Article Type: Research
J Name: Modern Phytomorphology
Short name: MP
ISSN: ISSN 2226-3063/ eISSN 2227-9555
Year: 2023
Volume: 17
Page numbers: 118-124
DOI: 10.5281/zenodo.200121
(10.5281/zenodo.Year-Volume-PDFNo.)
Short Title: Using regional normalized difference vegetation index for the large-scale yield prediction of potato, vegetables, fruits, and berries, cultivated in Kherson region of Ukraine

RESEARCH ARTICLE

Using regional normalized difference vegetation index for the large-scale yield prediction of potato, vegetables, fruits, and berries, cultivated in Kherson region of Ukraine

Pavlo Lykhovyd[1*], Raisa Vozhehova[2], Oleksandr Averchev[3], Oleksandr Rudik[4], Liudmyla Hranovska[5], Sergiy Lavrenko[6], Nataliia Avercheva[7], Grigoriy Latiuk[8]

[1]Institute of Climate-Smart Agriculture of NAAS, Department of Irrigated Agriculture and Decarbonization of Agroecosystems, 01010 Kyiv, Ukraine

[2]Institute of Climate-Smart Agriculture of NAAS, Director, 01010 Kyiv, Ukraine

[3]Kherson State Agrarian and Economic University, Head of the Department of Agriculture, 73006 Kherson, Ukraine [4]Odesa State Agrarian University, Head of the Department of Field and Vegetable Crops, 65039 Odesa, Ukraine

[5]Institute of Climate-Smart Agriculture of NAAS, Head of the Department of Irrigated Agriculture and Decarbonization of Agroecosystems, 01010 Kyiv, Ukraine

[6]Kherson State Agrarian and Economic University, Vice-Rector for Science and International Cooperation, 73006 Kherson, Ukraine

[7]Kherson State Agrarian and Economic University, Department of Social and Behavioral Sciences, 73006 Kherson, Ukraine [8]Odesa State Agrarian University, Department of Field and Vegetable Crops, 65039 Odesa, Ukraine

***Corresponding author:** Pavlo Lykhovyd, Institute of Climate-Smart Agriculture of NAAS, Department of Irrigated Agriculture and Decarbonization of Agroecosystems, 01010 Kyiv, Ukraine; Email: pavel.likhovid@gmail.com

Received: 06.10.2023, Manuscript No.: mp-23-115965| **Editor Assigned**: 15.10.2023, Pre-QC No. mp-23-115965(PQ) | **Reviewed**: 21.10.2023, QC No. mp-23-115965(Q) | **Revised:** 28.10.2023, Manuscript No. mp-23-115965(R) | **Accepted**: 05.11.2023 | **Published**: 15.11.2023

Abstract

Large-scale yield prediction of major crops is important to ensure rational science-based policy in agricultural economic activity, especially in planning import and export of plant products and food security assessment. Remote sensing is flexible and convenient tool for the evaluation and prediction of crop yields on large areas. In this study, remote sensing data on the normalized difference vegetation index, calculated for the croplands of Kherson region, were applied to create regression models of potato, vegetables, fruits, and berries yields in the region. The normalized difference vegetation index values were calculated using raw MODIS Terra images for the croplands of the region, retrieved from the service of the University of Natural Resources and Life Sciences (Vienna, Austria) and GIMMS Global Agricultural Monitoring System for the period 2005-2021. The average annual yields of the crops in the region were retrieved from the Ukrainian State Statistical Service for the stipulated period. As a result, linear regression models and models based on artificial neural networks were created to predict yields based on the values of the normalised difference vegetation index. The strongest relationship between the remote sensing data and yield was established for vegetables (in May, R=0.63), while the weakest relationship was established for fruits and berries (in August, R=0.33). The regression models developed have a reasonable to good prediction accuracy for potatoes and vegetables (MAPE=10.04—21.07%), while the prediction of fruits and berries yields has a low precision and reliability. The developed models could be further used in agrarian policy substantiation in Kherson region, as well as in scientific purposes.

Artificial neural network-based models provided better predictive accuracy but are less helpful in understanding the principles of regional crop yield prediction.

Keywords: artificial intelligence, crop, economy, plant production, regression, remote sensing.

Introduction

Remote sensing is a universal tool for various agricultural applications, e.g., environmental monitoring, crop mapping, vegetation cover monitoring, drought events monitoring, disease and insects spreading, floods and desertification scales, etc. (Teke et al., 2013). Today it is one of the irreplaceable constituents of global food security provision, as well as ecological safety (Karthikeyan et al., 2020). It is also widely implemented for irrigation and fertilization management, prediction of yields in plant science, using different spatial vegetation indices as predictors. Yield prediction is an important part of current agricultural science because it is crucial for rational agricultural planning, adjustment of agrotechnology's, plant products supply forecasting and provision of food security. Crop yield prediction should be an additional instrument in the field of agrarian economy, supporting strategic planning of plant production and food import-export policy (Nyéki & Neményi, 2022).

To satisfy the task mentioned above, methods for regional large-scale yield predictions for major crops should be developed. There are different ways for large-scale yield predictions, based on various empirical approaches and simulations, e.g., using photosynthetically active radiation amounts (PAR), leaf area indices (LAI), solar induced chlorophyll fluorescence parameters (SICF), light, water, and nutrients use efficiency, etc. (Karthikeyan et al., 2020). Simulation models are usually realised in the form of specialised software applications; the most popular of them are DSSAT (Jones et al., 2003), WOFOST (Van Diepen et al., 1989), APSIM (Holzworth et al., 2014), and CERES (Timsina & Humphreys, 2006). But simulation models and software are not limited to this list. However, remote sensing-based predictions are now gaining popularity and demand, as they are comparatively simple and reliable enough to provide relevant information on possible scenarios of crop production in a certain agricultural location.

Most scientists use vegetation indices to develop remote sensing-based crop yield predictions. The normalized difference vegetation index (NDVI) and Enhanced Vegetation Index (EVI) have the highest popularity in scientific community (Baret & Guyot, 1991; Liu, 1995). Other remote sensing-based approaches utilise normalised difference water index (NDWI), two-band EVI (EVI2), Green-Red Vegetation Index (GRVI), and vegetation condition index (VCI) as predictors (Gao, 1996; Jiang et al., 2008; Motohka et al., 2010).

Besides, current science has huge mathematical and statistical apparatus to build up predictive models with the highest accuracy, relevance, and fitting quality. The choice of certain mathematical approach in the yield prediction depends mainly on the number of inputs used in the model, their distribution pattern, and the aims of prediction. Sometimes, more sophisticated and mathematically strong methods are not applicable for predictions because of small sample size, or its great inequality, or unnormal distribution of data, etc. Artificial neural networks, for example, notwithstanding their great performance and accuracy, are inappropriate in many scientific and practical purposes because it is impossible to derive the way to solve the prediction task (so called "black box nature"), as in regression modelling, making the latter relevant even considering its relative "out-of-date" status (Karthikeyan et al., 2020).

The main purpose of this study was to establish the relationship between the values of the regional Normalised Difference Vegetation Index (NDVI), calculated for the croplands of the Kherson region, and the average annual yields of potatoes, vegetables, fruits and berries, cultivated there. Mathematical models were also developed to predict the yields of the crops mentioned above on a regional scale.

Materials and Methods

The average annual yields of potatoes, vegetables, fruits, and berries harvested in the Kherson region in the period 2005-2021 were taken from official statistical reports, presented by the Ukrainian State Statistical Body in statistical yearbooks. Monthly values of the regional NDVI for the Kherson region croplands were calculated using the raster analysis toolkit from raw images (MODIS Terra, 250 m resolution, smoothed series), downloaded from the satellite monitoring service at the University of Natural Resources and Life Sciences (Vienna, Austria); partly NDVI images were analysed on a regional scale using the plot analysis toolkit of the GIMMS Global Agricultural Monitoring System.

The relationship between NDVI values and the yields of the studied crops was established through the common Pearson correlation analysis procedure (Yadav, 2018). The interpretation of the correlation relationship was made by the Evans guidelines (1996). The relationship was established for each month of vegetation of active crops, so that it could further be determined which term of NDVI screening fits best for yield prediction.

Yield prediction models were developed using linear regression analysis. The reason for the choice in favour of linear function is that the sample size is medium (N=17), the variation of the inputs (assessed using the coefficient of variation CV) is moderately high, therefore, polynomial functions could not be used without the risks of overfitting (Draper & Smith, 1998; Jenkins & Quintana-Ascencio, 2020). Linear regression analysis was performed using the common procedure within the BioStat v.7 software. The models for the prediction of the yield were developed using the calculated values of the regression coefficient and interception. The precision of the models was evaluated through the calculation of the mean absolute percentage error (MAPE), and their fit quality was determined by the values of the predicted coefficient of determination (Chicco et al., 2021).

Artificial neural network-based models were developed using Tiberius software. The back propagation of errors algorithms was used to train the networks. The neural networks had five hidden neurones, the learning rate was 0.80, training was performed in 1000 epochs. Neural networks are a prospective and promising way of environmental and biological modelling, though they have a great drawback in the impossibility of clear derivation of the algorithms, used by the network to achieve the solution (Vozhehova et al., 2019a).

Results and Discussion

The initial data used in the study are presented in Tab. 1 and 2. The values of the coefficient of variation (CV) testify about the moderate dispersion of the inputs. The lowest dispersion was found for the NDVI values in June, whereas the highest dispersion was found in the values in April. The greatest variation instability was in the yields of fruits and berries, while the lowest was in potato yields. This factor greatly affected the accuracy and reliability of the yield prediction models, as will be seen below.

Table 1. Average monthly NDVI values for the croplands of Kherson region

Year	Month					
	March	April	May	June	July	August
2005	0.29	0.30	0.45	0.50	0.52	0.50
2006	0.34	0.35	0.42	0.54	0.57	0.47
2007	0.33	0.40	0.45	0.40	0.37	0.33
2008	0.32	0.40	0.56	0.58	0.47	0.47
2009	0.30	0.33	0.45	0.57	0.40	0.40
2010	0.31	0.32	0.45	0.56	0.56	0.55
2011	0.25	0.25	0.43	0.54	0.50	0.45
2012	0.33	0.39	0.46	0.50	0.51	0.50
2013	0.43	0.47	0.50	0.51	0.50	0.49
2014	0.45	0.49	0.52	0.52	0.49	0.46
2015	0.44	0.50	0.55	0.57	0.55	0.51
2016	0.45	0.51	0.56	0.57	0.56	0.53
2017	0.42	0.49	0.53	0.54	0.52	0.49
2018	0.46	0.50	0.53	0.53	0.53	0.51
2019	0.46	0.52	0.56	0.56	0.54	0.51
2020	0.38	0.42	0.50	0.54	0.46	0.42
2021	0.35	0.35	0.57	0.61	0.53	0.52
CV	18.36%	20.24%	10.56%	9.00%	10.62%	10.98%

Table 2. Average annual yields of potato, vegetables, fruits, and berries in Kherson region

Year	Potato	Vegetables	Fruits and berries
2005	9.60	13.90	7.67
2006	9.70	14.50	1.86
2007	9.00	12.10	4.66
2008	10.20	16.70	5.45
2009	10.20	22.10	5.40
2010	10.40	19.40	7.45
2011	11.00	24.00	10.69
2012	10.70	28.80	10.11
2013	10.00	27.20	10.35
2014	11.20	28.90	9.66
2015	12.10	30.10	8.52
2016	11.80	31.30	7.61
2017	11.00	30.40	7.87
2018	11.90	31.60	8.78
2019	11.60	32.20	6.91
2020	12.60	31.79	5.62
2021	19.10	30.20	4.96
CV	19.80%	28.40%	32.73%

The relationship between the yield of the crops studied and the vegetation index by each month is presented in Tab. 3. The best correlation was established for the yields of vegetable crops and the regional NDVI in May, while the lowest correlation was determined for fruits and berries. The Pearson correlation coefficient value for fruits and berries was 0.09-0.33, making it almost impossible to predict reliable yields for these crops.

Table 3. Correlation between the yields of the studied crops and monthly NDVI

Crop	March	April	May	June	July	August
Potato	N/A	0.56	0.57	0.57	0.26	N/A
Vegetables	N/A	0.62	0.63	0.42	0.33	0.41
Fruits & berries	N/A	0.20	0.09	-0.09	0.17	0.33

Note: *N/A means "not applicable" for this period of time because of absence of actively vegetating crop in the field.

The regression models for the crops' yields prediction, developed based on the results of linear regression analysis, are presented in Tab. 4. The regression statistics for the evaluation of their accuracy and quality of fitting are provided in Tab. 5.

Table 4. Linear regression models for potato, vegetables, fruits, and berries yields prediction

Crop	Model
Potato	Yield=-3.5970+27.7080×NDVI
Vegetables	Yield=-18.2710+86.6660×NDVI
Fruits & berries	Yield=0.3748+14.4510×NDVI

Table 5. Regression statistics for the linear models of potato, vegetables, fruits, and berries yields prediction

Statistical criteria	Potato	Vegetables	Fruits & berries
R	0.5718	0.6346	0.3278
MSE	3.5931	32.1500	5.3885
S	1.8955	5.6701	2.3213
RSQ	0.3270	0.4027	0.1074
RSQ adjusted	0.2821	0.3629	0.0479

RSQ predicted	0.1375	0.2226	0.0452
MAPE	10.04%	21.07%	34.53%

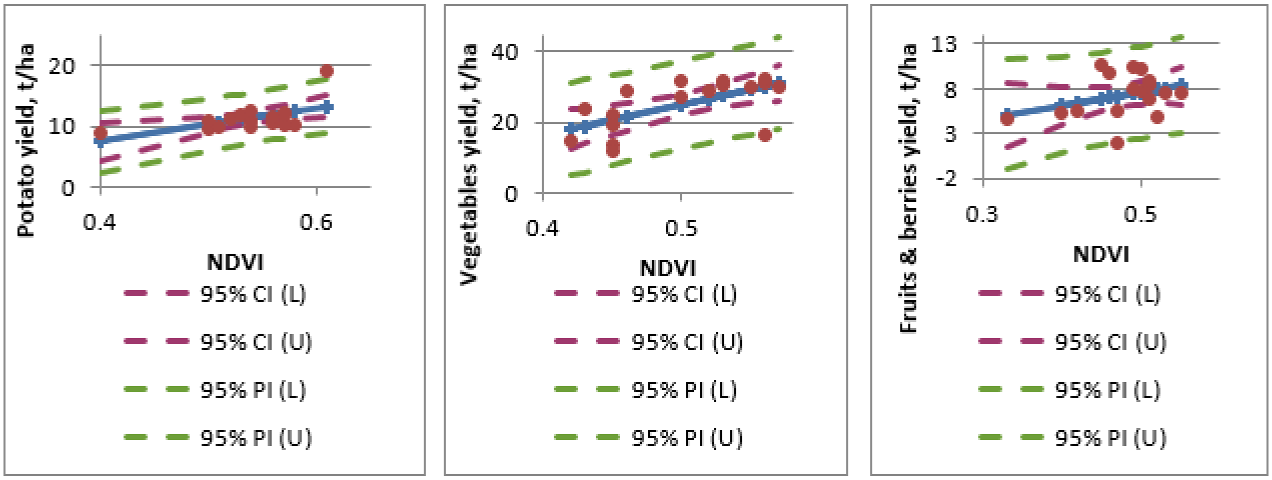

Figure 1. Fitting quality and the graph of dispersion of the regression models for potato, vegetables, fruits, and berries yields prediction

From the results, described in Table 5, it is evident, that the best fitting quality (Fig. 1) is attributed to the model of vegetables yield prediction, while the highest accuracy is associated with the model of potato yield prediction. According to current classification (Moreno et al., 2013), the model of potato yield prediction belongs to good forecasting models, while the models for vegetables, fruits, and berries yields are of reasonable quality. The model for fruits and berries yields prediction has low fitting quality along with high MAPE; therefore, it could not be recommended for practical use. The regression statistics for the neural network-based models are given in Tab. 6.

Table 6. Regression statistics for the neural network-based models of potato, vegetables, fruits, and berries yields prediction

Statistical criteria	Potato	Vegetables	Fruits & berries
R	0.9323	0.7705	0.4349
RSQ	0.8691	0.5937	0.1891
MAPE	1,39%	1.34%	4.52%

As expected, neural network-based models perform much better in terms of precision and fitting quality. However, they are only of theoretical usefulness, as it is impossible to derive the function of yield dependence on NDVI. The poorest performance was associated with fruits and berries, while the best fitting quality was for potato, and the least MAPE value was attributed to vegetables.

Discussion

Large-scale yield prediction based on remote sensing is a prospective and relevant approach to ensure timely forecasts, which is essential for rational agrarian policy and provision of food security. Most studies in this direction were targeted on the prediction of yields of such staple food crops as winter wheat (Ren et al., 2008), barley (Weissteiner & Kühbauch, 2005), rice (Huang et al., 2013), corn (Mkhabela et al., 2005), some oil crops (Lykhovyd, 2021), and soybeans (Andrade et al., 2022). Models, used for regional crop yield prediction, differ in their accuracy and fitting quality depending on the initial data sets and mathematical approaches applied to create the forecasts. However, most of the NDVI-based models provide sufficient accuracy to be used in science and practise.

Regarding NDVI-based potato yield prediction, the study by (Vannoppen & Gobin 2022) provided a moderately good predictive model, which could be applied in the conditions of Northern Belgium. The model has higher fitting quality, but the mean square error values are comparatively high. As for vegetable crops, fruits, and berries, there are just a few studies, where the yields of particular crops (not cumulative yields as in our case) were successfully estimated using remote-sensing NDVI data (Maselli et al., 2012; Bai et al., 2019; Suarez et al., 2020). Therefore, our models are the first to predict cumulative yields of vegetables and fruits. However, we must admit that the quality of the predictions for fruits and berries is much lower than expected; therefore, it is not recommended to apply the model for practical purposes. This could be put upon the fact that the crops, included in both groups, are quite different by their morpho-biological qualities, especially, if we are talking about fruits and berries (represented by trees, bushes, shrubs, and herbaceous plants).

Besides, it was testified one more time that artificial intelligence, represented by neural networks, is superior to regression modeling. Models based on neural networks are 10-20 times more accurate in yield prediction, while the fit quality is almost twice better. However, some studies claim that the quality of modelling does not differ significantly in the case of neural networking approach and regression (Grzesiak et al., 2003). However, it may depend not only on the architecture of neural networks, but also on the activation functions applied, as well as on the software used to build the model (Lykhovyd, 2018; Vozhehova et al., 2019b). Generally, artificial neural networks are preferable method of nonlinear modelling in life sciences, the main drawback is their "black-box nature", which is difficult to open and analyse for non-specialists in data science (Castelvecchi, 2016).

To sum up the discussion it is necessary to state that

i) Remote sensing NDVI data could be used to predict yields of major crops on the regional scale when used in appropriate time and directly for the crop mask or croplands mask;
ii) Linear regression models provide satisfactory forecasts, however, if we need rougher and more accurate prognosis, it is better to pass the ball to artificial neural networks.

Conclusions

The results of the current study provide new information on the prediction of regional yields in the South of Ukraine. Regression models, built up for potato and vegetables, have moderate fitting quality and good predictive accuracy, so that they could be applied both for scientific and practical purposes. NDVI-based forecasting of the yields of major crops on the regional scale is of high importance nowadays, and this approach should be further studied and developed to ensure rational agrarian policy and food security.

References

Andrade T.G., Andrade Jr A.S.D., Souza M.O., Lopes J.W.B., Vieira P.F.D.M.J. 2022. Soybean yield prediction using remote sensing in southwestern Piauí state, Brazil. *Rev. Caatinga*, **35**, 105-116.

Bai T., Zhang N., Mercatoris B., Chen Y. 2019. Jujube yield prediction method combining Landsat 8 Vegetation Index and the phenological length. *Computers and Electronics in Agriculture*, **162**, 1011-1027.

Baret F., Guyot G. 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, **35**, 161-173.

Castelvecchi D. 2016. Can we open the black box of AI?. *Nature News*, **538**, 20.

Chicco D., Warrens M.J., Jurman G. 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, **7**, e623.

Draper N.R., Smith H. 1998. Applied Regression Analysis **326**. *John Wiley & Sons*.

Evans J.D. 1996. Straightforward Statistics for the Behavioral Sciences. *Thomson Brooks/Cole Publishing Co*.

Gao B.C. 1996. NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, **58**, 257-266.

Grzesiak W., Lacroix R., Wójcik J., Blaszczyk P. 2003. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. *Can. J. Anim. Sci.*, **83**, 307-310.

Holzworth D.P., Huth N.I., deVoil P.G., Zurcher E.J., Herrmann N.I., McLean G., Chenu K., van Oosterom E.J., Snow V., Murphy C., Moore A.D., Brown H., Whish J.P.M., Verrall S., Fainges J., Bell L.W., Peake A.S., Poulton P.L., Hochman Z., Thorburn P.J., Gaydon D.S., Dalgliesh N.P., Rodriguez D., Cox H., Chapman S., Doherty A., Teixeira E., Sharp J., Cichota R., Vogeler I., Li F.Y., Wang E., Hammer G.L., Robertson M.J., Dimes J.P., Whitbread A.M., Hunt J., van Rees H., VcClelland T., Carberry P.S., Hargreaves J.N.G., MacLeod N., McDonald C., Harsdorf J., Wedgwood S., Keating B.A. 2014. APSIM–evolution towards a new generation of agricultural systems simulation. *Environ. Modell. Softw.*, **62**, 327-350.

Huang J., Wang X., Li X., Tian H., Pan Z. 2013. Remotely sensed rice yield prediction using multi-temporal NDVI data derived from NOAA's-AVHRR. *PloS One*, **8**, e70816.

Jenkins D.G., Quintana-Ascencio P.F. 2020. A solution to minimum sample size for regressions. *PloS One*, **15**(2), e0229345.

Jiang Z., Huete A.R., Didan K., Miura T. 2008. Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, **112**, 3833-3845.

Jones J.W., Hoogenboom G., Porter C.H., Boote K.J., Batchelor W.D., Hunt L.A., Wilkens P.W., Singh U., Gijsman A.J., Ritchie J.T. 2003. DSSAT Cropping System Model. *Eur. J. Agron.*, **18**, 235-265.

Karthikeyan L., Chawla I., Mishra A.K. 2020. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *J. Hydrol.*, **586**, 124905.

Kogan F.N. 1995. Application of vegetation index and brightness temperature for drought detection. *Adv. Space Res.*, **15**, 91-100.

Liu, H.Q. 1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.*, **33**, 457-465.

Lykhovyd P. 2021. Forecasting oil crops yields on the regional scale using normalized difference vegetation index. *J. Ecol. Eng.*, **22**, 53-57.
Lykhovyd P.V. 2018. Prediction of sweet corn yield depending on cultivation technology parameters by using linear regression and artificial neural network methods. *Biosyst. Divers.*, **26**, 11-15.
Maselli F., Chiesi M., Brilli L., Moriondo M. 2012. Simulation of olive fruit yield in Tuscany through the integration of remote sensing and ground data. *Ecol. Model.*, **244**, 1-12.
Mkhabela M.S., Mkhabela M.S., Mashinini, N.N. 2005. Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. *Agric. For. Meteorol.*, **129**, 1-9.
Moreno J.J.M., Pol A.P., Abad A.S., Blasco B.C. 2013. Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema*, **25**, 500-506.
Motohka T., Nasahara K.N., Oguma H., Tsuchida S. 2010. Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sens.*, **2**, 2369-2387.
Nyéki A., Neményi M. 2022. Crop yield prediction in precision agriculture. *Agronomy*, **12**, 2460.
Ren J., Chen Z., Zhou Q., Tang H. 2008. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *Int. J. Appl. Earth Obs. Geoinf.*, **10**, 403-413.
Suarez L.A., Robson A., McPhee J., O'Halloran J., van Sprang C. 2020. Accuracy of carrot yield forecasting using proximal hyperspectral and satellite multispectral data. *Precision Agric.*, **21**, 1304-1326.
Teke M., Deveci H.S., Haliloğlu O., Gürbüz S.Z., Sakarya U. 2013. A short survey of hyperspectral remote sensing applications in agriculture. *2013 6th Int. Conf. Recent Adv. Space Technol.*, 171-176.
Timsina J., Humphreys E.J.A.S. 2006. Performance of CERES-Rice and CERES-Wheat models in rice–wheat systems: A review. *Agron. Syst.*, **90**, 5-31.
Van Diepen C.V., Wolf J.V., Van Keulen H., Rappoldt C. 1989. WOFOST: a simulation model of crop production. *Soil Use Manag.*, **5**, 16-24.
Vannoppen A., Gobin A. 2022. Estimating yield from NDVI, weather data, and soil water depletion for sugar beet and potato in Northern Belgium. *Water*, **14**, 1188.
Vozhehova R.A., Lykhovyd P.V., Kokovikhin S.V., Biliaieva I.M., Markovska O.Y., Lavrenko S.O., Rudik O.L. 2019b. Artificial Neural Networks and Their Implementation in Agricultural Science and Practice. *Warsaw, Diamond Trading Tour*, 108 pp.
Vozhehova R.A., Lykhovyd P.V., Lavrenko S.O., Kokovikhin S.V., Lavrenko N.M., Marchenko T.Y., Sydyakina O.V., Hlushko T.V., Nesterchuk V.V. 2019a. Artificial neural network use for sweet corn water consumption prediction depending on cultivation technology peculiarities. *Res. J. Pharm. Biol. Chem. Sci.*, **10**, 354-358.
Weissteiner C.J., Kühbauch W. 2005. Regional yield forecasts of malting barley (*hordeum vulgare L.*) by NOAA-AVHRR remote sensing data and ancillary data. *J. Agron. Crop Sci.*, **191**, 308-320.
Yadav S. 2018. Correlation analysis in biological studies. *J. Pract. Cardiovasc. Sci.*, **4**, 116-121.