Retraction Notice

Title of retracted article: **A Participatory Iterative Mapping Approach and Evaluation of Three Machine Learning Algorithms for Accurate Mapping of Cropping Patterns in a Complex Agro-Ecosystems**

Author(s): Gudina Legese Feyisa1*, Leo Palao1, Andy Nelson1, Khin Thawda Win2, Kin Nge Htar2, Murali Krishna Gumma3, David E. Johnson1

* Corresponding author. Email: fgudina@gmail.com

Journal: Advances in Remote Sensing (ARS)
Year: 2016
Volume: 5
Number: 1
Pages (from - to): 1 - 17
DOI (to PDF): http://dx.doi.org/10.4236/ars.2016.51001
Paper ID at SCIRP: 2630157
Article page: http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=62970

Retraction date: 2016-02-15

Retraction initiative (multiple responses allowed; mark with **X**):
- **X** All authors
- ☐ Some of the authors:
- ☐ Editor with hints from
  - ○ Journal owner (publisher)
  - ○ Institution:
  - ○ Reader:
  - ○ Other:

Date initiative is launched: 2016-02-15

Retraction type (multiple responses allowed):
- ☐ Unreliable findings
  - ○ Lab error ○ Inconsistent data ○ Analytical error ○ Biased interpretation
  - ○ Other:
- ☐ Irreproducible results
- ☐ Failure to disclose a major competing interest likely to influence interpretations or recommendations
- ☐ Unethical research

- ☐ Fraud
  - ○ Data fabrication ○ Fake publication ○ Other:
- ☐ Plagiarism ☐ Self plagiarism ☐ Overlap ☐ Redundant publication *
- ☐ Copyright infringement ☐ Other legal concern:

- ☐ Editorial reasons
  - ○ Handling error ○ Unreliable review(s) ○ Decision error ○ Other:

- **X** Other:

Results of publication (only one response allowed):
- ☐ are still valid.
- **X** were found to be overall invalid.

Author's conduct (only one response allowed):
- ☐ honest error
- ☐ academic misconduct
- **X** none (not applicable in this case – e.g. in case of editorial reasons)

* Also called duplicate or repetitive publication. Definition: "Publishing or attempting to publish substantially the same work more than once."

History
Expression of Concern:
☐ yes, date: yyyy-mm-dd
X no

Correction:
☐ yes, date: yyyy-mm-dd
X no

Comment:

Free style text with summary of information from above and more details that can not be expressed by ticking boxes.

This article has been retracted according to COPE's Retraction Guidelines. Since authors have their personal reasons, they have to withdraw this paper from journal *Advances in Remote Sensing*.

Advances in Remote Sensing, 2016, 5, 1-17
Published Online March 2016 in SciRes. http://www.scirp.org/journal/ars
http://dx.doi.org/10.4236/ars.2016.51001

Scientific Research Publishing

A Participatory Iterative Mapping Approach and Evaluation of Three Machine Learning Algorithms for Accurate Mapping of Cropping Patterns in a Complex Agro-Ecosystems

RETRACTED

Gudina Legese Feyisa[1*], Leo Palao[1], Andy Nelson[1], Khin Thawda Win[2], Kin Nge Htar[2], Murali Krishna Gumma[3], David E. Johnson[1]

[1]International Rice Research Institute (IRRI), Metro Manila, Philippines
[2]International Rice Research Institute (IRRI), Myanmar Office, Nay Pyi Taw, Myanmar
[3]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India
Email: *fgudina@gmail.com

Received 6 December 2015; accepted 22 January 2016; published 25 January 2016

Abstract

We applied time-series analysis of vegetation indices (VIs) (NDVI and EVI) derived from the Moderate Resolution Imaging Spectrometer (MODIS) sensors to detect seasonal patterns of irrigated and rainfed cropping patterns in five townships in the central dry zone of Myanmar, an important agricultural region of the country, that is both poorly mapped for cropping practices and which faces environmental and climate related challenges to agriculture. To improve mapping accuracy of cropping pattern, we implemented a participatory iterative ground truthing and mapping approach and explored the efficiency of three state-of-the-art machine learning algorithms: Support Vector Machine (SVM), Random Forest (RF), and a classification tree and rule-based model (C5.0). We first collected reference data at random locations and run a preliminary supervised classification using the SVM algorithm. Based on the preliminary classification outputs, we invited township agricultural officers to assess the accuracies of the maps based on local knowledge and secondary statistical data and hence identify areas with high land cover heterogeneity, which enabled us to allocate more sample sizes in such areas. We compared accuracies achieved by use of increasing size of predictor layers of VIs (8-days, 16-days and monthly composite stacks of 1 to 3 years). Results show the combined effects of i) an iterative participatory approach to field data collection and map classification, ii) identification of superior algorithm and iii) appropriate size and type of predictor VIs, we were able to substantially improve mapping accuracy; depending on the models

*Corresponding author.

How to cite this paper: Feyisa, G.L., *et al.* (2016) A Participatory Iterative Mapping Approach and Evaluation of Three Machine Learning Algorithms for Accurate Mapping of Cropping Patterns in a Complex Agro-Ecosystems. *Advances in Remote Sensing*, **5**, 1-17. http://dx.doi.org/10.4236/ars.2016.51001

used, accuracy improvement ranged from 31% to 43%. Among the three algorithms we compared, SVM with Gaussian Radial Basis kernel function was superior in terms of all accuracy measures. Particularly, the accuracy difference was statistically significant ($p < 0.005$) when larger numbers of VIs layers were used and significance of difference diminishes with the increasing size of training data. Accuracy achieved by use of NDVI was consistently better than that of EVI. Though the difference is not significant, 8-days NDVI composite resulted in better accuracy than 16-days composite. Maximum overall accuracy of 94% was achieved using SVM and 8-days NDVI composites of three years. In conclusion, our findings highlight the value of participatory field validation approaches to improve classification accuracy especially in areas where land use patterns are temporally and spatially complex. We also show that choice of classifiers and size of predictor variables are essential and complementary to the participatory field approach in achieving desired accuracy of cropping patter mapping in areas where other sources of spatial information are scarce.

Keywords

Accuracy, Agriculture, Classification, MODIS, NDVI, Rice, Time-Series

1. Introduction

Timely detection and monitoring of spatiotemporal dynamics in agricultural land use are needed for various decision making processes and research works. Remote sensing technologies are increasingly being used in providing a timely and accurate picture of the agricultural sectors [1]. Due to its inherent nature, agricultural sector change over space and time depending of a number of factors including the changes in physical environment that agricultural plants heavily depend on (e.g. rainfall, temperature, soil). Food and Agriculture Organization of the United Nations (FAO) [2] stresses the importance of timeliness and accuracy of agricultural information as a major factor underlying agricultural statistics and associated monitoring systems (FAO, 2011).

Satellite data, such as imagery from Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat and SPOT satellite sensors have become important sources of information in agricultural sciences. Image classification is among the most widely applied image analysis in extracting useful spatial and temporal information using remotely sensed data [3]. The potentials of using remotely sensed data for crop identification and mapping cropping patterns over large areas has been widely explored [4]-[6]. The basis for remote sensing of cropping patterns is mainly the unique temporal patterns of crop phonology which could be characterized through analysis of spectral patterns of vegetation indices such as the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) [4] [6]-[8].

MODIS data is highly suitable for detection of seasonal agricultural crop phonology due to its high temporal resolution [5] [9] [10]. However, in highly heterogeneous agricultural landscapes, application of MODIS in identifying crop types and cropping patterns is a challenging task, and the usability of information extracted from MODIS is constrained by low accuracies. This is mainly due to spectral confusions among various land cover types associated with coarse spatial resolution of MODIS data (~250 m pixel size) [7] [11]. In particular, heterogeneity of cropping patterns is common in tropical regions of developing countries where small holder farming is dominated by diverse crop types [12]. The Central Dry Zone of Myanmar (CDZM), where we have undertaken this study, is a typical example of diverse cropping system. The area we studied is characterized by small land holding (often plots are less than half hectare) and cropping patterns is highly heterogeneous with a mixture of irrigated and rainfed systems where two or three crops are grown a year with varying crop rotations, hence accurate detection and identifying cropping patterns using MODIS data is a challenging task.

Several techniques of improving mapping accuracy have been proposed in the literature, including approaches to improve spatial resolution of MODIS through integration of data from different sensors [13] [14], and techniques of combining various classification algorithms, and techniques of applying object-based and pixel-based classification [15] [16]. Though a large number of comparative studies attempted to compare and recommend most suitable classification methods, the findings are often inconsistent [17] [18]. Studies such as Qian, Zhou [19], Liu, Wang [20], Shao and Lunetta [16] and Fernández-Delgado, Cernadas [17] have generally indicated that machine learning algorithms such as Random Forest (RF) and Support Vector Machine (SVM) are superior in accuracy. However, suitability of the classification algorithms depended on several factors including sizes and purity of training sample [16] [18], parameter tuning [18] [21], whether the classification is pixel-based or object-based [19] [21] [22], and performance criteria used in assessing the accuracy performances of the classifiers [23].

During a study of cropping patterns in the CDZM using field based ground truth data and class labeling of unsupervised classification of multitemporal MODIS VIs, we encountered problems of accuracy and the challenges of obtaining representative ground truth points for heterogeneous systems in these areas. This was compounded by the lack of spatially and temporal explicit agricultural information in the CDZM. The objectives in this study were to develop an accurate cropping pattern classification approach for complex agricultural landscapes using MODIS time-series data and appropriate sources of field and validation informan. We undertook a participatory field data collection and field campaign mapping approach that combined assessment of the performance of three different sets of multitemporal VI data classified through three machine learning algorithms (RF, SVM and C5.0) using the "caret" package implemented in R language. The underlying theory and detailed mathematical explanation of these algorithms have been documented in the literature [24]-[28]. Our approach combined the analyses of the remote sensing data with detailed knowledge possessed by the farmers and local agricultural officers. The local knowledge held by these individuals can provide an overview, and able to indicate cropping patterns that occur in one area as distinct from another while, at the same time, able to describe crops and practices at particular locations and how this might change over time. The cropping patterns of the CDZM are highly variable and governed by soil type, topography and the presence of irrigation. Fields are often small and choice of crop may be determined by relatively small changes in the toposequence. This is an area that has received relatively little research attention in recent decades and it is poorly described in the contemporary press.

2. Study Area

The CDZM covers 55,000 km^2 and comprises Magway, Mandelay and lower Sagaing regions. The area is transected north to south by the Chindwin and Ayeyarwady rivers, and the topography is gently undulating and about 300 masl. Most streams are dry for much of the year. The CDZM is tends to be resource poor, water is scarce and crop losses to drought are frequent, vegetation cover is thin and there is severe soil erosion. In some areas, in addition to drought, soil alkalinity is a constraint. There is a dearth of detailed current information on the cropping systems and agricultural potential of the CDZM. This presents a major challenge to the development of technologies which are likely to be well adapted to particular locations (e.g. upland or alkaline soil conditions) and certain cropping systems (e.g. short duration rice varieties for the rainfed monsoon crop to followed by legume). Greater information on the current cropping patterns in the region is required in order to support adaptive research and targeting of improved options for the farmers.

The study was undertaken in five townships of the CDZM; Ye'u, Monywa, Pakokku, Wundwin and Pyawbe. The study sites lie between 20.39 to 23.04 N latitude and 94.67 to 96.34 E longitude (**Figure 1(a)**) and the area is characterized by low rainfall, receiving an average annual rainfall of 833 mm across townships. In the CDZM, the monsoon is divided into; the pre-monsoon season which starts in May, and the major monsoon season which starts in August and ends in October (**Figure 1(b)**). Average minimum and maximum temperatures are 21˚C and 31.9˚C, respectively (source: local agro-meteorological stations of each Township). The altitude ranges between 62 to 658 meters above sea level.

The CDZM is characterized by small holder crop production systems. In many places of the studied Townships, irrigation facilities offer the possibility of producing a number of crop through the year. Double and even triple cropping systems are common in many villages of the studied areas, which could benefit from full irrigation or supplemental irrigation. In the rainfed upland areas areas two crops per year may be grown in some or areas comprising sunflower (*Helianthus annuus*), seasame (*Sesamum indicum*), groundnut (*Arachis hypogaea*), green gram (*Vigna radiata*), maize (*Zea mais*) or butter bean (*Phaseolus lunatus*) or a single crop of pigeon pea (*Cajanus cajan*), cotton (*Gossypium* sp.) or sugar cane (*Saccharum officinarum*). In the lowland rainfed areas monsoon rice (*Oryza sativa*) may be followed by either chickpea (*Cicer arietinum*), green gram or other pulse. In irrigated areas two crop of rice may be grown as well as a crop of chickpea, sunflower or other. Double or triple cropping in purely rainfed system is also possible due to local topography and rainfall resulting in sufficient soil moisture to support low water use crops such as sorghum (*Sorghum* bicolor), groundnuts and chickpea.

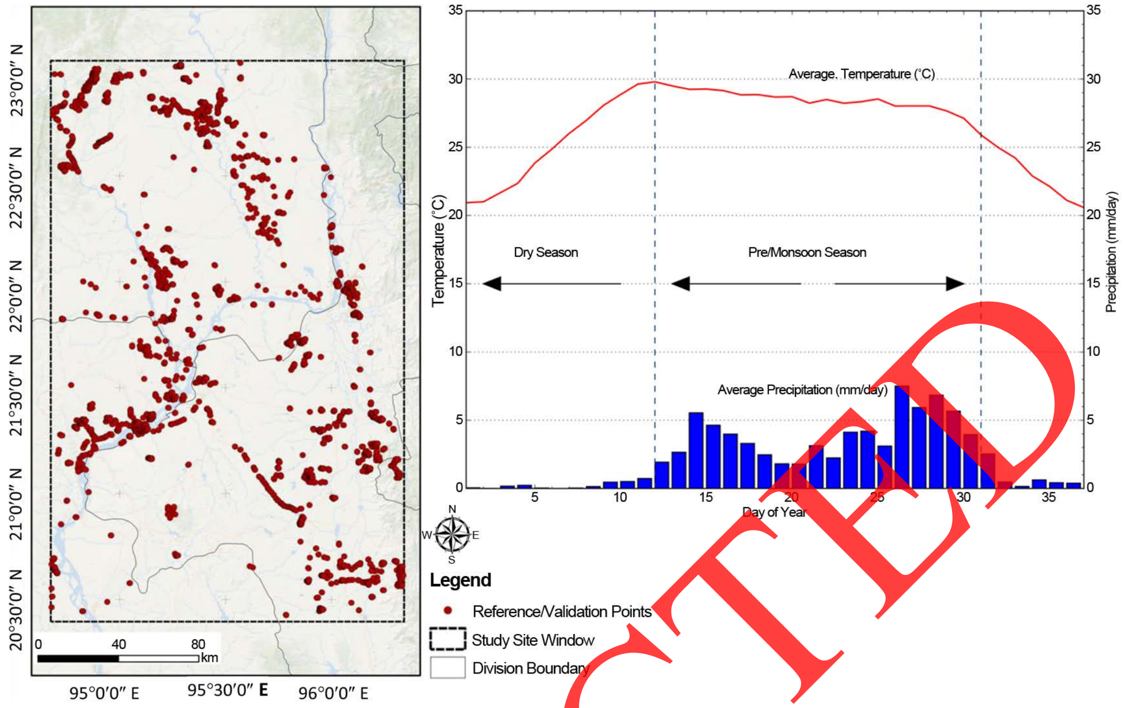

Figure 1. (a) Map showing studied Townships and extent of the processing window; (b) Mean monthly rainfall across in the studied five Townships of CDZM. Locations of reference data are shown in points and grouped into general land use/cover categories.

Areas that produce a single crop per year are also common; especially during years of low rainfall or late onset of rainfall.

3. Methods

3.1. Data Types and Sources

MODIS vegetation indices (NDVI and EVI) in 8-day, 16-day and monthly composites were acquired from http://earthexplorer.usgs.gov data portal. Each of the composite vegetation indices (VIs) was layer-stacked into 1-year (2012), 2-years (2012-2013) and 3-years (2012-2014). We collected reference data from three main sources:

1) An initial field surveys were undertaken in March 2013, based on which a preliminary map was produced. This was followed by participatory and iterative process of additional reference data collection during August 2014 and Nov 2015. The data collected during the field surveys included information such as current crop types, cropping calendars for each crop, water source for crop production and information regarding incidences of stresses such as drought, disease and flood. Information about historical cropping patterns were also collected through discussions with local farmers. We identified 10 major copping patterns and 4 non-agricultural land cover types (Table 1); we implemented a participatory mapping approach where we involved township officers from the Ministry of Agriculture and Irrigation and farmers in assessments of classified maps produced incrementally during the field data collection campaign and in the guided collection of field reference data to address map classification errors. In order to employ this strategy, we collected an initial reference data at random locations and based on which we produced a preliminary map of cropping patterns and LULC classes. The initial map was then used to identify areas dominated by classification errors through involvement of agricultural officers and farmers in identifying inaccurate areas. This procedure was repeated at least three times at each of the four visited townships. To further elaborate this procedure, an initial supervised classification was run using SVM and based on reference data collected through a conventional random sampling procedure and using one year stack of monthly composite of NDVI. Agricultural officers were then invited to evaluate mapping accuracy based on group discussions, where the communities were able to identify particular villages where classification

Table 1. Description of major cropping patterns and non-agricultural land use/cover types.

No.	LULC/cropping pattern	Descriptions
1	Irrigated double rice	Irrigated rice in monsoon, followed by rice in winter, fallow in summer
2	Irrigated rice-upland crops	Irrigated rice in monsoon, followed by upland crops in winter
3	Irrigated double upland crops	Irrigated non-rice crops in monsoon and winter
4	Irrigated rice followed by two upland crops	Irrigated rice in monsoon, followed by upland crops in winter and summer
5	Recession crops	A unique cropping patterns near rivers, where river water cover the crop land during monsoon and in dry season as water recedes different crops are produced using residual moisture
6	Mixed crop class	Crop lands that are characterized by heterogeneous cropping patterns less than the MODIS pixel size
7	Rainfed rice-upland crops	Rainfed rice during monsoon, followed by non-rice crops during winter/summer
8	Rainfed double upland crops	Rainfed non-rice crops during monsoon and winter/summer
9	Rainfed single rice	Rainfed rice during monsoon and fallow during winter and summer
10	Rainfed single upland crops	Rainfed non-rice crops during monsoon and fallow during winter and summer
11	Urban/settlements	varying sizes of towns and cities, villages where built-up surfaces predominate
12	Water bodies	Large rivers, lakes, dams
13	Forest/woodlands	Forest reserves, plantation forests, woodlands dominated by scattered trees and grasslands, with no crop production
14	Barren lands and other non-agricultural lands	Dry river beds, exposed rocky surfaces, with no/low vegetation cover

errors were large. On subsequent days, we incorporated more reference data from areas where major errors were spotted by local people. Ground-based reference data of 1868 points were collected, out of which 542 samples were various cropping patterns and 1326 were non-agricultural land cover types. 2) High spatial resolution data made available by Google® were used to identify non-agricultural land use/cover types such as water bodies, forests and settlement areas. 1786 points were identified from high resolution images; this resulted in total of 3654 reference points in the final dataset. 3) We obtained estimates of area coverage of major crop types from government agricultural statistics.

3.2. Spectral Profiles and Separability Analysis of Reference Data

Spectral separability of reference data was assessed through the Transformed Divergence separability index in ENVI v5.2 [29]. Non-agricultural land uses such as urban, forest/rangeland and water were easily separable from each other and from different cropping patterns of agricultural land uses. Rice-based cropping patterns followed similar spectral patterns cross across all years (**Figure 2**) and spectrally different from non-rice cropping systems but the temporal spectral patterns was similar within rice-based category. The lowest separability values were observed between rainfed double upland crops and irrigated double upland crops. Generally, the highest separability was achieved from the 3-year stack of 8-days VI composites data. It was particularly clear that the 1 year stack of VI composites was not adequately showing spectral separability among classes.

3.3. Classification and Accuracy Comparisons

We evaluated accuracy among 66 different combinations of four sets of variables: classifiers (SVM, RF and C5.0), vegetation indices (NDVI and EVI), size of predictors (one year, two years and three years layers of 8-days, 16-days and monthly composite of each VIs, and size of training data (12 sets of incremental reference data). Using these sets of variables, we compared performances of three state-of-the-art machine learning supervised classification algorithms: support vector machine (SVM), Random Forest (RF) and C5.0 (a decision trees and rule-based model). Model building, parameter tuning, classification and accuracy assessment were performed using "caret" R-package [30] implemented in R language and statistical software v.3.1.2 [31]. In addition to the

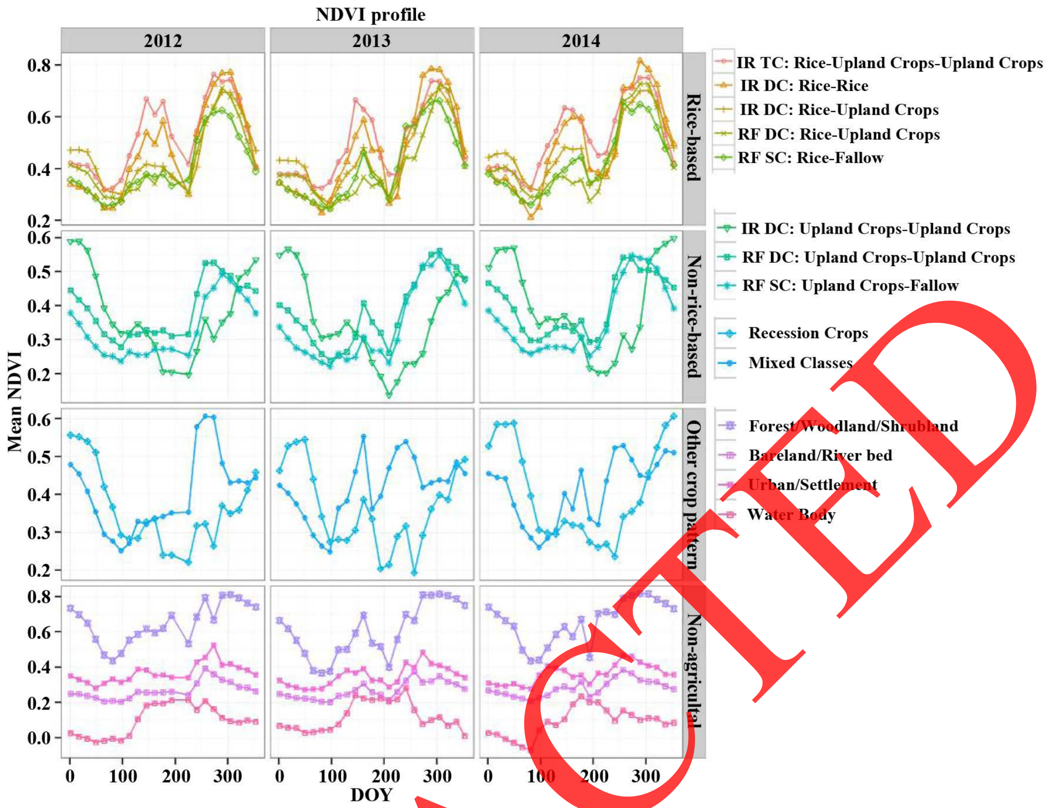

Figure 2. Temporal NDVI profile of cropping patterns and non-agricultural land cover types. IR = Irrigated, RF = Rainfed, TC = Triple crop, DC = Double crops, and SC = Single crop patterns. Upland crops are mainly pigeon pea, chickpea and sesame. DOY is referring to day of year.

"caret" package, a number of other packages were used for image processing and geospatial analysis: the "raster" package [32], "rgdal" [33], "lattice" [34] and "ggplot2" [35].

RF is a classifier consisting of a collection of tree-structured classifiers [36] that form an ensemble of classification and regression tree (CART)-like classifiers [37]. The algorithm can handle high-dimensional datasets and because of its Law of Large Numbers principle, it doesn't suffer from over-fitting [36]. The SVM algorithm is a classification using the Radial Basis Function (RBF) kernel. Although there are various kernels for use in SVM, RBF is considered to be a reasonable first choice due to its attributes outlined in [38]. Ranges of tuning parameters were tested for each of the three algorithms and the tuning parameter that yielded the highest accuracy was considered to be optimal value. For the RF model, "number of trees" were set to 500 as a default value based on suggestion by Breiman and Cutler [39].

The "optimal" tuning parameters of the three models compared in this study (SVM, RF and C5.0) were determined by applying 3-fold cross validation repeated six times. In order to evaluate and compare performances the three models, the predicted classes were compared against the "hold-out" test data. Kappa coefficient and overall accuracy were used as measures of accuracies, which are two of commonly used measures of assessing the accuracy of thematic maps [40]. User's and producer's accuracies were also assessed to examine accuracy of each class. Paired t-test was used to assess statistical differences among accuracies achieved by different algorithms [41] implementing the "caret" R-Package, which is based on Hothorn, Leisch [42] and Eugster, Hothorn [43]. A summary of classification procedures and accuracy assessment is shown in **Figure 3**.

4. Results

4.1. Accuracy Improvements through Participatory Iterative Mapping and Model Comparisons

Figure 4(a) shows class accuracies (user's and producer's) and the trends of accuracy changes for each class as

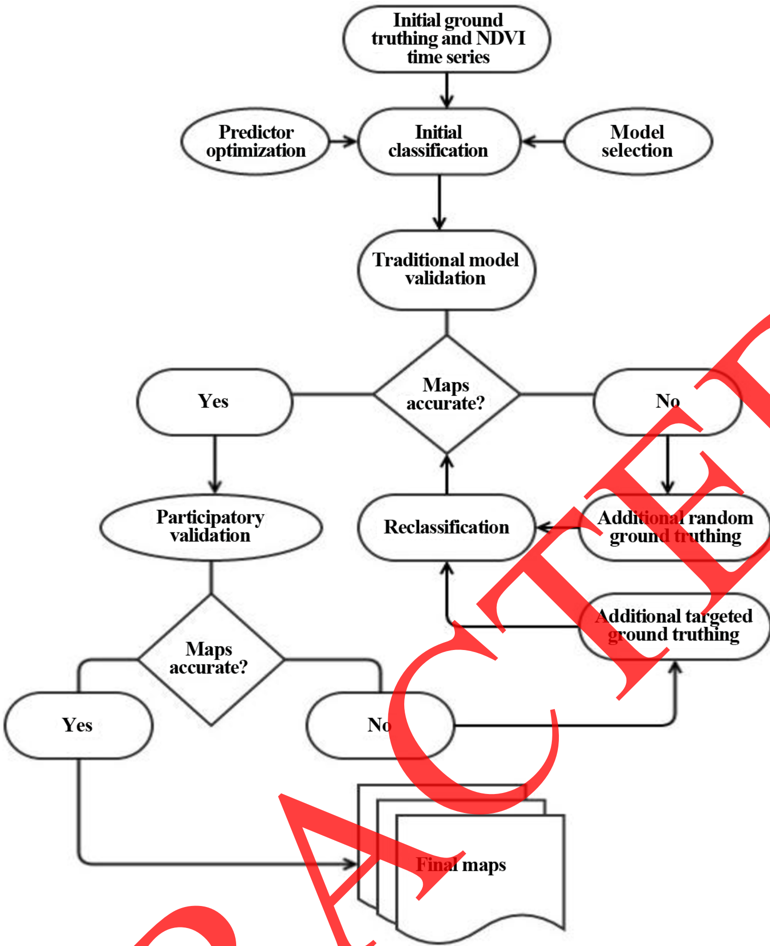

Figure 3. Flow chart of analysis procedures.

size of training data increases and the size of predictor layers being constant. The result is based on classification by SVM model and three years stack of 16-days NDVI composite. It should be noted from **Figure 4(a)** that the accuracy improvement achieved through increasing number of iteration (number of days) was not the same for different classes: large accuracy fluctuation was observed in non-agricultural land cover category and rice-based cropping patterns generally showed increasing trends of accuracy (**Figure 4(a)**). When small size of training data was used, RF resulted in relatively higher accuracy (**Figure 4(b)**). The significance of difference tended to diminish as the size of training data set gets larger. The accuracy difference between the initial classification, which was based on conventional random sampling of reference data, and the final reference data collected through iterative classification is about 31%, 32% and 43% for RF, SVM and C5.0 classifiers, respectively (**Figure 4(b)**). This was due to confusion between urban and bare surfaces such as dry river beds and exposed rocky areas.

Accuracy achieved by all models increased with size of predictor dataset and differences among the three classifiers depended on size of predictor layers (**Figure 5**). Highest accuracy was achieved by SVM using 8 days NDVI composite of 3 years, but not significantly different from RF and C5.0 models. Though the differences are small, SVM was shown to consistently outperform and in most cases; C5.0 model not only resulted in lowest accuracy (**Figure 4(b)**), but also consumed the largest computational time (**Figure 7**). Accuracy differences among the three classifiers was clearer and consistent when large number of predictor variables are used compared to comparisons using fewer number of predictors (12 stack layers versus 137 stack layers) (**Figure 6(a)** and **Figure 6(b)**). With largest number of predictors (137 layers), processing time consumed by C5.0 was nearly 8 times that of SVM. In all sets of data, SVM was shown to be computationally most efficient (**Figure 7**). Statistical significance of accuracy difference among classifiers was higher when fewer numbers of predictor layers

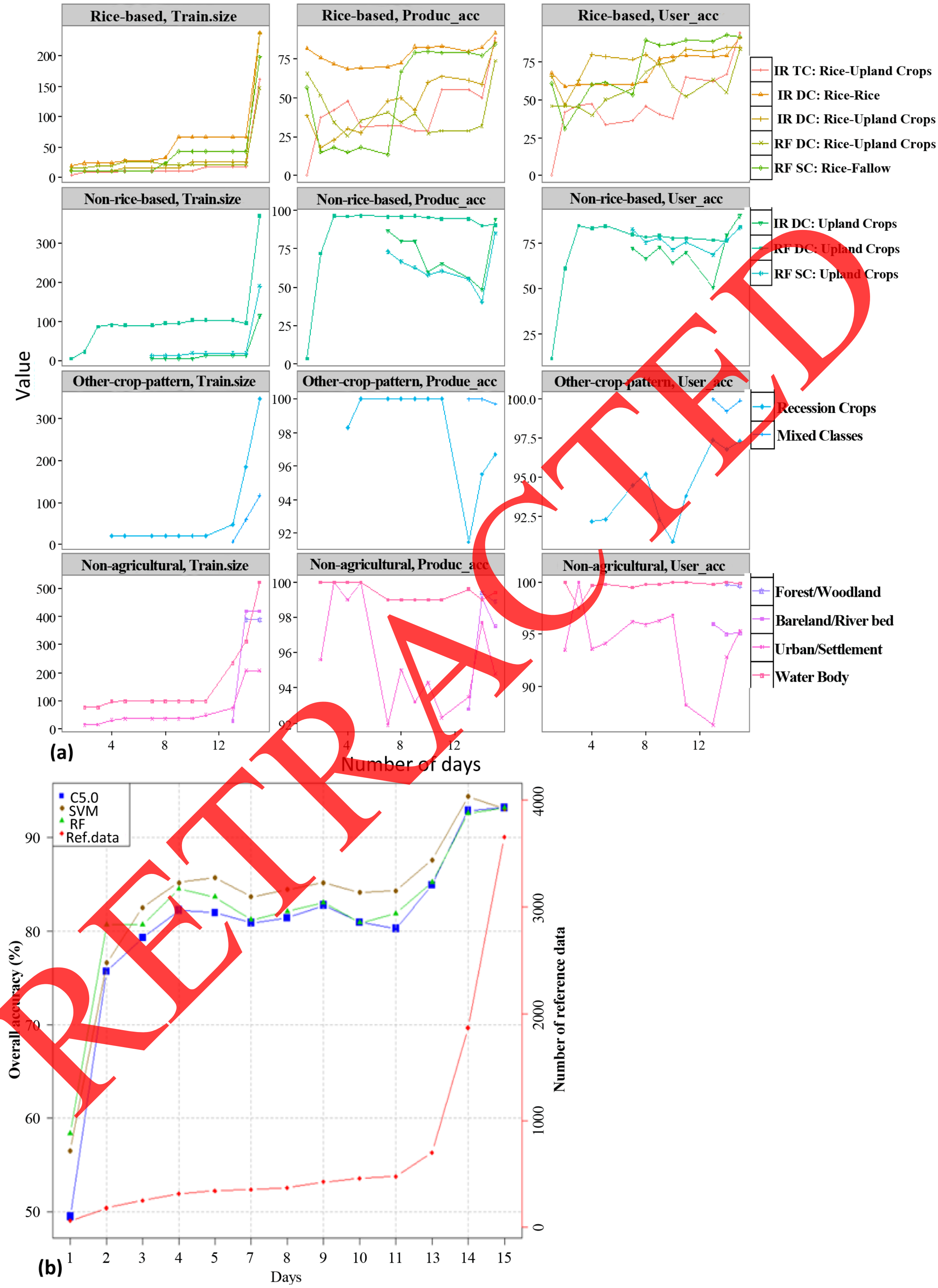

Figure 4. Accuracy improvement achieved through participatory iterative mapping procedures.

Figure 5. Accuracy differences among the three machine learning algorithms compared using NDVI layers. With increasing numbers of predictor variables, accuracy differences become more obvious; particularly, SVM model showed significantly higher accuracy as the size of predictor layers increases.

are used. Accuracy of the three models was higher when NDVI is used compared to same number of EVI layers and accuracy difference of up to 4% was observed between NDVI and EVI when SVM classifier was used (**Figures 8(a)-(c)**).

Accuracy also varied with type of vegetation indices used in the prediction. The lowest accuracy was observed when monthly EVI of 1 year was used and increasing the number of VIs layers (predictor variables) has

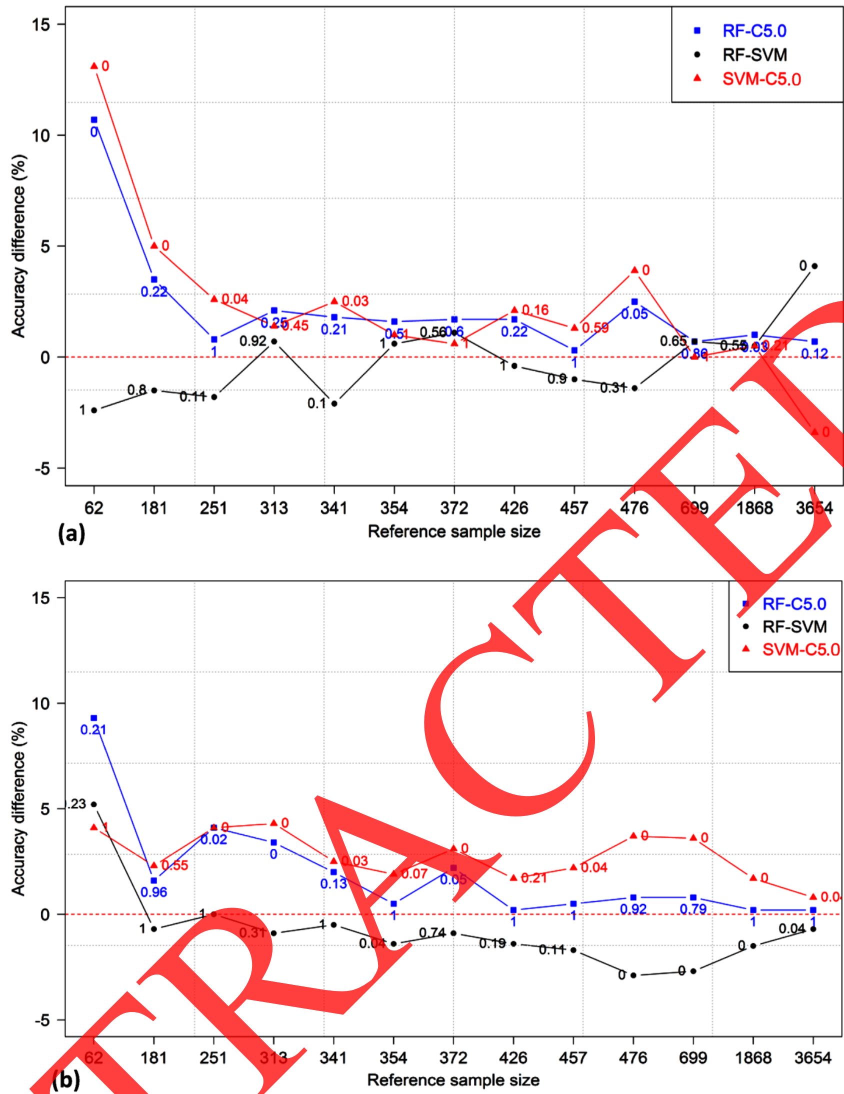

Figure 6. Accuracy differences among the models with increasing size of reference data and (a) the smallest number of predictor layers (1 year stack of NDVI monthly composite (12 layer stacks)), (b) the largest number of predictor layers (3 years stack of NDVI 8-days (137 layer stack)).

considerably improved accuracy (**Figure 8**).

4.2. Cropping Pattern Map

The final classification output developed using SVM algorithm is shown in **Figure 9**. An overall accuracy of 93% and Kappa coefficient of 0.92 were achieved when SVM was applied using the final training data set and 3 years layers of 16-days NDVI composite. Both user's and producer's accuracy of all classes exceeded 74% (**Table 2**). The smallest producer's and user's accuracies were due to large confusion among class "rainfed monsoon rice and winter/summer upland crops", "irrigated monsoon rice and winter/summer upland crops", and "rainfed monsoon upland crop and winter/summer upland crops" (**Table 2**).

5. Discussion

Active involvement of local agricultural officers and farmers in the iterative process of mapping was an efficient procedure for improving accuracy and identifying areas with high classification errors. It not only helped rapid

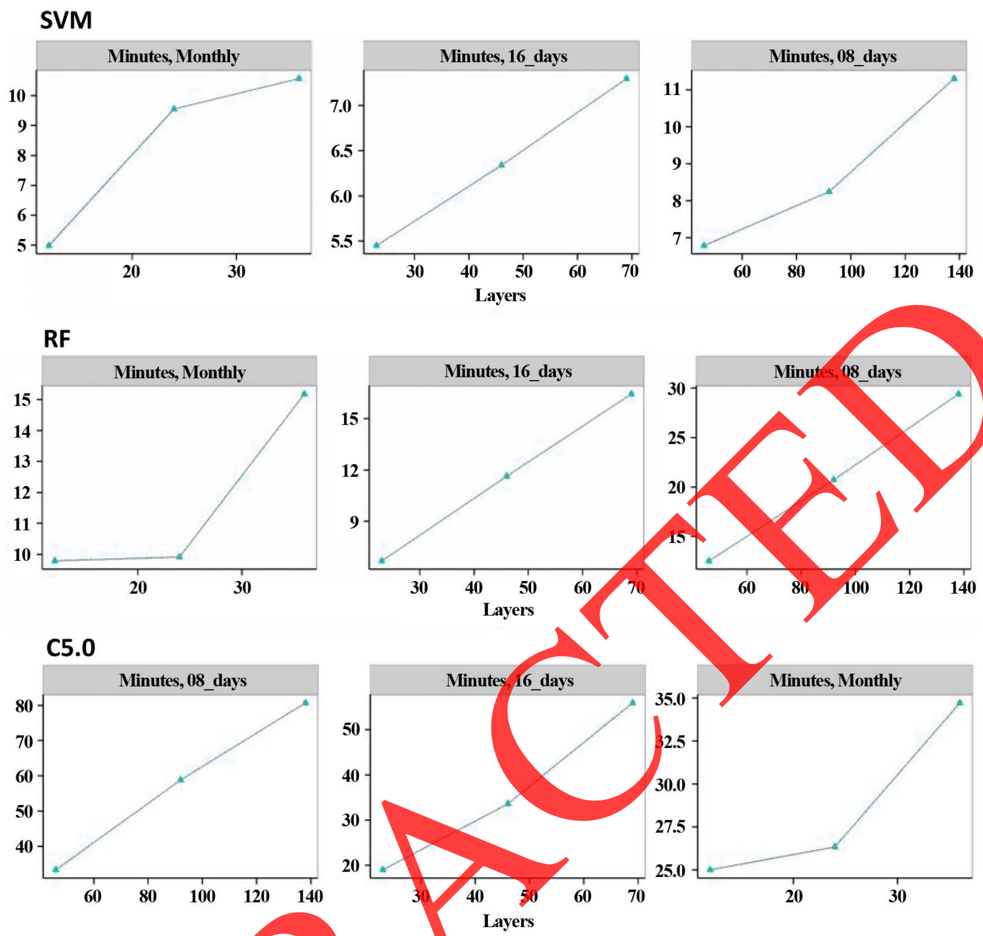

Figure 7. Comparison of computational time elapsed: within algorithm (row) and across algorithm (columns). Total time elapsed to complete model training and determining final tuning parameter is calculated in seconds (rounded to minutes).

assessment of accuracy, but also facilitated identification of priority sampling locations which enabled substantial error reduction. Information regarding cropping patterns may not be easy to collect by field observations, since once can see only one type of land cover at a time. This was made possible by iteratively evaluating the accuracy of each class as we added more reference data and involved local people to examine how the classification errors were minimized by including more samples from locations where predominant errors were detected by the people. The participatory reference data collection was an efficient way of collecting accurate reference data since local people have knowledge about the historical patterns and changes of such patterns in time.

Ground-based data collection is expensive and time consuming, particularly when the area to be mapped is large and accessibility is constrained by factors such as absence of roads and conflicts. Foody [44] emphasizes the need for consideration of sample size and economic feasibility of sample collection. Resources being limited, sampling strategy need to be economical and the quantity and quality of the data should allow for statistically meaningful analyses (Congalton, 1991). Recent study by Mialhe, Gunnell [45] has also highlighted the importance of participatory mapping approach in data-poor areas, and in areas such as conflict zones. Our approach signifies the importance of examining mapping accuracy in a participatory process particularly in conditions where sites are inaccessible or difficult to access. Arguably, participation of local community in the mapping process enables more efficient way of sampling for substantial accuracy improvements. The approach we applied has two-fold improvements: first, it minimizes time and resource needed for traveling and collection of reference data by facilitating an efficient and accurate way of generating data over large areas, and inaccessible areas; with the help of the discussion with local people, more ground samples were systematically allocated to areas that are highly heterogeneous, which could be caused by complexity of land cover patterns; second, we were able to assess location accuracy of mapping at locations for which we have no prior knowledge/reference data, which would not be possible through traditional accuracy assessment in remote sensing. This interactive and iterative process of mapping has resulted in substantial improvement of accuracy ranging from 31% to 43%.

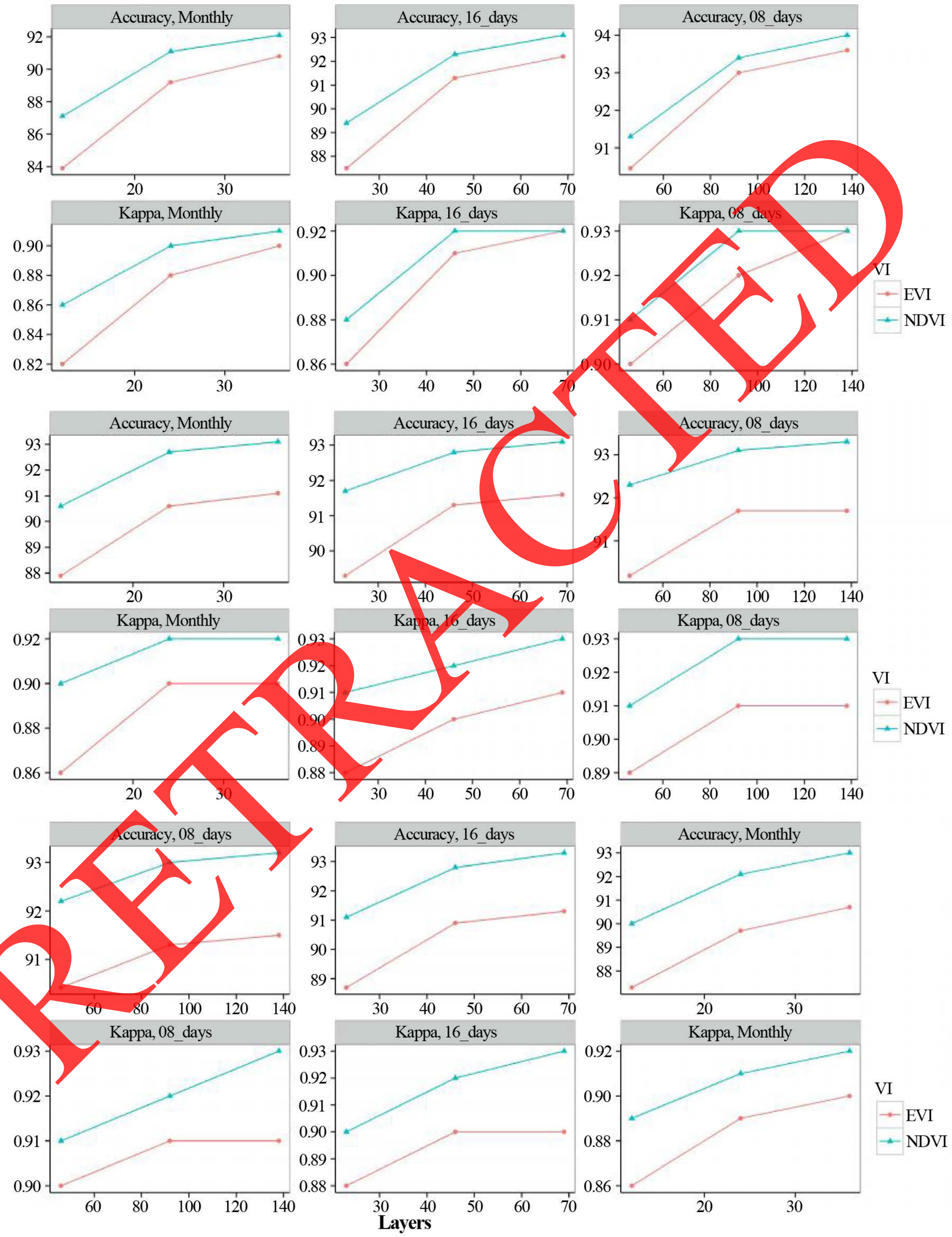

Figure 8. Accuracy variation with the use of NDVI and EVI as predictors: (a) SVM, (b) RF and (c) C5.0; (d) shows accuracy differences among the three models using NDVI 16-days composite of 3 years.

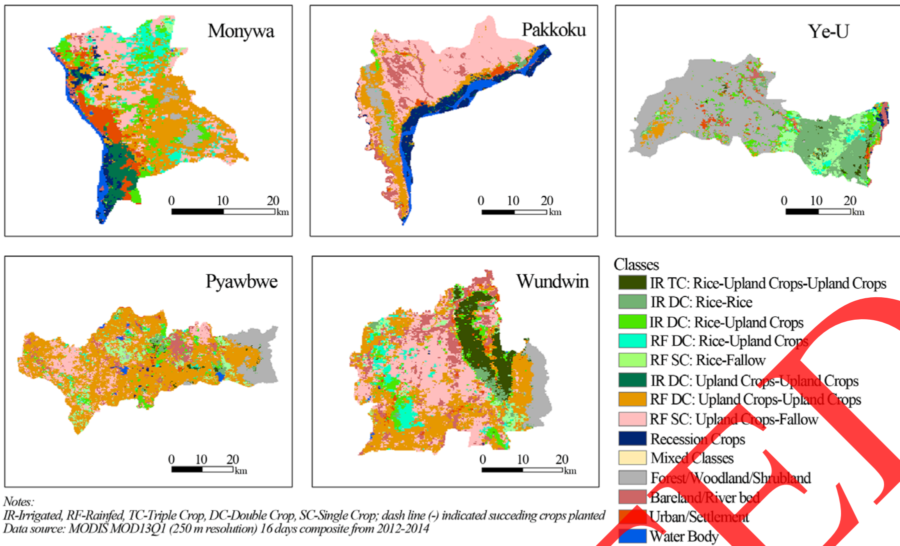

Figure 9. Final map of cropping patterns and major non-agricultural land use/cover for the five Townships.

Table 2. Error matrix of final map based on SVM algorithm using NDVI 16-days composite of 3 years stack computed applying cross-validation. One-third of the final 3654 reference data (combination of ground-based data and samples identified on screen using high resolution imagery provided by Google Earth®. Description of the classes is: 1 = Barren land and Other nonagricultural lands, 2 = Forest/Woodland/Scrubland, 3 = Irrigated monsoon rice and summer rice, 4 = Irrigated monsoon rice and winter/summer upland crops, 5 = Irrigated monsoon upland crop and winter/summer upland crop, 6 = Irrigated monsoon rice, winter upland crop and summer upland crops, 7 = Mixed crop class (this is areas where it was no unique cropping pattern, mixture of various cropping patterns were identified within an area of one MODIS pixel size, 8 = Recession crops (these are crop produced near rivers and creeks, where no crop is grown during monsoon and as water recedes, various crops are grown using residual moisture without/with supplemental irrigation, 9 = Rainfed monsoon rice and winter/summer upland crops, 10 = Rainfed monsoon upland crop and winter/summer upland crops, 11 = Rainfed single crop monsoon rice, 12 = Rainfed single crop monsoon upland crops, 13 = Urban/settlement areas, 14 = Water.

		Reference data															
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**Total**	**User**
SVM predicted	**1**	136	0	0	2	0	0	0	2	0	0	0	2	1	1	**144**	**95**
	2	0	128	1	0	0	0	0	0	0	0	0	0	0	0	**129**	**99**
	3	0	0	71	2	0	1	0	0	1	0	3	0	0	0	**78**	**92**
	4	0	1	1	67	0	3	0	0	4	1	1	1	1	0	**80**	**84**
	5	0	0	0	1	36	0	0	1	0	1	0	0	0	0	**39**	**91**
	6	0	0	2	1	0	48	0	0	0	0	0	0	0	0	**52**	**93**
	7	0	0	0	0	0	0	39	0	0	0	0	0	0	0	**39**	**100**
	8	0	0	0	0	1	0	0	111	0	1	0	0	0	1	**115**	**97**
	9	0	0	1	4	0	0	0	0	36	0	0	1	0	0	**43**	**84**
	10	0	1	1	1	0	0	0	0	4	112	6	5	1	0	**132**	**85**
	11	1	0	3	0	0	0	0	0	1	1	56	1	0	0	**63**	**90**
	12	2	0	0	0	0	0	0	0	2	6	1	54	0	0	**64**	**84**
	13	0	0	0	1	1	0	0	0	0	1	0	0	66	0	**69**	**96**
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	172	**172**	**100**
	Total	**140**	**129**	**79**	**79**	**38**	**54**	**39**	**115**	**49**	**123**	**66**	**63**	**69**	**174**	**1217**	
	Prod	**97**	**99**	**90**	**85**	**95**	**89**	**100**	**97**	**73**	**91**	**85**	**86**	**96**	**99**		
	Overall		**93**														

The accuracy performance of all machine learning algorithms compared in this study has increased with increasing size of predictor layers and training sample size, which is in agreement with studies such as Li, Wang [18]. With limited number of training samples, RF was slightly better in accuracy. This may imply the robustness of RF when reference data are limited. Similar comparative studies that attempted to examine accuracy differences among different machine learning algorithms showed different results. Studies by Adam, Mutanga [46] and Duro, Franklin [22] indicated machine learning algorithms such as SVM and RF performed similarly, and other studies such as that of Cracknell and Reading [47], Fernández-Delgado, Cernadas [17] and Kampichler, Wieland [48] showed RF to be the most accurate algorithm. Our result, which showed SVM to be superior in accuracy compared to both RF and C5.0, is yet in agreement with studies such as that of Shao and Lunetta [16] and Maroco, Silva [49] indicated SVM to be superior in accuracy.

Though accuracy comparisons could depend on a number of factors, machine learning algorithms are sensitive to tuning parameter. Tuning of learning parameters is often not straight forward [50]. In this study, the "best" values of tuning parameters were determined automatically based on highest accuracy achieved by the methods. Our results of computational time comparison also showed SVM is the most efficient method and both lowest accuracy and highest computational time was shown by C5.0 algorithm. This is an interesting finding particularly when classification of large data set over large spatial scales is of interest.

Comparisons of accuracy difference between NDVI and EVI are rare in the literature. A study by Wardlow and Egbert [51], which compared agreement of crop information derived from MODIS EVI and NDVI products with ground-based measurements indicated no considerable difference between the two indices. Our observation which indicated mapping accuracy to be relatively higher with the use of NDVI as predictor over EVI, could be explained by the sensitivity of EVI to variation in blue band compared to NDVI [52], which could result in more noises particularly in time series analysis, despite atmospheric correction applied to the images, varying magnitude of atmospheric effects could be manifested in values of EVI.

6. Conclusions

We implemented a unique participatory iterative mapping technique to minimize mapping errors in a complex agricultural landscape. Accuracy performances (overall accuracy and kappa coefficient) of three machine learning algorithms were evaluated after determining optimal tuning parameters for each of the algorithms. The accuracy changes with increasing size of training samples and predictor layers were examined. Variation in accuracy with the use of NDVI and EVI as predictors was also explored. In addition to accuracy measures, processing time was also used as a criterion to assess model performances.

The participation of local agricultural officers and farmers people in iterative mapping process substantially improved accuracy and enabled efficient sampling strategies. We were able to substantially improve accuracy by up to 43% through optimally identifying the most suitable algorithm and predictor data types and sizes. An overall accuracy of 93% was achieved and 14 major classes of cropping patterns and non-agricultural land cover types were detected. The use of larger number of VIs stacks increased accuracy by up to 4% compared to the smallest number of predictor layers (one year stack of monthly composite versus 3 years stack of 16-days or 8-days composite). NDVI resulted in higher accuracy compared to EVI of same number of layers.

Comparison based on overall accuracy and kappa coefficient shows that consistently higher, though small, accuracy was achieved by SVM compared to RF and C5.0. Statistical significance of accuracy differences depended mainly on number of predictors: with fewer numbers of predictors, less significance of differences were observed. When smallest number of training samples were used to predict classes using only 12 layers of VI, RF was slightly better in accuracy. As the number of training sample becomes larger and the size of predictor layers increases, all of the three algorithms performed well, though SVM was relatively superior. These results provide insights into the importance of size of training samples and predictor variables in deciding which classifier may need to be used in order to achieve high accuracy levels.

Acknowledgements

The authors would like to acknowledge the assistance of the farmers and the officers of the Ministry of Agriculture and Irrigation in the townships of Monywa, Ye'U, Pakoku, Wundwin and Pyawbwe.

Funding

This research was supported by the Livelihood and Food Security Trust Fund (LIFT) of the *United Nations Office* *for Project Services* (*UNOPS*) in Myanmar under a project titled: Reducing risks and improving livelihoods in the rice environments of Myanmar through better targeting of management options MOA Number R1.4/007/2012.

References

[1] Atzberger, C. (2013) Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sensing*, **5**, 949-981. http://dx.doi.org/10.3390/rs5020949

[2] Food and Agriculture Organization of the United Nations (FAO) (2011) Global Strategy to Improve Agricultural and Rural Statistics. FAO, Rome.

[3] Mather, P. and Tso, B. (2009) Classification Methods for Remotely Sensed Data. CRC Press, Boca Raton.

[4] Zhong, L. (2012) Efficient Crop Type Mapping Based on Remote Sensing in the Central Valley, California. PhD Thesis, University of California, Berkeley.

[5] Wardlow, B.D., Egbert, S.L. and Kastens, J.H. (2007) Analysis of Time-Series MODIS 250 m Vegetation Index Data for Crop Classification in the US Central Great Plains. *Remote Sensing of Environment*, **108**, 290-310. http://dx.doi.org/10.1016/j.rse.2006.11.021

[6] Long, J.A., Lawrence, R.L., Greenwood, M.C., Marshall, L. and Miller, P.R. (2013) Object-Oriented Crop Classification Using Multitemporal ETM+ SLC-Off Imagery and Random Forest. *GIScience & Remote Sensing*, **50**, 418-436.

[7] Sakamoto, T., Gitelson, A.A. and Arkebauer, T.J. (2014) Near Real-Time Prediction of US Corn Yields Based on Time-Series MODIS Data. *Remote Sensing of Environment*, **147**, 219-231. http://dx.doi.org/10.1016/j.rse.2014.03.008

[8] Hmimina, G., Dufrêne, E., Pontailler, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., de Grandcourt, A., Burban, B., Flechard, C. and Granier, A. (2013) Evaluation of the Potential of MODIS Satellite Data to Predict Vegetation Phenology in Different Biomes: An Investigation Using Ground-Based NDVI Measurements. *Remote Sensing of Environment*, **132**, 145-158. http://dx.doi.org/10.1016/j.rse.2013.01.010

[9] Gumma, M.K., Thenkabail, P.S., Maunahan, A., Islam, S. and Nelson, A. (2014) Mapping Seasonal Rice Cropland Extent and Area in the High Cropping Intensity Environment of Bangladesh Using MODIS 500m Data for the Year 2010. *ISPRS Journal of Photogrammetry and Remote Sensing*, **91**, 98-113. http://dx.doi.org/10.1016/j.isprsjprs.2014.02.007

[10] Peng, D.L., Huang, J.F., Li, C.J., Liu, L.Y., Huang, W.J., Wang, F.M. and Yang, X.H. (2014) Modelling Paddy Rice Yield Using MODIS Data. *Agricultural and Forest Meteorology*, **184**, 107-116. http://dx.doi.org/10.1016/j.agrformet.2013.09.006

[11] Busetto, L., Meroni, M. and Colombo, R. (2008) Combining Medium and Coarse Spatial Resolution Satellite Data to Improve the Estimation of Sub-Pixel NDVI Time Series. *Remote Sensing of Environment*, **112**, 118-131. http://dx.doi.org/10.1016/j.rse.2007.04.004

[12] Jain, M., Mondal, P., De Fries, R.S., Small, C. and Galford, G.L. (2013) Mapping Cropping Intensity of Smallholder Farms: A Comparison of Methods Using Multiple Sensors. *Remote Sensing of Environment*, **134**, 210-223. http://dx.doi.org/10.1016/j.rse.2013.02.029

[13] Gumma, M.K., Thenkabail, P.S., Hideto, F., Nelson, A., Dheeravath, V., Busia, D. and Rala, A. (2011) Mapping Irrigated Areas of Ghana Using Fusion of 30 m and 250 m Resolution Remote-Sensing Data. *Remote Sensing*, **3**, 816-835. http://dx.doi.org/10.3390/rs3040816

[14] Chen, C.-R., Chen, C.-F. and Son, N.-T. (2014) Rice Crop Monitoring with Multitemporal MODIS-Landsat Data Fusion. *Proceedings of the EGU General Assembly Conference*, Vienna, 27 April-2 May 2014.

[15] Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sanchez, J. (2012) An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification. *ISPRS*: *Journal of Photogrammetry and Remote Sensing*, **67**, 93-104. http://dx.doi.org/10.1016/j.isprsjprs.2011.11.002

[16] Shao, Y. and Lunetta, R.S. (2012) Comparison of Support Vector Machine, Neural Network, and CART Algorithms for the Land-Cover Classification Using Limited Training Data Points. *ISPRS*: *Journal of Photogrammetry and Remote Sensing*, **70**, 78-87. http://dx.doi.org/10.1016/j.isprsjprs.2012.04.001

[17] Fernández-Delgado, M., Cernadas, E., Barro, S. and Amorim, D. (2014) Do We Need Hundreds of Classifiers to Solve Real World Classification Problems? *The Journal of Machine Learning Research*, **15**, 3133-3181.

[18] Li, C., Wang, J., Wang, L., Hu, L. and Gong, P. (2014) Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. *Remote Sensing*, **6**, 964-983. http://dx.doi.org/10.3390/rs6020964

[19] Qian, Y., Zhou, W., Yan, J., Li, W. and Han, L. (2014) Comparing Machine Learning Classifiers for Object-Based Land Cover Classification Using Very High Resolution Imagery. *Remote Sensing*, **7**, 153-168. http://dx.doi.org/10.3390/rs70100153

[20] Liu, M., Wang, M., Wang, J. and Li, D. (2013) Comparison of Random Forest, Support Vector Machine and Back Propagation Neural Network for Electronic Tongue Data Classification: Application to the Recognition of Orange Beverage and Chinese Vinegar. *Sensors and Actuators B*: *Chemical*, **177**, 970-980. http://dx.doi.org/10.1016/j.snb.2012.11.071

[21] Sonobe, R., Tani, H., Wang, X., Kobayashi, N. and Shimamura, H. (2014) Parameter Tuning in the Support Vector Machine and Random Forest and Their Performances in Cross- and Same-Year Crop Classification Using TerraSAR-X. *International Journal of Remote Sensing*, **35**, 7898-7909. http://dx.doi.org/10.1080/01431161.2014.978038

[22] Duro, D.C., Franklin, S.E. and Dubé, M.G. (2012) A Comparison of Pixel-Based and Object-Based Image Analysis with Selected Machine Learning Algorithms for the Classification of Agricultural Landscapes Using SPOT-5 HRG Imagery. *Remote Sensing of Environment*, **118**, 259-272. http://dx.doi.org/10.1016/j.rse.2011.11.020

[23] Caruana, R. and Niculescu-Mizil, A. (2006) An Empirical Comparison of Supervised Learning Algorithms. *Proceedings of the* 23rd *International Conference on Machine Learning*, Pittsburgh, 25-29 June 2006. http://dx.doi.org/10.1145/1143844.1143865

[24] Ben-Hur, A. and Weston, J. (2010) A User's Guide to Support Vector Machines. In: Carugo, O. and Eisenhaber, F., Eds., *Data Mining Techniques for the Life Sciences*, Humana Press, New York, 223-239.

[25] Pal, M. (2005) Random Forest Classifier for Remote Sensing Classification. *International Journal of Remote Sensing*, **26**, 217-222. http://dx.doi.org/10.1080/01431160412331269698

[26] Pal, M. and Mather, P.M. (2005) Support Vector Machines for Classification in Remote Sensing. *International Journal of Remote Sensing*, **26**, 1007-1011. http://dx.doi.org/10.1080/01431160512331314083

[27] Lek, S. and Guégan, J.-F. (2012) Artificial Neuronal Networks: Application to Ecology and Evolution. Springer Science & Business Media, New York.

[28] Michie, D., Spiegelhalter, D.J. and Taylor, C.C. (1994) Machine Learning, Neural and Statistical Classification. Ellis Horwood Limited, Chichester.

[29] Exelis Visual Information Solutions, ENVI v. 5.2, Boulder, 2015.

[30] Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Team, R.C., Benesty, M., Lescarbeau, R., Ziem, A. and Scrucca, L. (2015) Caret: Classification and Regression Training. R Package Version 6.0-41. http://CRAN.R-project.org/package=caret

[31] R Core Team (2014) R: A Language and Environment for Statistical Computing. Vienna. http://www.R-project.org

[32] Hijmans, R.J. (2015) Raster: Geographic Data Analysis and Modeling. R Package Version 2.3-24. http://CRAN.R-project.org/package=raster

[33] Bivand, R., Keitt, T. and Rowlingson, B. (2014) Rgdal: Bindings for the Geospatial Data Abstraction Library. R Package Version 0.9-1. http://CRAN.R-project.org/package=rgdal

[34] Sarkar, D. (2008) Lattice: Multivariate Data Visualization with R. http://lmdvr.r-forge.r-project.org

[35] Wickham, H. (2009) ggplot2: Elegant Graphics for Data Analysis. http://had.co.nz/ggplot2/book

[36] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. http://dx.doi.org/10.1023/A:1010933404324

[37] Breiman, L., Friedman, J., Stone, C.J. and Olshen, R.A. (1984) Classification and Regression Trees. CRC Press, Boca Raton.

[38] Hsu, C.-W., Chang, C.-C. and Lin, C.-J. (2003) A Practical Guide to Support Vector Classification. http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[39] Breiman, L. and Cutler, A. (2007) Random Forests-Classification Description. Department of Statistics, Berkeley.

[40] Congalton, R.G. and Green, K. (1999) Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers, Boca Raton, London and New York.

[41] Foody, G.M. (2004) Thematic Map Comparison: Evaluating the Statistical Significance of Differences in Classification Accuracy. *Photogrammetric Engineering and Remote Sensing*, **70**, 627-633. http://dx.doi.org/10.14358/PERS.70.5.627

[42] Hothorn, T., Leisch, F., Zeileis, A. and Hornik, K. (2005) The Design and Analysis of Benchmark Experiments. *Journal of Computational and Graphical Statistics*, **14**, 675-699. http://dx.doi.org/10.1198/106186005X59630

[43] Eugster, M.J., Hothorn, T. and Leisch, F. (2008) Exploratory and Inferential Analysis of Benchmark Experiments. Technical Report Number 030, Department of Statistics, University of Munich, Munich.

[44] Foody, G.M. (2009) Classification Accuracy Comparison: Hypothesis Tests and the Use of Confidence Intervals in Evaluations of Difference, Equivalence and Non-Inferiority. *Remote Sensing of Environment*, **113**, 1658-1663. http://dx.doi.org/10.1016/j.rse.2009.03.014

[45] Mialhe, F., Gunnell, Y., Ignacio, J.A.F., Delbart, N., Ogania, J.L. and Henry, S. (2015) Monitoring Land-Use Change by Combining Participatory Land-Use Maps with Standard Remote Sensing Techniques: Showcase from a Remote Forest Catchment on Mindanao, Philippines. *International Journal of Applied Earth Observation and Geoinformation*, **36**, 69-82. http://dx.doi.org/10.1016/j.jag.2014.11.007

[46] Adam, E., Mutanga, O., Odindi, J. and Abdel-Rahman, E.M. (2014) Land-Use/Cover Classification in a Heterogeneous Coastal Landscape Using RapidEye Imagery: Evaluating the Performance of Random Forest and Support Vector Machines Classifiers. *International Journal of Remote Sensing*, **35**, 3440-3458. http://dx.doi.org/10.1080/01431161.2014.903435

[47] Cracknell, M.J. and Reading, A.M. (2014) Geological Mapping Using Remote Sensing Data: A Comparison of Five Machine Learning Algorithms, Their Response to Variations in the Spatial Distribution of Training Data and the Use of Explicit Spatial Information. *Computers & Geosciences*, **63**, 22-33. http://dx.doi.org/10.1016/j.cageo.2013.10.008

[48] Kampichler, C., Wieland, R., Calmé, S., Weissenberger, H. and Arriaga-Weiss, S. (2010) Classification in Conservation Biology: A Comparison of Five Machine-Learning Methods. *Ecological Informatics*, **5**, 441-450. http://dx.doi.org/10.1016/j.ecoinf.2010.06.003

[49] Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I. and de Mendonça, A. (2011) Data Mining Methods in the Prediction of Dementia: A Real-Data Comparison of the Accuracy, Sensitivity and Specificity of Linear Discriminant Analysis, Logistic Regression, Neural Networks, Support Vector Machines, Classification Trees and Random Forests. *BMC Research Notes*, **4**, 299. http://dx.doi.org/10.1186/1756-0500-4-299

[50] Snoek, J., Larochelle, H. and Adams, R.P. (2012) Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems*, **25**, 2960-2968.

[51] Wardlow, B.D. and Egbert, S.L. (2010) A Comparison of MODIS 250-m EVI and NDVI Data for Crop Mapping: A Case Study for Southwest Kansas. *International Journal of Remote Sensing*, **31**, 805-830. http://dx.doi.org/10.1080/01431160902897858

[52] Fensholt, R., Sandholt, I. and Stisen, S. (2006) Evaluating MODIS, MERIS, and VEGETATION Vegetation Indices Using *in Situ* Measurements in a Semiarid Environment. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 1774-1786. http://dx.doi.org/10.1109/TGRS.2006.875940