

Article

Optimal Solar Zenith Angle Definition for Combined Landsat-8 and Sentinel-2A/2B Data Angular Normalization Using Machine Learning Methods

Jian Li *  and Baozhang Chen

School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China; baozhang.chen@nuist.edu.cn

* Correspondence: jian.li@nuist.edu.cn

Abstract: Data from Landsat-8 and Sentinel-2A/2B are often combined for terrestrial monitoring because of their similar spectral bands. The bidirectional reflectance distribution function (BRDF) effect has been observed in both Landsat-8 and Sentinel-2A/2B reflectance data. However, there is currently no definition of solar zenith angle (θ_{sz}) that is suitable for the normalization of the BRDF-adjusted reflectance from the three sensors' combined data. This paper describes the use of four machine learning (ML) models to predict a global θ_{sz} that is suitable for the normalization of bidirectional reflectance from the combined data in 2018. The observed θ_{sz} collected globally, and the three locations in the Democratic Republic of Congo (26.622°E, 0.356°N), Texas in the USA (99.406°W 30.751°N), and Finland (25.194°E, 61.653°N), are chosen to compare the performance of the ML models. At a global scale, the ML models of Support Vector Regression (SVR), Multi-Layer Perception (MLP), and Gaussian Process Regression (GPR) exhibit comparably good performance to that of polynomial regression, considering center latitude as the input to predict the global θ_{sz} . GPR achieves the best overall performance considering the center latitude and acquisition time as inputs, with a root mean square error (RMSE) of 1.390°, a mean absolute error (MAE) of 0.689°, and a coefficient of determination (R^2) of 0.994. SVR shows an RMSE of 1.396°, an MAE of 0.638°, and an R^2 of 0.994, following GPR. For a specific location, the SVR and GPR models have higher accuracy than the polynomial regression, with GPR exhibiting the best performance, when center latitude and acquisition time are considered as inputs. GPR is recommended for predicting the global θ_{sz} using the three sensors' combined data.



Citation: Li, J.; Chen, B. Optimal Solar Zenith Angle Definition for Combined Landsat-8 and Sentinel-2A/2B Data Angular Normalization Using Machine Learning Methods. *Remote Sens.* **2021**, *13*, 2598. <https://doi.org/10.3390/rs13132598>

Academic Editor: David M. Johnson

Received: 21 May 2021

Accepted: 29 June 2021

Published: 2 July 2021

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



Copyright: © 2021 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

The polar orbit satellite Landsat-8, launched by National Aeronautics and Space Administration (NASA) [1,2] and the Sentinel-2A and 2B satellites, launched by ESA [3], have similar spectral bands. Together, these three satellites provide 10–30 m moderate spatial resolution multi-spectral global coverage. The combination of the three satellites presents a new solution for global moderate-resolution landcover monitoring. Compared with a single satellite, the combination of the three satellites, taking advantage of their complementary revisit interval patterns, provides a 2.9-day global median average revisit interval [4,5]. This would benefit numerous remote sensing applications, such as deforestation [6], fire monitoring [7], agriculture dynamics [8], and ice velocity detection [9].

Both Landsat-8 and Sentinel-2A/2B have sun-synchronous polar orbits. Their view angles are ±10.3° (Sentinel-2) and ±7.5° (Landsat-8) from the nadir view when acquiring observations, resulting in non-Lambertian surface directional reflectance effects. The magnitude of these effects varies as a function of geometry (i.e., the relation between the sun, the object, and the sensor), and is usually described by the bidirectional reflectance distribution function (BRDF). The BRDF effects should be corrected to provide stable and

consistent satellite datasets. To eliminate the BRDF effects in the Landsat-8 and Sentinel-2 datasets, a global semi-empirical approach, parameterized by external BRDF model parameters derived from the MODIS BRDF data product, has been used to normalize the Landsat-8 and Sentinel-2 data reflectance to the nadir view (0° view zenith) [10].

The solar zenith (θ_{sz}) used for normalization was defined in the following criteria: θ_{sz} can be modelled for any date and location, and the difference between the observed θ_{sz} and θ_{sz} used for normalization is minimized [11,12]. According to the above criteria, there are two ways to define solar zenith (θ_{sz}), namely defining θ_{sz} with respect to the scene acquisition center latitude or defining θ_{sz} with respect to the scene acquisition center latitude and scene acquisition time. A latitudinally fixed θ_{sz} was defined by a degree-six polynomial fitted on the basis of Landsat-8 data to retrieve constant latitudinal θ_{sz} , using latitude as an input variable for the combined Landsat-8 and Sentinel-2A/2B reflectance data [13]. To define θ_{sz} with respect to the scene acquisition center latitude and scene acquisition time, a specific normalized θ_{sz} was defined based on a polynomial model to derive the local overpass time against latitude, and a physical astronomical model was used to normalize the θ_{sz} to a specific angle and minimize the angular difference between the observed and normalized θ_{sz} for the Landsat-5 and Landsat-7 datasets [12]. However, for both solutions, the approximation of polynomial regression is usually limited by the explicit relationships they rely on [14]. Recent research has proven that machine learning (ML) models can directly learn parameters and functional forms for complex nonlinear relationships from data, which is different from parametric regression [15]. To date, the global θ_{sz} distribution by the combination of the Landsat-8 and Sentinel-2A/2B has not been investigated yet and there is no global θ_{sz} definition that is suitable for normalizing the BRDF time series given by the combination of the Landsat-8 and Sentinel-2A/2B reflectance data considering these three data sets together.

In this study, the global distribution of θ_{sz} was first quantified for Landsat-8 and Sentinel-2A/2B, and, in a different manner from previous research, ML models were explored to predict a θ_{sz} that is suitable for normalizing the BRDF time series from the combination of Landsat-8 and Sentinel-2A/2B. Specifically, global metadata records from 1 January to 31 December in 2018 for Landsat-8 and Sentinel-2A/2B were used to quantify the θ_{sz} variation between these three satellites. ML regression methods, namely regularized linear regression (RLR), support vector regression (SVR), Gaussian process regression (GPR), and a multi-layer perceptron (MLP), were then used to predict a suitable global θ_{sz} with respect to latitude and with respect to both latitude and acquisition time. Finally, θ_{sz} suitable for normalizing the BRDF effect for the combination of Landsat-8 and Sentinel-2A/2B were predicted using the suggested machine learning model. To compare the estimation accuracy between the ML models, test samples are compared with global datasets as well as the observed θ_{sz} collected from the Democratic Republic of Congo (26.622° E, 0.356° N), Texas, USA (99.406° W 30.751° N), and Finland (25.194° E, 61.653° N) in 2018. The regression accuracy of each model is evaluated in terms of the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) between the observed and normalized. Additionally, the ML model suitable for the estimation of θ_{sz} for the combination of Landsat-8 and Sentinel-2A/2B is recommended.

Section 2 describes the data used in this study. Sections 3 and 4 then introduce the methods used to analyze the data and present the results obtained, respectively. Section 5 discusses the results, before Section 6 concludes by stating several implications and recommendations for global time-series applications for the three sensors combined.

2. Data

2.1. Satellite Remote Sensing Configurations

Both Landsat-8 and Sentinel-2A/2B were launched into polar sun-synchronous orbits. Landsat-8 has an altitude of 705 km and an incline of 98.22° . The scanning angle for Landsat-8 is $\pm 7.5^\circ$ and the swath width is 185 km. Landsat-8 revisits the same location every 16 days and crosses the equator at $10:00 \pm 15$ min [16]. The Sentinel-2A and Sentinel-

2B satellites orbit at an altitude of 786 km and have an incline of 98.62°. The scanning angle for both the satellites is ±10.3° and the swath width is 290 km. Both sensors revisit the same location every 10 days, giving a combined revisit interval of five days. Sentinel-2 has an equatorial crossing time of 10:30 [17].

2.2. Global θ_{sz} Metadata Records for Landsat-8 and Sentinel-2A/2B

The Landsat-8 observation metadata records were bulk downloaded from the United States Geological Survey (USGS) Landsat archive metadata database [18]. Metadata records with approximate scene dimensions of 185 km × 180 km are defined in the Worldwide Reference System (WRS-2) [19]. The metadata records acquired from January 1–December 31, 2018, were extracted. Only the images acquired during the daytime were used in this study.

The following information in each Landsat-8 metadata record was used: “sceneStart,” “sceneStop,” “sunEle,” “ce_x,” and “ce_y.” The scene center acquisition time (Act) for each record was computed as the average of the “sceneStart” and “sceneStop” times, and the scene center θ_{sz} was derived as 90°—“sunEle.” The scene center latitude (Lat) and scene center longitude (Lon) coordinates were defined as “ce_x” and “ce_y”, respectively, in each metadata record.

The Sentinel-2A/2B metadata records were downloaded from the USGS Earth explorer [20]. Each metadata record is defined in a fixed 109 km × 109 km tile projected in the Universal Transverse Mercator (UTM) map projection [3,21]. The Sentinel-2 projected tiles are stored as Standard Archive Format for Europe (SAFE) files [17] and cut along orbit swaths. All the metadata records acquired from 1 January–31 December, 2018, were extracted. Only the images acquired in descending orbits were used in this study.

For each Sentinel-2A/2B metadata record, the following information was used: “Acquisition Start Date,” “Acquisition End Date,” “Sun Zenith Angle Mean,” “Center Latitude dec,” and “Center Longitude dec.” The Act for each Sentinel-2 metadata record was computed as the average of the “Acquisition Start Date” and “Acquisition End Date,” and the “Sun Zenith Angle Mean” was defined as the θ_{sz} value for each scene. “Center Latitude dec” and “Center Longitude dec” were used as the Lat and Lon for each metadata record, respectively.

There are duplicated observations for Landsat-8 and Sentinel-2A/2B during 2018 because the satellites have a repeat circle of 16 days (Landsat) and 10 days (Sentinel). Large amounts of redundant data pose a challenge to the training of ML models. A large redundant dataset requires more memory to fit the data, and more consumption time is required to train and to extract useful features from the data. An increase in the size of the redundant data volume would decrease the predictive ability and effectiveness of a machine learning algorithm. [22,23]. Thus, a global dataset for the study was established by selecting every tenth line of the Landsat-8 and Sentinel-2A/2B datasets collected during 2018. The three datasets were then merged together from top to bottom. In total, 361,826 metadata records were used in this study, with 25,737 from Landsat-8, 161,799 from Sentinel-2A, and 174,290 from Sentinel-2B. As θ_{sz} varies smoothly in space and time [12], it constitutes a good representation for the global θ_{sz} dataset for the combination of Landsat-8 and Sentinel-2A-2B in 2018. The number of metadata records for Sentinel-2 are larger than that of Landsat-8. The proposed ML model is first trained from the combination of the three sensors and then predicts θ_{sz} ; thus, this, not even available, dataset would infect the derived θ_{sz} values slightly and become more inclined to Sentinel-2 data.

2.3. Local θ_{sz} Metadata Records for Landsat-8 and Sentinel-2A/2B

To examine the ML models’ performance on local θ_{sz} , observations collected by Landsat-8 and Sentinel-2A/2B over three locations during 2018 were selected, namely, northeast of the Democratic Republic of Congo (26.622° E 0.356° N, over path/row 175/60 for Landsat-8 and tile number T35NMA for Sentinel-2); TX, USA (99.406° W 30.751° N, over path/row 28/39 for Landsat-8 and tile number T14RMV for Sentinel-2); and south

of Finland (25.194° E 61.653° N, over path/row 188/17 or 189/17 for Landsat-8 and tile number T35NMA for Sentinel-2) (henceforth referred to, for brevity, as Congo, Texas, and Finland). These three locations were selected as they have a different latitude/longitude but span a large latitudinal range. Locations with different latitudes are used to examine the ML models' prediction performance because θ_{sz} will change with latitude for a given local time.

3. Methodology

Following the optimal θ_{sz} definition criteria described in [12], Lat, Lat and Act, and Lat and Lon and Act were used as the input variables into different ML models to predict θ_{sz} respectively, ensuring that θ_{sz} can be modeled at any location and date to produce consistent BRDF normalized reflectance data for the combination of Landsat-8 and Sentinel-2A -2B. This requirement enables a large volume of data auto processing for the three sensors combined. For generating multi-temporal composite produce, a consistent θ_{sz} definition is representative of the same day in the compositing period, if it can be modelled at any location and any date [24]. Statistical metrics were used to evaluate the fitting performance between the predicted θ_{sz} values given by each model and the testing datasets, ensuring that the difference between the observed and normalized θ_{sz} was minimized, because it will introduce unreliability into the semi-empirical BRDF model to normalize reflectance data at the θ_{sz} that are different from the θ_{sz} used to invert BRDF model parameters [11]. These statistical metrics are the root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2).

The process flow of the θ_{sz} retrieval based on ML models is illustrated in Figure 1, and the process can be summarized as preprocessing, training, and prediction. In the preprocessing step, the input variables, e.g., Lat, Lon, and Act, were normalized and scaled, in which values were shifted and rescaled to ensure that they ended up with a value between 0 and 1. After this the normalized variables were fitted for the ML models as inputs. In addition, the global θ_{sz} datasets were split randomly into training and testing datasets, with 70% of the data for training and 30% for testing in each set. In the training process, ten-fold cross-validation [25] was used to optimize the hyper-parameters of each model and check the overall performance of each regression method. In the cross-validation, the datasets were randomly and equally dispatched into k groups ($k = 10$). In each validation process, $(k - 1)$ groups were used as the training instances and one group was treated as the test instance. An evaluation score was obtained for the model in each validation process, giving a total of k evaluation scores after looping for every test instance. Finally, the hyper-parameters for each model were optimized in the cross-validation process using the average of the evaluation score. In the prediction process, the predicted θ_{sz} was estimated by each of the ML models using the best hyper-parameters obtained from the cross-validation process. Particularly, polynomial regression and four ML models were used to predict θ_{sz} for the normalization.

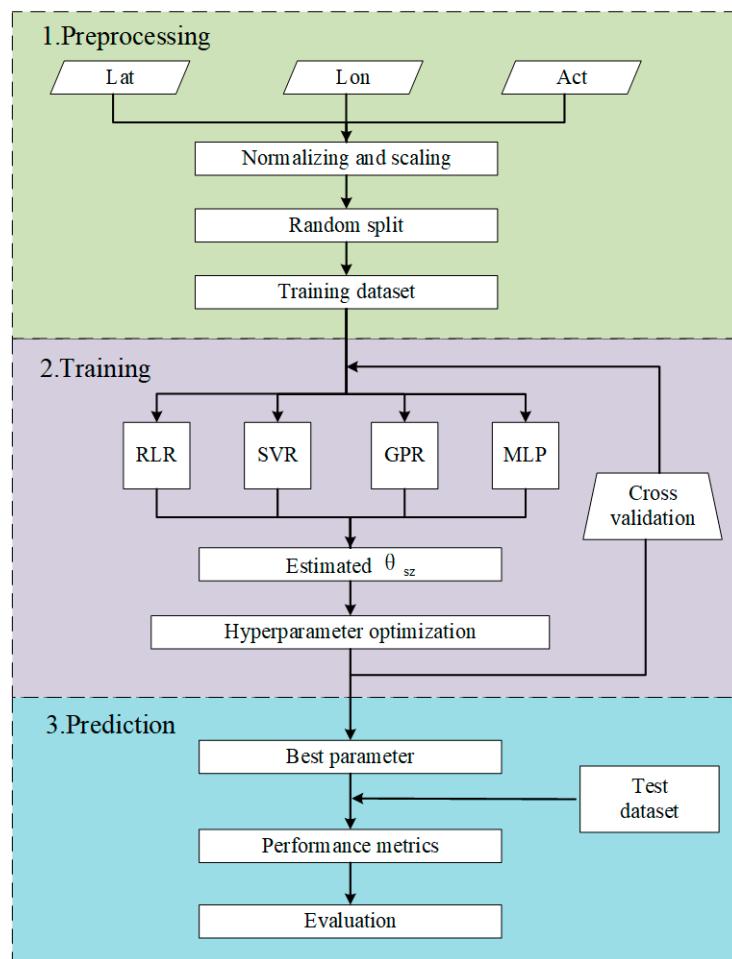


Figure 1. The flow chart of the proposed ML models to estimate global θ_{sz} of 2018 for the combination of Landsat-8, Sentinel-2A, and Sentinel-2B.

3.1. Polynomial Regression Model

For reference, the single variable Lat was used as the input to a 6th-degree polynomial to retrieve constant latitudinal θ_{sz} . In this study, the following 6th-degree polynomial regression was used as a benchmark model to compare with ML models, for both the latitudinal fix θ_{sz} model and the physical astronomical model, described in Section 1, built upon the following polynomial regression [12,13]:

$$y = p_0x^6 + p_1x^5 + p_2x^4 + p_3x^3 + p_4x^2 + p_5x + p_6 \quad (1)$$

where x is the Lat for each metadata record and y is the predicted θ_{sz} from the polynomial regression.

3.2. ML Regression Models

ML has been successfully applied in many regression [14,15,26–28] and classification [29–31] tasks. Despite the good fitting performance of parametric models, i.e., polynomial regression, the approximation of these models is usually limited by the explicit relationships they rely on. Unlike parametric regression, ML models learn the parameters and functional form from the data. In this study, four ML models were used to predict θ_{sz} for the combination of Landsat-8 and Sentinel-2A/2B.

3.2.1. Regularized Linear Regression

RLR was used to predict θ_{sz} from different input variables. The output results were assumed to be a linear sum according to the weighted input variables, $x = [x_1, x_2, \dots, x_n]^T$, then, $\theta_{sz} = x^T w$. The maximum likelihood of the output results was obtained by minimizing the squared errors of the weights. Additionally, the weights $w = [w_1, w_2, \dots, w_n]^T$ can be estimated by least squares minimization.

3.2.2. Support Vector Regression

SVR is based on support vector machines, which are nonlinear ML models for classification and regression [32,33]. In the SVR model, a linear model was used to estimate the output θ_{sz} after transforming the input variables into a high-dimensional space. A kernel function was used to transform the input data to ensure that a linear model can be fitted. After mapping the training data into hyperspace, Vapnik's ε -insensitive cost function was used to optimize the parameters of the linear model, as follows:

$$L_\varepsilon(e) = C \max(0, |e| - \varepsilon), \quad C > 0 \quad (2)$$

where an error $e = y - \hat{y}$ within a ε -defined margin is ignored, while the influence of the samples outside this linear margin is penalized. The parameter C regularizes the trade-off between model complexity and error frequency. In this work, we use the SVR implementation of the Python package Scikit-Learn [34], which is based on LIBSVM [35]. There are the following two key parameters: the margin distance, defined by ε , is used to evaluate the sample data that fall outside some boundaries, and the C factor is used to balance the weighting between complexity and accuracy. As well as the ε distance and C factor, the kernel type must be selected; a radial basis function kernel was used. To optimize the hyperparameters, we considered values of 1, 10, and 100 for the C factor and values of 0.001, 0.01, 0.1, and 0.2 for the ε distance in the experiments.

3.2.3. Gaussian Process Regression

Gaussian process models [22,36] are probabilistic ML models designed for regression and classification problems. They offer powerful regression ability in vegetation biophysical parameter retrieval [37,38], and landcover classification [39]. A GPR model assumes that the observed θ_{sz} is a function of the input variables using a joint Gaussian distribution of the available observations with zero mean and covariance matrix K , as follows:

$$\begin{bmatrix} y \\ \hat{y} \end{bmatrix} \sim N\left(0, \begin{bmatrix} K + \sigma^2 I_n & k_* \\ k_*^T & k_{**} + \sigma^2 \end{bmatrix}\right) \quad (3)$$

where k_* is the covariance between the training inputs and the testing inputs, k_{**} is the autocovariance for the testing inputs, and $K + \sigma^2 I_n$ defines the covariance matrix of the noise for training inputs. The GPR model uses Bayes' principle to define a posterior distribution and likelihood function over the output predicted \hat{y} given the new input and the training dataset. The mean value of the posterior distribution is used as the prediction, and the confidence intervals of the prediction are derived from the likelihood function. In our implementation, GPR implementation of the Python package Scikit-Learn was used to derive the predicted θ_{sz} . When learning a GPR model, some parameters related to the covariance or kernel functions also need to be specified. The RBF kernel (Radial-basis function kernel) was chosen, which is a stationary squared exponential kernel. The hyperparameters to be tuned for RBF kernel include magnitude, characteristic length, and noise variance. Additionally, the maximum-likelihood method was applied for parameter tuning in this study. Compared with SVR, GPR not only gives a prediction mean for the new observation, but also provides a full probabilistic posterior distribution.

3.2.4. Multi-Layer Perception

Multi-layer perception network (MLP) is a classical type of feedforward artificial neural network and has been widely used in different fields of remote sensing [15,27]. There are the following three parts in the MLP: an input layer, a hidden layer, and an output layer. A non-linear activation function was used in each node of neuron in the hidden and output layers. Additionally, there can be multiple layers of neurons in each hidden layer. The backpropagation iterative optimization process is used to adjust the connection weight between the hidden layers of the neurons by minimizing the estimation error [40]. The MLP model used a representation of neurons to achieve highly precise estimates of the nonlinear relationship between the input variable and the output result. Since the network is fully connected, each node in one layer connects to every node in the next layer with a certain weight. Therefore, the output c of each neuron is as follows:

$$c = \varphi(\sum w_i a_i + b_i) \quad (4)$$

where a_i and w_i are the input and weight for the neuron, respectively, b_i is the bias of the neuron, and φ is the activation function used. In our implementation, MLP Regressor in the Python Scikit-Learn package was used. The parameters for MLP include the number of hidden layers, number of neurons for each hidden layer, optimizer, and activation function. The ReLU activation function was used, and the Adam optimizer was chosen. To find the optimum parameters, the number of hidden layers is tested with 1, 2 and 3, and the number of neurons for each hidden layer is tested with (200), (200, 100) and (200, 140, 70).

4. Results

4.1. Global θ_{sz} Distribution and Variations for Landsat-8 and Sentinel-2A/2B

Figure 2 shows the observed θ_{sz} plotted against the Lat and Lon for the global data in 2018 by Landsat-8, Sentinel-2A, Sentinel-2B, and the three sensors combined. For all three sensors, the variation of θ_{sz} with latitude is apparent. As can be seen in Figure 2, θ_{sz} increases from the equator area to the polar regions. Compared with Sentinel-2, Landsat-8 acquires more observations around the two polar regions; this is because of the different satellite imagery acquisition strategies [17,41]. There are few observed θ_{sz} collected over the ocean, an abrupt change for θ_{sz} was observed in the region at about 30° N, −30° E. This is because there is an interval of a few months between the observations. Additionally, for each grid point in the graph, if there were more than one θ_{sz} sensed, the first was chosen to plot.

Across the whole year, θ_{sz} varies because the satellite local overpass time changes with the latitude and because of seasonal changes in the position of the sun. Figure 3 illustrates θ_{sz} as a function of Lat and Act for Landsat-8, Sentinel-2A, Sentinel-2B, and the three sensors combined in 2018. As both Landsat-8 and Sentinel-2 have a polar orbit, a similar θ_{sz} variation with respect to the Lat and Act for Landsat-8 and Sentinel-2 can be seen from Figure 3.

Table 1 summarizes the global mean average, median average, minimum, maximum, and standard deviation of θ_{sz} for each sensor and the three sensors combined in 2018. Specifically, the global maximum and minimum θ_{sz} for Landsat-8 are 89.99° and 20.96°, with a mean average of 49.97° and standard deviation of 18.58°. Compared with Landsat-8, Sentinel-2 gives a lower θ_{sz} , with a maximum, minimum, and mean of 83.38, 14.74, and 43.89° for Sentinel-2A and 88.87, 14.76, and 44.89° for Sentinel-2B. This is because the average equatorial crossing time of Landsat-8 is 30 min earlier than Sentinel-2 (10:00 for Landsat-8, 10:30 for Sentinel-2). There are slight differences between the θ_{sz} of Sentinel-2A and Sentinel-2B because they have a phase delay of 180° between them [3]. The global mean average difference in θ_{sz} between Landsat-8 and Sentinel-2A/2B is 6.09°/5.09°.

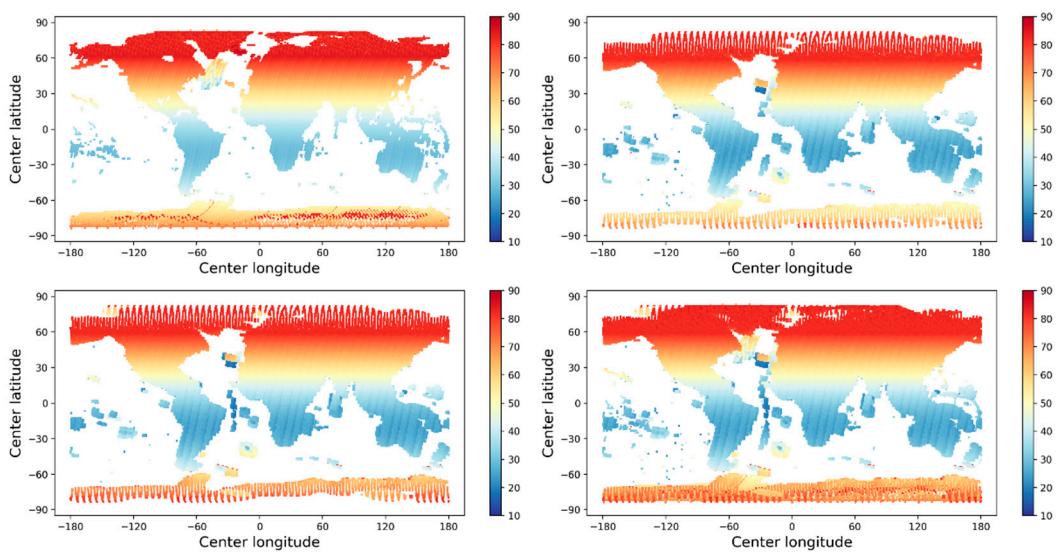


Figure 2. Observed θ_{sz} plotted against Lat and Lon in 2018 for Landsat-8 (**top left**), Sentinel-2A (**top right**), Sentinel-2B (**bottom left**), and all three sensors combined (**bottom right**).

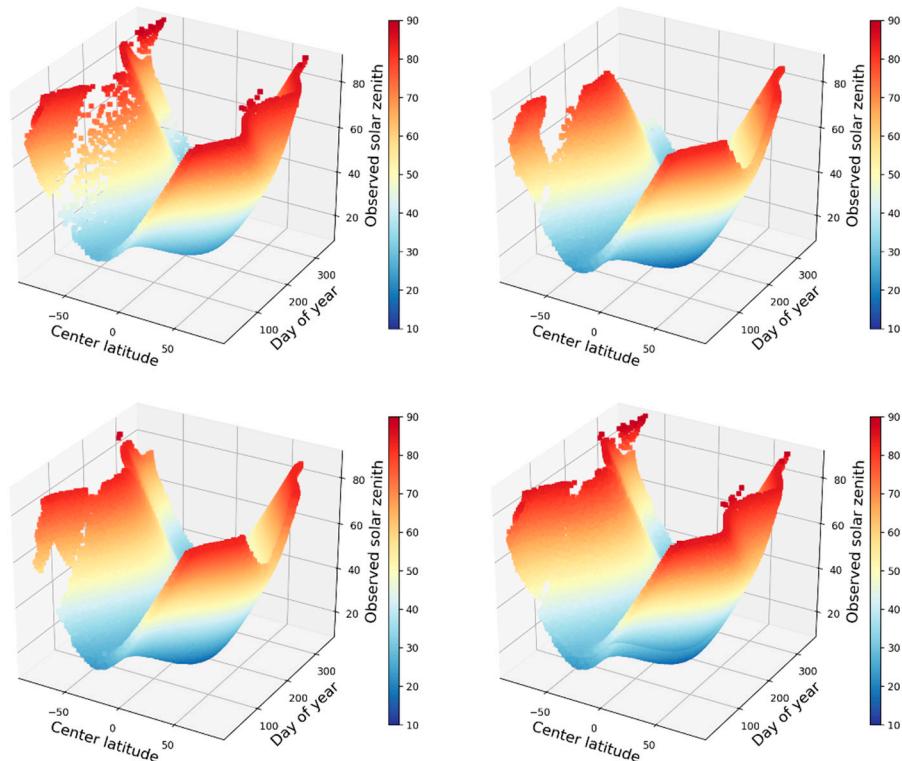


Figure 3. Observed θ_{sz} plotted against Lat and Act in 2018 for Landsat-8 (**top left**), Sentinel-2A (**top right**), Sentinel-2B (**bottom left**), and all three sensors combined (**bottom right**).

Table 1. Summary statistics of observed θ_{sz} for the global data of 2018 for Landsat-8, Sentinel-2A, Sentinel-2B, and all three sensors combined. Results are given to three decimal places.

	Mean	Median	Standard Deviation	Maximum	Minimum
Landsat-8	49.973°	48.145°	18.581°	89.985°	20.963°
Sentinel-2A	43.890°	41.573°	17.951°	83.377°	14.739°
Sentinel-2B	44.885°	42.665°	18.346°	88.872°	14.759°
Three sensors combined	44.802°	42.520°	18.250°	89.985°	14.739°

Figure 4 shows a boxplot of the observed θ_{sz} distribution for each of the sensor and the three sensors combined. The dotted orange lines show the median average for each θ_{sz} data set. It is clear that the medium of θ_{sz} for Landat-8 is higher than for Sentinel-2 and the three sensors combined due to the different equatorial crossing times of Landsat-8 and Sentinel-2. The medium value of θ_{sz} for Sentinel-2 and the three sensors combined are on the same level because the number of observed θ_{sz} for Sentinel-2 accounts for a large proportion of the three sensors combined.

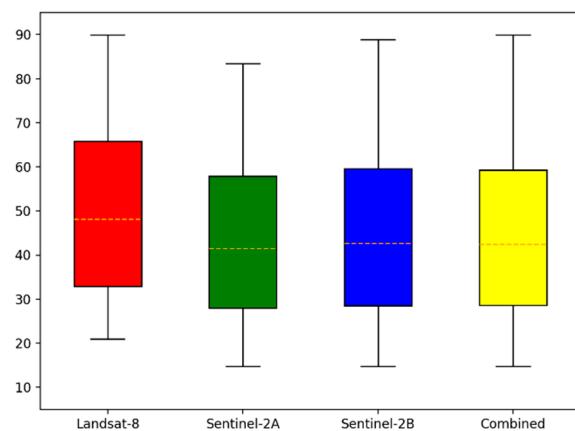


Figure 4. Boxplot of θ_{sz} for Landsat-8 (red), Sentinel-2A (green), Sentinel-2B (blue), and all three sensors combined (yellow).

4.2. Performance of ML Models for Global θ_{sz} Prediction

Table 2 summarizes the accuracy metrics, with respect to the observed θ_{sz} for the three sensors combined, for all five methods when using different combinations of input variables. When only using one input variable (Lat or Act) to predict θ_{sz} , the SVR and MLP models achieve a comparable performance to that of polynomial regression. The GPR model is slightly better than polynomial regression, with $R^2 = 0.526$, MAE = 10.590°, and RMSE = 12.463° with Lat as the input variable, compared with values of 0.525, 10.597°, and 12.473°, respectively, for polynomial regression. When more than one variable is used for the fitting process, the nonlinear ML models exhibit a significant improvement over the accuracy of the linear RLR model, as can be seen for the combination of Lat and Act, and Lat and Lon and Act in Table 2. Considering Lat and Act as inputs, all three nonlinear ML models achieve acceptable results. Specifically, GPR gives the best results in terms of R^2 (0.994), MAE (0.689°), and RMSE (1.390°), followed by SVR with values of 0.994, 0.638°, and 1.396°, respectively. Note that using three variables as inputs (Lat and Lon and Act) does not improve the global-scale prediction results of the GPR and SVR models compared to the use of two input variables (Lat and Act). This is consistent with the results discussed in Section 4.1, as Lat and Act are the two main factors determining the variation of θ_{sz} .

Table 2. Regression accuracy obtained with polynomial fitting and ML models on the observed θ_{sz} for global data over all 12 months of 2018 for the three sensors combined. Results are given to three decimal places.

Input	Metric	Polyfit	RLR	SVR	GPR	MLP
Lat	R^2	0.525	0.067	0.516	0.526	0.525
	MAE	10.597°	14.720°	10.581°	10.590°	10.619°
	RMSE	12.473°	17.484°	12.597°	12.463°	12.470°
Act	R^2	0.114	0.000	0.113	0.116	0.113
	MAE	14.633°	15.512°	14.593°	14.620°	14.635°
	RMSE	17.038°	18.101°	17.044°	17.023°	17.048°
Lat and Act	R^2	-	0.067	0.994	0.994	0.993
	MAE	-	14.720°	0.638°	0.689°	0.873°
	RMSE	-	17.484°	1.396°	1.390°	1.504°
Lat and Lon	R^2	-	0.070	0.993	0.994	0.992
Lon and Act	MAE	-	14.692°	0.711°	0.691°	1.052°
	RMSE	-	17.454°	1.489°	1.391°	1.598°

Figure 5 compares the regression performance of different models using Lat (left) and Act (right). The dotted points denote the data used for training (yellow) and testing (gray). The polynomial regression parameterized single input, colored in blue, is plotted as the reference line. The plot on the left shows a bowl-shaped observed θ_{sz} against Lat for all three nonlinear ML models, representing the polar solar-synchronous orbit geometry. The RLR model (orange) is a straight line, which does not fit much of the data along the center latitude in the x-axis. The GPR, MLP, and SVR models exhibit a regular regression fitting line of θ_{sz} against latitude, with GPR (red) and SVR (green) providing a better fit than MLP (purple). Some differences between GPR and SVR appear at latitudes above 60° north and below 60° south, where GPR gives a better fit to the θ_{sz} in high-latitude areas, with an RMSE of 12.463° (GPR) compared to 12.597° (SVR). The gray area illustrates the GPR θ_{sz} fitting for the 95% confidence interval. There are few data located around 60° S for the observed θ_{sz} , because around this latitude most parts of the earth are occupied by ocean and a very small proportion of land (Figure 2). Both of the two ML models predicted a smooth regression fitting line in this area, reflecting the ML models' prediction ability when the data were sparse or missed. The right-hand plot compares the regression performance of the ML models against the polynomial regression for θ_{sz} against Act. The RLR model (orange) still follows a straight line, whereas all three nonlinear ML models give a sinusoidal shape. The performance of these three models is quite similar, with an R^2 around 0.1, indicating a relatively weak relationship between θ_{sz} and Act.

Figure 6 shows the results of an importance test of the input variables with respect to the predicted θ_{sz} for the four ML regression models. The importance of each input variable in a prediction was measured using the MAE, which was computed by leaving out each input variable in turn and performing a test, and then computing the average value. After each test, a stable variable importance was established. A higher MAE value indicates that the input variable that has been left out has more importance to the model. For the RLR model, all three input variables have nearly the same importance to the fitting model because of the linear form of the combination of input variables. For all three nonlinear ML models, Lat has a strong correlation to the predicted θ_{sz} , as shown in Figure 6. Act has a moderate correlation to the fitted θ_{sz} , while there is only a weak correlation between Lon and θ_{sz} . The results agree with our physical intuition regarding how each of the input variables contributes to the predicted θ_{sz} , as discussed in Section 4.1.

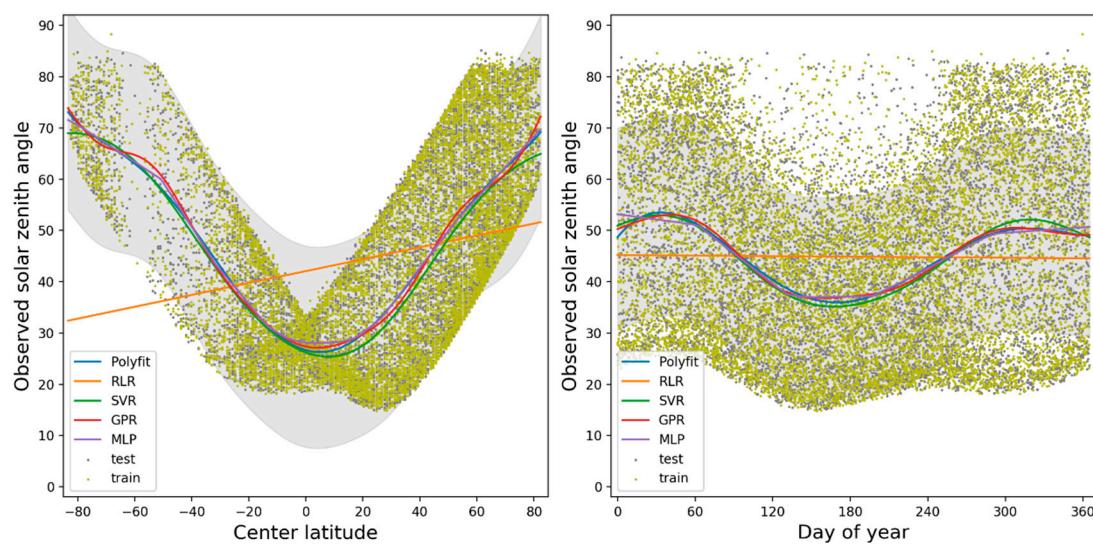


Figure 5. Observed θ_{sz} plotted against Lat (left) and Act (right) for all 12 months of 2018 for the three sensors combined. Comparison of the performance of the polynomial regression and ML models using the plotted data.

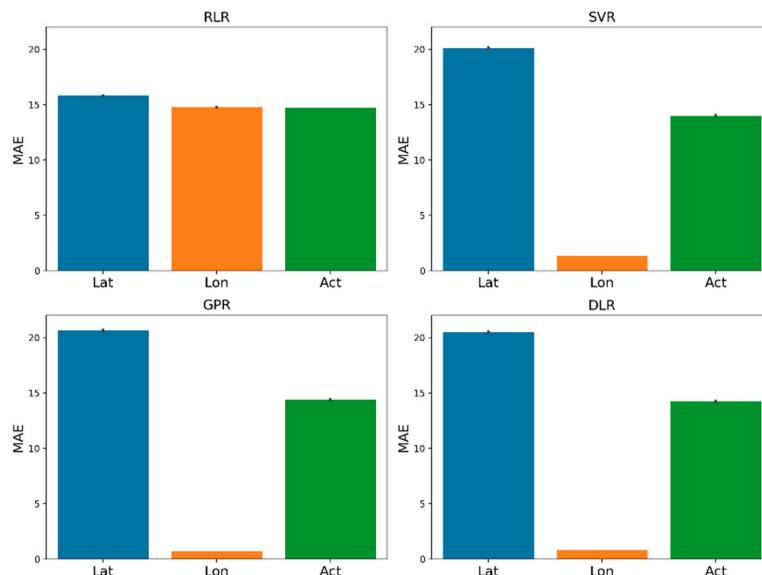


Figure 6. Input variable importance plots of θ_{sz} for the following four ML regression models: RLR, SVR, GPR, and MLP. The x-axis denotes the input variable, Lat, Lon, or Act, and the y-axis measures the difference in MAE calculated as the average from several tests when leaving one variable out.

Figure 7 shows a scatterplot of the observed θ_{sz} against the predicted θ_{sz} for the four ML models, taking Lat and Act as the input variables on a global scale. The RLR model presents the most disperse scatterplot, implying that relatively large errors could occur if a linear model was used to derive θ_{sz} . This is because complex nonlinear relationships exist between the input variables and θ_{sz} . It can be observed that the three nonlinear ML models produce a distribution of data around the 1:1 line (gray line). All of them achieve MAE values below 1° (SVR: 0.638°, GPR: 0.689°, and MLP: 0.873°), which is less than the global mean average difference of θ_{sz} between Landsat-8/Sentinel-2A (6.09°) and Landsat-8/Sentinel-2B (5.09°), implying that the ML models have an excellent ability to fit this nonlinear complex relationship. There are some outliers when θ_{sz} was larger than 70° for all the SVR, GPR, MLP models. This is because when polar-orbiting Landsat-8 and Sentinel-2 transit over polar regions, the observed θ_{sz} varied greatly over the area with

same Lat but different Lon [4] and also because Lat and Act were used as inputs. We do not include Lon to predict θ_{sz} in Figure 7.

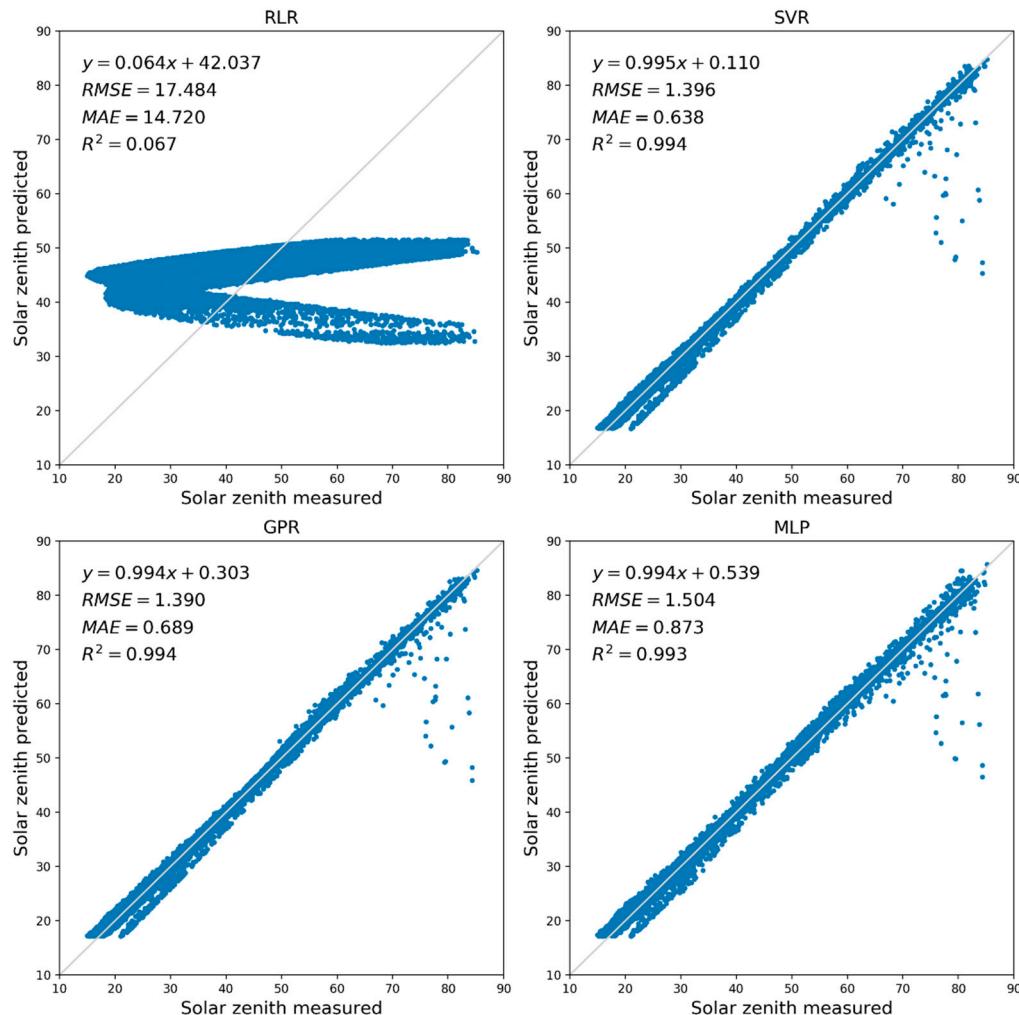


Figure 7. Scatterplot of observed θ_{sz} against predicted θ_{sz} given by RLR (top left), SVR (top right), GPR (bottom left), and MLP (bottom right) for the three sensors combined using Lat and Act as input variables in 2018.

Figure 8 shows the distribution of the predicted θ_{sz} using Lat and Act as the input variables for different ML models. The RLR model appears in the form of a fitted plane. The nonlinear fitted models have a very similar distribution to the observed θ_{sz} (see Figure 3), with smooth variations in space and time, indicating good θ_{sz} predictions for the three sensors combined.

As all three nonlinear ML models achieve good accuracy in terms of estimating the observed θ_{sz} for the three sensors combined at the global scale, the computation time and RAM consumption were also investigated. A desktop computer with an Intel Core i7-8700 3.20 GHz CPU and 16 GB RAM running Windows 10 was used as the computing environment. Table 3 summarizes the computation times and RAM consumption of the different models for θ_{sz} regression using Lat and Act as inputs. Compared with the other three nonlinear models, RLR has the lowest runtime and memory consumption (1.36 s and 501.7 MB, respectively). When the regression models become more complex, the computation time and RAM consumption increase (1747.27 s and 5118.5 MB for the GPR model, 7090.03 s and 4999.9 MB for the SVR model).

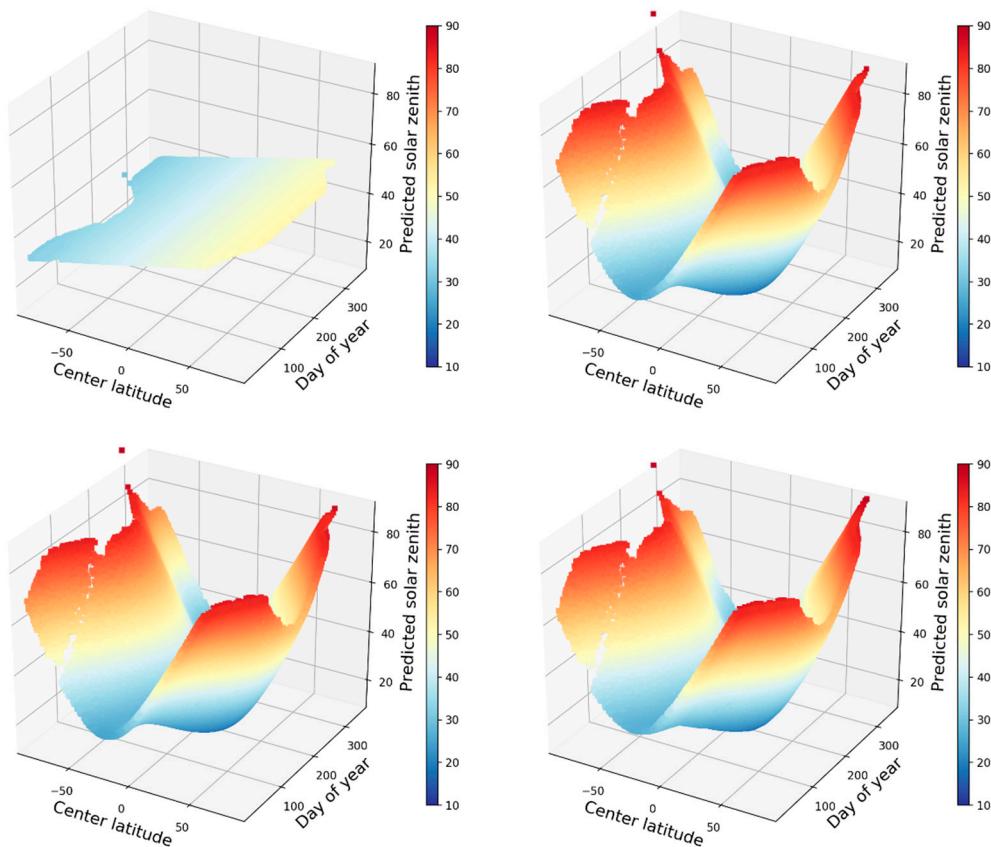


Figure 8. θ_{sz} predicted by RLR (**top left**), SVR (**top right**), GPR (**bottom left**), and MLP (**bottom right**) for the three sensors combined using Lat and Act as input variables in 2018.

Table 3. Comparison of computational time and RAM consumption in the training process using different ML models.

Model	Time(s)	RAM(MB)
RLR	1.36	501.7
SVM	7090.03	4999.9
GPR	1747.27	5118.5
MLP	2152.50	3512.2

4.3. Performance of ML Models for Local θ_{sz} Prediction

The left-hand side of Figures 9–11 shows θ_{sz} from the three sensors combined plotted as a function of Act for 2018 over the northeast of the Congo (Figure 9), Texas (Figure 10) and Finland (Figure 11). θ_{sz} against Lat was not shown, as there are nearly no variations for observed Lat over the three selected locations through 2018.

The variations in θ_{sz} from the three sensors combined appear to be sinusoidal for the Congo because the solar geometric position swings back and forth around the equator. In 2018, the mean average θ_{sz} values for Landsat-8 and Sentinel-2A/2B were 31.023, 26.977, and 27.204°, respectively. The maximum variation in θ_{sz} over the Congo in 2018 for the three sensors combined was 17.006°, with a maximum of 35.666° observed on June 23 from Landsat-8 and a minimum of 18.660° observed on October 1 from Sentinel-2B.

The θ_{sz} variations appear to be bowl-shaped over Texas (Figure 10, left), with the smallest θ_{sz} occurring in the summer when the sun is closest to being directly overhead at the time of satellite overpass in the northern hemisphere. Over 2018, the mean average θ_{sz} values for Landsat-8 and Sentinel-2A/2B were 38.013, 35.056, and 34.580°, respectively. The maximum variation in θ_{sz} for the three sensors combined over Texas in 2018 was 41.741°,

with a maximum of 57.757° from Landsat-8 on December 26 and a minimum of 16.016° from Sentinel-2B on June 12.

Finland (Figure 11, left) shows a similar bowl-shaped θ_{sz} variations to Texas, but spans even greater ranges due to the higher latitude. The mean average θ_{sz} values for Landsat-8 and Sentinel-2A/2B were 60.344, 56.505, and 56.932 $^\circ$, respectively, over 2018. The maximum variation in θ_{sz} for the three sensors combined over Finland in 2018 was 46.514 $^\circ$, with a maximum of 84.859 $^\circ$ from Landsat-8 on December 11 and a minimum of 38.345 $^\circ$ from Sentinel-2B on June 23.

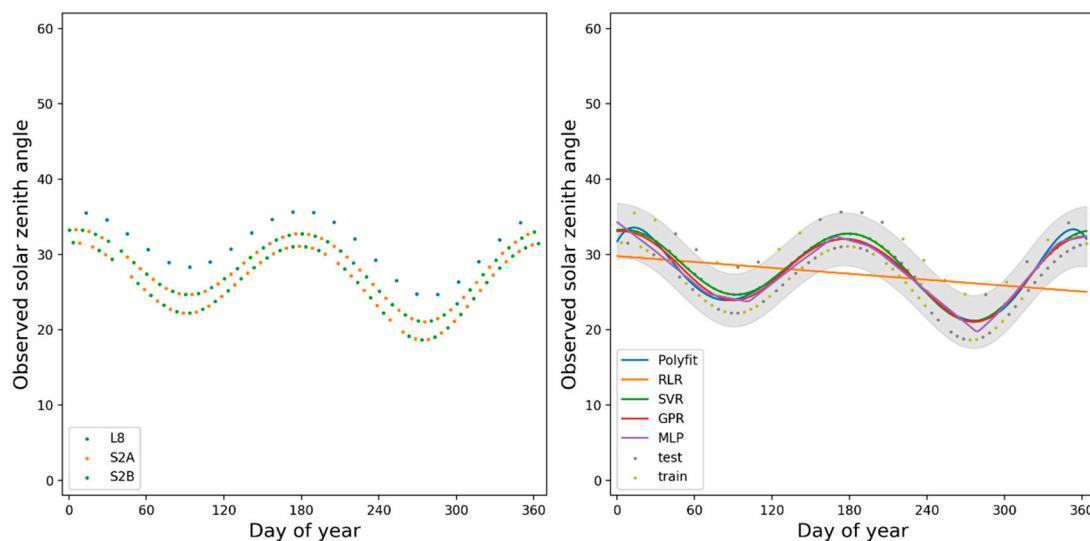


Figure 9. Observed θ_{sz} over the Congo (26.622° , 0.356°) plotted against Act (left) for all 12 months of 2018 for the three sensors combined. Comparison of the performance (right) for the polynomial regression and different ML models on the observed θ_{sz} against Act in 2018 for the three sensors combined.

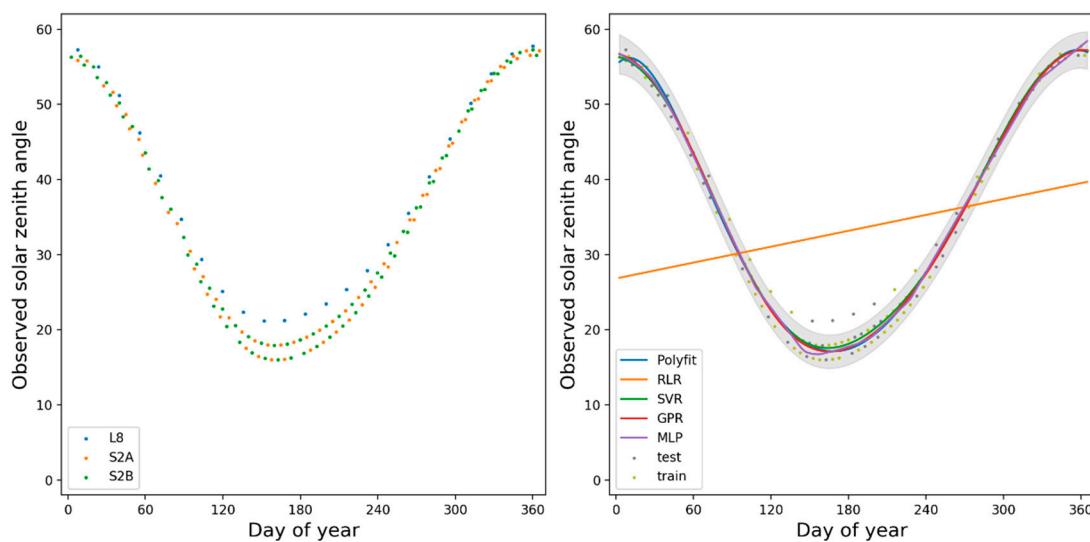


Figure 10. Observed θ_{sz} over Texas (-99.406° , 30.751°) plotted against Act (left) for all 12 months of 2018 for the three sensors combined. Comparison of the performance (right) for the polynomial regression and different ML models on the observed θ_{sz} against Act in 2018 for the three sensors combined.

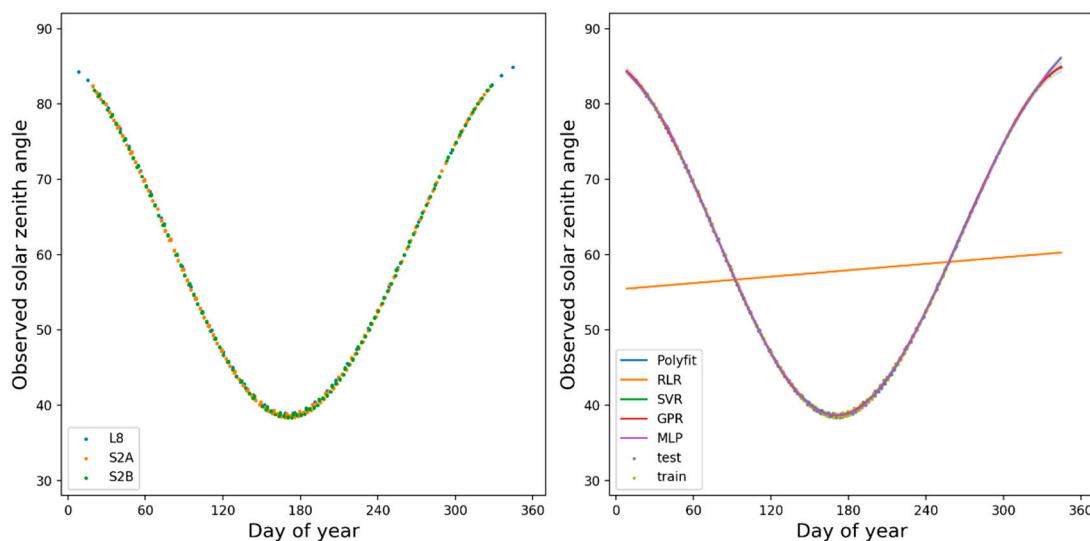


Figure 11. Observed θ_{sz} over Finland (25.194° , 61.653°) plotted against Act (left) in 2018 for the three sensors combined. Comparison of the performance (right) of the polynomial regression and ML models on the observed θ_{sz} against Act in 2018 for the three sensors combined.

The right-hand sides of Figures 9–11 compare the regression performance of the different ML models for predicting θ_{sz} , considering Act as the input variable for the Congo, Texas, and Finland, respectively. Polynomial regression is shown as a reference. For all cases, the nonlinear ML models fit the change in θ_{sz} with respect to Act. The RLR model produces a straight line for both cases. The regression results of SVR, GPR, and MLP reflect the variation of θ_{sz} in the three locations.

Tables 4–6 present summary statistics for the predicted θ_{sz} considering the inputs of Act, Lat and Act, and Lat and Lon and Act given by the different ML models for the Congo, Texas and Finland in 2018, respectively. Considering Act as the input variable, all of the nonlinear ML models achieve better results than the RLR linear model. SVR and GPR obtain better results than polynomial regression in all cases, achieving RMSEs of 1.769 and 1.810° for the Congo, 1.099 and 1.162° for Texas, and 0.187 and 0.181° for Finland, respectively, compared with 1.943° (Congo), 1.228° (Texas), and 0.207° (Finland) for polynomial regression. Compared with polynomial regression, SVR and GPR reduced the RMSE by 9.0 and 6.8% for Congo, 10.5 and 5.4% for Texas and 9.7 and 12.6% for Finland, respectively. The major difference between SVR and GPR occurs when Lat and Act, and Lat and Lon and Act are considered for training. For the Congo, GPR achieves an RMSE of 0.987° when Lat and Act are considered as inputs, compared with 1.279° for SVR; similar differences appear at the other two locations. This demonstrates that GPR achieves better regression performance when Lat and Act are considered. Additionally, the same results were found when Lat and Lon and Act were considered as inputs.

Table 4. Regression accuracy obtained with polynomial regression and ML models on the observed θ_{sz} for the Democratic Republic of Congo ($26.622^\circ, 0.356^\circ$) in 2018 for the three sensors combined. Results are given to three decimal places.

	Metric	Polyfit	RLR	SVR	GPR	MLP
Act	R^2	0.778	-0.068	0.816	0.808	0.784
	MAE	1.550°	3.575°	1.334°	1.401°	1.520°
	RMSE	1.943°	4.263°	1.769°	1.810°	1.918°
Lat and Act	R^2	0.036	0.904	0.943	0.902	
	MAE	-	3.508°	0.974°	0.907°	1.129°
	RMSE		4.051°	1.279°	0.987°	1.294°
Lat and Lon and Act	R^2	0.020	0.753	0.943	0.851	
	MAE	-	3.527°	1.464°	0.905°	1.291°
	RMSE		4.084°	2.051°	0.986°	1.594°

Table 5. Regression accuracy obtained with polynomial regression and ML models on the observed θ_{sz} for Texas ($-99.406^\circ, 30.751^\circ$) in 2018 for the three sensors combined. Results are given to three decimal places.

	Metric	Polyfit	RLR	SVR	GPR	MLP
Act	R^2	0.992	-0.104	0.994	0.993	0.992
	MAE	0.844°	13.059°	0.724°	0.790°	0.857°
	RMSE	1.228°	14.763°	1.099°	1.162°	1.240°
Lat and Act	R^2	-0.127	0.997	0.998	0.993	
	MAE	-	13.093°	0.625°	0.560°	0.972°
	RMSE		14.920°	0.823°	0.632°	1.156°
Lat and Lon and Act	R^2	-0.144	0.982	0.998	0.995	
	MAE	-	13.183°	1.480°	0.543°	0.766°
	RMSE		15.030°	1.899°	0.620°	0.991°

Table 6. Regression accuracy obtained with polynomial regression and ML models on the observed θ_{sz} for Finland ($25.194^\circ, 61.653^\circ$) in 2018 for the three sensors combined. Results are given to three decimal places.

	Metric	Polyfit	RLR	SVR	GPR	MLP
Act	R^2	1.000	-0.039	1.000	1.000	1.000
	MAE	0.164°	12.726°	0.153°	0.149°	0.183°
	RMSE	0.207°	14.443°	0.187°	0.181°	0.230°
Lat and Act	R^2	-0.034	1.000	1.000	0.999	
	MAE	-	12.677°	0.140°	0.125°	0.291°
	RMSE		14.411°	0.185°	0.159°	0.351°
Lat and Lon and Act	R^2	-0.037	0.991	1.000	0.999	
	MAE	-	12.703°	0.994°	0.126°	0.238°
	RMSE		14.429°	1.317°	0.165°	0.377°

5. Discussion

The combination of Landsat-8 and Sentinel-2A/2B has been widely used to monitor landcover use and for landcover mapping [42]. The BRDF effect has been demonstrated in both Landsat-8 and Sentinel-2 reflectance data [10]. Semi-empirical approaches have been advocated for the normalization of Landsat-8 and Sentinel-2 directional reflectance data to the nadir view. However, no suitable definition of θ_{sz} has yet been presented for normalizing the combined Landsat-8 and Sentinel-2A/2B directional reflectance data to produce reliable time series. θ_{sz} has the potential to change greatly over space and time because of the latitudinal variation of the local cross time and also because of seasonal changes in solar position. The annual reflectance variations caused by changes in θ_{sz} can

be as large as 0.053 for the red band and 0.065 for the NIR band at latitudes of 50° [43]. This is nontrivial and constitutes a serious issue for remote sensing applications.

In this paper, Landsat-8 and Sentinel-2A/2B metadata records collected in 2018 were analyzed. Combining the three sensors, the minimum and maximum θ_{sz} are 14.739 and 89.985°, and the mean value is 44.802° with a standard deviation of 18.250°. The differences in the global mean average θ_{sz} between Landsat-8 and Sentinel-2A and between Landsat-8 and Sentinel-2B are 6.09 and 5.09°, respectively. These differences are caused by the 30-minute difference in the mean equator overpass time (Landsat-8: 10:00, Sentinel-2: 10:30), which necessitates the normalization of θ_{sz} for the three sensors to eliminate the effects of BRDF from θ_{sz} variations.

The novelty of this research lies in that the four ML models were used to learn and optimize the functional forms of these models and then predict the values of θ_{sz} for the combination of the three sensors. The results showed that the GPR model achieved a higher accuracy to estimate the θ_{sz} compared to the polynomial and astronomical physical models. Further, research on the quantification of θ_{sz} for the combination of Landsat-8 and Sentinel-2A -2B is lacking and defining a θ_{sz} suitable for normalizing the BRDF time series given the data set is required.

The four ML models, i.e., RLR, SVR, GPR, and MLP, were selected to predict θ_{sz} for a combination of these three sensors. In the RLR model, the predicted θ_{sz} is assumed to be a linear weighted sum of the input variables; thus, the performance of the RLR model is limited by its simple function form. The SVR model is a regression version of the traditional support vector classification model, and it estimates θ_{sz} values by optimizing the penalty function after delivering at a sparse solution. GPR is a probabilistic approximation to non-parametric data distribution regression models; in the prediction, both a predictive mean value and result variance can be learned. The MLP model is a basic type of an artificial neural network, it can fit any nonlinear relationship through a non-linear activation function and back-propagation iterative optimization. We noted that the advantage of the MLP model compared to SVR and GPR is not substantial in θ_{sz} prediction. This is possible because that MLP-base model presents a better performance when the datasets are non-linear, complex, and redundant, etc., image and natural language data [44,45].

The optimal θ_{sz} definition for normalizing combined Landsat-8 and Sentinel-2A/2B reflectance data was considered as the following criteria: θ_{sz} can be modelled at any location and date, and the difference between the observed θ_{sz} and θ_{sz} for normalization are minimized. Lat, Lat and Act, and Lat and Lon and Act were used as the input variables into different ML models to predict θ_{sz} , respectively. RMSE, MAE, and R^2 metrics were used to evaluate the regression accuracy of the different ML models. Instead of an astronomical physical model [12], which builds upon polynomial and uses physical knowledge and mathematical relationships to convert geometrical coordinates to θ_{sz} , four ML models were used to fit the values of θ_{sz} . The functional forms of these models were learned and optimized from the global θ_{sz} data from all three sensors. The performance of the astronomical physical model depends on the estimation accuracy of polynomial regression. The research results showed that the GPR and SVR models are slightly better than the polynomial for global θ_{sz} data, and an improvement was made at all of the three locations' θ_{sz} datasets.

With Lat as the input variable, the nonlinear ML models achieved RMSEs of 12.597° (SVR), 12.463° (GPR), and 12.470° (MLP) and R^2 values of 0.516 (SVR), 0.526 (GPR), and 0.525 (MLP) when comparing the predicted θ_{sz} with the test θ_{sz} at the global scale. This prediction accuracy is comparable to that of the reference polynomial regression, which achieved an RMSE of 12.473° and an R^2 of 0.525, with GPR slightly better. SVR, GPR, and MLP achieved RMSEs of 1.396, 1.390, and 1.504° and R^2 values of 0.994, 0.994, and 0.993, respectively, when considering Lat and Act as the input variables; these values compare favorably with the linear RLR model (an RMSE of 17.484° and an R^2 of 0.067). There is little further improvement when Lat and Lon and Act are used as input variables. A relative importance test showed that Lat is most closely correlated with the variation of θ_{sz} ,

followed by Act. Lon has a weak relationship with θ_{sz} at a global scale. This is consistent with our physical knowledge that θ_{sz} varies with the latitudinal variation of local cross time and the seasonal changes in solar position.

The performance of the ML models was further evaluated at specific locations (Congo, Texas, and Finland). In the case of Finland, the predicted θ_{sz} against the Act gave RMSEs of 0.187° (SVR), 0.181° (GPR), and 0.230° (MLP) and R^2 values of 1.000 (SVR), 1.000 (GPR), and 1.000 (MLP), all of which are better than the linear RLR values of 14.443° (RMSE) and -0.039 (R^2). Compared with the precision statistics from the polynomial regression (0.207° for RMSE and 1.000 for R^2), MLP achieved comparable accuracy, and SVR and GPR performed better, reducing the RMSE by 9.7% (SVR) and 12.6% (GPR), compared with the polynomial. For the Congo and Texas, the nonlinear models (SVR, GPR, and MLP) exhibited more favorable performances than the linear RLR model. Compared with the well-established polynomial model, MLP produced a performance comparable to that of polynomial regression, and SVR and GPR gave greater accuracy in all cases. These results indicate that, although polynomial regression performs well, it is constrained by its explicit and simple parametric relations. On the contrary, for ML models, functional parameters are learned and optimized from data; thus, an improved performance can be achieved, compared with polynomial regression, in the case of complex nonlinear relations. GPR performs better than SVR, especially when Lat and Act, and Lat and Lon and Act were considered as input variables.

The resource consumption of each of the ML models was also investigated. The resource consumption rises with the model complexity, with training times of 7090.03 s for SVM, 1747.27 s for GPR, and 2152.50 s for MLP, compared with 1.36 s for RLR. All of the nonlinear models have a greater computational cost than the linear model. However, once the learning model has been trained, the model can be stored for further use; this is a clear advantage of the ML models over physical models.

6. Conclusions

In this study, the variation in θ_{sz} for the combination of Landsat-8, Sentinel-2A, and Sentinel-2B was quantified for the year 2018. Throughout the year, the minimum and maximum θ_{sz} for the three sensors combined are 14.739 and 89.985° , respectively, giving a mean value of 44.802° with a standard deviation of 18.250° . As Landsat-8 crosses the equator 30 min before Sentinel-2, the global mean average θ_{sz} difference is 6.09° (between Landsat-8 and Sentinel-2A) and 5.09° (between Landsat-8 and Sentinel-2B).

The four ML models were explored to train the forms and to optimize the model parameters and then used to estimate the values of θ_{sz} for a combination of three sensors. The ML models exhibit a performance comparable to that of polynomial regression when Lat is used to predict the global θ_{sz} . When Lat and Act were considered as the inputs, the nonlinear ML models achieve an obvious improvement in prediction accuracy compared with the linear model. The GPR model achieved the best overall model performance when using Lat and Act were used as the inputs, with an RMSE of 1.390° , MAE of 0.689° , and R^2 of 0.994, followed by the SVR model with an RMSE of 1.396° , MAE of 0.638° , and R^2 of 0.994. In addition, comprehensive analysis of the model regression for specific locations (Congo, Texas, and Finland) were discussed. Considering Act as input variable, the SVR and GPR models achieve more accurate estimations than the polynomial regression in all cases, implying that ML models can stratify θ_{sz} definition criteria better than polynomial regression. GPR achieved the best results, especially when Lat and Act, and Lat and Lon and Act were considered as the input variables. The GPR model is recommended for the prediction of global θ_{sz} for the three sensors combined in 2018.

In this study, more than 350,000 metadata records were obtained by down-sampling one-tenth of the combined dataset of Landsat-8 and Sentinel-2A/2B during 2018. A scalable Gaussian process [23] that has been specially designed for huge data volumes could be used to handle the full dataset. In the future, the proposed optimal θ_{sz} could be used for the normalization of bidirectional reflectance to produce consistent global time series for

the combination of Landsat-8 and Sentinel-2A -2B and quantify the effects of normalized θ_{sz} on the reflectance data. Landsat-9, proposed for launch in mid-2021, will be placed into the current Landsat-7 orbit and will carry a refined version of the Landsat-8 sensor payload [46]. Thus, Landsat-9 data could be added to the experiment, giving a combination of Landsat-8/9 and Sentinel-2A/2B from which to derive θ_{sz} values suitable for applying BRDF normalization, thus resulting in more consistent time series data.

Author Contributions: J.L. conceived, designed, and performed the experiments; J.L. analyzed the results; J.L. wrote the paper; B.C. wrote and revised the paper. Both authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key R&D Program of China (Grant number 2018YFA0606001), the Startup Foundation for Introducing Talent of Nanjing University of Information Science and Technology (Grant number 2018r071), and the National Natural Science Foundation of China (Grant number 41771114 and 41977404).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Landsat-8 data used for analysis can be found at here: <http://landsat.usgs.gov/consumer.php> (accessed on 8 June 2019). The Sentinel-2 data used for analysis can be found at here: <https://earthexplorer.usgs.gov> (accessed on 8 June 2019).

Acknowledgments: The authors would like to thank the USGS Earth Explorer team for the provision of the Landsat and Sentinel-2 metadata.

Conflicts of Interest: No potential conflict of interest was reported by the authors.

References

1. Knight, E.J.; Kvaran, G. Landsat-8 operational land imager design, characterization and performance. *Remote Sens.* **2014**, *6*, 10286–10305. [[CrossRef](#)]
2. Loveland, T.R.; Irons, J.R. Landsat 8: The plans, the reality, and the legacy. *Remote Sens. Environ.* **2016**, *185*, 1–6. [[CrossRef](#)]
3. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P. Sentinel-2: ESA’s optical high-resolution mission for GMES operational services. *Remote Sensing Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
4. Li, J.; Roy, D.P. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sens.* **2017**, *9*, 9.
5. Li, J.; Chen, B. Global Revisit Interval Analysis of Landsat-8–9 and Sentinel-2A-2B Data for Terrestrial Monitoring. *Sensors* **2020**, *20*, 6631. [[CrossRef](#)]
6. Poortinga, A.; Tenneson, K.; Shapiro, A.; Nquyen, Q.; San Aung, K.; Chishtie, F.; Saah, D. Mapping plantations in Myanmar by fusing landsat-8, sentinel-2 and sentinel-1 data along with systematic error quantification. *Remote Sens.* **2019**, *11*, 831. [[CrossRef](#)]
7. Roy, D.P.; Huang, H.; Boschetti, L.; Giglio, L.; Yan, L.; Zhang, H.H.; Li, Z. Landsat-8 and Sentinel-2 burned area mapping-A combined sensor multi-temporal change detection approach. *Remote Sens. Environ.* **2019**, *231*, 111254. [[CrossRef](#)]
8. 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](#)]
9. Derkacheva, A.; Mouginot, J.; Millan, R.; Maier, N.; Gillet-Chaulet, F. Data Reduction Using Statistical and Regression Approaches for Ice Velocity Derived by Landsat-8, Sentinel-1 and Sentinel-2. *Remote Sens.* **2020**, *12*, 1935. [[CrossRef](#)]
10. Roy, D.P.; Li, J.; Zhang, H.K.; Yan, L.; Huang, H.; Li, Z. Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. *Remote Sens. Environ.* **2017**, *199*, 25–38. [[CrossRef](#)]
11. Lucht, W.; Lewis, P. Theoretical noise sensitivity of BRDF and albedo retrieval from the EOS-MODIS and MISR sensors with respect to angular sampling. *Int. J. Remote Sens.* **2000**, *21*, 81–98. [[CrossRef](#)]
12. Zhang, H.K.; Roy, D.P.; Kovalskyy, V. Optimal Solar Geometry Definition for Global Long-Term Landsat Time-Series Bidirectional Reflectance Normalization. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1410–1418. [[CrossRef](#)]
13. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [[CrossRef](#)]
14. Ruescas, A.B.; Hieronymi, M.; Mateo-Garcia, G.; Koponen, S.; Kallio, K.; Camps-Valls, G. Machine learning regression approaches for colored dissolved organic matter (CDOM) retrieval with S2-MSI and S3-OLCI simulated data. *Remote Sens.* **2018**, *10*, 786. [[CrossRef](#)]

15. Shen, H.; Jiang, Y.; Li, T.; Cheng, Q.; Zeng, C.; Zhang, L. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. *Remote Sens. Environ.* **2020**, *240*, 111692. [[CrossRef](#)]
16. Irons, J.R.; Dwyer, J.L.; Barsi, J.A. The next Landsat satellite: The Landsat data continuity mission. *Remote Sens. Environ.* **2012**, *122*, 11–21. [[CrossRef](#)]
17. Gascon, F.; Bouzinac, C.; Thepaut, O.; Jung, M.; Francesconi, B.; Louis, J.; Fernandez, V. Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sens.* **2017**, *9*, 584. [[CrossRef](#)]
18. WWW1. Landsat Metadata. Available online: <http://landsat.usgs.gov/consumer.php> (accessed on 8 June 2019).
19. USGS, Landsat Collection-1 Product Definition. Available online: <https://www.usgs.gov/media/files/landsat-collection-1-level-1-product-definition> (accessed on 16 April 2019).
20. WWW2. Sentinel-2 Metadata. Available online: <https://earthexplorer.usgs.gov> (accessed on 8 June 2019).
21. Roy, D.P.; Li, J.; Zhang, H.K.; Yan, L. Best practices for the reprojection and resampling of Sentinel-2 Multi Spectral Instrument Level 1C data. *Remote Sens. Lett.* **2016**, *7*, 1023–1032. [[CrossRef](#)]
22. Camps-Valls, G.; Verrelst, J.; Muñoz-Marí, J.; Laparra, V.; Mateo-Jiménez, F.; Gómez-Dans, J. A survey on Gaussian processes for earth-observation data analysis: A comprehensive investigation. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 58–78. [[CrossRef](#)]
23. Liu, H.; Ong, Y.S.; Shen, X.; Cai, J. When Gaussian process meets big data: A review of scalable GPs. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 4405–4423. [[CrossRef](#)]
24. Roy, D.P.; Ju, J.; Kline, K.; Scaramuzza, P.L.; Kovalskyy, V.; Hansen, M.; Zhang, C. Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sens. Environ.* **2010**, *114*, 35–49. [[CrossRef](#)]
25. Bengio, Y.; Grandvalet, Y. No unbiased estimator of the variance of k-fold cross-validation. *J. Mach. Learn. Res.* **2004**, *5*, 1089–1105.
26. Lee, Y.; Han, D.; Ahn, M.H.; Im, J.; Lee, S.J. Retrieval of total precipitable water from Himawari-8 AHI data: A comparison of random forest, extreme gradient boosting, and deep neural network. *Remote Sens.* **2019**, *11*, 1741. [[CrossRef](#)]
27. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Hartling, S.; Esposito, F.; Fritsch, F.B. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* **2020**, *237*, 111599. [[CrossRef](#)]
28. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* **2012**, *118*, 127–139. [[CrossRef](#)]
29. Fernández-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* **2014**, *15*, 3133–3181.
30. Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote Sens.* **2014**, *6*, 964–983. [[CrossRef](#)]
31. Qian, Y.; Zhou, W.; Yan, J.; Li, W.; Han, L. Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sens.* **2015**, *7*, 153–168. [[CrossRef](#)]
32. Brereton, R.G.; Lloyd, G.R. Support vector machines for classification and regression. *Analyst* **2010**, *35*, 230–267. [[CrossRef](#)]
33. Deka, P.C. Support vector machine applications in the field of hydrology: A review. *Appl. Soft Comput.* **2014**, *19*, 372–386.
34. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Duchesnay, E. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
35. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–27. [[CrossRef](#)]
36. Rasmussen, C.E.; Nickisch, H. Gaussian processes for machine learning (GPML) toolbox. *J. Mach. Learn. Res.* **2010**, *11*, 3011–3015.
37. Verrelst, J.; Camps-Valls, G.; Muñoz-Marí, J.; Rivera, J.P.; Veroustraete, F.; Clevers, J.G.; Moreno, J. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—A review. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 273–290. [[CrossRef](#)]
38. Verrelst, J.; Alonso, L.; Camps-Valls, G.; Delegido, J.; Moreno, J. Retrieval of vegetation biophysical parameters using Gaussian process techniques. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 1832–1843. [[CrossRef](#)]
39. Morales-Alvarez, P.; Pérez-Suay, A.; Molina, R.; Camps-Valls, G. Remote sensing image classification with large-scale Gaussian processes. *IEEE Trans. GeoScience Remote Sens.* **2017**, *56*, 1103–1114. [[CrossRef](#)]
40. Hecht-Nielsen, R. Theory of the backpropagation neural network. In *Neural Networks for Perception*; Academic Press: Cambridge, MA, USA, 1992; pp. 65–93.
41. Wulder, M.A.; White, J.C.; Loveland, T.R.; Woodcock, C.E.; Belward, A.S.; Cohen, W.B.; Fosnight, E.A.; Shaw, J.; Masek, J.G.; Roy, D.P. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* **2016**, *185*, 271–283. [[CrossRef](#)]
42. ED Chaves, M.; CA Picoli, M.; D Sanches, I. Recent Applications of Landsat 8/OLI and Sentinel-2/MSI for Land Use and Land Cover Mapping: A Systematic Review. *Remote Sens.* **2020**, *12*, 3062. [[CrossRef](#)]
43. Gao, F.; He, T.; Masek, J.G.; Shuai, Y.; Schaaf, C.B.; Wang, Z. Angular effects and correction for medium resolution sensors to support crop monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4480–4489. [[CrossRef](#)]
44. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
45. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)] [[PubMed](#)]
46. Masek, J.G.; Wulder, M.A.; Markham, B.; McCorkel, J.; Crawford, C.J.; Storey, J.; Jenstrom, D.T. Landsat 9: Empowering open science and applications through continuity. *Remote Sens. Environ.* **2020**, *248*, 111968. [[CrossRef](#)]