR index, which was also one of the most frequently selected predictors in all models. This index combines information from the RED, NIR and SWIR regions and tends to decrease the effects of bright and pronounced soil background in situations with sparse canopy cover like our study area. This high relevance of the RSR index agrees well with findings of several earlier studies focusing on similar target variables (e.g., Brown et al., 2000; Zhu et al., 2010).

Earlier studies proved that climatic variables are important for mapping woody vegetation and can improve remote sensing models (e.g., Liu et al., 2017; Brandt et al., 2018). Therefore, to account for potential influences of the large environmental gradient covered by the Zagros area on the relationship between remote sensing data and woody cover, we also included selected bioclim variables (BIO12, BIO15, BIO16 and BIO17). The annual precipitation (BIO12) was selected in all models, while precipitation seasonality (Bio15) was selected in all

models except the model with 40 m resolution. Precipitation of wettest and driest quarters (BIO16, BIO17) were also selected for all models except for the model with 20 m resolution. Even though, the bioclim variables were at a notably lower spatial resolution as compared to the S2 based predictors, it was remarkable to observe they were still selected in almost all models. We assume that, the bioclim variables were helpful to pre-stratify the feature space into coarse environmental regions which may explain a part of the variability in the reference data which was collected across the entire Zagros area. Topography metrics were also explored in some preliminary analyses but were found to not increase the model performance. The result of this analysis is shown in Supplementary Material 2. This was surprising but may partly be explained with the large environmental gradient of our study area. This gradient may challenge a stable relation between topography and woody cover as for example in the Northern areas with more precipitation, a shaded slope may be disadvantageous (due to less light and hence photosynthetic activity) while in the Southern areas with less precipitation a shaded slope may be beneficial due to the reduced transpiration stress. A more detailed analyses would be required to understand how the patterns of differing woody cover values relate to environmental variables and land-use trajectories. Our woody cover map could be a good starting point to analyze this question in future studies and corresponding findings may in turn be helpful to further improve the presented workflow by considering regional differences in the drivers of woody-cover patterns.

4.3. Relevance of the created woody cover map

For environmental science and natural resource management, accurate and timely maps of woody cover are critical (Karlson et al., 2015). The Zagros Forests, which have a documented history for presence of human settlements of beyond 5500 years, are one of Iran's most valuable cultural landscapes and play an important role in soil and water conservation. Climate change, population development, and people's reliance on these woodlands for their livelihoods (by harvesting wood and forest sub-products, turning woodlands to agricultural fields, and grazing areas) have all contributed to the decline of the woodland ecosystems in Zagros over the last 30 years (Jazirei, Ebrahimi, 2013). Despite the high environmental importance of these woodlands, they are currently not part of a national-scale protection plan. As a result, developing accurate maps of the Zagros Forests (cover, density, and distribution) are an important contribution to monitor the area. Providing an accurate fractional woody cover map can therefore support forest managers in determining (1) the exact distribution of wooded areas in the vast Zagros region and (2) support the regular assessment of potential changes induced by for example changes in land-use practices and climate-change (Eskandari et al., 2020).

Our woody cover product and the straightforward workflow developed here are particularly relevant for semi-arid and arid regions as many existing large-scale tree/woody cover products based on satellite data have clear limitations for characterizing areas with sparse tree canopies (e.g., Brandt et al., 2016; Cunningham et al., 2019; Fagan, 2020). To further illustrate this, we validated the Global Forest Change map (Hansen et al., 2000) (GFC), a benchmark for many similar studies, using our reference dataset. The corresponding results show a severe underestimation of the woody cover across the entire Zagros Forests (Fig. 11). This backs up prior findings of Bastin et al. (2017), Brandt et al. (2016) and Cunningham et al. (2019) who also reported that global forest/woody cover products are underestimating the woody cover in arid and semi-arid forests and woodlands. Our final woody cover map returned visually much more plausible estimates across the majority of Zagros region compared with all available global products, including the GFC and a recently published map of the Zagros area (Ghorbanian et al., 2020).

In our product, we dealt with the problem of bright soil background by adding bare soil samples with corresponding zero woody cover

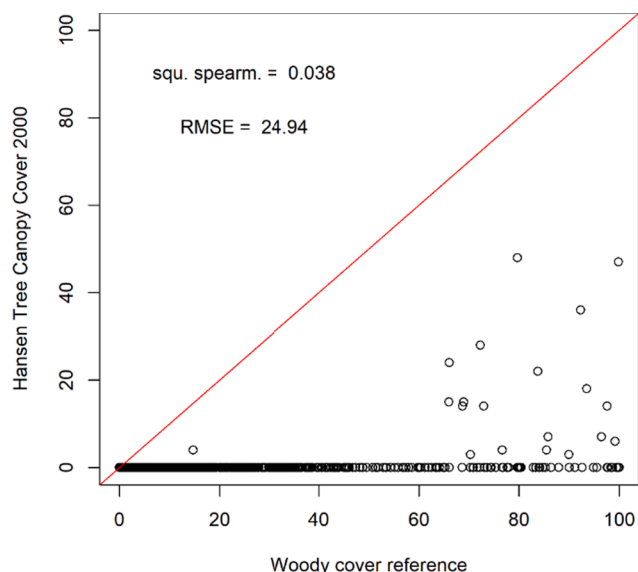


Fig. 11. Validation graph of Hansen global forest cover product at 30 m resolution with our reference woody cover dataset. It is worth mentioning that the Hansen global forest cover product is estimated for trees with heights greater than 5 m. While these heights are not achieved in all areas of Zagros, there is still a large fraction of the area with notably higher trees.

reference values to our reference dataset and by including the RSR index as an input variable. Both measures improved the performance of our model. We also followed the suggestion by [Symeonakis et al. \(2018\)](#) and used S2 images of the dry/summer season, i.e. the time of maximum difference amongst the spectral signals amongst vegetation, crops and grasses. In terms of additional remote sensing data from active sensors, the application of Sentinel 1 (S1) SAR data could be examined in the future. However, some studies already showed that adding S1 data is not in all cases improving the results ([Heckel et al., 2020](#)). In some preliminary analyses, the addition of S1 backscatter variables did not improve our models for estimating woody cover (results not shown here). One reason for this might be that the Zagros area is topographically complex and the standard pre-processing of SAR data to account for terrain shadows, foreshortening and related effects in GEE may have not been sufficient to obtain a signal that increases the information content already available in the S2 data. As discussed below, S1 data was, however, helpful to improve the differentiation of urban and bare soil areas in the land-cover classification. This may relate to the comparably extreme behavior of urban areas in the SAR signal (double-bounce effect) which may have led to a comparably clear signal despite the discussed challenges related to the pre-processing of the SAR data.

4.4. Land cover mapping

Studies on the status quo and changes of ecosystems and land-use systems in the Zagros region are partly constrained by the lack of reliable land cover maps ([Eskandari et al., 2020](#)). Recently published land cover maps for the entire Iran based on Sentinel-1 (S1) and S2 data resulted in overall accuracy and kappa coefficient of 91.35 % and 0.91, respectively ([Ghorbanian et al., 2020](#)). However, these maps showed severe under-estimation of forested areas in the Zagros region, mostly due to mis-classification as rangelands. This may also relate to the definition of the classes forest and rangeland. Most of the Zagros Forests are indeed used as rangeland by nomads and villagers living in the area. On the other hand, extensive temperate forest regions exist in the Northern parts of Iran and hence, in a national-scale land cover map, the class forest may have been exclusively used for these ecosystems. Other products covering the entire Zagros area were not available, excluding global land cover products which are typically either available at a comparably coarse spatial grain or with only few land-cover classes. Our land cover product achieved high accuracies for the seven considered classes and we deem it useful for masking out irrelevant land cover classes from our woody cover product. The use of this map for other purposes may be limited due to the comparably sparse number of land-cover classes considered in the work-flow. Furthermore, despite the overall high accuracies, we still observed some obvious mis-classifications in our land cover map. One of the major problems were misclassification of orchards with natural woody cover. Particularly, in areas where orchards are surrounded by forest stands, it is very challenging to discriminate the two classes. One example is given in [Supplementary Material 3](#), where it is shown that very dense natural woodlands tend to be misclassified as plantation areas. One approach to improve related misclassifications in future studies may be to integrate spatial context into the analysis. In most cases plantations and orchards occur in the immediate neighborhood of settlements and nearby roads. Hence, integrating a layer representing the distance to settlements and roads could be helpful to mitigate the over-classification of plantation areas. Further confusions occurred between the build-up class and bare soil. This is also a commonly occurring problem that has been discussed before (e.g., [Piyooosh, Ghosh, 2018](#)). In the Zagros area, besides bare soil areas, also areas covered with seasonally occurring herbaceous vegetation and crops (especially rainfed crops) can be confused with built-up areas. Combining optical and radar data can help mitigating this problem. While optical data is able to capture phenological characteristics of the vegetation, the backscatter behavior of SAR data can further contribute to reliably identify built-up surfaces as the scattering

behavior typically differs notably from vegetated and bare soil areas ([Luti et al., 2021](#)). Some earlier studies showed the potential of jointly using Sentinels 1 and 2 data to extract built-up areas (e.g., [Luti et al., 2021](#), [Dong et al., 2020](#), [Abdikan et al., 2016](#)). However, in topographical complex areas such as the Zagros region, the pre-processing of SAR data to account for terrain shadows, foreshortening and related effects is rather challenging. In our study, we addressed this issue by applying filters to the available Sentinel-1 data catalogue in GEE and this comparably straightforward approach resulted in reasonable SAR mosaics that contributed to an improved separation of built-up and bare soil areas.

5. Conclusions

In this study we proposed a work-flow to map woody cover in an extensive semi-arid woodland region at comparably high spatial resolution using exclusively freely available datasets. The presented work-flow based on a combination of VHR images from Google Satellite, Bing and Sentinel-2 multispectral satellite data, implemented mostly in the Google Earth Engine, resulted in a reliable woody cover product for the entire Zagros Forest area. An additional land cover classification was used to mask out irrelevant land cover classes. The woody cover product was compared to existing global forest cover datasets and proved to be of notably higher quality. The approach is particularly interesting because no costly field data is required and the work-flow is straightforward to implement and could hence also be beneficial for estimating woody cover in other arid and semi-arid regions of the world.

CRedit authorship contribution statement

Elham Shafeian: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Fabian Ewald Fassnacht:** Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – review & editing. **Hooman Latifi:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The first author acknowledges the German Academic Exchange Service (DAAD) for supporting her research with a PhD scholarship. The authors are grateful to the Research Institute of Forests and Rangelands of Iran, which enabled a visit to parts of the Zagros Forest region via the National Zagros Monitoring Project (Project No. 01-09-09-047-97012) during 2019. Parts of this study were conducted within the FORZA project at the Karlsruhe Institute of Technology and K. N. Toosi University of Technology. The project is co-funded by the German National Space Agency DLR (Deutsches Zentrum für Luft- und Raumfahrt e.V.) on behalf of the German Federal Ministry of Education and Research based on the Bundestag resolution 01DK20085 and by the Iranian Ministry of Science, Research and Technology.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jag.2021.102621>.

References

- Abdikan, S., Sanli, F.B., Ustuner, M., Calò, F., 2016. Land cover mapping using sentinel-1 SAR data, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B7, XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.
- Abdolahi, H., Shataee Joybari, S.H., 2012. Comparative evaluation of IRS-P6-LISS-III and LISS IV images for canopy cover mapping of Zagros forests (Case Study: Javanroud Forests). *J. Wood For. Sci. Technol.* 19 (1) (in Farsi).
- Abdollahnejad, A., Panagiotidis, D., Bílek, L., 2019. An integrated GIS and remote sensing approach for monitoring harvested areas from very high-resolution, low-cost satellite images. *Remote Sens.* 11 (21), 2539. <https://doi.org/10.3390/rs11212539>.
- Ahmadi Sani, N., Darvishsefat, A.A., Zobieri, M., Farzaneh, A., 2007. Potentiality of aster images for forest density mapping in Zagros forests of Iran, case study: Marivan forests. *Iranian J. Natural Resour.* 61 (3), 603–614.
- Alibakhshi, S., Hovi, A., Rautiainen, M., 2019. Temporal dynamics of albedo and climate in the sparse forests of Zagros. *Sci. Total Environ.* 663, 596–609.
- Anchang, J.Y., Prihodko, L., Ji, W., Kumar, S.S., Ross, C.W., Yu, Q., Lind, B., Sarr, M.A., Diouf, A.A., Hanan, N.P., 2020. Toward operational mapping of woody canopy cover in tropical savannas using google earth engine. *Front. Environ.-Sci.* 8 <https://doi.org/10.3389/fenvs.2020.0000410.3389/fenvs.2020.00004.s001>.
- Andersen, H.-E., McGaughey, R.J., Reutebuch, S.E., 2005. Estimating forest canopy fuel parameters using LIDAR data. *Remote Sens. Environ.* 94 (4), 441–449.
- Arsalani, M., Pourtahmasi, K., Azizi, G., Bräuning, A., Mohammadi, H., 2018. Tree-ring based December–February precipitation reconstruction in the southern Zagros Mountains, Iran. *Dendrochronologia* 49, 45–56.
- Ashraf Vaghefi, S., Keykhai, M., Jahanbakhshi, F., Sheikholeslami, J., Ahmadi, A., Yang, H., Abbaspour, K.C., 2019. The future of extreme climate in Iran. *Sci. Rep.* 9, 1464.
- Attarod, P., Rostami, F., Dolatshahi, A., Sadeghi, S.M.M., Amiri, G.Z., Bayramzadeh, V., 2016. Do changes in meteorological parameters and evapotranspiration affect declining oak forests of Iran? *J. For. Sci.* 62 (12), 553–561.
- Bai, L., 2010. Comparison and Validation of Five Global Land Cover Products over the African Continent. Degree-thesis in Physical Geography and Ecosystem Analysis. Lund University, Sweden.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., et al., 2007. LAI, fAPAR and fCover CYCLOPS global products derived from VEGETATION Part 1: Principles of the algorithm. *Remote Sens. Environ.* 110, 275–286.
- Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., Patriarca, C., Picard, N., Sparrow, B., Abraham, E.M., Aloui, K., Atesoglu, A., Attore, F., Bassüllü, Ç., Bey, A., Garzuglia, M., García-Montero, L.G., Groot, N., Guerin, G., Laestadius, L., Lowe, A.J., Mamane, B., Marchi, G., Patterson, P., Rezende, M., Ricci, S., Salcedo, I., Diaz, A.-P., Stolle, F., Surappaeva, V., Castro, R., 2017. The extent of forest in dryland biomes. *Science* 356 (6338), 635–638. <https://doi.org/10.1126/science.aam6527>.
- Baumann, M., Ozdogan, M., Kummerle, T., Wendland, K.J., Espinova, E., Radeloff, V.C., 2012. Using the Landsat record to detect forest-cover changes during and after the collapse of the Soviet Union in the temperate zone of European Russia. *Remote Sens. Environ.* 124, 174–184. <https://doi.org/10.1016/j.rse.2012.05.001>.
- Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Bivand, Roger, Keitt, Tim, Rowlingson, Barry, 2021. rgdal: Bindings for the ‘Geospatial’ Data Abstraction Library. R package version 1.5-23. <https://CRAN.R-project.org/package=rgdal>.
- Bivand, R., Rundel, C., Pebesma, E., Stuetz, R., et al., 2020. R package rgeos. Available at <https://cran.r-project.org/web/packages/rgeos/index.html>.
- Brandt, M., Hiernaux, P., Rasmussen, K., Mbou, C., Kergoat, L., Tagesson, T., Ibrahim, Y. Z., Wélé, A., Tucker, C.J., Fensholt, R., 2016. Assessing woody vegetation trends in Sahelian drylands using MODIS based seasonal metrics. *Remote Sens. Environ.* 183, 215–225.
- Brandt, M., Rasmussen, K., Hiernaux, P., Herrmann, S., Tucker, C.J., Tong, X., Tian, F., Mertz, O., Kergoat, L., Mbou, C., David, J.L., Melocik, K.A., Dendoncker, M., Vincke, C., Fensholt, R., 2018. Reduction of tree cover in West African woodlands and promotion in semi-arid farmlands. *Nat. Geosci.* 11 (5), 328–333.
- Breiman, L., 2001. Random forests, Machine learning, Issue 1, Vol 45:5–32, <https://link.springer.com/article/10.1023/A:1010933404324>.
- Brown, L., Chen, J.M., Leblanc, S.G., Cihlar, J., 2000. A Shortwave Infrared Modification to the Simple Ratio for LAI Retrieval in Boreal Forests: An Image and Model Analysis. *Remote Sens. Environ.*, 71, 16–25.
- Bucini, G., Saatchi, S., Hanan, N., Boone, R.B., Smit, I., 2009. Woody cover and heterogeneity in the Savannas of the Kruger national park, South Africa. Conference: Geoscience and Remote Sensing Symposium, 2009 IEEE International IGARSS 2009, 4. <https://doi.org/10.1109/IGARSS.2009.5417381>.
- Coburn, C.A., Roberts, A.C.B., 2004. A multiscale texture analysis procedure for improved forest stand classification. *Int. J. Remote Sens.* 25 (20), 4287–4308.
- Corcoran, J., Knight, J., Pelletier, K., Rampi, L., Wang, Y., 2015. The effects of point or polygon based training data on randomforest classification accuracy of wetlands. *Remote Sens.* 7, 4002–4025. <https://doi.org/10.3390/rs70404002>.
- Cunningham, D., Cunningham, P., Fagan, M.E., 2019. Identifying biases in global tree cover products: a case study in Costa Rica. *Forests* 10 (10), 853. <https://doi.org/10.3390/f10100853>.
- Darvishsefat, A., Saroei, S., 2003. Evaluation of the potential of Landsat ETM+ for forest density mapping in Zagros forests of Iran. In: International GIS and Remote Sensing Symposium, 21–25 July 2003, Toulouse, France. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=1294498>.
- Deka, J., Tripathi, O.P., Khan, M.L., 2012. Implementation of forest canopy density model to monitor tropical deforestation. *J. Indian Soc. Remote Sens.* 41, 2.
- Deur, M., Gašparović, M., Balenović, I., 2020. Tree Species Classification in Mixed Deciduous Forests Using Very High Spatial Resolution Satellite Imagery and Machine Learning Methods. *Remote Sens.*, 12, 3926; doi:10.3390/rs12233926.
- Dobrowski, S.Z., Safford, H.D., Cheng, Y.B., Ustin, S.L., 2008. Mapping mountain vegetation using species distribution modeling, image-based texture analysis, and object-based classification. *Appl. Veg. Sci.* 11, 499–508. <https://doi.org/10.3170/2008-7-18560>.
- Dong, D., Wang, C., Yan, J., He, Q., Zeng, J., Weia, Z., 2020. Combining Sentinel-1 and Sentinel-2 image time series for invasive *Spartina alterniflora* mapping on Google Earth Engine: a case study in Zhangjiang Estuary. *J. Appl. Remote Sens.* 14 (4).
- Erfanifard, Y., Khodaei, Z., Fallah Shamsi, R., 2014. A robust approach to generate canopy cover maps using UltraCam-D derived orthoimagery classified by support vector machines in Zagros woodlands. *West Iran, European Journal of Remote Sensing* 47 (1), 773–792. <https://doi.org/10.5721/EuJRS2014474>.
- Eskandari, S., Jaafari, M.R., Oliva, P., Ghorbanzadeh, O., Blaschke, T., 2020. Mapping land cover and tree canopy cover in Zagros forests of Iran: application of sentinel-2. Google Earth, and Field Data. *Remote Sens.* 12, 1912. <https://doi.org/10.3390/RS12121912>.
- Fadaei, H., Sakai, T., Yoshimura, T., Kazuyuki, M., 2010. Estimation of tree density with high-resolution imagery in the Zarin forest of north Iran (cupressus sempervirens var. horizontalis, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Volume XXXVIII, Part 8, Kyoto Japan 2010, 679–684.
- Fagan, M.E., 2020. A lesson unlearned? Underestimating tree cover in drylands biases global restoration maps. *Glob. Change Biol.* 26 (9), 4679–4690. <https://doi.org/10.1111/gcb.15187>.
- Fassnacht, F.E., Latifi, H., Hartig, F., 2018. Using synthetic data to evaluate the benefits of large field plots for forest biomass estimation with LIDAR. *Remote Sens. Environ.* 213, 115–128. <https://doi.org/10.1016/j.rse.2018.05.007>.
- Fassnacht, F.E., Schmidt-Riese, E., Kattenborn, T., Hernández, J., 2021. Explaining Sentinel 2-based dNBR and RdNBR variability with reference data from the bird’s eye (UAS) perspective. *Int. J. Appl. Earth Obs. Geoinf.* 95 (102262).
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al., 2007. The shuttle radar topography mission. *Rev. Geophys.* 45, RG2004. <https://doi.org/10.1029/2005RG000183>.
- Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, Muchoney, X.Y.D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Bacchini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. *Remote Sens. Environ.* 83, pp. 287–302.
- Gao, Q., Zribi, M., Escorihuela, M.J., Baghdadi, N., 2017. Synergetic Use of sentinel-1 and sentinel-2 data for soil moisture mapping at 100 m resolution. *Sensors* 17, 1966.
- García Criado, M., Myers-Smith, I.H., Björkman, A.D., Lehmann, C.E.R., Stevens, N., 2020. Woody plant encroachment intensifies under climate change across tundra and savanna biomes. *Glob. Ecol. Biogeogr.* 29 (5), 925–943. <https://doi.org/10.1111/gcb.13072>.
- Genuer, R., Poggi, J.M., Tuleau-Malot, C., 2015. VSURF: An R package for variable selection using random forests. *The R J.* 7/2.
- Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., Hasanlou, M., 2020. Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. *ISPRS J. Photogramm. Remote Sens.* 167 (276–288), 276–288. <https://doi.org/10.1016/j.isprsjprs.2020.07.013>.
- Gislason, P., Benediktsson, J.A., Sveinsson, J.R., 2006. Random Forests for land cover classification. *Pattern Recogn. Lett.* 27, 294–300.
- Gonsamo, A., D’Odorico, P., Pellikka, P., 2013. Measuring fractional forest canopy element cover and openness - Definitions and methodologies revisited. *Oikos* 122, 1283–1291.
- Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* 21 (6), 1331–1364.
- Heckel, K., Urban, M., Schratz, P., Mahecha, M.D., Schmulius, C., 2020. Predicting forest cover in distinct ecosystems: the potential of multi-source sentinel-1 and -2 data fusion. *Remote Sens.* 12 (302) <https://doi.org/10.3390/rs12020302>.
- Higginbottom, T.P., Symeonakis, E., Meyer, H., van der Linden, S., 2018. Mapping fractional woody cover in semi-arid savannas using multi-seasonal composites from Landsat data. *ISPRS J. Photogramm. Remote Sens.* 139, 88–102.
- Hijmans, R.J., Etten, J.V., 2012. raster: Geographic analysis and modeling with raster data. R package version 2.0-12, <http://CRAN.R-project.org/package=raster>.
- Hosseini, Z., Latifi, H., Naghavi, H., Bakhtiarvand Bakhtiari, S., Fassnacht, F.E., 2020. Influence of plot and sample sizes on aboveground biomass estimations in plantation forests using very high-resolution stereo satellite imagery. *Forestry: An Int. J. Forest Res.* 94 (2), 278–291. <https://doi.org/10.1093/forestry/cpaa028>.
- Jazirei, M.H., Ebrahimi Rostaghi, M., 2013. Zagros Forestry. University of Tehran press (UTP).
- Jongman, R.H.G., Skidmore, A.K., Múcher, C.A., Bunce, R.G.H., Metzger, M.J., 2017. Global Terrestrial Ecosystem Observations: Why, Where, What and How? In: Walters, M., Sholes, R.J. (Eds.). 2017. The GEO Handbook on Biodiversity Observation Networks. Springer. ISBN: 978-3-319-27286-3. 330 p.
- Joshi, C., De Leeuw, J., Skidmore, A.K., van Duren, I.C., van Oosten, H., 2006. Remotely sensed estimation of forest canopy density: A comparison of the performance of four methods. *Int. J. Appl. Earth Observation Geo-information* 8, 84–95.
- Karlson, M., Ostwald, M., Reese, H., Sanou, J., Tankoano, B., Mattsson, E., 2015. Mapping tree canopy cover and aboveground biomass in sudano-sahelian woodlands using landsat 8 and random forest. *Remote Sens.* 7, 10017–10041. <https://doi.org/10.3390/rs70810017>.

- Kattenborn, T., Lopatin, J., Förster, M., Braun, A.C., Fassnacht, F.E., 2019. UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data. *Remote Sens. Environ.* 227, 61–73.
- Khabazi, M., 2020. Satellite data as indicators of oak forests canopy cover change (Case study Kamfrouze oak forests). *J. Wildlife Biodiversity* 4 (1), 66–72. <https://doi.org/10.22120/jwb.2019.119131.1106>.
- Korhonen, L., Petteiri Packalen, Hadi, Packalen, P., Rautiainen, M., 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote Sens. Environ.* 195, 259–274.
- Latifi, H., Fassnacht, F.E., Hartig, F., Bergere, C., Hernández, J., Corvalán, P., Koch, B., 2015. Stratified aboveground forest biomass estimation by remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* 38, 229–241. <https://www.worldclim.org/data/bioclim.html>.
- Lee, A.C., Lucas, R.M., 2007. A LiDAR-derived canopy density model for tree stem and crown mapping in Australian forests. *Remote Sens. Environ.* 111, 493–518.
- Lesiv, M., See, L., Laso Bayas, J.C., Sturn, T., Schepaschenko, D., Karner, M., Moorhy, I., McCallumFritz, S., 2018. Characterizing the spatial and temporal availability of very high-resolution satellite imagery in google earth and microsoft bing maps as a source of reference data. *Land* 7 (4). <https://doi.org/10.3390/land7040118>.
- Liao, Z., Van Dijk, A.J.J.M., He, B., Larraondo, P.R., Scarth, P.F., 2020. Woody vegetation cover, height and biomass at 25-m resolution across Australia derived from multiple sites, airborne and satellite observations. *Int. J. Appl. Earth Obs. Geoinf.* 93, 1022092.
- Liu, X., Liu, H., Qiu, S., Wu, X., Tian, Y., Hao, Q., 2017. An Improved Estimation of Regional Fractional Woody/Herbaceous Cover Using Combined Satellite Data and High-Quality Training Samples. *Remote Sens.* 9, 32.
- López, G., J., J. Prado M., L. Manzo D. y A. Peralta Higuera., 2016. Monitoring changes of forest canopy density in a temperate forest using high-resolution aerial digital photograph. *Investigaciones Geográficas, Boletín, num. 90, Instituto de Geografía, UNAM, Mexico, 59-74, dx.doi.org/10.14350/ig.47360*.
- Ludwig, M., Morgenthal, T., Detsch, F., Higginbottom, T.P., Valdes, M.L., Nauß, T., Meyer, H., 2019. Machine learning and multi-sensor-based modelling of woody vegetation in the Molopo Area, South Africa. *Remote Sens. Environ.* 222, 195–203.
- Luti, T., Fioravante, P.D., Marinosci, I., Strollo, A., Riitano, N., Falanga, V., Mariani, L., Congedo, L., Munafò, M., 2021. Land consumption monitoring with SAR data and multispectral indices. *Remote Sens.* 13, 1586. <https://doi.org/10.3390/rs13081586>.
- Mahdavi, A., Mirzayi, J., Karami, O., 2014. Status of dead trees in Zagros forests (case study: forests of Beyore district of Ilam province). *J. Forest Sustainable Development* 1 (4), 329–340. http://fsdj.guilan.ac.ir/article_1253.html.
- Mellor, A., Boukir, S., Hayward, A., Jones, 2015. Exploring issues of training data imbalance and mislabeling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS J. Photogrammetry Remote Sens.*, 105, 155–168.
- Meyer, L.H., Heurich, M., Beudert, B., Premier, J., Pflugmacher, D., 2019. Comparison of landsat-8 and sentinel-2 data for estimation of leaf area index in temperate forests. *Remote Sens.* 11, 1160.
- Mirzaeizadeh, V., Niknejad, M., Hojjati, S.M., 2015. Estimation of forest canopy density using FCD. *Ecol. Iranian Forests* 3 (5) (in Farsi).
- Nagelkirk, R.L., Dahlin, K.M., 2020. Woody cover fractions in African Savannas from landsat and high-resolution imagery. *Remote Sens.* 12, 813.
- Nandy, S., Joshi, P.K., Das, K.K., 2003. Forest canopy density stratification using biophysical modeling. *J. Indian Soc. Remote Sens.* 31 (4).
- Piyooch, A.K., Ghosh, S.K., 2018. Development of a modified bare soil and urban index for Landsat 8 satellite data. *Geocarto Int.* 33 (4), 423–442.
- Pourmoghadam, K., Pourmoghadam, K., Khosropour, E., Haidari, H., 2013. Identifying forest types associated with physiological factors in middle Zagros forests in Iran. *Int. J. Adv. Biol. Biomed. Res.* 1 (8), 830–834. http://www.ijabbr.com/article_7817_c4714aae64b4f165c6dcbdbcf7305e85.pdf.
- R Development Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/>.
- Rahimi, S.H., Sharifia, Z., Mastrodonardob, G., 2020. Comparative Study of the Effects of Wildfire and Cultivation on Topsoil Properties in the Zagros Forest. *Iran, Eurasian Soil Science* 53 (11), 1655–1668.
- Rautiainen, M., Heiskanen, J., 2013. Seasonal contribution of understory vegetation to the reflectance of a boreal landscape at different spatial scales. *IEEE Geosci. Remote Sens. Lett.*, 10 (4), 923 Seasonal Contribution.
- Rautiainen, M., Möttö, M., Heiskanen, J., Akujärvi, A., Majasalmi, T., Stenberg, P., 2011. Seasonal reflectance dynamics of common understory types in a northern European boreal forest. *Remote Sens. Environ.* 115, 3020–3028.
- Rikimaru, A., Roy, P.S., Miyatake, S., 2002. Tropical Forest Density Mapping *Tropical Ecology*, 43 (1), 39–47.
- Sagheb-Talebi, K., Sajedi, T., Pourhashemi, M., 2014. *Forests of Iran (A Treasure from the Past, a Hope for the Future)*. Springer.
- Sankaran, M., Hanan, Niall P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S., Gignoux, J., Higgins, S.I., Le Roux, X., Ludwig, F., Ardo, J., Banyikwa, F., Bronn, A., Bucini, G., Caylor, K.K., Coughenour, M.B., Diouf, A., Ekaya, W., Feral, C.J., February, E.C., Frost, P.G.H., Hiernaux, P., Hrabar, H., Metzger, K.L., Prins, H.H.T., Ringrose, S., Sea, W., Tews, J., Worden, J., Zambatis, N., 2005. Determinants of woody cover in African savannas. *Nature* 438.
- Schlerf, M., Atzberger, C., 2006. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sens. Environ.* 100, 281–294.
- Shahvali Kouhshour, A., Pir Bavaghar, M., Fatehi, P., 2012. Forest cover density mapping in sparse and semi dense forests using forest canopy density model (Case study: Marivan forests). *J. Appl. RS & GIS Techniques in Natural Resour. Sci.* 3 (3) (in Farsi).
- Smith, W.K., Dannenberg, M.P., Yan, D., Herrmann, S., Barnes, M.L., Barron-Gafford, G. A., Biederman, J.A., Ferrenberg, S., Fox, A.M., Hudson, A., Knowles, J.F., MacBean, N., Moore, D.J.P., Nagler, P.L., Reed, S.C., Rutherford, W.A., Scott, R.L., Wang, X., Yang, J., 2019. Remote sensing of dryland ecosystem structure and function: Progress, challenges, and opportunities. *Remote Sens. Environ.* 233 (July), 111401 <https://doi.org/10.1016/j.rse.2019.111401>.
- Soleimannejad, L., Bonyad, A.E., Naghdi, R., 2018. Remote sensing-assisted mapping of quantitative attributes in Zagros open forests of Iran. *Caspian J. Environ. Sci.* 16 (3), 215–230.
- Symeonakis, E., Higginbottom, T.P., Petroulaki, K., Rabe, A., 2018. Optimisation of savannah land cover characterisation with optical and SAR data. *Remote Sens.* 10, 499 <https://doi.org/10.3390/rs10040499>.
- Twery, M.J., Weiskittel, A.R., 2013. "Forest-Management Modelling." In *Environmental Modelling*, edited by John Wainwright and Mark Mulligan, John Wiley & Sons, pp. 379–398.
- Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L.P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., Woodcock, C., 2015. Free and open access satellite data are key. *Biol. Conserv.* 182, 173–176.
- Van Deventer, A.P., Ward, A.D., Gowda, P.H., Lyon, J.G., 1997. Using thematic mapper data to identify contrasting soil plains and tillage practices. *Photogramm. Eng. Remote Sens.* 63, 87–93.
- Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. *Remote Sens. Environ.* 16, 125–141.
- Villegas, J.C., Dominguez, F., Barron-Gafford, G.A., Adams, H.D., Guardiola-Claramonte, M., Sommer, E.D., Selvey, A.W., Espeleta, J.F., Zou, C.B., Breshears, D. D., Huxman, T.E., 2015. Sensitivity of regional evapotranspiration partitioning to variation in woody plant cover: Insights from experimental dryland tree mosaics. *Glob. Ecol. Biogeogr.* 24 (9), 1040–1048. <https://doi.org/10.1111/gcb.12349>.
- Wagenseil, H., Samimi, C., 2007. Woody vegetation cover in Namibian savannahs. *erdkunde* 10 (3112), 325–334.
- Wang, C., Qi, J., Cochrane, M., 2005. Assessment of tropical forest degradation with canopy fractional cover from Landsat ETM+ and IKONOS imagery. *Earth Interact* 9 (22), 1–18. <https://doi.org/10.1175/EI133.1>.
- Waśniewski, A., Hościło, A., Zagajewski, B., Moukétou-Tarazewicz, D., 2020. Assessment of Sentinel-2 Satellite Images and Random Forest Classifier for Rainforest Mapping in Gabon. *Forests* 11 (9), 941. <https://doi.org/10.3390/f11090941>.
- Wingate, V.R., Kuhn, N.J., Phinn, S.R., van der Waal, C., 2019. Mapping trends in woody cover throughout Namibian savanna with MODIS seasonal phenological metrics and field inventory data. *Biogeosciences Discuss.* <https://doi.org/10.5194/bg-2019-28>.
- Wood, E.M., Pidgeon, A.M., Radeloff, V.C., Keuler, N.S., 2012. Image texture as a remotely sensed measure of vegetation structure. *Remote Sens. Environ.* 121, 516–526.
- Yang, J., Weisberg, P.J., Bristow, N.A., 2012. Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. *Remote Sens. Environ.* 119, 62–71.
- Yang, L., Jia, K., Liang, S., Wei, X., Yao, Y., Zhang, X., 2017. A robust algorithm for estimating surface fractional vegetation cover from Landsat data. *Remote Sensing* 9 (8), 1–20. <https://doi.org/10.3390/rs9080857>.
- Yang, X., Crews, K.A., 2019. Fractional woody cover mapping of Texas savanna at Landsat scale. *Land* 8 (1). <https://doi.org/10.3390/land8010009>.
- Zandler, H., Brenning, A., Samimi, C., 2015. Quantifying dwarf shrub biomass in an arid environment: comparing empirical methods in a high dimensional setting. *Remote Sens. Environ.* 158, 140–155.
- Zhang, M., Huang, H., Li, Z., Hackman, K.O., Liu, C., Andriamiarisoa, R.L., Rahevelo, T. A.N.A., Li, Y., Gong, P., 2020. Automatic High-Resolution Land Cover Production in Madagascar Using Sentinel-2 Time Series, Tile-Based Image Classification and Google Earth Engine. *Remote Sens.* 12, 3663. <https://doi.org/10.3390/rs12213663>.
- Zhang, W., Brandt, M., Wange, Q., Prishchepov, A.V., Tucker, C.J., Lia, Y., Lyua, H., Fensholt, R., 2019. From woody cover to woody canopies: How Sentinel-1 and Sentinel-2 data advance the mapping of woody plants in savannas. *Remote Sens. Environ.* 234.
- Zhu, G., Ju, W., Chen, J.M., Zhou, Y., Li, X., Xu, X., 2010. Comparison of Forest Leaf Area Index Retrieval Based on Simple Ratio and Reduced Simple Ratio, 2010 18th International Conference on Geoinformatics, 18–20 June 2010. doi: 10.1109/9/GEOINFORMATICS.2010.5568204.