EEET ECOLOGICAL ENGINEERING
& ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology 2023, 24(4), 223–228
https://doi.org/10.12912/27197050/162698
ISSN 2719-7050, License CC-BY 4.0

Received: 2023.02.20
Accepted: 2023.03.27
Published: 2023.04.10

The Use of Spatial Normalized Difference Vegetation Index for Determination of Humus Content in the Soils of Southern Ukraine

Pavlo Lykhovyd[1]

[1] Institute of Climate-Smart Agriculture of NAAS, Mikhail Omelyanovich-Pavlenko, 9, 01010, Kyiv, Ukraine
E-mail: pavel.likhovid@gmail.com

ABSTRACT

Spatial normalized difference vegetation index finds various applications in crop monitoring and prediction. Although this index is mainly aimed to represent the state of vegetation cover, it is suggested that it could be utilized for other remote monitoring purposes, for example, soil humus content monitoring. The study was carried out in 2022–2023 fallow-field period in Kherson oblast, the South of Ukraine, to establish the relationship between the values of bare-soil normalized difference vegetation index and content of humus in the soils of the region. Statistical modeling was performed using the best subsets regression analysis in BioStat v.7 and artificial neural network with back propagation of error algorithm in Tiberius XL. The best performance was recorded for the combined model of cubic regression and artificial neural network, with moderate fitting quality (coefficient of determination is 0.29), and good prediction accuracy (mean average percentage error is 13.22%). The results approve the suggestion of possibility of spatial vegetation index use in soil state monitoring, especially, if further scientific work enhances the fitting quality of the model.

Keywords: modeling, normalized difference vegetation index, best subsets regression analysis, satellite imagery, soil cover.

INTRODUCTION

Remote sensing is an integrative union of modern science and technology, applied for getting information about dislocation, dimensions, and various properties of the Earth-based objects of natural and artificial origin without direct measurements and contact with the studied objects, mainly by the means of satellite-based sensors (Death, 2008). Spectral images, obtained as a result of "photographing" the Earth surface in different bands from space, are further used by scientists and practitioners to assess features of land-based objects. As far as remote sensing technique is rapidly developing worldwide, its application is also spreading to more and more branches of science and economy. For example, it is hardly believable that modern populational ecology, environmental science, precision agriculture, economics, military science, climatology, geology, logistics, etc., would have been developed so rapidly and provide mankind with novel insights on the life on the Earth and innovative high-quality services with no remote sensing technique applied in the studies (Lykhovyd et al., 2020a).

Speaking about agriculture, remote sensing brought such amazing opportunities for scientists and practitioners as remote control for crop growth and development, machinery use, environmental conditions, crop yield prediction, phenology, infestation of crops with pests, insects, and diseases, etc. (Bastiaanssen et al., 2000). Environmental monitoring for vegetation cover is widely assisted by remote sensing technologies within geoinformation systems (Lykhovyd, 2021b). Wide horizons are open for crop modeling and simulation of their productive processes based on various information obtained from satellite spectral imagery (Bouman, 1995; Maselli et al., 2000; Tripathy et al., 2013; Lykhovyd et al., 2020b; Lykhovyd, 2021a). Nowadays, decision support systems for agricultural producers are implementing remote sensing data and scientifically grounded ways of crop control by the means of satellite-based imagery to enhance agricultural land productivity and make technological decisions of practitioners more weighed and reasonable (McBratney et al., 2005).

As it has been mentioned before, science applies satellite imagery for its purposes, mainly not in its raw state, but through processing the images from different spectral bands and calculating various indices to represent the properties of the Earth-based objects. For example, agronomy applies about 150 different vegetation indices to describe plant, soil cover, and water availability conditions (Lykhovyd, 2022). However, only few of these indices found wide practical implementation. Nowadays, agricultural practice mostly applies normalized difference vegetation index (NDVI), proposed in 1974, to support crop producers (Rouse et al., 1974). Most digital farming platforms are based on NDVI images. Therefore, it is reasonable to investigate this vegetation index thoroughly to find its suitability for something more than just crop monitoring. For example, it could be a prospective index for soil properties monitoring on the areas, which are free from living plants, especially, in the areas subjected to various negative environmental impacts (Lykhovyd et al., 2019). So, quantitative remote sensing could be used for soil properties estimation (Ben-Dor, 2002), besides there were attempts to adopt it for degraded land identification by NDVI values (Abdulhussein & Mihalache, 2021).

We find it interesting from theoretical and practical points of view to assess NDVI for suitability to estimate soil fertility, especially, such index as humus content. The hypothesis of our study is that NDVI could be used for assessment of humus content in soils if it is screened in the fallow-field (bare field after harvesting or in pre-sowing period). If it is so, reasonable use of NDVI could be helpful to reduce financial and time expenditures for soil surveys. The hypothesis is based on the fact that humus is an organic matter of the soil, and it has its unique spectral reflectance, while soil organic matter content was proved to be precisely recognized through the near-infrared (NIR) spectral band (Heil & Schmidhalter, 2021). As NDVI computation engages NIR spectral band, it is highly likely that this vegetation index might be applied for remote soil humus content determination in the surface soil layer (up to the depth of 0–30 cm).

MATERIALS AND METHODS

The study was carried out in the fallow-field (which lasts from November to February) period of 2022–2023 using bare-soil cloud-free NDVI imagery for the fields in different agricultural districts of Kherson oblast, Southern Ukraine. The screens of NDVI for the studied fields were obtained at OneSoil AI platform, as well as averaged index values for each field. In total, 1478 individual bare-soil NDVI values for the fields, located in different districts of Kherson oblast, were analyzed and generalized. The soil humus content by the fields or agricultural districts were taken from the results of the regional soil surveys (Bychkov et al., 1987) and soil surveys, conducted in the framework of field experiments carried out at the Institute of Irrigated Agriculture of NAAS (now reorganized into the Institute of Climate-Smart Agriculture of NAAS). After the NDVI data generalization and association with corresponding humus content using geotagging, 34 data pairs "bare-soil NDVI (pts.) – humus content (%)" were created and processed in statistical analysis by the means of the best subsets regression in BioStat v.7. In total, nine regression functions were tested, as it is described in the Table 1.

The location of the studied fields (for bare-soil NDVI derivation) in Kherson oblast is depicted in the Figure 1. The agglomerations of the fields are marked with the symbol "✓" on the background of OneSoil AI NDVI platform screen

Table 1. Regression functions evaluated within the best subsets regression test in BioStat v.7 to predict the content of humus in the soils by the values of bare-soil NDVI in Kherson oblast

Function name	Equation
Linear	$Y=a*x+b$
Quadratic	$Y=a*x^2+b*x+c$
Cubic	$Y=a*x^3+b*x^2+c*x+d$
Stepwise	$Y=a*x^b$
Exponential-1 (Composit)	$Y=a*b^x$
Hyperbolic	$Y=a+b/x$
Logarithmic	$Y=a+b*\ln(x)$
Exponential-2	$Y=e^{a+bx}$
Sigmoid	$Y=e^{a+b/x}$

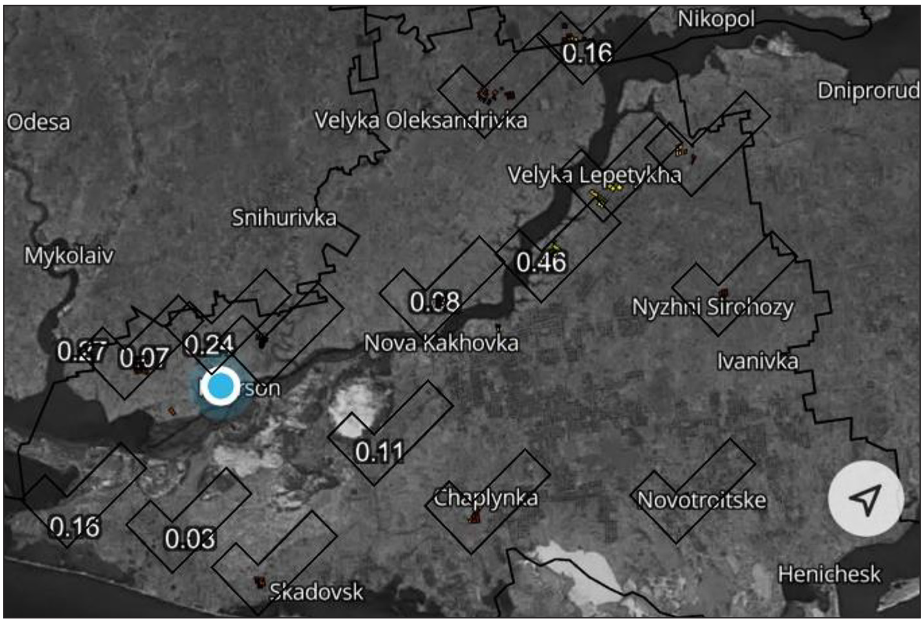

Figure 1. The screen of OneSoil AI platform for barre-soil NDVI values derivation with the marks of the fields and checkpoints used in the study

Further, artificial neural networks (ANN) with back propagation learning algorithm of different architecture and learning rate were applied to derive soil humus content by NDVI. The ANN were created and assessed using Tiberius XL software (Brierley, 1998). The types of ANN (training was conducted during 1000 epochs), used in the study, are described in the Table 2. We tried to test the ANN with the highest and the lowest possible number of neurons at both maximum (1.0) and average (0.7) learning rate.

In the end, the combined ANN-regression model was derived to estimate the content of humus in the soils of Kherson oblast using the values of spatial NDVI. The results of the ANN-derived model were additionally processed using the best subsets regression analysis in BioStat v.7 to create the combined cubic regression model with the best fitting quality and prediction accuracy (Lavrenko et al., 2022).

The quality of the model fitting, its adequacy to the input dataset, was assessed using Pearson's correlation coefficient (R) and the coefficient of determination (R^2). The evaluation scale we referred to in this study was as proposed by Evans (1996). Accuracy testing was performed through the computation of mean absolute percentage error (MAPE) and its score interpretation using the grades by Blasco et al. (2013). The computations were performed using Microsoft Excel 365 and BioStat v.7 software.

Table 2. Artificial neural networks used to predict the content of humus in the soils by the values of bare-soil NDVI in Kherson oblast

Variant	Number of neurons; learning rate
1	5; 1.0
2	5; 0.7
3	1; 1.0
4	1; 0.7

RESULTS AND DISCUSSION

Nine regression models built within the best subsets analysis were evaluated by their fitting quality first. The best fitting quality (moderate strength of relationship) is observed for cubic model (R = 0.4973, R^2 = 0.2473), although it is impossible to deny overfitting. The best fitting quality among other models with no overfitting is attributed to linear model (R = 0.3776, R^2 = 0.1426). The graphs of the proposed regression models fitting are presented in the Figure 2. The values of fitting quality assessment coefficients are presented in the Table 3.

Apart from fitting adequacy, regression models were assessed in terms of the prediction accuracy by the values of MAPE. The best accuracy by the mentioned index is attributed to cubic model (MAPE = 13.77%), with slightly less accuracy of quadratic (15.26%) and linear (15.41%) ones, respectively. All the mentioned models are of good

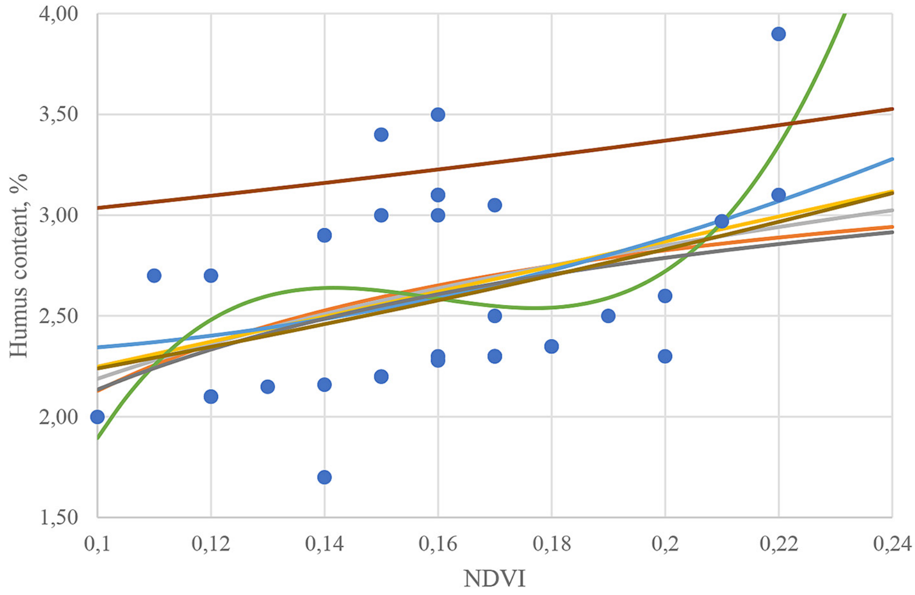

Figure 2. Approximation graphs for nine regression models used in the study of humus content in the soil derivation by spatial NDVI

Table 3. Evaluation of regression models fitting quality

Function name	R	R^2
Linear	0.3776	0.1426
Quadratic	0.3843	0.1477
Cubic	0.4973	0.2473
Stepwise	0.3751	0.1407
Exponential-1 (composit)	0.3791	0.1438
Hyperbolic	0.3621	0.1311
Logarithmic	0.3707	0.1374
Exponential-2	0.3791	0.1438
Sigmoid	0.3690	0.1362

prediction accuracy according to the Blasco et al. (2013) gradation.

Somewhat better fitting quality and prediction accuracy was obtained using ANN with 5 neurons in the hidden layer and learning rate of 0.8. Less number of neurons in the hidden layer and increased learning rate resulted in substantial decrease in the fitting quality and precision. The best ANN model overscored the cubic regression both in terms of approximation (R^2 0.29 vs. 0.25) and accuracy (MAPE 12.28% vs. 13.77%). The statistics for ANN-based modeling is in the Table 4.

Although ANN-based models are better than regression ones, it is impossible to apply them out of the software environment they were built in. In this regard, a combined model, which is the improved cubic regression model adjusted by the ANN simulation, is used. The fitting performance of such a model in our study is superior to regular cubic model (R^2 = 0.29 vs 0.25), as well as the accuracy (MAPE = 13.22% vs. 13.77%). The model looks like the Eq. 1, where Y represents the content of humus in the soil (%), and x is the average value of bare-soil NDVI (pts.):

$$Y = 43.164 - 1168.8x + 12122x^2 - 54272x^3 + 89044x^4 \quad (1)$$

The proposed model has good prediction accuracy, moderate fitting quality, and it could be used for computations in any working environment in contrast to the ANN-based model.

The study by Kumar et al. (2016) was conducted on quite similar subject. The authors discovered strong regression relationship (R^2 = 0.7254)

Table 4. Basic evaluation of ANN-based modeling of soil humus content using bare-soil NDVI values

ANN type	R	R^2	MAPE
5 neurons; learning rate 1.0	0.45	0.20	13.27%
1 neuron; learning rate 1.0	0.48	0.23	12.45%
1 neuron; learning rate 0.8	0.48	0.23	14.24%
5 neurons; learning rate 0.8	0.54	0.29	12.28%

between bare-soil NDVI and soil organic carbon content (SOC). As far as SOC is strongly related to humus content in the soil, this study is another supporting one for our theory. Herbei et al. (2022) also established high strength of the relationship between NDVI and humus content in soils with R^2=0.538. Larkin et al. (2020) established strong linear relationship between NDVI and humus content with R=0.84.

Therefore, notwithstanding the fact tat there is a lack of scientific evidence for prediction of soil humus content based on the values of spatial NDVI now, the topic is promising, and relevant, and further scientific investigations are required to draw final conclusions about the strength, direction, and reliability of the relationship between these parameters.

CONCLUSIONS

The results of the preliminary study discovered an opportunity to use bare-soil NDVI values to derive the amount of humus content in the soils of Kherson oblast. The combined approach to prediction of humus content using an integrated ANN and cubic regression model showed good prediction accuracy (MAPE = 13.22%) and moderate fitting quality (R^1 = 0.29). Considering that the results are encouraging, further investigation of the approach to soil humus content derivation from bare-soil NDVI is to be conducted.

REFERENCES

1. Abdulhussein, A., Mihalache, M. 2021. Use of remote sensing techniques and geographic information systems to identify degraded land in Dhi Qar region from Irak. Scientific Papers. Series A. Agronomy, LXIV(2), 13–21.
2. Bastiaanssen, W., Molden, D., Makin, I. 2000. Remote sensing for irrigated agriculture: examples from research and possible applications. Agricultural Water Management, 46(2), 137–155.
3. Ben-Dor, E. 2002. Quantitative remote sensing of soil properties. Advances in Agronomy, 75, 173–243.
4. Blasco, B.C., Moreno, J.J.M., Pol, A.P., Abad, A.S. 2013. Using the R-MAPE index as a resistant measure of forecast accuracy. Psicothema, 25(4), 500–506.
5. Bouman, B. 1995. Crop modelling and remote sensing for yield prediction. NJAS wageningen journal of life sciences, 43(2), 143–161.
6. Brierley, P.D. 1998. Some practical applications of neural networks in the electricity industry. Cranfield University (United Kingdom).
7. Bychkov, V.V., Ostapov, V.I., Zhuravlev, A.I., Kotliar, N.M., Lirnyk, V.A., Lomonosov, P.I., Pisarenko, V.A., Funtov, A.P. 1987. Scientifically based system of agriculture of Kherson oblast. Kherson, 448.
8. Death, R. 2008. Encyclopedia of Ecology. Elsevier, Amsterdam, The Netherlands.
9. Evans, J.D. 1996. Straightforward statistics for the behavioral sciences. Thomson Brooks/Cole Publishing Co.
10. Heil, K., Schmidhalter, U. 2021. An evaluation of different NIR-spectral pre-treatments to derive the soil parameters C and N of a humus-clay-rich soil. Sensors, 21(4), 1423.
11. Herbei, M.V., Bertici, R., Sala, F. 2022. The use of remote sensing images in order to characterize the soil agrochemical indexes in relation to the agricultural crops. Agriculture & Forestry/Poljoprivreda i Sumarstvo, 68(2), 23–33.
12. Kumar, P., Pandey, P.C., Singh, B.K., Katiyar, S., Mandal, V.P., Rani, M., Tomar, V., Patairiya, S. 2016. Estimation of accumulated soil organic carbon stock in tropical forest using geospatial strategy. The Egyptian Journal of Remote Sensing and Space Science, 19(1), 109–123.
13. Larkin, M.A., Gubarev, D.I., Medvedev, I.F. 2020. Using the results of soil-agrochemical survey and remote sensing data in the formation of working areas in the field. In: Scientific Dialogue in Language Environment, 66–70.
14. Lavrenko, S., Lykhovyd, P., Lavrenko, N., Ushkarenko, V., Maksymov, M. 2022. Beans (Phaseolus vulgaris L.) yields forecast using normalized difference vegetation index. International Journal of Agricultural Technology, 18(3), 1033–1044.
15. Lykhovyd, P. 2021a. Forecasting oil crops yields on the regional scale using normalized difference vegetation index. Journal of Ecological Engineering, 22(3), 53–57.
16. Lykhovyd, P. 2022. Theoretical Bases of Crop Production on the Reclaimed Lands in the Conditions of Climate Change. RS Global Sp. z O.O, Poland, Warsaw, 259.
17. Lykhovyd, P.V. 2021b. Study of climate impact on vegetation cover in Kherson oblast (Ukraine) using normalized difference and enhanced vegetation indices. Journal of Ecological Engineering, 22(6), 126–135.
18. Lykhovyd, P., Biliaieva, I., Boitseniuk, K. 2020a. Remote sensing applications in precision agriculture. In: Problems and Innovations in Science. Abstracts of the 1st International scientific and practical conference, 9–12.
19. Lykhovyd, P., Lavrenko, S., Lavrenko, N. 2020b. Forecasting grain yields of winter crops in Kherson

oblast using satellite-based vegetation indices. Bioscience Research, 17(3), 1912–1920.

20. Lykhovyd, P.V., Lavrenko, S., Lavrenko, N., Dementiieva, O. 2019. Agro-environmental evaluation of irrigation water from different sources, together with drainage and escape water of rice irrigation systems, according to its impact on maize. Journal of Ecological Engineering, 20(2), 1–7.
21. Maselli, F., Romanelli, S., Bottai, L., Maracchi, G. 2000. Processing of GAC NDVI data for yield forecasting in the Sahelian region. International Journal of Remote Sensing, 21(18), 3509–3523.
22. McBratney, A., Whelan, B., Ancev, T., Bouma, J. 2005. Future directions of precision agriculture. Precision Agriculture, 6(1), 7–23.
23. Rouse, Jr J., Haas, R., Deering, D., Schell, J., Harlan, J. 1974. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation (No. E75-10354).
24. Tripathy, R., Chaudhari, K., Mukherjee, J., Ray, S., Patel, N., Panigrahy, S., Parihar, J. 2013. Forecasting wheat yield in Punjab state of India by combining crop simulation model WOFOST and remotely sensed inputs. Remote Sensing Letters, 4(1), 19–28.