Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 30, 2018, Sofia

DOI: https://doi.org/10.3897/arb.v30.e07

REMOTE ESTIMATION OF CROP CANOPY PARAMETERS BY STATISTICAL REGRESSION ALGORITHMS FOR WINTER RAPESEED USING SENTINEL-2 MULTISPECTRAL IMAGES

Dessislava Ganeva, Eugenia Roumenina

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: d.ganeva@space.bas.bg

Abstract

Estimation of crop canopy parameters is important task for remote sensing monitoring of agriculture and constructing strategies for within-field management. The main objective of this study is to evaluate the retrieval from Sentinel-2 images by parametric and non-parametric statistical models several crop canopy parameters for monitoring before winter and after winter rapeseed crop in real farming conditions of North East Bulgaria. For the calibration of the models in-situ data from three field campaigns is used. For most of the studied parameters models with good accuracy were identified, except for aboveground fresh biomass. The best identified model for vegetation fraction (RMSEcv = 0.14 %) and plant density (RMSEcv = 9 nb/m^2) were parametric models with three band vegetation index (3BSI-Tian) and linear fitting function for the first, three band vegetation index (3BSI-Verreslt) and polynomial for the second parameter. For aboveground dry biomass (RMSEcv = 52 g/m^2), mean plant height (RMSEcv = 4 cm) and nitrogen content in fresh biomass (RMSEcv = 2 g/kg) the best models were non-parametric, Gaussian Processes Regression for the first parameter and Variational Heteroscedastic variant of the Gaussian Processes Regression for the other two.

Introduction

Quick and accurate retrieval of crop canopy parameters is of importance for remote sensing monitoring of agriculture. Rapeseed is one of the most important oilseed crops worldwide [1] and it is monitored by traditional methods [2] as well as by remote sensing [3–5] .

One of the important periods in the development of winter rapeseed crop is the before winter and after winter period [6]. Before winter, the condition of the crop is evaluated regarding its preparation to withstand the winter meteorological conditions. The rapeseed crop should be developed enough before winter, but the plants should not be developed beyond a certain threshold because they are more sensitive to low temperatures and frost. The rapeseed plants compete for light and when there is an uneven plant density or weed development, the plants develop in height and again become more susceptible to frost. After winter, the condition of the crop is evaluated regarding its plant density, phenological phase and frost damage. A decision is often made to keep the rapeseed crop because enough plants survived the winter and they have good change for a satisfactory yield, or to destroy the rapeseed crop and plant the field with spring crop.

The traditional monitoring is based on parameters like plant density, phenological phase, biomass. This study is focused on remote estimation of several crop canopy parameters of winter rapeseed with the purpose to be able to use them for before winter and after winter monitoring and within-field management.

Historically, Vegetation indices (VI) with regression functions are used to retrieve biophysical parameters, such as aboveground biomass [7] or vegetation fraction [8], agronomical parameters, such as plants density [9], or growth parameters, such as plant height [3]. More recently, VI types of formulas are used with all available wavelength and different regression functions for retrieval of biophysical parameters, such as leaf area index and leaf chlorophyll content, called "spectral index optimization" [10] and retrieval of nitrogen concentration, was obtained by non-parametric models [11–15]. All those methods use in-situ sampled, or simulated, data for the modeled canopy parameters and associated surface reflectance from remote sensing sensor. In our approach we use in-situ sampled data and Sentinel-2 associated surface reflectance.

In order to determine the variability of the fields before the actual collection of the field data [16], the sample locations were determined by calculating VI, Table 1, that correlate with the rapeseed canopy parameters that we are interested in. A VI that was not tested with rapeseed but was specially tested with Sentinel-2 images [17] was also included in the list. The selected VI, Table 1, were calculated on the Sentinel-2 image from 12.11.2017, downloaded from Copernicus Data Open Hub in 2A product (https://scihub.copernicus.eu). The samples were positioned manually to capture the maximum vegetation variability of the test fields and in relatively homogeneous surrounding area of 20 m^2 in terms of vegetation density and growth phase

The main objective of this study is to evaluate the retrieval from Sentinel-2 images by parametric and non-parametric statistical models several crop canopy parameters for monitoring before winter and after winter rapeseed crop.

Table 1. Vegetation Indices used to position the location of the samples for the field campaign

Estimated parameter	Vegetation Index	Formula	Reference
Aboveground biomass, number of plants per m^2 after emergence	RVI (Ratio Vegetation Index)	NIR / Red	[9]
Aboveground biomass	OSAVI (Optimized Soil Adjusted Vegetation Index)	(1 + L)(NIR − Red) / (NIR + Red + L) (L = 0.16)	[7]
LAI	SAVI (Soil Adjusted Vegetation Index)	(1 + L)(NIR − Red) / (NIR + Red + L) (L = 0.5)	[18]
Canopy chlorophyll and nitrogen	CiredEdge	R783/R705 − 1	[17]
Plan height	EVI (Enhanced Vegetation Index)	EVI = G × ((RNIR − Rred) / (RNIR + C1 × Rred − C2 × Rbleu +L)); G = 2.5; C1 = 6; C2=7.5; L = 1	[3]
Vegetation Fraction	VARIgreen (Visible Atmospherically Resistant Index)	(R550 – R670) / (R550 + R670)	[8]
Number of plants per square meter after emergence	NDVI (Normalized Difference Vegetation Index)	(NIR − Red) / (NIR + Red)	[9]

Materials and Methods

Study area

This study area is part of the East Danube plain in Bulgaria, Fig. 1, and the study period was over one growing season, from September 2017 to July 2018, on three mass fields sown with different hybrids of winter rapeseed. The area is mostly flat, the soil has mainly sandy loam texture, the climate in this region is Moderate Continental with cold winters and hot summers (mean daily temperature 10.2º C), and an annual cumulative rainfall of 540 mm.

The bigger of the field plots (P1) is 137 ha and was planted for three days, from 03.09.2017, with sowing rate of 80 plants/m^2. The other two plots are smaller, one (P2) is 10 ha, planted on 28.08.2017 at sowing rate of 56 plants/m^2 and the other (P3) is 15 ha, planted on 04.09.2017 at sowing rate of 76 plants/m^2.

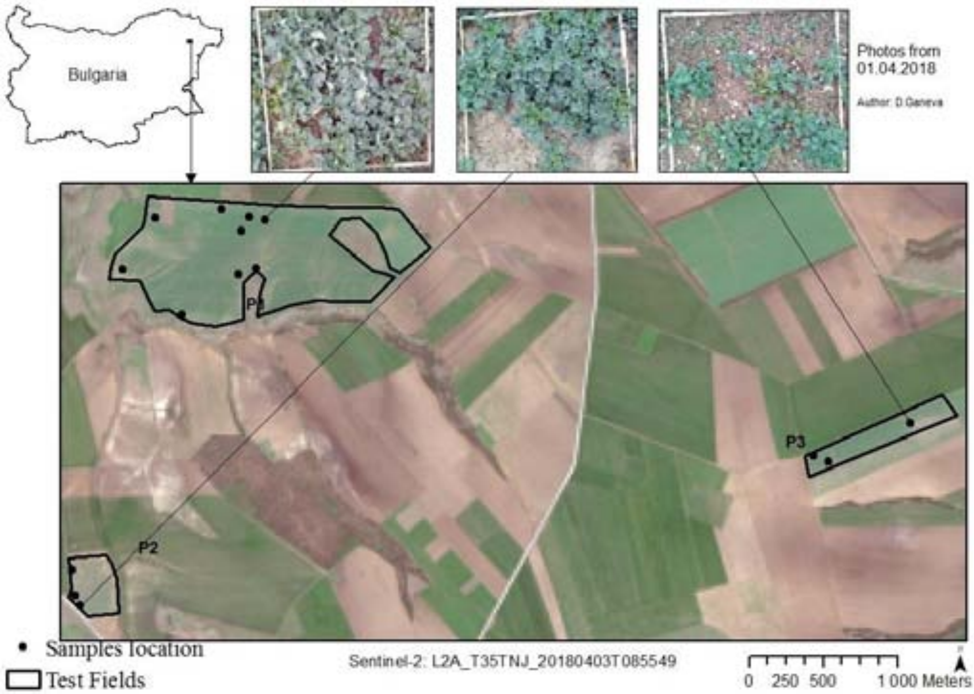


Fig. 1. Location of the field sites near Dobrich, North East Bulgaria

In-situ data

The crop canopy parameters were measured within 1 m^2 elementary sampling units (ESU) at the pre-defined samples location. Each ESU was geo-located by means of a GPS, with an accuracy of ±3–5 m. For the aboveground fresh biomass (FBM), all plants in the ESU were manually cut, stored in paper bags and transported to a laboratory.

In the laboratory the FBM is weighted (in g/m^2) and then, a sample of it is oven-dried at 105° C until constant weight to obtain the dry biomass (DBM), measured in g/m^2. In each ESU the density (number winter rapeseed per m^2, NbPlant) was counted and recorded. The mean plant height of all plants (PlantH) in an ESU was recorded as well, in cm. The vegetation fraction (VF) as described by [19] was expertly estimated and photographed. The recorded VF includes the rapeseed plants and the weeds. The nitrogen content in g/kg (N) was measured from a sample of the FBM by Kjeldahl method [20]. The crop canopy parameters were measured during three field campaigns, two before winter and one after winter, Table 3. At each field campaign, Table 3, the ESU at a given sample location was positioned near the one from the previous field campaign.

Table 2. Descriptive statistics for the measured crop canopy parameters

Parameters	Number of ESU	Mean	SE Mean	StDev	Minimum	Maximum
FBM (g/m^2)	30	1058.00	151.00	829.00	146	2 761
DBM (g/m^2)	30	124.50	17.80	97.60	18	332
VF (%)	30	0.58	0.05	0.28	0.15	1.00
NbPlant (nb/m^2)	30	36.13	3.18	17.42	10	95
PlantH (cm)	45	15.67	0.83	5.57	9	28
N (g/kg)	30	10.28	0.26	1.39	7.54	13.46

Because of the meteorological conditions, some of the plants started growing immediately after sowing but many had more than a month delay. Particularly the plots P2 and P3 were with plants in very different phonological phases, from BBCH13 to BBCH19 [21], during the before winter field campaign. This difference in the phenological phase was completely reduced after winter, where all plots were at BBCH50/BBCH51.

Remote sensing data

Sentinel-2, sensor S2A and S2B, data was used for the study. Cloud free images closest to the field data collection were selected at Copernicus Open Hub Access (https://scihub.copernicus.eu). All spectral bands of 10 m and 20 m spatial resolution are used, from level 2A products. It provided 10 spectral bands from 490 nm to 2 190 nm, resampled at 10 m spatial resolution.

Methods

Two types of models for crop canopy retrieval are studied: (1) the parametric univariate regression models with a Vegetation index (VI) as independent variable and (2) the non-parametric multivariate regression models with all the 10 spectral bands from the remote sensing data as independent variables. In both models the crop canopy parameter is the dependent variable.

Table 3. Timing of the field campaigns and satellite acquisitions with corresponding measured parameters

Field Plot	Nb of ESU	Before Winter			After Winter		
		Sampling date	Crop Canopy param.	Sentinel-2 image date (sensor)	Sampling date	Crop Canopy param.	Sentinel-2 image date (sensor)
P1	9	23.11.2017	FBM, DBM, VF, NbPlant, PlantH	29.11.2017 (S2A)	01.4.2018	FBM, DBM, VF, NbPlant, PlantH, N	03.04.2018 (S2B)
		13.12.2017	PlantH, N	12.12.2017 (S2A)			
P2	3	23.11.2017	FBM, DBM, VF, NbPlant, PlantH	29.11.2017 (S2A)	01.4.2018	FBM, DBM, VF, NbPlant, PlantH, N	03.04.2018 (S2B)
		13.12.2017	PlantH, N	12.12.2017 (S2A)			
P3	3	24.11.2017	FBM, DBM, VF, NbPlant, PlantH	29.11.2017 (S2A)	01.4.2018	FBM, DBM, VF, NbPlant, PlantH, N	03.04.2018 (S2B)
		13.12.2017	PlantH, N	12.12.2017 (S2A)			

Model Selection

The first step consists in model conceptualization and selection. The selected models are parametric univariate and non-parametric multivariate regression models, as described in [22], and they are applied using the Automated Radiative Transfer Models Operator (ARTMO) package (http://ipl.uv.es/artmo/).

In the first type of models it is not really VI that is selected but rather general formula to be used, because all the top of canopy reflectance of the spectral bands are tested with all formulae, Table 4. Also, different fitting functions are used for the regression: linear, exponential, logarithmic, power and polynomial. The first six VI from Table 4, are the same one used in the preliminary study to determine the samples location. Two more 3 bands VI were added to the list because they are reported [23–24] to perform better for retrieval of biophysical variables and leaf nitrogen concentration.

The evaluated non-parametric multivariate regression models are Least Square Linear Regression (LSLR), Principal Components Regression (PCR), Partial Least Squares Regression (PLSR), Kernel Ridge Regression (KRR), Gaussian Processes Regression (GPR) and the Variational Heteroscedastic variant of the Gaussian Processes Regression (VHGP). For non-parametric multivariate regression models the "curse of dimensionality" [25] could represent a problem. Therefore, some of the selected models have dimension reduction, like PCR, PLSR, KRR. PCR and PLSR have been developed for cases, as is of this study, where there are many, possibly correlated, predictor variables and relatively few samples [26]. GPR is reported [27] to be a robust model for biophysical variables retrieval. GPR and VHGP are especially valuable [23] because they calculate a Coefficient of Variation ($CV = \sigma/\mu$), where σ is the Standard Deviation (SD) around the estimated variable and μ the mean estimated variable. CV provides relative uncertainty of the estimated variable in %. Finally, LSLR was selected as the oldest method for comparison with the others.

Model calibration/fitting evaluation and validation

The model calibration, evaluation and selection will be treated together, because model selection is an integral part of the model fitting process and furthermore in the present study, the goal is not to evaluate one model alone, but rather compare and select a model [28].

In the calibration step the parameters will be adjusted to make the model as consistent as the available data [29]. Model evaluation according to [30] consists in preliminary experiments and test of alternative scenarios. This part of the work was done previously in [31] and the results will be used in the development of the models in the present study. Namely, bare soil pixels are included in the input data for the models and the reflectance data is from 1 pixel, corresponding to the sample, without averaging with the pixels around. Because outliers in the data could cause distortion in the models [32], all crop canopy variables were tested for outliers with Grubbs test [33]. No outliers were detected at 5 % level of significance for smallest and largest data values, apart from one sample for NbPlant. The outlier sample is from plot P1, where the sowing rate is 80 plants/m^2 and the measured sample had 95 plants/m^2. The decision was to remove the sample, because it was probably due to local malfunctioning of the sowing machine.

Table 4. Type of VI testes in the study. Ra, Rb and Rc are the reflectance for a given wavelength.

Type index	The formula is from the VI	Formula	Reference
2 bands Simple ratio of reflectances	RVI (Ratio Vegetation Index)	Ra / Rb	[9]
2 bands normalized difference ratios of reflectances	OSAVI (Optimized soil adjusted index)	1.16 × (Ra – Rb) / (Ra + Rb + 0.16)	[7]
2 bands normalized difference ratios of reflectances	SAVI (Soil adjusted vegetation index)	1.5 × (Ra – Rb) / (Ra + Rb + 0.5)	[18]
2 bands Simple ratio of reflectances	CIredEdge (Red-edge Chlorophyll Index)	(Ra / Rb) − 1	[17]
3 bands normalized difference ratios of reflectances	EVI (The enhanced vegetation index)	2.5 × ((Ra – Rb) / (Ra + 6 × Rb – 7.5 × Rc + 1))	[3]
2 bands normalized difference ratios of reflectances	NDVI (Normalized Difference Vegetation Index)	(Ra – Rb) / (Ra + Rb)	[9]
3 bands normalized difference ratios of reflectances	3BSI-Verrelst	(Ra – Rc) / (Rb + Rc)	[23]
3 bands normalized difference ratios of reflectances	3BSI-Tian	(Ra – Rb – Rc) / (Ra + Rb + Rc)	[24]

In model evaluation there is also the notion of evaluation of uncertainty and accuracy. In the present study there is an uncertainty because of the way the field data is collected (from only 1m^2, without repetition of the measurements and geo-located with GPS with an accuracy of ± 3–5 m), that could be considered as part of the dependent variables noise, but this uncertainty is not considered in the present study.

Each goodness-of-fit statistic gives certain aspect of the model accuracy [34]. In this study the accuracy of the model will be based on Root Mean Square Error (RMSE), for the magnitude of error, and Coefficient of determination (R^2), for the spatial patterns [34].

Leave-one-out cross-validation (LOOCV) will be used to get the expected prediction error, to help select the best model and to ensure that the model does not overfit. All models will be ranked by Normalized RMSE (NRMSE in %), NRMSE% = 100 × (RMSE / Range(Obs)), because it is a goodness-of-fit measurement that is suitable for comparison [34].

*Table 5. Best parametric and non-parametric models for each studied crop canopy parameter. For all models, pixels with bare soil are included, except model marked by *.*

Variable	**Model**	**Best Bands**	**RMSEcv**	**R^2cv**	**NRMSEcv %**
FBM (g/m^2)	EVI / Polynomial	Ra:842, Rb:560, Rc:665	421	0.75	15
	PLSR		429	0.74	16
DBM (g/m^2)	EVI / Polynomial	Ra:842, Rb:560, Rc:665	52	0.73	16
	GPR	560;740;490;842	52	0.73	16
VF (%)	3BSI-Tian / Linear	Ra:842, Rb:705, Rc:560	0.14	0.81	14
	PCR		0.15	0.78	16
NbPlant (nb/m^2)	3BSI-Verrelst / Polynomial	Ra:665; Rb:842;Rc:560	9	0.73	14
	PLSR		10	0.68	15
PlantH (cm)	3BSI-Verrelst / Power*	Ra:490, Rb:842, Rc:560	4	0.53	20
	VHGR	560;842;665;490	4	0.72	15
N (g/kg)	VHGR	665;842	2	0.74	15

Results and discussion

The best parametric univariate regression models with a VI as independent variable are ranked by NRMSEcv and R^2cv > 0.5, for every variable, Table 5. However, the models for N and PlantH did not perform well with the added bare soil pixels. It was clearly visible in the scatter plot of “residuals vs predicted value” and “measured vs estimated”. New models for N and PlantH are fitted without the bare soil pixels. No satisfactory model was found for N. A weaker model was found for PlantH, Table 5.

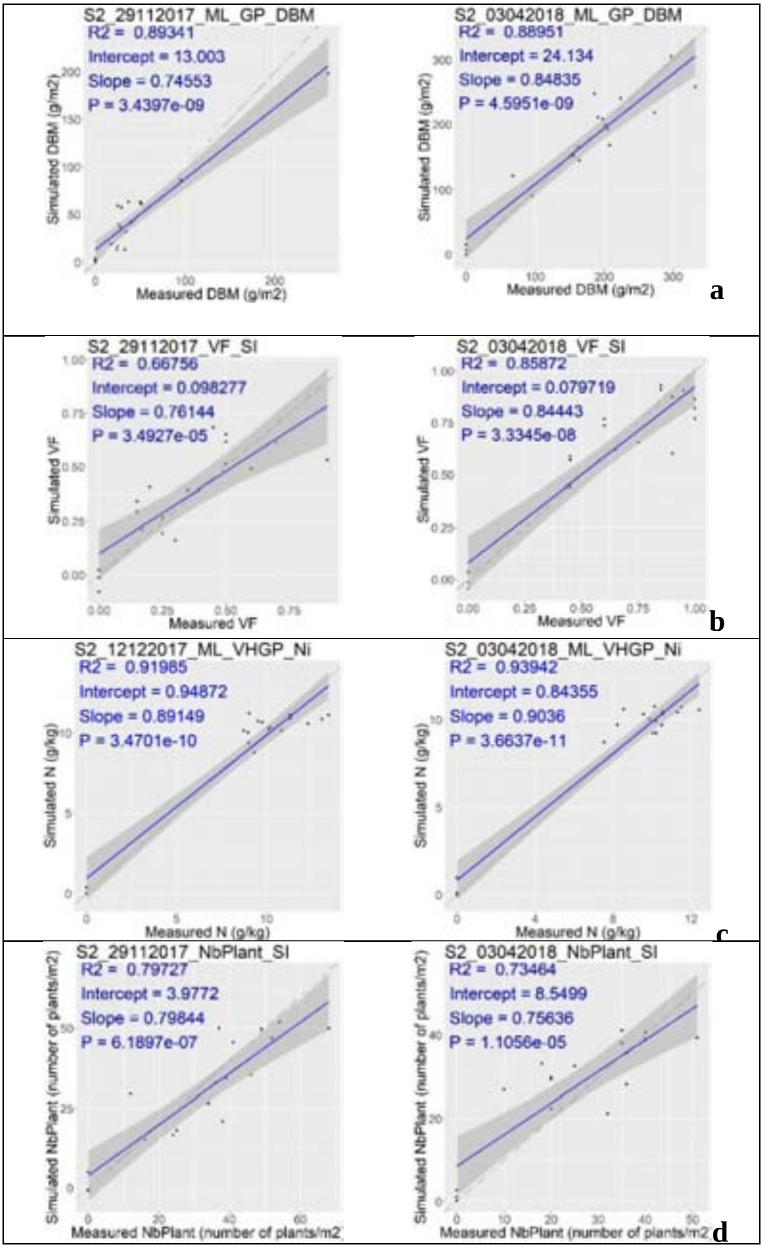
The best non-parametric multivariate regression models with all the spectral bands from the remote sensing data as independent variables, ranked by NRMSEcv and R^2cv > 0.5, for every variable are in Table 5. All non-parametric models behave correctly with the bare soil pixels included.

The comparison between parametric and non-parametric models for every crop canopy parameter shows, Table 5, that in terms of goodness-of-fit cross-validation statistic the models have comparable performance. To be able to further distinguish between their performance, each model is applied to the remote sensing images and the estimated values are compared to the measured values. The models that meet the criteria of $R^2 > 0.514$ [36] and slop > 0.6 are considered only, Table 6 and Fig. 2.

Table 6. Comparison between measured/estimated values from models, where $R^2>0.514$ and slop > 0.6

Model	R2	Intercept	Slop	P
S2_03042018_FBM_SI	0.84290	250.43	0.80541	7.86E-08
S2_03042018_ML_PLS_FBM	0.82123	316.59	0.78671	2.23E-07
S2_29112017_ML_GP_DBM	**0.89341**	**13.003**	**0.74553**	**3.44E-09**
S2_03042018_ML_GP_DBM	**0.88951**	**24.134**	**0.84835**	**4.60E-09**
S2_03042018_DBM_SI	0.82875	31.317	0.79268	1.58E-07
S2_29112017_VF_SI	**0.66756**	**0.098277**	**0.76144**	**3.49E-05**
S2_03042018_VF_SI	**0.85872**	**0.079719**	**0.84443**	**3.33E-08**
S2_12122017_ML_VHGP_Ni	**0.91985**	**0.94872**	**0.89149**	**3.47E-10**
S2_03042018_ML_VHGP_Ni	**0.93942**	**0.84355**	**0.9036**	**3.66E-11**
S2_29112017_ML_PLS_NbPlant	0.78664	5.9884	0.81215	9.37E-07
S2_03042018_ML_PLS_NbPlant	0.77134	7.2667	0.70727	3.55E-06
S2_29112017_NbPlant_SI	**0.79727**	**3.9772**	**0.79844**	**6.19E-07**
S2_03042018_NbPlant_SI	**0.73464**	**8.5499**	**0.75636**	**1.11E-05**
S2_29112017_ML_VHGP_PlantH	**0.86076**	**2.2378**	**0.76445**	**2.96E-08**
S2_12122017_ML_VHGP_PlantH	**0.70192**	**3.7626**	**0.63536**	**1.43E-05**
S2_03042018_ML_VHGP_PlantH	**0.87996**	**2.8455**	**0.81848**	**8.96E-09**

No model could estimate, according to the proposed criteria, the FBM before winter, however the DBM is estimated for before and after winter with the GPR model. On the other hand, FBM and DBM are highly correlated with Pearson's correlation (r) of 0.99. For both periods the VF is estimated by 3BSI-Tian/linear, N and PlantH by VHGP. The NbPlant is estimated by the two types of models, the simpler will be considered, namely 3BSI-Verrelst/Polynomial.


S2_29112017_ML_GP_DBM
R2 = 0.89341
Intercept = 13.003
Slope = 0.74553
P = 3.4397e-09
Simulated DBM (g/m2)
Measured DBM (g/m2)
S2_03042018_ML_GP_DBM
R2 = 0.88951
Intercept = 24.134
Slope = 0.84835
P = 4.5951e-09
Simulated DBM (g/m2)
Measured DBM (g/m2)
a
S2_29112017_VF_SI
R2 = 0.66756
Intercept = 0.098277
Slope = 0.76144
P = 3.4927e-05
Simulated VF
Measured VF
S2_03042018_VF_SI
R2 = 0.85872
Intercept = 0.079719
Slope = 0.84443
P = 3.3345e-08
Simulated VF
Measured VF
b
S2_12122017_ML_VHGP_Ni
R2 = 0.91985
Intercept = 0.94872
Slope = 0.89149
P = 3.4701e-10
Simulated N (g/kg)
Measured N (g/kg)
S2_03042018_ML_VHGP_Ni
R2 = 0.93942
Intercept = 0.84355
Slope = 0.9036
P = 3.6637e-11
Simulated N (g/kg)
Measured N (g/kg)
c
S2_29112017_NbPlant_SI
R2 = 0.79727
Intercept = 3.9772
Slope = 0.79844
P = 6.1897e-07
Simulated NbPlant (number of plants/m2)
Measured NbPlant (number of plants/m2
S2_03042018_NbPlant_SI
R2 = 0.73464
Intercept = 8.5499
Slope = 0.75636
P = 1.1056e-05
Simulated NbPlant (number of plants/m2)
Measured NbPlant (number of plants/m2
d

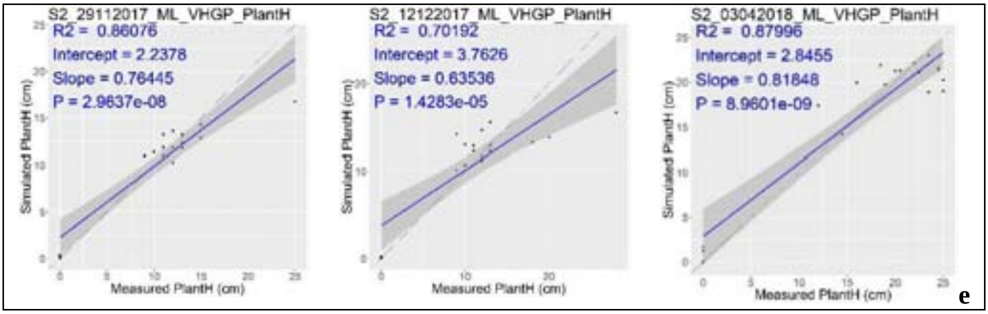


Fig. 2. Scatter plots of measured/estimated values from models, where $R^2 > 0.514$ and slop > 0.6. a) DBM; b) VF; c) N; d) NbPlant; e) PlantH

It is difficult to make comparison between the retrieval methods from the literature, because the different studies did not evaluate the same variables or used other measurement for accuracy [3, 7]. The DBM was evaluated by fitting parametric models and VIs [7], but the DBM is sampled from the whole growing season of the rapeseed, not only until Inflorescence Emergence stage (BBCH50). The NbPlant was sampled only at emergence and no measurement of accuracy is published [9] or Unmanned Aerial Vehicule (UAV) is used with shape feature recognition or classification [36–37]. The VF is studied with UAV with spatial resolution of 2.5 cm [8]. The model for retrieval of PlantH was evaluated by standard error of the estimate (SEE) [3], which is not one of the goodness-of-fit measurements of the present study. PlantH was also evaluate with RGB camera mounted on a UAV [38–39].

Selected model for each variable

Aboveground Fresh Biomass (FBM) in g/m^2: no model could estimate, according to the proposed criteria, the FBM before and after winter.

Dry Biomass (DBM) in g/m^2: for DBM retrieval, the GPR model is selected with the most relevant bands being the blue, green, middle red-edge and near infrared (λ = 560; 740; 490; 842). In related studies [3, 7, 40, 41] it was found that for crops or grass, DBM is best retrieved with visible and near-infrared reflectance which is in accordance with our results.

The relative uncertainty, expressed by CV, is higher than 10 % for all periods, Fig. 3, meaning that those areas are retrieved with high uncertainty. The higher is the DBM the lower is the relative uncertainty of the model.

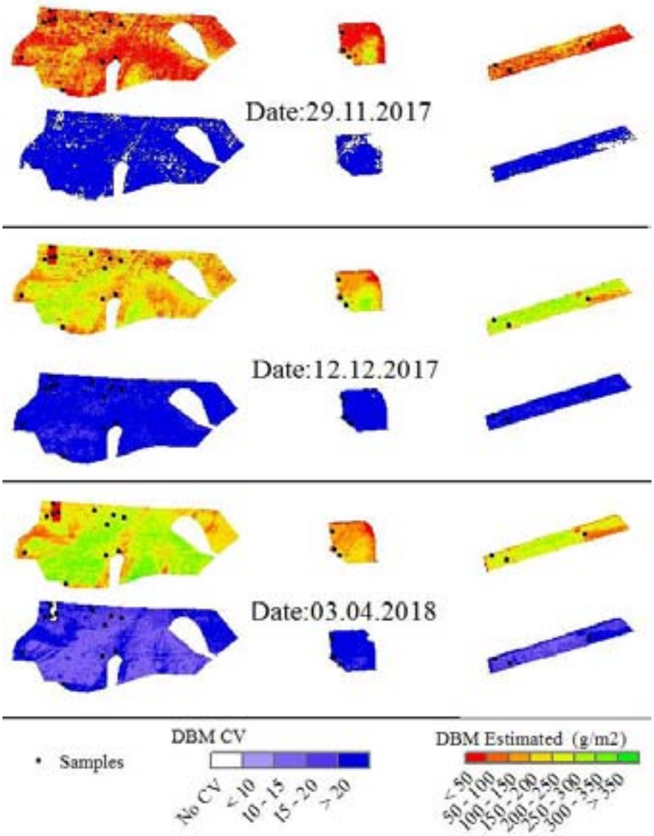


Fig. 3. GPR model for DBM retrieval, applied to the selected remote sensing images before winter and after winter

Vegetation Fraction (VF): for VF retrieval, the 3BSI-Tian VI and linear fitting function is selected with the most relevant bands being the green, shortest red-edge and near infrared (λ = 560; 705; 842). In related studies [8] it is reported that the VF for rapeseed crop before flowering is best correlated with green and red reflectance. For vegetation in general, VF is reported to be characterized [42] by green, red, shortest infra-red reflectance, red and near infra-red [43] or even only by visible reflectance [8]. It was remarked [43] that for higher vegetation cover the estimation is poorer. However, we cannot verify this statement in our study as no model with uncertainty calculation was selected.

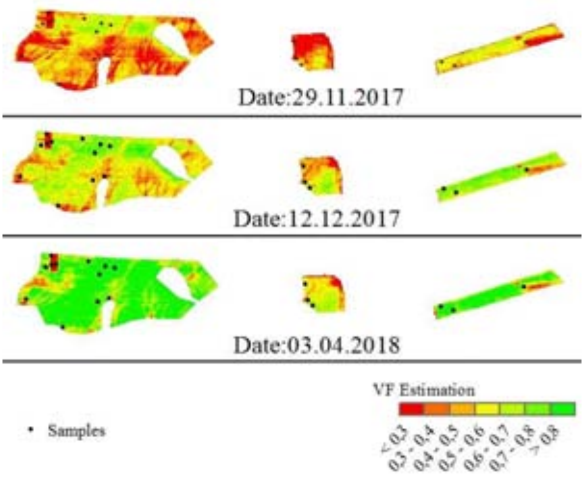


Fig. 4. 3BSI-Tian VI / linear model for VF retrieval, applied to the selected remote sensing images before winter and after winter

Plant density (NbPlant) in nb/m^2: for NbPlant retrieval, the 3BSI-Verrelst VI and polynomial fitting function is selected with the most relevant bands being the green, red and near infrared (λ = 560; 665; 842). In related studies [9, 44] it is reported that NbPlant is best correlated to red, green and near infrared reflectance which is in accordance with our results. The after winter period shows decrease of the NbPlant, Fig. 5, compared to the before winter period. The sampled measurements show the same trend.

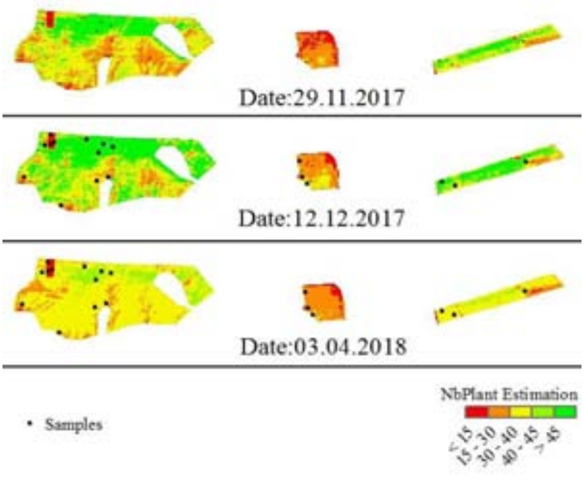


Fig. 5. 3BSI-Verrelst VI / polynomial model for NbPlant retrieval, applied to the selected remote sensing images before winter and after winter

Mean Plant Height (PlantH) in cm: for PlantH retrieval, the VHGP model is selected with the most relevant bands being the visible and near infrared (λ = 490; 560; 665; 842). In related studies [3] it is reported that PlantH is best correlated to same reflectance as calculated by our model, Fig. 6.

Nitrogen content in fresh biomass (N) in g/kg: for N (g/kg) retrieval, the VHGP model is selected with the most relevant bands being the red and near infrared (λ = 665; 842). In related studies [12–14, 41] it is reported that N is best correlated to visible and near infra-red. Some of those studies report also correlation to the longer red-edge, others to the shorter red-edge. One of the relevant bands selected by our model is close to the absorption of chlorophyll, located around 675 nm, and reported to characterize the nitrogen status of leaves [45] and the other is the near infrared that is reported [47] to determine the canopy structure. Crop phenology is found to be cause for substantial difference in canopy reflectance vs canopy chlorophyll content [46] and therefore canopy N. This is probably why in our study no parametric model could fit the before and after winter crop stage. However, the selected VHGP model gives almost constant uncertainty, between 10–20 %, Fig. 7, for before and after winter period where the rapeseed crop is in different phenological phases.

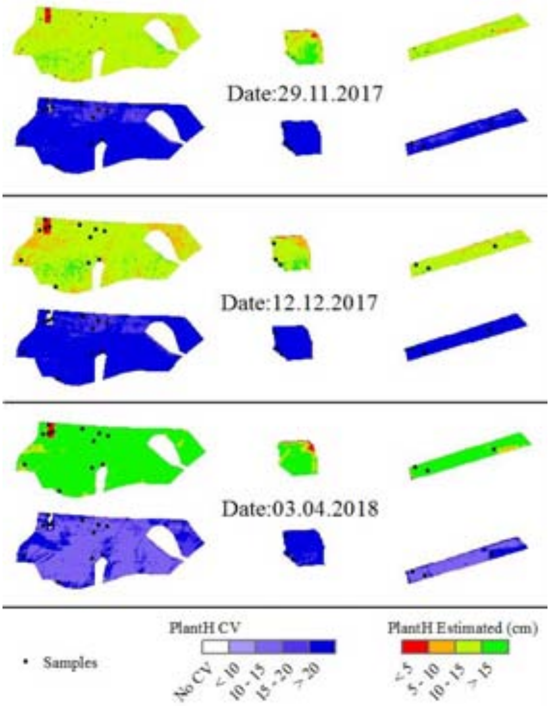


Fig. 6. VHGP model for PlantH retrieval, applied to the selected remote sensing images before winter and after winter

Conclusion

This study demonstrated that rapeseed crop canopy parameters before and after winter, such as, aboveground dry biomass (DBM), vegetation fraction (VF), plant density (NbPlant), mean plant height (PlantH), nitrogen content (g/kg) in fresh biomass (N), can be retrieved directly from remote sensing measurement from Sentinel-2 with good accuracy in real farming conditions of North East Bulgaria.

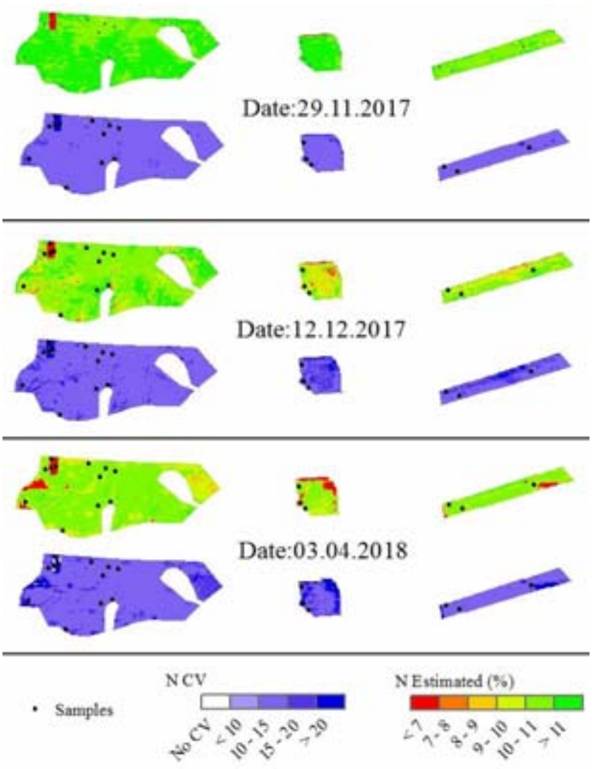


Fig. 7. VHGP model for N retrieval, applied to the selected remote sensing images before winter and after winter

For some of the parameters, VF (RMSEcv = 0.14 %), NbPlant (RMSEcv = 9 nb/m^2), the best models were parametric and for others, DBM (RMSEcv=52 g/m^2), PlantH (RMSEcv = 4 cm), N (RMSEcv = 2 g/kg), non-parametric. It also demonstrated that for FBM no model could be selected for the both periods: before and after winter. The retrieved parameters and classification techniques can be used for delineating within-field units for which to develop different management.

Some of the studies parameters are not independent from one another. For example, FBM, VF and NbPlant, or PlantH, NbPlant and FBM. To be able to increase their retrieval accuracy multi-output regression algorithms can be used [50] in a future work.

Acknowledgment

This paper has been partially supported by the Research Grant Award № ДФНП-17-43/26.07.2017 to the project "Assessment of the condition of winter rapeseed crops in Northeastern Bulgaria through satellite and terrestrial data" of the Program for Supporting Young Scientists and PhD Students – 2017 from the Bulgarian Academy of Sciences.

We would like to thank Mr. Georgi Gerdjikov, agronomist at Pioneer™, for his valuable agronomical information and help with the field sites.

References

1. FAOSTAT, 2014. Online statistical database of the Food and Agricultural Organization of the United Nations. http://faostat.fao.org
2. Diepenbrock, W., Yield analysis of winter oilseed rape (Brassica napus L.): A review. Field Crops Research, 2000, 67, 1, 35–49.
3. Bartoszek, K., Usefulness of MODIS data for assessment of the growth and development of winter oilseed rape, Zemdirbyste-Agriculture, 2014, 101, 4, 445–52.
4. She, B., J. F. Huang, D. Y. Zhang, and L. S. Huang., Assessing and Characterizing Oilseed Rape Freezing Injury Based on MODIS and MERIS Data. International Journal of Agricultural and Biological Engineering, 2017, 10, 3, 143–57.
5. Wojtowicz, M., A. Wojtowics, and J. Piekarczyk. Application of Remote Sensing Methods in Agriculture. Communication in Biometry and Crop Science. 2015, 11, 1, 31–50.
6. Иванова, Р., Рапицата - култура на настоящето и бъдещето. Второ допълнително и преработено издание. Издателство «Виденов и син», София, 2012. стр. 306. (In Bulgarian)
7. Han, J., C. Wei, Y. Chen, *et. al*. Mapping Above-Ground Biomass of Winter Oilseed Rape Using High Spatial Resolution Satellite Data at Parcel Scale under Waterlogging Conditions. Remote Sensing, 2017, 9, 3, 1–21.
8. Fang, S., W. Tang, Y. Peng, *et. al*. Remote Estimation of Vegetation Fraction and Flower Fraction in Oilseed Rape with Unmanned Aerial Vehicle Data. Remote Sensing, 2016, 8, 416, 1–19.
9. Piekarczyk, J., M. Wójtowicz, and A. Wójtowicz.. Estimation of Agrinomic Parameters of Winter Oilseed Rape from Field Reflectance Data. Acta Agrophysica. 2006, 8, 1, 205–218.
10. Rivera, J. P., J. Verrelst, J. Delegido, F. Veroustraete, and J. Moreno. On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization. Remote Sensing, 2014, 6, 6, 4927–51.

11. Yu, X., H. Lu, and Q. Liu. Deep-Learning-Based Regression Model and Hyperspectral Imaging for Rapid Detection of Nitrogen Concentration in Oilseed Rape (Brassica Napus L.) Leaf. Chemometrics and Intelligent Laboratory Systems, 2018, 172, 188–93.
12. Li, L., S. Wang, T. Ren, et al. Ability of Models with Effective Wavelengths to Monitor Nitrogen and Phosphorus Status of Winter Oilseed Rape Leaves Using in Situ Canopy Spectroscopy. Field Crops Research, 2018, 215, 1, 173–186.
13. Li, L., J. Lu, S. Wang, et al. Methods for Estimating Leaf Nitrogen Concentration of Winter Oilseed Rape (Brassica Napus L.) Using in Situ Leaf Spectroscopy. Industrial Crops and Products, 2016, 9, 194–204.
14. Zhang, X., F. Liu, Y. He, and X. Gong. Detecting Macronutrients Content and Distribution in Oilseed Rape Leaves Based on Hyperspectral Imaging. Biosystems Engineering, 2013, 115, 1, 56–65.
15. Wang, F., J. Huang, Y. Wang, Z. Liu, and F. Zhang. Estimating Nitrogen Concentration in Rape from Hyperspectral Data at Canopy Level Using Support Vector Machines. Precision Agriculture, 2013, 14, 2, 172–183.
16. Bohnam, C. D. Measurements for Terrestrial Vegetation Measurements, 2nd edition ed. Wiley-Blackwell, 2013. 246 p.
17. Clevers, J. G. P. W, and A. A Gitelson. Remote Estimation of Crop and Grass Chlorophyll and Nitrogen Content Using Red-Edge Bands on Sentinel-2 and -3. International Journal of Applied Earth Observations and Geoinformation, 2013, 23, 344–351.
18. Hatfield, J. L., and J.H. Prueger. Value of Using Different Vegetative Indices to Quantify Agricultural Crop Characteristics at Different Growth Stages under Varying Management Practices. Remote Sensing, 2010, 2, 2, 562–78.
19. Gitelson, A. A., Y. J. Kaufman, R. Stark, and D. Rundquist. Novel Algorithms for Remote Estimation of Vegetation Fraction. Remote Sensing of Environment, 2002, 80, 1, 76–87.
20. AOAC, Nitrogen (Total) in Fertilizers: Kjeldahl Method No. 978 1990, 02, 18–19
21. Lancashire, P. D., H. Bleiholder, T. Van Den Boom, et al. A Uniform Decimal Code for Growth Stages of Crops and Weeds. Annals of Applied Biology, 1991, 119, 3, 561–601.
22. Verrelst, J., G. Camps-Valls, J. Muñoz-Marí et al.. Optical Remote Sensing and the Retrieval of Terrestrial Vegetation Bio-Geophysical Properties - A Review. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 108, 273–290.
23. Verrelst, J., J. Pablo, F. Veroustraete, et al. Experimental Sentinel-2 LAI Estimation Using Parametric, Non-Parametric and Physical Retrieval Methods – A Comparison. ISPRS J. of Photogrammetry and Rem. Sens., 2015, 108, 260–272.
24. Tian, Y. C., K. J. Gu, X. Chu, et al. Comparison of Different Hyperspectral Vegetation Indices for Canopy Leaf Nitrogen Concentration Estimation in Rice. Plant and Soil, 2013, 376, 1, 193–209.
25. Bellman, R. Adaptive Control Processes: A Guided Tour. Princeton, NJ: Princeton. University Press. 2016. 276 p.
26. Mevik, B-H., and R. Wehrens. The Pls Package: Principal Component and Partial Least Squares Regression in R. Journal of Statistical Software, 2007, 18, 2, 1–22.

27. Verrelst, J., L. Alonso, G. Camps-valls, J. Delegido, and J. Moreno, Retrieval of Vegetation Biophysical Parameters Using Gaussian Process Techniques, IEEE Transactions on Geoscience and Remote Sensing, 2012, 50, 5, 1832–43.
28. Cawley, G. C., and N. L. Talbot, On Over-Fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation, Journal of Machine Learning Research, 2010, 11, 2079–2107.
29. Rykiel, E. J, Testing Ecological Models : The Meaning of Validation, Ecological Modelling, 1996, 90, 229–244.
30. Gardner, R. H., and R. H., Urban. Models in Ecosystem Science, Canham C. D., Cole J. J. and Lauenroth W. K. Ed. Princeton University Press, Princeton, New Jersey, 2003, Ch. ”Model validation and testing: past lessons, present concerns, future prospects”, 184–203.
31. Ganeva, D., ‘Semiautomatic Retrieval of Biomass Based on Vegetation Index Optimization and Learning Machine Methods for Winter Rapeseed Crops’, in Proceedings of the Fourteenth International Scientific Conference – SPACE, ECOLOGY, SAFETY, Sofia, Bulgaria, 7–9 November 2018, 299–305.
32. Obsborne, J. W., and A. Overbay, The Power of Outliers (and Why Researchers Should Always Check for Them), Journal of Chemical Information and Modeling, 2013, 53, 9, 1689–99.
33. Grubbs, F. E., G., Beck, Extension of sample sizes and percentage points for significance tests of outlying observations. Technometrics, 1972, 14, 847–854.
34. Richter, K., C. Atzberger, T. B. Hank, and W. Mauser, Derivation of Biophysical Variables from Earth Observation Data : Validation and Statistical Measures, *Journal of Applied Remote Sensing*, 2012, 6, 1, 1–24.
35. Rogerson, P.A. *Statistical method for Geography*. First edit. London: Sage Publications. 2001. 236 p.
36. Zhao, B., J. Zhang, C. Yang, G. Zhou, and Y. Ding, Rapeseed Seedling Stand Counting and Seeding Performance Evaluation at Two Early Growth Stages Based on Unmanned Aerial Vehicle Imagery, Frontiers in Plant Sci., 2018, 9, 1362, 1–17.
37. Jin, X., S. Liu, F. Baret, M. Hemerlé, and A. Comar, Remote Sensing of Environment Estimates of Plant Density of Wheat Crops at Emergence from Very Low Altitude UAV Imagery, Remote Sensing of Environment, 2017, 198, 105–114.
38. Bendig, J., K. Yu, H. Aasen, et al., Combining UAV-Based Plant Height from Crop Surface Models, Visible, and near Infrared Vegetation Indices for Biomass Monitoring in Barley, International Journal of Applied Earth Observation and Geoinformation, 2015, 39, 79–187.
39. Li, W., Z. Niu, H. Chen, et al., Remote Estimation of Canopy Height and Aboveground Biomass of Maize Using High-Resolution Stereo Images from a Low-Cost Unmanned Aerial Vehicle System, Ecological Indicators, 2016, 67, 637–648.
40. Näsi, R., N. Viljanena, J. Kaivosojab, et al., Assessment of Various Remote Sensing Technologies in Biomass and Nitrogen Content Estimation Using an Agricultural Test Field, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3/W3, 2017 Frontiers in Spectral Imaging and 3D Technologies for Geospatial Solutions, 25–27 October 2017, Jyväskylä, Finland, XLII, 2017, 25–27.

41. Hansen, P. M., and J. K. Schjoerring, Reflectance Measurement of Canopy Biomass and Nitrogen Status in Wheat Crops Using Normalized Difference Vegetation Indices and Partial Least Squares Regression, Remote Sensing of Environment, 2003, 86, 4, 542–553.
42. Gitelson, A. A., R. Stark, U. Grits, et al., Vegetation and Soil Lines in Visible Spectral Space: A Concept and Technique for Remote Estimation of Vegetation Fraction', International Journal of Remote Sensing, 2002, 23, 13, 2537–62.
43. Purevdorj, T., R. Tateishi, T. Ishiyama, and Y. Honda, International Journal of Relationships between Percent Vegetation Cover and Vegetation Indices, International Journal of Remote Sensing, 1998, 19, 18, 3519–35.
44. Warren, G., and G. Metternicht, Agricultural Applications of High-Resolution Digital Multispectral Imagery, Photogrammetric Engineering & Remote Sensing, 2005, 71, 5, 595–602.
45. Penuelas, J., and I. Filella, Technical Focus: Visible and Nearinfrared Reflectance Techniques for Diagnosing Plant, Trends in Plant Science, 1998, 3, 4, 151–156.
46. Peng, Y., A. Nguy-Robertson, T. Arkebauer, and A. A. Gitelson, Assessment of Canopy Chlorophyll Content Retrieval in Maize and Soybean : Implications of Hysteresis on the Development of Generic Algorithms, Remote Sens, 2017, 9, 226, 1–18.
47. Baret, F., and T. Fourty. Radiometric estimates of nitrogen status of leaves and canopies. In: Lemaire G, ed. Diagnosis of nitrogen status in crops. Heidelberg, Germany: Springer-Verlag, 1997, 201–227.
48. Tuia, D., J. Verrelst, L. Alonso, F. Pérez-Cruz, and G. Camps-Valls, Multioutput Support Vector Regression for Remote Sensing Biophysical Parameter Estimation, IEEE Geoscience and Remote Sensing Letters, 2011, 8, 4, 804–808.

ДИСТАНЦИОННО ОПРЕДЕЛЯНЕ НА ПАРАМЕТРИ НА ПОСЕВА НА ЗИМНА РАПИЦА ЧРЕЗ СТАТИСТИЧЕСКИ РЕГРЕСИОННИ АЛГОРИТМИ С ПОМОЩТА НА МУЛТИСПЕКТЪРНИ ИЗОБРАЖЕНИЯ ОТ SENTINEL-2

Д. Ганева, Е. Руменина

Резюме

Определянето на параметри на посева на земеделски култури е важна задача за дистанционното наблюдение на селското стопанство и за изготвяне на стратегии за управление на полетата. Основната цел на това изследване е да се оцени извличането на параметри на посева на зимна рапица за два периода от растежа: преди презимуване и след презимуване, чрез параметрични и непараметрични статистически модели и Sentinel-2 изображения, в реални условия на земеделие в Североизточна България. За калибрирането на моделите се използват in-situ данни от три полеви кампании. За повечето от

изследваните параметри бяха идентифицирани модели с добра точност, с изключение за надземната свежа биомаса. Най-добрите модели за общо площно покритие (RMSEcv = 0.14 %) и гъстота на посева (RMSEcv = 9 nb/m^2) са параметрични модели с три-канален вегетационен индекс (3BSI-Tian) и линейна функция за първия, 3-канален вегетационен индекс (3BSI-Verrelst) и полиномиялна за втория параметър. За надземна суха биомаса (RMSEcv = 52 g/m^2), средна височина на растенията (RMSEcv = 4 cm) и съдържание на азот в свежа биомаса (RMSEcv = 2 g/kg) най-добрите модели са непараметрични, *Gaussian Processes Regression* за първия параметър и *Variational Hetero-scedastic* вариант на *Gaussian Processes Regression* за другите две.