

## Article

# Actual Evapotranspiration Estimation Using Sentinel-1 SAR and Sentinel-3 SLSTR Data Combined with a Gradient Boosting Machine Model in Busia County, Western Kenya

Peter K. Musyimi <sup>1,2,\*</sup>, Ghada Sahbeni <sup>1,\*</sup>, Gábor Timár <sup>1</sup>, Tamás Weidinger <sup>3</sup> and Balázs Székely <sup>1</sup>

<sup>1</sup> Department of Geophysics and Space Science, Institute of Geography and Earth Sciences, ELTE Eötvös Loránd University, Pázmány Péter prom 1/C, H-1117 Budapest, Hungary

<sup>2</sup> Department of Humanities and Languages, Karatina University, Karatina P.O. Box 1957-10101, Kenya

<sup>3</sup> Department of Meteorology, Institute of Geography and Earth Sciences, ELTE Eötvös Loránd University, Pázmány Péter prom 1/A, H-1117 Budapest, Hungary

\* Correspondence: musyimpeter@student.elte.hu (P.K.M.); gsahbeni@caesar.elte.hu (G.S.)

† These authors contributed equally to this work.

**Abstract:** Kenya is dominated by a rainfed agricultural economy. Recurrent droughts influence food security. Remotely sensed data can provide high-resolution results when coupled with a suitable machine learning algorithm. Sentinel-1 SAR and Sentinel-3 SLSTR sensors can provide the fundamental characteristics for actual evapotranspiration (AET) estimation. This study aimed to estimate the actual monthly evapotranspiration in Busia County in Western Kenya using Sentinel-1 SAR and Sentinel-3 SLSTR data with the application of the gradient boosting machine (GBM) model. The descriptive analysis provided by the model showed that the estimated mean, minimum, and maximum AET values were 116, 70, and 151 mm/month, respectively. The model performance was assessed using the correlation coefficient ( $r$ ) and root mean square error (RMSE). The results revealed a correlation coefficient of 0.81 and an RMSE of 10.7 mm for the training dataset (80%), and a correlation coefficient of 0.47 and an RMSE of 14.1 mm for the testing data (20%). The results are of great importance scientifically, as they are a conduit for exploring alternative methodologies in areas with scarce meteorological data. The study proves the efficiency of high-resolution data retrieved from Sentinel sensors coupled with machine learning algorithms, focusing on GBM as an alternative to accurately estimate AET. However, the optimal solution would be to obtain direct evapotranspiration measurements.

**Keywords:** actual evapotranspiration; Busia County; gradient boosting machine; Kenya; estimation modeling; Sentinel-1 SAR; Sentinel-3 SLSTR



**Citation:** Musyimi, P.K.; Sahbeni, G.; Timár, G.; Weidinger, T.; Székely, B. Actual Evapotranspiration Estimation Using Sentinel-1 SAR and Sentinel-3 SLSTR Data Combined with a Gradient Boosting Machine Model in Busia County, Western Kenya. *Atmosphere* **2022**, *13*, 1927. <https://doi.org/10.3390/atmos13111927>

Academic Editors: Demetrios

E. Tsesmelis, Nikolaos Skondras and Ippokratis Gkotsis

Received: 3 October 2022

Accepted: 15 November 2022

Published: 18 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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

Water fluxes are fundamental for many theoretical, practical, and applied disciplines of climatology, hydrometeorology, and agriculture. Thus, the estimation and quantification of actual evapotranspiration (AET) are equally indispensable in realizing the global, regional, and country-specific 2030 17 sustainable development goals of the common agenda of the UN. The sustainable development goals (SDG) include SDG1—no poverty, SDG6—clean water and sanitation, SDG13—climate action, and SDG2—zero hunger [1]. Moreover, the Food and Agriculture Organization (FAO) SDG indicator 6.4.1—the change in water use [2] efficiency, as well as 6.4.2—the level of water stress [3] over time, is essential for several sectors related to the global economy, such as agriculture, industry, mining, and power production. In this context, agriculture services represent the most significant share of water consumption [4], with a higher percentage in the semiarid and Mediterranean climatic regions [5]. In this regard, AET estimation is crucial for adequately implementing sustainable agricultural systems and achieving food security in developing countries.

The importance of actual evapotranspiration cannot be undervalued, particularly in light of the current global challenges such as food insecurity and climate change. It is a major hydrological cycle component [6], along with precipitation, depicting the primary constituents of the surface energy budget [7]. Moreover, it is a crucial component of regional and global environmental phenomena associated with meteorological, agricultural, and hydrological applications [8–10]. Not only do its changes influence precipitation, streamflow, and surface temperature, among other hydro-climatological variables [11,12], but it also plays a vital role in the climate system, coupled with water, carbon, and energy cycles [13]. Despite the creation of global initiatives aiming to directly measure AET, such as the FLUXNET project [14], short and inconsistent field surveys exist. In view of this, previous studies on machine learning applications have progressively been used in various environment-related fields. A study by Malik et al. [15] demonstrated the potential of the gradient boosting machine (GBM) model for pan-evaporation process prediction in Iran and India. In addition, Shrivastav and Jha [16] successfully used the GBM to explore the effects of temperature and humidity on COVID transmission. Furthermore, Frey [17] demonstrated that the GBM has a significant predictive performance in natural resource management aiming to enhance ecological sustainability. Hailstorm prediction and forecasting and severe weather forecasting have been performed efficiently, proving the suitability of the GBM and other machine learning models [18,19]. An accurate estimation of AET remains a challenging scientific problem [20–23] due to the associated prediction uncertainty when quantifying the actual evapotranspiration [14,24]. Although the gradient boosting machine has been used extensively in this field, limited research has been conducted to investigate its efficiency for AET estimation, which is the aim of the present study.

It is also significant that studies have extensively explored these differences, revealing significant uncertainties in AET estimations using various modeling techniques and approaches [25–28]. The authors of [29] provided insight regarding the vast disparities in evapotranspiration assessment using multiple theoretical methods compared to other global variable uncertainties, many of which are retrieved using satellite remote sensing systems [30–32]. Nevertheless, [33] found that AET products show the lowest uncertainties in the case of the LSM (land surface model) and moderate uncertainties when using a moderate resolution imaging spectroradiometer (MODIS), with the highest uncertainties observed when using the water budget approach. However, in recent years, remote sensing has made significant progress in estimating and assessing AET variation over time and space. Studies have employed remote-sensing-based models in AET-related studies. Using the surface energy balance index (SEBI), two-source model (TSM), surface energy balance algorithm for land (SEBAL), surface energy balance system (SEBS), Eta mapping algorithm (ETMA), and atmosphere–land exchange inverse model (ALEXI), many researchers have successfully assessed and predicted AET spatiotemporal variation in different regions across the globe [34–40]. In the same context, [41] suggested a more sophisticated analysis aiming to reduce the range of uncertainty in observation-based AET estimations based on a combination of the remote sensing and machine learning tools discussed in the present study.

In practice, combining remote sensing data and machine learning models can efficiently improve water flux balance modeling and management strategies to create a more sustainable future. This can be explained by the spatiotemporal variability in water fluxes, which is highly influenced by the heterogeneity of the land surface, topography, lithology climate, meteorological conditions, soil moisture content characteristics, and vegetation vigor and density [42,43]. As evapotranspiration represents a vital component of the hydrosphere and atmosphere, this induces complex land–atmosphere feedback processes and drivers. Recent developments achieved by the European Space Agency (ESA), such as the Sentinel-1 SAR (synthetic aperture radar) and Sentinel-3 sea and land surface temperature radiometer (SLSTR), are revolutionary in terms of their provision of free data access to the public. It is worth noting that SLSTR involves the acquisition of TIR (thermal infrared) data [44] and its comparison with the NDVI (normalized difference vegetation index) tem-

poral variation [45], which is an objective of this present research study. Remotely sensed data retrieved from Sentinel sensors are the prerequisites for the potential application of spectral and spatial-temporal characteristics [46] in agriculture [45,47].

Studies on AET estimation and quantification have gained momentum in recent years. For instance, [48] found that Sentinel-2 and Sentinel-3 data offer the most essential and suitable spectral information required for AET estimation, despite the remarkable differences in the spatial resolution from 10 m to 1 km. Furthermore, various projects, such as the Sen-ET project (<https://www.esa-sen4et.org/>, accessed on 5 July 2022), have demonstrated that the high-spatial-resolution remotely sensed data (10–60 m) retrieved from Sentinel-2 and medium-spatial-resolution (1 km) thermal data captured by Sentinel-3 produce reliable AET estimates with great accuracy. Similarly, [10] successfully estimated AET using multispectral data, proving the potential of remote sensing for drought monitoring studies. Although remotely sensed data have been widely and extensively used for earth observation, limited research has been conducted in Kenya through the combination of machine learning and remotely sensed data and the investigation of its efficiency in AET estimation and prediction. In this context, to ensure the continuous monitoring of water resource consumption and balance, a combination of variables, especially VH, VV, VV–VH, and VH/VV, are of interest since they have rarely been used in previous studies [45]. These variables, combined with Sentinel-3 data such as the NDVI and land surface temperature (LST), therefore, can provide valuable spectral information associated with AET, unlike other processes, such as conventional pan-evaporation and lysimetric and eddy covariance, which require enormous datasets acquired from field campaigns [49]. The scarcity of meteorological data has led to an urgent need to explore alternative approaches for estimating AET in regions. Therefore, this study is of critical scientific benefit because it provides the basis for utilizing up-to-date alternative approaches and methodologies and integrating them with a machine learning model GBM for AET prediction in local and regional areas with insufficient meteorological data. Furthermore, the study demonstrates the potential of remote sensing to estimate actual evapotranspiration. The proposed approach provides a foundation for future local and regional research applications. This is because evapotranspiration estimates are fundamental parameters for water balance modeling, an essential aspect of Kenya's dominated rainfed agriculture, which is its economic mainstay [50,51].

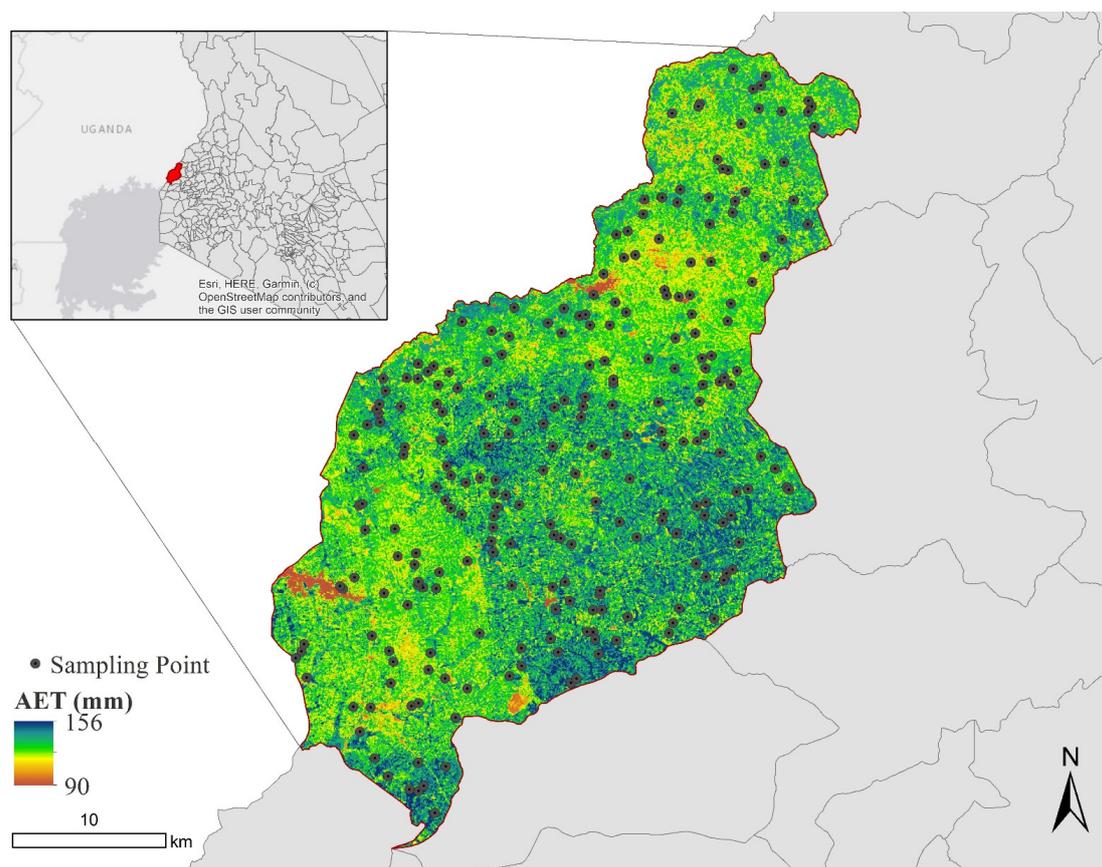
In Kenya, remote sensing for the estimation of AET is critical for drought monitoring [52] since the high evapotranspiration potential leads to hydric stress and, consequently, lower crop yields [50]. Our study establishes a basis for research in other regions in Kenya, because around 89% of its total landmass (29 out of 47 counties) is influenced by arid and semiarid climates [53]. These areas are prone to water scarcity, a global challenge recognized by the UN that has led to calls for action to manage water [54,55], as well as food insecurity due to low agricultural production, the low adaptive capacity of households, and high vulnerability to climate extremes, which have strong negative socio-economic impacts [56]. For instance, Sorre [57] indicated that the increased frequency and amount of temperature and precipitation anomalies have led to recurring droughts in Busia County. Deficiencies in precipitation and fluctuations in evapotranspiration influence water availability [58]. In this regard, our study aims to investigate the efficiency of the machine learning model, i.e., the GBM, combined with remotely sensed data retrieved from Sentinel-1SAR and Sentinel-3 SLSTR sensors in AET estimation and to determine the main variables influencing its spatial distribution over Busia County in Western Kenya.

## 2. Materials and Methods

### 2.1. Study Area

Kenya lies between the latitudes of 4.5° N and 4.5° S and longitudes 34° E and 42° E in Eastern Africa, covering 582,646 km<sup>2</sup> of the land surface, with a population of 47,564,296 million [59]. As Kenya is administratively divided into forty-seven counties, it has a diverse climate and is a prosperous country with geographical features such as the famous Great Rift Valley and iconic Mount Kenya, with a height of 5199 m above sea

level, and Lake Victoria. Busia County, presented in Figure 1, is in the west and divided into seven administrative sub-counties: Funyula, Budalangi, Butula, Matayos, Nambale, Teso North, and Teso South, lying on latitudes of  $0^{\circ}27'$  to  $38.7684''$  north and longitudes of  $34^{\circ}6'$  to  $41.2632''$  east. It borders Bungoma to the north, Kakamega to the east, and Siaya to the southwest. The study area has a tropical climate with an average temperature of  $22^{\circ}\text{C}$  and an average rainfall of 1691 mm annually [60]. It has an annual mean maximum temperature range of  $26^{\circ}\text{C}$  to  $30^{\circ}\text{C}$  [61] and a mean minimum temperature range of  $14^{\circ}\text{C}$  to  $22^{\circ}\text{C}$  [57,62]. Busia County experiences a bimodal rainfall distribution with an extended rainy season in April–May and a short rainy season in October [63]. It is also prone to flooding, specifically in the Budalangi Constituency, Teso North Sub-County, situated in the low-lying swampy zone [64,65]. The altitude varies from 1130 m on the shores of Lake Victoria to approximately 1500 m in Funyula and the North Teso Hills. Overall, the study area has a complex terrain along the Samia Hills, with the Kavirondo Rocks, granitic hills in Amukura, and Chelelemuk representing a conspicuous topographic stretch. Busia is characterized by sandy loam soils with dark clay domination in the northern and central parts, making it agriculturally prosperous [66], with diverse food and cash crops, including tobacco, cotton, maize, robusta coffee, sugarcane cultivation, and various horticultural crops [62].



**Figure 1.** Busia County in western Kenya: The geographic location of the sampling points. The AET raster map was retrieved from the WaPOR official website.

## 2.2. Used Data and Processing

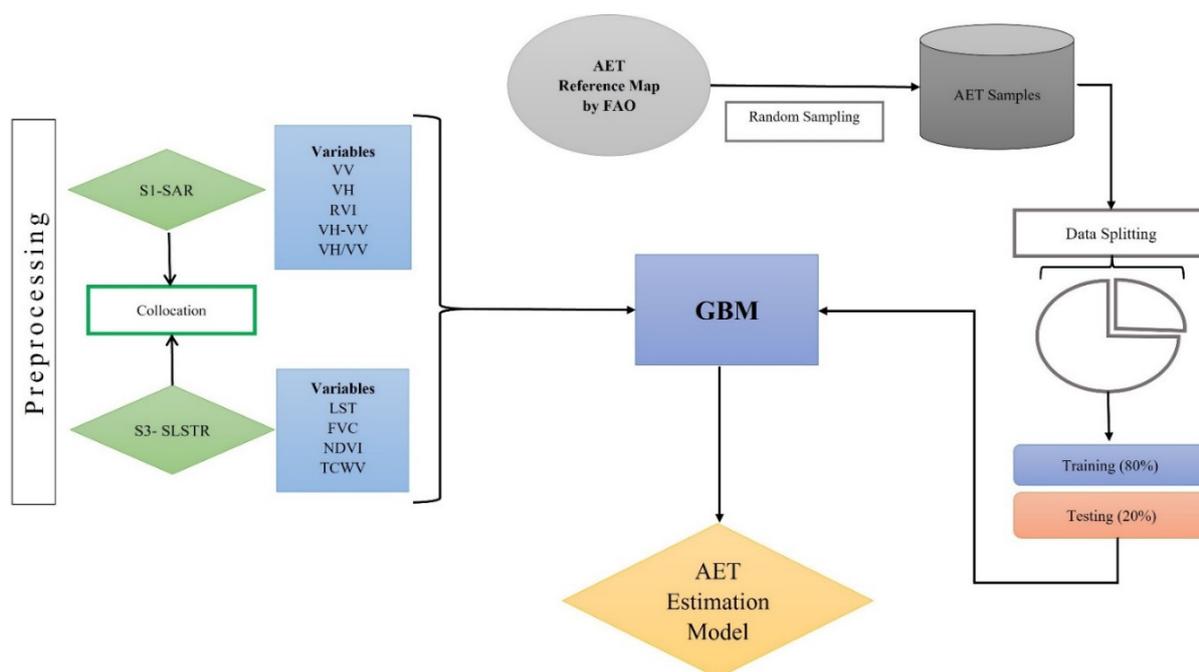
Sentinel-1 SAR ground-range-detected (GRD) data, acquired on 24 September 2021, and Sentinel-3 SLSTR Level 2 data, acquired on 29 September 2021, were downloaded from the Copernicus Open Access Hub Portal (<https://scihub.copernicus.eu/dhus/#/home>, accessed on 8 September 2022). Detailed information about Sentinel products is available online in the user guides [67]. A 30 m raster map of the actual evapotranspiration

(mm) in Busia County in September 2021 was retrieved from the WaPOR (the FAO portal to monitor Water Productivity through Open access of Remotely sensed derived data) (<https://wapor.apps.fao.org/>, accessed on 8 September 2022). More details about the reference data can be found in the metadata file available on the WaPOR 2.1 official website. SLSTR, referring to the sea and land surface temperature radiometer, is a dual-scan temperature radiometer selected for the ESA Sentinel-3 mission in low Earth orbit as a part of the Copernicus Programme [67]. It provides a full range of applications related to earth observation, the most prominent of which are the sea surface temperature (SST) assessment and land monitoring [68,69]. SLSTR products offer highly accurate global and regional sea and land surface temperatures (SST and LST) for climatological and meteorological applications. The Sentinel-3 mission provides images of a high frequency and resolution [70]. It has sufficient complexity to interpret data due to its dependence on many factors, such as moisture content, surface heterogeneity, and vegetation cover monitoring, among others. SAR comprises high-resolution returns of radar frequency energy from terrain illuminated by a sensor-generated directed beam of pulses. It monitors both geophysical and biophysical components [71]. The physical characteristics of the surface features include surface roughness, geometric structure, and digital elevation models [72].

Sentinel-1 ground-range-detected (GRD) data were co-registered, radiometrically calibrated, and then geometrically corrected using range Doppler terrain correction and filtered through the speckle effect using a three-by-three Lee filter [73]. Once the data were converted to a decibel scale, the ratio (VH/VV), the difference (VH-VV), and the radar vegetation index (RVI) were derived using the band math tool on the Sentinel Application Platform (SNAP). Once the Sentinel-3 SLSTR Level 2 data were co-registered, they were geometrically corrected. Then, the LST, TCWV, NDVI, and FVC features were extracted. After the LST values were converted from Kelvin to degrees Celsius ( $^{\circ}\text{C}$ ), the products were resampled to 30 m, stacked, and clipped using ArcMap 10.3. The derived covariates are presented in Table 1. For the reference map, random sampling was conducted using the “create random points” function in ArcMap 10.3 manufactured by ESRI in San Diego, USA. As a result, 250 sampling points were created (Figure 1). Then, we extracted the corresponding multiple values from the Sentinel-1 and Sentinel-3 derivative variables. Once the database had been created in ArcMap 10.3, it was imported to RStudio to train, calibrate, and test the gradient boosting model (GBM) [74]. Only 80% of the data were used for the training, while the remaining 20% were used for the testing. Figure 2 describes the flow chart of the study, the variables from the satellite image, and the gradient boosting machine estimation model.

**Table 1.** Variables used in the study and their descriptions.

Variable	Description
VH	Sigma naught ( $\sigma^{\circ}$ ) backscatter intensity in decibels (dB)
VV	Sigma naught ( $\sigma^{\circ}$ ) backscatter intensity in decibels (dB)
Diff	The difference between VH and VV (dB)
ratio	The ratio between VH and VV (dB)
RVI	Radar vegetation index in (dB), $\frac{4 \times VH}{VH + VV}$ [75]
LST	Land surface temperature in degrees Celsius ( $^{\circ}\text{C}$ )
FVC	Fractional vegetation cover
TCWV	Total column of water vapor ( $\text{kg}/\text{m}^2$ )
NDVI	Normalized difference vegetation index



**Figure 2.** Flow chart of the study.

### 2.3. Gradient Boosting Machine (GBM)

The 'gbm' R package was used to train and calibrate the model (RDocumentation). The gbm R package implements extensions to [76] the AdaBoost algorithm and Friedman's gradient boosting machine [77]. The AdaBoost algorithm trains a decision tree, whereby each observation is assigned an equal variable weight. After that, the weight of the difficult-to-classify observations aims to improve the prediction, so that the final ensemble represents the weighted sum of the previous tree models. GBM is a predictive modeling algorithm that leads to the decision making of tree-like structures to reduce residual errors from the previous iteration [78]. Hence, it is highly competitive with random forest algorithms. In addition, boosting improves the trees' accuracy [79]. This machine learning algorithm was used because of its robust characteristics that produce better predictions than the simpler ones [80,81]. Several studies have used this model for applications such as sentiment classification [82], where GBM performed better than random forest. GBM has been used for predictive functions [83] in related clinical research and produces better results in cases of complex relationships. Khoi et al. [84] also demonstrated its good performance in predicting the water quality index. The model has further shown an exclusive potential to predict the impact of air quality on urban areas and an impressive performance in predicting pollution caused by human activities [85].

Furthermore, GBM is an ensemble-based model that can be used for regression and classification purposes for decision making [86] and builds on weak successive trees to improve the previous tree. It is reliable when fitting new models to produce accurate estimates [87]. GBM portrays superior results when combined with other techniques to minimize the prediction error. Its basic concept is presented as follows:

Inputs:

The input data are  $(x, y)N_{i=1}$ , where  $N_{i=1}$  is the sampling dataset.  $x = (x_1, \dots)$  refers to the input variables, and  $y$  refers to the response variable.

Number of iterations  $M$ .

The choice of loss of functions is  $\Psi(y, f)$ , where  $\Psi$  is the loss function,  $y$  is the response variable, and  $f$  is the function expressed as follows:

$$\hat{f}(x) = y, \hat{f}(x) = \operatorname{argmin}[f(x)\Psi(y, f(x))]. \quad (1)$$

where  $\hat{f}(x)$  is the estimate or approximation function (predictive learning) of  $x$ ,  $f(x)$  is the functional dependence and  $\hat{f}(x)$  is the function estimate (predictive learning), and  $\Psi(y, f)$  is the loss function.

The choice of base learner model is  $h(x, \theta)$ , a custom base learner function, which implies a node regression tree induced in a best-first manner.

Algorithm:

Initialize  $\hat{f}(0)$ , where  $\hat{f}$  is a function and  $(0)$  is a constant, the initial constant value prediction.

For  $t = 1$  to  $M$ , do:

Compute the negative gradient  $g_t(x)$ , where  $g_t(x)$  is the negative gradient of the loss function associated with the whole ensemble.

Fit a new base-learner function  $h(x, \theta_t)$ , which is a simple parameterized function of the input variables  $x$ , and  $h$  is a regression tree.

Find the best gradient descent step size  $\rho t$ :

$$\rho t = \operatorname{argmin}_{\rho} \sum_{i=1}^N \Psi[y_i, \hat{f}^{t-1}(x_i) + \rho h(x_i, \theta_t)] \quad (2)$$

where  $\rho t$  is the gradient descent.

Update the function estimate. For the function estimate at the  $t^{\text{th}}$  iteration, the optimization rule is, therefore, defined as:

$$\hat{f}^t \leftarrow \hat{f}^{t-1} + \rho t h(x, \theta_t) \quad (3)$$

An updated model ensures framework overfitting, which is restrained by the end number of gradient boosting repetitions number.

Source: Friedman's gradient boost algorithm [76,77].

#### 2.4. Limitations of the Applied Datasets and Methodology

Data derived from different sensors with different properties, e.g., radiometric, spatial, and spectral resolutions, might lead to uncertainties during modeling, mainly because some preprocessing steps might lead to the loss of spectral characterization and information in the pixels. In addition, the machine learning model applied is predictive; therefore, it cannot generalize and outline the exact relationship between AET and the spectral information retrieved from Sentinel-1 SAR and Sentinel-3 SLSTR Level 2. These issues can only be resolved when a more extensive database size is accessible and additional variables are integrated into the model. Moreover, the machine learning model was semi-automatically calibrated. Future research will focus on the optimization algorithms for the model hyperparameters, which will reduce the estimation errors and further improve its accuracy. However, the present work satisfactorily demonstrates the applicability of the methodology used and elucidates the importance of machine learning in modeling hydrological and environmental processes using remotely sensed data.

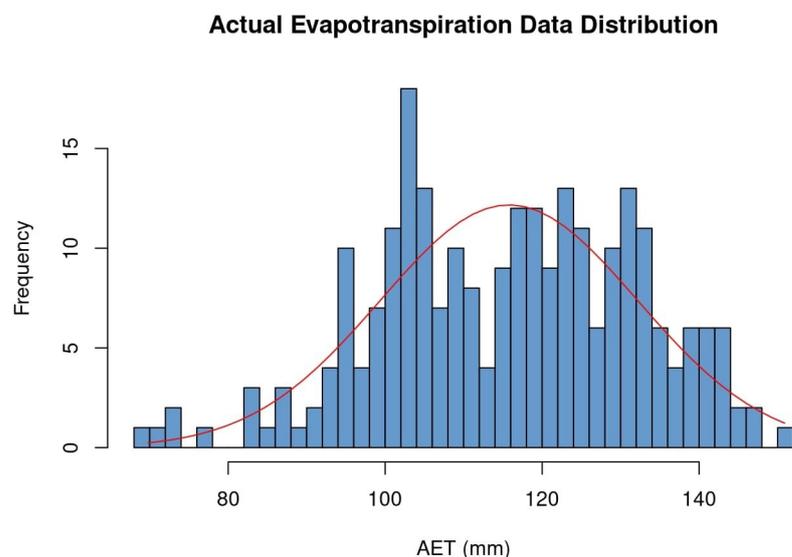
### 3. Results and Discussion

#### 3.1. Descriptive Statistics

A vast difference between a minimum value of 69.8 mm and a maximum value of 150.9 mm was revealed, indicating spatial variability in the actual evapotranspiration distribution in the study area, as shown in Table 2 and Figure 3. Moreover, the normality test demonstrated that the AET estimates were normally distributed, with a slight negative skewness of  $-0.24$ , which was further proved by the disparity between the mean of 115.8 mm/month and the median of 117 mm/month.

**Table 2.** Descriptive statistical estimates of actual evapotranspiration in Busia County for September 2020.

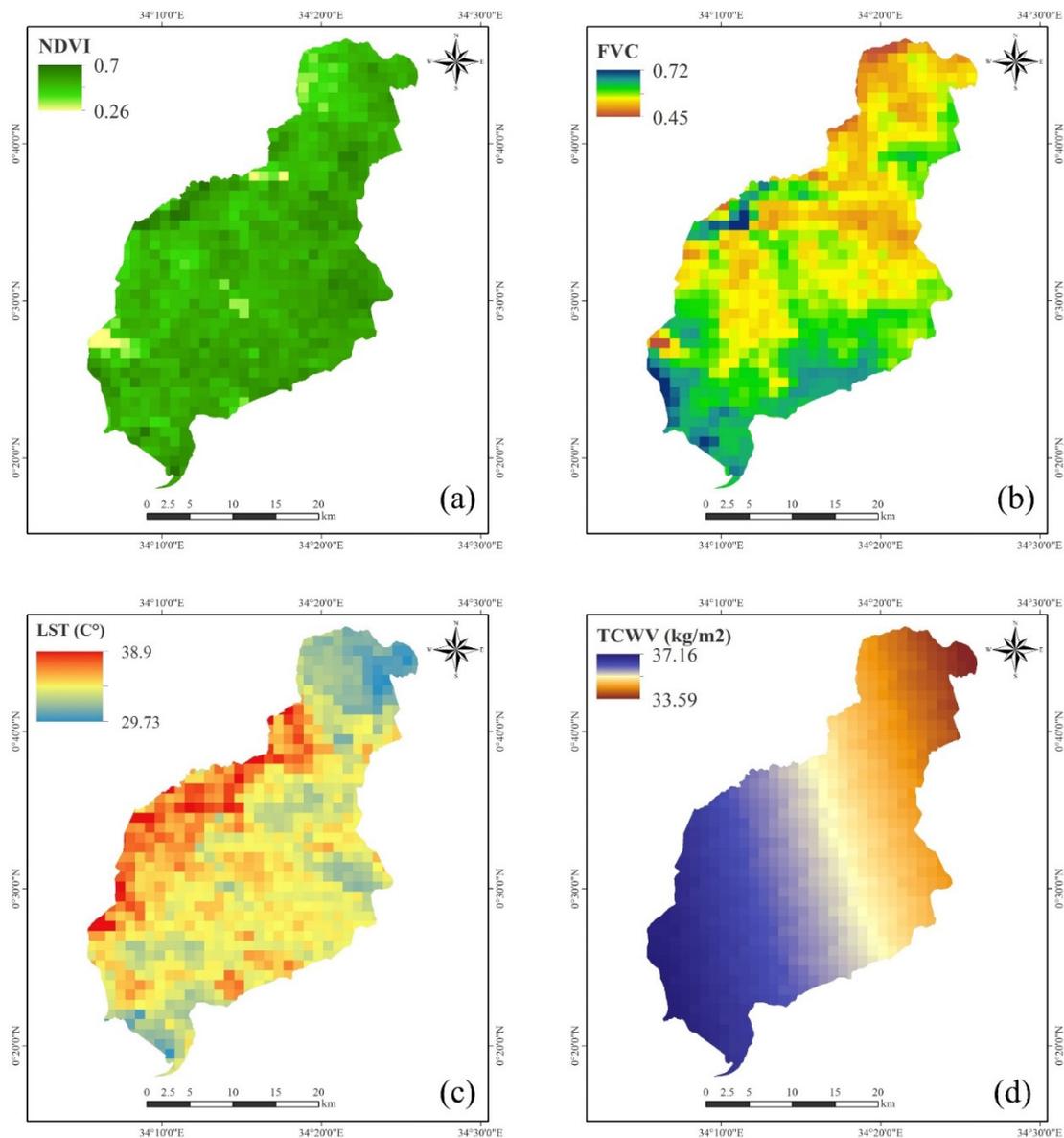
Actual Evapotranspiration (mm/month)				Unitless
Mean	Median	Minimum	Maximum	Skewness
115.8	117	69.8	150.9	−0.24

**Figure 3.** Histogram of the data distribution with the normality curve.

### 3.1.1.1. Sentinel-1 and Sentinel-3 Used Variables

The data derived from Sentinel-1 and Sentinel-3 are shown in Figure 4 below. The NDVI varied from 0.26 to 0.70, while the FVC varied from 0.45 to 0.72. Overall, these variables showed a non-homogenous distribution pattern across Busia County, which agrees with [88], who revealed the remarkable influences of the NDVI and FVC on AET, in addition to topographical characteristics. This can be explained by the significant contribution of vegetation to the increase in the actual evapotranspiration due to the increased available energy absorbed by the canopy, as identified by Zhao et al. [89]. To further support this finding, Klisch and Atzberger [90] demonstrated the applicability of the NDVI derived from MODIS data for drought monitoring, since low estimated NDVI values indicate stressed-out vegetation, which is an indicator of drought occurrence in most scenarios [91]. Therefore, NDVI assessment constitutes the basis for early drought warnings [92]. In addition, the land surface temperature (LST) values ranged from 29.7 °C to 38.9 °C, while the TCWV estimates ranged from 33.6 kg/m<sup>2</sup> to 37.2 kg/m<sup>2</sup>. The LST and TCWV values were relatively high, directly contributing to high energy availability, indicating higher AET estimates, as suggested by [93]. Lower values indicate that the AET and LST negatively relate to air surface temperature changes [94]. The AET was found to be proportionally increased in regions with a high net solar radiation and air surface temperature, greatly influenced by the increase in evapotranspiration intensity, coupled with increased atmospheric evaporative demands, thus further increasing the frequency of droughts [58]. Nonetheless, compared to the tropics, the AET and LST have a positive relationship in high-altitude regions [95]. Although LST variation can be fundamental in selecting the wettest and driest pixels, as stated by Wang et al. [96], this may also introduce uncertainties and increase prediction errors. The authors of [97] found that Sentinel-1 and Sentinel-3 sensors can provide estimate values of the NDVI and LST to ascertain spatiotemporal vegetation dynamics, droughts, and water availability in water stress conditions, respectively, which are essential factors influencing, and significant driving forces of, the AET distribution. This is consistent with Arast et al. [39], who

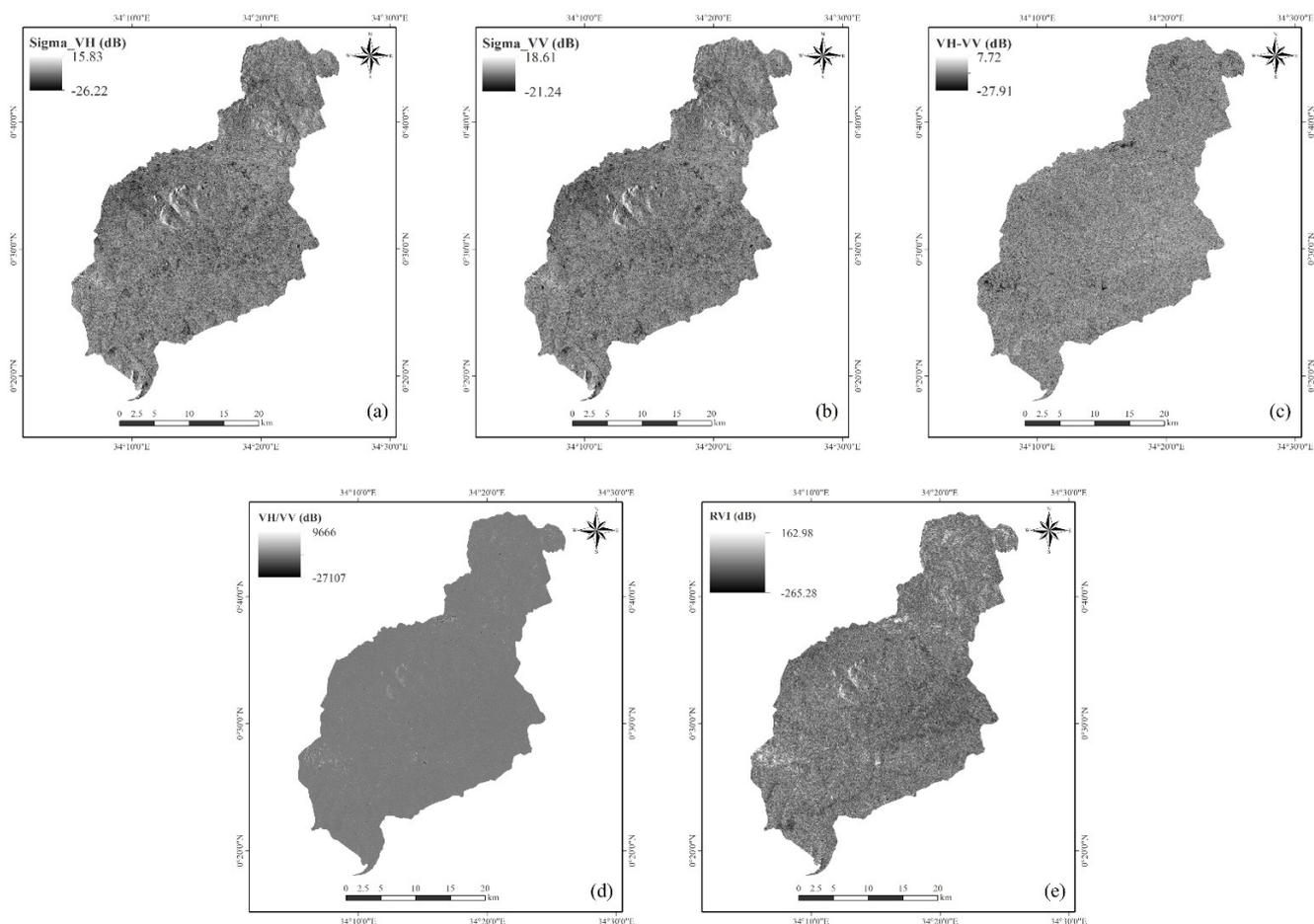
demonstrated that the NDVI, net solar radiation, and other meteorological parameters influence AET. The data from the two sensors are efficiently sufficient to derive the variables under investigation and explore their associations with AET.



**Figure 4.** Variables derived from Sentinel-3 SLSTR data: (a) NDVI, (b) FVC, (c) LST, and (d) TCWV.

The radar variables, i.e., the VH, VV, the ratio (VH/VV), the difference (VV–VH), and RVI, derived from the Sentinel-1 SAR data, are presented in Figure 5. Since an electromagnetic signal received by radar sensors is highly influenced by the surface [98], the backscatter intensity of VH polarization ranged from  $-26.2$  dB to  $15.8$  dB, while the backscatter intensity of VV polarization was slightly more robust, ranging from  $-21.2$  dB to  $18.6$  dB. Thus, this shows a more substantial variation compared to the VV backscatter values since the more robust the co-polarization (HH or VV) reflection is, the brighter the SAR image will be [99]. As former studies indicate, a decrease in the backscatter intensity is chiefly attributed to vegetation growth, causing volume scattering [45,100]. The difference (VV–VH) values ranged from  $-27.9$  dB to  $7.7$  dB, while the backscatter estimated values for the ratio (VH/VV) ranged from  $-27,107$  to  $9666$ . Moreover, the ratio (VH/VV) decreased with a change in the land cover to scarce vegetation and non-vegetated rocky terrains, e.g.,

the Kavirondo Rock series in Busia. According to many studies, the value increases during the vegetation growth season [70,99], leading to the more significant influence of vegetation biomass [45]. Further analysis showed that the RVI estimated values ranged from  $-265.38$  to  $163.0$  dB. In general, the reflected energy drastically varied in Busia County according to the vegetation vigor and density, which are proportionally associated with the soil moisture content and canopy physiology in various growth stages [70]. Based on Figure 5, the bright features can probably be attributed to riparian vegetation [99] along the water bodies, such as the River Mosokoto and River Sio, and swampy areas, such as the Yala swamp, one of Kenya's most extensive freshwater wetlands.



**Figure 5.** Remotely sensed variables derived from Sentinel-1 SAR data: (a) VH, (b) VV, (c) difference (VH–VV), (d) ratio, and (e) RVI.

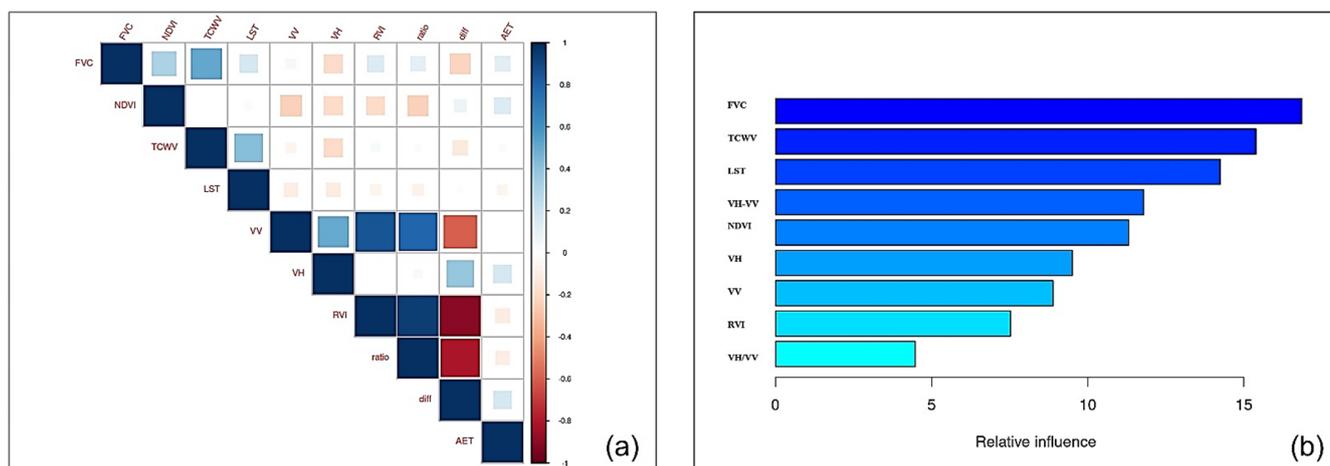
### 3.1.2. Model Training Using a Random Search

Hyperparameter optimization enables decision making concerning the most important hyperparameters and tuning spaces. Therefore, we used the random grid search method to calibrate the model and optimize its hyperparameters by defining the search space as a bounded domain of hyperparameter values and randomly sampling points in that domain [101]. The authors of [102] indicated that a random search could tremendously improve the model accuracy by successfully probing a larger configuration space. When the random search is compared with the grid search, according to Larochelle et al. [103], although the grid search is one of the most extensively utilized hyperparameter optimization algorithms [102], the random search, over the same domain, effectively identifies accurate models with minimal processing. The authors of [104] also indicated that hyperparameters must be established before starting the process.

For GBM calibration, three hyperparameters are supposed to be optimized, including the number of trees, learning rate, and depth of each tree. The number of trees represents the total number of trees in the sequence or ensemble. The averaging of separately growing trees in bagged and random forests renders overfitting with too many trees exceedingly tricky. GBMs, on the other hand, work differently since each tree is built in sequence so as to correct the flaws of the previous tree. The learning rate determines the extent to which each tree contributes to the final output and affects how rapidly the algorithm descends the gradient descent. Typical values vary from 3 to 8, yet a tree depth of 1 is not uncommon [78]. A detailed explanation of the method used to calibrate a GBM is presented in [105]. In this study, the best hyperparameters determined using a random search were the ntree (Number of trees) of 800, shrinkage or learning rate of 0.01, and interaction depth (depth of each tree) of 3.

### 3.1.3. Relative Influences of the Variables on the Model

The calibrated model was statistically significant in estimating the AET over the study area based on the existing reference’s physical background and AET values estimated by the FAO in Busia County. A split criterion was applied so as to better understand and visualize the explanatory variable’s influence on the model prediction of AET. In Figure 6, a demonstration of the most to least influential variables affecting the AET prediction model is shown. The more substantial the influence of the response and explanatory variable is, the larger the value is. Figure 6b illustrates that the FVC is the most influential explanatory variable, with the most significant impact on the modeling and estimation of the AET in Busia. This variable is consistent with the drought patterns in the same area since low values indicate insufficient amounts of precipitation. The total column water vapor and land surface temperature influenced the AET on various scales, including daily and monthly, according to Rocha et al. [96] and Wu et al. [106], revealing their great potential for application in many agronomy-based studies [96,107,108]. These variables are imperative for estimating regional AET, provided that soil moisture content and related variables are available in the remote sensing area [109].



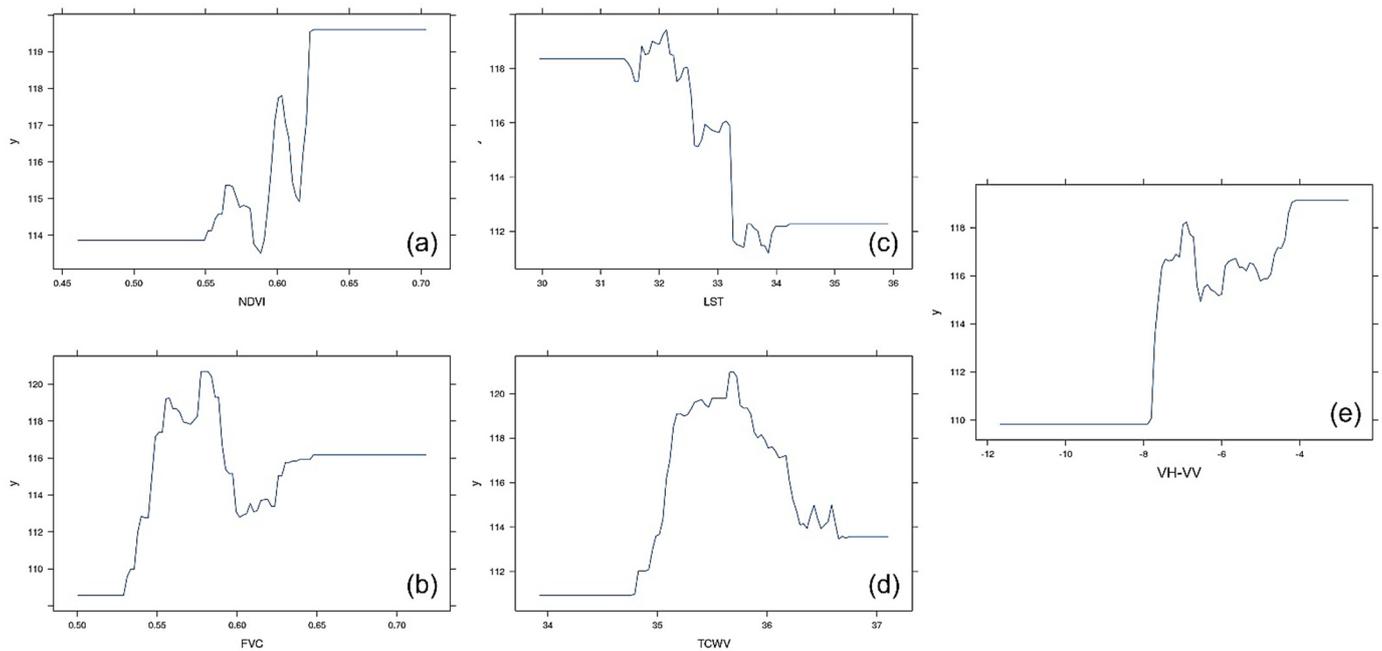
**Figure 6.** (a) Pearson’s correlation between the AET and remotely sensed data and (b) variables’ importance based on the GBM model.

Similarly, Probst et al. [94] found that LST variation is strongly associated with vegetation evapotranspiration and energy balance in the case of wet soil and plants. Furthermore, the LST enhances evapotranspiration in cold air and unlimited soil water in inadequate precipitation conditions. In addition, droughts, high temperatures, and stronger radiative forcing lead to the drying propensity of the surface due to high evapotranspiration rates and low soil moisture, inducing an increase in the heat flux and high temperatures [58].

Furthermore, Yang et al. [69] demonstrated that NDVI patterns are usually consistent with the AET spatial distribution, while [45] demonstrated that the SAR backscatter and NDVI can be used in various physical environmental conditions because they have suitable optical plant properties. The least influential variables were the radar vegetation index (RVI), which measures the randomness of the scattering [110], and the ratio VH/VV, both associated with vegetation conditions. A study by Szigarski et al. [111] indicated that the correlation between the RVI and other indices depends on the other indices' independence from the surface roughness and soil moisture.

Furthermore, as stated by Rosenqvist et al. [99], surface roughness and hilly terrain may cause strong reflection, and VH polarization demonstrates multiple scattering and, hence, a low influence on AET. From Figure 6a, the correlation matrix demonstrates variations between the AET and remotely sensed data, indicating estimated positive correlations between the AET and FVC, NDVI, VH, and VH-VV. The findings agree with Yan et al. [112], who used cloud-free MODIS images from 2000 to 2014 with the ETWatch system and found that the NDVI positively correlated with the AET. Ma et al. [113] also demonstrated the significance of the FVC as a driving parameter that affected the AET and influenced its variation. In addition, there was an estimated weak negative correlation between the AET, VH/VV, and RVI, while there appeared to be no correlation between the AET and VV, LST, and TCWV. In agreement with our finding, Yan et al. [112] found that AET and LST are negatively correlated in water-scarce areas at various spatiotemporal extents.

After identifying the most to least influential variables (Figure 6b), partial dependence plots (PDPs) were created to visualize and understand the response variable changes, as presented in Figure 7. Evidently, they demonstrated the change in the average predicted AET ( $y$ ) values. As demonstrated by the PDPs, the AET estimated values increased as the majority of the variables used in the model increased. For instance, the predicted AET estimated values increased with the increase in the NDVI (Figure 7a), which agrees with [95,114–116], who found that this variable has a significant correlation with the AET and is closely linked to green-leaf-area- and vegetation-based indices. A varying trend in the AET, with the NDVI, cover varying between 55% and 62.5%, occurred, which does not necessarily demonstrate water availability but the greenery characteristic of vegetation [116]. The results further concur with Arast et al. [39], who found that larger NDVI values usually indicate increased AET estimates over different spatial and temporal timescales. The NDVI pattern did not entirely harmonize with the FVC (Figure 7b) response change characterization. For instance, the AET estimates were the highest for a vegetation density of 55% to 60%. In addition, the LST (Figure 7c) demonstrated varying trends in its response change, with the highest AET estimates recorded for temperatures ranging between 32 °C and 33 °C, whereas the highest response change in the TCWV variable (7d) demonstrated the highest predicted AET estimates (Figure 7d). The smaller the VH–VV (Figure 7e) index response was, the higher the predicted AET estimates for the radar variables were. Specifically, the PDPs satisfactorily demonstrated the change in the predicted average AET estimates as the variables varied in their distribution.



**Figure 7.** Variation in the AET(y) with the most significant covariates in Busia County: (a) NDVI, (b) FVC, (c) LST, (d) TCWV, and (e) VH – VV.

3.2. Accuracy Assessment Using Correlation Coefficient (R) and Root-Mean-Squared Error (RMSE)

The spatial distribution of the AET can be estimated using the GBM final model, which contributes to the improvement and development of the proper utilization of water systems [117]. The correlation coefficient R and root-mean-squared error (RMSE) were used to evaluate the model’s statistical performance. The two-assessment metrics can be determined using Equations (1) and (2):

$$RMSE = \sqrt{\sum_1^n (\hat{y}_I - y_i)^2 / n} \tag{4}$$

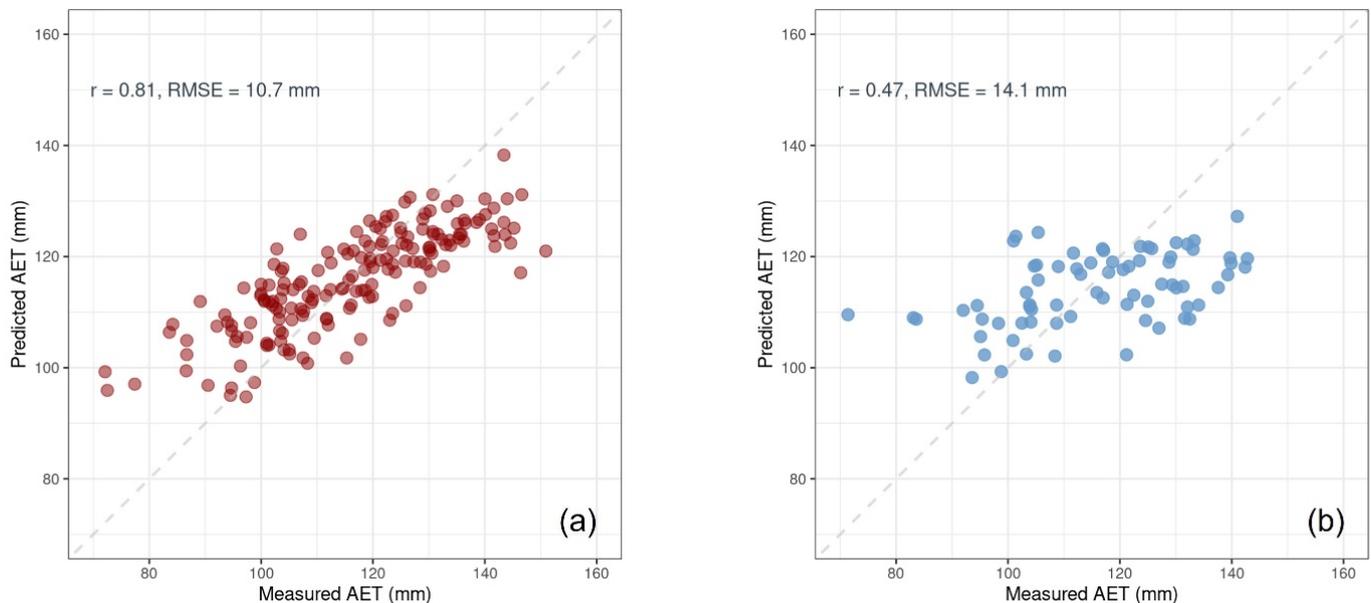
where  $\hat{y}_i$  is the predicted value for the  $i^{th}$  observation,  $y_i$  is the observed value, and  $n$  is the total number of observations:

$$R = \sum(x_i - \bar{x})(y_i - \bar{y}) / \sqrt{\sum(x_i - \bar{x})^2 (y_i - \bar{y})^2} \tag{5}$$

where  $x_i$  is the  $x$ -variable in a sample,  $\bar{x}$  is the mean of the  $x$ -variable values,  $y_i$  is the  $y$ -variable in a sample, and  $\bar{y}$  is the mean of the  $y$ -variable values.

The model established a relationship between the remotely sensed data retrieved from Sentinel-1 and Sentinel-3 and the AET reference data. The calibrated model yielded a correlation coefficient  $r$  of 0.81 and an RMSE value of 10.7 mm, indicating its efficiency in predicting the AET. Figure 8a illustrates the relationship between the measured and estimated AET values (in mm) using the calibrated GBM model based on only two satellite images. The model established a statistical significance, yet an overestimation occurred in a few cases. The reason for this is that the model was semi-automatically calibrated. However, it was satisfying, and future work will entail the application of more processed sensor data for further investigations regarding climatological and ecological dynamics, as evidenced in a study by the authors of [118]. Based on Figure 8b, the GBM testing results showed a reasonably moderate association between the measured and predicted AET values (in mm), with a correlation  $R$  coefficient of 0.47 and an RMSE of 14.1 mm. This revealed the applied approach’s potential statistical significance, which can be used as a

timeless basis for future studies aiming to model and map the AET spatial distribution in similar environmental and climatic conditions.



**Figure 8.** Relationship between the observed and predicted AET values for (a) the training set (80%) and (b) the testing set (20%).

#### 4. Conclusions

This research aimed to estimate the monthly mean actual evapotranspiration using Sentinel-1 SAR ground-range-detected (GRD) and Sentinel-3 SLSTR Level 2 data in a typical tropical climate in Busia County. The gradient boosting machine was trained and tested using reference data acquired from the WaPOR. The model showed a strong correlation ( $r = 0.81$ ) between the observed and estimated AET data for the training and a moderate correlation ( $r = 0.47$ ) for the testing, revealing the superiority of the applied method. The FVC was a highly influential explanatory variable, with the most significant impact on the prediction model for AET estimation, while the ratio VV/VH was the least influential variable regarding the AET estimation model. Although the remotely sensed data and GBM application undoubtedly yielded promising results, further examination is highly recommended using other machine learning algorithms to optimize the approach's efficiency and explore the nature of the statistical relationships between the AET and applied variables. This could enable the timely and consistent monitoring of actual evapotranspiration, water deficiencies, and agricultural sustainability, as well as ensure food security. This research further enhances our understanding of AET assessment and the potential of using Sentinel-1 and Sentinel-3 data for regional drought monitoring and natural resources management. With a relatively successful estimation of the AET, drought events can easily be predicted in future studies since the AET is one of the primary factors of drought magnitude and occurrence. Further research will be carried out on larger scales and in different climatic regions to validate the applicability of the proposed methodology.

**Author Contributions:** Conceptualization, P.K.M. and G.S.; data processing and code writing, G.S.; writing—original draft preparation P.K.M.; writing—review and editing, P.K.M. and G.S.; data interpretation, manuscript revision, G.T., T.W. and B.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research was supported by OTKA No. K-138176, GINOP-2.3.2-15-2016-00007, GINOP-2.3.2-15-2016-0005. The Stipendium Hungaricum doctoral research scholarship of the Hungarian government is financially supporting Peter K. Musyimi and Ghada Sahbeni's research.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that supported this research can be found on the Copernicus Open Access Hub (<https://scihub.copernicus.eu/dhus/#/home>) and the FAO WaPOR (<https://wapor.apps.fao.org/>). R code used in this study can be shared upon request.

**Acknowledgments:** The authors would like to thank the Editors, Assistant Editors, and anonymous reviewers for their support and constructive comments that helped improve the quality of this scientific article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. UN. The 2030 Agenda and the Sustainable Development Goals: An opportunity for Latin America and the Caribbean (LC/G.2681-P/Rev.3), Santiago. 2018. Available online: [https://www.cepal.org/sites/default/files/events/files/2030\\_agenda\\_and\\_the\\_sdgs\\_an\\_opportunity\\_for\\_latin\\_america\\_and\\_the\\_caribbean.pdf](https://www.cepal.org/sites/default/files/events/files/2030_agenda_and_the_sdgs_an_opportunity_for_latin_america_and_the_caribbean.pdf) (accessed on 18 September 2022).
2. UN-Water. *Step-by-Step Methodology for Monitoring Water Use Efficiency (6.4.1)*; Annual Technical Report; UN-Water: Geneva, Switzerland, 2019. Available online: <https://www.unwater.org/publications/step-step-methodology-monitoring-water-use-efficiency-641/> (accessed on 18 September 2022).
3. UN-Water. *Step-by-Step Methodology for Monitoring Water Stress (6.4.2)*; Annual Technical Report; UN-Water: Geneva, Switzerland, 2019. Available online: <https://www.unwater.org/publications/step-step-methodology-monitoring-water-stress-64-2/> (accessed on 18 September 2022).
4. ET4FAO 2022. Increasing Crop Water Use Efficiency at Multiple Scales Using Sentinel Evapotranspiration (ET4FAO). Available online: <https://eo4society.esa.int/projects/et4fao/> (accessed on 10 September 2022).
5. Food and Agriculture Organization of the United Nations. AQUASTAT-FAO. FAO's Information System on Water and Agriculture. 2018. Available online: <http://www.fao.org/aquastat/en/> (accessed on 18 September 2022).
6. Budagovskiy, A.I.; Novák, V. Theory of evapotranspiration: Transpiration and its quantitative description. *J. Hydrol. Hydromech.* **2011**, *59*, 3–23. [[CrossRef](#)]
7. Baldocchi, D.D.; Falge, E.; Wilson, K. A spectral analysis of biosphere-atmosphere trace gas flux densities and meteorological variables across hour to multi-year time scales. *Agric. For. Meteorol.* **2001**, *107*, 1–27. [[CrossRef](#)]
8. Trenberth, K.E.; Smith, L.; Qian, T.; Dai, A.; Fasullo, J. Estimates of the global water budget and its annual cycle using observational and model data. *J. Hydrometeorol.* **2007**, *8*, 758–769. [[CrossRef](#)]
9. El-Shirbeny, M.A.; Abutaleb, K.A. Monitoring of Water-Level Fluctuation of Lake Nasser Using Altimetry Satellite Data. *Earth Syst. Environ.* **2018**, *2*, 367–375. [[CrossRef](#)]
10. El-Shirbeny, M.A.; Saleh, S.M. Actual evapotranspiration evaluation based on multi-sensed data. *J. Arid. Agric.* **2021**, *7*, 95–102. [[CrossRef](#)]
11. El-Shirbeny, M.A.; Ali, A.M.; Khder, G.A.; Saleh, N.H.; Afify, N.M.; Badr, M.A.; Bauomy, E.M. Monitoring agricultural water in the desert environment of New Valley Governorate for sustainable agricultural development: A case study of Kharga. *Euro-Mediterranean J. Environ. Integr.* **2021**, *6*, 56. [[CrossRef](#)]
12. Koster, R.D.; Dirmeyer, P.A.; Guo, Z.; Bonan, G.; Chan, E.; Cox, P.; Gordon, C.T.; Kanae, S.; Kowalczyk, E.; Lawrence, D.; et al. Regions of strong coupling between soil moisture and precipitation. *Science* **2004**, *305*, 1138–1140. [[CrossRef](#)] [[PubMed](#)]
13. Seneviratne, S.I.; Corti, T.; Davin, E.L.; Hirschi, M.; Jaeger, E.B.; Lehner, I.; Orlowsky, B.; Teuling, A.J. Investigating soil moisture–climate interactions in a changing climate: A review. *Earth Sci. Rev.* **2010**, *99*, 125–161. [[CrossRef](#)]
14. Jung, M.; Reichstein, M.; Ciais, P.; Seneviratne, S.I.; Sheffield, J.; Goulden, M.L.; Bonan, G.; Cescatti, A.; Chen, J.; de Jeu, R.; et al. Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature* **2010**, *467*, 951–954. [[CrossRef](#)]
15. Malik, A.; Saggi, M.K.; Rehman, S.; Sajjad, H.; Inyurt, S.; Bhatia, A.S.; Farooque, A.A.; Oudah, A.Y.; Yaseen, Z.M. Deep learning versus gradient boosting machine for pan evaporation prediction. *Eng. Appl. Comput. Fluid Mech.* **2022**, *16*, 570–587. [[CrossRef](#)]
16. Shrivastav, L.K.; Jha, S.K. A gradient boosting machine learning approach in modeling the impact of temperature and humidity on the transmission rate of COVID-19 in India. *Appl. Intell.* **2021**, *51*, 2727–2739. [[CrossRef](#)] [[PubMed](#)]
17. Frey, U.J. Putting machine learning to use in natural resource management-improving model performance. *Ecol. Soc.* **2020**, *25*, 45. [[CrossRef](#)]
18. Gagne, D.J.; Haupt, S.E.; Nychka, D.W.; Thompson, G. Interpretable deep learning for spatial analysis of severe hailstorms. *Mon. Weather. Rev.* **2019**, *147*, 2827–2845. [[CrossRef](#)]
19. McGovern, A.; Elmore, K.L.; Gagne, D.J., II; Haupt, S.E.; Karstens, C.D.; Lagerquist, R.; Smith, T.; Williams, J.K. Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bull. Amer. Meteor. Soc.* **2017**, *98*, 2073–2090. [[CrossRef](#)]
20. Castellvi, F.; Snyder, R.L. A comparison between latent heat fluxes over grass using a weighing lysimeter and surface renewal analysis. *J. Hydrol.* **2010**, *381*, 213–220. [[CrossRef](#)]

21. Faharani, H.J.; Howell, T.A.; Shuttleworth, W.J.; Bausch, W.C. Evapotranspiration; Progress in measurement and modeling in agriculture. *Trans. ASABE* **2007**, *50*, 1627–1638. [[CrossRef](#)]
22. Loos, C.; Gayler, S.; Priesack, S. Assessment of water balance simulations for large-scale weighing lysimeters. *J. Hydrol.* **2007**, *335*, 259–270. [[CrossRef](#)]
23. Wohlfahrt, G.; Ischick, C.; Thalinger, B.; Hörtnagl, L.; Obojes, N.; Hammerle, A. Insights from independent evapotranspiration estimates for closing the energy balance: A grassland case study. *Vadose Zone J.* **2010**, *9*, 1025–1033. [[CrossRef](#)]
24. Wegehenkel, M.; Gerke, H. Comparison of real evapotranspiration measured by weighing lysimeters with simulations based on the Penman formula and a crop growth model. *J. Hydrol. Hydromech.* **2013**, *61*, 161–172. [[CrossRef](#)]
25. Chen, Y.; Xia, J.; Liang, S.; Feng, J.; Fisher, J.B.; Li, X.; Li, X.; Liu, S.; Ma, Z.; Miyata, A.; et al. Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China. *Remote Sens. Environ.* **2014**, *140*, 279–293. [[CrossRef](#)]
26. Jiménez, C.; Prigent, C.; Mueller, B.; Seneviratne, S.I.; McCabe, M.F.; Wood, E.F.; Rossow, W.B.; Balsamo, G.; Betts, A.K.; Dirmeyer, P.A.; et al. Global intercomparison of 12 land surface heat flux estimates. *J. Geophys. Res. Atmos.* **2011**, *116*, D02102. [[CrossRef](#)]
27. Miralles, D.G.; Holmes, T.R.H.; De Jeu, R.A.M.; Gash, J.H.; Meesters, A.G.C.A.; Dolman, A.J. Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 453–469. [[CrossRef](#)]
28. Mueller, B.; Seneviratne, S.I.; Jimenez, C.; Corti, T.; Hirschi, M.; Balsamo, G.; Ciais, P.P.; Dirmeyer, P.; Fisher, J.B.; Guo, Z.; et al. Evaluation of global observation-based evapotranspiration datasets and IPCC AR4 simulations. *Geophys. Res. Lett.* **2011**, *38*, L06402. [[CrossRef](#)]
29. McMahan, T.A.; Peel, M.C.; Lowe, L.; Srikanthan, R.; McVicar, T.R. Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: A pragmatic synthesis. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 1331–1363. [[CrossRef](#)]
30. Kim, S.; Sharma, A. The role of floodplain topography in deriving basin discharge using passive microwave remote sensing. *Water Resour. Res.* **2019**, *55*, 1707–1716. [[CrossRef](#)]
31. Libertino, A.; Sharma, A.; Lakshmi, V.; Claps, P. A global assessment of the timing of extreme rainfall from TRMM and GPM for improving the hydrologic design. *Environ. Res. Lett.* **2016**, *11*, 54003. [[CrossRef](#)]
32. Zhang, R.; Kim, S.; Sharma, A. A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes. *Remote Sens. Environ.* **2019**, *223*, 82–94. [[CrossRef](#)]
33. Long, D.; Longueuegne, L.; Scanlon, B.R. Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. *Water Resour. Res.* **2014**, *50*, 1131–1151. [[CrossRef](#)]
34. Bastiaanssen, W.G.M.; Menenti, M.; Feddes, R.A.; Holtslag, A.A.M. A remote sensing surface energy balance algorithm for land (SEBAL), Part 1: Formulation. *J. Hydrol.* **1998**, *212–213*, 198–212. [[CrossRef](#)]
35. Su, Z. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.* **2002**, *6*, 85–100. [[CrossRef](#)]
36. Loheide, S.P.; Gorelick, S.M. A local-scale, high-resolution evapotranspiration mapping algorithm (ETMA) with hydroecological applications at riparian meadow restoration sites. *Remote Sens. Environ.* **2005**, *98*, 182–200. [[CrossRef](#)]
37. Anderson, M.C.; Norman, J.M.; Mecikalski, J.R.; Otkin, J.A.; Kustas, W.P. A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *J. Geophys. Res. Atmos.* **2007**, *112*, 1–17. [[CrossRef](#)]
38. Velpuri, N.M.; Senay, G.B.; Singh, R.K.; Bohms, S.; Verdin, J.P. A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET. *Remote Sens. Environ.* **2013**, *139*, 35–49. [[CrossRef](#)]
39. Arast, M.; Ranjbar, A.; Mousavi, S.H.; Abdollahi, K. Assessment of the Relationship between NDVI-Based actual evapotranspiration by SEBS. *Iran. J. Sci. Technol. Trans. A Sci.* **2020**, *44*, 1051–1062. [[CrossRef](#)]
40. Nagler, P.L.; Glenn, E.P.; Nguyen, U.; Scott, R.L.; Doody, T. Estimating riparian and agricultural actual evapotranspiration by reference evapotranspiration and MODIS enhanced vegetation index. *Remote Sens.* **2013**, *5*, 384913871. [[CrossRef](#)]
41. McCabe, M.F.; Wood, E.F.; Wójcik, R.; Pan, M.; Sheffield, J.; Gao, H.; Su, H. Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies. *Remote Sens. Environ.* **2008**, *112*, 430–444. [[CrossRef](#)]
42. Conroy, J.W.; Wu, J.; Elliot, W. Modification of the evapotranspiration routines in the WEPP model: Part I. In Proceedings of the ASAE Annual International Meeting, Las Vegas, NV, USA, 27–30 July 2003; pp. 1–16. Available online: <http://www.pubs.asce.org/WWWdisplay.cgi?8801815> (accessed on 1 October 2022).
43. Builes, V.H.R.; Porch, T.G.; Harmsen, E.W. Genotypic differences in water use efficiency of common bean under drought stress. *Agron. J.* **2011**, *103*, 1206–1215. [[CrossRef](#)]
44. Donlon, C.; Berruti, B.; Buongiorno, A.; Ferreira, M.H.; Féménias, P.; Frerick, J.; Goryl, P.; Klein, U.; Laur, H.; Mavrocordatos, C.; et al. The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. *Remote Sens. Environ.* **2012**, *120*, 37–57. [[CrossRef](#)]
45. Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* **2017**, *199*, 415–426. [[CrossRef](#)]
46. Guzinski, R.; Nieto, H. Evaluating the feasibility of using Sentinel-2 and Sentinel-3 satellites for high-resolution evapotranspiration estimations. *Remote Sens. Environ.* **2019**, *221*, 157–172. [[CrossRef](#)]
47. Seelan, S.K.; Laguetta, S.; Casady, G.M.; Seielstad, G.A. Remote sensing applications for precision agriculture: A learning community approach. *Remote Sens. Environ.* **2003**, *88*, 157–169. [[CrossRef](#)]

48. Guzinski, R.; Nieto, H.; Sandholt, I.; Karamitlios, G. Modeling High-Resolution Actual Evapotranspiration through Sentinel-2 and Sentinel-3 Data Fusion. *Remote Sens.* **2020**, *12*, 1433. [CrossRef]
49. Nsiah, J.J.; Gyamfi, C.; Anornu, G.K.; Odai, S.N. Estimating the spatial distribution of evapotranspiration within the Pra River Basin of Ghana. *Heliyon* **2021**, *7*, e06828. [CrossRef] [PubMed]
50. Shilenje, Z.W.; Murage, P.; Ongoma, V. Estimation of Potential Evaporation Based on Penman Equation under Varying Climate, for Murang'a County, Kenya. *Pak. J. Meteorol.* **2015**, *12*, 33–42. Available online: <https://www.researchgate.net/publication/297917775> (accessed on 1 September 2022).
51. Ogallo, L. The mainstreaming of climate change and variability information into planning and policy development for Africa. *Procedia Environ. Sci.* **2010**, *1*, 405–410. [CrossRef]
52. Marshall, M.T.; Funk, C.; Michaelsen, J. Agricultural Drought Monitoring in Kenya Using Evapotranspiration Derived from Remote Sensing and Reanalysis Data. *USGS Staff. Publ. Res.* **2012**, *978*, 1–30. Available online: <http://digitalcommons.unl.edu/usgsstaffpub/978> (accessed on 1 September 2022).
53. Akuja, T.E.; Kandagor, J. A review of policies and agricultural productivity in the arid and semiarid lands (ASALS), Kenya: The case of Turkana County. *J. Appl. Biosci.* **2019**, *140*, 14304–14315. [CrossRef]
54. UN-Water and Sanitation. Water Action Decade 2018–2028. Available online: <https://www.un.org/sustainabledevelopment/water-and-sanitation/> (accessed on 5 July 2022).
55. Bhaduri, A.; Bogardi, J.; Siddiqi, A.; Voigt, H.; Vörösmarty, C.; Pahl-Wostl, C.; Bunn, S.E.; Shrivastava, P.; Lawford, R.; Foster, S.; et al. Achieving sustainable development goals from a water perspective. *Front. Environ. Sci.* **2016**, *4*, 64. [CrossRef]
56. Carleton, T.A.; Hsiang, S.M. Social and economic impacts of climate. *Science* **2016**, *353*, aad9837. [CrossRef]
57. Sorre, A.M. Effects of climate change on rural livelihoods in Busia County, Kenya. *Int. J. Agric. Sci.* **2017**, *3*, 75–89. Available online: <https://www.iardjournals.org/> (accessed on 5 July 2022).
58. Masson-Delmotte, V.P.; Zhai, A.; Pirani, S.L.; Connors, C.; Péan, S.; Berger, N.; Caud, Y.; Chen, L.; Goldfarb, M.I.; Gomis, M.; et al. (Eds.) *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2021. Available online: [https://report.ipcc.ch/ar6/wg1/IPCC\\_AR6\\_WGI\\_FullReport.pdf](https://report.ipcc.ch/ar6/wg1/IPCC_AR6_WGI_FullReport.pdf) (accessed on 5 July 2022).
59. KNBS 2019. 2019 Kenya Population and Housing Census. Volume IV: Distribution of Population by Socio-Economic Characteristics. December 2019. Available online: <https://www.knbs.or.ke/?wpdmpro=2019-kenya-population-and-housing-census-volume-iv-distribution-of-population-by-socio-economic-characteristics> (accessed on 5 July 2022).
60. Juma, G.S.; Kituni, N.; Makokha, J.W. Advection, and its applications: Trajectories over Busia County in Kenya. *Clim. Change* **2020**, *6*, 186–190. Available online: [http://www.discoveryjournals.org/climate\\_change/current\\_issue/v6/n22/A3.pdf](http://www.discoveryjournals.org/climate_change/current_issue/v6/n22/A3.pdf) (accessed on 5 July 2022).
61. Makori, A.J.; Abuom, P.O.; Kapiyo, R.; Anyona, O.D.; Dida, O.G. Effects of water Physico-chemical parameters on tilapia (*Oreochromis niloticus*) growth in earthen ponds in Teso North Sub-County, Busia County. *Fish. Aquat. Sci.* **2017**, *20*, 30. [CrossRef]
62. Sorre, A.M.; Kurgat, A.; Musebe, R.; Sorre, B. Adaptive capacity to climate change among smallholder farmers in Busia County, Kenya. *J. Agric. Vet. Sci.* **2017**, *10*, 40–48. Available online: <https://www.iosrjournals.org/iosr-javs/papers/Vol10-issue11/Version-1/H1011014048.pdf> (accessed on 5 July 2022).
63. Kebeney, S.J.; Msanya, B.M.; Ng'etich, W.K.; Semoka, J.M.R.; Serrem, C.K. Pedological Characterization of Some Typical Soils of Busia County, Western Kenya: Soil Morphology, Physico-chemical Properties, Classification, and Fertility Trends. *Int. J. Plant Soil Sci.* **2014**, *4*, 29–44. [CrossRef]
64. Okeyo, B.; Wamugi, S.M. Climate Change Effects and the Resulting Adaptation Strategies of Smallholder Farmers in Three Different Ecological Zones (Kilifi, Embu, and Budalangi) in Kenya. *J. Environ. Earth Sci.* **2018**, *8*, 40–70. Available online: <https://www.iiste.org/Journals/index.php/JEES/article/view/43354/44682> (accessed on 5 July 2022).
65. Wekesa, B.M.; Ayuya, O.I.; Lagat, J.K. Effect of climate-smart agricultural practices on household food security in smallholder production systems: Micro-level evidence from Kenya. *Agric. Food Secur.* **2018**, *7*, 80. [CrossRef]
66. Government of Kenya. *Busia County Integrated Development Plan*; Government Printers: Nairobi, Kenya, 2013.
67. ESA. User Guides. 2022. Available online: <https://sentinels.copernicus.eu/web/sentinel/user-guides> (accessed on 5 July 2022).
68. Coppo, P.; Ricciarelli, B.; Brandani, F.; Delderfield, J.; Ferlet, M.; Mutlow, C.; Munro, G.; Nightingale, T.; Smith, D.; Bianchi, S.; et al. SLSTR: A high accuracy dual scan temperature radiometer for sea and land surface monitoring from space. *J. Mod. Opt.* **2010**, *57*, 1815–1830. [CrossRef]
69. Yang, L.; Li, J.; Sun, Z.; Liu, J.; Yang, Y.; Li, T. Daily actual evapotranspiration estimation of different land use types based on SEBAL model in the agro-pastoral ecotone of northwest China. *PLoS ONE* **2022**, *17*, e0265138. [CrossRef]
70. Harfenmeister, K.; Spengler, D.; Weltzien, C. Analyzing Temporal and Spatial Characteristics of Crop Parameters Using Sentinel-1 Backscatter Data. *Remote Sens.* **2019**, *11*, 1569. [CrossRef]
71. Ishitsuka, K.; Tsuji, T.; Matsuoka, T. Surface Change of the Soil Liquefaction Caused by the 2011 Great East Japan Earthquake Derived from SAR Data. In Proceedings of the International Symposium on Engineering Lessons Learned from the 2011 Great East Japan Earthquake, Tokyo, Japan, 1–4 March 2012. Available online: <https://www.jaee.gr.jp/event/seminar2012/eqsymposium/pdf/papers/26.pdf> (accessed on 5 July 2022).

72. Henderson, F.M.; Lewis, A.J. *Principles and Applications of Imaging Radar. Manual of Remote Sensing*, 3rd ed.; John Wiley and Sons: New York, NY, USA, 1998. Available online: <https://www.osti.gov/biblio/293027> (accessed on 2 October 2022).
73. Filippini, F. Exploitation of Sentinel-2 Time Series to Map Burned Areas at the National Level: A Case Study on the 2017 Italy Wildfires. *Remote Sens.* **2019**, *11*, 622. [CrossRef]
74. Hosen, M.S.; Amin, R. Significant of Gradient Boosting Algorithm in Data Management System. *Eng. Int.* **2021**, *9*, 85–100. [CrossRef]
75. Saatchi, S. SAR methods for mapping and monitoring forest biomass. In *SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation*; Flore, A., Herndon, K., Thapa, R., Cherrington, E., Eds.; Servir Global: Huntsville, AL, USA, 2019; pp. 207–246. [CrossRef]
76. Freund, Y.; Schapire, R. A decision-theoretic generalization of online learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]
77. Friedman, J. Greedy boosting approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
78. Bhagat, S.K.; Tiyasha, T.; Tung, T.M.; Mostafa, R.R.; Yaseen, Z.M. Manganese (Mn) removal prediction using extreme gradient model. *Ecotoxicol. Environ. Saf.* **2020**, *204*, 111059. [CrossRef] [PubMed]
79. Naganna, S.R.; Beyaztas, B.H.; Bokde, N.; Armanuos, A.M. On the evaluation of the gradient tree boosting model for groundwater level forecasting. *Knowl. Based Eng. Sci.* **2020**, *1*, 48–57. [CrossRef]
80. Landry, M.; Bartz, A. *Machine Learning with R and H2O*; H2O.ai: Mountain View, CA, USA, 2022. Available online: <https://h2o.ai/resources/booklet/machine-learning-with-r-and-h2o/> (accessed on 5 August 2022).
81. Landry, M.; Erlinger, P.T.; Patschke, D.; Varrichio, C. Probabilistic gradient boosting machines for GEFCom2014 wind forecasting. *Int. J. Forecast.* **2016**, *32*, 1061–1066. [CrossRef]
82. Sridharan, K.; Komarasamy, G. Sentiment classification using harmony random forest and harmony gradient boosting machine. *Soft Comput.* **2020**, *24*, 7451–7458. [CrossRef]
83. Zhang, Z.; Zhao, Y.; Canes, A.; Steinberg, D.; Lyashevskaya, O. written on behalf of AME Big-Data Clinical Trial Collaborative Group. Predictive analytics with gradient boosting in clinical medicine. *Ann. Transl. Med.* **2019**, *7*, 152. [CrossRef]
84. Khoi, D.N.; Quan, N.T.; Linh, D.Q.; Nhi, P.T.T.; Thuy, N.T.D. Using Machine Learning Models for Predicting the Water Quality Index in the La Buong River, Vietnam. *Water* **2022**, *14*, 1552. [CrossRef]
85. Zalakeviciute, R.; Rybarczyk, Y.; Alexandrino, K.; Bonilla-Bedoya, S.; Mejia, D.; Bastidas, M.; Diaz, V. Gradient Boosting Machine to Assess the Public Protest Impact on Urban Air Quality. *Appl. Sci.* **2021**, *11*, 12083. [CrossRef]
86. Bhagat, S.K.; Tung, T.M.; Yaseen, Z.M. Heavy metal contamination prediction using ensemble model: Case study of Bay sedimentation, Australia. *J. Hazard. Mater.* **2021**, *403*, 123492. [CrossRef] [PubMed]
87. Natekin, A.; Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurobot.* **2013**, *7*, 21. [CrossRef] [PubMed]
88. Gibson, L.A.; Münch, Z.; Engelbrecht, J. Particular uncertainties encountered in using a pre-packaged SEBS model to derive evapotranspiration in a heterogeneous study area in South Africa. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 295–310. [CrossRef]
89. Zhao, F.; Ma, S.; Wu, Y.; Qiu, L.; Wang, W.; Lian, Y.; Chen, J.; Sivakumar, B. The role of climate change and vegetation greening on evapotranspiration variation in the Yellow River basin, China. *Agric. For. Meteorol.* **2022**, *316*, 108842. [CrossRef]
90. Klisch, A.; Atzberger, C. Operational Drought Monitoring in Kenya Using MODIS NDVI Time Series. *Remote Sens.* **2016**, *8*, 267. [CrossRef]
91. Javadnia, E.; Mobasher, M.R.; Kamali, G.A. MODIS NDVI quality enhancement using ASTER images. *J. Agr. Sci. Tech.* **2009**, *2*, 549–558. Available online: <https://www.researchgate.net/publication/228947571> (accessed on 5 July 2022).
92. Barrett, A.B.; Duivenvoorden, S.; Salakpi, E.E.; Muthoka, J.M.; Mwangi, J.; Oliver, S.; Rowhani, P. Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya. *Remote Sens. Environ.* **2020**, *248*, 111886. [CrossRef]
93. Schirmbeck, J.; Fontana, D.C.; Roberti, D.R. Evaluation of OSEB and SEBAL models for energy balance of a crop area in a humid subtropical climate. *Bragantia* **2018**, *77*, 609–621. [CrossRef]
94. Sun, Z.; Wang, Q.; Batkhisig, O.; Ouyang, Z. Relationship between Evapotranspiration and Land Surface Temperature under Energy- and Water-Limited Conditions in Dry and Cold Climates. *Adv. Meteorol.* **2016**, *2016*, 1–9. [CrossRef]
95. Wang, K.; Dickinson, R.E. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Rev. Geophys.* **2012**, *50*, 2. [CrossRef]
96. Rocha, N.S.d.; Käfer, P.S.; Skokovic, D.; Veeck, G.; Diaz, L.R.; Kaiser, E.A.; Carvalho, C.M.; Cruz, R.C.; Sobrino, J.A.; Roberti, D.R.; et al. The Influence of Land Surface Temperature in Evapotranspiration Estimated by the S-SEBI Model. *Atmosphere* **2020**, *11*, 1059. [CrossRef]
97. Saadi, S.; Boulet, G.; Bahir, M.; Brut, A.; Delogu, É.; Fanise, P.; Mougnot, B.; Simonneaux, V.; Lili Chabaane, Z. Assessment of actual evapotranspiration over a semiarid heterogeneous land surface by means of coupled low-resolution remote sensing data with an energy balance model: Comparison to extra-large aperture scintillometer measurements. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 2187–2209. [CrossRef]
98. Brown, C.E.; Fingas, M.F.; Hawkins, R. Synthetic aperture radar sensors: Viable for marine oil spill response? In Proceedings of the Arctic and Marine OIL SPILL Program Technical Seminar, Ottawa, ON, Canada, 10 January 2003; Volume 1, pp. 299–310. Available online: [https://inis.iaea.org/search/search.aspx?orig\\_q=RN:34059651](https://inis.iaea.org/search/search.aspx?orig_q=RN:34059651) (accessed on 5 July 2022).

99. Rosenqvist, A.; Perez, A.; Olfindo, N. *A Layman's Interpretation Guide to L-Band and C-Band Synthetic Aperture Radar Data*; Committee on Earth Observation Satellites: Washington, DC, USA, 2018. Available online: [https://ceos.org/ard/files/Laymans\\_SAR\\_Interpretation\\_Guide\\_2.0.pdf](https://ceos.org/ard/files/Laymans_SAR_Interpretation_Guide_2.0.pdf) (accessed on 5 July 2022).
100. Vreugdenhil, M.; Wagner, W.; Bauer-Marschallinger, B.; Pfeil, I.; Teubner, I.; Rüdiger, C.; Strauss, P. Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study. *Remote Sens.* **2018**, *10*, 1396. [[CrossRef](#)]
101. Liashchynskiy, P.; Liashchynskiy, P. Grid search, random search, genetic algorithm: A big comparison for NAS. *arXiv* **2019**, arXiv:1912.06059.
102. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305. Available online: <https://www.researchgate.net/journal/Journal-of-Machine-Learning-Research-1532-4435> (accessed on 2 August 2022).
103. Larochelle, H.; Erhan, D.; Courville, A.; Bergstra, J.; Bengio, Y. An empirical evaluation of deep architectures on problems with many factors of variation. In Proceedings of the Twenty-Fourth International Conference on Machine Learning (ICML'07), Corvallis, OR, USA, 20–24 June 2007; Ghahramani, Z., Ed.; pp. 473–480. [[CrossRef](#)]
104. Probst, P.; Boulesteix, A.L.; Bischl, B. Tunability: Importance of hyperparameters of machine learning algorithms. *J. Mach. Learn. Res.* **2019**, *20*, 1934–1965. Available online: <https://jmlr.org/papers/volume20/18-444/18-444.pdf> (accessed on 1 September 2022).
105. Boehmke, B.; Greenwell, B.M. *Chapter 12: Gradient Boosting, Hands-On Machine Learning with R*, 1st ed.; Chapman and Hall/CRC: London, UK, 2019. [[CrossRef](#)]
106. Wu, B.; Zhu, W.; Yan, N.; Xing, Q.; Xu, J.; Ma, Z.; Wang, L. Regional actual evapotranspiration estimation with land and meteorological variables derived from multi-source satellite data. *Remote Sens.* **2020**, *12*, 332. [[CrossRef](#)]
107. Courault, D.; Seguin, B.; Olioso, A. Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches. *Irrig. Drain. Syst.* **2005**, *19*, 223–249. [[CrossRef](#)]
108. Cristóbal, J.; Jiménez-Muñoz, J.C.; Prakash, A.; Mattar, C.; Skoković, D.; Sobrino, J.A. An improved single-channel method to retrieve land surface temperature from the Landsat-8 thermal band. *Remote Sens.* **2018**, *10*, 431. [[CrossRef](#)]
109. García-Santos, V.; Sánchez, J.M.; Cuxart, J. Evapotranspiration Acquired with Remote Sensing Thermal-Based Algorithms: A State-of-the-Art Review. *Remote Sens.* **2022**, *14*, 3440. [[CrossRef](#)]
110. Kim, Y.; Jackson, T.; Bindlish, R.; Lee, H.; Hong, S. Radar vegetation index for estimating the vegetation water content of rice and soybean. *IEEE Geosci. Remote Sens. Lett.* **2011**, *9*, 564–568. [[CrossRef](#)]
111. Szigarski, C.; Jagdhuber, T.; Baur, M.; Thiel, C.; Urbazaev, M.; Parrens, M.; Entekhabi, D. Analysis of the radar vegetation index and assessment of potential for improvement. In Proceedings of the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8143–8146. [[CrossRef](#)]
112. Yan, N.; Tian, F.; Wu, B.; Zhu, W.; Yu, M. Spatiotemporal analysis of actual evapotranspiration and its causes in the Hai Basin. *Remote Sens.* **2018**, *10*, 332. [[CrossRef](#)]
113. Ma, Z.; Yan, N.; Wu, B.; Stein, A.; Zhu, W.; Zeng, H. Variation in actual evapotranspiration following changes in climate and vegetation cover during an ecological restoration period (2000–2015) in the Loess Plateau, China. *Sci. Total Environ.* **2019**, *689*, 534–545. [[CrossRef](#)] [[PubMed](#)]
114. Suzuki, R.; Masuda, K. Interannual co-variability found in evapotranspiration and satellite-derived vegetation indices over northern Asia. *J. Meteorol. Soc. Jpn.* **2004**, *82*, 1233–1241. [[CrossRef](#)]
115. Li, S.-G.; Eugster, W.; Asanuma, J.; Kotani, A.; Davaa, G.; Oyunbaatar, D.; Sugita, M. Energy partitioning and its biophysical controls above a grazing steppe in central Mongolia. *Agric. For. Meteorol.* **2006**, *137*, 89–106. [[CrossRef](#)]
116. Nzioka, J.M.; Njeri, J.K.; Karanja, F.K.; Manene, M.M. On the Relationship between Satellite-based Evapotranspiration and Normalized Difference Vegetation Index, Case Study: Narok County of Kenya. *Afr. J. Phys. Sci.* **2014**, *1*, 2313–3317. Available online: <https://core.ac.uk/download/pdf/236173325.pdf> (accessed on 2 October 2022).
117. Glen, A.S.; Dickman, C.R.; Soule, M.E.; Mackey, B.G. Evaluating the role of the dingo as a trophic regulator in Australian ecosystems. *Austral Ecol.* **2007**, *32*, 492–501. [[CrossRef](#)]
118. Muir, C.; Southworth, J.; Khatami, R.; Herrero, H.; Akyapı, B. Vegetation Dynamics and Climatological Drivers in Ethiopia at the Turn of the Century. *Remote Sens.* **2021**, *13*, 3267. [[CrossRef](#)]