Forecast of Yield of Major Crops in Ukraine in War Conditions 2022 Based on MODIS and Sentinel-2 Satellite Data

Nataliia Kussul[1,2,3], Sophia Drozd[1] and Hanna Yailymova [1,2]

[1]Educational and Research Institute of Physics and Technology, Igor Sikorsky Kyiv Polytechnic Institute, Peremohy Avenue 37, Kyiv, Ukraine

[2]Department of Space Information Technologies and System, Space Research Institute National Academy of Science of Ukraine an State Space Agency of Ukraine, Glushkov Avenue 40, Kyiv, Ukraine

[3]Anhalt University of Applied Sciences, Bernburger Str. 57, Köthen, Germany

nataliia.kussul@gmail.com, sofi.drozd.13@gmail.com, anna.yailymova@gmail.com

Abstract: Ukraine was one of the main exporters of plant products. However, as a result of the aggression, the country's agriculture has suffered greatly, export volumes are decreasing, which may provoke a shortage of agricultural products on world markets. It is impossible to assess crop yield and forecast the harvest volume locally, as the collection of information has become difficult due to the active conduct of hostilities and the occupation of a large part of the territories. Therefore, it is necessary to use land remote sensing data to assess crop yield. In this research, we will build regression models based on a random forest for each region of Ukraine to estimate crop yield based on 16-day composites of the NDVI time series during the summer vegetation period from Sentinel-2 (10m) and MODIS (500m) satellites, involving in the calculation NDVI crop maps. The official yield of maize, sunflower, soybean, rapeseed, and wheat for the years 2016-2021 was used as training data. According to the results of the analysis, models based on NDVI from the MODIS satellite showed better accuracy (relative error within 8-18%), but models based on NDVI data from Sentinel-2 better described the variance of the predicted yield. During the research, we found a sharp drop in land productivity indicators compared to the productivity of 2021 for the territories of central, southern and eastern Ukraine. According to our estimates based on MODIS data, the average yield at the country level is expected to be 40.98 t/ha for wheat, 57.66 t/ha for maize, 23.57 t/ha for sunflower, 21.06 t/ha for soybeans, 21.15 t/ha for rapeseed. Estimates based on Sentinel-2 data: 43.22 t/ha for wheat, 71.93 t/ha for maize, 26.86 t/ha for sunflower, 22.94 t/ha for soybeans, 28.23 t/ha for rapeseed.

1 INTRODUCTION

For many years in a row, Ukraine has maintained the status of a leading exporter of agricultural products on world markets. According to the USDA, as of 2021, the state provided 46% of the world's sunflower oil exports, 9% of wheat exports, 17% of barley exports, and 12% of maize exports [1].

However, the country's agrarian well-being is now under threat. With the start of hostilities in February 2022, thousands of hectares of fields were damaged by rocket explosions and airstrikes [1]. The physical, chemical and biological characteristics of the soil were affected by shelling and explosions. Many territories, even those that managed to survive, remained unplowed, untreated from pests and weeds. Thus, the official information on sown areas throughout Ukraine as of the beginning of the year is no longer relevant. Previous forecasts regarding the volume of harvesting of the main agricultural crops are also not relevant, and it is very difficult to make new forecasts based on information from farmers and agronomists.

The yield forecasting process was previously based on a combination of data collected by local farmers directly on the ground with remote sensing data from satellites. Thanks to the combination and comparison of biophysical and satellite models [3], [3], [5], it was possible to achieve high accuracy of land productivity forecasts even in the early stages of vegetation [6] and to calculate the amount of harvest from certain land plots in advance. However, now, in wartime, the process of obtaining information about growth dynamics directly from local specialists has

become much more complicated. The problem of data availability is particularly acute in the occupied territories and territories of active hostilities in the east and south of Ukraine, which were the main suppliers of agricultural crops before the war.

Thus, the task of assessing the productivity of the country's agrarian sector has largely shifted to satellite monitoring and remote sensing data.

Many of the world's leading researchers have already begun to assess the volume of harvested crops in the conditions of war on the territory of Ukraine. Among them are NASA Harvest scientists [7]. Using data from Planet Labs satellites and the European Space Agency's Sentinel-2 mission, they set out to determine how the Russian-Ukrainian conflict is disrupting the global food system. To make the predictions, the scientists used vegetation index data from the Landsat-8 satellite, combining it with many factors of environmental conditions, such as precipitation, soil moisture and temperature during the growing season, obtained from the MERRA-2 NASA reanalysis data set. According to research results, according to NASA Harvest experts, the yield of winter wheat in Ukraine in 2021 will decrease to 4.1 t/ha instead of last year's 4.65 t/ha in 2021. Official data released by the Ministry of Agrarian Policy and Food of Ukraine as of October 2022 fully confirms NASA Harvest forecasts. However, these forecasts give an estimate of the yield at the level of the entire country, and do not allow tracing the situation in each region separately. In addition, the results of research are provided only on the yield of winter wheat, bypassing the spring agricultural crops, in the production of which Ukraine specializes, namely: maize, sunflower, soy, and rapeseed.

The Ministry of Agrarian Policy and Food of Ukraine provided its estimates of the yield of spring crops, but based mainly on data from controlled territories, and the question of the productivity of a large part of the country's cultivated areas remains open and needs an answer.

In this study, we want to estimate the yield of the main spring agricultural crops of Ukraine at the regional level according to the remote sensing data, covering all controlled and occupied agricultural lands of the state.

According to the experience of past studies, good yield predictions at a regional scale can be achieved using a vegetation index calculated from high spatial resolution satellite data such as Sentinel-2 in combination with Landsat missions [8], [9], [10]. However, when using the Sentinel-2 satellite, due to the low frequency of repeated visits of 5 days, there is often a problem of gloom and missing data. It is sometimes impossible to overcome this problem even with composites constructed from long time interval data. Therefore, yield forecasting is also practiced based on the data of a smaller spatial, but denser time step. For example, the Terra MODIS sensor [11], [12], with an average spatial resolution (from 250 m) and a daily frequency of repeated visits is successfully used to solve the problems of crop forecasting.

In this work, we decided to conduct experiments based on satellite data of both satellites - Sentinel-2 and MODIS, with further comparison and assessment of the reliability of the results of yields forecasts.

The purpose of the study is to obtain an accurate yields forecasting of Ukraine's land at the regional level for winter wheat and the main summer crops such as maize, sunflower, soybeans and rapeseed.

As the main method for conducting the research, we use regression analysis based on satellite data of the NDVI time series for the summer vegetation period and official statistical data on the yield of selected crops for previous years.

As a result of the successful conduct of the experiment, we will be able to provide accurate yield forecasts and estimate the volume of production of agricultural plant products in war conditions both in the controlled and non-controlled territories of Ukraine at the regional level.

2 DATA AND MATERIALS

To conduct an experiment on predicting crop yield as input parameters, we used time series of the NDVI index in 16-day composites for the summer growing season (June-August) 2016-2022 based on the data of the wide-area Sentinel mission satellites of the Copernicus program and Terra MODIS sensors.

The Multispectral Instrument S2 (MSI) provides high-resolution optical images from 13 spectral bands (10-20m), and captures atmospheric bands with a spatial resolution of 60m. For forecasting, we independently calculated NDVI from the products of the harmonized Sentinel-2 MSI, level A2, as the normalized difference of the red and infrared ranges.

Collections of 16-day Terra MODIS product composites (500m) already include ready-to-use NDVI products that generated based on near-field bands infrared and red range of each scene.

In preprocessing for input data analysis, NDVI binary masks were used for each studied culture, prepared based on classification maps of Ukraine [13], [14] (Figure 1).
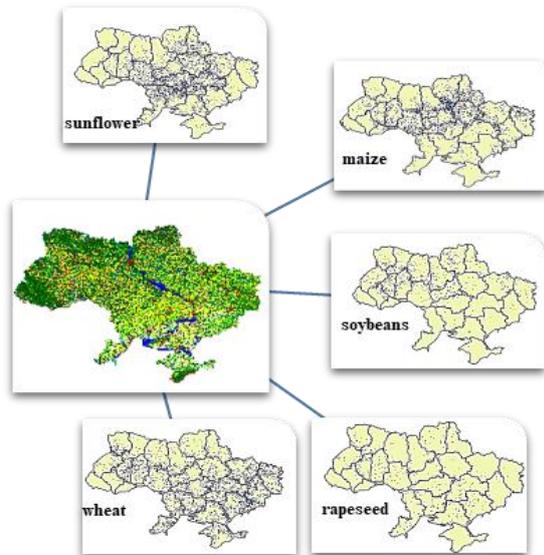
Figure 1: Binary maps of crops from the land cover classification map of Ukraine, 2022.

For each 16-day composite, an area-averaged NDVI was calculated, selecting for calculations only those vegetation index values that covered the target crop pixels. Due to this approach, the quality of input data was significantly improved. Thus, for each region of Ukraine, we obtained a time series that included six NDVI values, two for each summer month. In this time series, we determined the maximum and minimum for each region of Ukraine NDVI during the summer growing season. The total average NDVI for the entire study period was also calculated.

For each region, using correlation analysis, the period that most affects land productivity was also determined for each crop.

As a result, the maximum, average, minimum NDVI and the most correlated with land productivity NDVI for each region and each studied crop were selected as regressors.

3 METHODOLOGY

To predict crop yield for each region of Ukraine, regression models were built based on a random forest ensemble.

Crop yield was defined as the dependent variable, NDVI values as the independent variables.

The defining characteristic of the constructed regression models was the approximation reliability coefficient (1):

$$R^2 = 1 - \sum \frac{(y_t - y_f)^2}{(y_f - \Delta y_{ft})^2} . \qquad (1)$$

where:

y_t − actual yield, according to official statistics,

y_f − predicted yield,

y_{ft} − the average value of all true values of crop yield.

Absolute (2), mean relative error (3), mean square error (4) and root mean square error (5) were the main metrics for assessing forecast accuracy.

$$\Delta y = y_{t-y_f} , \qquad (2)$$

$$MRE = \frac{\sum \frac{\Delta y}{y_t}}{n}, \qquad (3)$$

$$MSE = \frac{(\sum y_i - \bar{y}_i)^2}{n}, \qquad (4)$$

$$RMSE = \sqrt{MSE} . \qquad (5)$$

where:

n – the number of observations,

y_i – the ith observed value of actual yield,

$\bar{y}_i$ – the ith observed value of predicted yield,

Δy – absolute error,

MRE – mean relative error,

MSE – mean square error,

$RMSE$ – root mean square error.

When forecasting, we alternately took one year from 2016 to 2021 for validation, and others, by removing the validation year, for training the model, until we went through all the years. At each iteration, before changing the validation year, the accuracy and errors of the regression model were calculated.

The 2022 data were independently fed into the trained model to generate predictions.

4 EXPERIMENTAL RESULTS

4.1 Assessment of Model Accuracy

After performing the regression analysis, we conducted a general comparison of the accuracies of the models based on the NDVI data from the Sentinel-2 mission and the MODIS sensor, based on the regionally averaged results of the regression models when forecasting for 2021. Models built based on NDVI MODIS demonstrated better accuracy. In particular, MODIS-models showed smaller errors of MAE, RMSE, Relative error (Table 1) compared to models based on NDVI with Sentinel-2 (Table 2) for all crops except rapeseed (MODIS RMSE is 4.56 vs. 3.84 Sentinel-2 RMSE).

However, the MODIS models explained yield variances worse, with the agreement between

predicted and observed yield values for three of the five crops in 2021 (Figure 2).
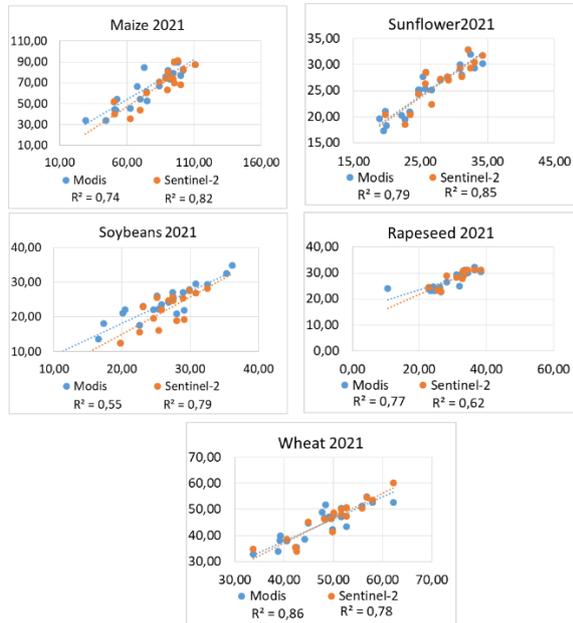

Figure 2: Comparative performance of Random Forests regression models based on Sentinel-2 NDVI and MODIS for 2021 yield test datasets. Plots are shown for five crops: maize, sunflower, soybean, rapeseed, wheat. Dashed lines represent the linear relationship between observations and predictions.

Table 1: Estimates of model errors based on NDVI data from MODIS satellites.

MODIS				
Crop	R2	MAE (c\ha)	RMSE (c\ha)	RE (%)
Maize	0,74	11,90	13,63	0,18
Sunflower	0,79	1,92	2,22	0,08
Soybeans	0,55	2,83	3,53	0,14
Rapeseed	0,77	3,38	4,56	0,16
Wheat	0,86	3,62	4,55	0,09

Table 2: Estimates of model errors based on NDVI data from Sentinel-2 satellites.

Sentinel-2				
Crop	R2	MAE (c\ha)	RMSE (c\ha)	RE (%)
Maize	0,8	17,14	19,17	0,22
Sunflower	0,9	2,15	2,51	0,09
Soybeans	0,8	4,54	5,45	0,2
Rapeseed	0,6	3,44	3,84	0,12
Wheat	0,8	3,54	4,33	0,09

The reliability coefficient of the approximation of forecasts based on MODIS data against forecasts based on Sentinel-2 data was 0.74 vs. 0.82 (maize), 0.79 vs. 0.85 (sunflower), 0.55 vs. 0.79 (soy). However, when forecasting the yield of wheat and rape, the coefficient of approximation of forecasts based on NDVI MSHDIS dominated (0.77 vs. 0.62 (rapeseed), 0.86 vs. 0.78 (wheat)).

In general, the forecasting results for each of the crops based on the data from both satellites were satisfactory with a maximum relative prediction error of 22% (Maize, models based on NDVI with Sentinel-2). Thus, the approach chosen by us is suitable for obtaining reliable yield forecasts and can be applied to calculate crop yield in 2022.

When constructing regression models based on Sentinel-2 satellites, we encountered the problem of missing data due to cloudiness caused by infrequent updates of Sentinel-2 data, even though we used a long time series. Therefore, when forecasting the harvest at the regional level for 2022, a greater weighting factor will be given to the results of model predictions based on NDVI from MODIS sensors.

4.2 Forecasting the Yield of the Land of Ukraine in 2022

The results of crop yield of each region of Ukraine, obtained by forecasting based on regression models based on NDVI MODIS and Sentinel-2, are provided in Table 3.

To assess the accuracy of the forecasts, we collected information on the average yield of the land of Ukraine and compared the official data with the yield averaged by region based on our predictions (Figure 3).
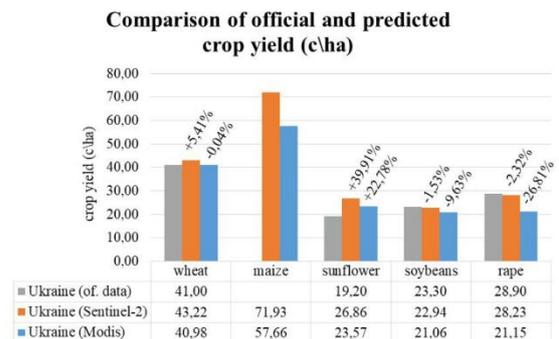

Figure 3: Comparison of official and predicted data on the average yield of land in Ukraine in 2022. The percentages in the chart show the relative deviation of the predicted crop yield from the MODIS and Sentinel-2 satellite data from the official statistics. There were no official data for maize at the time of the research.

Table 3: Predicted crop yield based on NDVI data from MODIS and Sentinel-2 satellites.

Crop / Region	Wheat		Sunflower		Soybeans		Rapeseed		Maize	
	MOD	S2	MOD	S2	MOD	S2	MOD	S2	MOD	S2
Vinnytska	4,81	4,72	2,95	3,26	2,08	2,75	2,08	3,08	6,99	9,00
Volynska	4,55	4,46	2,69	2,90	2,69	2,64	2,69	3,23	8,42	8,98
Dnipropetrovska	3,27	NaN	1,98	NaN	1,52	NaN	1,53	NaN	2,96	NaN
Donetska	3,10	NaN	1,77	NaN	0,00	NaN	0,00	NaN	2,98	NaN
Zhytomyrska	4,31	4,30	2,46	2,46	2,29	2,64	2,31	2,84	6,83	6,80
Zakarpatska	3,34	3,24	2,03	2,15	2,25	2,31	2,28	NaN	5,10	5,27
Zaporizka	2,84	NaN	1,42	NaN	3,18	NaN	3,17	NaN	3,42	NaN
Ivano-Frankivska	5,02	4,68	2,79	2,77	2,41	2,76	2,42	NaN	6,51	6,78
Kyivska	3,90	4,06	2,49	2,78	1,64	2,16	1,62	2,69	5,62	7,77
Kirovohradska	3,53	4,43	2,03	2,50	1,18	1,20	1,22	2,40	3,99	4,95
Luhanska	3,25	NaN	1,74	NaN	0,00	NaN	0,00	NaN	3,15	NaN
Lvivska	4,97	4,78	2,73	2,73	2,66	2,70	2,69	2,95	7,31	7,33
Mykolaivska	3,01	3,11	1,61	1,86	1,05	1,35	1,06	2,08	3,23	3,82
Odeska	2,28	2,96	1,54	1,72	1,54	2,02	1,56	2,37	3,38	4,59
Poltavska	4,24	NaN	2,44	NaN	1,59	NaN	1,57	NaN	5,45	NaN
Rivnenska	4,69	4,72	2,69	2,77	2,56	2,58	2,55	NaN	7,35	8,39
Sumska	5,26	NaN	2,84	NaN	NaN	NaN	NaN	NaN	7,85	NaN
Ternopilska	5,53	5,36	2,82	3,23	2,71	2,92	2,74	3,31	8,17	9,48
Kharkivska	3,75	NaN	2,67	NaN	1,24	NaN	1,33	NaN	4,04	NaN
Khersonska	3,33	3,85	1,49	NaN	3,29	NaN	3,29	NaN	6,82	7,95
Khmelnytska	5,74	5,10	3,41	3,45	2,77	2,66	2,76	3,11	8,95	9,29
Cherkaska	4,16	4,58	2,56	2,82	1,59	1,67	1,53	2,86	5,65	6,84
Chernivetska	5,04	4,09	2,72	2,70	1,90	2,10	1,93	NaN	6,17	5,63
Chernihivska	4,44	5,04	2,70	2,89	2,17	2,24	2,17	2,97	8,04	9,39
Ukraine (mean)	4,10	4,32	2,36	2,69	1,93	2,29	1,94	2,82	5,77	7,19

According to the results of the comparisons, the largest forecast error was recorded in the assessment of sunflower yield. The predicted values exceed the official indicators by 22.78% (MODIS) and 39.91% (Sentinel-2). This can be explained by an incomplete harvest. The smallest error, which was 0.02 t/ha, was found when forecasting wheat yield based on MODIS data (Sentinel-2 – 2.22 t/ha). That is, the regression model absolutely confirmed the official data. Good results were found when predicting the yield of rapeseed. The error of the model based on Sentinel-2ya-2 data is 0.67 t/ha, which is about 2.32% yield. The MODIS-based model, on the contrary, had a high forecast error, which was -26.81% of the official data. Deviations from official data in predicting soybean yields were -1.53% and -9.63% for Sentinel-2 and MODIS-based NDVI models, respectively.

It was not possible to assess the accuracy of the maize yield forecast due to the lack of official statistics.

After comparing the predicted yields with the official data, maps of land productivity dynamics were constructed using the MODIS NDVI predictions to estimate the change in yields relative to the 2021 figures (Figure 4).

As can be seen from the maps, for each crop, the greatest yield decline is expected in the central, eastern and northern parts of the country. An extreme drop in yield is expected in Odesa region and is -46.09% for maize, -45.06% for soybeans, -34.48% for sunflower, -38.57% for rapeseed and -43.72% for wheat. In general, yield declines are observed for every crop in almost every region.
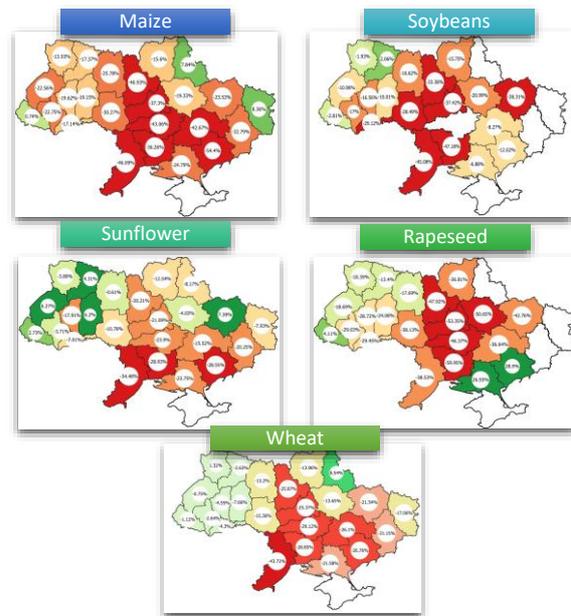

Figure 4: Assessing crop yield dynamics against 2021 yield data.

The exception, according to forecasts, is rapeseed in the Zaporizhzhia and Kherson regions (+28.9% and 26.55% productivity, respectively). Weak positive trends are also observed for the western part of Ukraine, which is certainly explained by the smaller number of active hostilities in the regions.

5 CONCLUSIONS

This study evaluated the effectiveness of crop yield forecasting based on the NDVI time series from MODIS and Sentinel-2 satellites, and forecasted crop yield of Ukraine at the regional level for 2022.

During the experiment, when validating the models on the 2021 test data, it was found that predictions based on NDVI from the Sentinel-2 satellite better describe the variance when comparing predictions with actual yield values, but models based on MODIS sensor data have better accuracy rates.

Although Sentinel-2 has better spatial resolution, due to the wide step of temporal updating of the data, it is difficult to correctly estimate the yield in certain areas, which were overcast during the survey period.

According to our estimates based on MODIS data, the average yield at the country level is expected to be 40.98 t/ha for wheat, 57.66 t/ha for maize, 23.57 t/ha for sunflower, 21.06 for soy, 21.15 t/ha for rapeseed. Estimated based on dataSentinel-2: 43.22 t/ha for wheat, 71.93 t/ha for maize, 26.86 t/ha for sunflower, 22.94 t/ha for soybean, 28.23 t/ha for rapeseed. These indicators are significantly lower than the indicators of crop yield in 2021.

Thus, according to the results of forecasts based on regression models for 2022, we found a sharp deterioration of crop yield in most regions of Ukraine, which is especially pronounced in central Ukraine, in the east and in the south. According to our estimates, the greatest drop in yield is expected in Odesa region (-46.09% for maize, -45.06% for soybeans, -34.48% for sunflower, -38.57% for rapeseed and -43.72% for wheat), as well as in Mykolaiv, Zaporizhzhia, Kherson, Kyiv, Kirovohrad, and Cherkasy. The negative dynamics of yield in these territories is explained by active military operations there, as a result of which large areas of agricultural fields were damaged and the territories were occupied. Thus, the war had an extremely negative impact on crop production in Ukraine, provoking a sharp deterioration in land productivity and, accordingly, decrease in the yield of agricultural crops.

ACKNOWLEDGMENTS

REFERENCES

[1] Svit perebuvaie na pochatku prodovolchoi kryzy, abo yakyi vrozhai vanhuie NASA Harvest Ukraini [The globe is at the beginning of food crisis or what yield NASA Harvest forecasts for Ukraine], 2022, [Online]. Available: https://kurkul.com/ spetsproekty/1338-svit-perebuvaye-na-pochatkoviy-stadiyi-prodovolchoyi-krizi-yakiy-vrojay-vanguye-nasa-harvest-ukrayini.

[2] D. Rawtani, G. Gupta, N. Khatri, P. K. Rao, and C. M. Hussain, "Environmental damages due to war in Ukraine: A perspective," Sci. Total Environ., vol. 850, no. 157932, p. 157932, 2022, doi: 10.1016/j.scitotenv. 2022.157932.

[3] A. Kolotii and et al., "Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XL-7/W3, pp. 39-44, 2015, doi: 10.5194/isprsarchives-xl-7-w3-39-2015.

[4] F. Kogan and et al., "Winter wheat yield forecasting: A comparative analysis of results of regression and biophysical models," J. Autom. Inf. Sci., vol. 45, no. 6, pp. 68-81, 2013, doi: 10.1615/JAutomat InfScien.v45.i6.70.

[5] F. Kogan and et al., "Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models," Int. J. Appl. Earth Obs. Geoinf., vol. 23, pp. 192–203, 2013, doi: 10.1016/j.jag.2013.01.002.

[6] M. E. Holzman and R. E. Rivas, "Early maize yield forecasting from remotely sensed temperature/ vegetation index measurements," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 9, no. 1, pp. 507-519, 2016, doi: 10.1109/JSTARS.2015. 2504262.

[7] T. Ben Hassen and H. El Bilali, "Impacts of the Russia-Ukraine war on global food security: Towards more sustainable and resilient food systems?," Foods, vol. 11, no. 15, p. 2301, 2022, doi: 10.3390/foods 11152301.

[8] S. Skakun and et al., "Winter wheat yield assessment from Landsat 8 and Sentinel-2 data: Incorporating surface reflectance, through phenological fitting, into regression yield models," Remote Sens. (Basel), vol. 11, no. 15, p. 1768, 2019, doi: 10.3390/ rs11151768.

[9] S. Skakun and et al., "Winter wheat yield assessment using Landsat 8 and sentinel-2 data," in IGARSS 2018 - IEEE International Geoscience and Remote Sensing Symposium, 2018, doi: 10.1109/IGARSS.2018. 8519134.

[10] S. Skakun, E. Vermote, J.-C. Roger, and B. Franch, "Combined use of Landsat-8 and sentinel-2A images for winter crop mapping and winter wheat yield assessment at regional scale," AIMS Geosci., vol. 3, no. 2, pp. 163-186, 2017, doi: 10.3934/geosci. 2017.2.163.

[11] I. Becker-Reshef, E. Vermote, M. Lindeman, and C. Justice, "A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data," Remote Sens. Environ., vol. 114, no. 6, pp. 1312-1323, 2010, doi: 10.1016/j.rse.2010. 01.010.

[12] O. Kussul, N. Kussul, S. Skakun, and et al., "Assessment of relative efficiency of using MODIS data to winter wheat yield forecasting in Ukraine," 2013 IEEE International Geoscience and Remote Sensing Symposium, 2013, doi: 10.1109/IGARSS. 2013. 6723516.

[13] A. Shelestov, M. Lavreniuk, V. Vasiliev, and et al., "Cloud approach to automated crop classification using Sentinel-1 imagery," IEEE Transactions on Big Data, 2019, vol. 6. no. 3, pp. 572-582, doi: 10.1109/TBDATA.2019.2940237.

[14] N. Kussul, L. Mykola, A. Shelestov, and S. Skakun, "Crop inventory at regional scale in Ukraine: Developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery," European Journal of Remote Sensing, vol. 51(1), pp. 627-636, doi: 10.1080/22797254.2018. 1454265.