land

MDPI

Article

Innovative Fusion-Based Strategy for Crop Residue Modeling

Solmaz Fathololoumi [1], Mohammad Karimi Firozjaei [2] and Asim Biswas [1,*]

[1] School of Environmental Sciences, University of Guelph, Guelph, ON N1G 2W1, Canada
[2] Department of Remote Sensing and GIS, Faculty of Geography, University of Tehran, Tehran 1417853933, Iran
* Correspondence: biswas@uoguelph.ca; Tel.: +1-519-731-6252

Abstract: The purpose of this study was to present a new strategy based on fusion at the decision level for modeling the crop residue. To this end, a set of satellite imagery and field data, including the Residue Cover Fraction (RCF) of corn, wheat and soybean was used. Firstly, the efficiency of Random Forest Regression (RFR), Support Vector Regression (SVR), Artificial Neural Networks (ANN) and Partial-Least-Squares Regression (PLSR) in RCF modeling was evaluated. Furthermore, to increase the accuracy of RCF modeling, different algorithms results were combined based on their modeling error, which is called the decision-based fusion strategy. The R2 (RMSE) between the actual and modeled RCF based on ANN, RFR, SVR and PLSR algorithms for corn were 0.83 (3.89), 0.86 (3.25), 0.76 (4.56) and 0.75 (4.81%), respectively. These values were 0.81 (4.86), 0.85 (4.22), 0.78 (5.45) and 0.74 (6.20%) for wheat and 0.81 (3.96), 0.83 (3.38), 0.76 (5.01) and 0.72 (5.65%) for soybean, respectively. The error of corn, wheat and soybean RCF estimating decision-based fusion strategy was reduced by 0.90, 0.96 and 0.99%, respectively. The results showed that by implementing the decision-based fusion strategy, the accuracy of the RCF modeling was significantly improved.

Keywords: crop residue; fusion; machine learning algorithm; reflective and radar bands

Citation: Fathololoumi, S.; Karimi Firozjaei, M.; Biswas, A. Innovative Fusion-Based Strategy for Crop Residue Modeling. *Land* **2022**, *11*, 1638. https://doi.org/10.3390/land11101638

Academic Editors: Carmine Serio, Guido Masiello and Sara Venafra

Received: 22 August 2022
Accepted: 20 September 2022
Published: 23 September 2022

1. Introduction

Modern agricultural activities, such as plowing and using heavy machinery known as tillage, can damage soil health [1,2]. In this case, the soil is more easily leaching by rain and loses its top layer, which is crucial for crop growth. The leached soil will flow downstream into the rivers and pollute the water due to elements such as phosphorus [3]. On the other hand, with decreasing soil quality, precipitated carbon is released [4]. The release of carbon from the soil plays an important role in increasing the carbon dioxide in the Earth's atmosphere [5,6]. Gases are one of the parameters affecting climate change. Therefore, maintaining soil quality in the agricultural process is very important [7,8].

One of the possible, cheap and feasible ways to reduce the damage caused by wind and water erosion and increase water storage to soil productivity is to maintain the remaining vegetation on the soil surface of agricultural lands at harvest time [9–11]. Crop residues consist of various components of the crop, including leaves, seeds, stems, etc., after harvest on agricultural land [12,13]. The presence of these residues on the soil surface can strengthen soil organic matter, better the absorption of nutrients by the plant and increase the efficiency of chemical fertilizers. Crop residues also have a large effect on soil, crop and environmental factors, such as water permeability, evaporation, crop yield and erosion [7,14–16]. They can improve the physical, chemical and biological condition of the soil and ultimately lead to a healthier crop due to its desirable and nutritious composition. Preserving residues at the soil surface by preventing the emission of gases, such as NH_3, CO_2 and SO_2, can reduce air pollution, while burning plant residues emits these gases [17,18].

Due to the importance of preserving crop residues on the soil surface, modeling and mapping residues as an indicator of tillage intensity are of great importance in agricultural management and achieving sustainable agricultural goals, including maintaining environmental health [13,19]. Mapping crop residues for agricultural areas can be a criterion for

evaluating the efficiency and quality of methods and tools used in harvesting. It is practically impossible to use traditional methods, such as field visits and sampling, to determine the amount of residues on a large scale and in a short time [12,20]. Utilizing the capabilities of remote sensing techniques and data can be useful in quantifying crop residues on a large spatial and temporal scales and higher accuracy [18,21]. Previous studies have used various satellite imagery to model the amount of residue cover fraction (RCF), including reflective multispectral imagery, such as Landsat; Sentinel 2 [16,22,23]; radar imagery, such as RADARSAT [13,24]; and reflective hyper Spectral imagery, such as Probe-1 [25]. Each of these types of images has advantages and disadvantages [13].

In previous studies, several remote sensing-based indices, such as Normalized Difference Senescent Vegetation Index (NDSVI) [26], Normalized Difference Residue Index (NDRI) [27], Normalized Difference Tillage Index (NDTI) [28], Shortwave Green Normalized Difference Index (SGNDI) [29], Shortwave Infrared Normalized Difference Residue Index (SINDRI) [30], Broadband spectral Angle Index (BAI) [8], Dead Fuel Index (DFI) [31], Normalized Difference Index (NDI) [32], Normalized Difference Vegetation Index (NDVI) [8], Simulated crop residue cover (MCRC) [29], Simple tillage index (STI) [28], Simulated cellulose absorption index (3BI1) [33], Simulated lignin Cellulose Absorption Index (3BI2) [33], Simulated NDRI (3BI3) [33] and Short-wave near-infrared Normalized Difference residue Index (SRNDI) [34], etc., have been proposed to identify and quantify the RCF [14,34]. In some studies, the efficiency of different spectral indices was compared [29,33]. The results showed that each of these indicators can have different performances, some of them are suitable for dry areas and some for wet areas. A number of indicators do not perform well in areas with high vegetation. Each of the developed indicators has advantages and disadvantages. Yue, Tian, Dong, Xu and Zhou [29] showed that single indices are not highly capable of modeling RCF in the complex surface conditions of agricultural areas. Hence, in some studies, multivariate regression based methods, such as Random Forest Regression (RFR), Support Vector Regression (SVR), Artificial Neural Networks (ANN), etc., were proposed to model RCF [33]. The researchers used experimental regression methods to examine the linear or nonlinear relationship between actual RCF and remote sensing-based indices related to RCF [7,14,33], the spectral angle [8] or spectral unmixing [35] used to estimate the amount of RCF. Raoufat, Dehghani, Abdolabbas, Kazemeini and Nazemossadat [9] utilized Landsat 8 and drone data for RCF mapping and found that Landsat 8 data was more accurate than drone data although the drone data had its own advantages. Yue and Tian [36] used the spectral and laboratory data for RCF mapping. They evaluated RS data and triangle technique using RFR method in their study. They concluded that their proposed method was very effective in the accurate modeling of RCF and decreased the negative effect of soil moisture on it. Wang et al. [37] used MODIS and ground data to quantify some crop-related indices in a large-scale area using ground data and building a linear regression relationship between them. They showed that their used method was successful in monitoring soil conditions, including soil erosion.

Although the research in RCF modeling is limited, summarizing previous studies show that several models have been developed over the years to estimate the RCF, each with its own advantages and disadvantages. Selecting the appropriate model to estimate the amount of RCF has a high impact on the modeling accuracy of this parameter. Therefore, providing an integrated model based on using the capabilities of different models and indicators in estimating RCF can be useful in improving the modeling accuracy of this parameter. Sentinel 1 satellite imagery is one of the well-known and widely used radar data in various agricultural applications. However, our knowledge shows that the capability of this radar image's bands in estimating the amount of RCF has not been evaluated. Therefore, evaluating the performance of the satellites and combining the capabilities of these bands with the indicators presented in previous studies to improve the accuracy of modeling RCF can be useful and crucial.

The purpose of this study was to present a new strategy based on fusion at the decision level for modeling the RCF. In this study, (1) the efficiency of spectral indices

based on reflective multispectral images presented in previous studies in modeling RCF were compared. (2) The importance of using Sentinel 1 radar satellite imagery bands in improving the accuracy of RCF modeling was assessed. (3) The efficiency of RFR, SVR, ANN and Partial-Least-Squares Regression (PLSR) algorithms in modeling RCF was evaluated and compared. A new strategy was proposed to integrate the results obtained from different algorithms at the decision level to improve the accuracy of modeling RCF.

2. Study Area

Ontario is one of Canada's most important agricultural centers, with a wide variety of agricultural products and orchards. Major agricultural products in Ontario include corn, wheat and soybeans, barley, forage, oat, canola, etc. A region in southern of Ontario was selected as the study area. This region is located at 82.5 degrees West and 42.5 degree North. The agricultural products of this area include the three main crops of corn, wheat and soybeans (Figure 1). The area of the study region is about 4324 km^2 and the area under corn, wheat and soybean cultivation in this area is 1218, 284 and 1087 km^2, respectively. The study area in terms of climatic division, has a warm summer humid continental based on the Köppen climate classification. The average annual temperature and rainfall are 17.0 °C and 424 mm, respectively. Therefore, due to the weather conditions and long winters in this area, most agricultural products are harvested by September. Criteria for selecting the study area were (1) high diversity of agricultural products, especially important crops in Canada; (2) proximity of the harvest time of the three types of agricultural products available in the region and the different farms available for each product; (3) access to the suitable cloud-free satellite image corresponding to the harvest date of all three crops for that area; and (4) easy access of the area in order to conduct field studies.

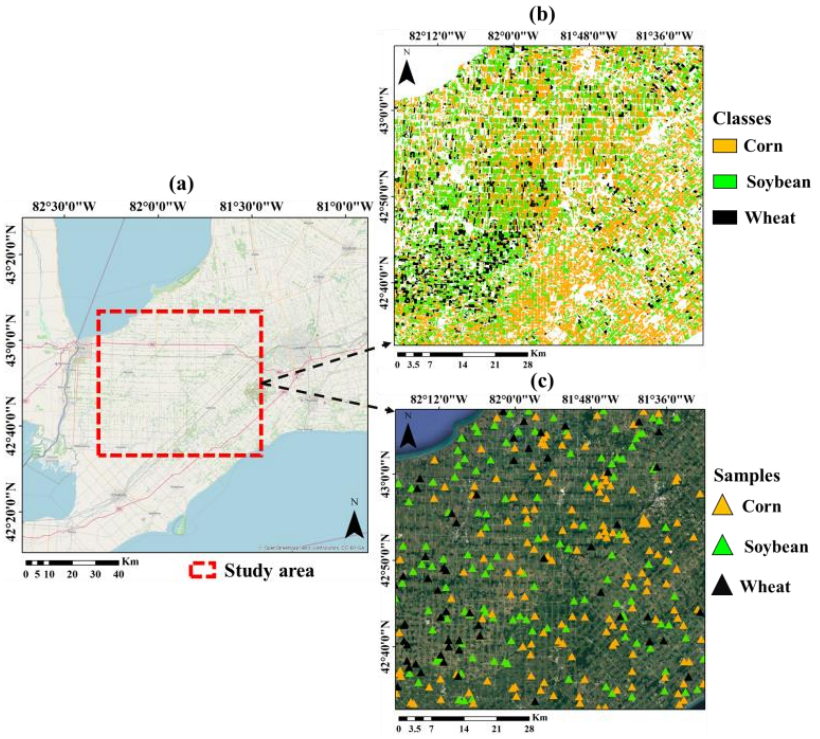

Figure 1. Maps of (**a**) the geographical location of the study area, (**b**) the land crop classes of the study area and (**c**) the geographical location and spatial distribution of RCF samples of different land crops.

3. Data and Methods

3.1. Data

A set of satellite images and ground data were used to evaluate different models and data in RCF modeling. Satellite imagery included the Landsat 8 Collection 2 Surface Reflectance image for 13 October 2020 and the Sentinel 1 Ground Range Detected (GRD)

image for 7 October 2020. The spatial resolution of the Landsat 8 and Sentinel 1 image bands is 30 and 10 m, respectively. The Landsat 8 bands, including Blue, Green, Red, NIR, SWIR 1 and SWIR 2, and Sentinel 1 bands, including VV and VH, were used in this study. Some criteria were considered when selecting the date of the images: (1) a lack of cloud cover in the area during the Landsat 8 overpass time, (2) the absence of precipitation in the study area a few days before the satellite overpass time and (3) proximity to the date of harvest of agricultural products. Landsat 8 image downloaded from https://earthexplorer.usgs.gov/ (accessed on 13 February 2021) and sentinel 1 downloaded from https://scihub.copernicus.eu/ (accessed on 16 February 2021). The land crop map prepared by Agriculture and Agri-Food Canada (AAFC) in 2020 with the spatial resolution of 30 m was used to mask various agricultural products. This data is produced annually and can be downloaded from https://open.canada.ca/data/en/dataset?q (accessed on 2 March 2021) website.

Ground data collection includes determining the RCF in autumn from a number of corn, soybean and wheat fields that were performed after harvest on the 9 October 2020, 10 October 2020 and 9 November 2020 dates. For this purpose, RCF values were determined at 57 land points for wheat, 149 points for corn and 128 points for soybeans. The selection of land areas was carried out in such a way that there was a suitable distribution in the fields of these crops in the whole study area (Figure 1b). The minimum distance between sampling points was 500 m. Additionally, based on the initial and complete field visit of the study area, the suitable distribution of absolute values between the highest and lowest actual values of RCF in the study area was also considered when selecting the sampling points. Little to no rain (less than 0.25 cm) had fallen the week before the sampling, and the soil moisture levels were constantly dry. When field sampling took place, the weather was identical to the weather when taking images the day before. Ordinary camera images were used for sampling. A Phantom 3 SE drone was used to produce images from each selected sampling point. The flight height of the drone for imaging was 20 m, and a digital orthophoto with a spatial resolution of 20 cm was prepared for each sampling point. Then, for each image, the position of product residues was manually digitized. After processing the images taken by the camera, the remaining coverage fraction was calculated for an area of 900 m^2 around each sampling point. Agisoft Metashape 1.8.3, developed by Agisoft LLC (St. Petersburg, Russia) and ArcMap 10.6.1, developed by Esri (Environmental Systems Research Institute), (Redlands, California, United States) software were used to determine the RCF based on the images prepared with the drone (digital orthophoto preparation and digitization).

3.2. Methods

Firstly, a map of common spectral indices in RCF modeling was prepared based on the Landsat 8 and Sentinel 1 VV and VH spectral bands. Then, these maps were masked to areas with NDVI < 0.3 to limit analysis to areas without significant green vegetative ground cover [16] and was also masked to agricultural crop fields using the AAFC map. Then, the efficiency of each of these indicators in modeling the RCF was evaluated for each agricultural product. Secondly, the efficiency of different algorithms, including RFR, SVM, ANN and PLSR in modeling RCF, was evaluated and they were compared with each other. Furthermore, different algorithms' results were combined based on the modeling error in order to increase the accuracy of receipt modeling. This strategy used the fusion capability at the decision level to improve the accuracy of RCF modeling.

3.2.1. Effective Variables

In previous studies, various spectral indices, including MCRC, SGNDI, DFI, STI, NDSVI, NDTI, NDI5, NDI7, NDVI, 3BI1, 3BI2, 3BI3, BAI and SRNDI, have been developed for RCF modeling [16,25,30,32–34]. Reflective band information was used to calculate these indices. In this study, in addition to spectral indices based on reflective band information, backscatter data obtained from VV and VH bands of Sentinel 1 were also used in RCF

modeling. The details of spectral indices and radar bands used as independent and primary variables in RCF modeling are shown in Table 1. To compare the efficiency of each of these independent variables in RCF modeling, their correlation coefficient with the values of RCF at the validation sampling points was calculated.

Table 1. Spectral indices and radar bands used as independent and primary variables in RCF modeling.

Type	Spectral Indices	Equation	Reference
Reflective-based index	MCRC	(OLI6 − OLI3)/(OLI6 + OLI3)	[29]
	SGNDI	(OLI3 − OLI7)/(OLI3 + OLI7)	[29]
	DFI	100 × (1 − OLI7/OLI6) × (OLI4/OLI5)	[31]
	STI	OLI6/OLI7	[28]
	NDSVI	(OLI6 − OLI4)/(OLI6 + OLI4)	[26]
	NDTI	(OLI6 − OLI7)/(OLI6 + OLI7)	[28]
	NDI5	(OLI5 − OLI6)/(OLI5 + OLI6)	[32]
	NDI7	(OLI5 − OLI7)/(OLI5 + OLI7)	[32]
	SRNDI	(OLI7 − OLI4)/(OLI7 + OLI4)	[34]
	NDVI	(OLI5 − OLI4)/(OLI5 + OLI4)	[33]
	3BI1	100 × (0.5 × (OLI2 + OLI7) − OLI4)	[33]
	3BI2	(OLI2 − OLI4)/(OLI2 − OLI7)	[33]
	3BI3	(OLI7 − OLI4)/(OLI7 + OLI6)	[33]
	BAI	-	[8]
Backscatter bands	VV	-	-
	VH	-	-

3.2.2. Machine Learning Methods

Multivariate modeling based on four common algorithms in agricultural and environmental modeling, including RFR, SVM, ANN and PLSR, was used to model the RCF. In the first scenario, all reflective band-based spectral indices were used as independent variables. In the second scenario, in addition to spectral indices based on reflective bands, VV and VH band information was also used in the modeling process. Each RFR, SVM, ANN and PLSR algorithm was calibrated based on training data (96 samples for corn, 38 samples for wheat and 84 samples for soybean). Then the efficiency of each of these algorithms in estimating the RCF fraction using test data (53 samples for corn, 20 samples for wheat and 43 samples for soybean) was evaluated.

PLSR breaks down both dependent and independent variables into a number of major components. PLSR is a two-line calibration algorithm that converts a large number of correlated linear variables into several non-correlated variables based on data compression. Hence, this algorithm can solve the challenges of high correlation between variables and overfitting in the modeling process [38].

In recent years, artificial neural networks have been widely used to estimate various environmental variables based on satellite data [39,40]. In this study, a back-propagation ANN was used to model the residue. This algorithm consists of input, hidden and output layers. Sigmoid and linear functions were used for activation in hidden and output nodes, respectively. To calibrate synaptic coefficients, the Levenberg–Marquardt minimization algorithm was used [41]. The number of nodes in the hidden layers varied between 4 to 8. To optimize the structural parameters of the ANN algorithm for the network, we changed the momentum coefficient and learning rate from 0.1 to 1.0 with a step of 0.05. The number of nodes in the hidden layer varied from 3 to 7. Mean squared error was used as a measure of the performance threshold and the determination of a network with optimal structure in receipt fraction modeling. The optimal network was selected in terms of mean absolute error between validation and predictions data.

The SVR model is a widely used algorithm for solving nonlinear problems [42]. In the SVR method, n-dimensional input variables are transferred to the new feature space with higher dimensions using the core functions and, as a result, optimal separator super planes are developed [43]. In this study, different Gaussian, linear, nonlinear quadratic, cubic, etc.,

kernels were evaluated, and finally the Gaussian kernel was selected and used as a function in the receipt fraction estimation process. The optimal values of box constraint, kernel scale and epsilon were set to 909, 857 and 0.04, respectively.

RFR is an ensemble-learning algorithm that combines a large set of decision trees to improve the accuracy of estimating a variable [44]. RFR has several advantages in modeling environmental variables, including (1) low sensitivity to noise and over-fitting, and (2) the use of a large number of quantitative and qualitative variables in the modeling process [45,46]. To implement this algorithm, two parameters, the number of trees and the number of attributes, must be set. The number of trees varied from 30 to 300 with step size 30 and the number of trees 150 was selected as the optimal value.

The optimal model for each of these four algorithms was selected to estimate the RCF and the mean absolute error between the validation data and the predictions.

3.2.3. Decision-Based Fusion Approach

To reduce the error of modeled RCF based on remote sensing data due to the weaknesses of different algorithms, in the proposed strategy, the results of four RFR, SVM, ANN and PLSR algorithms were combined based on Equation (1).

$$RCF_f = \sum_{i=1}^{n} W_i RCF_{model(i)} \quad (1)$$

In Equation (1), RCF_f is the modeled RCF based on the remote sensing data by combining the results of different algorithms, $RCF_{model(i)}$ is the fraction of the modeled RCF based on the remote sensing data obtained from the ith algorithm, Wi is the degree of importance of the ith algorithm and n is the number of used algorithms. Equation (2) is used to calculate the significance of the ith algorithm.

$$W_i = \frac{RMSE_{model(i)}}{\sum_{i=1}^{n} RMSE_{model(i)}} \quad (2)$$

In Equation (2), $RMSE_{model(i)}$ is the mean squares root of the of the estimated fraction and is based on the ith algorithm. The lower the RMSE of an algorithm in estimating the RCF, the greater its impact and importance in the result of the RCF estimation. MATLAB 2019a software was used to implement various indices and algorithms for RCF modeling.

4. Results

The mean (sd) of the RCF for the calibration data of corn, wheat and soybean were 17.8 (9.6), 20.1 (11.2) and 19.2 (7.7)%, respectively (Figure 2). These values were 17.4 (9.4), 18.2 (10.3) and 18.6 (8.5)% for validation data, respectively. The RCF means and sd values based on the calibration and validation datasets were close to each other. The highest frequency of calibration data for corn, wheat and soybean crops was in the 10.9–18.3, 10.5–17.3 and 17.4–23.4% categories, respectively. For validation data, the highest frequency for these products was in the 12.2–18.4, 11.1–16.6 and 18.8–25.0% categories, respectively.

The efficiency of spectral indices based on reflective bands in RCF modeling was different. The efficiency of NDI5, NDI7, NDTI, NDVI, STI, DFI and BAI was higher than other indices, such as 3BI1, 3BI2, MCRC, NDSVI, SGNDI, and SRNDI (Table 2). The R2 between BAI and corn, wheat and soybean residues were 0.63, 0.66 and 0.61, respectively, which was higher than other spectral indices. The R^2 between the VV (VH) bands and the corn, wheat and soybean residues were 0.25 (0.29), 0.28 (0.36) and 0.20 (0.25), respectively. The efficiency of radar bands in RCF modeling was less than the spectral indices.

In the first scenario (dataset including reflective band-based spectral indices), based on calibration data, the R^2 (RMSE) between the actual and modeled RCF using ANN, RFR, SVR and PLSR algorithms for corn crop were 0.86 (3.13), 0.91 (2.63), 0.82 (3.92) and 0.79 (4.22%), respectively (Figure 3). These values were 0.84 (4.22), 0.88 (3.72), 0.81 (5.25) and 0.77 (5.85%) for wheat and 0.85 (3.14), 0.89 (2.73), 0.79 (3.81) and 0.75 (4.04%), respectively,

for soybean. The efficiency of RCF modeling using different machine learning algorithms based on spectral indices was different. The RFR and PLSR algorithms had the highest and lowest accuracy in forming an optimal network for RCF modeling, respectively.

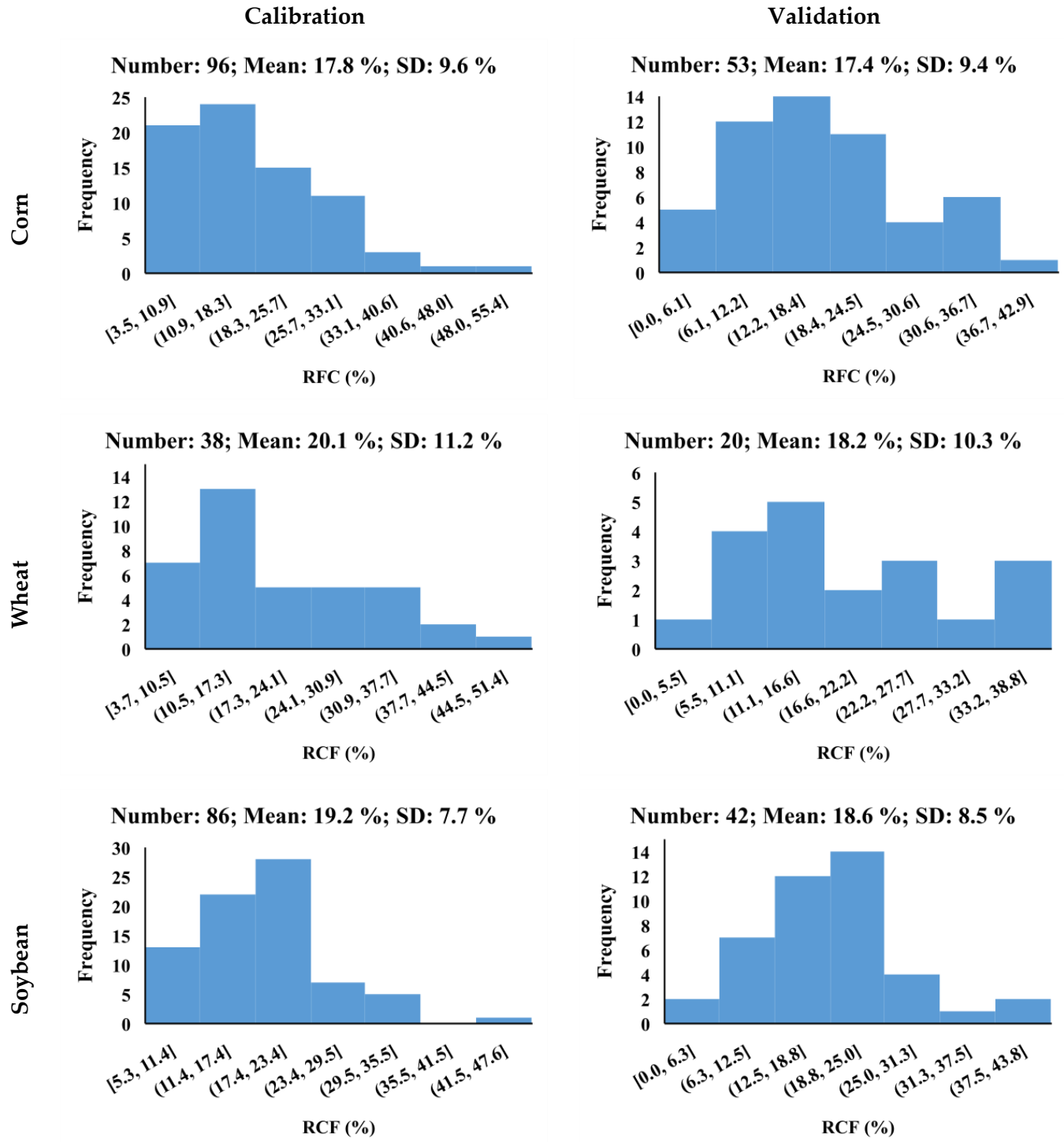

Figure 2. Frequency distribution of RCF values of corn, wheat and soybean crops for calibration and validation data in different classes.

Table 2. The R^2 between the effective variables and the residues for different crops.

	Corn	Wheat	Soybean
3BI1	0.09	0.08	0.11
3BI2	0.15	0.10	0.33
3BI3	0.46	0.35	0.56
MCRC	0.13	0.38	0.13
NDI5	0.54	0.48	0.41
NDI7	0.60	0.51	0.45
NDSVI	0.07	0.11	0.11
NDTI	0.42	0.48	0.61
NDVI	0.43	0.51	0.22
SGNDI	0.20	0.08	0.22
SRNDI	0.24	0.15	0.21
STI	0.43	0.50	0.60
DFI	0.55	0.59	0.52
BAI	0.63	0.66	0.61
VV	0.25	0.28	0.20
VH	0.29	0.36	0.25

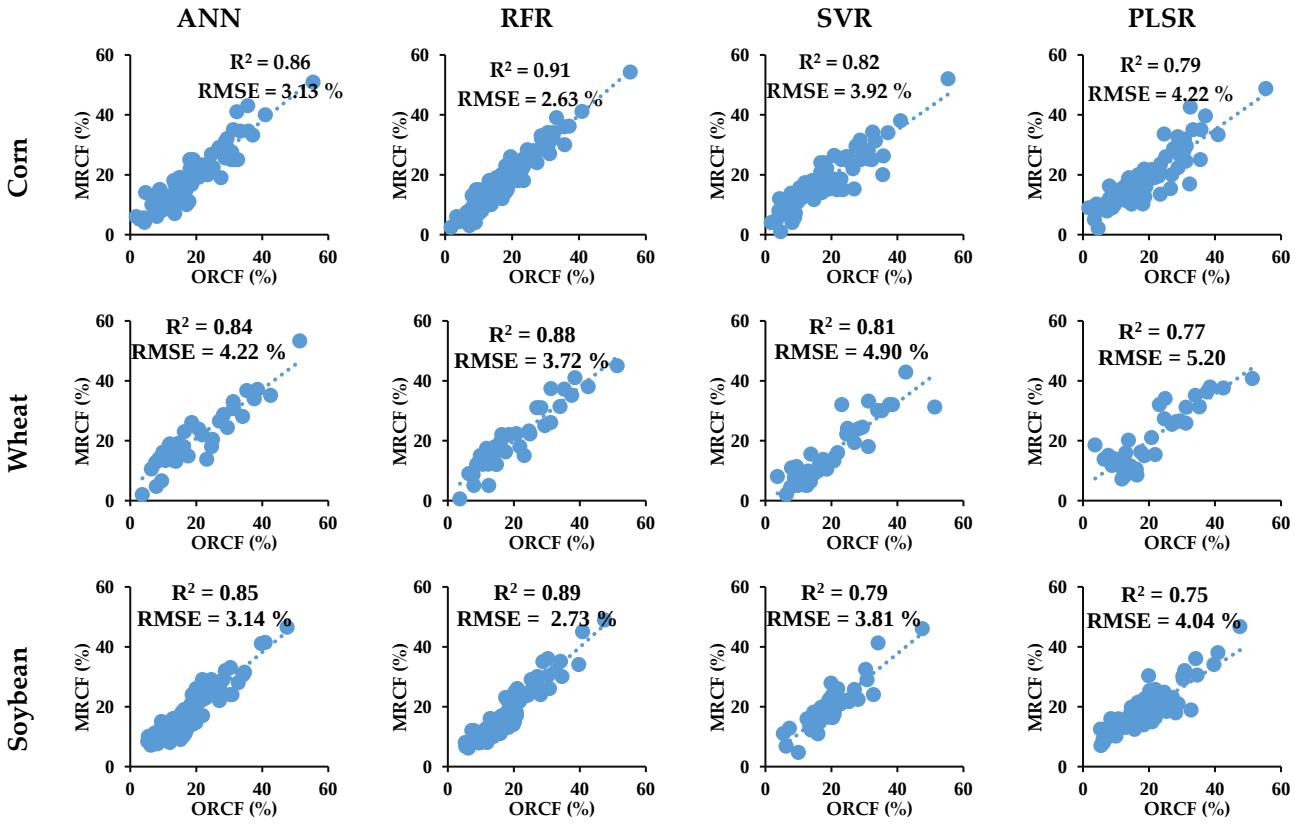

Figure 3. R^2 and RMSE between real and modeled RCF based on calibration data.

For the validation data, the R^2 (RMSE) between the actual and modeled RCF based on ANN, RFR, SVR and PLSR algorithms for corn were 0.83 (3.89), 0.86 (3.25), 0.76 (4.56) and 0.75 (4.81%) (Figure 4). These values were 0.81 (4.86), 0.85 (4.22), 0.78 (5.45) and 0.74 (6.20%) for wheat and 0.81 (3.96), 0.83 (3.38), 0.76 (5.01) and 0.72 (5.65%), respectively, for soybean. The results showed that the RFR algorithm had the highest accuracy in RCF modeling. The efficiency of this algorithm in corn RCF modeling was higher than soybean and wheat.

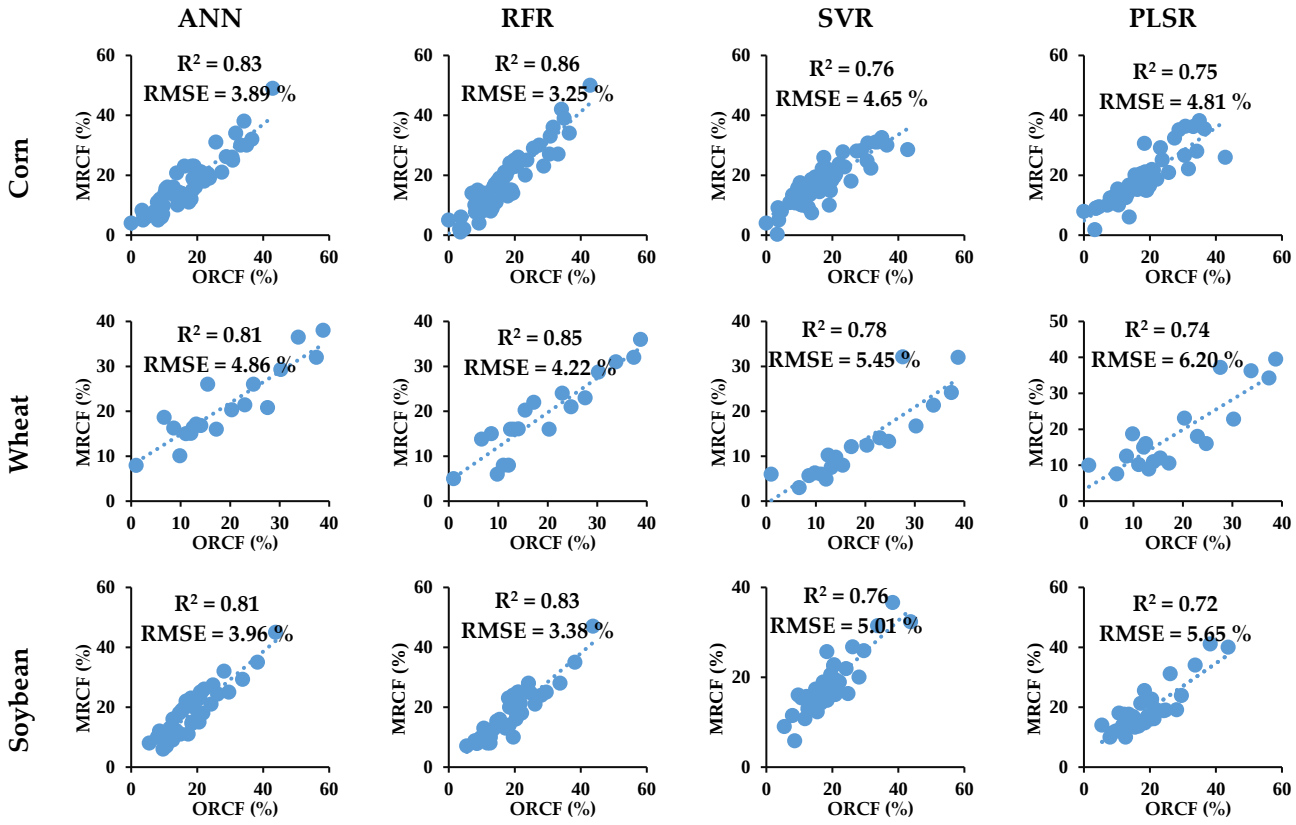

Figure 4. R^2 and RMSE between actual and model RCF based on validation data.

The addition of radar bands to the spectral indices dataset in the corn RCF modeling, caused an increase in the accuracy of RCF estimation using machine learning algorithms (Table 3). Considering the radar bands, the RMSE of corn RCF modeling using ANN, RFR, SVR and PLSR decreased by 0.44, 0.57, 0.54 and 0.30%, respectively. The reduction rates of RMSE for wheat (soybean) were 0.71 (0.37), 0.61 (0.49), 0.55 (0.51) and 0.38 (0.64), respectively.

Table 3. R^2 (RMSE) between the actual and modeled values of the RCF based on different machine learning algorithms, considering spectral indices and radar bands as dependent variables.

	Corn	Wheat	Soybean
ANN	0.85 (3.45)	0.84 (4.15)	0.85 (3.59)
RFR	0.89 (2.68)	0.87 (3.61)	0.89 (2.89)
SVR	0.80 (4.02)	0.80 (4.90)	0.80 (4.50)
PLSR	0.77 (4.51)	0.77 (5.58)	0.75 (5.01)

The R2 between the actual and modeled RCF based on the fusion strategy at the decision level for corn, wheat and soybean crops was 0.92, 0.89 and 0.88, respectively (Figure 5). RMSE values were 1.78, 2.65 and 1.90%, respectively. The error of estimating the RCF of corn, wheat and soybean products based on the proposed strategy was reduced by 0.90, 0.96 and 0.99%, respectively, compared to the results of the best machine learning algorithm.

The RCF map of corn, wheat and soybean crops prepared based on the fusion strategy at the decision level showed that the spatial distribution of the residue varied across the study area (Figure 6). The RCF of three crops varied between 0 and 62%. The RCF of corn on farms located in the eastern parts of the study area was less than the western part. Corn fields located in the northwestern parts of the study area had the highest values of residue. The lowest values of soybean RCF were in farms located in the central parts of the study area. The number of wheat fields in the study area was less than corn and soybean fields. The number of wheat fields with low RCF was lower than wheat fields with high RCF.

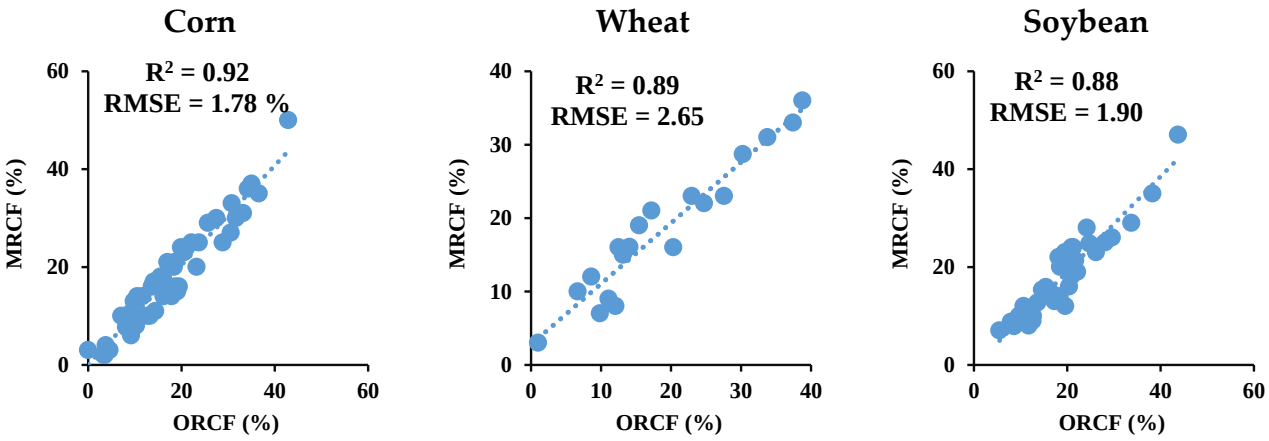

Figure 5. R^2 and RMSE between the actual and modeled RCF based on the fusion strategy at the decision level for corn, wheat and soybean crops. ORCF: observed residual cover fraction; MRCF: modeled residual cover fraction.

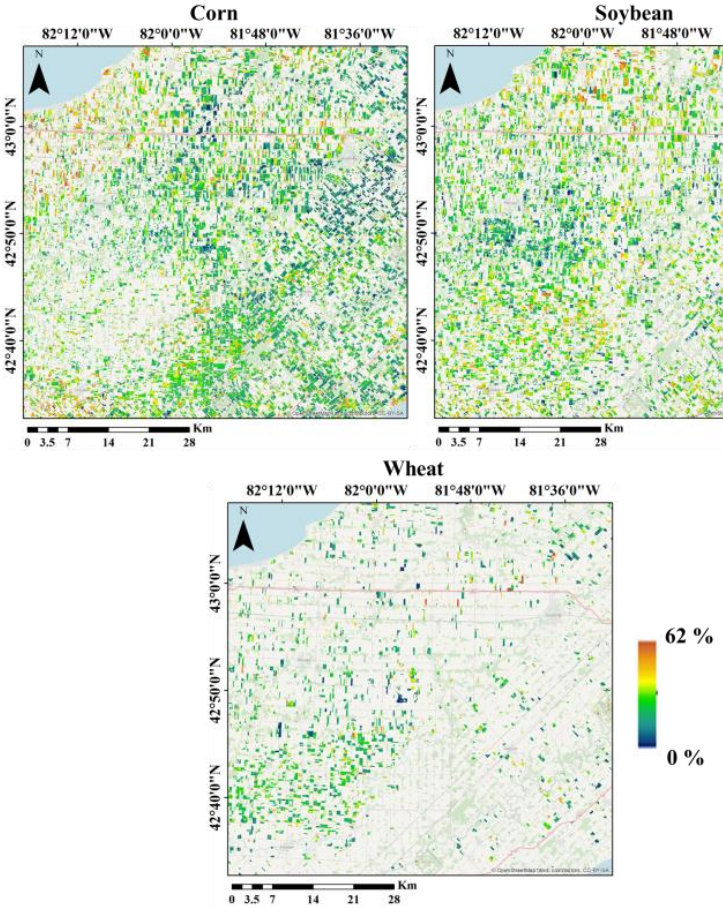

Figure 6. RCF maps for different land crop in the study area.

The mean RCF for corn, wheat and soybean crops in the study area were 18.2%, 19.39% and 17.7%, respectively (Figure 7). RCF was higher in wheat fields than in corn and soybean fields. The values of the standard deviation (Sd) of the RCF for corn, wheat and soybean fields in the study area were 8.3%, 10.23% and 7.4%, respectively. The highest and lowest variation of RCF in this study area was related to wheat and soybean crops. The range of variation in the RCF amount for corn fields as greater than wheat and corn crops.

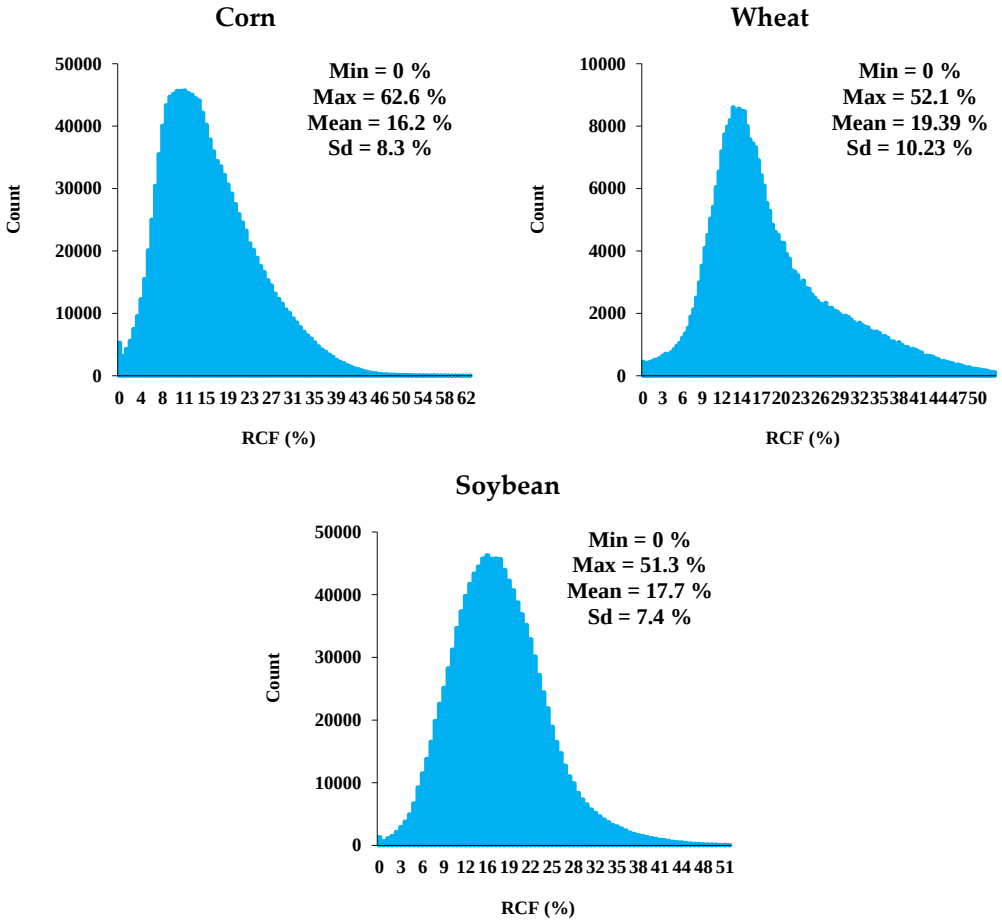

Figure 7. Frequency histogram and statistical parameters of the modeled RCF values of corn, wheat and soybean crops in the study area based on the proposed strategy.

5. Discussion

An accurate RCF map is crucial in agricultural planning and management [19,22,33]. Unlike terrestrial methods, satellite images have high application and high efficiency in preparing these maps due to their extensive spatiotemporal coverage and their low cost [8,29,30,34]. However, the accuracy of different crops' RCF mapping using satellite data is dependent on various factors, including (1) dependent variables used in the modeling process, (2) the quality of calibration and validation data and (3) the algorithms used to construct the appropriate model between the effective variables and the RCF [8,29,33]. In this study, the effect of dependent variables and algorithms used in the RCF modeling process was investigated.

In previous studies, various spectral indices based on satellite images have been provided to prepare this map [29,33]. In a number of these studies, the efficiencies of these spectral indices in RCF modeling were evaluated and compared [8,33]. The results showed a different performance of each of these indicators in different conditions. In this study, it was also shown that the efficiency of a number of indices, such as NDI5, NDI7, NDTI, NDVI, STI, DFI and BAI, was higher than other spectral indices when preparing the RCF map. Even the efficiency of each spectral index in preparing a RCF map for different crops is different. The results showed that for modeling the RCF of corn, wheat and soybean crops, BAI efficiency was higher than other indicators. The BAI is designed to minimize the effect of background soil moisture in the process of RCF estimating. Yue, Tian, Dong and Xu [8] showed that BAI had a high efficiency in estimating the RCF of wheat and maize products by reducing the effect of field soil moisture. In general, the spectral behavior of different crops at different wavelengths and their differences with the background soil and vegetation in the study area can affect the accuracy of the developed spectral indices. NDVI

preparing the receipt map and agricultural products are highly sensitive to vegetation in the study area. As a result, for areas with high vegetation cover, such as forests and pastures, the RCF values are overestimated.

The focus of previous studies has been on the development of optimal indicators and methods in RCF modeling and on the use of reflective bands of satellite images, including Landsat, Sentinel 2, ASTER, etc. [16,30,33,34]. In a limited number of studies, the efficiency of the information obtained from radar images, including RADARSAT, in RCF modeling was evaluated [24]. However, in this study, for the first time, the efficiency of Sentinel 1 VV and VH bands in modeling this parameter was evaluated and compared with the efficiency of spectral indices. The results showed that considering spectral indices and VV and VH radar bands simultaneously as effective variables increases the accuracy of RCF modeling based on machine learning algorithms. Due to the simple and free access to Sentinel 1 images with a high frequency for different agricultural regions around the world, the use of these images in RCF modeling can be of great practical importance. In general, the accuracy of multivariate RCF modeling is higher than univariate modeling using machine learning algorithms. The results of this study showed that the efficiencies of machine learning algorithms in RCF modeling were different to one another. The efficiency of RFR algorithm in modeling this parameter was higher than ANN, SVR and PLSR. Ding, Zhang, Wang, Xie, Wang, Liu and Hall [33] also showed that RFR was highly efficient in RCF mapping.

Each of these algorithms may have high or low performance under different conditions. Therefore, providing an integrated model based on the results of these algorithms can be useful. In various fields, such as land cover classification, improving the spatial resolution of land surface temperature, etc., the strategy of combining the results of different algorithms called fusion at the decision level has been used to improve the modeling accuracy of target variable. In this study, the results of the RCF estimation obtained from different machine learning algorithms were combined based on the degree of importance of each algorithm to improve the modeling accuracy of this variable. The results showed that by implementing the fusion strategy at the decision level, the accuracy of the RCF map was significantly improved.

6. Conclusions

In this study, to improve the accuracy of RCF modeling and mapping, a new strategy based on the fusion of different machine learning algorithms' results at the decision level was developed. The results showed that by considering both spectral indices based on reflective bands and radar bands as dependent variables in machine learning algorithms, the RCF modeling error is reduced by an average of 15%. Among the various machine learning algorithms in RCF modeling, RF accuracy is higher than other algorithms, including ANN, SVR and PLSR. The results of the proposed strategy showed that the integration of the capabilities of different machine learning algorithms increases the accuracy of RCF modeling. With the fusion of the results of different machine learning algorithms at the decision level, the accuracy of RCF modeling for corn, wheat and soybean crops compared to the most optimal algorithm has increased by more than 33, 25 and 34%, respectively. It is suggested that in future studies, the efficiency of deep learning algorithms in RCF modeling be evaluated. It could also be very useful to use the proposed algorithm to prepare a more accurate RCF map in agricultural areas around the world and implement optimal programs to improve the agricultural situation and conserve soil and environmental quality.

Author Contributions: Conceptualization, A.B. and S.F.; methodology, S.F., M.K.F. and A.B.; software, S.F. and M.K.F.; validation, S.F. and M.K.F.; formal analysis, S.F. and M.K.F.; data curation, S.F., M.K.F. and A.B.; writing—original draft preparation, S.F. and M.K.F.; writing—review and editing, A.B.; supervision, A.B.; project administration, A.B.; funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This project is supported by the Canada First Research Excellence Fund (CFREF)–Food from Thought project at the University of Guelph and Natural Sciences and Engineering Research Council of Canada (NSERC) (RGPIN-2014-4100). Funding for open access charge: is from Natural Sciences and Engineering Research Council of Canada (NSERC) (RGPIN-2014-4100).

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon reasonable request.

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

References

1. Birkás, M.; Jolánkai, M.; Gyuricza, C.; Percze, A. Tillage effects on compaction, earthworms and other soil quality indicators in Hungary. *Soil Tillage Res.* **2004**, *78*, 185–196. [CrossRef]
2. Reicosky, D.; Allmaras, R. Advances in tillage research in North American cropping systems. *J. Crop Prod.* **2003**, *8*, 75–125. [CrossRef]
3. Yuce, G.; Pinarbasi, A.; Ozcelik, S.; Ugurluoglu, D. Soil and water pollution derived from anthropogenic activities in the Porsuk River Basin, Turkey. *Environ. Geol.* **2006**, *49*, 359–375. [CrossRef]
4. Jiang, Z.; Bian, H.; Xu, L.; Li, M.; He, N. Pulse effect of precipitation: Spatial patterns and mechanisms of soil carbon emissions. *Front. Ecol. Evol.* **2021**, *9*, 302. [CrossRef]
5. Isson, T.T.; Planavsky, N.J.; Coogan, L.; Stewart, E.; Ague, J.; Bolton, E.; Zhang, S.; McKenzie, N.; Kump, L. Evolution of the global carbon cycle and climate regulation on earth. *Glob. Biogeochem. Cycles* **2020**, *34*, e2018GB006061. [CrossRef]
6. Lal, R.; Kimble, J.; Follett, R. Pedospheric processes and the carbon cycle. In *Soil Processes and the Carbon Cycle*; CRC Press: Boca Raton, FL, USA, 2018; pp. 1–8.
7. Hively, W.D.; Lamb, B.T.; Daughtry, C.S.; Shermeyer, J.; McCarty, G.W.; Quemada, M. Mapping crop residue and tillage intensity using WorldView-3 satellite shortwave infrared residue indices. *Remote Sens.* **2018**, *10*, 1657. [CrossRef]
8. Yue, J.; Tian, Q.; Dong, X.; Xu, N. Using broadband crop residue angle index to estimate the fractional cover of vegetation, crop residue, and bare soil in cropland systems. *Remote Sens. Environ.* **2020**, *237*, 111538. [CrossRef]
9. Raoufat, M.H.; Dehghani, M.; Abdolabbas, J.; Kazemeini, S.A.; Nazemossadat, M.J. Feasibility of satellite and drone images for monitoring soil residue cover. *J. Saudi Soc. Agric. Sci.* **2020**, *19*, 56–64.
10. Page, K.L.; Dang, Y.P.; Dalal, R.C.; Reeves, S.; Thomas, G.; Wang, W.; Thompson, J.P. Changes in soil water storage with no-tillage and crop residue retention on a Vertisol: Impact on productivity and profitability over a 50 year period. *Soil Tillage Res.* **2019**, *194*, 104319. [CrossRef]
11. Alberts, E.E.; Neibling, W.H. Influence of crop residues on water erosion. *Manag. Agric. Residues* **1994**, *13*, 19–44.
12. Cai, W.; Zhao, S.; Wang, Y.; Peng, F.; Heo, J.; Duan, Z. Estimation of winter wheat residue coverage using optical and SAR remote sensing images. *Remote Sens.* **2019**, *11*, 1163. [CrossRef]
13. Zheng, B.; Campbell, J.B.; Serbin, G.; Galbraith, J.M. Remote sensing of crop residue and tillage practices: Present capabilities and future prospects. *Soil Tillage Res.* **2014**, *138*, 26–34. [CrossRef]
14. Najafi, P.; Navid, H.; Feizizadeh, B.; Eskandari, I. Remote sensing for crop residue cover recognition: A review. *Agric. Eng. Int. CIGR J.* **2018**, *20*, 63–69.
15. Dvorakova, K.; Shi, P.; Limbourg, Q.; van Wesemael, B. Soil organic carbon mapping from remote sensing: The effect of crop residues. *Remote Sens.* **2020**, *12*, 1913. [CrossRef]
16. Hively, W.D.; Shermeyer, J.; Lamb, B.T.; Daughtry, C.T.; Quemada, M.; Keppler, J. Mapping crop residue by combining Landsat and WorldView-3 satellite imagery. *Remote Sens.* **2019**, *11*, 1857. [CrossRef]
17. Bégué, A.; Arvor, D.; Bellon, B.; Betbeder, J.; De Abelleyra, D.; Ferraz, R.P.D.; Lebourgeois, V.; Lelong, C.; Simões, M.; Verón, S.R. Remote sensing and cropping practices: A review. *Remote Sens.* **2018**, *10*, 99. [CrossRef]
18. Sonmez, N.K.; Slater, B. Measuring intensity of tillage and plant residue cover using remote sensing. *Eur. J. Remote Sens.* **2016**, *49*, 121–135. [CrossRef]
19. Daughtry, C.S.; Hunt, E.; Doraiswamy, P.; McMurtrey, J. Remote sensing the spatial distribution of crop residues. *Agron. J.* **2005**, *97*, 864–871. [CrossRef]
20. Zhuang, Y.; Chen, D.; Li, R.; Chen, Z.; Cai, J.; He, B.; Gao, B.; Cheng, N.; Huang, Y. Understanding the influence of crop residue burning on PM2. 5 and PM10 concentrations in China from 2013 to 2017 using MODIS data. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1504. [CrossRef]
21. Quemada, M.; Hively, W.; Daughtry, C.; Lamb, B.; Shermeyer, J. Improved crop residue cover estimates obtained by coupling spectral indices for residue and moisture. *Remote Sens. Environ.* **2018**, *206*, 33–44. [CrossRef]
22. Zheng, B.; Campbell, J.B.; de Beurs, K.M. Remote sensing of crop residue cover using multi-temporal Landsat imagery. *Remote Sens. Environ.* **2012**, *117*, 177–183. [CrossRef]
23. Yue, J.; Fu, Y.; Guo, W.; Feng, H.; Qiao, H. Estimating fractional coverage of crop, crop residue, and bare soil using shortwave infrared angle index and Sentinel-2 MSI. *Int. J. Remote Sens.* **2022**, *43*, 1253–1273. [CrossRef]
24. McNairn, H.; Wood, D.; Gwyn, Q.; Brown, R.; Charbonneau, F. Mapping tillage and crop residue management practices with RADARSAT. *Can. J. Remote Sens.* **1998**, *24*, 28–35. [CrossRef]

25. Bannari, A.; Pacheco, A.; Staenz, K.; McNairn, H.; Omari, K. Estimating and mapping crop residues cover on agricultural lands using hyperspectral and IKONOS data. *Remote Sens. Environ.* **2006**, *104*, 447–459. [CrossRef]
26. Qi, J.; Marsett, R.; Heilman, P.; Bieden-bender, S.; Moran, S.; Goodrich, D.; Weltz, M. RANGES improves satellite-based information and land cover assessments in southwest United States. *Eos Transact. Am. Geophys. Union* **2002**, *83*, 601–606. [CrossRef]
27. Gelder, B.; Kaleita, A.; Cruse, R. Estimating mean field residue cover on midwestern soils using satellite imagery. *Agron. J.* **2009**, *101*, 635–643. [CrossRef]
28. Van Deventer, A.; Ward, A.; Gowda, P.; Lyon, J. Using thematic mapper data to identify contrasting soil plains and tillage practices. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 87–93.
29. Yue, J.; Tian, Q.; Dong, X.; Xu, K.; Zhou, C. Using hyperspectral crop residue angle index to estimate maize and winter-wheat residue cover: A laboratory study. *Remote Sens.* **2019**, *11*, 807. [CrossRef]
30. Serbin, G.; Hunt, E.R.; Daughtry, C.S.; McCarty, G.W.; Doraiswamy, P.C. An improved ASTER index for remote sensing of crop residue. *Remote Sens.* **2009**, *1*, 971–991. [CrossRef]
31. Bocco, M.; Sayago, S.; Willington, E. Neural network and crop residue index multiband models for estimating crop residue cover from Landsat TM and ETM+ images. *Int. J. Remote Sens.* **2014**, *35*, 3651–3663. [CrossRef]
32. McNairn, H.; Protz, R. Mapping corn residue cover on agricultural fields in Oxford County, Ontario, using Thematic Mapper. *Can. J. Remote Sens.* **1993**, *19*, 152–159. [CrossRef]
33. Ding, Y.; Zhang, H.; Wang, Z.; Xie, Q.; Wang, Y.; Liu, L.; Hall, C.C. A comparison of estimating crop residue cover from sentinel-2 data using empirical regressions and machine learning methods. *Remote Sens.* **2020**, *12*, 1470. [CrossRef]
34. Jin, X.; Ma, J.; Wen, Z.; Song, K. Estimation of maize residue cover using Landsat-8 OLI image spectral information and textural features. *Remote Sens.* **2015**, *7*, 14559–14575. [CrossRef]
35. Yue, J.; Tian, Q.; Tang, S.; Xu, K.; Zhou, C. A dynamic soil endmember spectrum selection approach for soil and crop residue linear spectral unmixing analysis. *Int. J. Appl. Earth Observ. Geoinf.* **2019**, *78*, 306–317. [CrossRef]
36. Yue, J.; Tian, Q. Estimating fractional cover of crop, crop residue, and soil in cropland using broadband remote sensing data and machine learning. *Int. J. Appl. Earth Observ. Geoinf.* **2020**, *89*, 102089. [CrossRef]
37. Wang, G.; Wang, J.; Zou, X.; Chai, G.; Wu, M.; Wang, Z. Estimating the fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil from MODIS data: Assessing the applicability of the NDVI-DFI model in the typical Xilingol grasslands. *Int. J. Appl. Earth Observ. Geoinf.* **2019**, *76*, 154–166. [CrossRef]
38. Hansen, P.; Schjoerring, J. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542–553. [CrossRef]
39. 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.* **2015**, *108*, 273–290. [CrossRef]
40. Baret, F.; Weiss, M.; Lacaze, R.; Camacho, F.; Makhmara, H.; Pacholcyzk, P.; Smets, B. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. *Remote Sens. Environ.* **2013**, *137*, 299–309. [CrossRef]
41. Hagan, M.T.; Menhaj, M.B. Training feedforward networks with the Marquardt algorithm. *IEEE Transact. Neural Netw.* **1994**, *5*, 989–993. [CrossRef] [PubMed]
42. Vapnik, V.; Golowich, S.; Smola, A. Support vector method for function approximation, regression estimation and signal processing. *Adv. Neural Inf. Process. Syst.* **1996**, *9*, 281–287.
43. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000.
44. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
45. Shah, S.H.; Angel, Y.; Houborg, R.; Ali, S.; McCabe, M.F. A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sens.* **2019**, *11*, 920. [CrossRef]
46. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm.* **2016**, *114*, 24–31. [CrossRef]