CHEMICAL ENGINEERING TRANSACTIONS

VOL. 96, 2022

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: David Bogle, Flavio Manenti, Piero Salatino

**ISBN** 978-88-95608-95-2; **ISSN** 2283-9216

DOI: 10.3303/CET2296070

# Digitalization of Crop Production for Transition to Climate-Optimized Agriculture Using Spring Wheat in East Kazakhstan as an Example

Marzhan Anuarbekovna Sadenova[a], Nail Alikuly Beisekenov*[,a], Petar Sabev Varbanov[a], Natalya Anatolyevna Kulenova[a], Farid Abitaev[b], Yuri Kamenev[b]

[a]Priority Department Centre «Veritas» D. Serikbayev East Kazakhstan technical university, 19 Serikbayev str. 070000, Ust-Kamenogorsk, Kazakhstan
[b]Experimental farm of oilseeds (EFoO), 070000, Ust-Kamenogorsk, Kazakhstan
MSadenova@ektu.kz

This paper presents the results of testing certain remote sensing indices (vegetation health index, temperature index, normalized drought difference index, and others). Multitemporal data of nine indices from Sentinel-2 and MODIS satellites for the vegetation period over 5 years were used to simulate the spring wheat yield of the Experimental farm of oilseeds. Results of spring wheat yield modeling have shown that the range of mean absolute error (MAE) values was 3.2 to 3.7, and fit index (d) values were 0.78 to 0.86. It was found that the best model $R^2$ = 0.53 based on NDWI, VSDI, and NDVI indices can serve as the most effective predictor for monitoring and forecasting spring wheat yields in eastern Kazakhstan.

## 1. Introduction

Conventional agriculture is a source of significant greenhouse gases emitted by crop and livestock production and contributes to global climate warming. Introducing elements of 'carbon farming' into modern agriculture while actively using GIS technology and remote sensing (RS) techniques creates a realistic case for reducing the carbon (C) footprint of crop production. Zero-tillage techniques, which have the potential to reduce greenhouse gas emissions by 20.6-23.7% compared to conventional tillage (Balafoutis et al., 2017) may have unintended and undesirable impacts on other sources or sinks of greenhouse gases. Soil moisture retention associated with no-tillage should provide more moisture for nitrifying and denitrifying bacteria as well as plants. At the same time, $N_2O$ production may increase, partly or fully offsetting the mitigation potential of carbon sequestration. Agricultural production is an important topic for every country in the world. As agriculture is vulnerable to the effects of climate variability, many studies have been conducted to assess the impact of climate change on the yields of different crops (Rossato et al., 2017). Satellite remote sensing data provides real-time estimates of the magnitude and variation of crop condition parameters (Doraiswamy et al., 2003). Accurate and timely estimation of wheat yield at regional and national scales is essential for crop planning and agricultural development in general (Franch et al., 2021). The development of remotely sensed yield prediction models allows for accurate, reliable, and timely estimates over large areas. In particular, this information is needed to ensure an adequate food supply in the country and to help policymakers and farmers (Sayago et al., 2018).
The use of various multivariate statistical analysis methods such as partial correlation analysis (PCorrA), principal component analysis (PCA), partial least square (PLS), and principal component regression (PCR) algorithms using MATLAB software (Shamsuddin et al., 2021) for the development of smart crop production is not new, but the issue of practical examples for this approach and accumulation of more practical data in different countries remains important. Tuvdendorj et al. (2019) reported that the best model would use a coefficient of determination based on NDWI, VSDI, and NDVI, which are the most effective predictors and reliable remote sensing indexes for monitoring spring wheat yield in northern Mongolia. The agricultural sector occupies one of the main roles in Kazakhstan's economy after mining. The eastern part of Kazakhstan has the

Paper Received: 11 January 2022; Revised: 2 August 2022; Accepted: 19 July 2022
Please cite this article as: Sadenova M.A., Beisekenov N.A., Varbanov P.S., Kulenova N.A., Abitaev F., Kamenev Y., 2022, Digitalization of Crop Production for Transition to Climate-optimized Agriculture Using Spring Wheat in East Kazakhstan as an Example, Chemical Engineering Transactions, 96, 415-420 DOI:10.3303/CET2296070

most favorable natural conditions and a more suitable area for rain-fed crops. A significant part of spring wheat is grown in the eastern provinces with average annual precipitation of 300 to 600 mm. Variations in average annual rainfall largely determine the yield of spring wheat. Vegetation cover, crop yields, and crop growth strongly depend on precipitation and associated soil moisture. Kazakhstan has a sharply continental climate with a short growing season, high evaporation, and low precipitation, which imposes serious constraints on the development of Kazakhstani agriculture. Due to the impact of climate change, more extreme and prolonged droughts occurred in the western part of Kazakhstan, which directly affected vegetation and crop growth, biodiversity, and the socio-economic situation in Kazakhstan. Weather information is commonly used for yield forecasting, but continuous measurements are lacking due to economic factors, among others. The use of Earth observation satellite images to monitor temporal and spatial changes in combination with point observations as joint monitoring has advantages. In addition, satellite images are much more practical than traditional methods and are more affordable to use. The use of remotely sensed data helps to assess the crop status of various fields at regional and country levels, even in remote areas, as it provides timely and accurate measurements. Several field studies have shown that models based on remotely sensed data enable the estimation of crop yields in many countries. Typically, remotely sensed indices are associated with yields using empirical regression models. In recent years, satellite data such as Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat, and Sentinel data have been used for yield forecasting and monitoring, and significant results have been obtained. The linear regression models used in the studies have some limitations for determining soil organic C content. The most obvious shortcoming is that they cannot model non-linear relationships between soil properties. Additional field yield data are needed to develop and validate crop models that are used for production forecasting or as inputs to statistical models for predicting yield anomalies using remote sensing (Fritz et al., 2019).

## 2. Objective and methodology of research

The purpose of this study is to try to assess the contribution of digitalization of crop production in the transition to climate-optimized agriculture. Elements of precision farming have already been introduced in East Kazakhstan to reduce the carbon footprint of crop production in terms of differentiated fertilizer application based on the content of major macronutrients in the soil. The methodology of the current study consists of three independent parts. Nine remote sensing indices were calculated from Sentinel-2, Landsat-8, and MODIS satellite data. Also, a map of soil organic carbon content in the study area is constructed based on satellite data. Based on agrochemical analysis data, soil macroelement content maps were constructed. All these data are used to correlate the actual spring wheat yield as input data for testing the spring wheat yield estimation model being created on the experimental farm in East Kazakhstan.

The research was conducted on the example of spring wheat crops. The plot is located at coordinates 82.644140°E and 50.115674°N and occupies 94.7 ha on the farm "Experimental farm of oilseeds" (EFoO) in Eastern Kazakhstan. According to soil taxonomy, chernozem soil type is the dominant soil unit on the territory of this plot. Processing of the field mask shown in Figure 1, a field mask of spring wheat was developed by remote sensing data from the Sentinel-2 satellite.

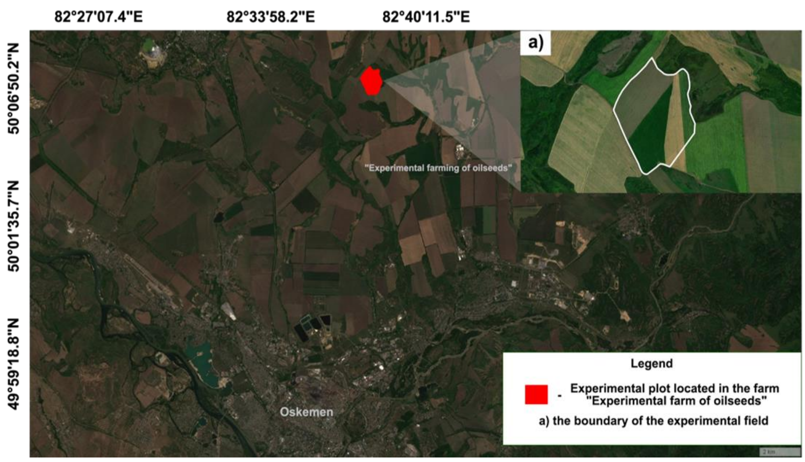


*Figure 1: Study area and experimental field (all data was obtained from Sentinel-2 satellite images)*

To obtain information on NDVI, the algorithms of satellite data processing obtained using API Leaflet and API Sentinel Hub were used, and the normalized values of NDVI per calendar week were obtained and calculated as the average for 5 years of observations. Seven-day (weekly) NDVI index values (cloudless composites MODIS, Sentinel-1-3, Landsat 8) for the calendar year, calculated using the plough mask, were considered. A total of 10 time series of normalized indices corresponding to 2017-2021 were generated. The maximum value of NDVI was reached in calendar weeks 27-31, corresponding to late July-early August (Sadenova et al., 2022). Also, in this study, MODIS/MOD1B time series data with a spatial resolution of 1 km and daily intervals were used to estimate spring wheat yields. MODIS data provided high temporal resolution and wide coverage but low spatial resolution. MODIS time series data for the period (May to October) were obtained for the study area from the Atmospheric Array and Distribution System (LAADS) (NASA) for 2017-2021. The study area covered the area of the Glubokoe District of East Kazakhstan Oblast using MODIS satellite data and its system modules. The reflectance bands (NIR, red, blue, and SWIR) were calculated using calibrated MOD021KM level data. In the processing section, all collected images were re-projected, merged into a mosaic, and calibrated for atmospheric and geometric correction. Nine indices were used in this study: normalized vegetation index (NDVI), normalized difference drought index (NDDI), normalized difference water index (NDWI), vegetation condition index (VCI), temperature condition index (TCI), vegetation health index (VHI), normalized multi-band drought index (NMDI), visible and shortwave infrared drought index (VSDI), and vegetation supply water index (VSWI). Soil agrochemical parameters were assessed by taking soil samples under field conditions. Soil sampling was done in spring 2022 by envelope method at 2 points of the experimental plot of the EFoO farm. It was placed under an agricultural crop: spring wheat. The sampled soil was analyzed by Machigin's method modified by CINAO (GOST 26205-91, 2020) for exchangeable potassium, mobile phosphorus, total nitrogen, and hydrogen index ranking, as well as humus and sulfur content. Meteorological data played a significant role in this study. Meteorological data were collected for 5 years on the experimental farm. The data was used to find correlations and dependencies between agrochemical indicators and spring wheat yields

## 3. Results and discussion

This paper presents the results of vegetation and drought indices (NDVI, NMDI, NDWI, VCI, TCI, VHI, NDDI, VSDI, and VSWI) which are calculated from cloudless and corrected reflectance bands. The organic carbon content of the soil of the experimental farm was investigated. Land degradation dynamics were assessed using the computational module “Trends.Earth”. It was found that soil organic carbon stock changes are related to soil changes. "Negative" changes in soil cover lead to a decrease in soil organic C and vice versa. This method probably cannot provide accurate data for most areas of Kazakhstan because soil C change often occurs without a change in soil cover. This condition is characteristic of cropland, where dehumidification occurs through ploughing of the area and erosion of the topsoil. There can be positive dynamics within a single land class, e.g. with good land management and fertilizer application. The organic C content of the soil must be calculated by taking into account indicators such as relief class, productivity, and exchangeability.

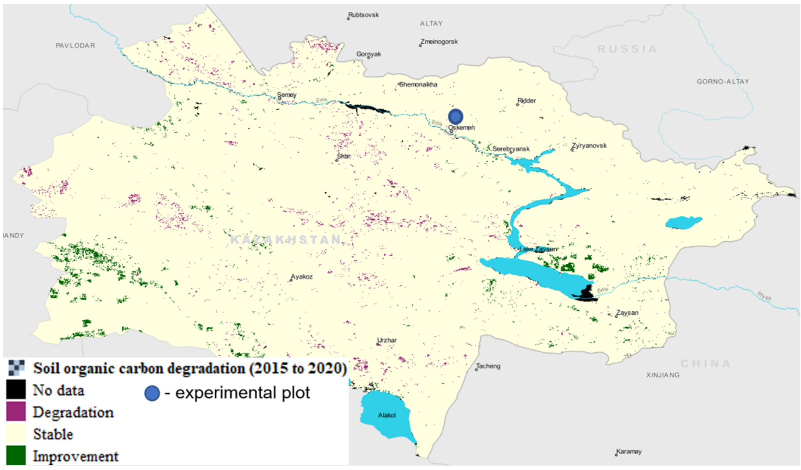


*Figure 2: Changes in soil organic C composition in 2015-2020 in East Kazakhstan*

Figure 2 shows the change in soil organic C composition from 2015 to 2020 by the example of fields in East Kazakhstan. The units of the imported raster layer are metric tonnes of organic C per hectare.

All chemical elements are very important for productive growth and soil enrichment. Nitrogen is vital for the proper development of the root system. All metabolic processes in the plant, from chlorophyll synthesis to vitamin assimilation, are activated by nitrogen. Lack of nitrogen can lead to incomplete yields or even crop failure. Based on the QGIS survey data, an agrochemical cartogram was developed as shown in Figure 3, where a) cartogram of soils' exchangeable potassium content (mg/100 g); b) cartogram of soils' mobile phosphorus content (mg/100 g); c) cartogram of total nitrogen content mg/100 g of soils; d) Ph ranking cartogram; e) humus content cartogram; g) cartogram of soils' sulfur content (mg/kg).

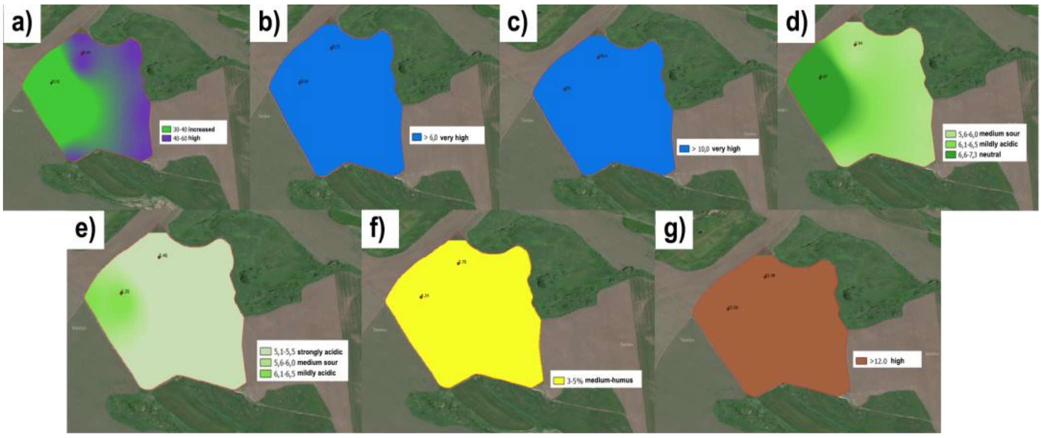


*Figure 3: Cartograms of agrochemical soil analysis of the experimental field of the EFoO (red dots indicate soil sampling in the field of spring wheat) (Image obtained using Sentinel-2 satellite)*

The possibility of creating a predictive measure of spring wheat yield calculated based on satellite data is considered. The methodology consists of three main parts. First, nine remote sensing indices are calculated using MODIS data. Second, nine remote sensing indices are used to correlate the actual spring wheat yield as input data for testing the spring wheat yield estimation model. Thirdly, regression models of spring wheat yield are developed, and the estimated spring wheat yield data for the experimental plot are compared with the 5 y. period. The general flow chart of this study, the processing method, and the individual steps are shown in Figure 4.

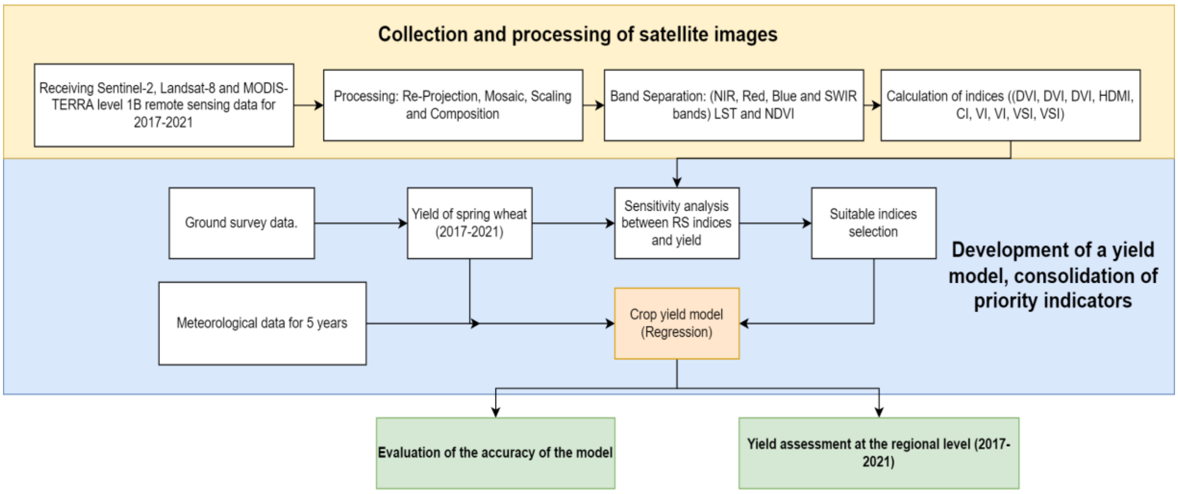


*Figure 4: Flowchart of processing method*

Yields are markedly dependent on growth conditions at each stage of cultivation. Nine typical vegetation and drought indices were calculated and compared with the yield of spring wheat. These indices are commonly used for crop yield estimation and drought monitoring, and only optical spectrum bands are required for calculation. For example, NDVI is most commonly used for vegetation monitoring, yield estimation, and forecasting. Growth and yield conditions are closely linked at the pixel level. The value of each pixel of the remote sensing indices was taken directly from the field mask locations. The relationship between NDVI, NDWI, NMDI, TCI, VCI, VHI, NDDI, VSDI, and VSWI with actual spring wheat yield during the growing season (July-August) for 2017-2021 in Eastern Kazakhstan was investigated. 10 days and monthly, nine remote sensing indices (NDVI, NMDI, NDWI, VCI, TCI, VHI, NDDI, VSDI, and VSWI) derived from red, NIR, blue, and SWIR spectral band transformations were used. a time curve of spring wheat growth indices for each pixel in the study area is presented to a continuous time series of data. In the next step, the month that gave the highest coefficients to the definition was selected to develop multilinear regression models based on all 5 y. of data. Table 1 was used to calculate the remote sensing indices.

*Table 1: Equations of tested nine remote sensing indices*

| N | Remote Sensing Based Indices | Equation |
|---|---|---|
| 1 | Normalized Difference Vegetation Index | $NDVI = \frac{NIR - RED}{NIR + RED}$ |
| 2 | Normalized Difference Water Index | $NDWI = \frac{NIR - SWIR}{NIR + SWIR}$ |
| 3 | Vegetation Condition Index | $VCI = \frac{(NDVIj - NDVI_{min})}{(NDVI_{max} + NDVI_{min})} \times 100\%$ |
| 4 | Temperature Condition Index | $TCI = \frac{T_{max} - Tj}{T_{max} - T_{min}} \times 100\%$ |
| 5 | Vegetation Health Index | $VHI = a \times VCI + (1 - a) \times TCI)$ |
| 6 | Normalized Multi-Band Drought Index | $NMDI = \frac{NIR - (SWIR1 - SWIR2)}{NIR - (SWIR1 + SWIR2)}$ |
| 7 | Vegetation Supply Water Index | $VSWI = Ts/NDVI$ |
| 8 | Normalized Difference Drought Index | $NDDI = \frac{(NDVI - NDWI)}{(NDVI + NDWI)}$ |
| 9 | Visible and Shortwave Infrared Drought Index | $VSDI = 1 = [(SWIR - BLUE) + (RED - BLUE)]$ |

A comprehensive method of model validation is to compare the measured values with the predicted values. We used fitting and model performance statistics: coefficient of determination (R2) root-mean-square error (RMSE) Eq(1), mean absolute error (MAE) Eq(2), bias and agreement index (d) Eq(3).

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Y} - Y)^2}{n}} \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{Y} - Y\right| \tag{2}$$

$$d = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(\left|\hat{Y}_i - Y\right| + |Y_i - \bar{Y}|)^2} \tag{3}$$

where $Y$ - observed values; $\hat{Y}$ - modelled values; $\bar{Y}$- mean value of observed values; n - number of data on crop yields and remote sensing data.

The correlation between nine monthly remote sensing indices, weather data, and agrochemical indices to spring wheat yields during the monitoring period was revealed. The results show the absolute value of the correlation coefficients (R) of each month for the nine indices, which are presented in Table 2. The results were statistically significant $p < 0.01$ with yield, except for NMDI. The relationship between NMDI and spring wheat yields showed the lowest results (0.13-0.14), indicating that this index is not suitable for estimating spring wheat yields in East Kazakhstan. Table 2 shows that monthly NDWI had the highest correlation with yield in June (0.52) and monthly VSDI had the highest correlation with yield in July (-0.58), and they were statistically significant at $p < 0.001$, respectively. This indicates that soil and crop moisture and water content are most important for yield. The results of the spring wheat yield simulation showed that the range of mean absolute error (MAE) values was 3.2 to 3.7, determination coefficient values were R2 = 0.53, and agreement index (d) values were 0.78 to 0.86, respectively. The best model was found to be R2 = 0.53. The low accuracy of the coefficient of determination is due to the specifics of the spectral indices used as well as the lack of a large number of correlated indicators to yield in the machine learning algorithm. This study focuses on reducing carbon footprint in crop production by identifying "hot spots" on fields that do not require fertilization due to high concentration of agrochemical indicators as well as based on a predictive model of yields from remote sensing data.

*Table 2: Multiyear correlation between monthly values and spring wheat yields in June-August (2017-2021)*

| N | Index | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | NDVI | NMDI | NDWI | TCI | VCI | VHI | NDDI | VSDI | VSWI |
| June | 0.41 | 0.2 | 0.52 | 0.48 | 0.32 | 0.44 | -0.32 | -0.52 | -0.40 |
| July | 0.45 | 0.13 | 0.4 | 0.30 | 0.33 | 0.34 | -0.41 | -0.58 | -0.3 |
| August | 0.27 | 0.14 | 0.36 | 0.18 | 0.37 | 0.31 | -0.36 | -0.34 | -0.22 |

Wheat yields are significantly affected by rainfall, soil type, soil moisture, and air temperature changes. In particular, drought and soil moisture deficiency affect the maximum reduction in yield and vegetation size. Thus, based on the results obtained, it is recommended to develop an irrigation system for spring wheat cultivation and to increase the number of yield observation trials in this experimental area. These results obviously show the promising application of NDWI and VSDI data for yield estimation at relatively low cost and timeliness.

## 4. Conclusion

The work establishes that the efficient use of organic carbon, meteorological data, and soil moisture content are the determining factors affecting the yield of spring wheat. All the above parameters, along with field data, are proposed to be verified by remote sensing data. For this purpose, the best and most appropriate indices were first identified by testing the correlations between the nine indices and the actual spring wheat yield. Having summarized all the parameters, we optimized the climatic indices for spring wheat yield prediction using mathematical modeling based on algorithms: coefficient of determination (R2), root mean square error (RMSE), mean absolute error (MAE), and agreement index (d). The results show that NDVI, NDWI, VCI, TCI, VHI, and NMDI indices with spring wheat yield have a positive correlation (0.45, 0.52, 0.37, 0.48, 0.44, and 0.14), while NDDI, VSWI, and VSDI have negative correlation (-0.41, -0.40 and -0.52) respectively. The results show an average correlation to the yields, indicating insufficient data in the machine learning model. In the future, it is planned to improve this model as well as to carry out a more detailed analysis of the indicator dependencies. Moreover, the results confirmed the importance of the integration of satellite and ground data for yield estimation. Based on the results, a limiting indicator was identified for optimizing the conditions as well as further agrarian decisions. The combination of these indicators in the developed methodology allowed not only to monitor and manage crop yields, but also to contribute to the reduction of $CO_2$ emissions into the atmosphere.

**Acknowledgements**

This research has been supported by Project BR10865102 "Development of technologies for remote sensing of the earth (RSE) to improve agricultural management", funded by the Ministry of Agriculture of the Republic of Kazakhstan.

**References**

Balafoutis A., Beck B., Fountas S., Vangeyte J., Van der Wal T., Soto I., Eory V., 2017, Precision agriculture technologies positively contributing to GHG emissions mitigation, farm productivity and economics. Sustainability, 9(8), 1339.

Doraiswamy P.C., Moulin S., Cook P.W., Stern A., 2003, Crop yield assessment from remote sensing. Photogrammetric engineering and remote sensing, 69(6), 665-674.

Franch B., Vermote E., Skakun S., Santamaria-Artigas A., Kalecinski N., Roger J.C., Sobrino J.A., 2021, The ARYA crop yield forecasting algorithm: Application to the main wheat exporting countries. International Journal of Applied Earth Observation and Geoinformation, 104, 102552.

Fritz S., See L., Bayas J.C.L., Waldner F., Jacques D., Becker-Reshef I., McCallum I., 2019, A comparison of global agricultural monitoring systems and current gaps. Agricultural systems, 168, 258-272.

Rossato L., Alvala R.C.D.S., Marengo J.A., Zeri M., Cunha A.P.D.A., Pires L.B., Barbosa H. A., 2017, Impact of soil moisture on crop yields over Brazilian semiarid. Frontiers in Environmental Science, 5, 73.

Sadenova M.A., Beisekenov N.A., Varbanov P.S., 2022, Forecasting Crop Yields Based on Earth Remote Sensing Methods. Chemical Engineering Transactions, 92, 691-696.

Sayago S., Bocco M., 2018, Crop yield estimation using satellite images: Comparison of linear and non-linear models. AgriScientia, 35(1), 1-9.

Shamsuddin A., Rashid N.A., Abd Hamid M.K., Ibrahim N., 2021, Refined Palm Oil Product Quality Predictor for Supporting Palm Oil Refinery Energy Management System, Chemical Engineering Transactions, 89, 115-120, DOI:10.3303/CET2189020.

Tuvdendorj B., Wu B., Zeng H., Batdelger G., Nanzad L., 2019, Determination of appropriate remote sensing indices for spring wheat yield estimation in Mongolia. Remote Sensing, 11(21), 2568.