

Article

Winter Wheat Yield Assessment from Landsat 8 and Sentinel-2 Data: Incorporating Surface Reflectance, Through Phenological Fitting, into Regression Yield Models

Sergii Skakun ^{1,2,3,*} , Eric Vermote ³, Belen Franch ^{1,3}, Jean-Claude Roger ^{1,3} ,
Nataliia Kussul ⁴, Junchang Ju ^{5,6} and Jeffrey Masek ⁶

¹ Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA

² College of Information Studies (iSchool), University of Maryland, College Park, MD 20742, USA

³ NASA Goddard Space Flight Center Code 619, 8800 Greenbelt Road, Greenbelt, MD 20771, USA

⁴ Space Research Institute NAS Ukraine & SSA Ukraine, 03680 Kyiv, Ukraine

⁵ Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD 20742, USA

⁶ NASA Goddard Space Flight Center Code 618, 8800 Greenbelt Road, Greenbelt, MD 20771, USA

* Correspondence: skakun@umd.edu; Tel.: +1-301-405-2179

Received: 9 July 2019; Accepted: 25 July 2019; Published: 27 July 2019



Abstract: A combination of Landsat 8 and Sentinel-2 offers a high frequency of observations (3–5 days) at moderate spatial resolution (10–30 m), which is essential for crop yield studies. Existing methods traditionally apply vegetation indices (VIs) that incorporate surface reflectances (SRs) in two or more spectral bands into a single variable, and rarely address the incorporation of SRs into empirical regression models of crop yield. In this work, we address these issues by normalizing satellite data (both VIs and SRs) derived from NASA’s Harmonized Landsat Sentinel-2 (HLS) product, through a phenological fitting. We apply a quadratic function to fit VIs or SRs against accumulated growing degree days (AGDDs), which affects the rate of crop development. The derived phenological metrics for VIs and SRs, namely peak, area under curve (AUC), and fitting coefficients from a quadratic function, were used to build empirical regression winter wheat models at a regional scale in Ukraine for three years, 2016–2018. The best results were achieved for the model with near infrared (NIR) and red spectral bands and derived AUC, constant, linear, and quadratic coefficients of the quadratic model. The best model yielded a root mean square error (RMSE) of 0.201 t/ha (5.4%) and coefficient of determination $R^2 = 0.73$ on cross-validation.

Keywords: agriculture; crop yield; wheat; Landsat 8; Sentinel-2; HLS; phenological fitting; growing degree days (GDD)

1. Introduction

The combination of data acquired by Landsat 8 and Sentinel-2 remote sensing satellites can provide a high temporal resolution (3–5 days) [1], which is critical for various applications requiring dense satellite data time series. Previously, such (or better) high temporal resolution was available mainly for Earth observation sensors, which acquire daily data over Earth’s surface, but at a coarser spatial resolution (>250 m) [2,3]. The latter, for example, includes space-borne remote sensing sensors, such as Moderate Resolution Imaging Spectroradiometer (MODIS), Visible Infrared Imaging Radiometer Suite (VIIRS), Advanced Very High Resolution Radiometer (AVHRR), and SPOT-VEGETATION. Taking into account an increased frequency of observations at moderate spatial resolution (<30 m), the assumption

is that methods for generating products from coarse spatial resolution sensors can be ported to moderate (Landsat 8/OLI, Sentinel-2/MSI) or high (Planet/PlanetScope) spatial resolution sensors. However, the practice shows that such a transition is not always straightforward due to larger data gaps because of clouds and uneven coverage, sensor characteristics, and increased spatial resolution (at least at the order of 100, when going from 250 to 30 m).

Consider, for example, a crop yield assessment/forecasting application [4]. The hypothesis is that satellite-based features, such as vegetation indices (VIs) or biophysical parameters derived at a single date or accumulated over some time period, can be correlated to crop yields [5,6]. Since the reference data on crop yields are mainly available at the regional scale, the corresponding empirical models are built by averaging satellite-based features over those regions and correlating these derived variables with crop yields [7–9]. It is assumed that there is a homogeneity within the region in terms of crops grown and agricultural practices applied and, therefore, the averaging should be performed for satellite data acquired at the same (or approximately the same) stages of crop growth, meaning that the data are normalized. This is usually the case for coarse spatial resolution remote sensing sensors, which enable a higher likelihood of obtaining a high temporal frequency of cloud-free data over the Earth's surface [10,11]. This is also evidenced by multiple successful applications of coarse spatial resolution satellite data to crop yield assessment and forecasting [5,7–9,12–15].

However, this is not the case for moderate spatial resolution satellite data (<30 m). Irregular spatial coverage, when the area in question is covered by several “stripes” sensed at different times, and lower revisit cycles lead to discrepancies in dates of cloud-free observations. For example, Azzari et al. (2017) [16] encountered this problem, when they used Landsat data for mapping the yields of corn and soybean in the USA. The direct application of their model to available Landsat observations resulted in obvious artefacts, which were due not to spatial variability of crop yields but rather to irregularities in satellite observations at a moderate spatial resolution. To address this problem, they used MODIS observations to further rectify (normalize) Landsat data, so to avoid the effect of irregular Landsat observations on the resulting product. MODIS data are frequently used to assist with the processing of Landsat-like sensors to generate near-daily observations at a 30 m resolution [17]. These studies showed promising results and have already been applied to assessing yields at the field scale [18,19]. However, the effect of combining various spatial resolutions, i.e., >250 m MODIS data with 30 m Landsat data, usually is not quantified in terms of crop yield model performance. For example, Gao et al. (2019) [20] indicated that the error of Landsat-MODIS fusion can be up to 0.083 for the normalized difference vegetation index (NDVI), 0.026 for the red band, and 0.052 for the near infrared (NIR) band. Moreover, previous studies showed saturation of the commonly applied NDVI at a 30 m spatial resolution for winter wheat at high yields (>4 t/ha) or for leaf-area index (LAI) > 2 [21,22], something that was not usually observed for MODIS due to mixed pixels. This means that the behavior of commonly applied VIs at coarse and moderate spatial resolutions might be different.

Recent studies took advantage of Landsat and Sentinel-2 data to address crop yield assessment at a moderate spatial resolution. For example, Lambert et al. (2018) [23] used Sentinel-2 data and a peak LAI approach to estimate the yields of cotton, maize, millet, and sorghum in Mali. The R^2 ranged from 0.48 to 0.80 for different crops on training data. However, they did not address the issues of proving that they are actually capturing the peak. Also, this study was based on a single year, therefore the model's robustness in time was not explored. Lai et al. (2018) [24] used time-integrated Landsat NDVI for wheat yield prediction in Australia. They used an asymmetric bell-shaped growth model to fit NDVI against time. Their approach yielded a root mean square error (RMSE) of 0.79 t/ha on cross-validation. However, they did not go beyond NDVI and did not explore surface reflectance values.

A frequently used approach to smooth satellite data is through phenological fitting. Multiple fitting functions have been proposed, tested, and validated [24–32]. The current general consensus is that no single model would fit all requirements and selection of the smoothing method and/or model is predicated by the application, and prior knowledge of the land–surface dynamics [25]. Also, almost all

existing remote sensing-based studies for crop yields are focused on using vegetation indices and do not exploit spectral information.

Though a combination of Landsat 8 and Sentinel-2 offers a high frequency of observations, discrepancies in available cloud-free data, however, will still exist. With more and more applications transitioning to a higher spatial resolution, it is important to explore and analyze the effect of those discrepancies on the quality of the resulting products. In the case of crop yield assessment, one of the important questions is as follows: How do discrepancies in the dates of satellite observations influence the performance of empirical crop yield models at the regional scale? The present paper aims to address this question by documenting the results of a study building empirical models for winter wheat yield assessment with Landsat 8 and Sentinel-2 data in Ukraine. Not performing satellite data smoothing may lead to poorer performance of the empirical crop yield model, which would be attributed not to the lack of correlation with satellite-derived variables but rather to observation irregularities. We will not limit satellite data fitting to vegetation indices only, and will also explore surface reflectance values in multiple spectral bands for the assessment of winter wheat yields. Finally, we will evaluate the use of multi-source data, namely Landsat 8 and Sentinel-2, for crop yield assessment. In particular, we assess quantitatively how the performance of an empirical model of crop yield depends on the combined use of Landsat 8 and Sentinel-2 data, compared to a single-sensor usage.

2. Materials

2.1. Study Area and Reference Data

The study was performed for Kirovohradska oblast in Ukraine for 2016 to 2018. An “oblast” is a high-level administrative division of the country, and each oblast is further divided into districts. There are 24 oblasts in Ukraine and Autonomous Republic of Crimea. Kirovohradska oblast is located in the central part of Ukraine and consists of 21 districts with a geographical area ranging from 65 to 165 thousand ha and cropland area ranging from 27 to 112 thousand ha (Figure 1).

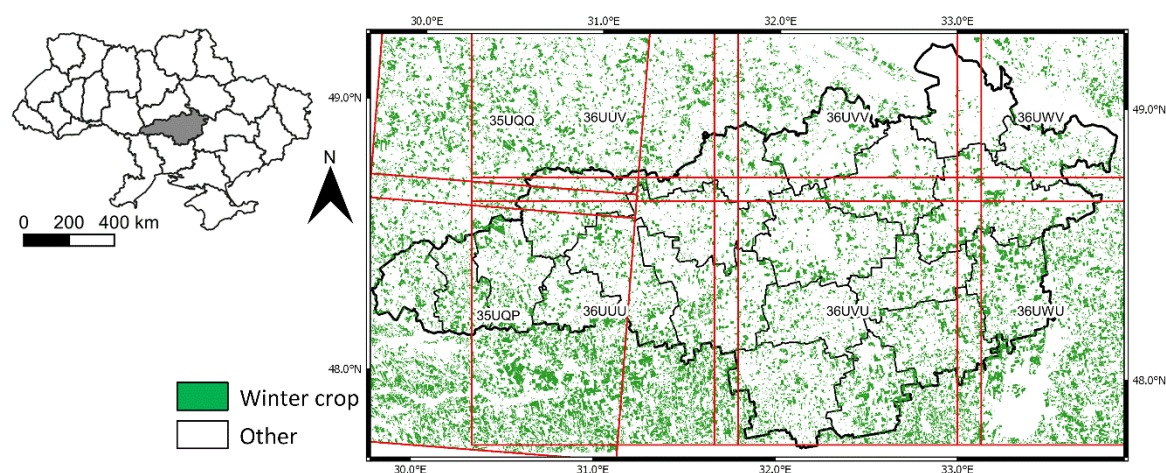


Figure 1. Location of Kirovohradska oblast in Ukraine and its division into districts. Shown also is the location of Sentinel-2 tiles covering the area. The background image shows a winter crop mask derived from Landsat 8 and Sentinel-2 for 2018.

The reason for selecting this region is that it is a top 10 wheat producer in Ukraine and because of the availability of reference crop yield and harvested area data at the district scale for 2016 to 2018. Winter wheat is one of the major crops in Kirovohradska oblast, accounting for 20% of the production of all crops in the region. Winter wheat is mainly rain-fed in the region and usually planted in September to October. After dormancy during the winter, it re-starts its growth in early spring, reaching maturity by the end of June. The harvest of winter crops is typically undertaken in July.

Reference data on the crop yield and harvested area at the district level were collected from the Department of Agro-Industry Development of Kirovohrad State Administration (<http://apk.kr-admin.gov.ua>). The data were made available online as the harvest progressed and were based on farm surveys of all large agricultural enterprises (that account for more than 90% of all winter crops produced in the region) and samples of household farms, the same way that official statistics are collected [33]. The final estimates for winter crop yields and harvested areas were available at the end of November and were used as a reference in this study. The reference data went through a quality-control procedure, which resulted in two districts in 2017 and five districts in 2018, being excluded from the analysis, because the reported production was $>+/-10\%$ than the area multiplied by yields. The average winter wheat yield among the districts was 4.0 ± 0.3 t/ha in 2016, 3.5 ± 0.4 t/ha in 2017, and 3.6 ± 0.4 t/ha in 2018.

2.2. Landsat-8/OLI and Sentinel-2A/MSI Datasets

Remote sensing images acquired by the Operational Land Imager (OLI) instrument aboard the Landsat-8 satellite and by the Multi-Spectral Instrument (MSI) aboard the Sentinel-2A/B satellites were used in the study. Landsat 8/OLI captures images of the Earth's surface in nine spectral bands at a 30 m spatial resolution (15 m for panchromatic band) [34], while Sentinel-2A/B/MSI captures images of the Earth's surface in 13 spectral bands at a 10, 20, and 60 m spatial resolution [35]. We used NASA's Harmonized Landsat Sentinel-2 (HLS) product [36], which provides a Level-2 Nadir BRDF (Bidirectional Reflectance Distribution Function)-Adjusted surface Reflectance (NBAR) at a 30 m spatial resolution. HLS combines, in a single data set, observations of the land surface from the Landsat 8's OLI [37] and Sentinel-2's MSI [38]. MSI data are also spectrally adjusted to OLI spectral bands, so to generate a single harmonized dataset. We used HLS version 1.4. A detailed description of HLS product generation is described in [37].

The HLS product is organized using the Sentinel-2 110 km $\times$ 110 km tiles. The following tiles cover the study region (Figure 1): 35UQQ, 35UQP, 36UUV, 36UUU, 36UVV, 36UVU, 36UWV, and 36UWU. Overall, 3565 scenes acquired by Landsat 8 and Sentinel-2A/B and covering eight tiles were used in the study. The scene covered the period of March to July within 2016 to 2018 to capture the main development stages of wheat. It should be noted that the availability of Sentinel-2 data varied through 2016 to 2018. In 2016 and 2017, only Sentinel-2A data were available, and only in 2017 did Sentinel-2A start operational acquisitions on a 10-day revisit cycle. In 2018, both Sentinel-2A/B were available with a combined five-day revisit cycle.

HLS data were used to extract surface reflectances (SRs) and to calculate corresponding vegetation indices (VIs). We used spectral bands common to Landsat 8 and Sentinel-2A/B, namely: Green (wavelength ~ 0.560 μ m), red (~ 0.660 μ m), near infrared (NIR) (~ 0.865 μ m, we used band 8A from Sentinel-2), and two short-wave infrared (SWIR) (~ 1.5 and ~ 2.2 μ m). This is an advantage of using analysis ready data (ARD), such as HLS, since data have been harmonized and as if they are acquired by a single sensor. We used a quality assessment (QA) layer provided in HLS to disregard unreliable pixels identified either as cloud, adjacent to cloud, shadow, snow/ice, or water.

Red and NIR bands were used to calculate the following commonly used VIs:

- Normalized difference vegetation index (NDVI) [39]:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}}. \quad (1)$$

- Enhanced vegetation index (EVI) [40]:

$$EVI2 = 2.5 \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + 2.4\rho_{red} + 1}. \quad (2)$$

- Difference vegetation index (DVI) [41]:

$$DVI = \rho_{NIR} - \rho_{red}, \quad (3)$$

where ρ_{NIR} and ρ_{red} are SR values in NIR and red spectral bands, respectively.

2.3. Meteorological Data

We used air temperature derived from NASA's Modern-Era Retrospective analysis for Research and Applications (MERRA2) reanalysis data set [42] to compute the growing degree days (GDDs) for winter wheat. Santamaria-Artigas et al. (2019) [43] showed that MERRA2 is one of the best data sources for computing GDD for winter wheat in terms of error (<2K compared to in situ measurements), spatial resolution, and low delay in data availability. The data are provided on a regular grid that has 576 points in the longitudinal direction and 361 points in the latitudinal direction, corresponding to a resolution of $0.625^\circ \times 0.5^\circ$. We used the time-averaged two-dimensional data collection (M2SDNXSLV) to extract daily averaged two-meter air temperature (T2MMEAN). The data sets were extracted from the netCDF format, transformed to the geo-referenced raster, subset for study areas, and linearly interpolated to the 30-m spatial resolution.

3. Methods

3.1. General Overview

Figure 2 shows main components and dataflows of the study with generation of winter crop masks and winter wheat yield estimation.

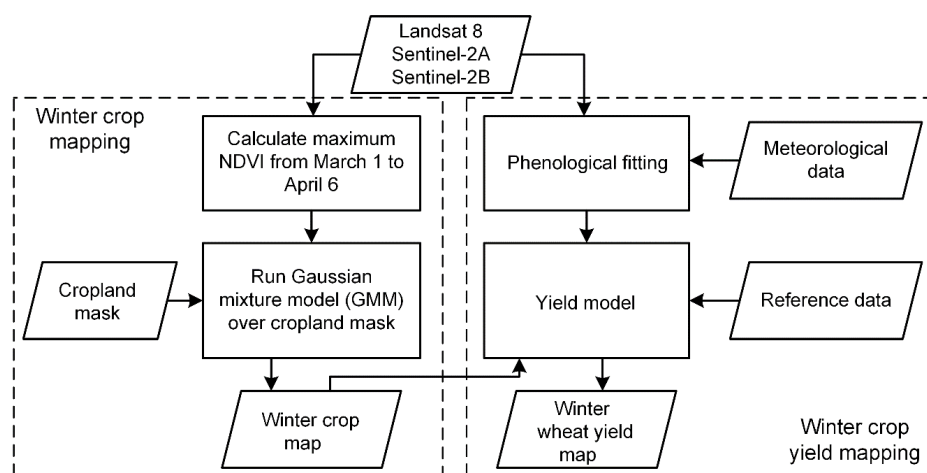


Figure 2. General workflow for winter wheat mapping and yield assessment.

The crop-specific masks were generated on a yearly basis, so the signal for building empirical yield models was extracted for winter crops only. Satellite data, namely SRs and VIs, were first smoothed before incorporating them into the crop yield model. This was performed through phenological fitting against meteorological data, namely accumulated GDD (AGDD), since temperature is a primary factor affecting the rate of crop development [44]. The following subsections describe in detail the methodologies used.

3.2. Winter Crop Type Mapping

In order to provide in-season winter crop maps (2016–2018), we implemented an automatic approach, which was initially developed for the coarse spatial resolution MODIS sensor [45] and further extended and prototyped for Landsat 8/OLI and Sentinel-2/MSI at a 30 m spatial resolution [22].

The approach uses a phenological metric during the green-up stage of winter crop development, namely the maximum NDVI, which is used as a main feature to discriminate winter crops from spring and summer crops. A cropland mask [46] is an essential data input for the method, to constrain the analysis to cultivated areas only, and to reduce potential confusion with natural vegetation and forested areas. To perform discrimination in an automatic way, a Gaussian mixture model (GMM) [47] is applied with the number of GMM components set to 2. Setting this number is based on the assumption that there will be two modes in the distribution of maximum NDVI (phenological metric), which would correspond to winter and non-winter crops, respectively. This method was run in an automatic unsupervised way during the vegetation season for 2016 to 2018 without the need to provide ground truth. However, this method requires ancillary data inputs and a priori knowledge about the crop calendar. We refer readers to [22,45] for details.

3.3. Winter Wheat Yield Assessment

Crop yield assessment was done through the development of an empirical regression model between satellite-derived features (regressors) and reference crop yields at the district scale (Figure 1). Satellite-derived features were averaged over the crop mask (Section 3.2) for each district and correlated with the available reference winter wheat yields (Section 2.1). Before averaging satellite-based features over pixels identified as winter crops, one has to ensure that satellite data are phenologically normalized, so averaging is performed for the same stages of crop growth development. Though a combination of Landsat 8 and Sentinel-2 offers a high frequency of observations, discrepancies in available cloud-free data, however, will still exist (Figure 3).

We used a quadratic model between satellite data (VI or SR) and AGDD [26,48,49]:

$$y_{sat} = a_0 + a_1 * AGDD + a_2 * AGDD^2, \quad (4)$$

where y_{sat} is the VI (NDVI, EVI2, or DVI) or SR (green, red, NIR, SWIR1, or SWIR2); and a_0 , a_1 , and a_2 are the constant, linear, and quadratic coefficients, respectively. GDD was defined as the average daily maximum (T_{max}) and minimum temperatures (T_{min}) minus a base temperature (T_{base}):

$$GDD = \frac{T_{max} + T_{min}}{2} - T_{base}, \quad (5)$$

where if $[(T_{max} + T_{min})/2 < T_{base}]$, then $GDD = 0$ [50], and with $T_{base} = 0$. AGDD was calculated by summing GDDs (Equation (5)).

Having derived per-pixel relationships between VI/SR and AGDD (Equation (4)), it is possible to derive various phenological metrics to correlate with yields or incorporate coefficients (a_0 , a_1 , a_2) directly into the yield model. The generalized relationship between crop yields and satellite-derived features can be expressed as follows:

$$y_{yield}^l = \omega_0 + \sum_{j=1}^{n_r} \omega_j r_j^l, \quad (6)$$

where y_{yield}^l is the crop yield for district l ; n_r is the number of regressors used; $\omega_j|_{j=0}^{n_r}$ are regression coefficients, and r_j^l are satellite-derived features averaged for district l over the crop mask:

$$r_j^l = \frac{1}{|A^l|} \sum_{p \in A^l} r_{j,p}^l \quad (7)$$

where A^l is the set of winter crop pixels; $r_{j,p}^l$ is the satellite-derived feature for the winter crop pixel at location p ; and $|\cdot|$ is the cardinality number, i.e., the number of elements in the set.

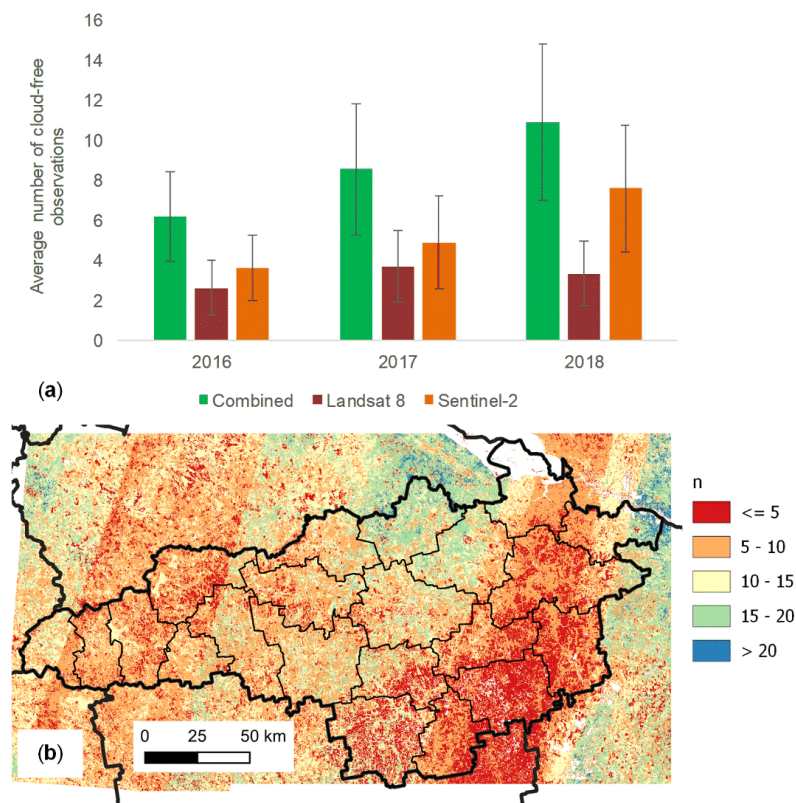


Figure 3. (a) Average number of cloud-free observations for winter crop pixels depending on satellite data usage. The number of pixels was taken from March until the end of June, which was the period of winter crop growth. (b) Spatial distribution of the number of cloud-free observations from Landsat 8 and Sentinel-2 from March through the end of June in 2018.

Multiple approaches were considered in the study that varied in the regressors used (VI, SR, or combination of SRs) and the number of regressors, as well as phenological metrics. For VIs, two phenological metrics were derived from the quadratics relationship: Peak and area under curve (AUC) (Figure 4).

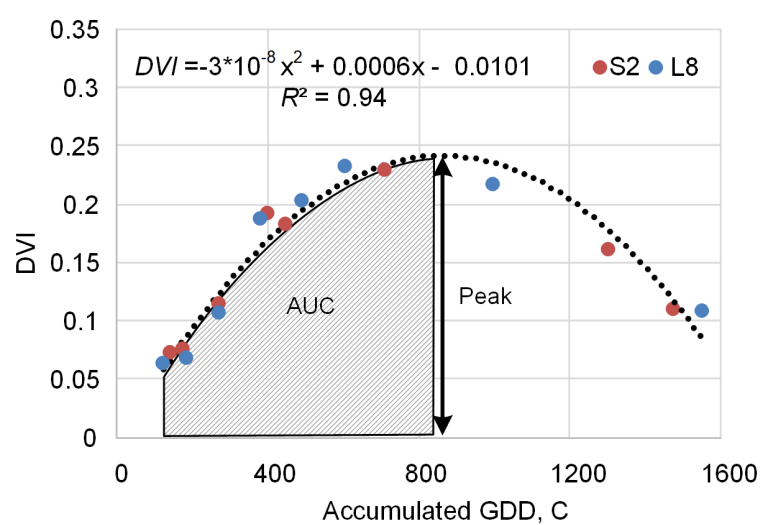


Figure 4. Peak and accumulation approaches for winter wheat yield assessment. Shown is the difference vegetation index (DVI), derived from Landsat 8 and Sentinel-2 data over a winter crop pixel, and fitted against accumulated growing degree days (AGDD) using a quadratic relationship.

The peak approach assumes that yield is correlated to the peak biomass during a crop growth season [7–9,23,37,51,52], which would result in the following regression (we removed l notations for simplicity):

$$y_{\text{yield}} = \omega_0 + \omega_1 r_1, \quad \text{where } r_1 \text{ is } p_{VI} = a_0 - \frac{a_1^2}{4a_2}, \quad (8)$$

where p_{VI} is the peak of the VI during the vegetation season.

The accumulation, or integrated, approach assumes that yield is proportional to the accumulated biomass over the crop growth season [18,53–56]:

$$y_{\text{yield}} = \omega_0 + \omega_1 r_1, \quad r_1 \equiv s_{VI} = \int_{AGDD_{\min}}^{AGDD_{\text{peak}}} y_{\text{sat},VI} dx = \int_{AGDD_{\min}}^{AGDD_{\text{peak}}} (a_0 + a_1 x + a_2 x^2) dx, \quad (9)$$

where $AGDD_{\min}$ and $AGDD_{\text{peak}}$ define ranges, for which accumulation is performed. $AGDD_{\min}$ was determined based on the minimum VI values, namely 0.05 for DVI and 0.15 for NDVI and EVI2. These thresholds for VIs were determined empirically to make sure that accumulation was performed for vegetated stages and not bare land. $AGDD_{\text{peak}}$ is the AGDD value, for which the VI's peak is achieved and is equal to $-\frac{a_1}{2a_2}$ (Equation (4)).

The p -value was also calculated to test the null hypothesis that a coefficient for this regressor is equal to zero, i.e., the regressor has no effect on the dependent variable.

It should be noted that almost all existing approaches dealing with satellite data and crop yields utilize VIs (e.g., Equations (1)–(3)), which incorporate two or more spectral bands into a single variable. With fitted SRs, one can explore trajectories in a multi-dimensional space of SRs for different spectral bands. Figure 5 shows examples of 1-D and 2-D trajectories for DVI and NIR-red, respectively, for different winter wheat yield values.

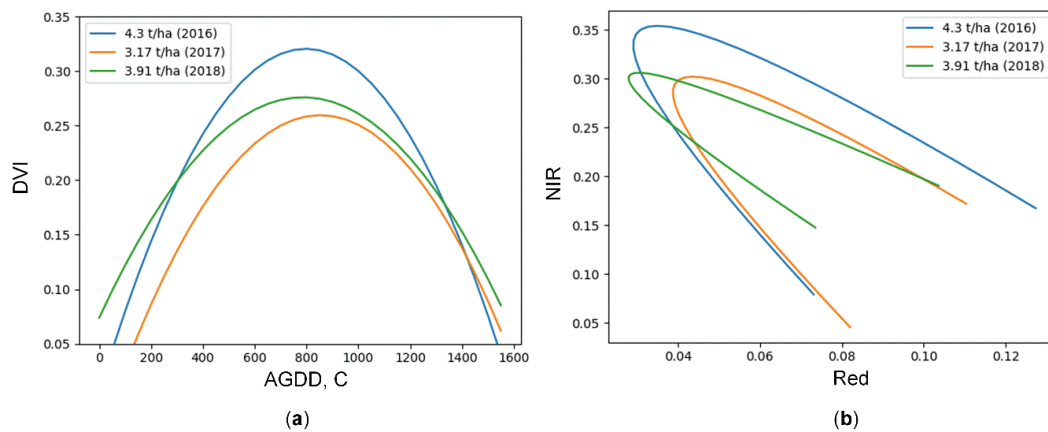


Figure 5. DVI (a) and NIR-red (b) dynamics for the same district over three years with variations in winter wheat yields.

Therefore, the following question arises: What is the added value that can be achieved, when incorporating surface reflectances into empirical regression crop yield models? However, adding more regressors in the model (Equation (6)), while having a limited number of reference data, might lead to model overfitting—the model's performance will improve on training (calibration) data, while failing to generalize, i.e., to provide a reasonable performance on independent (previously unseen) data. In order to prevent overfitting of the model (Equation (6)) with multiple regressors, we used a ridge regression [57] with the L_2 (or Tikhonov's) regularization [58,59]:

$$\hat{W} = \underset{W}{\operatorname{argmin}} (\|Y - WR\|^2 + \alpha \|W\|^2), \quad (10)$$

where Y , R , and W are vectors of yields, regressors, and weights of the regression, respectively, and α is the regularization coefficient. The regularization term, $\alpha \|W^2\|$, prevents coefficients of the regression increasing during the minimization process, and therefore prevents overfitting, and might improve the generalization ability of the model. Such a constraint can either force the coefficients, W , to be “sparse” or reflect prior knowledge, namely correlation between regressors. Table 1 summarizes the regression models used in the study. The optimal regularization parameter, α , is determined during the training (calibration) process by dividing calibration data into training and testing datasets, where α is selected in such a way that it maximizes a model’s performance for testing data. Note that validation data were not used for selecting the optimal regularization parameter.

Table 1. Summary of the types of regression models used in the study.

Parameters	Metric	Number of Regressors, n_r	Examples
Single VI	Peak	1	$y_{yield} = \omega_0 + \omega_1 p_{DVI}$
Single VI or SR	AUC	1	$y_{yield} = \omega_0 + \omega_1 s_{DVI}$
Multiple SRs	AUC	2–4	$y_{yield} = \omega_0 + \omega_1 s_{NIR} + \omega_2 s_{red}$
Multiple SRs	a_0, a_1, a_2 , AUC	4–16	$y_{yield} = \omega_0 + \omega_1 a_{0,NIR} + \omega_2 a_{1,NIR} + \omega_3 a_{2,NIR} + \omega_4 s_{NIR} + \omega_5 a_{0,red} + \omega_6 a_{1,red} + \omega_7 a_{2,red} + \omega_8 s_{red}$

3.4. Implementation and Performance Evaluation

VI (NDVI, EVI2, and DVI—Equations (1)–(3)) and SRs (green, red, NIR, SWIR1, SWIR2) were derived from HLS over the study area. Regression models of yields as a function of satellite-derived features were evaluated using a standard set of metrics, such as coefficient of determination (R^2), root mean square error (RMSE) between estimate and reference, and relative RMSE (RRMSE):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{est} - y_{ref})^2}{\sum_{i=1}^n (\bar{y}_{ref} - y_{ref})^2}, \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{est} - y_{ref})^2}, \quad (12)$$

$$RRMSE = 100\% \times \frac{RMSE}{\bar{y}_{ref}}, \quad (13)$$

where y_{est} and y_{ref} are estimate and reference values, respectively; $\bar{y}_{ref} = \frac{1}{n} \sum_{i=1}^n y_{ref}$ is the average of the reference values, and n is the number of samples.

The developed crop yield models were evaluated with two versions of leave-one-out cross-validations (CVs). Within spatial CV, we iteratively removed all samples for a particular district, which was reserved for testing purposes, and the model was calibrated on the rest of the districts using all three years (2016–2018). Within temporal CV, the model was trained on two years, while the third year was reserved for testing. Spatial CV allowed us to explore the spatial robustness of the empirical models, while temporal CV allowed us to explore the temporal robustness.

4. Results

4.1. Winter Crop Type Mapping

Figure 6 shows results of the validation for 2016 to 2018, when comparing the areas of winter crops from official statistics and those derived from satellite data, namely Landsat 8 and Sentinel-2, using the GMM approach.

Overall, there is a good correspondence between satellite-derived and reference areas, with an average R^2 of 0.91 and a bias of +4.3 thousand ha. This bias might be attributed to the confusion between winter crops and early spring cereal crops [60].

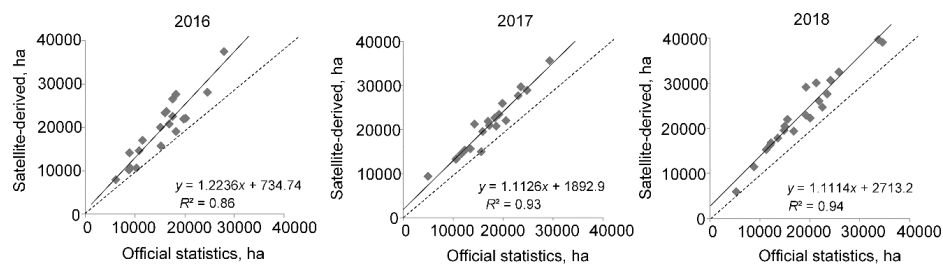


Figure 6. Comparison between winter crop areas from official statistics and satellite at the district scale in Kirovohradska oblast in 2016 to 2018.

4.2. Fitting VI and SR

Table 2 provides a summary of fitting metrics for various VI and SR values derived from HLS for 2016 to 2018.

Table 2. Summary (average and standard deviation) of metrics (RMSE and R^2), when fitting satellite-derived VI and SR values against AGDD using a quadratic model for 2016 to 2018. The percentages show a fraction of winter crop pixels with >6 valid HLS observations, for which metrics were calculated.

2016 (47%, Total 8,207,432 Winter Crop Pixels)						
Value	RMSE		Average y_{sat}		R^2	
	av.	std.	av.	std.	av.	std.
DVI	0.016	0.012	0.184	0.056	0.94	0.09
EVI2	0.025	0.020	0.393	0.083	0.94	0.10
NDVI	0.033	0.017	0.603	0.107	0.94	0.09
Green	0.005	0.002	0.053	0.007	0.61	0.25
Red	0.006	0.003	0.048	0.011	0.81	0.16
NIR	0.017	0.011	0.232	0.050	0.90	0.12
SWIR1	0.015	0.007	0.181	0.026	0.76	0.20
SWIR2	0.015	0.008	0.130	0.031	0.85	0.14
2017 (70%, Total 9,137,931 Winter Crop Pixels)						
Value	RMSE		Average y_{sat}		R^2	
	av.	std.	av.	std.	av.	std.
DVI	0.014	0.009	0.168	0.051	0.93	0.10
EVI2	0.022	0.015	0.350	0.076	0.93	0.11
NDVI	0.027	0.015	0.578	0.106	0.94	0.11
Green	0.004	0.002	0.053	0.006	0.60	0.26
Red	0.004	0.002	0.050	0.011	0.83	0.19
NIR	0.014	0.009	0.219	0.045	0.91	0.11
SWIR1	0.011	0.006	0.186	0.027	0.78	0.21
SWIR2	0.010	0.007	0.138	0.032	0.88	0.15
2018 (85%, Total 10,334,332 Winter Crop Pixels)						
Value	RMSE		Average y_{sat}		R^2	
	av.	std.	av.	std.	av.	std.
DVI	0.015	0.010	0.181	0.059	0.88	0.13
EVI2	0.025	0.016	0.349	0.089	0.88	0.13
NDVI	0.030	0.017	0.616	0.119	0.90	0.12
Green	0.005	0.002	0.052	0.007	0.59	0.27
Red	0.005	0.003	0.050	0.012	0.77	0.21
NIR	0.015	0.009	0.230	0.052	0.83	0.17
SWIR1	0.013	0.007	0.180	0.030	0.72	0.24
SWIR2	0.012	0.008	0.129	0.036	0.79	0.20

All VIs (DVI, EVI2, and NDVI) showed a very good correspondence to a quadratic model (Equation (4)), with an average $R^2 > 0.9$. Among SRs, the best performance was for NIR, with an average $R^2 = 0.88$. Other SRs, except green, showed a variable performance, with R^2 ranging from 0.7 to 0.9, with the green spectral band yielding an average R^2 of 0.6. Nevertheless, all VIs and SRs (except green) exhibited good performance with a relative error of fitting less than 10% (for the green band, the error was 10.2%). It should also be noted that the fitting results were stable across multiple years.

4.3. Comparison Between VI-Based Yield Models

A comparison of VI-based empirical models for different years is presented in Table 3.

Table 3. Winter wheat yield model performance (R^2 , RMSE, RRMSE) on calibration data depending on VI and peak/AUC feature for 2016 to 2018. Best models for each year are marked with italics.

Model	R^2	RMSE, t/ha	RRMSE, %	p -Value
2016				
Peak-DVI (data)	0.179	0.308	7.7	5.61×10^{-2}
Peak-DVI (fitting)	0.332	0.278	7.0	6.29×10^{-3}
<i>AUC-DVI</i>	<i>0.588</i>	<i>0.218</i>	5.5	5.02×10^{-5}
Peak-EVI2 (data)	0.056	0.330	8.3	3.03×10^{-1}
Peak-EVI2 (fitting)	0.282	0.288	7.2	1.32×10^{-2}
AUC-EVI2	0.209	0.302	7.6	3.71×10^{-2}
Peak-NDVI (data)	0.088	0.325	8.1	1.92×10^{-1}
Peak-NDVI (fitting)	0.485	0.244	6.1	4.55×10^{-4}
AUC-NDVI	0.057	0.330	8.3	2.98×10^{-1}
2017				
Peak-DVI (data)	0.422	0.247	7.1	2.63×10^{-3}
Peak-DVI (fitting)	0.400	0.252	7.2	3.67×10^{-3}
<i>AUC-DVI</i>	<i>0.589</i>	<i>0.208</i>	6.0	1.26×10^{-4}
Peak-EVI2 (data)	0.405	0.251	7.2	3.40×10^{-3}
Peak-EVI2 (fitting)	0.381	0.256	7.3	4.89×10^{-3}
AUC-EVI2	0.570	0.213	6.1	1.87×10^{-4}
Peak-NDVI (data)	0.388	0.254	7.3	4.38×10^{-3}
Peak-NDVI (fitting)	0.393	0.253	7.3	4.05×10^{-3}
AUC-NDVI	0.407	0.250	7.2	3.28×10^{-3}
2018				
Peak-DVI (data)	0.597	0.176	4.7	4.53×10^{-4}
Peak-DVI (fitting)	0.571	0.182	4.8	7.08×10^{-4}
<i>AUC-DVI</i>	<i>0.608</i>	<i>0.174</i>	4.6	3.66×10^{-4}
Peak-EVI2 (data)	0.565	0.183	4.9	7.81×10^{-4}
Peak-EVI2 (fitting)	0.507	0.195	5.2	1.97×10^{-3}
AUC-EVI2	0.532	0.190	5.1	1.34×10^{-3}
Peak-NDVI (data)	0.406	0.214	5.7	7.92×10^{-3}
Peak-NDVI (fitting)	0.349	0.224	6.0	1.60×10^{-2}
AUC-NDVI	0.202	0.248	6.6	8.07×10^{-2}

The biggest impact of smoothing satellite data for peak-VI models was observed in 2016, where fewest cloud-free data were available compared to 2017 and 2018 (Figure 3a). The coefficient of determination, R^2 , was 0.179, 0.056, and 0.088 for the peak-DVI, EVI2, and NDVI models, respectively, without phenological fitting, when the peak was calculated directly from available HLS datasets; those R^2 values increased to 0.332, 0.282, and 0.485, when calculating the peak p_{VI} (Equation (8)) from phenological fitting. The best performance for 2016 was, however, for the AUC-DVI approach, with an $R^2 = 0.588$. This approach gave the best performance among all VI-based approaches considered, providing a consistent performance of $R^2 = 0.589$ and $R^2 = 0.608$ for 2017 and 2018, respectively (Figure 7).

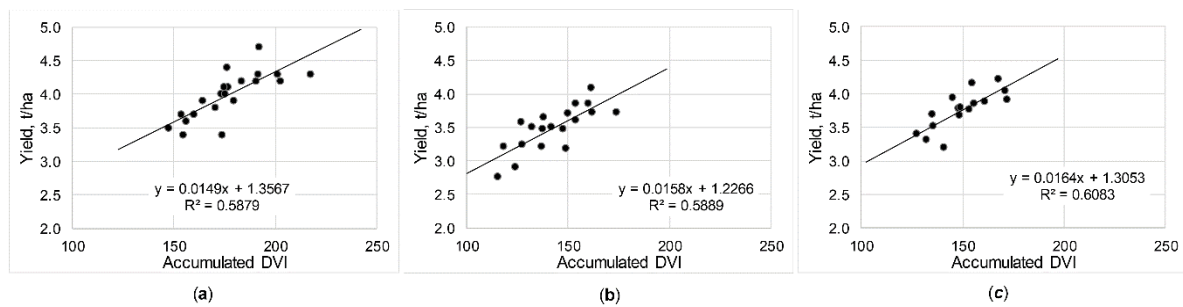


Figure 7. Regression crop yield models based on DVI and AUC for 2016 (a), 2017 (b), and 2018 (c).

Performance of EVI2 and NDVI varied within 2016 to 2018, yielding a good performance for one year and failing to produce good results for another year. Interestingly, all models consistently improved from 2016 to 2018 thanks, most probably, to the availability of more cloud-free data (Figure 3a). On average, the number of cloud-free observations for winter crops within the March to June period for 2018 was 1.3 and 1.8 times larger compared to 2017 and 2016, respectively. This shows the importance of having dense and high-frequency satellite observations at a moderate spatial resolution (10–30 m) for crop yield studies.

4.4. Impact of the Combined Use of Landsat 8 and Sentinel-2 Data

We used the AUC-DVI crop yield model as a benchmark to evaluate the combined use of Landsat 8/OLI and Sentinel-2/MSI (both A and B satellites) data as compared to a single sensor usage. Performance metrics, namely R^2 and $RRMSE$, are shown in Figure 8. As with the results from the previous subsection, the biggest impact was observed in 2016. With a single sensor usage (OLI or MSI), it was impossible to derive satellite-based features for some districts in 2016, because of not having enough cloud-free observations. In 2016 (March–June period), there were on average 2.6 ± 1.4 cloud-free observations from Landsat 8/OLI and 3.6 ± 1.6 cloud-free observations from Sentinel-2/MSI. With a single sensor usage in 2016, there was a poor correlation observed between crop yields and satellite-derived features, even with phenological fitting.

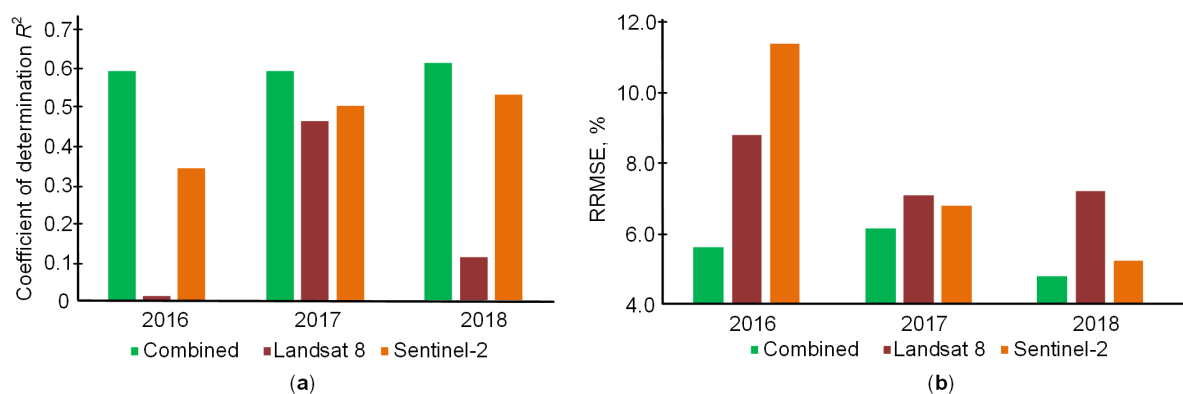


Figure 8. Performance of the AUC-DVI empirical winter wheat yield model on calibration data depending on the combined use of Landsat 8/OLI and Sentinel-2/MSI and a single usage. The coefficient of determination R^2 (a) and relative RMSE (b) are shown.

In 2017 and 2018, the combined use of Landsat 8/OLI and Sentinel-2/MSI outperformed a single sensor usage, though less impact was observed in 2018, when both Sentinel-2A and -2B provided a nominal five-day revisit cycle. Though Landsat 8/OLI has a 16-day revisit cycle (8 days for overlapping scenes) compared to 5 days from both Sentinel-2's, it provided a valuable contribution, accounting for, on average, 30% to 40% of total cloud-free observations in the March to June period in the study area. These results highlight the importance of multi-source satellite data usage for crop yield studies.

4.5. Comparison Between VI- and SR-Based Yield Models

Results of the cross-validation (CV) of different models incorporating VIs, SRs, and derived AUC (Equation (9)) and coefficients a_0 , a_1 , and a_2 of the parabola (Equation (4)) are presented in Table 4.

Table 4. Results of the cross-validation of different models. For each group, the best models are marked in italics.

Model	Spatial CV			Temporal CV		
	RMSE, t/ha	RRMSE, %	R^2	RMSE, t/ha	RRMSE, %	R^2
VI-based and AUC						
AUC–DVI	0.226	6.0	0.65	0.257	6.9	0.60
AUC–NDVI	0.334	8.9	0.24	0.408	10.9	0.15
AUC–EVI2	0.271	7.2	0.50	0.323	8.6	0.45
<i>AUC–NIR</i>	<i>0.226</i>	<i>6.0</i>	<i>0.65</i>	<i>0.236</i>	<i>6.3</i>	<i>0.63</i>
AUC–red	0.396	10.6	0.18	0.479	12.8	0.31
AUC–green	0.368	9.8	0.10	0.408	10.9	0.00
AUC–SWIR1	0.388	10.3	0.01	0.459	12.3	0.19
SR-based and AUC						
AUC–NIR + red	0.229	6.1	0.64	0.253	6.7	0.60
<i>AUC–NIR + green</i>	<i>0.229</i>	<i>6.1</i>	<i>0.64</i>	<i>0.244</i>	<i>6.5</i>	<i>0.62</i>
AUC–NIR + SWIR1	0.228	6.1	0.65	0.249	6.6	0.61
AUC–red + green	0.289	7.7	0.43	0.427	11.4	0.35
AUC–red + SWIR1	0.337	9.0	0.24	0.356	9.5	0.14
AUC–green + SWIR1	0.357	9.5	0.15	0.457	12.2	0.04
AUC–NIR + red + green + SWIR1	0.237	6.3	0.62	0.268	7.1	0.53
VI/SR-based, and a_0 , a_1 , a_2 , AUC						
AUC, coefficients–DVI	0.217	5.8	0.68	0.218	5.8	0.68
AUC, coefficients–NDVI	0.283	7.6	0.45	0.328	8.8	0.30
AUC, coefficients–EVI2	0.272	7.3	0.50	0.336	9	0.38
<i>AUC, coefficients–NIR + red</i>	<i>0.207</i>	<i>5.5</i>	<i>0.71</i>	<i>0.201</i>	<i>5.4</i>	<i>0.73</i>
AUC, coefficients–NIR + green	0.218	5.8	0.68	0.233	6.2	0.63
AUC, coefficients–NIR + SWIR1	0.222	5.9	0.67	0.221	5.9	0.67
AUC, coefficients–red + green	0.249	6.6	0.58	0.366	9.8	0.53
AUC, coefficients–red + SWIR1	0.283	7.6	0.45	0.291	7.8	0.43
AUC, coefficients–green + SWIR1	0.359	9.6	0.16	0.486	13.0	0.03
AUC, coefficients–NIR + red + green + SWIR1	0.212	5.7	0.70	0.218	5.8	0.73

Among the single regressor-based models, the best performance was achieved using an NIR spectral band, with an RMSE = 0.236 t/ha (6.3%) and $R^2 = 0.63$ for temporal CV. This model outperformed VI- and SR-based models. Among VIs, the best performance was achieved using DVI, with an RMSE = 0.257 t/ha (6.9%) and $R^2 = 0.6$.

Having two AUC-SR-based regressors did not improve on the model compared to the AUC–NIR based model. The best results were achieved for the AUC–NIR + green model (RMSE = 0.244 t/ha, 6.5%, $R^2 = 0.62$), which did not improve on the single regressor AUC–NIR; however, it did improve the AUC–VI-based models. This shows the importance of adding separate SR values, rather than using VIs with pre-calculated weights for spectral bands. Surprisingly, the AUC–NIR outperformed the AUC–NIR + red model and other combinations of spectral bands, meaning that AUC-derived features for green, red, and SWIR1 spectral bands did not add much information in terms of winter wheat yields. SWIR2 data were not included in the analysis, because of considerable correlation between SWIR1 and SWIR2.

The addition to the AUC of the coefficients of the quadratic relationship (Equation (4)) as regressors, which define the full dynamics in the multi-dimensional spectral space (Figure 5), improved estimations of winter wheat yields. In particular, the AUC, coefficients–NIR + red model showed the best overall results, with RMSE = 0.201 t/ha (5.4%) and $R^2 = 0.73$ (Figure 9). It shows that not only

is the AUC important but also the full trajectories itself, defined by coefficients a_0 , a_1 , a_2 . Since this model has eight regressors (AUC, a_0 , a_1 , a_2 for NIR and red), it easily overfits for a small sample size, such as in this study, and, therefore, the role of regularization (Equation (10)) becomes extremely important. The derived optimal regularization coefficient was $\alpha = 10^{-5}$. Without regularization, the AUC, coefficients–NIR + red model yields an RMSE = 0.296 t/ha (7.9%) and $R^2 = 0.49$ performance, which is at least 30% worse than with regularization. Adding all spectral bands (NIR, red, green, SWIR1) to the AUC, coefficients-based model did not improve the AUC, coefficients–NIR + red model; however, it showed the second best result among all the models considered in the study along with AUC, coefficients–DVI model. It is also worth noting that performance metrics within spatial CV were consistently better than metrics for temporal CV, meaning a better robustness of the models in the spatial context, rather than temporal.

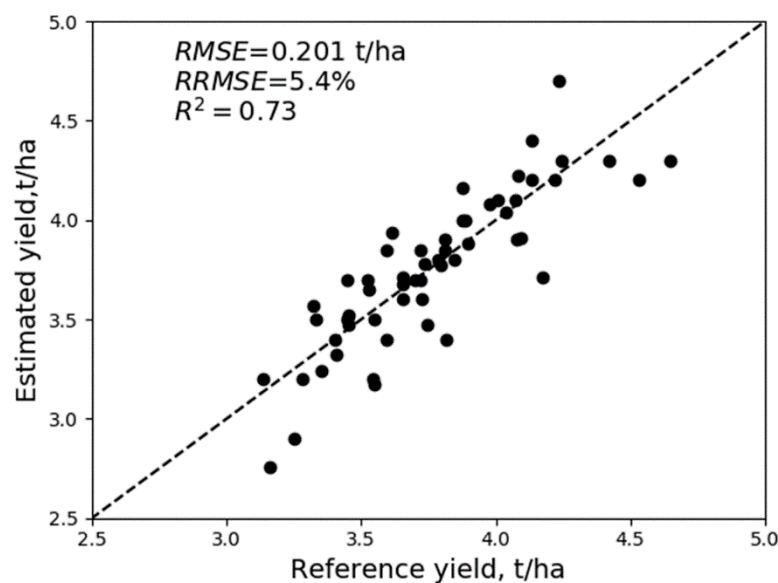


Figure 9. Estimates vs. reference winter wheat yields within the temporal cross-validation for the best AUC, coefficients–NIR + red model.

5. Discussion

Remote sensing has always been a valuable source of information for crop yield. Previous success has been achieved chiefly by using coarse spatial resolution satellite data, thanks to its high temporal resolution (daily), which is needed to capture the behavior of crop growth. Vegetation indices that incorporate two or more spectral bands with pre-defined weights have usually been applied to correlate with crop yields. With recent advancements in technologies, high temporal resolution satellite data have been available at moderate (10–30 m, combination of free Landsat 8 and Sentinel-2) and high (~3 m, commercial Planet) spatial resolutions. However, most existing studies continue to exploit vegetation indices, disregarding mainly surface reflectance values (SRs) in the available spectral bands. This study aimed to address this question and explored the use of SRs, namely green (~0.560 μm), red (~0.660 μm), NIR (~0.865 μm), and SWIR1 (~1.5 μm), when estimating winter wheat yields. In order to account for discrepancies in satellite data acquisitions, we applied a quadratic relationship (Equation (4)) to SRs of individual spectral bands against accumulated GDD (Equation (5)). This relationship was previously applied to VIs only and has never been applied to SRs. Our results showed that this quadratic relationship can be effectively applied to SRs as well, with a relative error of fitting ranging from 5% to 12%, and R^2 from 0.59 to 0.91. The best results of phenological fitting for SR values were observed for NIR (average $R^2 = 0.88$), while the green spectral band showed the worst performance of $R^2 = 0.6$. While these results show the adequacy of a quadratic function, we would like to emphasize the potential limitations of this fitting model. The use of a parabola implies that the upward phase

of crop growth is mirrored in the senescence, which is the case for our study area. However, in the regions, where water is more limiting to crop growth than sunlight, such an assumption might be incorrect. Therefore, when fitting the parabola to the satellite-derived VIs or SRs, one should always check the fitting performance, and in case of poor fitting, other fitting models should be explored to better capture crop phenology.

By having those quadratic relationships in place for SRs, multiple phenological metrics can be derived. A similar approach with multiple phenological metrics has been successfully applied to cropland and grassland mapping in the USA for multiple years using Landsat and MODIS data, though only EVI was used for fitting [50]. Therefore, phenological metrics, such as area under curve (AUC) (Figure 4) and coefficients of the quadratic relationship (Equation (4)), can be used to estimate crop yields. Since adding multiple regressors into a regression model might lead to model overfitting (especially in the case of a small sample size), we proposed the use of ridge regression [53], which incorporates an L_2 regularization [58,59]. Our best winter wheat yield model was a combination of AUC and constant, linear, and quadratic coefficients of the quadratic model for NIR and red spectral bands, with an RMSE = 0.201 t/ha (5.4%) and $R^2 = 0.73$ (Figure 9). This model outperformed VI-based models. Adding other spectral bands, namely green and SWIR1, did not improve the model performance for winter wheat yields. These results confirm previous findings for winter wheat yields studies, where NIR and red play the most important role in capturing the biophysical parameters of winter wheat [12,61]. For other crops, for example, corn and soybean, a similar approach can be applied to identify spectral bands that maximize the correlation with crop yields. However, these should be investigated further.

We also investigated—in a numerical way—the efficiency of the combined use of Landsat 8/OLI and Sentinel-2/MSI for winter wheat yield estimation. All three years featured in the study (2016–2018) had different scenarios of satellite data acquisitions, which benefited such an analysis. The most substantial impact was observed in 2016, when only one of the two Sentinel-2 satellites was available, however, was not fully operational (revisit cycle >10 days). Using a single sensor only in 2016, either Landsat 8/OLI or Sentinel-2A/MSI, did not give any reasonable results for the crop yield model, since each satellite contributed with 2.6 ± 1.4 and 3.6 ± 1.6 cloud-free observations in the March to June period, respectively, which was not enough to capture the growth dynamics of winter wheat. The single-sensor-only models provided $R^2 = 0.01$ and $R^2 = 0.34$ for Landsat 8 and Sentinel-2A, respectively. A combination of satellites increased the number of observations to 6.2 ± 2.2 , and the performance of the winter wheat yield model improved to $R^2 = 0.59$. For 2017 to 2018, when data from Sentinel-2B started to be available and both Sentinel-2A/B became operational with a combined five-day revisit cycle, that impact was not so dramatic; nevertheless, the combination of Landsat 8 and Sentinel-2 improved the performance of the winter wheat yield model by 18% and 15% in 2017 and 2018, respectively, compared to Sentinel-2A/B only. Though Landsat 8 has a nominal 16-day revisit cycle (8-day for overlapping scenes), which is by far less than two satellites Sentinel-2A/B, it consistently added on average 3.2 ± 1.6 cloud-free observations (during the March to June period), which constituted on average 39% of the total observations. These results further highlight the importance of multi-source data usage in these types of studies.

6. Conclusions

In this paper, we focused on the estimation of winter wheat yield at a regional scale in Ukraine, using the combined Landsat 8 and Sentinel-2 data at a 30-m spatial resolution. We emphasized the importance of normalizing (smoothing) satellite data in crop yield studies through, for example, phenological fitting. Not doing such normalization may result in a considerable decrease in crop yield model performance. Unlike most existing approaches utilizing vegetation indices only, we derived per-pixel phenological metrics directly from all spectral bands (green, red, NIR, and SWIR), when fitting them against AGDD. The derived phenological metrics were correlated against crop yields using a ridge regression (with L_2 regularization) in order to avoid model overfitting. The best results were

achieved for the model with NIR and red spectral bands and derived area under curve (AUC), constant, linear, and quadratic coefficients of the quadratic model. The best model yielded RMSE = 0.201 t/ha (5.4%) and $R^2 = 0.73$ on cross-validation.

We also analyzed the importance of multi-source data usage for crop yield estimation. For the area considered in this study, both Landsat 8/OLI and Sentinel-2A/B/MSI data were valuable, and the combined use yielded better performance than a single sensor usage for all three years (2016–2018). These results reiterate the need for high-frequency observations at moderate and higher spatial resolutions for crop yield studies.

Further works should be directed towards an investigation of other crops, such as corn and soybeans, as well as identification of efficient ways of incorporating synthetic aperture radar (SAR) observations to improve the frequency of data in areas with high cloud cover.

Author Contributions: Conceptualization and methodology, S.S., E.V. and B.F.; data curation, S.S., J.J. and J.M.; validation, S.S. and N.K.; writing—original draft preparation, S.S.; writing—review and editing, E.V., B.F., J.-C.R. and N.K.; supervision, E.V. and J.-C.R.; funding acquisition, S.S., B.F. and J.-C.R.

Funding: This research was funded by the NASA Land-Cover/Land-Use Change (LCLUC) Program, grant number 80NSSC18K0336, and Cooperative Agreement NASA Harvest, grant number 80NSSC18M0039.

Acknowledgments: We would like to thank NASA for providing support in generating the Harmonized Landsat Sentinel-2 (HLS) product. This study used modified Copernicus Sentinel data (2016–2018).

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

References

1. Li, J.; Roy, D.P. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sens.* **2017**, *9*, 902.
2. Justice, C.O.; Vermote, E.; Townshend, J.R.; DeFries, R.; Roy, D.P.; Hall, D.K.; Salomonson, V.V.; Privette, J.L.; Riggs, G.; Strahler, A.; et al. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1228–1249. [[CrossRef](#)]
3. Justice, C.O.; Román, M.O.; Csizsar, I.; Vermote, E.F.; Wolfe, R.E.; Hook, S.J.; Friedl, M.; Wang, Z.; Schaaf, C.B.; Miura, T.; et al. Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *J. Geophys. Res. Atmos.* **2013**, *118*, 9753–9765. [[CrossRef](#)] [[PubMed](#)]
4. Becker-Reshef, I.; Justice, C.; Sullivan, M.; Vermote, E.; Tucker, C.; Anyamba, A.; Small, J.; Pak, E.; Masuoka, E.; Schmaltz, J.; et al. Monitoring global croplands with coarse resolution earth observations: The Global Agriculture Monitoring (GLAM) project. *Remote Sens.* **2010**, *2*, 1589–1609. [[CrossRef](#)]
5. Johnson, D.M. A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 65–81. [[CrossRef](#)]
6. Meroni, M.; Fasbender, D.; Balaghi, R.; Dali, M.; Haffani, M.; Haythem, I.; Hooker, J.; Lahlou, M.; Lopez-Lozano, R.; Mahyou, H.; et al. Evaluating NDVI data continuity between SPOT-VEGETATION and PROBA-V missions for operational yield forecasting in North African countries. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 795–804. [[CrossRef](#)]
7. Becker-Reshef, I.; Vermote, E.; Lindeman, M.; Justice, C. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sens. Environ.* **2010**, *114*, 1312–1323. [[CrossRef](#)]
8. Franch, B.; Vermote, E.F.; Becker-Reshef, I.; Claverie, M.; Huang, J.; Zhang, J.; Justice, C.; Sobrino, J.A. Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. *Remote Sens. Environ.* **2015**, *161*, 131–148. [[CrossRef](#)]
9. Kogan, F.; Kussul, N.; Adamenko, T.; Skakun, S.; Kravchenko, O.; Kryvobok, O.; Shelestov, A.; Kolotii, A.; Kussul, O.; Lavrenyuk, A. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 192–203. [[CrossRef](#)]
10. Whitcraft, A.K.; Becker-Reshef, I.; Killough, B.D.; Justice, C.O. Meeting earth observation requirements for global agricultural monitoring: An evaluation of the revisit capabilities of current and planned moderate resolution optical earth observing missions. *Remote Sens.* **2015**, *7*, 1482–1503. [[CrossRef](#)]

11. Whitcraft, A.K.; Vermote, E.F.; Becker-Reshef, I.; Justice, C.O. Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sens. Environ.* **2015**, *156*, 438–447. [[CrossRef](#)]
12. Franch, B.; Vermote, E.F.; Skakun, S.; Roger, J.C.; Becker-Reshef, I.; Murphy, E.; Justice, C. Remote sensing based yield monitoring: Application to winter wheat in United States and Ukraine. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *76*, 112–127. [[CrossRef](#)]
13. Kolotii, A.; Kussul, N.; Shelestov, A.; Skakun, S.; Yailymov, B.; Basarab, R.; Lavreniuk, M.; Oliinyk, T.; Ostapenko, V. Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 39–44. [[CrossRef](#)]
14. López-Lozano, R.; Duveiller, G.; Seguí, L.; Meroni, M.; García-Condado, S.; Hooker, J.; Leo, O.; Baruth, B. Towards regional grain yield forecasting with 1km-resolution EO biophysical products: Strengths and limitations at pan-European level. *Agric. For. Meteorol.* **2015**, *206*, 12–32. [[CrossRef](#)]
15. Mkhabela, M.S.; Bullock, P.; Raj, S.; Wang, S.; Yang, Y. Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. *Agric. For. Meteorol.* **2011**, *151*, 385–393. [[CrossRef](#)]
16. Azzari, G.; Jain, M.; Lobell, D.B. Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sens. Environ.* **2017**, *202*, 129–141. [[CrossRef](#)]
17. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218.
18. Gao, F.; Anderson, M.C.; Zhang, X.; Yang, Z.; Alfieri, J.G.; Kustas, W.P.; Mueller, R.; Johnson, D.M.; Prueger, J.H. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sens. Environ.* **2017**, *188*, 9–25. [[CrossRef](#)]
19. Yang, Y.; Anderson, M.C.; Gao, F.; Wardlow, B.; Hain, C.R.; Otkin, J.A.; Alfieri, J.; Yang, Y.; Sun, L.; Dulaney, W. Field-scale mapping of evaporative stress indicators of crop yield: An application over Mead, NE, USA. *Remote Sens. Environ.* **2018**, *210*, 387–402. [[CrossRef](#)]
20. Gao, F.; Anderson, M.; Daughtry, C.; Johnson, D. Assessing the Variability of Corn and Soybean Yields in Central Iowa Using High Spatiotemporal Resolution Multi-Satellite Imagery. *Remote Sens.* **2018**, *10*, 1489. [[CrossRef](#)]
21. Herrmann, I.; Pimstein, A.; Karnieli, A.; Cohen, Y.; Alchanatis, V.; Bonfil, D.J. LAI assessment of wheat and potato crops by VEN μ S and Sentinel-2 bands. *Remote Sens. Environ.* **2011**, *115*, 2141–2151. [[CrossRef](#)]
22. Skakun, S.; Vermote, E.; Roger, J.C.; Franch, B. Combined Use of Landsat-8 and Sentinel-2A Images for Winter Crop Mapping and Winter Wheat Yield Assessment at Regional Scale. *AIMS Geosci.* **2017**, *3*, 163–186. [[CrossRef](#)] [[PubMed](#)]
23. Lambert, M.J.; Traoré, P.C.S.; Blaes, X.; Baret, P.; Defourny, P. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* **2018**, *216*, 647–657. [[CrossRef](#)]
24. Lai, Y.R.; Pringle, M.J.; Kopittke, P.M.; Menzies, N.W.; Orton, T.G.; Dang, Y.P. An empirical model for prediction of wheat yield, using time-integrated Landsat NDVI. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 99–108. [[CrossRef](#)]
25. Atkinson, P.M.; Jegathan, C.; Dash, J.; Atzberger, C. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sens. Environ.* **2012**, *123*, 400–417. [[CrossRef](#)]
26. De Beurs, K.M.; Henebry, G.M. Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sens. Environ.* **2004**, *89*, 497–509. [[CrossRef](#)]
27. Fisher, J.I.; Mustard, J.F.; Vadeboncoeur, M.A. Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sens. Environ.* **2006**, *100*, 265–279. [[CrossRef](#)]
28. Ganguly, S.; Friedl, M.A.; Tan, B.; Zhang, X.; Verma, M. Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sens. Environ.* **2010**, *114*, 1805–1816. [[CrossRef](#)]
29. Jönsson, P.; Cai, Z.; Melaas, E.; Friedl, M.A.; Eklundh, L. A Method for Robust Estimation of Vegetation Seasonality from Landsat and Sentinel-2 Time Series Data. *Remote Sens.* **2018**, *10*, 635. [[CrossRef](#)]
30. Roy, D.P.; Yan, L. Robust Landsat-based crop time series modelling. *Remote Sens. Environ.* **2019**. [[CrossRef](#)]
31. Zhang, X.; Friedl, M.A.; Schaaf, C.B.; Strahler, A.H.; Hodges, J.C.; Gao, F.; Reed, B.C.; Huete, A. Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* **2003**, *84*, 471–475. [[CrossRef](#)]

32. Zhang, X.; Friedl, M.A.; Schaaf, C.B. Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *J. Geophys. Res. Biogeosci.* **2006**, *111*, G04017. [\[CrossRef\]](#)
33. Gallego, F.J.; Kussul, N.; Skakun, S.; Kravchenko, O.; Shelestov, A.; Kussul, O. Efficiency assessment of using satellite data for crop area estimation in Ukraine. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *29*, 22–30. [\[CrossRef\]](#)
34. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [\[CrossRef\]](#)
35. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [\[CrossRef\]](#)
36. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Skakun, S.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [\[CrossRef\]](#)
37. Vermote, E.; Justice, C.; Claverie, M.; Franch, B. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* **2016**, *185*, 46–56. [\[CrossRef\]](#)
38. Doxani, G.; Vermote, E.; Roger, J.C.; Gascon, F.; Adriaensen, S.; Frantz, D.; Hagolle, O.; Hollstein, A.; Kirches, G.; Li, F.; et al. Atmospheric correction inter-comparison exercise. *Remote Sens.* **2018**, *10*, 352. [\[CrossRef\]](#)
39. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [\[CrossRef\]](#)
40. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [\[CrossRef\]](#)
41. Richardson, A.J.; Wiegand, C.L. Distinguishing vegetation from soil background information. *Photogramm. Eng. Remote Sens.* **1977**, *43*, 1541–1552.
42. Molod, A.; Takacs, L.; Suarez, M.; Bacmeister, J. Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geosci. Model Dev.* **2015**, *8*, 1339–1356. [\[CrossRef\]](#)
43. Santamaria-Artigas, A.E.; Franch, B.; Guillevic, P.; Roger, J.-C.; Vermote, E.F.; Skakun, S. Evaluation of Near-Surface Air Temperature from Reanalysis Over the United States and Ukraine: Application to Winter Wheat Yield Forecasting. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**. [\[CrossRef\]](#)
44. Hatfield, J.L.; Prueger, J.H. Temperature extremes: Effect on plant growth and development. *Weather Clim. Extrem.* **2015**, *10*, 4–10. [\[CrossRef\]](#)
45. Skakun, S.; Franch, B.; Vermote, E.; Roger, J.-C.; Becker-Reshef, I.; Justice, C.; Kussul, N. Early season large-area winter crop mapping using MODIS NDVI data, growing degree days information and a Gaussian mixture model. *Remote Sens. Environ.* **2017**, *195*, 244–258. [\[CrossRef\]](#)
46. Lavreniuk, M.; Kussul, N.; Skakun, S.; Shelestov, A.; Yailymov, B. Regional retrospective high resolution land cover for Ukraine: Methodology and results. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 3965–3968.
47. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
48. Alemu, W.G.; Henebry, G.M. Land surface phenologies and seasonalities using cool earthlight in mid-latitude croplands. *Environ. Res. Lett.* **2013**, *8*, 045002. [\[CrossRef\]](#)
49. McMaster, G.S.; Wilhelm, W.W. Growing degree-days: One equation, two interpretations. *Agric. For. Meteorol.* **1997**, *87*, 291–300. [\[CrossRef\]](#)
50. Nguyen, L.H.; Joshi, D.R.; Clay, D.E.; Henebry, G.M. Characterizing land cover/land use from multiple years of Landsat and MODIS time series: A novel approach using land surface phenology modeling and random forest classifier. *Remote Sens. Environ.* **2018**. [\[CrossRef\]](#)
51. Sun, L.; Gao, F.; Anderson, M.; Kustas, W.; Alsina, M.; Sanchez, L.; Sams, B.; McKee, L.; Dulaney, W.; White, W.; et al. Daily mapping of 30 m LAI and NDVI for Grape yield prediction in California Vineyards. *Remote Sens.* **2017**, *9*, 317. [\[CrossRef\]](#)
52. Manjunath, K.R.; Potdar, M.B.; Purohit, N.L. Large area operational wheat yield model development and validation based on spectral and meteorological data. *Int. J. Remote Sens.* **2002**, *23*, 3023–3038. [\[CrossRef\]](#)
53. He, M.; Kimball, J.; Maneta, M.; Maxwell, B.; Moreno, A.; Beguería, S.; Wu, X. Regional crop gross primary productivity and yield estimation using fused Landsat-MODIS data. *Remote Sens.* **2018**, *10*, 372. [\[CrossRef\]](#)

54. Monteith, J.L. Solar radiation and productivity in tropical ecosystems. *J. Appl. Ecol.* **1972**, *9*, 747–766. [[CrossRef](#)]
55. Prince, S.D.; Goward, S.N. Global primary production: A remote sensing approach. *J. Biogeogr.* **1995**, 815–835. [[CrossRef](#)]
56. Tucker, C.J.; Holben, B.N.; Elgin, J.H., Jr.; McMurtrey, J.E., III. Relationship of spectral data to grain yield variation. *Photogramm. Eng. Remote Sens.* **1980**, *46*, 657–666.
57. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
58. Phillips, D.L. A technique for the numerical solution of certain integral equations of the first kind. *J. ACM (JACM)* **1962**, *9*, 84–97. [[CrossRef](#)]
59. Tikhonov, A.N.; Leonov, A.S.; Yagola, A.G. *Nonlinear Ill-Posed Problems*; Chapman & Hall: London, UK, 1998.
60. Kussul, N.; Skakun, S.; Shelestov, A.; Lavreniuk, M.; Yailymov, B.; Kussul, O. Regional scale crop mapping using multi-temporal satellite imagery. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 45–52. [[CrossRef](#)]
61. Shelestov, A.; Kolotii, A.; Skakun, S.; Baruth, B.; Lozano, R.L.; Yailymov, B. Biophysical parameters mapping within the SPOT-5 Take 5 initiative. *Eur. J. Remote Sens.* **2017**, *50*, 300–309. [[CrossRef](#)]



© 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/>).