



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Tomppo, Erkki; Antropov, Oleg; Praks, Jaan Cropland classification using Sentinel-1 time series

Published in: Remote Sensing

*DOI:* 10.3390/rs11212480

Published: 01/11/2019

Document Version Publisher's PDF, also known as Version of record

Published under the following license: CC BY

Please cite the original version:

Tomppo, E., Antropov, O., & Praks, J. (2019). Cropland classification using Sentinel-1 time series: Methodological performance and prediction uncertainty assessment. *Remote Sensing*, *11*(21), Article 2480. https://doi.org/10.3390/rs11212480

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.



# Article Cropland Classification Using Sentinel-1 Time Series: Methodological Performance and Prediction Uncertainty Assessment

# Erkki Tomppo <sup>1</sup>,\*, Oleg Antropov <sup>2</sup> and Jaan Praks <sup>1</sup>

- <sup>1</sup> Department of Electronics and Nanoengineering, Aalto University, P.O. Box 11000, FI-00076 AALTO, 02150 Espoo, Finland; jaan.praks@aalto.fi
- <sup>2</sup> VTT Technical Research Centre of Finland, P. O. Box 1000, FI-00076 VTT, 02150 Espoo, Finland; oleg.antropov@vtt.fi
- \* Correspondence: erkki.tomppo@aalto.fi

Received: 19 September 2019; Accepted: 20 October 2019; Published: 24 October 2019



Abstract: Methods based on Sentinel-1 data were developed to monitor crops and fields to facilitate the distribution of subsidies. The objectives were to (1) develop a methodology to predict individual crop species or or management regimes; (2) investigate the earliest time point in the growing season when the species predictions are satisfactory; and (3) to present a method to assess the uncertainty of the predictions at an individual field level. Seventeen Sentinel-1 synthetic aperture radar (SAR) scenes (VV and VH polarizations) acquired in interferometric wide swath mode from 14 May through to 30 August 2017 in the same geometry, and selected based on the weather conditions, were used in the study. The improved k nearest neighbour estimation, ik-NN, with a genetic algorithm feature optimization was tailored for classification with optional Sentinel-1 data sets, species groupings, and thresholds for the minimum parcel area. The number of species groups varied from 7 to as large as 41. Multinomial logistic regression was tested as an optional method. The Overall Accuracies (OA) varied depending on the number of species included in the classification, and whether all or not field parcels were included. OA with nine species groups was 72% when all parcels were included, 81% when the parcels area threshold (for incorporating parcels into classification) was 0.5 ha, and around 90% when the threshold was 4 ha. The OA gradually increased when adding extra Sentinel-1 scenes up until the early August, and the initial scenes were acquired in early June or mid-May. After that, only minor improvements in the crop recognition accuracy were noted. The ik-NN method gave greater overall accuracies than the logistic regression analysis with all data combinations tested. The width of the 95% confidence intervals with ik-NN for the estimate of the probability of the species with the largest probability on an individual parcel varied depending on the species, the area threshold of the parcel and the number of the Sentinel-1 scenes used. The results ranged between 0.06–0.08 units (6–8% points) for the most common species when the Sentinel-1 scenes were between 1 June and 12 August. The results were well-received by the authorities and encourage further research to continue the study towards an operational method in which the space-borne SAR data are a part of the information chain.

**Keywords:** crops; classification; synthetic aperture radar; C-band; k-NN; genetic algorithm; multinomial logistic regression; boreal region; ESA Sentinel-1

# 1. Introduction

The Introduction section discusses the importance of studying Sentinel-1 time series data in the mapping of crops, and relates our work with the state-of-the-art in the context of the study.



The introduction is divided into two sub-sections. Firstly, the background and motivation are discussed and the objectives of the study are given. Examples of the earlier studies using synthetic aperture radar (SAR) data in crop classification are reviewed further in the second sub-section.

#### 1.1. Background and Motivation

The Land Parcel Identification System (LPIS) of the European Commission is used for registration of agricultural reference parcels considered eligible for annual payments of European Common Agricultural Policy (CAP) subsidies to farmers [1]. LPIS and the subsidies presume a system of controlling the quality of the management for each parcel. The control was carried out through field visits and using aerial photographs, and both methods are quite expensive. The European Union is interested in a more cost-efficient way of controlling the management of crop fields, e.g., the utilization of the space-borne remote sensing data. Reliable image acquisition is important in operative applications. Clouds and haze often prevent the acquisition of applicable optical area images in such a way that, e.g., even a single cloud free Sentinel-2 can not be acquired from the growing season from each region in most of the European countries. Space-borne SAR data are thus the only remote sensing data source suitable for rapid and near real-time assessment of crop fields during the growing season in European countries. Since the advent of the Copernicus programme Sentinel-1 SAR satellites that are capable to provide repeated acquisitions every six days (with two satellites), the potential for continuous monitoring of crops was established.

One key prerequisite for an operational method is a possibility to recognize the management operations and species early enough during the growing season. This requirement arises from the need to make decisions and provide the subsidy payments early enough. It is also important to know the uncertainty of the species prediction at an individual parcel level. An operational system may still need several auxiliary information sources, e.g., provided by methods using data acquired by space-borne or airborne sensors, and even field visits. Uncertainties of the species prediction at an individual field parcel level are therefore important to decide whether other information is needed or not, in addition to the predictions based on satellite observations.

To date, several approaches to use space-borne SAR data to monitor agricultural regimes, e.g., for detecting ploughing, moving, or sowing activities were reported in the literature (see Section 1.2). Studies in crop classification with fully polarimetric C-band SAR and optical area data have also been reported (Section 1.2). Many studies concentrated on tree orchards, grapes, or sugarcane. A few studies exist with detailed corn species recognition particularly using Sentinel-1 data. Section 1.2 reviews the aforementioned literature.

The objectives of this research are (1) to develop a method for recognition of individual crop species, or species groups, and management regimes; (2) to investigate the earliest time point in the growing season when the species predictions will be satisfactory; and (3) to present a method to assess the uncertainty of the crop species prediction at an individual parcel level. The improved k nearest neighbour method (ik-NN) with the feature optimization using a genetic algorithm was modified for estimation [2,3]. The uncertainty assessment of the parcel level predictions was based on the probabilities of the crop species that are outputs of the ik-NN method. A parametric method to estimate the confidence intervals of the largest probabilities was developed. Multinomial logistic regression was employed as an optional method for comparison with the ik-NN method. The methods and results were demonstrated on a test site located in Southern Finland, in Eura muninicipality. The number of crop field parcels was 10,287 with a total study area of 24,503.5 hectares.

#### 1.2. SAR Data in Crop Classification

SAR technology offers many advantages for crop monitoring activities, given that the radiowaves are generally unaffected by the presence of clouds and haze. The presence of multiple SAR instruments increases the opportunity to build temporally rich data sets in the early weeks of crop growth. Due to

the size of the crops, C-band is considered to be a wavelength most suitable for crop monitoring applications with various SAR techniques [4–7].

Multitemporal SAR data sets can be used to exploit changes in the crop structure as crops transition from one growth stage to another and thus to separate one crop type from another. In particular, SAR backscatter at C-band is very sensitive to changes in structure during seed and fruit development, stages which occur later in the growing season [8]. This means that high quality crop inventories can be readily delivered close to the end of the growing season. For mono-temporal SAR based classification of crops, in order to maximize classification accuracies, SAR image acquisition is recommended to be planned during seed and reproductive phenology phases [8,9]. With multitemporal SAR approaches, it remains still to be seen how early in the growing season the crop inventories can be produced, and what accuracies can be achieved as the growing season develops.

To date, the majority of research on cropland classification was done using multiparametric SAR data [10]. This includes mostly polarimetric and multitemporal SAR data [8,11–22], as well as multi-frequency SAR and fusion of satellite optical and SAR data [7,23]. In addition, the potential of interferometric SAR approaches was evaluated along with SAR backscatter data in crop monitoring [6,24,25]. While often SAR data acquired at X-band, and, to lesser extent, L-band [26], can still be used in crop monitoring and change detection, satellite C-band SAR data are assumed to be the preferred choice for crop mapping, particularly for agricultural areas with low vegetation [8]. There is a trade-off between the polarimetric information and the multitemporal information; however, results obtained using the multitemporal information tend to be better [5]. On the other hand, when only a few acquisitions are available, the polarimetric mode may perform better than the single- and dual polarization modes.

In Loosvelt et al. [15], the Random Forests approach was used for the probabilistic mapping of vegetation using fully polarimetric L- and C-band EMISAR data to assess and analyze classification uncertainty based on the local probabilities of class membership. Results showed that multi-configuration in the dataset decreases the classification uncertainty for the different agricultural crops when compared to the single-configuration alternative. Furthermore, the uncertainty assessment revealed lower confidence for the classification of (mixed) pixels at the field edges.

Currently, Sentinel-1 series satellites of European Space Agency (ESA) are the key resource for supplying freely available SAR data. The data are not fully polarimetric, but scenes are acquired every 6–12 days (depending on the geographical region) and can continuously cover the whole growing season. This enables monitoring at different times and allows access to the whole phenological cycle of the crops. It also brings forward the aspect of temporal variation of backscatter for different types of crops, as these can be useful in differentiating between crops.

Most recently, due to freely available ESA Sentinel-1 C-band SAR data, many research efforts concentrate on various crop classification with advanced methods using Sentinel-1 time series [21,23,26–29]. Likewise, we use SAR data and review several classification approaches in more detail.

In [26], repeat-pass Sentinel-1 data were used over North Dakota to classify individual agricultural land-cover types. In this approach, the times series forms the basis of a classification algorithm, where individual pixels are compared against a model of average crop backscatter response and classified as the crop with the least difference from the model.

Similarly, in [21], the temporal intensity models were used in K-means clustering approach for crop classification (4 to 5 types of crops including corn, soybean, rice, peanut, lotus, and grass), gaining accuracies on the order of 75 to 90% for different crops. Xu et al. [21] used Sentinel-1 time-series data to construct temporal intensity models employing K-means clustering. The introduced spectral similarity value measure (SSV) seemed to work better than the decision tree and the Bayesian classifier methods. The training data were produced through visual interpretation. The number of the species groups were five and four in the two study area. Overall accuracies were as high as 90–92%. The average sizes of the field parcels were not given.

Analysis of Sentinel-1 time series along with optical Sentinel-2 data in [23] has shown that SAR backscatter and NDVI (normalized difference vegetation index) may be complementary for agricultural applications. Particularly, the VH-to-VV ratio at C-band was shown to be a good discriminator and notably suitable for crop applications. Veloso et al. [23] analyzed and interpreted the temporal trajectory of remote sensing data for a variety of winter and summer crops that are widely cultivated in the world (wheat, rapeseed, maize, soybean and sunflower). Sentinel-1 data were used and compared to the temporal variation with NDVI derived from Sentinel-2 type optical data. The performances of different features in assessing the phenological stages of the crops were analyzed. The purpose was to also investigate the possibilities of estimating physical parameters, such as fresh biomass and green area index (GAI). The authors concluded that the dense time series allowed the capture of short phenological stages and to describe various crop developments. The authors concluded also that a better understanding of SAR backscatter and NDVI temporal behaviors under contrasting agricultural practices and environmental conditions help upcoming studies related to crop monitoring based on Sentinel-1 and Sentinel-2, such as dynamic crop mapping and biophysical parameters estimation.

In earlier work [27], a set of five Sentinel-1 scenes were used along with one multispectral Sentinel-2 image to classify six crop types (beans, beetroot, grass, maize, potato, and winter wheat). To assess the potential of crop classification with common off the shelf supervised learning models, a benchmarking using four different approaches (kernel-based extreme learning machine, multilayer feedforward neural networks, random forests, and support vector machine) was implemented. The first approach performed best with very high overall accuracy of 96.8%. Evaluation of the sensitivity of classification models and relative importance of data types using data-based sensitivity analysis showed that the two most important scenes (explanatory variables) was VV-pol channel of one of Sentinel-1 scenes and band 4 of the Sentinel-2 scene, indicating complementary of optical and SAR data in crop classification.

In yet another methodological study on crop-type mapping using a sequence of Sentinel-1 images [28], the dynamic conditional random fields were shown to effectively capture spatio-temporal phenological information for various crops. This kind of variation appears inherent in images and can be used for crop classification purposes. Not surprisingly, the final classification performance was higher for the multitemporal stack than for any of the separate scenes. The suggested approach was also shown to perform better than the conventional maximum likelihood approach with multitemporal images aggregated as composite bands.

Several conventional machine learning approaches were compared with deep learning approaches subject to crop classification performance in [30]. Ndikumana et al. [30] demonstrated the performances of the traditional k-NN, random forest (RF) and support vector machine (SVM) in a test site in France with eleven species categories and 25 Sentinel-1 scenes, ranging from from May 2017 to September 2017. The two most dominant categories, rice and wheat, comprised more than 50% of the area of the test sites. The average parcel area of rice was 3.3 hectares and that of wheat 1.7 hectares. The performances of the traditional methods were compared to those of two deep learning techniques, deep recurrent neural network (RNN)-based classifiers. Interestingly, while deep learning approaches proved superior, the performance in terms of overall accuracies was not so strong: 86–87% with more traditional methods and 89–90% with deep learning models. The k-NN method and its optimized version ik-NN has some advantages over the other methods, where we selected it as the main method here [31]. These advantages are discussed in Section 2.3.

While there is a considerable amount of research on the use of Sentinel-1 time series in crop mapping, considerable gaps seem to be present. Most notably, the prediction uncertainty assessment is typically missing. Furthermore, the connection of classification performance to the specific stage of phenological cycle of crops is barely established. However, examples of the earlier studies and references can be found, e.g., in Veloso et al. [23] and Song and Wang [32]. The latter one deals with one species in its phenology. In addition, the reported classification accuracy levels were not always satisfactory, particularly for a large number of crops, lacking in depth sensitivity analysis of

the classification performance on the number of crops used, size of the parcels, and the number of Sentinel-1 scenes employed in the experiment. In addition, in the majority of the studies, no selection of most suitable Sentinel-1 scenes was performed, and the study area is often very small. In this study, we address several of these issues.

# 2. Materials and Methods

The methods were developed and the results demonstrated and discussed using one test site from southwest Finland, located in the Eura municipality (Figure 1), crop data from the test site, provided by the Finnish Food Authority [33] and seventeen 17 Sentinel-1 scenes.

# 2.1. Sentinel-1 Scenes

Seventeen ground-range detected (GRD) Sentinel-1 images acquired in IW mode during the period 14 May–30 August were downloaded from ESA Open Access Hub and orthorectified.

Image orthorectification was done using a digital elevation model available from the National Land Survey of Finland. Scenes were aggregated in azimuth and range to obtain images with pixel dimensions approximately corresponding to the 20 m grid spacing. The bilinear interpolation method was used for resampling in connection with the ortho-rectification. Radiometric normalization of intensity was done using a projected pixel area based approach to minimize the effect of the topography [34,35]. The scenes were further re-projected to the ERTS89/ETRS-TM35FIN projection (EPSG:3067) and resampled to a final pixel size of 16 m.

The averages and standard deviations of the intensities of the both VV and VH polarizations were calculated for each agriculture parcel, as well as the ratios. Multi-temporal imagery shows the progress of the vegetation during the growing season and is necessary when identifying crop species. It also makes it possible to reduce the effect of the random scattering (speckle) on the predictions and error estimates and also to utilize the variation of the data acquisition conditions in the estimation through using multifaceted information. The dates of the images are shown in Table 1. An example of multitemporal color coded composition of Sentinel-1 time series at different stages of crop development are shown in Figure 2.



Figure 1. Study area location in the southwestern part of Finland.





**Figure 2.** An example of multitemporal Sentinel-1 colour composition, upper left image, Red, Green. Blue (RGB) composite of scenes acquired on 1 June (R), 7 June (G) and 13 June 13 (B); upper right image—RGB composite of scenes acquired on 25 July (R), 6 August (G) and 12 August (B); bottom image is the enlarged fragment of the upper right image showing the parcel boundaries in white color. The backscatter coefficient of VH polarization of each scene is used.

Image	Date	Mode	Polarization
1	14 May 2017	IW	VV, VH
2	26 May 2017	IW	VV, VH
3	1 June 2017	IW	VV, VH
4	7 June 2017	IW	VV, VH
5	13 June 2017	IW	VV, VH
6	19 June 2017	IW	VV, VH
7	25 June 2017	IW	VV, VH
8	1 July 2017	IW	VV, VH
9	7 July 2017	IW	VV, VH
10	13 July 2017	IW	VV, VH
11	19 July 2017	IW	VV, VH
12	25 July 2017	IW	VV, VH
13	6 August 2017	IW	VV, VH
14	12 August 2017	IW	VV, VH
15	18 August 2017	IW	VV, VH
16	24 August 2017	IW	VV, VH
17	30 August 2017	IW	VV, VH

Table 1. List of Sentinel-1 scenes used in the study.

# 2.2. Field Parcel Data

The Eura site included 10,287 agriculture parcels with the total area amounting to 24,503.5 hectares. The number of different crop species or land management categories in the 2017 data was 119. The species data together with the boundary data of the agriculture parcel were provided by the Finnish Food Agency [33]. Several species groupings were tested. The grouping with nine categories was used in the analyses and is shown in Table 2 together with the numbers of the parcels by species groups as well as the averages and standard deviations of the areas of the parcels by species groups. The final grouping was selected in collaboration with the agriculture experts from the Finnish Food Authority taking into account also the development of the crop species within the growing season [33]. The ranges of the dates of the Sentinel-1 scenes were also selected in collaboration with the same authorities and cover the growing seasons of the species. Overall, the average parcel area was 2.38 hectares and the standard deviation of the areas 2.83 hectares.

Code	Species Group	Number of the Parcels	Average Areas	Standard Deviation of the Areas
1	Spring wheat, Oats	2994	2.809	2.889
2	Broad bean, Vicia faba	104	2.742	2.901
3	Potato, sugar beet	1373	2.559	2.512
4	Spring Oilseed Rape	218	3.320	3.190
5	Pea, beetroot and some other	726	1.852	2.444
6	Grass, Hay	928	1.965	2.275
7	Winter wheat, Winter barley	295	4.121	3.919
8	Feed barley, Malting barley	1858	3.061	3.446
9	Shelter belt, Nature management field	1791	0.837	1.208
Total	All parcels	10,287	2.382	2.827

**Table 2.** The nine species groups, the codes used in the analyses together with number of the parcels, averages, and standard deviations of the areas in the test site.

#### 2.3. Methods

Commonly used methods when classifying multinomial variable observations with remote sensing data are, in addition to discriminant analysis, logistic regression and neural network based methods, Random Forests and Support Vector Machine (e.g., [36,37]). Discriminant analysis and logistic regression as parametric methods presume the acceptance of the distributional properties of the data.

Several studies have been published with all these methods. Nearest neighbour techniques have been employed since the early 1950s in pattern recognition. Nearest neighbour methods became popular in forest and other environmental applications since they were launched for forest inventory applications in the early 1990s in Finland (e.g., [31,38]). The feature optimization with a heuristic genetic algorithm approach for k-NN for forest remote sensing purposes was developed by Tomppo and Halme [3]. The advantages of k-NN include, in addition to being non-parametric and thus not presuming any distributional properties, (1) a possibility to simultaneous estimation of several parameters, (2) to keep the dependence structure of the variables on the training data also in the estimates, and (3) to keep the distribution of training data in small area estimates [31,38]. The non-parametric property means suitability to problems with complex dependencies between dependent and explanatory variables. Keeping the dependencies between different dependent variables is important when dealing with multivariate data, e.g., if we would be interested in estimating crop species and yield by species. The small area estimates can be derived directly from the training data using the weights of the training observations and thus avoiding the typical tendency towards mean that is present in many other methods. The k-NN methods suit thus well for operational applications [31].

Two methods were tested keeping in mind the applicability for operational work. The improved ik-NN estimation method with a genetic algorithm feature optimization was tailored for the study with optional Sentinel-1 data sets, species groupings, and thresholds for the minimum parcel area (cf. [2]). The genetic algorithm has turned out to suit complex optimization problems. It is popular in economic and ecological phenomena and machine learning (e.g., [39]). However, it often needs fine-turning and tailoring for different applications. Multinomial logistic regression for categorical variables were tested as an optional method. The number of species groups varied from 7 to as large as 41, the latter one requested by the agricultural authorities for the Eura site. Leave-one-out cross validation and splitting the data to training data (2/3) and validation data (1/3) were used in result validation. The flowchart of the work-flow of the study is shown in Figure 3. The first analyses were conducted with both methods. Only the ik-NN method was used in the deeper analyses because it turned out to give smaller prediction errors.



Figure 3. The flowchart of the study.

# 2.3.1. SAR Metrics Used in the Crop Species Assessment

The SAR features were calculated for each field parcel to remove the speckle effect and of the radiometric variation within the homogeneous parcel for both polarizations. From the pixel level intensities  $(I_{s,i}^{k,p})$ , the following features were calculated for each parcel, *s*, for both polarizations *p* and for each image, *k* 

(a) averages,

$$10\log_{10}\overline{I_s^{k,p}} = 10\log_{10}\frac{\sum_{i=1}^{n_s}I_{s,i}^{k,p}}{n_s}, k = 1, \dots, 17, p \in \{\text{VV}, \text{VH}\},$$
(1)

where  $n_s$  is the number of the pixels on parcel s,

(b) standard deviations

$$\sqrt{\sum_{i=1}^{n_s} (10\log_{10}I_{s,i}^{k,p} - \overline{10\log_{10}I_s^{k,p}})/(n_s - 1)}, k = 1, \dots, 17, p \in \{\text{VH}, \text{VV}\},$$
(2)

(c) ratios

$$1/n_s \sum_{i=1}^{n_s} I_{s,i}^{k_{1,p}} / I_{s,i}^{k_{2,p}}, k_1 = 2, 3, 4 \text{ and } k_2 = 5, 6, 7, 8, 9.$$
(3)

Several different sets of predictor variables were tested, parcel-averaged intensities (Equation (1)), standard deviations (Equation (2)), and different ratio features (Equation (3)).

#### 2.3.2. ik-NN Method for Crop Species Identification

The well-known k-NN estimation method was tailored for and employed in crop species prediction. The weights for the features selected were calculated using the genetic algorithm [3,31] and its variant for categorical variables [2]. This k-NN method is called the ik-NN method here. The advantages of the ik-NN method is the weighting of the explanatory variables based on their importance in prediction and thus smaller prediction errors compared to the ordinary k-NN method; thus, it is called improved.

Let's recall the main features of the ik-NN estimation with the genetic algorithm in the feature weighting. Denote the *k* nearest feasible parcel by  $i_1(p), \ldots, i_k(p)$  when the distance is calculated in the feature space. The weight  $w_{i,p}$  of parcel *i* to parcel *p* is defined as

$$w_{i,p} = \frac{1}{d_{p_{i},p}^{t}} \bigg/ \sum_{j \in \{i_{1}(p), \dots, i_{k}(p)\}} \frac{1}{d_{p_{j},p}^{t}}, \quad \text{if and only if } i \in \{i_{1}(p), \dots, i_{k}(p)\}$$

$$= 0 \quad \text{otherwise.}$$
(4)

The value of *k* was fixed to be 5 after preliminary tests using the overall accuracy as the criteria. The distance weighting power *t* is a real number, usually  $t \in [0, 2]$ . The value t = 1 was used here. A small quantity, greater than zero, is added to *d* when d = 0 and  $i \in \{i_1(p), ..., i_k(p)\}$ .

The distance metric *d* employed was

$$d_{p_j,p}^2 = \sum_{l=1}^{n_f} \omega_l^2 (f_{l,p_j} - f_{l,p})^2,$$
(5)

where

 $f_{l,p}$  is the *l*th SAR feature variable of parcel *p*,

 $f_{l,p_i}$  is the *l*th SAR feature variable of the nearest neighbour *j* of parcel *p*,

 $n_f$  the number of SAR feature variables,

and  $\omega_l$  the weight for the *l*th SAR feature variable.

The values of the elements  $\omega_l$  of the weight vector  $\omega$  were selected with a genetic algorithm. The details of the genetic algorithm employed are given in [3] and the modification to categorical variables in [2].

The fitness function for the categorical variables to be minimized with respect to  $\omega$  vector was

$$f[\boldsymbol{\omega},\boldsymbol{\gamma},\boldsymbol{B}(\mathbf{X}) = \sum_{j=1}^{n_m} \gamma_j [1 - B_j(\mathbf{X}_j)], \tag{6}$$

where  $\gamma_j > 0$  is a user defined coefficient,  $\mathbf{X}_j$  an error matrix,  $B_j$  the accuracy measure with response variable *j* whose classes are to be predicted,  $n_m$  is the number of response variables to be considered in

the optimization procedure, and  $\omega$  is the weight vector to be optimized (Equation (5)). The number of generations in the genetic algorithm optimization was selected to be 40 after the tests.

For categorical variables, the mode or median of the predicted classes for the nearest neighbors can be used as a prediction instead of a weighted average as is used for continuous variables. For this study, the mode gave more accurate results than the median, consistent with the earlier investigations [2]. The predicted category is the category that has the greatest sum of the weights,  $\omega_{i,p}$ , when summed up by classes over the k nearest neighbors. In theory, equal sums are very rare when real value weights are used; in fact, the probability is zero if rounding is not considered. In cases of equal sums for two or more classes, one class is selected randomly from among those with the greatest sum. This method was used for predicting the categorical variable obtaining the values, the value being the species or species group number.

#### 2.3.3. Multinomial Logistic Regression Method

Multinomial logistic regression was tested as one optional estimation method. The probability of the species group k on parcel p was estimated using the model

P(species group on parcel 
$$p = k | \mathbf{x}_p$$
) =  $\frac{e^{\beta_k \mathbf{x}_p}}{1 + \sum_{l=1}^{L-1} e^{\beta_k \mathbf{x}_p}}, k = 1, ..., L - 1$  and  
=  $\frac{1}{1 + \sum_{l=1}^{L-1} e^{\beta_k \mathbf{x}_p}}, k = L,$  (7)

where  $f(k, p) = \beta_k \mathbf{x}_p$  is a linear predictor function,  $\beta_k$  the vector of the regression coefficient associated with species group k,  $\mathbf{x}_p$  a vector the set of the explanatory variables associated with observation (parcel) p, and L the number of the species group, here 9. The data was split to the training data and validation data also in this case, two thirds of the parcels and one third of the parcels, respectively.

#### 2.3.4. Confidence Intervals of Probabilities for Individual Observations Using ik-NN

The uncertainty of the category prediction of an individual field parcel is important when assessing which field parcel needs additional efforts to get reliable enough information for decision-making. In order to do this, the confidence intervals of the predicted probability of the category for an individual field parcel are needed.

We propose the following procedure to assess the uncertainty of an individual prediction. The k-NN estimation, and its improved version ik-NN, produce probabilities for the predicted category on parcel p. These probabilities can be calculated using the weights  $w_{i,p}$  (Equation (4)) as follows

$$\widetilde{prob(k)}_p = P(\operatorname{category}(p) = k) = \sum_{i \in I_p} w_{i,p} \operatorname{Ind}_{(cat(i)=k)},$$
(8)

where *k* is the mode category based on the largest sum of the weights  $w_{i,p}$  by categories on parcels  $i \in I_p$  and  $Ind_{cat(i)=k}$  is an indicator function of the category in parcel *i*. The confidence intervals for the probabilities of the mode for the individual parcels were calculated using a linear model

$$\widehat{prob(k)}_p = a + \sum_{l=1}^{n_f} b_l f_{l,p} + c \cdot k + \varepsilon,$$
(9)

where  $f_{l,p}$  are the SAR features (Equation 5), k is the predicted crop category, a categorical variable (factor), and a, b, and c are the regression coefficients to be estimated (b being a vector), and  $\epsilon$  a normally distributed random error.

The confidence intervals of the predictions were calculated in a normal way using the estimator

$$\hat{V}_f = s^2 \, \mathbf{x}_0 \, (\mathbf{X}' \mathbf{X})^{-1} \, \mathbf{x}_0' + s^2 \,, \tag{10}$$

for the variance for the individual prediction with a predictor vector of  $x_0$ , residual sum of  $s^2$  and design matrix X consisting of the feature vectors f and predicted categories k.

#### 3. Results

#### 3.1. Classification Performance Using ik-NN and Logistic Regression

The results were calculated using several subsets of the Sentinel-1 scenes and also several minimum area thresholds of the agricultural field parcels. The data sets were split into training data (2/3) and validation data (1/3) both with the ik-NN method and the multinomial logistic regression. Two starting time points of the scenes were tested, the first scenes from 1 June, and the first scenes from 14 May. The Sentinel-1 data accumulation period varied with the ik-NN method. Only the full Sentinel-1 data sets were tested with the multinomial logistic regression, starting either on 1 June or 14 May. The area thresholds varied from zero hectares to four hectares in the first results with the both methods (Tables 3 and 4).

The overall accuracy OA with the ik-NN method employing the genetic algorithm and nine species groups was 71–73% when all parcels were included and the Sentinel-1 scenes until mid August, or beyond that, around 80% when the parcel area threshold (for incorporating parcels into classification) was 0.5 hectares, and around 89–90% when the parcel area threshold was two to four hectares (Tables 3 and 4).

The OAs with the multinomial logistic regression were somewhat smaller than with the ik-NN method with all minimum area thresholds (Tables 3 and 4). The differences of the OA proportions between ik-NN and multinomial logistic regression were statistically significant at the level of 5% for the area thresholds 0, 0.5 and 1 hectares, near significance for 2 hectares, and not significant for three and four hectares when using McNemar's test for dependent samples. The validation data are used in the tests. The number of the observations in the validation data affects the significance. The rest of the analyses were carried out therefore with the ik-NN method only.

Concerning the key question of a sufficient time period for Sentinel-1 data acquisition for crop species prediction, some detailed analyses were carried starting with the features from one scene and accumulating scene by scene until the features from all scenes were the explanatory variables. These analyses were carried out using the parcel area threshold of two and four hectares only. The OA increased with the number of Sentinel-1 scenes until the middle of August when the first scenes were from the early June. After that, only minor improvements in the crop recognition accuracy were noted (Table 3). When using all 17 Sentinel-1 scenes from 14 May, in addition to the scenes from 1 June onward, the OAs were somewhat greater when the latest scenes were from the middle of July and slightly greater when the latest scenes were at about the same level as both starting dates, 14 May and 1 June (Tables 3 and 4 and Figure 4).

When starting the scene accumulation from the end of the image acquisition period that is from the latest scene, 30 August, the increase in the OA was more rapid compared to starting from the early growing season (Figure 5, cf. Figure 4). The OA increased quite rapidly and the scenes from 30 August backwards to late or mid July gave almost as large OA as the case in which the June and May scenes were also included. A conclusion is that the data from the later period of the growing season are more important than from the early period.

We also tested some separate combinations of a few Sentinel-1 data sets in crop species groups classification. The OAs with three scenes, from 1 June, 1 July and 25 July, and with the area thresholds two and four hectares were 0.628 and 0.641, respectively. Using the three scenes, from 1 July, 24 August and 30 August, the OA was 0.690 and 0.730, respectively. That is, adding the scene from 1 July to the two late August scenes increased the overall accuracies (cf. Figure 5). The difference of the proportions of correctly classified for the case 2 hectares, 0.062, is statistically significant at 5% level when using McNemar's test for dependent samples ( $1.96 \times \sqrt{var_{dif}} = 0.061$ ), and the difference of 0.089 slightly

not statistically significant at the same level for the case of four hectares (0.091). The validation data are used in the tests here.

The results revealed, in addition to relatively good potentials of Sentinel-1 time series in distinguishing the crop species, that it is sufficient to acquire the images from early June until early or mid-August, or even from late June, or from mid July until mid- or late August, when the growing conditions are similar to that in the test site in South Finland.

An example of a confusion matrix is given in Table 5. The user's and producer's accuracies (UA and PA) as well as F1 scores were also quite promising, UA and PA for most of the crop categories between 75–90, in some cases even 95%, except for the category Shelter belt, nature management field (9) and producer's accuracy for the category 5, 'Pea, beetroot and some other species' when the Sentinel-1 scenes were from 1 June until 3 August, and the parcel area threshold was two hectares (Table 6). Both user's and producer's accuracies of the category 2 (Broad bean), the category with smallest number of parcels, increased to some extent when the two May scenes were included.

**Table 3.** Overall accuracies with the improved k nearest neigbor (ik-NN) method, maximum 15 Sentinel-1 scenes from the period 1 June–30 August, ascending geometry VV and HV polarizations, and, with multinomial logist regression (Log regr), nine species groups, the minimum area of the parcel varies from 0 ha to 4 ha.

Sentinel-1 Data	1.6–1.7	1.6–13.7	1.6-25.7	1.6-12.8	1.6-24.8	1.6-30.8	1.6-30.8
Accumulation Period			ik	NN			Log regr
Number of scenes	6	8	10	12	14	15	15
Area threshold (ha)			0	verall accu	racy		
0	0.517	0.600	0.667	0.710	0.719	0.722	0.672
0.5	0.562	0.659	0.748	0.803	0.809	0.813	0.741
1	0.606	0.686	0.782	0.828	0.838	0.870	0.766
2	0.612	0.703	0.809	0.853	0.873	0.870	0.807
3	0.601	0.700	0.824	0.871	0.874	0.876	0.829
4	0.613	0.709	0.839	0.894	0.888	0.894	0.829

**Table 4.** Overall accuracies with the ik-NN method, maximum 17 Sentinel-1 scenes from the period 14 May–30 August, ascending geometry VV and HV polarizations, and with multinomial logist regression (Log regr), nine species groups, the minimum area of the parcel varies from 0 ha to 4 ha.

Sentinel-1 Data	14.5–1.7	14.5–13.7	14.5-25.7	14.5-12.8	14.5-24.8	14.5-30.8	14.5-30.8
Accumulation Period			ik-	NN			Log regr
Number of scenes	8	10	12	14	16	17	17
Area threshold (ha)			0	verall accur	acy		
0	0.537	0.578	0.657	0.707	0.716	0.722	0.686
0.5	0.594	0.646	0.747	0.802	0.807	0.813	0.749
1	0.625	0.692	0.776	0.839	0.841	0.846	0.772
2	0.644	0.719	0.805	0.847	0.872	0.872	0.817
3	0.629	0.705	0.824	0.850	0.880	0.880	0.834
4	0.591	0.695	0.818	0.883	0.888	0.900	0.820

Predicted	Observed Group									
Group	1	2	3	4	5	6	7	8	9	Total
1	451	0	0	4	2	8	7	28	7	507
2	0	12	2	1	3	0	0	0	0	18
3	2	1	204	0	22	0	0	1	0	230
4	1	1	0	35	0	0	0	0	0	37
5	3	3	2	0	48	1	0	0	5	62
6	3	0	0	0	2	68	0	0	17	90
7	1	0	0	0	0	5	53	1	1	61
8	11	1	2	0	5	1	2	277	1	300
9	1	0	0	0	3	14	1	0	32	51
Total	473	18	210	40	85	97	63	307	63	1356

**Table 5.** An example of confusion matrix, 15 Sentinel-1 scenes, 1 June–30 August, nine species groups, minimum area 2 ha, observed on coloumns, predicted on rows.

**Table 6.** An example of user's and producer's accuracies (UA and PA) and F1 scores by species groups with all the Sentinel-1 scenes when the first scene was either from 1 June, 15 scenes, or 14 May, 17 scenes. With 15 scenes, the data correspond to Table 5.

Species Group											
	1	2	3	4	5	6	7	8	9		
UA, 15 scenes	0.890	0.667	0.887	0.946	0.774	0.756	0.869	0.923	0.628		
PA, 15 scenes	0.954	0.668	0.971	0.875	0.565	0.701	0.841	0.902	0.508		
F1, 15 scenes	0.921	0.667	0.927	0.909	0.653	0.727	0.855	0.912	0.562		
UA, 17 scenes	0.898	0.857	0.902	0.947	0.769	0.726	0.850	0.915	0.628		
PA, 17 scenes	0.954	0.667	0.967	0.900	0.588	0.711	0.810	0.912	0.508		
F1, 17 scenes	0.925	0.750	0.933	0.923	0.666	0.718	0.830	0.913	0.562		



Number of the Sentinel-1 scenes

**Figure 4.** Overall accuracies as a function of the stage of the phenology and the number of scenes calculated with the minimum sizes of the parcels of two and four hectares.



Number of the Sentinel-1 scenes

**Figure 5.** Overall accuracies as a function of the stage of the phenology and the number of scenes calculated with the minimum sizes of the parcels of two and four hectares. The time is reverse compared to Figure 4.

An example of a map of the predicted crop species is shown in Figure 6 and an example of the largest class probability (scale 0–100) in Figure 7 (see Section 3.3). The Sentinel-1 scenes from 1 June to 30 August have been used here, in 15 scenes total. The crop species has been predicted for all field parcels here, that is, with a minimum field parcel size of greater than zero. Still, the majority of the probabilities are large, in 90% of the parcel area larger than 50% and, in about 80%, larger than or equal to 80% (see also Table 3).



**Figure 6.** An example of a map of the predicted crop species with the field parcel boundaries, ©Finnish Food Authority, ©Google Earth.



**Figure 7.** An example of a map of the probability of the predicted crop species with the field parcel boundaries, ©Finnish Food Authority, ©Google Earth.

# 3.2. The Effect of VV and VH Polarizations in Crop Species Classification

Some experiments were carried to test the capability of different VV and VH polarizations in crop species classifications. Each of the 17 single Sentinel-1 scenes were analyzed separately due to the assumption that the the classification capabilities of VV and VH polarizations change during the growing season. Only one explanatory variable, one scene, and one polarization was thus employed in each classification. The area threshold of two hectares was used in this experiment. The OAs were low as can be expected (Figure 8). The tests revealed that the OA increases during the growing season with both polarization until early August and either stabilizes (VV polarization) or decreases slightly after that (VH polarization). The OA increases with VV polarization from about 25% to 52% and with VH also from about 25% near 50% but decreases with VH after that to the level of 43%. The OAs with the two scenes together from 12 and 18 August gave the OAs of 58% (VV) and 55% (VH) and with the three scenes from from 12, 18, and 24 August 62% (VV) and 59% (VH). Finally, when using all 17 scenes, the OAs with single polarizations were 85.2% (VV) and 82.7% (VH). Based on these experiments, VV performs slightly better than VH, although the differences are small. However, the differences were statistically significant at the 5% level, many at the 1% level, in all cases except with the scene 19 June when using McNemar's test for dependent samples.



**Figure 8.** Overall accuracies when using a single Sentinel-1 scene and either VV or VH polarization calculated and the minimum sizes of the parcels of two hectares.

# 3.3. Confidence Intervals for Individual Parcel Predictions

The probabilities of the predicted category in the validation data sets were assessed as given in Equation (8). The 95% confidence intervals of the predicted probabilities were calculated using Equation (10) and the Sentinel-1 data sets from the period 1 June–12 August, that is, in total 14 scenes. Basic statistics of the widths of the confidence intervals by the parcel area thresholds are shown in Table 7. The median and mean of the intervals with a threshold of two hectares, e.g., are about 9% points. Figure 9 shows statistics of the widths of predictions intervals by the species groups for the parcels in the validation data set together with the outliers. The statistics are the median, 25% and 75% percentiles and 1.5 times the interquartile range.

The averages of the widths of 95% confidence intervals of the predicted probabilities by the area classes (ha) and species group predictions are shown in Table 8 and the number of the parcels using the same grouping in Table 9. The average width of the interval is smaller than 10% points in most cases.

**Table 7.** Examples of the statistics of the 95% confidence intervals of the predictions of the probability of the category with the largest probability when the Sentinel-1 scenes from the period 1 June–12 August 2017 were used. The statistics are shown by the threshold of the minimum of the parcel area.

Area (ha)	Mean of Predictions	Std of Predictions	Min of Intervals	Median of Intervals	Mean of Intervals	Max of Intervals	Std of Intervals
0 ha	0.888	0.096	0.062	0.120	0.130	0.077	0.022
2 ha	0.889	0.092	0.049	0.088	0.094	0.274	0.026
4 ha	0.889	0.092	0.062	0.120	0.130	0.299	0.039



**Figure 9.** Distributions of the confidence intervals by the species groups, shown as boxplots, that is, median, 25% and 75% percentiles, 1.5 times of the interquartile ranges, and individual outliers. The width of the box is proportional to the square root of the number of the observations of the species group.

Area ha		Predicted Species Group											
Alca, Ila	1	2	3	4	5	6	7	8	9	Total			
[0-1)	0.066	0.180	0.080	0.125	0.097	0.084	0.113	0.070	0.070	0.074			
1–2)	0.057	0.177	0.073	0.108	0.100	0.082	0.106	0.063	0.060	0.067			
2–3)	0.056	0.176	0.072	0.112	0.092	0.083	0.105	0.065	0.061	0.067			
3–4)	0.058	0.176	0.067	0.104	0.097	0.081	0.107	0.066	0.065	0.068			
4–5)	0.057	0.176	0.066	0.105	0.099	0.087	0.104	0.069	0.080	0.069			
5–6)	0.058	0.175	0.079	0.111	0.102	0.082	0.104	0.066	0.061	0.073			
6–7)	0.057	0.174	0.063	0.113	0.104	0.086	0.101	0.066	0.074	0.068			
7–8)	0.062	-	0.066	0.108	-	0.110	0.112	0.067	-	0.071			
<b>[</b> 8–9)	0.063	0.183	0.074	0.108	-	0.102	0.108	0.070	-	0.075			
[ <sup>5</sup> –10)	0.063	-	0.077	0.107	-	0.091	0.113	0.076	-	0.076			
[10-	0.077	-	0.088	0.130	0.118	0.109	0.128	0.092	-	0.093			
Total	0.060	0.177	0.074	0.112	0.098	0.084	0.109	0.068	0.068	0.071			

**Table 8.** The averages of the widths of 95% confidence intervals of the probability predictions of the categories by the area classes (ha) and species groups.

Area			Prec	licted	Speci	es Gro	oup			
Alea	1	2	3	4	5	6	7	8	9	Total
[0-1)	349	4	143	13	88	79	16	196	496	1384
[1–2)	236	4	107	17	28	40	11	126	123	692
2–3)	177	3	79	12	14	27	13	84	49	458
3–4)	114	3	62	10	6	20	9	62	23	309
4–5)	75	2	34	6	4	7	5	35	7	175
5–6)	47	2	22	7	7	10	1	26	1	123
[6–7)	35	1	16	2	4	5	1	12	5	81
7–8)	19	0	12	1	0	2	4	14	0	52
8–9)	14	1	6	1	0	1	2	12	0	37
9–10)	12	0	7	2	0	1	2	7	0	31
[10–	26	0	12	3	5	5	5	31	0	87
Total	1104	20	500	74	156	197	69	605	704	3429

Table 9. The number of the parcels by the area classes (ha) and species group prediction.

## 4. Discussions

Each study concerning crop species prediction with multi-temporal SAR data has some unique features, including the classification approach, number, and specificity of crop species used, as well as the SAR data used. This generally complicates comparison, as no baseline mapping approach is established, often requiring testing with several approaches within each study, e.g., [30]. Other factors affecting the possible accuracy levels in mapping are the structure of the crop species, that is, whether only crop species are predicted or also root vegetables.

In this study, different types of species groups were predicted with high accuracy. Uncertainty levels of the prediction were also assessed. Achieved accuracy levels compare favourably to prior recent studies particularly with multitemporal Sentinel-1 data [21,23,30]. Here, OA was either at same level or higher, especially when using the parcel size threshold of four hectares. Importantly, the number of crops was considerably higher, indicating a potential for further development of an operational crop mapping method.

The overall accuracy (OA) with the ik-NN method varied depending on the number of species included in the classification, and whether all field parcels were included, or an area threshold for the field parcels was employed. The later one exhibits the dependence of OA on the size of the area threshold. The OAs have been calculated here using the observation as the weight, that is, each field parcel had the same weight. The area weighted OAs had been larger because the crop species were predicted more accurately for the larger field parcels than for the small field parcels.

Some tests with a very high number of the species groups showed also a great OA. The OA was at a level of 82% with 41 species and with an area threshold of 0.5 hectares. However, the results with 41 species groups are dominated by a few species with large areas explaining an unexpected large OA.

The uncertainty assessments of the predicted crop species also gave promising results. The certainty was over 90% in most cases when using 95% confidence intervals (Tables 7 and 8, see also Figure 7). The prediction uncertainty is largest in the rare species groups, such as 2, 4, and 7 (Broad bean Vicia faba, Spring Oilseed Rape and 7 = Winter wheat, Winter barley) as can be expected. Overall, the species can be predicted relatively accurately in most cases, even compared to the earlier studies.

Importantly, while SAR time series have been used recently, the analyses of the effect of the precise acquisition time of the SAR data in crop species predictions has been studied less. The scenes from the late growing season had the greatest explanatory power in predicting the crop species (Figures 4 and 5). The analyses separately with the VV and VH polarizations and with one Sentinel-1 scene only revealed the same thing; in addition to that, the VV polarization gave greater OAs than VH polarization in the later part of the growing season (Figure 8).

Establishing the number of Sentinel-1 scenes and the timing of SAR data that it takes to achieve acceptable overall accuracy can have important effects on operational LPIS. The time period from

which the Sentinel-1 scenes or other SAR data are needed depends naturally on the progress of the growing season and thus the latitude and other growing factors. The data acquisition period must be adapted to the local conditions. The proposed method to estimate the confidence intervals of the probabilities of the predicted species group is also new to the best of our knowledge.

# 5. Conclusions

The study demonstrated a strong potential of C-band Sentinel-1 time series in consistent monitoring crop cultivation and in distinguishing crop species with high levels of accuracy. The improved k-NN method, ik-NN, in which the feature weights are optimized using a genetic algorithm, gave somewhat higher overall accuracies with all the data sets tested than multinomial logistic regression analysis. It was also competitive compared to several prior studies keeping in mind larger number of crops, especially when larger parcels are considered in the prediction accuracy assessment.

Overall accuracies in predicting crop species with nine species groups and ik-NN method were remarkably high. The overall accuracy in a separate validation data set was around 70% when all cropland parcels were included, around 80% when the parcel area threshold was 0.5 ha, and around 90% when the threshold was 4.0 ha. The scenes acquired from early June until the middle of August turned out to be sufficient. The user's and producer's accuracies were also high, in most cases 90% or higher except the category 'Shelter belt, nature management field' (about 55%). The scenes from the mid and late growing season turned out to be important in species prediction in our study.

A method to assess the uncertainty of the crop prediction on an individual field parcel was presented. The widths of the 95% confidence intervals for the probability of the individual field parcels varied by the species groups and also by the area of the parcel. The narrowest average intervals were 0.05–0.06 units (5–6% points) and the widest ones less than 0.2 units (20% points). It is expected that this kind of information can be used in determining whether other auxiliary data and training data are needed for crop identification, such as, e.g., airborne observations, or even field visits.

The presented analysis of SAR data acquisition timing on crop species predictions can be used in other similar studies and incorporated into operational LPIS. The latter can also benefit from the developed methodology to estimate the prediction uncertainty of the various species group. The results of this study were well-received by the authorities and encourage the continuation of the study towards an operational method in which the space-borne SAR data are a critical part of the information chain.

**Author Contributions:** E.T. was primarily responsible for performing the analysis and acquiring reference data used in the study, as well as writing the paper. O.A. preprocessed the SAR data and co-wrote the manuscript. J.P. co-wrote a part of the text and designed a presentation of some results. All authors participated in the discussions and reviewed the results.

**Funding:** The study was funded by Aalto University and in the beginning partly by Bitcomp Oy. The work by O.A. was partly funded by VTT Technical Research Centre of Finland LTD.

**Acknowledgments:** The species data with field parcel boundaries were provided by the Finnish Food Authority. The work began in a collaboration with Bitcomp Oy and Aalto University. Bitcomp Oy partly funded the beginning of the work. Three anonymous reviewers provided constructive comments that improved the article significantly. Majella Clarke edited the English. We thank all institutions and individuals that have contributed to the article.

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

#### References

- European Court of Auditors. The Land Parcel Identification System: A Useful Tool to Determine the Eligibility of Agricultural Land—However, Its Management Could Be Further Improved; Publications Office of the European Union: Luxembourg, 2016.
- Tomppo, E.O.; Gagliano, C.; Natale, F.D.; Katila, M.; McRoberts, R.E. Predicting categorical forest variables using an improved k-Nearest Neighbour estimator and Landsat imagery. *Remote Sens. Environ.* 2009, 113, 500–517. [CrossRef]

- 3. Tomppo, E.; Halme, M. Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: A genetic algorithm approach. *Remote Sens. Environ.* **2004**, *92*, 1–20. [CrossRef]
- 4. McNairn, H.; Brisco, B. The application of C-band polarimetric SAR for agriculture: A review. *Can. J. Remote Sens.* **2004**, *30*, 525–542. [CrossRef]
- Skriver, H.; Mattia, F.; Satalino, G.; Balenzano, A.; Pauwels, V.R.N.; Verhoest, N.E.C.; Davidson, M. Crop Classification Using Short-Revisit Multitemporal SAR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2011, 4, 423–431. [CrossRef]
- 6. Wegmuller, U.; Werner, C. Retrieval of vegetation parameters with SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **1997**, 35, 18–24. [CrossRef]
- McNairn, H.; Champagne, C.; Shang, J.; Holmstrom, D.; Reichert, G. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. *ISPRS J. Photogramm. Remote Sens.* 2009, 64, 434–449. [CrossRef]
- 8. McNairn, H.; Shang, J. A Review of Multitemporal Synthetic Aperture Radar (SAR) for Crop Monitoring. In *Multitemporal Remote Sensing*; Ban, Y., Ed.; Springer: Cham, Switzerland, 2016; pp. 317–340.
- 9. Deschamps, B.; McNairn, H.; Shang, J.; Jiao, X. Towards operational radar-only crop type classification: Comparison of a traditional decision tree with a random forest classifier. *Can. J. Remote Sens.* **2012**, *38*, 60–68. [CrossRef]
- Bégué, A.; Arvor, D.; Bellon, B.; Betbeder, J.; De Abelleyra, D.; Ferraz, R.P.D.; Lebourgeois, V.; Lelong, C.; Simões, M.; R. Verón, S. Remote Sensing and Cropping Practices: A Review. *Remote Sens.* 2018, 10, 99. [CrossRef]
- Shuai, G.; Zhang, J.; Basso, B.; Pan, Y.; Zhu, X.; Zhu, S.; Liu, H. Multi-temporal RADARSAT-2 polarimetric SAR for maize mapping supported by segmentations from high-resolution optical image. *Int. J. Appl. Earth Obs. Geoinf.* 2019, 74, 1–15. [CrossRef]
- 12. Hutt, C.; Waldhoff, G. Multi-data approach for crop classification using multitemporal, dual-polarimetric TerraSAR-X data, and official geodata. *Eur. J. Remote Sens.* **2018**, *51*, 62–74. [CrossRef]
- Jiao, X.; Kovacs, J.M.; Shang, J.; McNairn, H.; Walters, D.; Ma, B.; Geng, X. Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. *ISPRS J. Photogramm. Remote Sens.* 2014, *96*, 38–46. [CrossRef]
- 14. Liu, C.A.; Chen, Z.X.; Shao, Y.; Chen, J.S.; Hasi, T.; PAN, H.Z. Research advances of SAR remote sensing for agriculture applications: A review. *J. Integr. Agric.* **2019**, *18*, 506–525. [CrossRef]
- Loosvelt, L.; Peters, J.; Skriver, H.; Lievens, H.; Van Coillie, F.; De Baets, B.; Verhoest, N. Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification. *Int. J. Appl. Earth Obs. Geoinf.* 2012, 19, 173–184. [CrossRef]
- 16. Chen, S.W.; Li, Y.Z.; Wang, X.S. Crop discrimination based on polarimetric correlation coefficients optimization for PolSAR data. *Int. J. Remote Sens.* **2015**, *36*, 4233–4249. [CrossRef]
- 17. Schmullius, C.; Thiel, C.; Pathe, C.; Santoro, M. Radar time series for land cover and forest mapping. *Remote Sens. Digit. Image Process.* **2015**, *22*, 323–356. [CrossRef]
- 18. Blaes, X.; Vanhalle, L.; Defourny, P. Efficiency of crop identification based on optical and SAR image time series. *Remote Sens. Environ.* **2005**, *96*, 352–365. [CrossRef]
- Steele-Dunne, S.; McNairn, H.; Monsivais-Huertero, A.; Judge, J.; Liu, P.W.; Papathanassiou, K. Radar Remote Sensing of Agricultural Canopies: A Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2017, 10, 2249–2273. [CrossRef]
- Guo, J.; Wei, P.L.; Liu, J.; Jin, B.; Su, B.F.; Zhou, Z.S. Crop Classification Based on Differential Characteristics of Hα Scattering Parameters for Multitemporal Quad- and Dual-Polarization SAR Images. *IEEE Trans. Geosci. Remote Sens.* 2018, 56, 6111–6123. [CrossRef]
- 21. Xu, L.; Zhang, H.; Wang, C.; Zhang, B.; Liu, M. Crop Classification Based on Temporal Information Using Sentinel-1 SAR Time-Series Data. *Remote Sens.* **2018**, *11*, 53. [CrossRef]
- 22. McNairn, H.; Kross, A.; Lapen, D.; Caves, R.; Shang, J. Early season monitoring of corn and soybeans with TerraSAR-X and RADARSAT-2. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 252–259. [CrossRef]
- 23. Veloso, A.; Mermoz, S.; Bouvet, A.; Toan, T.L.; 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]

- 24. Zalite, K.; Antropov, O.; Praks, J.; Voormansik, K.; Noorma, M. Monitoring of Agricultural Grasslands with Time Series of X-Band Repeat-Pass Interferometric SAR. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3687–3697. [CrossRef]
- 25. Tamm, T.; Zalite, K.; Voormansik, K.; Talgre, L. Relating Sentinel-1 interferometric coherence to mowing events on grasslands. *Remote Sens.* **2016**, *8*, 802. [CrossRef]
- 26. Whelen, T.; Siqueira, P. Time-series classification of Sentinel-1 agricultural data over North Dakota. *Remote Sens. Lett.* **2018**, *9*, 411–420. [CrossRef]
- 27. Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K.I. Assessing the suitability of data from Sentinel-1A and 2A for crop classification. *GISci. Remote Sens.* **2017**, *54*, 918–938. [CrossRef]
- 28. Kenduiywo, B.K.; Bargiel, D.; Soergel, U. Crop-type mapping from a sequence of Sentinel 1 images. *Int. J. Remote Sens.* **2018**, *39*, 6383–6404. [CrossRef]
- 29. Clauss, K.; Ottinger, M.; Kuenzer, C. Mapping rice areas with Sentinel-1 time series and superpixel segmentation. *Int. J. Remote Sens.* **2018**, *39*, 1399–1420. [CrossRef]
- Ndikumana, E.; Ho Tong Minh, D.; Baghdadi, N.; Courault, D.; Hossard, L. Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sens.* 2018, 10, 1217. [CrossRef]
- 31. Tomppo, E.; Haakana, M.K.M.P.J. *Multi-Source National Forest Inventory—Methods and Applications, Managing Forest Ecosystems*; Springer: Dordrecht, The Netherlands, 2008.
- 32. Song, Y.; Wang, J. Mapping Winter Wheat Planting Area and Monitoring Its Phenology Using Sentinel-1 Backscatter Time Series. *Remote Sens.* **2019**, *11*, 449. [CrossRef]
- 33. Authority, F.F. Finnish Food Authority. Available online: https://www.ruokavirasto.fi/en/about-us/ services/ (accessed on 22 October 2019).
- 34. Small, D. Flattening Gamma: Radiometric Terrain Correction for SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3081–3093. [CrossRef]
- 35. Tomppo, E.; Antropov, O.; Praks, J. Boreal Forest Snow Damage Mapping Using Multi-Temporal Sentinel-1 Data. *Remote Sens.* **2019**, *11*, 384. [CrossRef]
- 36. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- 37. Vapnik, V.N. The Nature of Statistical Learning Theory; Springer: New York, NY, USA, 1995.
- Chirici, G.; Mura, M.; McInerney, D.; Py, N.; Tomppo, E.O.; Waser, L.T.; Travaglini, D.; McRoberts, R.E. A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. *Remote Sens. Environ.* 2016, *176*, 82–294. [CrossRef]
- 39. Mitchell, M. An Introduction to Genetic Algorithms; The MIT Press: Cambridge, MA, USA, 1996.



© 2019 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 (http://creativecommons.org/licenses/by/4.0/).