Mapping and Biomass Estimation of Menthol Mint (*Mentha arvensis L.*) crop using Sentinel 2 data

KK Choudhary*[a], CS Murthy[a], B Laxman[a], JJ Varghese[b], PT Lepcha[b] and P Chakraborty[b]

[a]Remote Sensing Applications Area, National Remote Sensing Centre,
Indian Space Research Organisation, Balanagar, Hyderabad - 500037
Emails: *karunkumar_choudhary@nrsc.gov.in; murthy_cs@nrsc.gov.in, laxman_boggarapu@nrsc.gov.in
[b]Regional Office, Spices Board, Barabanki, UP- 225001
Emails: john.jovarghese@nic.in; phurba.lepcha@nic.in; pritamchakraborty.sb@gov.in

KEY WORDS: Menthol mint; Object oriented; Narrow Band vegetation indices, Multivariate analysis; Red Edge

ABSTRACT

Spatial statistics on Area and production of high value crop like Menthol Mint is much needed for various stakeholders since India is the largest producer of Mentha oil (73% of world production) and the crop is an export earner for the country. The present study has two fold objectives namely – mapping menthol crops and assessing its above ground fresh biomass which is the economical part of the crop. Barabanki district of Uttar Pradesh state, a prominent mint growing district is the study area. Object oriented classification approach was applied to overcome the limitations of pixels based approach. Multi-resolution segmentation was performed on Single scene of Sentinel 2 data for the period corresponding to the maximum vegetative stages with suitable scale, shape and compactness factors. Segmentations were assigned classes based on the threshold value of Normalised difference vegetation Indices (NDVI) and Grey Level Concurrence Matrix (GLCM) homogeneity derived from Ground truth data for the study region. The object based accuracy assessment was made for the classes. Sentinel-2 multispectral data were utilized in terms of Narrowband Vegetation Indices (NVI) and angle or area based vegetation index (AVI) to estimate the fresh Biomass of the crop as these do not have saturation problem like broadband Indices. To overcome the multicollinearity among the Vegetation indices the multivariate analysis were performed for Biomass prediction using multiple linear regression (MLR), Partial Least Square regression (PLSR); Principal component regression (PCR) and support vector regression (SVR). Further, twenty one NVI and six AVI were computed using visible, Red-edge and Near Infra Red bands of Sentinel-2 data. Univariate and MultiVariate prediction of Biomass were made. Seventyfive percent of field measured biomass data was used for developing the model and rest is used for validating the model. Result of univariate model validation reveals the highest coefficient of determination (R^2=0.71) for narrowband Green-NDVI (GNDVI-2) followed by Narrow NDVI with R^2 = 0.67. The best model performance in Multivariate analysis was obtained using PLSR with VI with R^2= 0.72 and RMSE= 3.2 t/ha. The study demonstrates the importance of Sentinel-2 VNIR bands for biomass estimation of Menthol Mint crop.

1. INTRODUCTION

Growing of High value crops are highly encouraged among the farmers for their maximum profitability. Mentha arvensis (commonly known as menthol mint, corn mint or Japanese mint) is one of the high value crop having widely applications in food, flavourings, pharmaceutical and cosmetic industries, mainly grown in *zaid* season in the Indo-Gangetic plains of India (Singh et al., 2007). The major districts in Uttar Pradesh where this crop is being cultivated are Badaun, Bareilly, Sahajanpur, Pilibhit, Lakhimpur Khiri, Barabanki and Ambedker Nagar. Although statistics on area and production of Menthol mint is available at district level but there is dearth of information on the spatial area and production estimate. Satellite remote sensing because of its synoptic coverage and multispectral in nature have the potential to map crop type as well as estimating crop vigour at different scales (Howard and Wylie, 2014).

Above ground green biomass is the economic part of the Menthol Mint and traditional methods for mapping green biomass involve direct measurements, which are time-consuming, expensive and require extensive field work. Furthermore, reliable estimates are restricted to local scales only, whereas sutiability analysis and policy makers require estimates at the landscape scale. One of the major sources of information for the study of landscapes and for estimating biomass over large areas is remote sensing (Kumar *et al*. 2001, Wulder *et al*. 2004). Numerous studies has shown significant relationships between satellite derived vegetation Indices (VI) and above ground biomass (Anaya et al., 2009; Ren et al., 2011; Yan et al., 2013 and Jin et al., 2014). Normalized reflectance ratio indices between red and NIR i.e., NDVI is being widely used as univariate proxy of the green vegetation quantity (Gitelson et al, 2014; Ogutu et al., 2014). However NDVI tends to lose its sensitivity at higher biomass conditions because of its saturation. Using several VI together and Narrowband VI may solve the problem of saturation. With the recent advances in multispectral satellite like Worldview-2, Geoeye, Rapideye, Sentinel-2 etc various ratio

vegetation indices can be generated and can be applied in combination for estimation of aboveground biomass. In recent years, the application VI based on red edge bands has increased for crop biophysical parameter estimation.
European Space Agency (ESA) twin-satellite mission Sentinel-2 is characterized by its high spatial resolution (10 m), wider swath (300 KM) and frequent revisiting capability which make it very useful for monitoring and inventorying purposes. Sentinel-2 bands details are provided in Table 1.

Table 1: Details of Sentinel-2 bands

Band No	B1	B2	B3	B4	B5	B6	B7	B8	B8a	B9	B10	B11	B12
Region	Coastal Blue	Blue	Green	Red	Red Edge 1	Red Edge 2	Red Edge 3	NIR 1	NIR 2	NIR 3	SWIR1	SWIR 2	SWIR3
Center Wavelength (nm)	443	490	560	665	705	740	783	842	865	945	1380	1610	2190
Band width (nm)	20	65	35	30	15	15	20	115	20	20	30	90	180
Spatial resolution (m)	60	10	10	10	20	20	20	10	20	60	60	20	20

The availability of high resolution satellite data has increased the importance of Object based mapping. The image object provide a more suitable scales to map crops at multiple scales which is the base of object base classification and its more acceptable process than individual pixel base approach (Gamanya et al., 2007).

2. OBJECTIVE

In the light of the above background the following objectives have been proposed for the study:

- Mapping of Menthol mint applying spectral and texture information from sentinel-2 bands using Object Oriented classification method.
- Estimation of above Ground fresh Biomass of Menthol mint using sentinel-2 derived vegetation indices.

3. STUDY AREA

Figure 1 shows the location of study area wich is the part of Barabanki district, Uttar Pradesh, India. The study area is a part of Indo-gangetic plains with sub-humid climate and sandy loam soils. The area is well known for its Menthol mint cultivation in the zaid season (March to June) which is a lean season for other field crops. The study has been done for zaid 2016 on total geographical area of around 1.43 lakh ha. The 85% of the area is Irrigated using tubewell and canals.

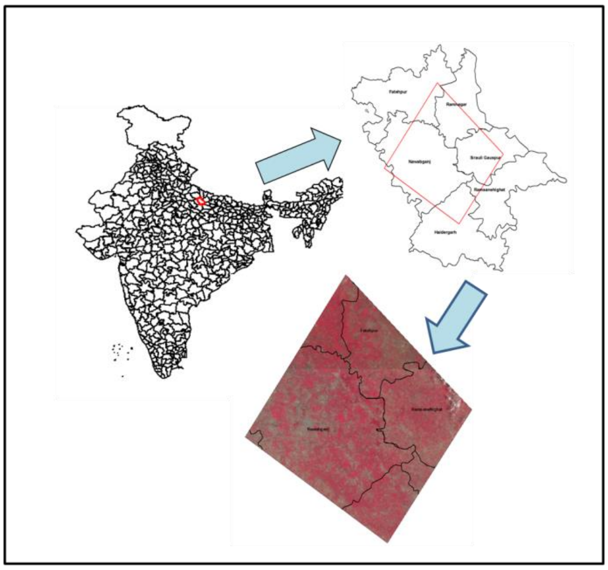

Fig 1: Study Area- Parts of Barabanki District, Uttar Pradesh

4. DATA USED

4.1 SATELLITE DATA

Sentinel-2 data top of atmosphere reflectance data corresponding to the maximum vegetative stage in the study region was downloaded. The cloud free area were subset and used for further analysis. The details of the sentinel-2 scenes used in the study are shown in table 2.

Table 2: Sentinel-2 satellite data used in the study

Scene Details	Date of Pass	Selected Bands	Spatial Resolution
S2A_OPER_PRD_MSIL1C_PDMC (level 1C)	2016/05/29	Bands 2, 3, 4, 5, , 7, 8 and 8a	10m, 20m

4.2 ANCILLARY DATA/SOFTWARE

The following data/software have been used in the study:

a. Vector layer of study area
b. Smart mobile for ground truth data collection
c. eCognition – for object oriented classification
d. ERDAS Imagine - for Vegetation Indices generation
e. ArcGIS – for extracting vector-based statistics
f. Unscrambler X (Evaluation version): For Multivariate regression analysis
g. Statistics on mango and oil palm plantations/area from State Horticultural Department

5. METHODOLOGY

The methodology flowchart is presented in Figure 2. The study had two major components namely classification of Menthol Mint using Object Oriented approach and estimation of Menthol mint fresh biomass using combination of Vegetation Indices. Multiresolution Image segmentation were performed on 4-band stack image (B2, B3, B4 and B8) with weightages of 1, 1, 2 and 3 respectively with a scale parameter of 20. The shape and compactness were 0.8, 0.5 respectively.

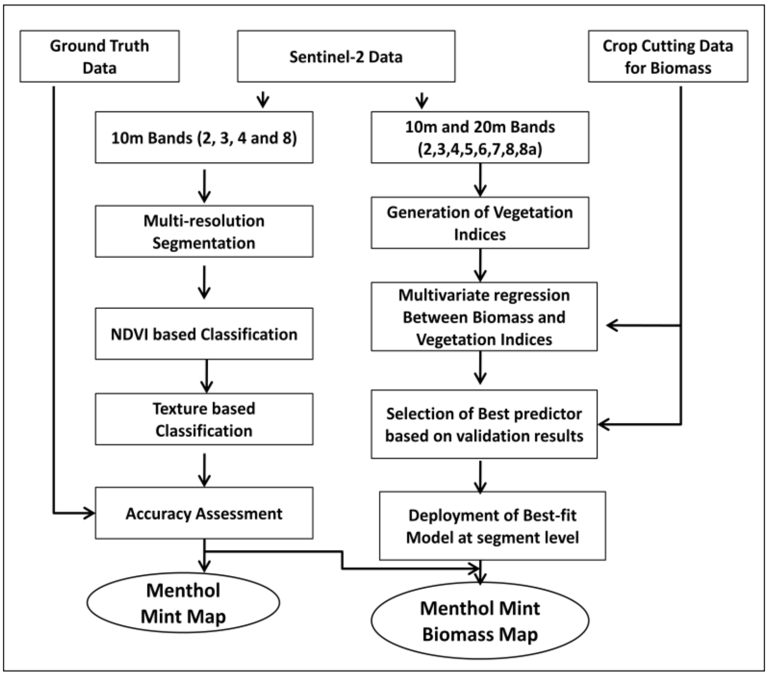

Fig 2: Outline of the methodology adopted in the study

For Biomass estimation 21 narrowband Vegetation Indices (NVI) and Six Angle/Area based Vegetation Indices (AVI) were computed from the selected bands. The lists of all the VI used in the study are provided in the Table 3. Two approach has been followed in the multivariate estimation of Menthol mint fresh biomass firstly only NVI were used and secondly NVI were used in combination with AVI. Figure 3 shows the diagrammatical representation of six AVI which is proposed in this study.

Table 3: Vegetation Indices used in the study

Sr. No	Vegetation Indices	Formula	Sr. No	Vegetation Indices	Formula
1	NDVI1	$\frac{B8-B4}{B8+B4}$	15	MTVI2	$\frac{1.5*(1.2*(B7-B3)-2.5*(B4-B3))}{\sqrt{\sqrt{(2*B7+1)^2-0.5-(6*B7-5*\sqrt{B4})}}}$
2	NDVI2	$\frac{B8a-B4}{B8a+B4}$	16	TCARI	$3*((B5-B4)-0.2*(B5-B3)*\frac{B5}{B4})$
3	NDRE 1a	$\frac{B8-B5}{B8+B5}$	17	GNDVI1	$\frac{B8-B3}{B8+B3}$
4	NDRE 1b	$\frac{B8-B6}{B8+B6}$	18	GNDVI2	$\frac{B8a-B3}{B8a+B3}$
5	NDRE 1c	$\frac{B8-B7}{B8+B7}$	19	IRECI	$\frac{B7-B4}{\frac{B5}{B6}}$
6	NDRE 2a	$\frac{B8a-B5}{B8a+B5}$	20	OSAVI	$\frac{B8-B4}{B8+B4+0.16}$
7	NDRE 2b	$\frac{B8a-B6}{B8a+B6}$	21	TCARI/OSAVI	Ratio of TCARI and OSAVI
8	NDRE 2c	$\frac{B8a-B7}{B8a+B7}$	22	Area1	Area of triangle formed by Band 5, 6 and 7
9	MCARI1	$\frac{((B5-B4)-0.2*(B5}{(B8+B6)*B}$	23	Area2	Area of triangle formed by Band 6, 7 and 8
10	MCARI2	$\frac{1.5*(2.5*(B7-B4)}{\sqrt{\sqrt{(2*B7+1)^2-0.5}}}$	24	Area3	Area of triangle formed by Band 7, 8 and 8a
11	RE-NDVI	$\frac{B6-B5}{B6+B5}$	25	A (B6)	Angle at Band 6
12	Mod-RE-NDVI	$\frac{B6-B5}{B6+B5-2*B1}$	26	A (B7)	Angle at Band 7
13	Mod RE Simple Ratio	$\frac{B6-B1}{B6+B1}$	27	A (B8)	Angle at Band 8
14	MTVI1	$1.2*(1.2*(B7-B3)-2.5*(B4-B3))$			

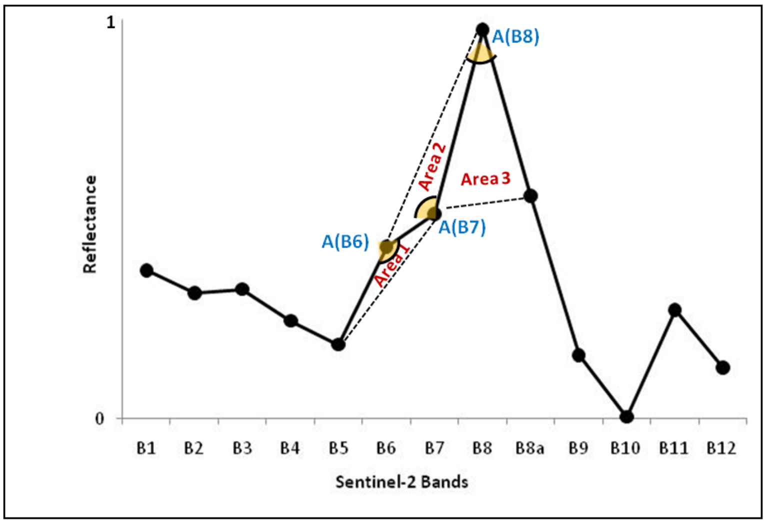

Fig 3: Angle and area based Vegetation Indices from Sentinel-2 data

6. COLLECTION OF CROP CUTTING DATA

Crop cutting experiments (CCE) were performed in the study region during 30th may to 05th Jun 2016. A total of 60 locations were manually harvested in a 5m X 5m area and weighing done with minimum time gap. The yield data were converted to per hectare basis. CCE data were recorded through BHUVAN CCE mobile app. The CCE sites were chosen such that to account maximum variability in the Menthol biomass. The above ground fresh biomass range were 10 to 40 t per ha in the conducted CCE. Figure 4 shows the locations of CCE sites and the spectral signature of different biomass class on sentinel-2 data.

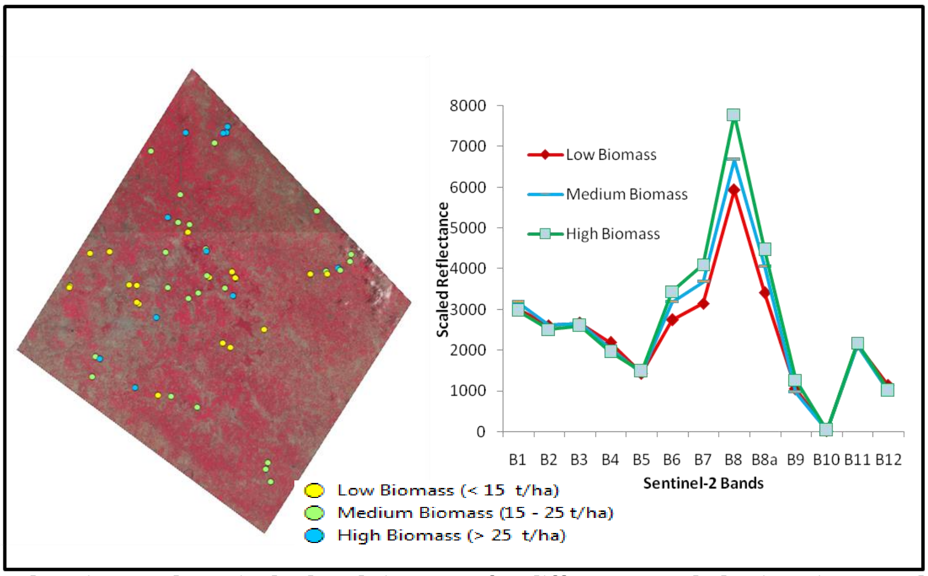

Fig 4: CCE locations and Sentinel-2 band signature for different Menthol Mint Biomass Classes

7. RESULTS:

7.1 SEGMENTATION AND CLASSIFICATION

Figure 5 shows different steps of image segmentation and classification. After multi-resolution segmentation were a two step method were adopted for Classification of Menthol Mint. First of all NDVI range threshold were applied for segregating Menthol mint segments and in the second step the Grey Level Co-occurrence Matrix (GLCM) homogeneity threshold were provided. The range of NDVI and GLCM homogeneity were computed from around 200 field data point collected in the study region. After classification the stray segments were cleared using geometry parameter i.e., Border Index. Finally the segments were smoothened using polynomial function. The accuracy of the menthol mint map were accessed using GT points. For the study region the mapping accuracy was found to be more than 90% for Menthol mint.

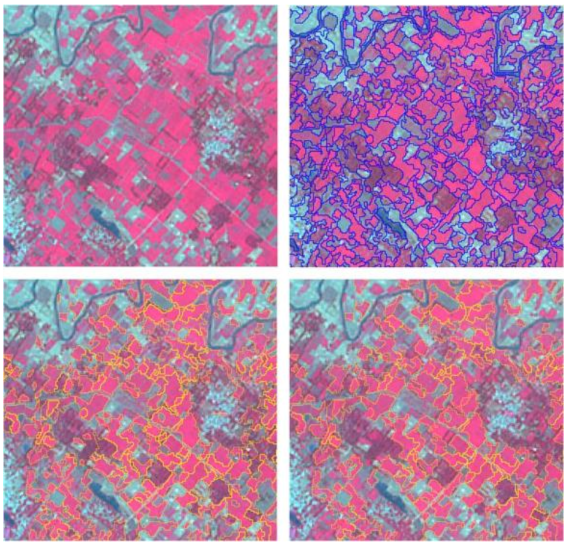

Fig 5: Multi-resolution segmentation and classification of Menthol Mint on Sentinel-2 satellite data

7.2 CORRELATION ANALYSIS

Table 4 represts the univariate Correlation coefficients matrix between Sentinel-2 Band reflectance and Biomass. Band 8a (Narrow NIR) was found to be highest correlated with the Green Biomass of Menthol Mint followed by Band 7 (Red Edge 3). Also the bands were significantly correlated with each other hence band were transformed into several Vegetation Indices and again correlation marix were generated for VI and Biomass.

Table 4: Correlation matrix of Sentinel 2 Bands and Above ground green Biomass of Menthol Mint

	Biomass	*Band 1*	*Band 2*	*Band3*	*Band4*	*Band5*	*Band6*	*Band7*	*Band8*	*Band8a*	*Band9*	*Band10*	*Band11*	*Band12*
Biomass	1.00													
Band 1	0.09	1.00												
Band 2	-0.40	0.06	1.00											
Band3	-0.39	0.07	0.99	1.00										
Band4	-0.41	0.20	0.96	0.94	1.00									
Band5	0.21	0.56	0.12	0.13	0.29	1.00								
Band6	0.50	0.01	-0.53	-0.46	-0.56	0.01	1.00							
Band7	0.54	-0.06	-0.55	-0.48	-0.59	0.01	0.98	1.00						
Band8	-0.12	-0.05	0.83	0.86	0.70	0.10	-0.14	-0.15	1.00					
Band8a	0.55	-0.03	-0.51	-0.44	-0.55	0.03	0.98	0.99	-0.12	1.00				
Band9	0.04	-0.19	0.02	0.03	0.02	0.14	0.22	0.23	0.14	0.20	1.00			
Band10	0.00	0.18	0.10	0.04	0.17	0.18	-0.15	-0.17	-0.04	-0.14	0.04	1.00		
Band11	0.33	0.37	-0.04	-0.01	0.08	0.80	0.30	0.32	0.12	0.32	0.19	-0.07	1.00	
Band12	-0.02	0.37	0.35	0.32	0.49	0.77	-0.46	-0.44	0.22	-0.44	0.02	0.09	0.67	1.00

Among all 27 VI the heighest correlation was found for GNDVI1 followed by GNDVI2 and NDVI 2 (Table 5). Significant correlations were found between different VI and hence Multivariate analysis need to be performed to remove multi-collinearity among VIs.

Table 5: Correlation coefficients between menthol Mint Biomass and Vegetation Indices used in study

Vegetation Indices	**Correlation with Biomass**	**Vegetation Indices**	**Correlation with Biomass**
NDVI1	0.66	MTVI2	0.24
NDVI2	0.67	TCARI	0.61
NDRE1a	0.63	GNDVI1	0.71
NDRE1b	0.37	GNDVI2	0.70
NDRE1c	0.28	IRECI	0.66
NDRE2a	0.57	OSAVI	0.66
NDRE2b	0.50	TCARI/OSAVI	0.57
NDRE2c	0.18	A(B6)	0.47
MCARI1	0.59	A(B7)	0.48
MCARI2	0.59	A(B8)	0.61
RE-NDVI	0.54	area1	0.56
Mod-RE-NDVI	0.54	area2	0.53
Mod RE SR	0.53	area3	0.59
MTVI1	0.64		

7.3 BIOMASS ESTIMATION

Four Multivariate regression methods i.e., Multiple Linear Regression (MLR), Partial Least Square Regression (PLSR), Pricipal component regression (PCR) and Support vector Regression (SVR) were tested in two conditions (i) With all NVI and (ii) With all NVI and AVI, for the data sets. The 75 % of the data were used for calibration of

the regression models and rest 255 were used for validation. The Table 6 shows the validation summary of the multivariate analysis.
The inclusion of AVI into the regression has improved the coefficient of determination and RSME in all the four methods. PLSR has performed better than other methods in terms of least RMSE i.e., 3.2 t ha^{-1}followed by SVR with RMSE of 3.8 t ha^{-1}. Detail scatterplots can be seen in Figure 6.

Table 6: Coefficient of determination and root mean square error for different Multivariate regressions

	All Narrowband VIs		All Narrowband and Angle/Area based VIs	
	R^2	**RMSE**	**R^2**	**RMSE**
Multiple Linear Regression (MLR)	0.81	6.4	0.85	5.2
Partial Least Square Regression (PLSR)	0.65	3.6	0.72	3.2
Pricipal component regression (PCR)	0.51	6.2	0.58	4.7
Support vector Regression (SVR)	0.56	4.5	0.61	3.8

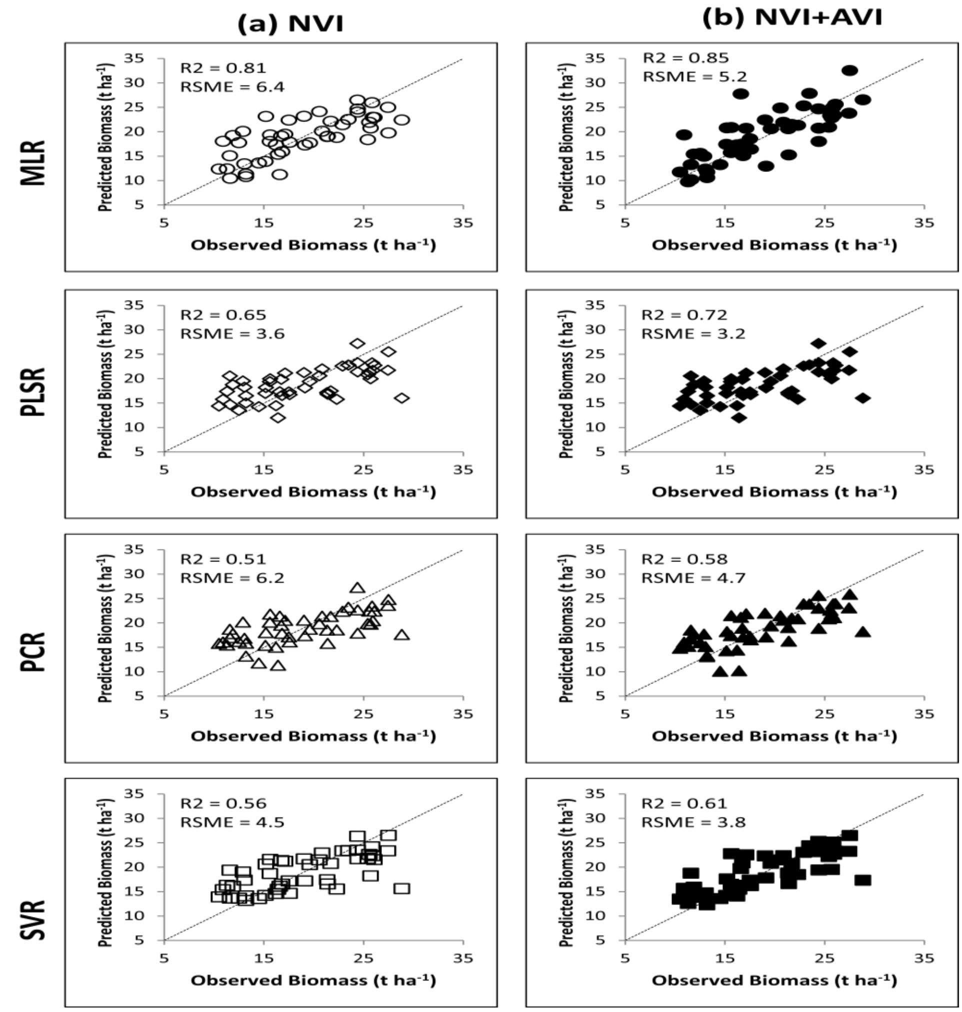

Fig 6: Results of different Multivariate regressions performed using (a) Only Narrowband VIs and (b) narrowband VI and Area/Angle based VIs

8. CONCLUSIONS

Menthol Mint can be mapped semi automatically using Multi-resolution segmentation and spectral and texture based classification performed on Single scene of Sentinel 2 data for the period corresponding to the maximum vegetative stages with suitable scale, shape and compactness factors. Sentinel-2 multispectral data were utilized in terms of Narrowband Vegetation Indices (NVI) and angle or area based vegetation index (AVI) to estimate the fresh Biomass of the crop as these do not have saturation problem like broadband Indices. Result of univariate model validation reveals the highest coefficient of determination (R^2=0.71) for narrowband Green-NDVI (GNDVI-2) followed by Narrow NDVI with R^2 = 0.67. The best model performance in Multivariate analysis was obtained using PLSR with VI with R^2= 0.72 and RMSE= 3.2 t/ha. The study demonstrates the importance of Sentinel-2 VNIR bands for biomass estimation of Menthol Mint crop.

9. ACKNOWLEDGEMENT

The authors would like to thank Director, NRSC and Deputy Director-Application Area, NRSC for their constant support and encouragement to this study. We would also like to place on record our appreciation for the Farmers of Barabanki district, UP for their support in carrying out crop cutting experiments at their fields and providing informations about the Menthol mint crop.

10. REFERENCES

Singh, A.K. and Khanuja, S.P.S. 2007. CIMAP initiatives for menthol mint. Spice India, (December): 14-17.

DM Howard and BK Wylie. 2014. Annual crop type classificatio of the US great plains for 2000 to 2011, photogramm. Eng. Remote Sens., 80(6), 537-549.

Anaya, J.A.; Chuvieco, E.; Palacios-Orueta, A. 2009. Aboveground biomass assessment in Colombia: A remote sensing approach. *For. Ecol. Manag.*, *257*, 1237–1246.

Ren, H.R.; Zhou, G.S.; Zhang, X.S. 2011. Estimation of green aboveground biomass of desert steppe in Inner Mongolia based on red-edge reflectance curve area method. *Biosyst. Eng.*, *109*, 385–395.

Yan, F.; Wu, B.; Wang, Y.J. 2013. Estimating aboveground biomass in Mu Us Sandy Land using Landsat spectral derived vegetation indices over the past 30 years. *J. Arid Land*, *5*, 521–530.

Jin Y, Yang X, Qiu J, Li J, Gao T, Wu Q, Zhao F, Ma H, Yu H and Xu, B. 2014. Remote Sensing-Based Biomass Estimation and Its Spatio-Temporal Variations in Temperate Grassland, Northern China. Remote Sens. 6, 1496-1513.

Gitelson , A. A., Y. Peng, and K. F. Huemmrich. 2014. "Relationship between fraction of radiation absorbed by photosynthesizing maize and soybean canopies and NDVI from remotely sensed data taken at close range and from MODIS 250 m resolution data," Remote Sens. Environ., vol. 147, pp. 108–120.

Ogutu ,B., J. Dash, and T. P. Dawson. 2014. "Evaluation of the influence of two operational fraction of absorbed photosynthetically active radiation (FAPAR) products on terrestrial ecosystem productivity modelling," Int.J. Remote Sens., vol. 35, pp. 321–340, 2014.

Gamanya, R., P.D. Maeyer, and M.D. Dapper, 2007. An automated satellite image classification design using object oriented segmentation algorithms: A move towards standardization, *Expert Systems with Applications*, 32:616–624