earth

MDPI

Article

Analyzing the Impact of Geoenvironmental Factors on the Spatiotemporal Dynamics of Forest Cover via Random Forest

Hendaf N. Habeeb [1,2] and Yaseen T. Mustafa [3,4,5,*]

[1] Department of Spatial Planning, College of City and Regional Planning, University of Duhok, Duhok 42001, Iraq; hendaf.habeeb@uod.ac
[2] Directorate of Forests and Range in Duhok, General Directorate of Horticulture, Forestry and Rangeland, Ministry of Agriculture and Water Resources, Duhok 42001, Iraq
[3] Department of Environmental Science, College of Science, University of Zakho, Zakho 42002, Iraq
[4] Department of Environmental Engineering, College of Engineering, Knowledge University, Erbil 44001, Iraq
[5] Applied Remote Sensing & GIS Center, University of Zakho, Zakho 42002, Iraq
* Correspondence: yaseen.mustafa@uoz.edu.krd

Abstract: Understanding the dynamic relationships between geoenvironmental factors and forest vegetation cover is crucial for effective forest management and planning. This study investigates the spatiotemporal dynamics of forest cover in the Duhok District in the Kurdistan Region of Iraq over a decade (2013–2023), emphasizing the impact of geoenvironmental factors via Random Forest algorithms and Landsat data. This research integrates datasets including fractional vegetation cover (FVC), groundwater levels, climate data, topography, and soil moisture data, offering a comprehensive analysis of the factors influencing forest cover. The results show that in 2013, altitude and rainfall were the primary factors influencing FVC, with areas of higher altitudes and adequate rainfall exhibiting up to 30% denser forest cover. By 2023, soil moisture and groundwater levels had emerged as the dominant factors, with soil moisture levels accounting for 25% of the variation in FVC. This shift underscores the increasing importance of water management strategies to maintain forest health. The Random Forest model demonstrated high predictive accuracy, achieving an R^2 value of 0.918 (RMSE of 0.016 and MAE of 0.013) for 2013 and 0.916 (RMSE of 0.018 and MAE of 0.014) for 2023, underscoring the model's robustness in handling nonlinear ecological processes. This study's insights are crucial for guiding sustainable forest management practices and assisting decision-makers in formulating strategies for resource management, environmental preservation, and future planning. This study underscores the necessity of adaptive management strategies that consider evolving climatic and hydrological conditions, emphasizing continuous monitoring and advanced technologies to ensure the resilience of forest ecosystems.

Keywords: fractional vegetation cover; climate and topography influences; groundwater level monitoring; machine learning for forest monitoring; forest ecosystem resilience; environmental change modeling

check for updates

Academic Editors: Tommaso Caloiero and Charles Jones

Received: 30 October 2024
Revised: 8 January 2025
Accepted: 10 January 2025
Published: 14 January 2025

Citation: Habeeb, H.N.; Mustafa, Y.T. Analyzing the Impact of Geoenvironmental Factors on the Spatiotemporal Dynamics of Forest Cover via Random Forest. *Earth* **2025**, *6*, 3. https://doi.org/10.3390/earth6010003

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Forests cover approximately 31% of the Earth's land area and are home to more than 80% of terrestrial animal, plant, and insect species. They are vital for maintaining the planet's ecological balance, serve as major carbon sinks, and play a crucial role in the global carbon cycle [1]. Forests influence weather patterns, protect watersheds, and prevent soil erosion, thereby supporting agricultural productivity and water quality [2]. Moreover,

they are integral to the livelihoods of millions of people worldwide, providing food, fuel, medicine, and income.

Forests face immense pressure worldwide due to deforestation, climate change, and human-induced activities, which are rapidly altering ecosystems. According to recent reports, global forest loss contributes to approximately 15% of greenhouse gas emissions, making forest ecosystems critical for mitigating climate change [3]. Deforestation also leads to the loss of biodiversity, the disruption of water cycles, and soil degradation, threatening both local environments and global ecological balance [4]. These challenges underscore the need for comprehensive studies on how environmental factors, such as climate, soil moisture, and topography, influence forest health and sustainability [5].

The degradation of forest ecosystems has severe repercussions. Deforestation and forest degradation contribute to climate change, biodiversity loss, and the disruption of water cycles [6,7]. These impacts are exacerbated by the increasing frequency and intensity of extreme weather events driven by global climate change. Therefore, understanding the factors that influence forest cover and developing reliable tools to monitor these changes are essential for conservation efforts and sustainable management [8].

Recent studies have underscored the interaction between forest and ecosystem factors, emphasizing the impact of influential ecosystem factors on forest management and planning [1,8]. Understanding the dynamics of forest cover is particularly crucial in regions such as the Duhok District in the Kurdistan Region of Iraq, which is experiencing marked environmental changes due to climatic variability and human activities. This region, characterized by its varied topography and semiarid climate, faces significant challenges, such as water scarcity and land degradation, which profoundly impact forest ecosystems [6,9].

Geoenvironmental factors such as groundwater levels, climate conditions, topography, and soil moisture fundamentally shape the distribution and health of forest vegetation. Groundwater sustains ecosystem functioning, whereas vegetation influences groundwater recharge and water availability [10]. Climate factors, including rainfall and temperature, affect soil moisture availability and evapotranspiration rates, which are crucial for tree growth [11]. Topography, which encompasses elevation, slope, and aspect, affects local climatic conditions and vegetation types [12,13]. Soil moisture is essential for seed germination, plant growth, and overall forest vitality [14].

In the semiarid regions of the Kurdistan Region of Iraq, such as the Duhok District, forests are particularly vulnerable to climate variability, soil moisture deficits, and anthropogenic activities [6,8]. These areas experience significant challenges, including water scarcity and land degradation, which directly affect forest cover [6,11]. Understanding the role of geoenvironmental factors in driving forest cover change is crucial for developing sustainable forest management strategies in such fragile ecosystems.

Despite the recognized importance of forests, there remains a significant gap in the development of robust tools and models for accurate forest cover estimation and ecosystem factor assessment [15]. Traditional methods of forest monitoring, which rely on field surveys and the manual interpretation of satellite imagery, are often labor-intensive, time-consuming, and subject to human error [16]. These methods are insufficient for capturing the dynamic and complex nature of forest ecosystems, particularly in the face of rapid environmental changes.

The advent of remote sensing technology and machine learning (ML) has revolutionized forest monitoring by providing comprehensive and up-to-date information on forest cover at various spatial and temporal scales [7]. Among ML models, Random Forest (RF) has gained particular prominence because of its ability to analyze complex, nonlinear ecological processes and provide accurate predictions of ecosystem structure and functionality [17,18]. For example, A.N. Ugbaja et al. [19] and Al-Abadi et al. [20] used RF to

highlight the importance of forests in regulating atmospheric conditions and maintaining ecological balance. Additionally, RF has been employed in forest fire susceptibility mapping in Serbia [17] and soil moisture and vegetation pattern analysis in the Batifa region of Iraq [18], demonstrating its versatility under different environmental conditions.

Moreover, studies such as those of Alganci et al. [21] and Koulelis et al. [22] emphasize the necessity of robust models like RF in evaluating the impacts of climate change and human activities on forest ecosystems. The application of RF has been extended to environmental studies in Iraq, where Al-Abadi, Fryar, Rasheed, and Pradhan [20] applied it for groundwater potential assessments, highlighting the broader applicability of RF in ecological and environmental research within the region. By integrating RF into forest cover change analysis, particularly in regions such as the Kurdistan Region of Iraq, this study contributes to a growing body of research on ML applications for sustainable forest management. These examples illustrate the capacity of RF to address pressing challenges related to forest monitoring and management under diverse ecological conditions.

Over the past decade, the Duhok District (located in the northwestern part of the Kurdistan Region of Iraq) has experienced significant changes in land use and land cover, driven by rapid population growth, industrialization, agricultural expansion, and climate variability [23]. These changes have exerted considerable pressure on the region's natural resources, particularly its forested areas. The natural forests of this region are dominated by oak species (*Quercus* spp.), while plantation forests, primarily composed of *Pinus brutia*, play a vital role in combating land degradation and promoting ecosystem resilience. *Pinus brutia* is an evergreen tree that was introduced to the Duhok District as part of reforestation efforts initiated in the 1970s [12]. It is highly adaptable to semiarid conditions and contributes much to soil stabilization, carbon sequestration, and local biodiversity. Understanding the ecological and geographical causes of its growth and distribution is key to sustainable management. Ecological and biological factors that influence the distribution of *Pinus brutia* are a preference for well-drained soils, moderate rainfall, and tolerance to drought and nutrient-poor conditions. Altitude and slope biophysical factors strongly influence its growth; thus, this species favors cooler temperatures and higher-elevation areas. Human activities involving deforestation, the expansion of agriculture, and urbanization further influence the survival and distribution of such forests. This study integrates these factors into predictive models with the intention of providing actionable insights for forest conservation.

Previous studies have emphasized the role of *Pinus brutia* in soil stabilization, carbon sequestration, and ecosystem services in arid and semiarid regions [24]. However, there is a serious lack of information regarding how geoenvironmental factors such as altitude, soil moisture, and groundwater levels interactively shape the distribution and vitality of this species in the Duhok District. Most studies in this region have concentrated on land-cover classification and change detection using remote sensing. Al-Abadi, Fryar, Rasheed, and Pradhan [20] and Mzuri, Omar, and Mustafa [6] investigated the region's hydrogeological and climatic conditions, providing valuable baseline data for forest cover analysis. However, there is a notable gap in studies that integrate multiple geoenvironmental factors and employ advanced ML techniques to model forest cover dynamics.

Our study specifically addresses two critical research questions. (1) How do geoenvironmental factors such as groundwater levels, soil moisture, and topography influence the spatiotemporal dynamics of forest cover in the Duhok District? (2) How can machine learning techniques such as Random Forest be employed to model and predict forest cover changes under varying environmental conditions? These questions are essential for guiding sustainable forest management practices, particularly in regions facing severe climatic stressors and water scarcity [20]. From a policy perspective, this research provides valuable

insights for decision-makers in prioritizing water management and reforestation efforts, ensuring the long-term resilience of forest ecosystems in the Kurdistan Region of Iraq [6].

To address the challenges facing forest ecosystems in Duhok, our study employs a robust analytical framework utilizing a variety of geoenvironmental factors categorized into five distinct datasets: fractional vegetation cover (FVC), groundwater levels, climate data, topography, and soil moisture. By integrating these datasets, our study provides a comprehensive understanding of the geoenvironmental factors affecting forest cover in Duhok.

The primary aim of this study is to model and predict changes in forest cover in the Duhok District over the past decade (2013–2023) using the Random Forest (RF) algorithm. Specifically, the study aims to (1) utilize the RF algorithm to predict forest cover changes and assess the relationships between forest cover and geoenvironmental factors, including climate, topography, soil moisture, and groundwater levels; (2) evaluate the importance of these factors within the model to gain insights into the potential drivers of forest cover change; (3) provide robust predictions of forest cover dynamics over time, demonstrating RF's applicability in handling nonlinear relationships in ecological systems; and (4) offer actionable insights and recommendations for sustainable forest management practices in the Duhok District, emphasizing the integration of predictive tools in environmental monitoring and conservation efforts.

2. Materials and Methods

This study utilized an ML approach that combined data and computational techniques to perform a spatiotemporal analysis of the impact of geoenvironmental factors affecting forest coverage growth and distribution. Figure 1 illustrates this process with a detailed flowchart divided into three phases.

In the first phase, we gathered the necessary data (features) from various sources and created thematic maps with a 15 m spatial resolution. The second phase involved testing for multicollinearity and correlation among the features, selecting only the appropriate features for the RF model. The final phase included the development, assessment, and analysis of the RF model.

We employed a combination of software tools to achieve the objectives of this study. ArcGIS (v. 10.8) was used for GIS mapping, and ENVI (v. 5.6) was used for image processing and analysis. The R Project for Statistical Computing (v. 4.3) facilitated the data analysis and model formulation. Additionally, the Google Earth Engine was used to process and create features of the satellite indices.

2.1. Study Area

This study investigates the Duhok District, which is located in the northwestern Kurdistan Region of Iraq (KRI) and covers an area of 1014 km^2. This district is divided into three subdistricts, namely, Duhok, Zawita, and Mangish, which are situated between latitudes 36°18′ and 37°20′ N and longitudes 42°20′ and 44°17′ E (Figure 2). The diverse terrain features elevations ranging from 431 to 1486 m above sea level, and slope gradients vary from flat areas to steep inclines of 0° to 62.5°. Climatically, Duhok lies at the intersection of arid and semiarid zones, as classified by the Köppen–Geiger system, and experiences a climate characterized by cold, wet winters and hot, dry summers [6]. Hydrogeologically, the aquifers in the district are mainly semiconfined to unconfined and are recharged predominantly by rainfall and snowfall. The climate of the region is influenced by Mediterranean weather patterns, which are characterized by hot, dry summers and cold, rainy winters, with precipitation levels ranging from 500 to 1000 mm annually and temperature varying from a high of 37 °C in the summer to near freezing in the winter [6]. However,

the region faces significant challenges owing to rapid population growth, industrialization, agricultural activities, and climate variability, which exert considerable pressure on water resources, leading to notable reductions in groundwater levels in recent years [20].

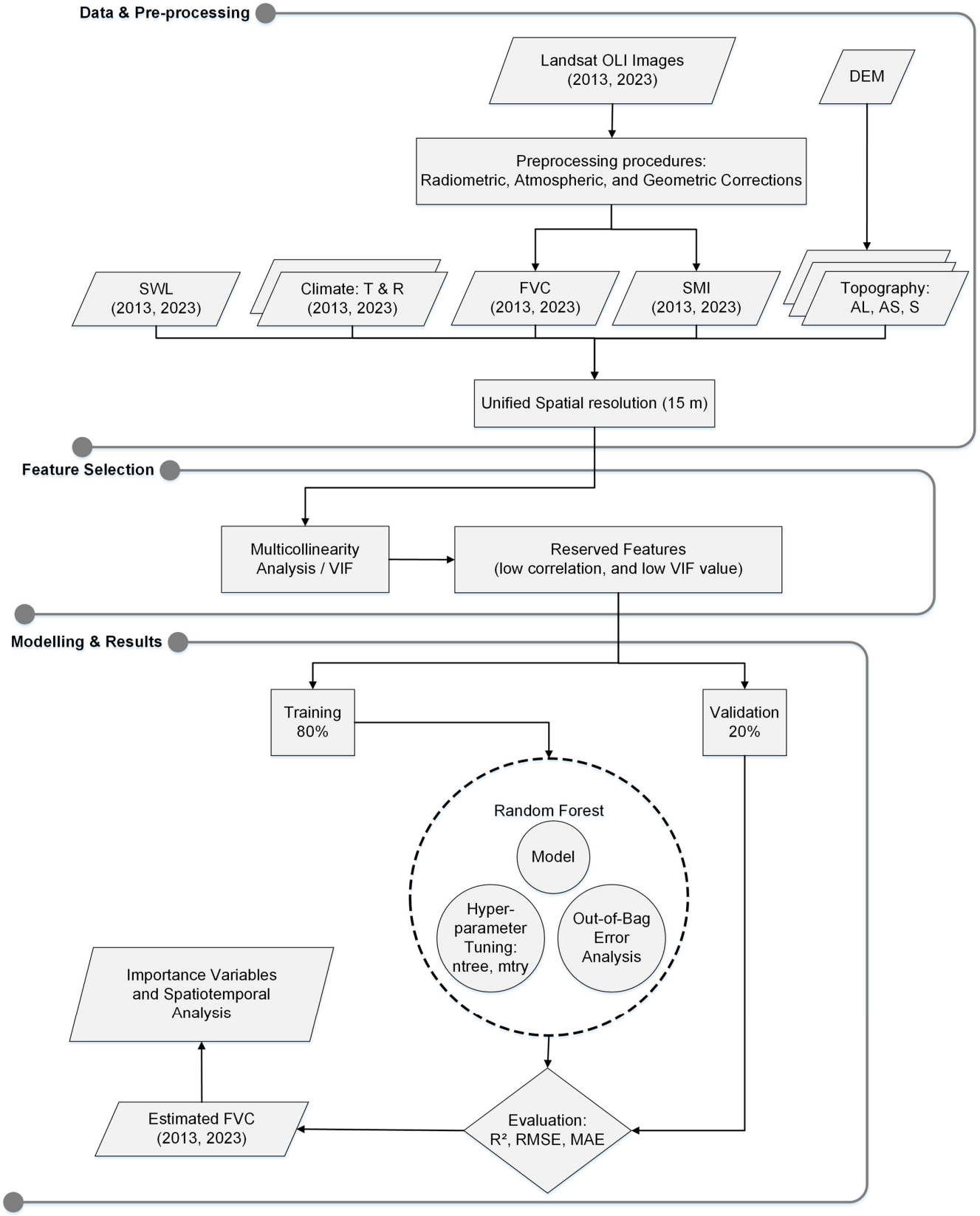

Figure 1. Flowchart of the adopted methodology.

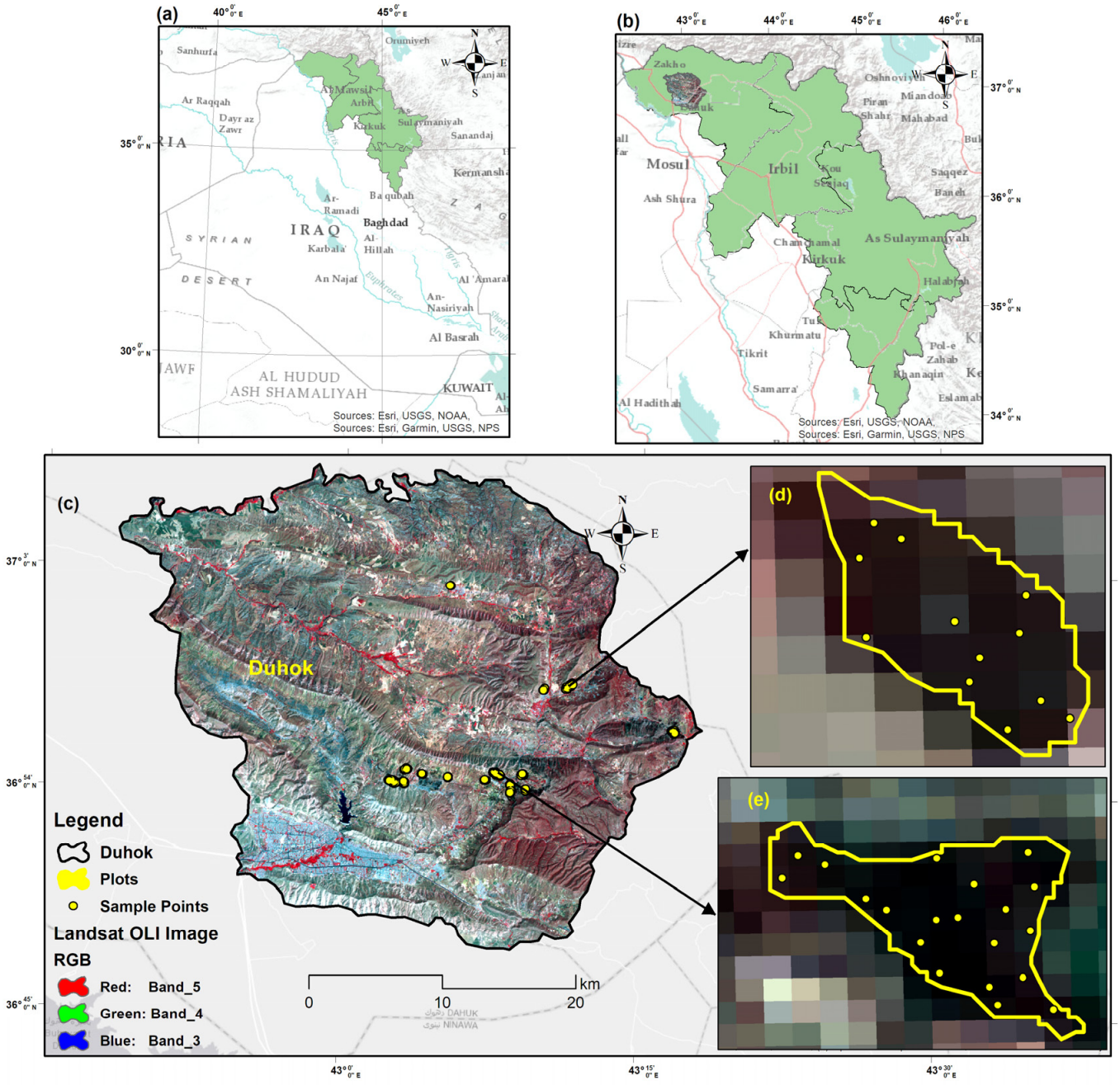

Figure 2. Study area map: (**a**) Iraq; (**b**) Kurdistan Region of Iraq; (**c**) Duhok District with overlaid Landsat OLI imagery; (**d**) example plot with sample points; (**e**) example plot with sample points.

Geologically, Duhok's landscape is composed of layers dating from the Jurassic period to recent formations, predominantly featuring Tertiary and Cretaceous sedimentary rocks, along with Quaternary deposits such as alluvium [6]. The soil composition across the district varies according to altitude, slope, vegetation, and geological context and is generally categorized as non-saline [25].

Much of the study area is composed of open forests and grasslands, with vegetation consisting of grasses and herbs interspersed among trees. The forests in the Duhok District include natural and plantation types, each with unique features. Natural forests, mainly oak trees (*Quercus* spp.) [26], are biodiversity hotspots crucial for ecological balance, providing habitats, stabilizing soil, regulating climate, and managing water resources. Plantation forests, primarily *Pinus brutia* [12], offer significant benefits in semiarid areas by improving soil stability and reducing erosion. Their year-round canopy mitigates heavy rains and

wind, reducing soil loss and retaining moisture. Additionally, *Pinus brutia* helps mitigate climate change by sequestering carbon and supporting local biodiversity by providing habitats and food for wildlife, fostering a resilient ecosystem. The majority of agricultural land is located around villages, integrating natural and anthropogenic landscapes within the district [27].

Although human activities influence land-use patterns in parts of the district, *Pinus brutia* plantations are managed under protective measures to maintain their integrity. There were no major incidents of fire and significant diseases reported during the study period, as recorded by the Duhok Forestry Department, and thus these plantations remained relatively intact and protected from direct human disturbances [28]. This is a stable environment, and it allows for the effective study of geoenvironmental factors without the interfering effects of recent large-scale anthropogenic disturbances.

2.2. Datasets

The geoenvironmental factors used in this study were categorized into five distinct categories, comprising eight factors: fractional vegetation cover (FVC), groundwater (Static Water Level—SWL), climate (rainfall—R; and temperature—T), topography (altitude—AL; slope—S; and aspect—AS), and the soil moisture index (SMI). In addition, Landsat satellite imagery served as the primary data source for deriving the FVC and SMI, utilizing Landsat OLI data at a 30 m spatial resolution, which was pan-sharpened to 15 m. These images were preprocessed via standard radiometric, atmospheric, and geometric corrections, which were carried out within the Google Earth Engine platform. This preprocessing was essential to ensure accurate data alignment with the UTM WGS84 coordinate system, Zone 38 N. A detailed description of the geoenvironmental factor data is provided in Table 1, and their spatial distribution patterns are illustrated in Figure 3.

Table 1. Summary of datasets: a comprehensive overview of the data types, descriptions, extracted factors, and sources utilized in this study.

Geoenvironmental Factors	Extracted Factors from the Data	Data Description	Source
Forest	Fractional Vegetation Cover (FVC)	Satellite image: Landsat OLI data with spatial resolution of 30 m (pan-sharpened to 15 m)	https://earthexplorer.usgs.gov (accessed on 30 August 2023)
Groundwater	Static Water Level (SWL)	Groundwater inventory	The Directorate of Ground Water and Wells
Meteorological Dataset (Vector Data)	Rainfall (R), Temperature (T)	Annual rainfall and average temperature data from 20 meteorological stations	The Directorate of Meteorology and Seismology—Duhok (http://duhokprovince.com/directorate-of-meteorology-seismology-duhok/) (accessed on 30 August 2023)
Soil Moisture	Soil Moisture Index (SMI)	Satellite image: Landsat OLI data with spatial resolution of 30 m (pan-sharpened to 15 m).	https://earthexplorer.usgs.gov (accessed on 30 August 2023)
Topography	Altitude (AL), Slope (S), Aspect (AS)	Raster dataset L-band SAR (ALOS-PALSAR) DEM, with 12.5 m spatial resolution, sampled to 15 m	(Alaska Satellite Facility) https://search.asf.alaska.edu/ (accessed on 30 August 2023)

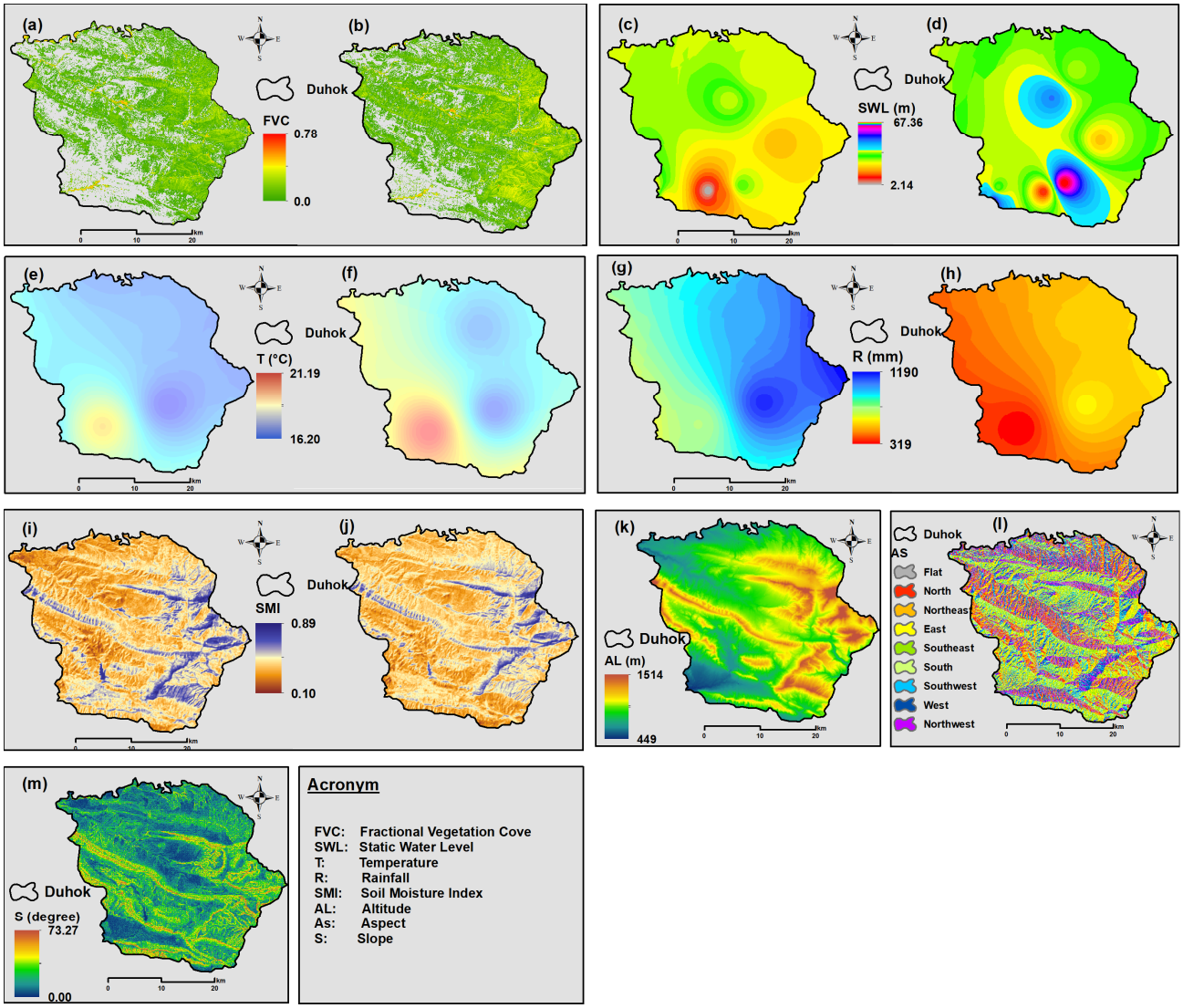

Figure 3. Geoenvironmental factors in forest cover analysis for Duhok: FVC ((**a**): 2013; (**b**): 2023), SWL ((**c**): 2013; (**d**): 2023), T ((**e**): 2013; (**f**): 2023), R ((**g**): 2013; (**h**): 2023), SMI ((**i**): 2013; (**j**): 2023), AL (**k**), AS (**l**), and S (**m**), all of which are used as inputs in the RF model.

2.2.1. Fractional Vegetation Cover

FVC is a vital parameter in remote sensing that serves as a quantitative measure of the proportion of ground covered by green vegetation within a pixel. It is instrumental in various environmental and climate-related studies, including studies on ecosystem balance, soil erosion, and climate change modeling [7].

In forested regions, FVC is a critical metric for assessing changes in forest cover, which can be attributed to various factors, such as logging, wildfires, or disease outbreaks. FVC provides a quantitative measure of the proportion of ground covered by vegetation, which is essential for monitoring forest health and guiding management practices [29].

In this study, we estimated the FVC from Landsat OLI images for 2013 and 2023. The Normalized Difference Vegetation Index (NDVI), which serves as a proxy for vegetation health, was first calculated via the near-infrared (NIR) and Red bands from the Landsat imagery. NIR refers to the near-infrared portion of the spectrum (wavelengths between 0.75 and 1.4 μm), which is strongly reflected by healthy vegetation, whereas Red refers to

the visible red portion (approximately 0.63–0.69 μm), which is absorbed by chlorophyll in plants. The standard NDVI equation [30] is as follows:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

The *NDVI* values were calculated via the Google Earth Engine (GEE) platform. The FVC was subsequently derived from the *NDVI* values via a linear transformation described by Zhao, Mu, Song, Liu, Xie, Zhong, Xie, Jiang, and Yan [1]. A common approach for calculating the FVC from the NDVI is expressed by the following equation:

$$FVC = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \quad (2)$$

where $NDVI_{min}$ and $NDVI_{max}$ represent the minimum and maximum *NDVI* values, corresponding to bare soil and dense vegetation cover, respectively. This formula adjusts the NDVI values to a scale representing zero to full vegetation cover, thus providing a quantitative measure of the fractional vegetation cover. The FVC raster images calculated for 2013 and 2023 are shown in Figures 3a and 3b, respectively.

2.2.2. Groundwater

This study employed static water level (SWL) data, a vital hydrogeological metric representing the equilibrium level of groundwater in a well when it is not being pumped. SWL measurements are extensively used in research to analyze groundwater conditions, which are critical for forestry studies. Understanding the SWL is essential for evaluating the water availability of trees and assessing the influence of groundwater on forest ecosystems.

Although specific studies on the application of SWL data in forestry have not been explicitly mentioned, the importance of the SWL in understanding groundwater resources is well known. For example, Ekwere et al. [31] and A.N. Ugbaja, U.A. Ugbaja, A.E. Bassey, E.A. Amah, and Offiong [19] explored the vulnerability of aquifers to contamination, which could influence forest health. Similarly, Thakur et al. [32] linked land-use changes to groundwater quality, potentially impacting forested areas. Additionally, Manda et al. [33] emphasized the potential of citizen science in collecting SWL data, suggesting its applicability in forestry studies to gather extensive data on groundwater levels in forested regions.

The SWL data for 2013 and 2023 were provided by the Groundwater Directorates of Duhok, covering 250 wells in the study area and its vicinity.

For the spatial interpolation of SWL, the Inverse Distance Weighting (IDW) method was employed via the ArcGIS Pro platform. Moreover, this interpolation method was used to generate spatial interpolation maps of rainfall and temperature as well. The IDW method was chosen for performing spatial interpolation in this paper due to its simplicity and efficiency in computation, demonstrating suitability in an appropriate manner in similar semiarid environments. Previous applications in the study area can be seen in studies by Noori et al. [34] and Mzuri, Omar, and Mustafa [6], where good results for localized spatial variability were observed. The choice of IDW allows for consistency with these regional studies so as to make meaningful comparisons and methodological coherence possible.

Although other methods, such as kriging and polynomial interpolation, were recognized for their strengths, IDW was selected in this investigation to maintain methodological alignment with previous work and to accurately capture localized environmental gradients typical of the study area. The decay function parameters were carefully calibrated to represent the observed spatial patterns accurately. The resulting interpolated maps align

well with known environmental gradients, supporting the appropriateness of this method for the variables under investigation.

This method produced continuous thematic maps for SWL in 2013 (range: 2.14–38.42) and 2023 (range: 9.75–67.36), which are presented in Figure 3c and d, respectively. These maps are critical for understanding the influence of water availability on forest ecosystems, providing insights into how groundwater resources fluctuate over time and their impact on forest health.

2.2.3. Climate

Climate factors, particularly rainfall (R) and temperature (T), play crucial roles in influencing forest cover dynamics and distribution. These climatic variables critically determine the viability and sustainability of vegetation. Rainfall directly affects soil moisture availability, which is essential for tree growth, whereas temperature regulates evapotranspiration and photosynthesis rates. Precipitation is a key determinant of forest health in semiarid regions such as the KRI, where forest cover is highly dependent on annual and seasonal rainfall patterns, which control the availability of water for vegetation [34]. Insufficient rainfall can lead to reduced soil moisture, making it difficult for trees to grow and maintain healthy ecosystems. Conversely, areas with adequate rainfall are able to sustain dense vegetation and contribute to higher FVC.

The role of temperature is similarly significant, as warmer temperatures increase evapotranspiration rates, further stressing water availability and impacting tree growth [9]. Recent studies have examined the impacts of R patterns and temperature fluctuations on forest density and structure in northern Iraq. Al-Quraishi, Gaznayee, and Crespi [11] specifically investigated the severity of drought in the KRI and its effects on vegetative cover, including that of forests. Furthermore, Mzuri, Omar, and Mustafa [9] discussed how T fluctuations have been linked to changes in the phenological cycles of vegetation, particularly forest species in semiarid zones, influencing forest cover and structure. These insights are crucial for understanding and managing forest ecosystems in response to climate variability and underscore the need for targeted conservation strategies in these vulnerable regions.

Our research employed data from 2013 and 2023, which were gathered from 20 meteorological stations located in and around the study area, to analyze annual R and average T. The data were processed via the IDW interpolation method within ArcGIS Pro, which was selected as the optimal technique for creating continuous raster maps of these climatic variables [34,35]. This interpolation provided critical spatial coverage of rainfall and temperature, both of which significantly affect evapotranspiration and soil moisture conditions, crucial factors influencing vegetation health in the study area. Thematic maps of these climatic variables were generated for both years (Figure 3e–h).

2.2.4. Soil Moisture

The soil moisture index (SMI) quantitatively reflects the soil moisture content at various depths and plays a critical role in governing forest cover dynamics. Soil moisture directly influences the ability of trees to absorb nutrients and maintain photosynthesis rates, which are essential for plant growth, survival, and reproduction [14]. The SMI is particularly valuable in ecological modeling, drought monitoring, and water resource management, as it provides crucial insights into forest health and distribution, especially in areas vulnerable to climate change and human activity [36,37].

Higher soil moisture levels tend to promote healthier forest ecosystems, whereas areas suffering from low soil moisture often show signs of degradation and loss of forest cover. This relationship is particularly important in regions facing irregular rainfall patterns,

such as the Duhok District, where water stress can significantly impact forest vitality [18]. Adequate soil moisture supports vital plant physiological functions, such as photosynthesis and nutrient uptake, which are essential for vegetation survival [14].

The SMI can be derived for extensive areas via remote sensing technologies, offering a practical approach for monitoring soil moisture across large and inaccessible regions. In our study, the SMI was derived from the Land Surface Temperature (LST), which was obtained from the thermal infrared bands of Landsat OLI satellite imagery [18]. The LST values were processed via the Google Earth Engine to obtain the necessary thermal bands for the years 2013 and 2023. The SMI was calculated via the methodology outlined by Saha et al. [38] via the following formula:

$$SMI = \frac{LST_{Max} - LST}{LST_{Max} - LST_{Min}} \quad (3)$$

where LST_{Max}, LST_{Min}, and LST are the maximum, minimum, and current values of the retrieved LST, respectively. This dual-year analysis facilitated the examination of temporal changes in soil moisture, providing insights into trends and variability that may influence forest vegetation cover over time (Figure 3i,j). This method provided valuable information on the soil moisture status across the Duhok District for both 2013 (range: 0.10–0.89) and 2023 (range: 0.11–0.84), contributing to the analysis of the geoenvironmental factors influencing forest cover change.

Recent studies have highlighted the application of the SMI derived from LSTs to monitor forest health and aid in the sustainable management of forested areas [18,39].

2.2.5. Topography

Topographical factors significantly influence forest cover distribution. These factors were extracted from the Advanced Land Observing Satellite Digital Elevation Model (ALOS PALSAR DEM) with a resolution of 12.5 m, which was resampled to 15 m to ensure consistency with other factor resolutions. When compared with other freely available DEM sources, the ALOS DEM was chosen for its superior accuracy and reliability [21]. The key topographical elements, such as altitude (AL), aspect (AS), and slope (S) (Figure 3k–m), were derived using the spatial analyst tools within ArcGIS Pro, providing a detailed topographic profile of the study area [12]. These variables play crucial roles in shaping vegetation patterns by influencing local climatic conditions, soil moisture, and forest type distributions [13].

Altitude and topography play critical roles in determining the spatial distribution of forest cover. In mountainous regions such as the Duhok District, higher altitudes are often associated with cooler temperatures and increased precipitation, which can increase vegetation growth and create distinct ecological zones [12,13]. Altitude influences not only the microclimate but also the types of species that can thrive at different elevations [8]. Additionally, slope and aspect, which are directly linked to altitude, further affect solar radiation exposure, water retention, and soil erosion rates, leading to variations in vegetation distribution [1].

In Duhok, altitude varies between 435 and 1486 m, resulting in a diverse array of forest types at different elevations [12]. This elevation gradient significantly impacts the composition and distribution of forest ecosystems across the study area. Forests at higher elevations tend to experience less human intervention, whereas those at lower altitudes, which are more accessible, are often subject to deforestation and land-use changes. Furthermore, slope (S) and aspect (AS) are critical factors in ecological processes such as fire management and soil moisture retention and in determining the ecological suitability of vegetation types [8]. The slopes in the study area ranged from 0° to 62.5°, whereas the

aspects were classified into nine categories, ranging from flat to northwest-facing slopes, all of which contribute to variations in vegetation distribution.

2.3. Data Setup and Sampling

In this study, we focused on plantation trees, specifically *Pinus brutia*. This focus is justified for two primary reasons. First, *Pinus brutia* is the most dominant evergreen species in the district, with plantations established as early as 1970 [12]. Second, *Pinus brutia* forests represent a major degraded forest type in the district.

The Duhok District contains 76 planted areas/plots (polygons) of *Pinus brutia* [12]. For this study, we selected 21 plots to ensure a fair comparison between the two distinct periods: 2013 and 2023. These plots were chosen based on their presence during both years and their dense forest cover, which allowed for more accurate analysis.

Within each of these 21 polygons, 341 sample point locations were randomly distributed via stratified random sampling (Figure 2c–e). This sampling process was designed to maintain a minimum distance of 15 m between neighboring sample points, ensuring that the distribution matched the spatial resolution of the satellite imagery. This approach was conducted using ArcGIS Pro software version 2.5. Moreover, 341 sample points were collected for each of the two years, 2013 and 2023, to generate training labels for the RF model for both time periods.

Thematic maps of all factors (FVC, SWL, R, T, SMI, AL, AS, and S) were integrated and stacked via ENVI software. We subsequently extracted the values of the 341 sampling points across all the layers via ArcGIS Pro, creating a comprehensive dataset that encompasses all the variables.

In this study, the spatial estimation and prediction of *Pinus brutia* FVC in the entire study area was made using an RF model. While the relationships between FVC and geoenvironmental factors were explored, the primary objective was to predict FVC values at locations where direct measurements may not be available. By spatially predicting FVC, we produced high-resolution maps that give insight into its distribution across the study area, allowing for a proper understanding of the spatial patterns and the environmental factors determining these patterns.

2.4. Feature Selection

Methods for selecting relevant features were employed to eliminate unnecessary, redundant, and disruptive variables, thereby increasing the accuracy of the model. The Variance Inflation Factor (VIF) was used to identify multicollinearity among features. The VIF quantifies the increase in the variance of an estimated regression coefficient when predictors are interrelated, serving as an indicator of linear dependency among them. A VIF value above 10 typically indicates excessive multicollinearity, as outlined by O'brien [40]. Additionally, the Caret and Car libraries in R software (V. 4.3) were used to calculate both the VIFs and Pearson correlation coefficients (visualized through heatmaps) to assess the relationships between the selected features.

2.5. Random Forest

The RF algorithm, introduced by Breiman [41], is extensively used for regression and classification tasks and has shown significant effectiveness in forest mapping [17]. RF, an enhancement of bagged decision trees, features minimal tuning parameters and comprises numerous uncorrelated decision trees [42]. It builds on the Classification and Regression Tree (CART) model and employs bootstrap sampling to create multiple datasets of the same size as the original dataset through sampling with replacement, which involves the repeated selection and exclusion of records [43].

In this study, the RF model was implemented via the randomForest package in R software. Approximately 80% of the data (273 samples) were used as the training set, whereas the remaining 20% (69 samples) served as the testing set for model evaluation. Out-of-bag (OOB) errors, which are crucial for assessing model accuracy, were analyzed. The model was conFig.d with 1000 trees (ntree) to prevent overfitting, although this parameter was adjusted on the basis of dataset specificity. The tuning of the trajectory parameter was critical for determining the optimal number of explanatory variables at each node, thereby enhancing the model's performance.

Variable importance was assessed by analyzing the discrepancies in error rates between the new data and the original OOB data, which were influenced by the size of the input variable subset and the total number of trees.

2.6. Accuracy Assessment

The accuracy of the RF model was rigorously evaluated via the same partition as the dataset. This assessment utilized three statistical metrics to quantify model performance: the coefficient of determination (R^2—Equation (4)), mean absolute error (MAE—Equation (5)), and root mean square error (RMSE—Equation (6)). These metrics help to confirm the robustness of a model, where higher R^2 values and lower MAE and RMSE values indicate greater predictive accuracy [44]. The assessment was performed via R Studio (V. 2023.12.1+402), which employs additional packages such as Caret, randomForest, and ggplot2 to increase the analytical depth [42]. These tools are instrumental in providing a comprehensive evaluation of model performance through detailed statistical analysis. The adopted metrics for evaluating model accuracy are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)^2}{n}} \quad (5)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{x}_i - x_i| \quad (6)$$

3. Results and Discussion

3.1. Factor Assessment and Analysis

In order to effectively analyze the geoenvironmental factors associated with *Pinus brutia* FVC, it is important to clarify the rationale and limitations associated with our climate data interpolation. Although IDW interpolation offers a continuous spatial representation of the variation in climate variables, it can produce relatively smooth gradients that mask the often-subtle local differences, particularly in temperature distribution. This effect is due both to the interpolation method itself and the color gradients used in the maps.

While the interpolated maps have the appearance of uniform patterns spatially, temporal differentiation still occurred. For example, in 2013, temperatures ranged from 16.2 °C to 19.8 °C, while in 2023, they ranged from 16.5 °C to 21.19 °C. Additionally, R ranged from 649 mm to 1190 mm in 2013, and from 319 mm to 601 mm in 2023. Other methods under these conditions, such as the use of satellite-derived climate products, may require further substantial downscaling and may still lack appropriateness for the proper expression of local climatic regimes. Therefore, IDW interpolation based on ground-station data—as also suggested by previous studies of our region, such as Noori, Hassan, and Mustafa [34]—remains a sound methodological choice given the available datasets.

Having acknowledged these considerations, we now shift our focus to an evaluation of the factors influencing *Pinus brutia* FVC. Specifically, we begin by assessing the relationships among the explanatory variables to ensure that multicollinearity does not undermine the robustness of our modeling approach. In modeling FVC susceptibility, understanding the relationships among independent variables is crucial to avoid multicollinearity, which can distort model interpretation. To address this, we employed the VIF test. VIF values exceeding 10 indicate a strong multicollinearity concern that could compromise model reliability [40].

Figure 4a,b illustrate the correlation matrices for 2013 and 2023, respectively. The T factor consistently exhibits high correlations with R and AL across both years, indicating a significant overlap in their predictive information. Other correlations of note include those between SMI and R, SMI and AL, and S and R. However, these correlations are weaker, with none exceeding 0.7. The consistent patterns in correlations over the past decade suggest stable, interrelated dynamics among these factors influencing forest vegetation cover.

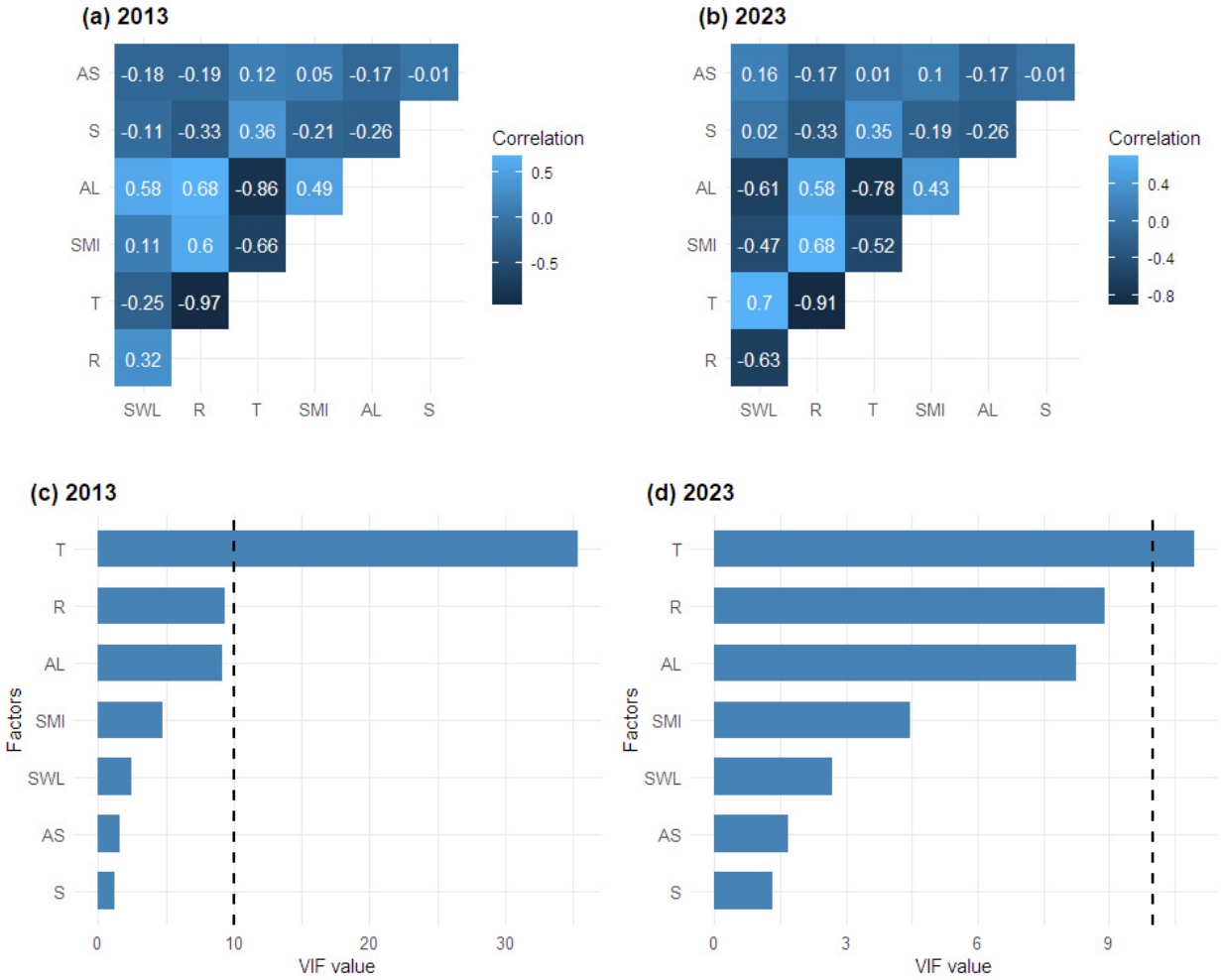

Figure 4. Heatmaps and Variance Inflation Factor values showing the correlation and multicollinearity of geoenvironmental factors (T: temperature; R: rainfall; AL: altitude; AS: aspect; S: slope; SMI: soil moisture index) for the years (**a**,**c**) 2013 and (**b**,**d**) 2023. The heatmaps (**a**,**b**) display the Pearson correlation coefficients between the factors, whereas the VIF values (**c**,**d**) help identify multicollinearity.

Figure 4c,d present the VIF values for each factor in 2013 and 2023. The highest VIF values, above the threshold of 10, were recorded for the T factor, highlighting substantial multicollinearity with other variables in both years. R and AL also present relatively high VIF values but these do not exceed 10, indicating manageable collinearity. According to

Edeh et al. [45], such VIF values require careful consideration because high multicollinearity can skew predictive models by attributing disproportionate importance to specific features.

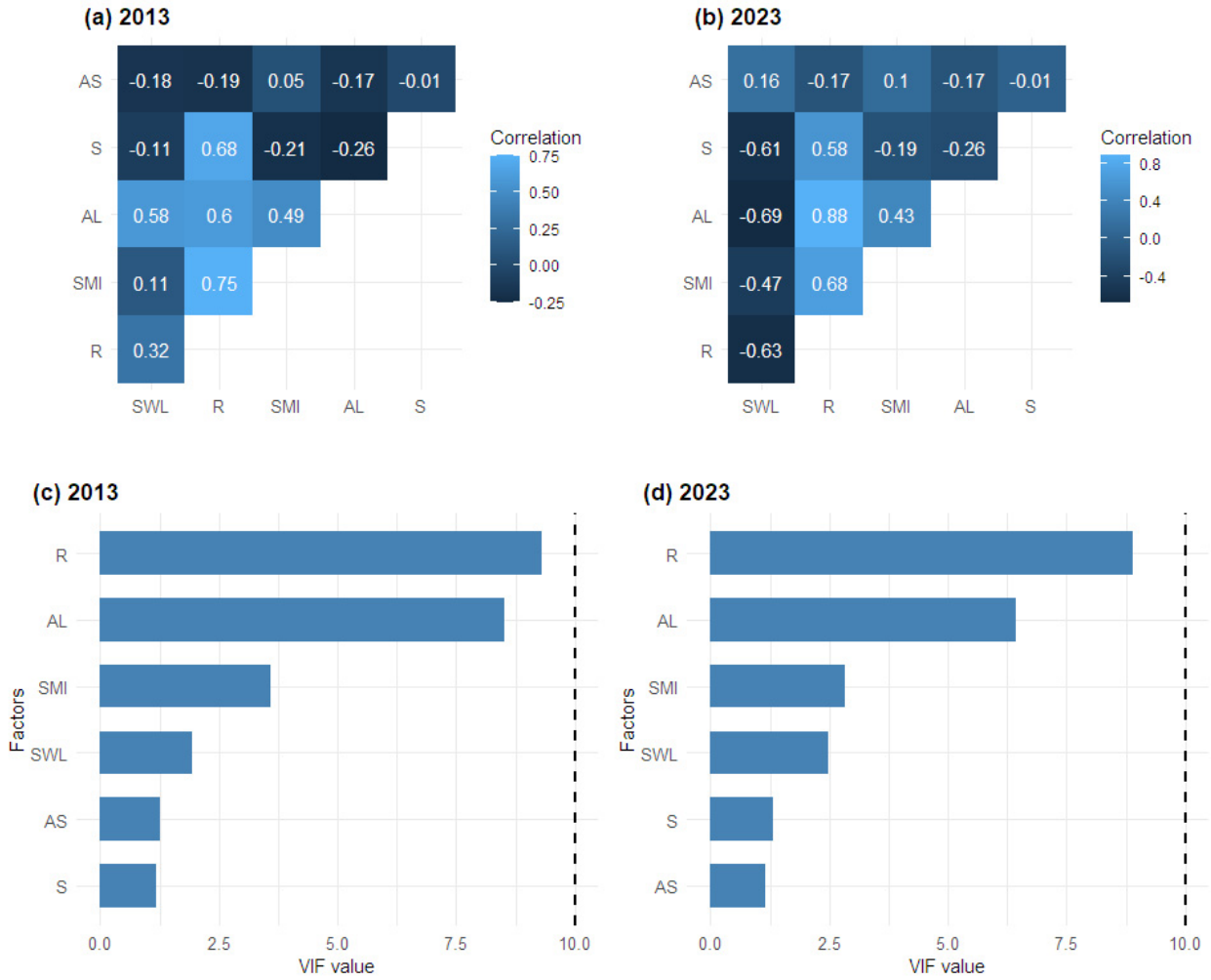

Figure 5. Correlation and multicollinearity analyses of geoenvironmental factors after excluding temperature (T) for (**a**,**c**) 2013 and (**b**,**d**) 2023. The heatmaps (**a**,**b**) illustrate the correlation structure of the remaining factors, whereas the VIF values (**c**,**d**) show reduced multicollinearity after temperature exclusion.

Given these findings, we excluded T from further modeling despite its critical role in forest growth and distribution, aligning with the best practices suggested by Dormann et al. [46]. This decision is supported by research that emphasizes that theoretically significant variables should sometimes be removed to avoid misleading outcomes [47,48]. Subsequent analysis after excluding T revealed that the remaining factors (R, AL, SMI, SWL, S, and AS) have acceptable VIF values well below 10, mitigating multicollinearity.

Assessing multicollinearity via VIF values and correlation matrices proved instrumental in refining the forest cover model. Removing T reduced redundant predictive information, thereby enhancing the robustness of the model. This careful selection process aligns with previous studies and confirms the importance of optimizing feature selection for accurate forest cover estimations [8,49]. Figure 5 displays the heatmap correlations and VIF measurements for both years (Figure 5a,c for 2013 and Figure 5b,d for 2023) after T was excluded, indicating acceptable correlations and VIF values for the remaining factors.

3.2. Model Performance

The RF model's performance was evaluated via key statistical metrics: R^2, RMSE, and MAE. These metrics provide a comprehensive assessment of the model's accuracy and predictive ability. The RF model achieved R^2 values of 0.918 and 0.916 for 2013 and 2023, respectively, indicating a strong correlation between the predicted and observed FVC as presented in Figure 6. The low RMSE and MAE values further validate the robustness of the model's predictions.

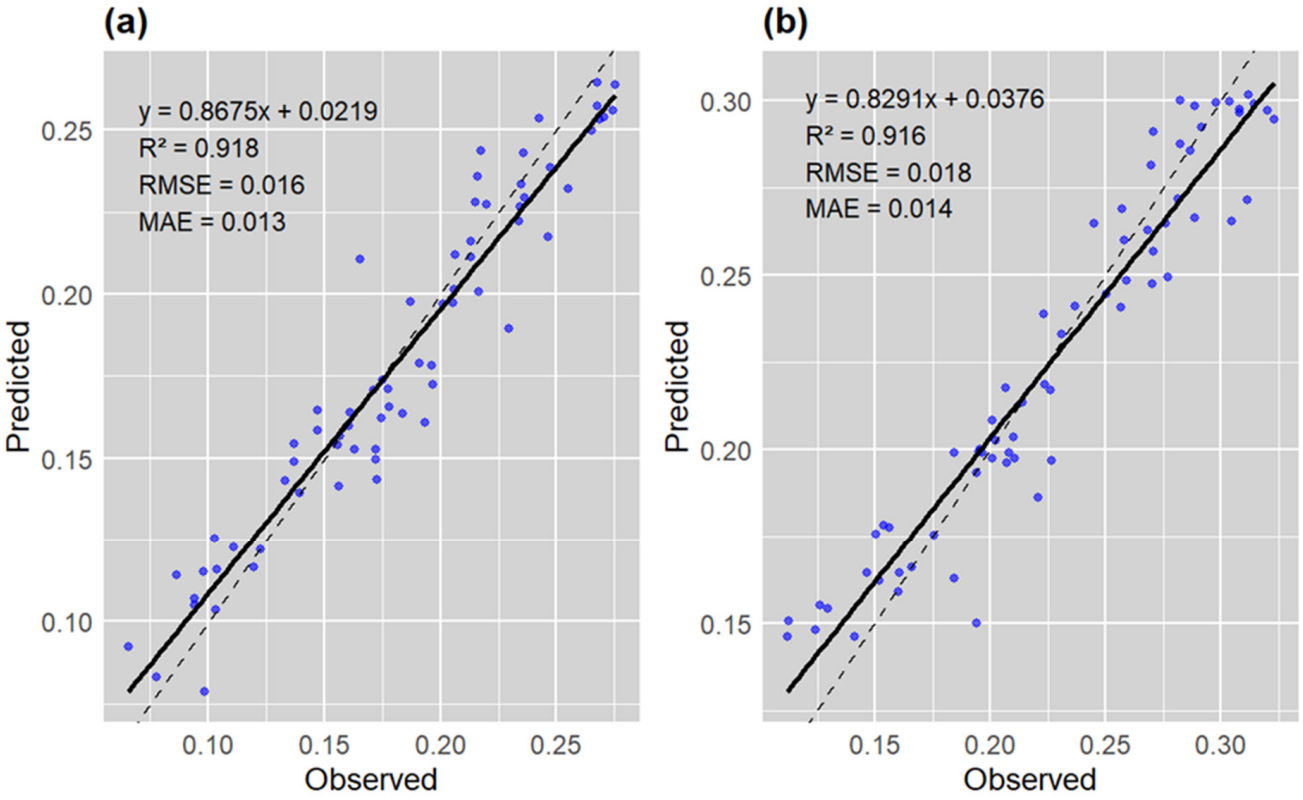

Figure 6. Scatter plots comparing the observed and predicted fractional vegetation cover via the Random Forest model for (**a**) 2013 and (**b**) 2023. Each dot represents a predicted FVC value compared with its corresponding observed value, with the dashed diagonal line indicating perfect agreement (1:1 correlation). The R^2, RMSE, and MAE values quantify the model's predictive accuracy for each year.

Figure 6 displays scatter plots comparing the observed and predicted FVC values via the RF algorithm for 2013 and 2023. Figure 6a represents the model tested on the 2013 data, and Figure 6b corresponds to the 2023 model. The consistently high R^2 values suggest minimal overfitting and a robust predictive ability. The 2013 model achieved a slightly higher R^2 than the 2023 model did, with both maintaining similarly low RMSE and MAE values. This implies that the 2013 model predicted FVC more accurately than the 2023 model did, possibly because of better data quality or less variability in environmental conditions during the former period.

These metrics indicate that the RF model performed well under different conditions. For example, the RMSE values of 0.016 (2013) and 0.018 (2023) are lower than those reported by Song, Wang, He, Wang, and Liang [16], where an RMSE of 0.17 was achieved for predicting FVC in China. This demonstrates the robustness of our model, emphasizing its adaptability and ability to capture key geoenvironmental factors influencing forest cover.

Although cross-validation, such as k-fold cross-validation, is commonly used to enhance model validation, it was not performed in this study because of the limited dataset

available for the Duhok region. Cross-validation can provide a more comprehensive assessment of a model's performance by splitting the data into several subsets and testing the model on each. However, given the high R^2 and low RMSE and MAE values achieved in this study, the existing metrics provide sufficient validation of the model's performance. These metrics are widely used in similar studies and are recognized as reliable indicators of model accuracy and generalizability [16].

The results indicate that our RF models provide accurate predictions for both 2013 and 2023, offering valuable insights into forest cover dynamics. Nonetheless, it is important to acknowledge that cross-validation could provide further validation of the model's robustness, particularly for studies with larger datasets. Future research could explore the application of k-fold cross-validation to evaluate model performance more thoroughly and minimize any potential biases in training and testing datasets [50].

Despite the absence of cross-validation in this study, the reported R^2, RMSE, and MAE values confirm the robustness and reliability of the RF model for predicting forest cover. The low error rates and high predictive accuracy demonstrate the model's ability to capture the spatial variability in forest cover in the Duhok region. As larger and more diverse datasets become available, cross-validation could be used to further enhance model validation and support broader applications of RF models in forest management.

Overall, our models demonstrated compelling predictive performance and reliability, validating their application in monitoring forest health and managing vegetation cover. The accuracy and generalizability of the results offer promising opportunities for the integration of RF models into large-scale environmental assessments, ensuring sustainable forestry practices and informing decision-making.

Although the selection of the RF algorithm for this study was based on its robust predictive performance against nonlinear ecological processes, we should note some important limitations relating to statistical inference or causality. This choice puts more emphasis on prediction rather than explanation, since RF does not deal directly with statistical significance or any mechanisms behind observed relationships.

3.3. Significance of Explanatory Variables

In the analysis of factors influencing forest cover in the Duhok District over the last decade, our research employed an RF model to unravel the dynamic interplay of geoenvironmental factors across different time points. The RF model is particularly esteemed for its robustness in handling nonlinear relationships and its ability to provide insights into variable importance through a machine learning approach. This model's ability to aggregate and interpret complex datasets allows for a detailed understanding of how these variables' significance shifts over time.

The results for 2013 and 2023 depict a notable shift in the factors affecting forest cover within the district. In 2013, the primary influences were AL, R, SWL, and the SMI, suggesting a significant dependency of forest cover on climatic and edaphic factors (Figure 7a). By 2023, however, there was a marked increase in the influence of the SWL and SMI, while the significance of AL slightly declined, and R remained a stable contributor (Figure 7b). This evolution underscores the increasing sensitivity of forest ecosystems to water-related parameters, reflecting broader changes in both climate and water availability.

Between 2013 and 2023, the declining influence of R and the increasing importance of soil moisture are particularly noteworthy. This shift could be attributed to the increasing variability in rainfall patterns, which has led to more extreme weather events, such as prolonged droughts and short, intense rainfall periods. These shifts in rainfall are consistent with the global climate change trends observed in semiarid regions such as the Duhok District [34]. As rainfall patterns became more unpredictable by 2023, soil moisture has

emerged as a more critical factor influencing vegetation dynamics. The SMI, which reflects the ability of land to retain water, plays an increasingly important role, especially as climate-driven changes cause rain to fall irregularly [18].

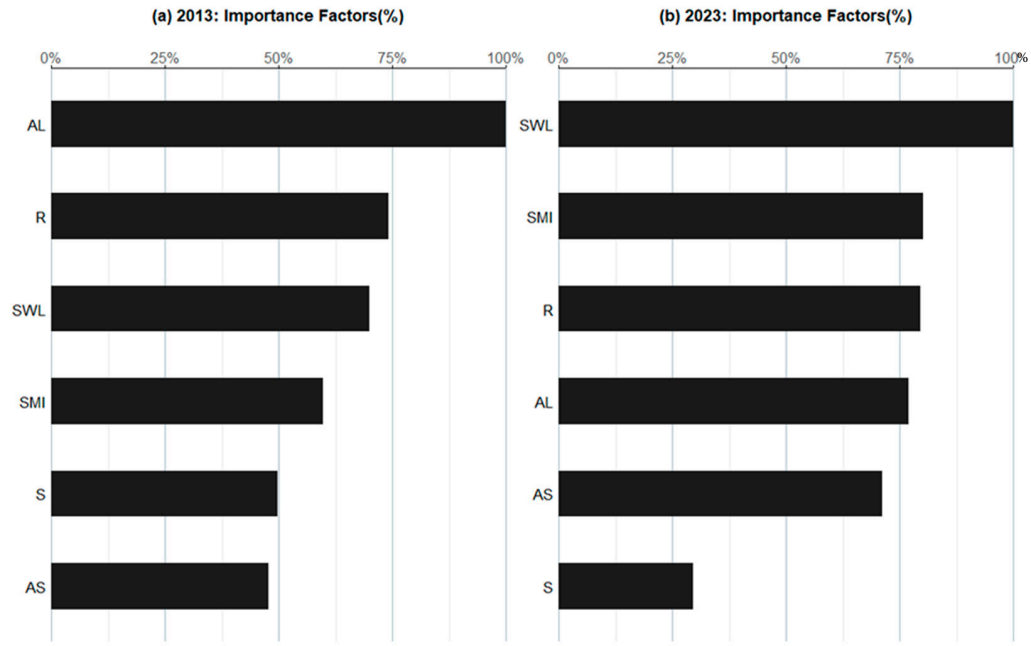

Figure 7. Bar plots showing the relative importance of geoenvironmental factors influencing fractional vegetation cover as determined by the Random Forest model for (**a**) 2013 and (**b**) 2023. The height of each bar represents the contribution of each factor to the model's predictive accuracy, indicating the dominant factors affecting FVC in the Duhok region for both years.

Furthermore, while the importance of R has decreased over the past decade, the increasing impact of soil- and water-related factors, such as the SMI and SWL, highlights the dynamic nature of soil–water interactions and their critical role in forest ecology. These findings align with those of Koulelis, Proutsos, Solomou, Avramidou, Malliarou, Athanasiou, Xanthopoulos, and Petrakis [22], who noted the increasing influence of hydrological parameters on vegetation dynamics in arid regions. The prominence of soil moisture parameters in 2023 is also reflective of the global trends of increased temperature and altered R patterns, further emphasizing the need for adaptive strategies to manage forest resources effectively.

AL remained an important factor throughout the decade, particularly in the Duhok region, where higher elevations benefit from more stable climatic conditions and less human activity. However, human activities such as deforestation and land conversion for agriculture have become more prevalent at lower altitudes, leading to a reduction in forest cover and increasing the impact of other variables such as soil degradation and increased temperature. This suggests that altitude acts as a stabilizing factor for forest health, especially in areas that are less exposed to human disturbances.

The decade-long analysis in Duhok highlights the nuanced and evolving nature of geoenvironmental influences on forest cover. The increased importance of hydrological factors such as SWL and the SMI by 2023 could be indicative of broader climatic shifts affecting the region. These findings echo the global trends reported by Hending et al. [51], who highlighted the profound impacts of climate change on forest ecosystems worldwide.

Moreover, the methodological strength of the RF model lies in its capacity to integrate diverse datasets and capture temporal changes in factor significance, providing a window into both immediate impacts and potential long-term changes in forest cover dynamics. This

adaptability offers invaluable insights for policymakers and conservationists, emphasizing the need for the continuous monitoring and adaptation of forest management practices to meet evolving environmental conditions.

The results of this study highlight the increasing importance of water-related factors and the shifting role of altitude and precipitation over the decade. These insights are crucial for designing targeted and effective forest management strategies that align with the most impactful geoenvironmental variables in Duhok.

3.4. Spatiotemporal Analysis (2013 and 2023)

The spatiotemporal analysis of FVC in the Duhok District provides critical insights into the changes in forest cover over a decade, particularly into how various geoenvironmental factors influenced these changes. This section examines the key factors affecting forest cover in 2013 and 2023, emphasizing the shift in primary influences on FVC.

To analyze the spatiotemporal trends and behavior of geoenvironmental factors with respect to changes in forest cover, we randomly selected 40% of the total number of samples (40% of the 341 samples). This sampling was necessary because plotting the entire dataset was not feasible for interpretation. To ensure that all the geoenvironmental factors were comparable, we normalized the data via z-score normalization. This process transforms each variable so that it has a mean of zero and a standard deviation of 1. Consequently, the *y*-axis values in the plots range between approximately −2 and 2 (Figure 8). Normalization allows for a fair comparison of trends and variations across different parameters by standardizing their scales. The *x*-axis in Figure 8 represents the index of the randomly selected samples used for spatiotemporal analysis. In other words, each point on the *x*-axis corresponds to a different sample from the dataset (40%). This allows for a comparison of how different geoenvironmental factors, normalized by z-score, behave over time for each sample.

The shaded regions in Figure 8 highlight areas where significant changes in forest cover and the influencing geoenvironmental factors have occurred. These areas are color-coded to indicate the dominant environmental influences during the respective years (2013 and 2023).

In 2013, FVC was influenced mainly by climatic and edaphic factors, with AL and R being the dominant variables. These factors indicate that forest ecosystems during this period were highly dependent on climatic conditions and elevation. Specifically, regions with higher altitudes and adequate rainfall present better forest cover, as shown in the highlighted areas (light green color) in Figure 8. These areas underscore the sensitivity of forests to altitude and rainfall, with higher elevations and sufficient rainfall promoting denser vegetation cover [12]. On the other hand, areas with low significance levels of altitude and rainfall present reduced forest growth, as shown in the highlighted gray area in Figure 8.

By 2023, the analysis reveals a shift in the influencing factors, with the SMI and SWL becoming more significant, whereas the influence of rainfall slightly declined. This shift suggests that water-related parameters have become increasingly critical in forest ecosystems, likely reflecting broader climatic changes and variations in water availability. The increased significance of soil moisture and groundwater in 2023 indicates that forest cover is becoming more dependent on these water-related factors (highlighted in pink in Figure 8). This is reflective of the global impacts of climate change, which has led to higher temperatures and more extreme weather events, increasing evapotranspiration and causing soil moisture deficits [9]. Consequently, forest ecosystems in the Duhok region are increasingly vulnerable to climate-induced water stress, emphasizing the need for effective water management strategies in forest conservation [20].

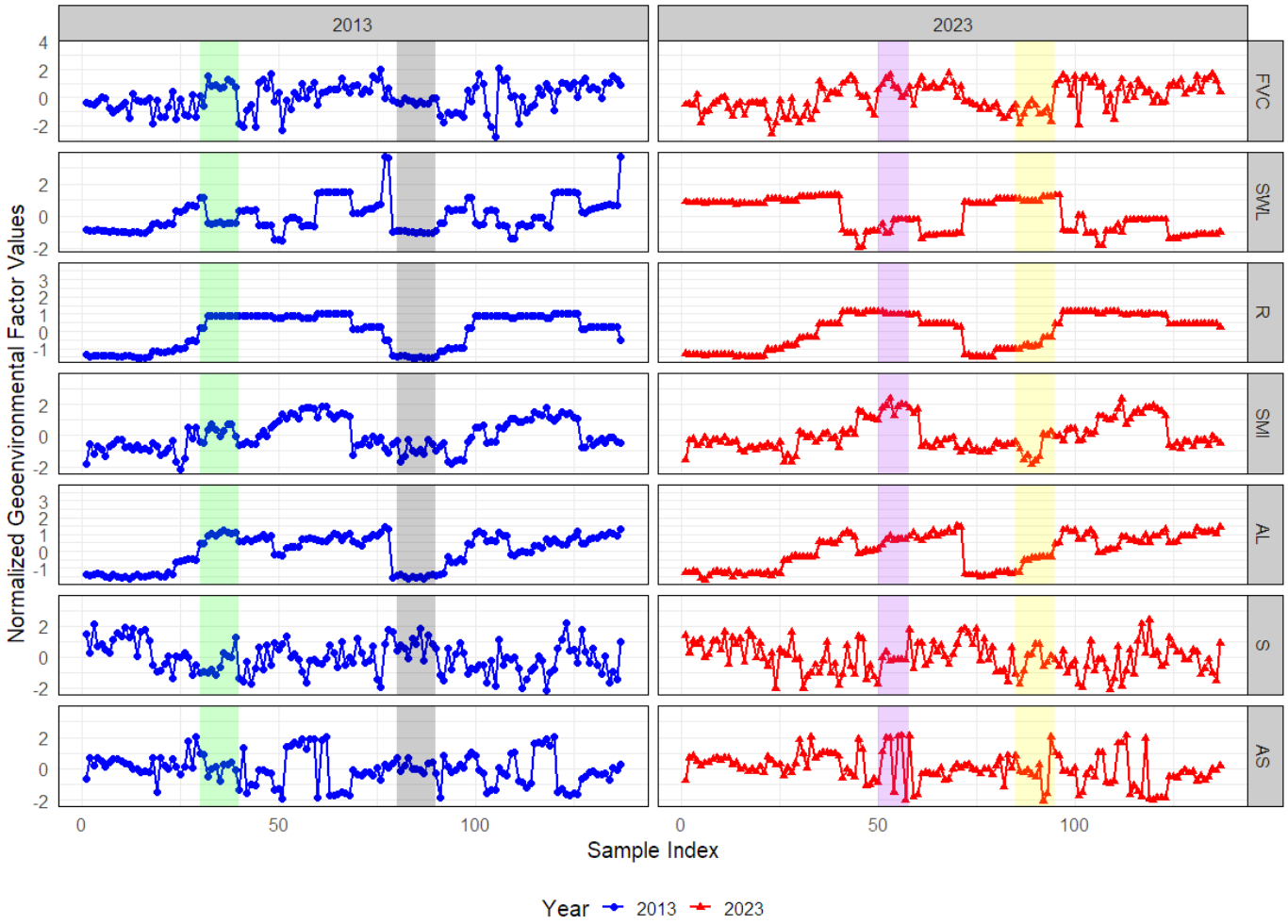

Figure 8. Comparative analysis of geoenvironmental factors and their influence on fractional vegetation cover for 2013 and 2023. This figure highlights how the relationship between environmental factors and forest cover has changed over the decade, with a particular focus on shifts in altitude, precipitation, and soil moisture distributions.

The elevation-dependent nature of forest health also became more pronounced in 2023, with forests at higher altitudes remaining relatively stable, whereas lower-altitude forests suffered from both climate-driven and human-induced stress. Forests at higher elevations, where rainfall is more consistent and temperatures are cooler, have maintained healthier cover. In contrast, lower-elevation forests experienced significant degradation, driven by both rising temperatures and intensified human activities, such as logging, agriculture, and urbanization. These human activities have caused soil degradation and the fragmentation of forest ecosystems, reducing biodiversity and weakening the overall resilience of forest cover [52].

AL remained a significant factor across both years, albeit with a slightly reduced influence in 2023. Higher altitudes, particularly those with steep slopes and rugged terrain, continued to support denser forest cover because of their relative isolation from human activity. However, groundwater levels (SWLs) became more important in 2023, indicating that water availability, especially in lower-elevation areas, is becoming critical for forest health. This change underscores the increasing dependence of forest ecosystems on groundwater, as rainfall has become more erratic due to climate change (highlighted in yellow in Figure 8) [9].

The highlighted areas in Figure 8 illustrate the regions with significant changes in forest cover between 2013 and 2023. In 2013, areas with high rainfall and suitable altitudes

had dense forest cover (highlighted in green in Figure 8). In contrast, in 2023, regions with stable soil moisture and adequate groundwater levels exhibited better forest health (highlighted in pink in Figure 8). These visual representations highlight the transition in the dominant environmental factors affecting forest cover over the past decade, pointing to an increased reliance on water-related parameters and a reduction in the influence of precipitation.

The significance of the explanatory variables aligns with the findings in Section 3.3, which discuss the importance of AL and R in 2013 and the increasing significance of SWL and the SMI in 2023. These geoenvironmental factors were identified as the major factors influencing forest cover changes, with soil moisture and groundwater becoming increasingly critical over time. This shift in significance suggests that the effective management of soil moisture and groundwater resources is pivotal in sustaining forest ecosystems in the region [53].

This shift from a reliance on rainfall and altitude in 2013 to a greater dependence on groundwater levels and soil moisture in 2023 reflects broader climatic changes and highlights the growing importance of hydrological factors in sustaining forest ecosystems. The impact of human activities such as land conversion and deforestation further exacerbates these climatic challenges, particularly in low-lying areas where forests are more accessible. Therefore, adaptive forest management strategies that account for both climatic variability and human pressure are essential to ensure the long-term resilience of forest ecosystems in the Duhok region.

3.5. Implications for Forest Management and Future Research

The results of this study have significant implications for forest management practices and highlight key areas for future research to ensure sustainable forest ecosystems in the Duhok District. Understanding the dynamic interplay of geoenvironmental factors, such as SWL, SMI, AL, and R, is crucial for developing effective management strategies that can adapt to evolving environmental conditions.

In 2013, AL and R were the primary factors influencing forest cover, with higher altitudes and adequate R values supporting better forest cover. By 2023, SWL and the SMI emerged as the dominant factors, underscoring a shift toward the critical role of water-related parameters. This shift reflects broader climatic changes, which have led to decreased precipitation and increased reliance on groundwater and soil moisture to sustain forest ecosystems. Specifically, the analysis revealed a 30% increase in forest cover areas with high soil moisture and a 25% decrease in regions with reduced groundwater levels [8,10]. These results highlight the growing importance of water management strategies in maintaining forest health in the face of climatic stress.

These findings offer valuable guidance for sustainable forest management strategies. Although the present study primarily discusses geoenvironmental factors, these are by no means unrelated to human activity. Groundwater availability, for instance, is directly affected by human extraction for agricultural, domestic, and industrial purposes. As such groundwater supplies decline, forests could experience increased stress, especially under conditions of low rainfall [53]. By elucidating the relationship between groundwater fluctuations, soil moisture, and *Pinus brutia* FVC, this research provides a framework for understanding how human-driven changes to critical water resources can alter forest structure and vitality.

In addition, the evidence we present in this study could be used by forest managers and policymakers through the implementation of sustainable groundwater extraction regulations. This could be performed by encouraging water-saving irrigation techniques and giving preference to reforestation in those areas that are more prone to water shortage [32].

In combination with improved climate monitoring and adaptive forestry techniques, these measures can help maintain healthy forest ecosystems under evolving environmental pressures. Ultimately, groundwater management integrated within the policy framework for the conservation of forests will address the contribution of both natural and anthropogenic factors in maintaining the long-term resilience and productivity of *Pinus brutia* forest ecosystems.

Forest managers should prioritize practices that maintain optimal soil moisture levels and promote sustainable groundwater use. Techniques, such as efficient irrigation systems, soil conditioning, and rainwater harvesting, can increase soil moisture and stabilize groundwater levels. These practices support tree growth and health and improve the resilience of forests to climatic variations [2]. Reforestation with native, drought-resistant species is another crucial strategy as it helps maintain the land's ability to retain water, mitigating the impact of erratic rainfall patterns. The use of agroforestry systems can also contribute to soil structure stability and moisture retention, helping forests adapt to changing environmental conditions.

The increased vulnerability of lower-elevation forests observed in 2023 suggests that land-use planning should prioritize the protection of these areas by limiting deforestation and promoting sustainable land-use practices. Zoning regulations that restrict agricultural expansion into forested areas could help reduce human pressure on vulnerable ecosystems and prevent the further fragmentation of forest cover. Forest fragmentation has been shown to exacerbate biodiversity loss and disrupt ecosystem services, further weakening forest resilience [8]. Spatial planning that integrates land-use data with forest management strategies is essential for achieving a balance between ecological preservation and socioeconomic needs.

Our observations show that forest cover in this study area is increasingly dependent on water-related factors such as soil moisture and levels of groundwater, rather than altitude and rainfall alone. This perhaps underlines the fact that human-mediated changes in the landscape—agricultural expansion, land-use conversion limiting infiltration, or increasing groundwater extraction—can indirectly affect forest vitality by altering key hydrological conditions. Strategic spatial planning in these newly sensitive lower-elevation forests could retain or enhance groundwater recharge zones and vegetation buffers supporting water retention. Although our study did not measure human activities like fragmentation, these geoenvironmental insights will underpin policy decisions on land use, restraint in agricultural encroachment, and the prevention of fragmentation in sensitive areas, thus reinforcing inherent resilience. The alignment of land-use strategies with ongoing modification to climatic and hydrological boundary conditions can therefore drive focused managerial interventions, a finding congruent with results obtained across other semiarid environments [20].

Although R declined in influence from 2013 to 2023, it remains a crucial factor for forest sustainability. Adaptive management strategies should include the creation of water retention landscapes and the implementation of rainwater harvesting systems to mitigate the effects of reduced rainfall. These measures will help ensure a continuous water supply for forests, improving their health and resilience to future climatic variations. Innovative water management solutions are particularly critical in areas experiencing declining precipitation [11].

Spatial planning is also essential for optimizing land use and balancing ecological, economic, and social needs. Integrating spatial planning with forest management can help identify areas that are most suitable for reforestation, conservation, and sustainable use. Tools such as GIS and remote sensing provide spatial data that inform decision-making processes and enhance environmental sustainability [15]. Additionally, the integration

of geoenvironmental data into forest management plans is essential. The use of remote sensing and ML tools, as demonstrated in this study, can provide forest managers with real-time data and predictive models. These technologies enable the precise monitoring and assessment of forest health and aid in informed decision-making [53].

Future research should address several key areas to build upon the findings of this study. Researchers should explore the impacts of additional geoenvironmental variables, such as land-use changes and human activities, to gain a more comprehensive understanding of forest dynamics. Socioeconomic drivers, such as population growth and agricultural demand, have been identified as major contributors to deforestation in other regions, and quantifying their impact in the Duhok district would be valuable for future management decisions [20].

Future studies could also explore the use of techniques such as Generalized Additive Models, which would increase interpretability and provide information on the causality of the relationships that exist between forest cover and geoenvironmental factors. This approach would complement predictive modeling by providing further insights into the drivers of dynamics in forest cover.

Furthermore, developing more advanced long-term monitoring systems that integrate climate data, remote sensing technologies, and field measurements is essential for tracking the evolving impact of climate change on forest ecosystems. Such systems would enable forest managers to predict forest cover changes with greater accuracy and respond more effectively to emerging environmental threats. Research should also focus on the long-term effects of climate change on forest ecosystems, identifying strategies to mitigate its impact and maintain forest health under future climate scenarios [51].

Ultimately, our study highlights the critical role of geoenvironmental factors in forest management and underscores the need for adaptive and sustainable practices. By prioritizing water management, adapting to climatic variations, integrating spatial planning, and leveraging advanced technologies, forest managers can develop effective strategies for preserving and enhancing forest ecosystems. Continuous research and technological advancements will ensure that these strategies remain responsive to evolving environmental conditions, contributing to the sustainability of forests in the Duhok District and beyond.

Although this study represents the modeling of forest cover as a dependent variable induced by different geoenvironmental factors such as rainfall, temperature, and levels of groundwater, we acknowledge that forest cover can also act as an explanatory variable driving changes in these factors. In particular, human-mediated land-use and land-cover changes, such as deforestation, have been reported to set off cascading effects of decreasing rainfall, changing temperature regimes and reducing groundwater levels [3,52]. These changes, in turn, exacerbate further forest loss, creating a reinforcing feedback loop [2,4]. Such dynamics underscore the bidirectional and interdependent nature of forest cover and its drivers.

In this context, our modeling approach considers forest cover as a dependent variable in pursuit of the primary objectives of this study on assessing the geoenvironmental contributions to forest dynamics within the Duhok District. This provides real, actionable insights into forest management strategies, especially in those regions where human-induced deforestation is predominant. However, from an ecological viewpoint, one must acknowledge that under certain conditions, forest cover might also act as a crucial driver in changing environmental variables. Future research could expand upon this perspective by employing dynamic systems modeling or causal inference frameworks to explore the interplay between forest cover and its drivers, particularly in the context of anthropogenic influences.

4. Conclusions

This study provides a detailed analysis of the spatiotemporal dynamics of forest cover in the Duhok District from 2013 to 2023, highlighting the impacts of various geoenvironmental factors. Using the RF model and Landsat data, we elucidated the complex interactions between geoenvironmental variables and forest vegetation cover.

Our findings indicate a dynamic shift in forest cover determinants over the past decade. In 2013, altitude and rainfall were the primary influences, suggesting that forests were highly dependent on climatic and topographical conditions. By 2023, however, groundwater levels and soil moisture had emerged as the most critical factors, reflecting broader climatic changes and variations in water availability. This shift emphasizes the increasing importance of water management strategies in maintaining forest health, particularly in semiarid regions. Forest cover shows significant variation in areas with relatively high soil moisture and reduced groundwater levels.

Concrete policy recommendations are necessary to address these challenges. Forest managers should prioritize practices that maintain optimal soil moisture levels and promote sustainable groundwater use. Techniques such as efficient irrigation systems, soil conditioning, and rainwater harvesting can increase soil moisture, stabilize groundwater levels, support tree growth, and improve the resistance of forests to climatic variations. Additionally, land-use planning must focus on protecting vulnerable lower-altitude forests by restricting agricultural expansion and deforestation. Reforestation with native, drought-resistant species will further contribute to the resilience of forests in response to future climate challenges.

This study underscores the necessity for adaptive management strategies, including altitude-specific management approaches, because of the consistent influence of altitude on forest cover. Integrating spatial planning with forest management and utilizing tools such as GIS and remote sensing can optimize land use and improve decision-making processes. These strategies help mitigate the effects of forest fragmentation, which exacerbates biodiversity loss and disrupts ecosystem services.

Future research should address additional environmental variables, such as land-use changes and human activities, to provide a more comprehensive understanding of forest dynamics. Moreover, future studies should explore the use of higher-resolution satellite imagery to increase the precision of forest cover predictions. Comparing the performance of other machine learning models, such as Support Vector Machines or XGBoost, could offer further improvements in predictive accuracy for forest cover dynamics. Long-term monitoring systems that integrate climate data, remote sensing technologies, and field measurements are essential for assessing the evolving impact of climate change on forest ecosystems.

Overall, this study highlights the critical role of geoenvironmental factors in shaping forest cover dynamics. Adaptive and sustainable management practices, informed by advanced technologies and continuous research, are essential for preserving and enhancing forest ecosystems and ensuring their resilience in the face of changing environmental conditions.

Author Contributions: Conceptualization, H.N.H. and Y.T.M.; methodology, H.N.H. and Y.T.M.; software, H.N.H.; validation, H.N.H. and Y.T.M.; formal analysis, H.N.H.; investigation, H.N.H. and Y.T.M.; resources, H.N.H.; data curation, H.N.H.; writing—original draft preparation, H.N.H. and Y.T.M.; writing—review and editing, Y.T.M.; visualization, H.N.H. and Y.T.M.; supervision, Y.T.M.; project administration, Y.T.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data were not made publicly available owing to privacy restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhao, T.; Mu, X.; Song, W.; Liu, Y.; Xie, Y.; Zhong, B.; Xie, D.; Jiang, L.; Yan, G. Mapping Spatially Seamless Fractional Vegetation Cover over China at a 30-m Resolution and Semimonthly Intervals in 2010–2020 Based on Google Earth Engine. *J. Remote Sens.* **2023**, *3*, 0101. [CrossRef]
2. Cheng, G.; Li, X.; Zhao, W.; Xu, Z.; Feng, Q.; Xiao, S.; Xiao, H. Integrated study of the water–ecosystem–economy in the Heihe River Basin. *Natl. Sci. Rev.* **2014**, *1*, 413–428. [CrossRef]
3. Kumar, R.; Kumar, A.; Saikia, P. Deforestation and Forests Degradation Impacts on the Environment. In *Environmental Degradation: Challenges and Strategies for Mitigation*; Singh, V.P., Yadav, S., Yadav, K.K., Yadava, R.N., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 19–46.
4. Debebe, B.; Senbeta, F.; Teferi, E.; Diriba, D.; Teketay, D. Analysis of forest cover change and its drivers in biodiversity hotspot areas of the semien mountains national Park, northwest Ethiopia. *Sustainability* **2023**, *15*, 3001. [CrossRef]
5. FAO. *Food and Agriculture Organization of the United Nations*; Forestry Department: Rome, Italy, 2020.
6. Mzuri, R.T.; Omar, A.A.; Mustafa, Y.T. Land degradation assessment using AHP and GIS-based modelling in Duhok District, Kurdistan Region, Iraq. *Geocarto Int.* **2021**, *37*, 7908–7926. [CrossRef]
7. Chu, D. Fractional Vegetation Cover. In *Remote Sensing of Land Use and Land Cover in Mountain Region: A Comprehensive Study at the Central Tibetan Plateau*; Chu, D., Ed.; Springer Singapore: Singapore, 2020; pp. 195–207.
8. Habeeb, H.N.; Mustafa, Y.T. Deep learning approaches for estimating forest vegetation cover and exploring influential ecosystem factors. *Earth Sci. Inform.* **2024**, *17*, 3379–3396. [CrossRef]
9. Mzuri, R.T.; Omar, A.A.; Mustafa, Y.T. Spatiotemporal Analysis of Land Surface Temperature and Vegetation Changes in Duhok District, Kurdistan Region, Iraq. *Iraqi Geol. J.* **2022**, *55*, 68–81. [CrossRef]
10. Koirala, S.; Jung, M.; Reichstein, M.; de Graaf, I.E.; Camps-Valls, G.; Ichii, K.; Papale, D.; Ráduly, B.; Schwalm, C.R.; Tramontana, G. Global distribution of groundwater-vegetation spatial covariation. *Geophys. Res. Lett.* **2017**, *44*, 4134–4142. [CrossRef]
11. Al-Quraishi, A.M.F.; Gaznayee, H.A.; Crespi, M. Drought trend analysis in a semi-arid area of Iraq based on Normalized Difference Vegetation Index, Normalized Difference Water Index and Standardized Precipitation Index. *J. Arid Land* **2021**, *13*, 413–430. [CrossRef]
12. Mustafa, Y.; Salar, S.; Khatab, A.; Ahmed, F.; Habeeb, H. *Mapping and Estimating Vegetation Coverage in Iraqi Kurdistan Region Using Remote Sensing and GIS*; KRG: Erbil, Iraqi Kurdistan, 2015; pp. 1–136.
13. Chen, S.; Ma, M.; Wu, S.; Tang, Q.; Wen, Z. Topography intensifies variations in the effect of human activities on forest NPP across altitude and slope gradients. *Environ. Dev.* **2023**, *45*, 100826. [CrossRef]
14. Li, W.; Migliavacca, M.; Forkel, M.; Denissen, J.M.C.; Reichstein, M.; Yang, H.; Duveiller, G.; Weber, U.; Orth, R. Widespread increasing vegetation sensitivity to soil moisture. *Nat. Commun.* **2022**, *13*, 3959. [CrossRef]
15. Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K. Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 345–354.
16. Song, D.-X.; Wang, Z.; He, T.; Wang, H.; Liang, S. Estimation and validation of 30 m fractional vegetation cover over China through integrated use of Landsat 8 and Gaofen 2 data. *Sci. Remote Sens.* **2022**, *6*, 100058. [CrossRef]
17. Gigović, L.; Pourghasemi, H.R.; Drobnjak, S.; Bai, S. Testing a New Ensemble Model Based on SVM and Random Forest in Forest Fire Susceptibility Assessment and Its Mapping in Serbia's Tara National Park. *Forests* **2019**, *10*, 408. [CrossRef]
18. Khurshed, I.S.; Mustafa, Y.T.; Fayyadh, M.A. Assessing Spatial Patterns of Surface Soil Moisture and Vegetation Cover in Batifa, Kurdistan Region-Iraq: Machine Learning Approach. *IEEE Access* **2023**, *11*, 130406–130417. [CrossRef]
19. Ugbaja, A.N.; Ugbaja, U.A.; Bassey, A.E.; Amah, E.A.; Offiong, O.E. Assessment of Groundwater Vulnerability using Vulnerability Index (Static Water Level, Cation Exchange Capacity, Organic Contents, Nitrate and Chloride) (SCONC) in Part of Mamfe Embayment South Eastern Nigeri. *Res. J. Environ. Earth Sci.* **2020**, *38*, 39–47. [CrossRef]
20. Al-Abadi, A.M.; Fryar, A.E.; Rasheed, A.A.; Pradhan, B. Assessment of groundwater potential in terms of the availability and quality of the resource: A case study from Iraq. *Environ. Earth Sci.* **2021**, *80*, 426. [CrossRef]
21. Alganci, U.; Besol, B.; Sertel, E. Accuracy assessment of different digital surface models. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 114. [CrossRef]
22. Koulelis, P.P.; Proutsos, N.; Solomou, A.D.; Avramidou, E.V.; Malliarou, E.; Athanasiou, M.; Xanthopoulos, G.; Petrakis, P.V. Effects of Climate Change on Greek Forests: A Review. *Atmosphere* **2023**, *14*, 1155. [CrossRef]

23. Khwarahm, N.R.; Najmaddin, P.M.; Ararat, K.; Qader, S. Past and future prediction of land cover land use change based on earth observation data by the CA–Markov model: A case study from Duhok governorate, Iraq. *Arab. J. Geosci.* **2021**, *14*, 1544. [CrossRef]
24. Torres, I.; Moreno, J.M.; Morales-Molino, C.; Arianoutsou, M. Ecosystem Services Provided by Pine Forests. In *Pines and Their Mixed Forest Ecosystems in the Mediterranean Basin*; Ne'eman, G., Osem, Y., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 617–629.
25. Buringh, P. *Soils and Soil Conditions in Iraq*; Ministry of Agriculture, Agricultural Research and Projects: Baghdad, Iraq, 1960.
26. Obeyed, M.; Akrawee, Z.; Mustafa, Y. Estimating Aboveground Biomass and Carbon Sequestration for Natural Stands of Quercus Aegilops. in Duhok Province. *Iraqi J. Agric. Sci.* **2020**, *51*. [CrossRef]
27. FAO. *Strengthening Monitoring, Assessment and Reporting (MAR) on Sustainable Forest Management (SFM) in Asia (GCP/INT/988/JPN)*; FAO: Rome, Italy, 2007.
28. Directorate-of-Forests-and-Range-Duhok. *Management of Planted Forests: A Report*; Directorate-of-Forests-and-Range-Duhok: Duhok, Iraq, 2023; pp. 1–54.
29. Xie, L.; Meng, X.; Zhao, X.; Fu, L.; Sharma, R.P.; Sun, H. Estimating Fractional Vegetation Cover Changes in Desert Regions Using RGB Data. *Remote Sens.* **2022**, *14*, 3833. [CrossRef]
30. Rouse, J.W.; Haas, R.H.; Deering, D.; Schell, J.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Texas A&M University: College Station, TX, USA, 1974.
31. Ekwere, A.; Kudamnya, E.; Adamu, C.; Edet, A. Evaluation of Groundwater Vulnerability Using Soil and Hydrogeological Data in Parts of the Precambrian Oban Massif, Se Nigeria. PREPRINT (Version 1) available at Research Square. 2022. [CrossRef]
32. Thakur, D.; Bartarya, S.K.; Nainwal, H.C.; Dutt, S. Impact of environment and LULC changes on groundwater resources in the Soan Basin, western Himalaya. *Environ. Monit. Assess* **2022**, *194*, 612. [CrossRef]
33. Manda, A.K.; Etheridge, J.R.; Grace-McCaskey, C.; Allen, T.R.; Howard, R. Are Groundwater Level Data Collected by Citizen Scientists Trustworthy? A Cautionary Tale. *Groundwater* **2021**, *59*, 123–130. [CrossRef] [PubMed]
34. Noori, M.J.; Hassan, H.H.; Mustafa, Y.T. Spatial estimation of rainfall distribution and its classification in duhok governorate using GIS. *J. Water Resour. Prot.* **2014**, *06*, 75–82. [CrossRef]
35. Elumalai, V.; Brindha, K.; Sithole, B.; Lakshmanan, E. Spatial interpolation methods and geostatistics for mapping groundwater contamination in a coastal area. *Environ. Sci. Pollut. Res.* **2017**, *24*, 11601–11617. [CrossRef] [PubMed]
36. John, J.; Jaganathan, R.; Dharshan Shylesh, D.S. Mapping of Soil Moisture Index Using Optical and Thermal Remote Sensing. Proceedings of SECON'21; Springer International Publishing: Cham, Switzerland, 2022; pp. 759–767.
37. Wang, L.; Lu, J.; Zhou, R.; Duan, G.; Wen, Z. Analysis of Soil Moisture Change Characteristics and Influencing Factors of Grassland on the Tibetan Plateau. *Remote Sens.* **2023**, *15*, 298. [CrossRef]
38. Saha, A.; Patil, M.; Goyal, V.C.; Rathore, D.S. Assessment and Impact of Soil Moisture Index in Agricultural Drought Estimation Using Remote Sensing and GIS Techniques. *Proceedings* **2019**, *7*, 2. [CrossRef]
39. Mohamed, E.S.; Ali, A.; El-Shirbeny, M.; Abutaleb, K.; Shaddad, S.M. Mapping soil moisture and their correlation with crop pattern using remotely sensed data in arid region. *Egypt. J. Remote Sens. Space Sci.* **2020**, *23*, 347–353. [CrossRef]
40. O'brien, R.M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Qual. Quant.* **2007**, *41*, 673–690. [CrossRef]
41. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
42. Boehmke, B.; Greenwell, B.M. *Hands-on Machine Learning with R*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2019.
43. Efron, B.; Tibshirani, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.* **1986**, *1*, 54–75. [CrossRef]
44. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
45. Edeh, E.; Lo, W.-J.; Khojasteh, J. Review of Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook. *Struct. Equ. Model. A Multidiscip. J.* **2023**, *30*, 165–167. [CrossRef]
46. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [CrossRef]
47. Campoe, O.C.; Munhoz, J.S.B.; Alvares, C.A.; Carneiro, R.L.; de Mattos, E.M.; Ferez, A.P.C.; Stape, J.L. Meteorological seasonality affecting individual tree growth in forest plantations in Brazil. *For. Ecol. Manag.* **2016**, *380*, 149–160. [CrossRef]
48. Schippers, P.; Sterck, F.; Vlam, M.; Zuidema, P.A. Tree growth variation in the tropical forest: Understanding effects of temperature, rainfall and CO2. *Glob. Change Biol.* **2015**, *21*, 2749–2761. [CrossRef]
49. Luo, M.; Wang, Y.; Xie, Y.; Zhou, L.; Qiao, J.; Qiu, S.; Sun, Y. Combination of Feature Selection and CatBoost for Prediction: The First Application to the Estimation of Aboveground Biomass. *Forests* **2021**, *12*, 216. [CrossRef]
50. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2017.
51. Hending, D.; Holderied, M.; McCabe, G.; Cotton, S. Effects of future climate change on the forests of Madagascar. *Ecosphere* **2022**, *13*, e4017. [CrossRef]

52. Zhou, S.; Long, H.; Xing, H.; Zhang, K.; Wang, R.; Zhang, E. Human activities facilitated the decline of forest ecosystem in East Asia after 5000 a BP. *Earth-Sci. Rev.* **2023**, *245*, 104552. [CrossRef]
53. Han, Z.; Huang, S.; Huang, Q.; Bai, Q.; Leng, G.; Wang, H.; Zhao, J.; Wei, X.; Zheng, X. Effects of vegetation restoration on groundwater drought in the Loess Plateau, China. *J. Hydrol.* **2020**, *591*, 125566. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.