TOWARD METHODOLOGY OF PREDICTIVE MODELING OF THE GROWING SEASON DYNAMICS

Alyona Shagnieva
Evgeny Panidi
Saint Petersburg State University, **Russia**

ABSTRACT

We applied the Normalized Difference Vegetation Index (NDVI) as an indicator reflecting dynamics of vegetation cover growing. NDVI time series were derived from the Terra/MODIS satellite imagery (MOD09A1 product was used, available from the open archive of the US Geological survey). Data for 2006-2016 were used. The ground test plate is located in the Karelia Republic, Russia (61°,7218 N, 34°,3689 E). Moving average method with three inflection points was applied to ensure a piecewise monotone smoothing of NDVI time series. We also applied meteorological data (surface air temperature, average daily precipitation, average snow cover height – collected at Petrozavodsk meteorological station in 2006-2016). Used meteorological data were derived from the website of the Russian Federal Service for Hydrometeorology and Environmental Monitoring – ROSHYDROMET (http://meteo.ru).

As a result, we designed a methodology for selection and processing (including initial compilation, smoothing and interpolation of time series) of remote sensing and meteorological observations to generate training data and control datasets demanded by artificial neural network. Obtained results demonstrate that neural network trained using meteorological data can predict NDVI values. Less than 15% relative error of the NDVI prediction was gained at our ground test plate.

Keywords: Remote Sensing, NDVI, Artificial Neural Networks, Predictive Modeling

INTRODUCTION

Application of artificial neural networks to reconstruction of the vegetation indices time series have became popular study direction in recent years. Usually, authors [2,7,9] discover the technique as an approach to alignment, smoothing and direct analysis of vegetation indices data obtained from remote sensing observations.

Our study [6,12] is directed onto investigation of artificial intelligence technologies application to estimate and predict dynamics of vegetation cover. Such predictions are demanded in agricultural planning [11]. General goals of current stage of the described research activities are to design and verify artificial neural network able to predict vegetation cover dynamics. We use remote sensing and meteorological data to train the neural network, as a number of studies [4,5], including our previous studies, shown a potential of complex analysis of meteorological and remote sensing data.

DATA AND TOOLS

At the current study stage, we derived NDVI time series from the Terra/MODIS satellite imagery (radiance data, MOD09A1 product [13]). Data for 2006-2016 were used, as it was the period covered with continuous below described meteorological data time series.

In our experiments, we operated with the NDVI data produced using standard well-known NDVI formula:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

where, *NIR* – near infrared reflectance (841-876 nm wavelength), *RED* – red reflectance (620-670 nm wavelength) [10].

NDVI was calculated using ArcGIS for Desktop software. Python scripting was applied, scripts were run in built-in ArcPy interpreter. Time series NDVI (Fig. 1) were obtained in the ground test plot for 11 years from 2006 to 2016. The ground test plate is located in the Karelia Republic, Russia (61°,7218 N, 34°,3689 E).\
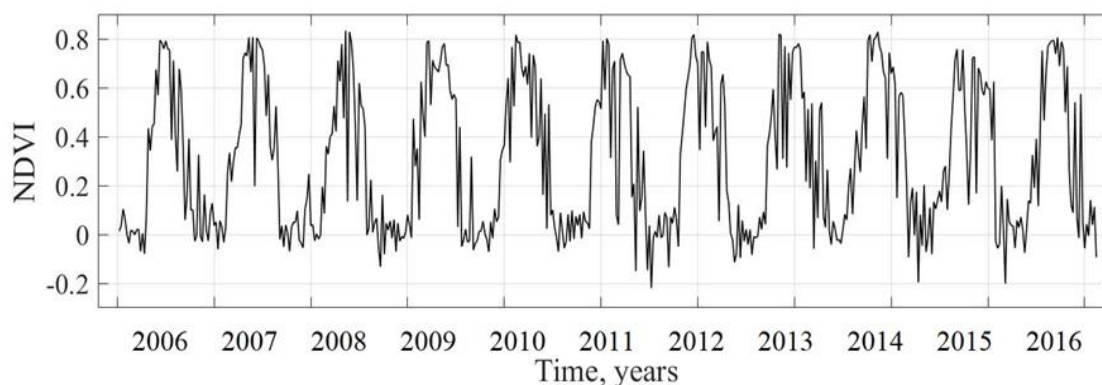
Figure 1. NDVI values at a measurement position in the ground test plot for 11 years

Geographic conditions and differences in illumination geometry and surface observation conditions cause noise in NDVI time series. As a result, preliminary processing of satellite data is required [1]. We propose to derive the NDVI time series from original reflectance data using a moving average method with three inflection points. This technique ensures piecewise monotonic smoothing of the NDVI time series. The moving average method involves replacement of each value in the time series with the averaged neighbouring values of the series separated by moving window of selected size. The selection of the window size was carried out empirically (Fig. 2). Minimal NDVI distortion was obtained with averaging by 7 neighbouring values (window size is 7). Smoothing was performed using the MATLAB software. The smoothed NDVI time series were interpolated to 1-day time resolution series using cubic spline function in MATLAB, while some other authors conclude that polynomial approximation is an appropriate technique when reconstructing remote sensing data [8].

In analysis, we applied also meteorological data collected at Petrozavodsk (ROSHYDROMET) meteorological station in 2006-2016. Daily average surface air temperature, precipitation, and snow cover height values were retrieved through the ROSHYDROMET website (http://meteo.ru) and composed into the 1-day time resolution time series.

Artificial neural networks (ANN) are posed as an effective tool for the consolidation and prediction of information. We used multilayer feedforward neural networks (multilayer perceptron) to build ANN-based predictive model. The AAN is implemented in the MATLAB [3]. Networks of this type have one-way connections from the input to the output layers of neurons, and are often used to predict and approximate nonlinear functions.
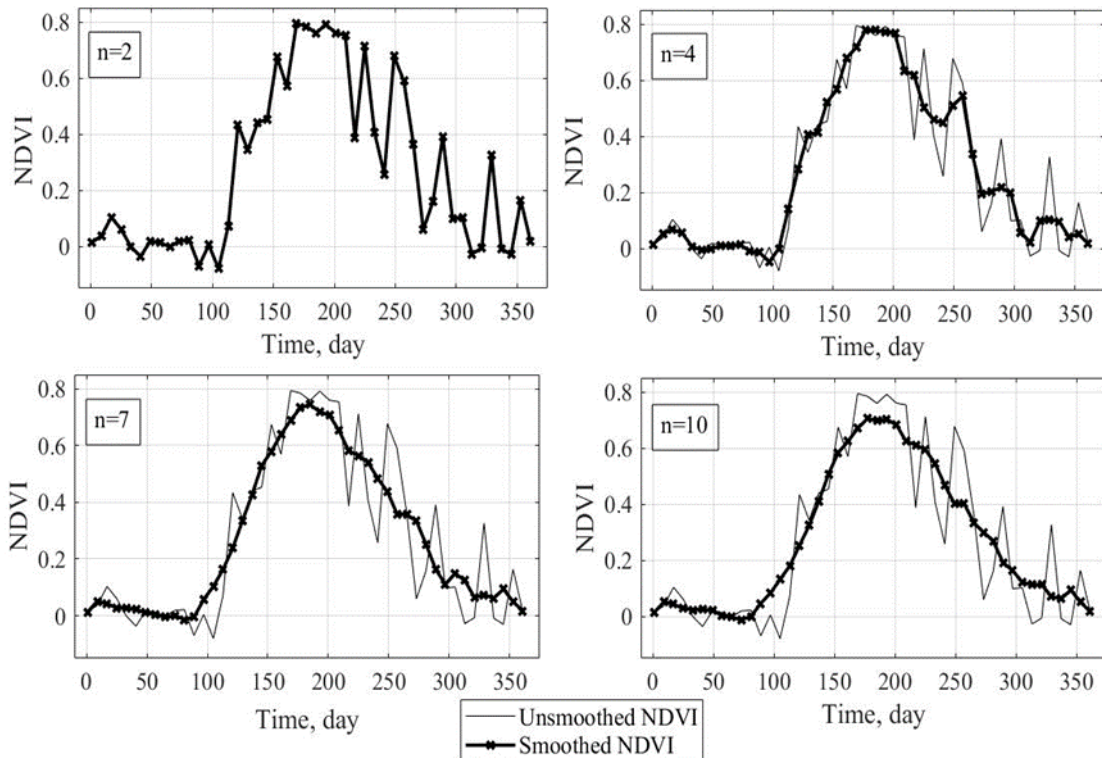
Figure 2. NDVI time series smoothing, selection of the moving window size

DATA PROCESSING AND RESULTS

A methodology is proposed for the selection and processing of remote sensing and meteorological observations data. It incorporates formation, smoothing and interpolation of time series datasets to generate data for training artificial neural networks (Fig. 3). It corresponds to the traditional processing chain of ANN-based analysis, and assumes selections of indicators (predictors) from the meteorological (climate) in-situ observations accordingly to our study context.
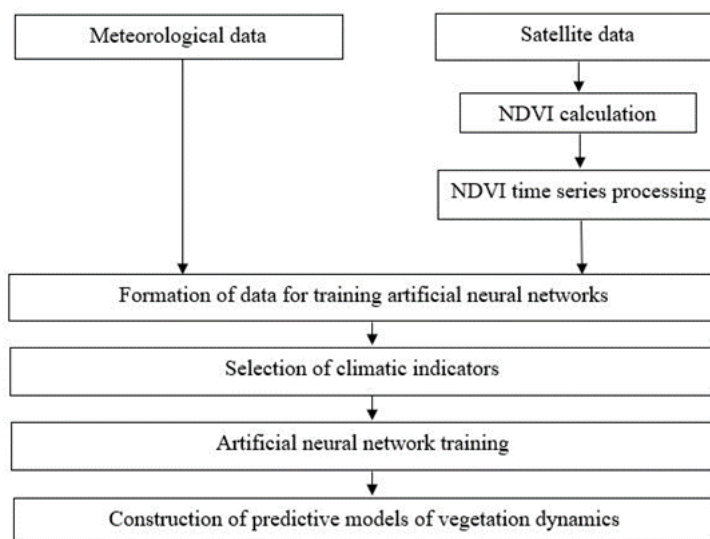

Figure 3. Data selection and processing methodology

The supervised learning was used to train the neural network. The training was carried out using error back propagation method [1]. Satellite data were used jointly with meteorological data at the training stage. A set of meteorological time series were placed into the model as the neural network input, and the preprocessed NDVI time series were placed as the reference output (neural network training scheme is shown in Fig. 4). Training on different type meteorological time series was performed repeatedly. Table 1 summarises the data used to train the ANN.

Figure 4. ANN training scheme

Table 1. Example of the data series used to train the ANN

		Input			Output
	Day of the year	Surface air temperature	Precipitation	Snow cover height	NDVI
2006	1	-2.3	0.0	35	0.0135
	2	-2.6	0.0	33	0.0222
	…	…	…	…	…
	365	-6.7	0.0	3	0.0539
2007	1	-0.2	7.1	5	-0.0266
	2	-0.1	1.0	7	-0.0072
	…	…	…	…	…
	365	1.9	0.0	4	0.0226
…					
2016	1	-12.4	1.1	3	0.0516
	2	-15.9	0.0	4	0.0523
	…	…	…	…	…
	365	0.4	0.7	13	-0.2040

Training on day of the year, surface air temperature, precipitation, and snow cover height time series was performed to select an optimal number of hidden layers of neurons in ANN. The quality of neural network training was determined by the average relative errors in NDVI prediction for the entire period and for the periods during which the index value exceeded 0.2. The time series example of the observed and predicted NDVI for 11 years with 10 hidden layers is shown in Fig. 5.

The average relative error for the entire period was 22%, and for the periods during which NDVI exceeded the 0.2 value was 11%. ANNs with 10 hidden layers were also trained on different combinations of meteorological indicators. Best results were obtained by ANN training on all four indicators (including the day number). However, its application can be difficult practically due to a possible time shift of the growing season.
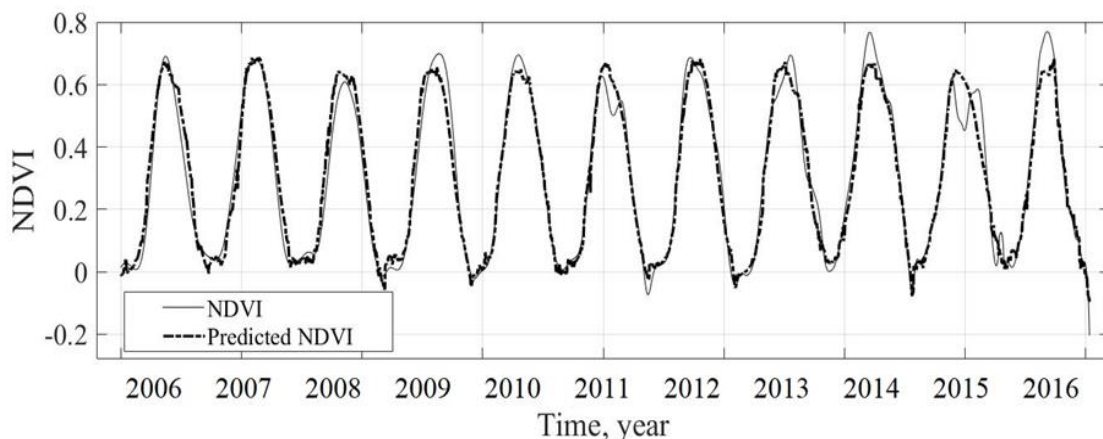
Figure 5. The time series of the observed and predicted NDVI

DISCUSSION

To test ANN-based models, NDVI prediction was carried out on the basis of continuous year-size meteorological time series outside of the training sample period. Year-size NDVI time series was predicted, thus the NDVI forecasting situation was modelled (Fig. 6).
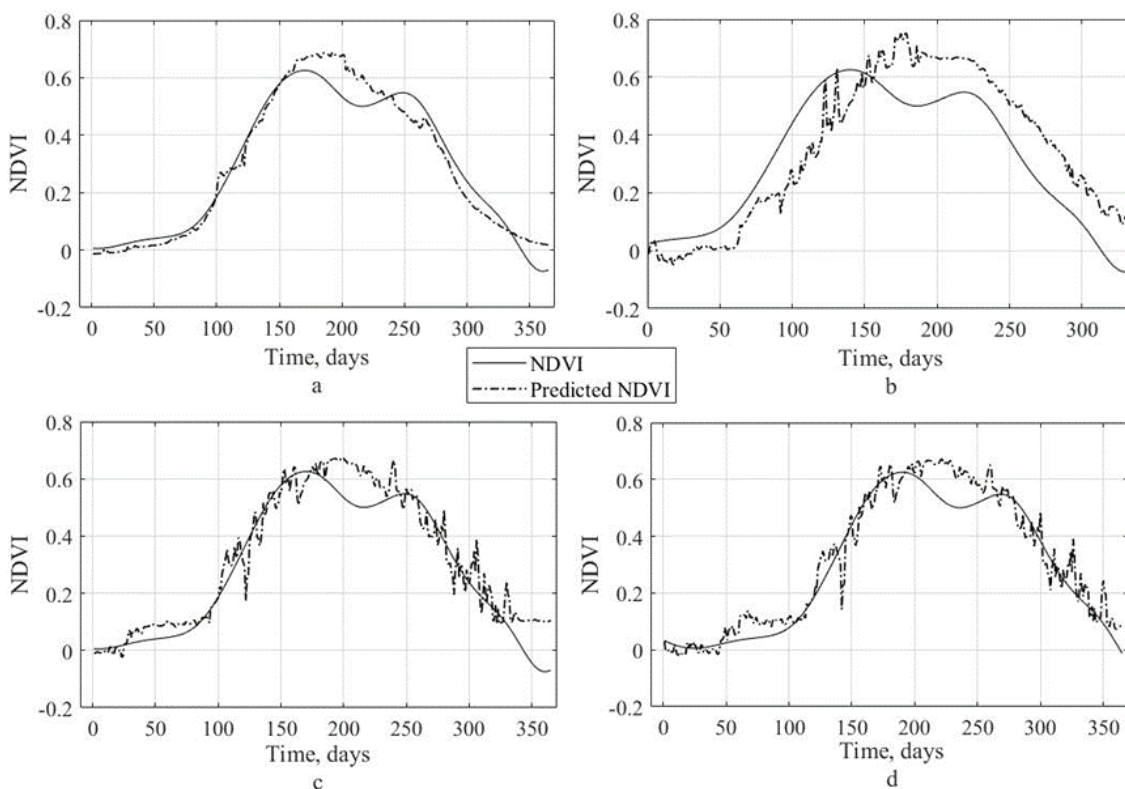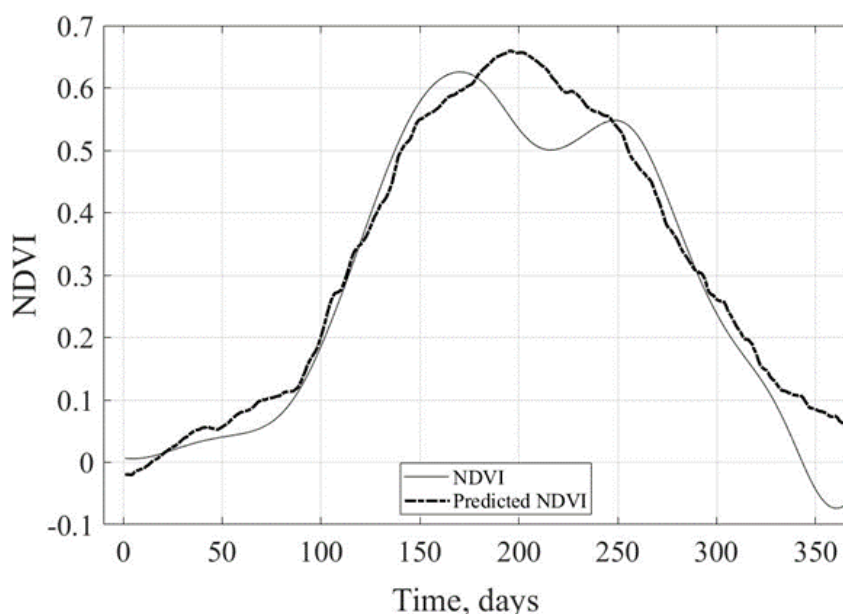

Figure 6. Observed and predicted NDVI time series in 2011

Figure 6a demonstrates the example of 2011 time series derived from satellite data and predicted by ANN based on four parameters (day of the year, daily average surface air temperature, daily average precipitation, average snow cover height). The average relative error in the period during which the NDVI exceeded 0.2 is 9%. However, if the

meteorological data time series are shifted in time (due to a weather anomaly, for example), the ANN prediction can be incorrect (Fig. 6b). Therefore, we can exclude the day number series and train ANN on three parameters. The time series predicted on three parameters is shown in Fig. 6c. The average relative error in the period during which the NDVI exceeded 0.2 is 14.1%. In this case, NDVI prediction becomes insensitive to the time shift of the meteorological parameter time series. Figure 6d demonstrates the forecasted NDVI time series shifted by thirty days to the right. The prediction quality is not changed that can be important when predicting the dynamics of vegetation cover in new areas.

Smoothing was proposed to reduce the noise level of the predicted NDVI time series. The smoothed computed and predicted NDVI time series are shown in Figure 7.

Figure 7. Observed and predicted NDVI time series in 2011

The average relative error in the period during which the NDVI exceeded 0.2 was 13.7%. The area under the time series graphs (for values above 0.2) can be used as an indicator of the total vegetation cover activity. The areas can be estimated as 106 and 115 units, respectively for the computed and predicted graphs. In this case, the error is estimated as 8%. Current results demonstrates (at least) partial applicability of proposed approach. However, additional testing and collection of modelling statistics on different test areas are needed to be realised as a part of future work.

Acknowledgement

REFERENCES

[1] Haykin S. Neural Networks and Learning Machines. 906 p., 2009.

[2] Kataev M.Yu., Bekerov A.A., Medvetskiy D.V. Method of smoothing NDVI time series. TUSUR reports, 1, pp. 85-88, 2017. (in Russian)

[3] Kim P. MATLAB Deep Learning: with Machine Learning, Neural Networks and Artificial Intelligence. 172 p., 2017.

[4] Medvedeva M.A., Bartalev S.A., Lupyan E.A., Matveev A.M., Tolpin V.A., Poida A.A. The possibility of estimation of the growing season onset basing on satellite and meteorological data. Sovremennye Problemy Distantsionnogo Zondirovaniya Zemli iz Kosmosa [Contemp. Probl. of Remote Sens. of the Earth from Space], 5(2), pp. 313-321, 2008. (in Russian)

[5] Miklashevich T.S., Bartalev S.A. Method for estimating vegetation cover phenological characteristics. Sovremennye Problemy Distantsionnogo Zondirovaniya Zemli iz Kosmosa [Contemp. Probl. of Remote Sens. of the Earth from Space], 13(1), pp. 9-24, 2016. doi:10.21046/2070-7401-2016-13-1-9-24 (in Russian)

[6] Panidi E., Tsepelev V., Torlopova N., Bobkov A. Mapping of the land cover spatiotemporal characteristics in northern Russia caused by climate change. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLI-B8, pp. 997–1002, 2016. doi:10.5194/isprs-archives-XLI-B8-997-2016

[7] Schowengerdt R.A. Remote Sensing, Models and Methods for Image Processing. 560 p., 2006.

[8] Plotnikov D.E., Miklashevich T.S., Bartalev S.A. Reconstruction of time series of remote sensing data by the method of polynomial approximation in a sliding window of variable size. Sovremennye Problemy Distantsionnogo Zondirovaniya Zemli iz Kosmosa [Contemp. Probl. of Remote Sens. of the Earth from Space], 2, pp. 103-110, 2014. (in Russian)

[9] Lambert J., Drenou C., Denux J.P., Gérard B., Chéret V. Monitoring forest decline through remote sensing time series analysis. GIScience and Remote Sensing, 50(4), pp. 437-457, 2013.

[10] Mazza A., Gargiulo M., Gaetano R., Scarpa G. Estimating the NDVI from SAR by convolutional neural networks. IEEE, pp. 1954-1957, 2018.

[11] Truhlyak E.V., Trubilin E.I. Precision Agriculture. 376 p., 2019. (In Russian)

[12] Tsepelev V., Panidi E., Torlopova N., Bobkov A. Study of climate change at the north of European Russia using land cover data. SGEM2015 Conference Proceedings, Book 3, Vol. 2, pp. 529-536. doi:10.5593/SGEM2015/B32/S14.071

[13] Vermote E. MOD09A1 MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC, 2015. doi:10.5067/MODIS/MOD09A1.006