Multinomial Logistics Regression for Digital Image Classification

Dr. Moe Myint, Chief Scientist, Mapping and Natural Resources Information Integration (MNRII), Switzerland
maungmoe.myint@mnrii.com

KEY WORDS: Image segmentation, Spectral Classes, Information Classes, Predictor and Response variables

ABSTRACT: Spectrally homogeneous and spatially contiguous image segments could be derived using the image segmentation algorithms such as K-mean. This paper suggests multinomial logistics regression models to interpret the spectrally homogeneous and spatially contiguous segments to fit them into the examined information classes instead of using traditional entire experience-based labeling systems to associate these cluster segments with meaningful information classes.

STUDY AREA

The study area is at west of Phnum Bokor National Park, North of Sihanouk valley Delta and East of Gulf of Thailand in Cambodia. The land cover dynamic is rapid because of the intensive expansion of oil palm. It is important to map; measure and monitor the land cover, land use and vegetation stratification dynamics for sustainable land management and value chain analyses of oil palm study. This study focus only general land cover, land use and oil palm classification using Multinomial Logistics Regression.

OBJECTIVES

The application objective is to classify the image into land use and land cover classes including mature oil palm and young oil palm. The theoretical objective is to apply image segmentation and multinomial logistic regression modeling in order to accomplish the application objective.

MATERIALS AND METHODS

Satellite Data

Digital satellite image acquired in December 2014 from the Landsat 8 satellite was applied for the study. Landsat 8 is an American Earth Observation Satellite launched on February 11, 2013. It has the two sensor payload, the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). Its radiometric resolution is 16bits. The following table describes the wavelength and spatial resolutions of individual spectral band of Landsat-8.

Spectral Bands	Wavelength	Resolution	Sensor payload
Band 1 – Coastal / Aerosol	0.433 – 0.453 µm	30 m	OLI
Band 2 - Blue	0.450 – 0.515 µm	30 m	OLI
Band 3 - Green	0.525 – 0.600 µm	30 m	OLI
Band 4 - Red	0.630 – 0.680 µm	30 m	OLI
Band 5 – Near Infrared	0.845 – 0.885 µm	30 m	OLI
Band 6 – Short Wavelength Infrared I	1.560 – 1.660 µm	30 m	OLI
Band 7 – Short Wavelength Infrared II	2.100 – 2.300 µm	30 m	OLI
Band 8 - Panchromatic	0.500 – 0.680 µm	15 m	OLI
Band 9 - Cirrus	1.360 – 1.290 µm	30 m	OLI
Band 10 – Long Wavelength Infrared I	10.30 – 11.30 µm	100 m	TIRS
Band 11 – Long Wavelength Infrared II	10.50 – 12.50 µm	100 m	TIRS

In this study Band -2 through Band 7 is selected because of their spectral sensitivity of land cover objects such as vegetation, water, soil. Moreover, the Normalized Difference Vegetation Index (NDVI) is calculated using Red and Near Infrared Bands as the derived. The NDVI value (-1 to +1) was stretched to 8 bits radiometric level. The NDVI is added as the additional layer to the image. Therefore, this study applied 7 layers of spectral information which represent blue, Green, Red, Near Infrared, Short wavelength infrared I and II, and NDVI.

Field Data

Field data collection and documentation of land cover, land use and oil palm plantations were carried out using GPS. Young oil palm (YPALM), mature oil palm (MPALM), agriculture (AG), waterbody (WATER), forests (FOREST), waterlogged or swampy areas (SWAMP), clearing of land within the forest Area (OPENFOR), clearing of land outside the forest area (OPENLAND), land development activities associated with agriculture, horticulture and industries activities (DEVELOP) were documented for classification. These land cover classes were assigned as the response variables for the Multinomial Logistics Regression. Therefore there are 9 response variable in total.

K-mean clustering

There are numerous clustering algorithms that can be used to determine the natural spectral groupings presented in the data set. One common form of clustering, called the "K-means" approach, accepts from the analyst the number of clusters to be located in the data. The algorithm then arbitrarily "seeds" or locates, that numbers of cluster centers in the multidimensional measurement space. Each pixel in the space is then assigned to the cluster whose arbitrary mean vector is closest. After all pixels have been classified in this manner, revised mean vectors for each of the clusters are computed. The revised means are then used as the basis to reclassify the image data. The procedure continues until there is no significant change in the location of class mean vectors between successive iterations of the algorithm. Once this point is reached, the analysts determine the land cover identity of each spectral class. (Lillesand M.T, Kiefer W. R, 2004). There are software that can derive the image segments into polygons such as Definiens and eCognition. However, due to the cost of these software and lack of the education about these tools, K-mean algorithm was selected to derive the spectral polygon segments.

Multinomial Logistic Regression

In statistics, multinomial logistic regression is a classification method that generalizes logistic regression to multiclass problems, i.e. with more than two possible discrete outcomes. (Green William, 1993). That is, it is a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable such as land use, land cover and forest stand structure, given a set of independent variables, which may be real-valued, binary-valued, categorical-valued, etc., such as spectral reflectance of spatial objects.

Multinomial logistic regression is known by a variety of other names, including multiclass LR, multinomial regression (Friedman, Jerome, 2010), softmax regression, multinomial logit, maximum entropy (MaxEnt) classifier, conditional maximum entropy model (ConLL, 2002).

The following statistical formula implements the multinomial logistics regression to predict the probability of the outcome of the response variables Pr (y_i = j) based on the predictor variables ($_{Xi}$) and their vectors coefficients (β_j)

$$\Pr(y_i = j) = \frac{\exp(X_i \beta_j)}{1 + \sum_{j=1}^{J} \exp(X_i \beta_j)}$$

In the context of image classifications, there will be J categories of outcome or response variables (y_i) as the discrete variables such Agriculture, Forests, Water etc. The β_j represents the vectors of coefficients of categories J. The X_i is the predictor variables such as mean reflectance of individual spectral channel and NDVI which are continuous variables. The multinomial logistic regression is applied to evaluate the probability of truly classification of spectral classes (21 spectral classes in this study) to the information classes (9 information classes in this study). In this context of multinomial logistic regression, spectral classes are the predictor variable and information classes are response variables. Spectral classes are category-valued predictors and information classes are categorical or category-values response variables.

The vglm – Vector Generalized Linear Models of VGAN (Yee, T. W. and Hastie, T. J. (2003)) Vector Generalized Linear and Addictive Models Package of R Statistics is applied to implement the models for image classification and evaluate the probability of truly classification of spectral classes to the information classes.

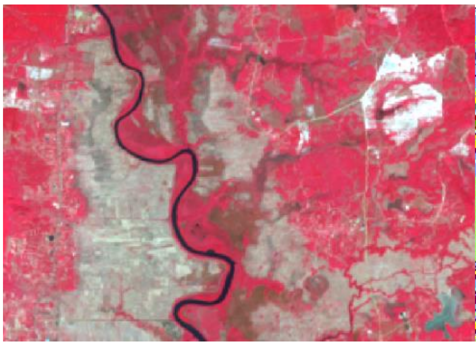
Object-based classification considers not only the individual pixels but also neighboring pixels for grouping individual pixels into information classes. With this approach, spatially contiguous pixels are usually first grouped into spectrally homogeneous objects, and then classification is conducted with objects as the minimum processing units (Yu et.al 2006). This study attempt to classify the spectral polygon segments as the objects to classify based on the means of spectral reflectance for each polygon segment using multinomial logistics regression.

Image Segmentation and Polygon Segments

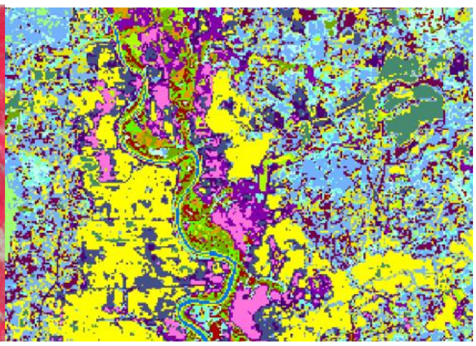
The 8 spectral layers from Landsat 8 image of December 2014 was segmented into 21 spectral Classes using K-mean Algorithms. Consequently, the local spectral variation appears as the salt and pepper effect exhibiting isolated individual pixels as a spectral class. The segments were converted into vector format.

Salt and Pepper pixels become the individual rectangular salt and pepper polygons or elongated sliver polygons. The minimum mapping unit for the study was determined as the 3 pixels by 3 pixels of the Landsat 8 image which equivalent to 90m by 90m or 8100 square meter. The minimum mapping unit is regarded as the minimum processing unit. The salt and pepper polygons and sliver polygons which area is less than or equal to minimum mapping unit is eliminated in order to reduce the local spectral variation for object based classification.

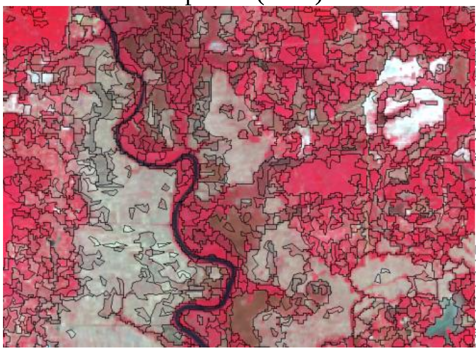
The segmented data as the polygons represents spectrally homogeneous and spatially contiguous areas as the spatial objects which area is larger than the minimum mapping unit. A small portion of the study area is illustrated for visualization of the image processing process.



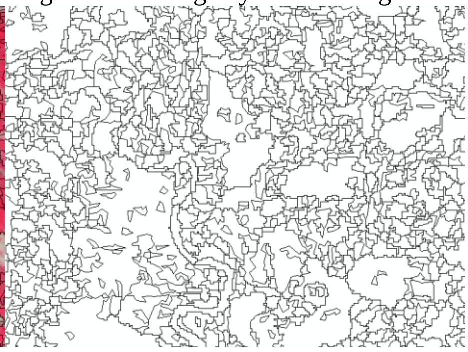
False Color Composite (FCC)



Segments of Image by K-mean Algorithm



Polygon Segments on the FCC image



Polygon Segments at Sample Area

The segmented data as the polygon represents spectrally homogeneous and spatially contiguous areas as the spatial objects. Individual segment polygon area is larger than the minimum mapping unit. Individual polygon segments have one of 1 to 21 spectral class values. Total 38978 polygon segments were derived in this study for classification.

Mean Spectral Reflectance and NDVI of Polygon Segments

Mean value of spectra reflectance and mean NDVI for each polygon segment - was derived using Zonal Statistics Algorithm. The following table exemplifies the selected 5 Polygon segments and describe the spectral class ID (CLID) and mean spectral reflectance of each polygon segments. These mean values (B1MEAN, B2MEAN, B3MEAN, B4MEAN, B5MEAN, B6MEAN and B7MEAN as the means reflectance of visible, near infrared, short wavelength infrared and NDVI) will be applied as the response variables for multinomial logistic regression modeling in the classification stage.

RecordID	CLID	B1MEAN	B2MEAN	B3MEAN	B4MEAN	B5MEAN	B6MEAN	B7MEAN
		Blue	Green	Red	NIR	SWIR1	SWIR2	NDVI
15040	16	9872.704	9573.222	9968	14647.22	14066.44	10523.667	191.889
15041	8	9245.6	8465.8	8324	11338.2	11091.8	8754	186
15042	8	9351	8669	8542.75	12045.25	10827.5	8568	188.5
15043	16	9534.308	9057.538	9103.846	13628.46	13658.39	10189.923	193.385
15044	12	9388.133	8857.433	8738.733	13964.47	11814.93	8993.067	198.167
15045	15	9262.857	8972.286	8587.714	15883	12185.86	8885.857	209.429

In this study, there are 38978 records of spatial segments with means of spectral reflectance of visible, near infrared, short wavelength infrared similar to the aforementioned table.

Field data is spatially joined to the polygon segments in order to link the field information class to the 21 spectral classes based on their locations. Moreover, sure to identify information classes such as waterbody and forest were also assigned to some polygon segments manually.

In this study, the 15045 polygon segments out of 38978 polygon segments were linked with field and interpreted information class for the classification with Multinomial Logistics Regression.

B1MEAN	B2MEAN	B3MEAN	B4MEAN	B5MEAN	B6MEAN	B7MEAN	GLUCODE	CLID
9645.92	9207.72	9535.76	14593.88	15905.8	11893.88	195.08	AG	20
9074.75	8401.25	8008.625	13698.875	12421.375	9141	203.5	DEVELOP	12
8567.465	7796.744	6706.07	15888.512	9113.023	6559.186	226.581	FOREST	6
8795.684	7961.474	7085.368	17788.895	10330.737	7226.211	230.421	MPALM	14
8873.971	8252.171	8020.257	12993.229	12726.457	9974.429	199.6	OPENFOR	8
9383.667	9086.667	9275	16201.667	15024	10960	205.333	OPENLAND	20
9151.227	8489.227	7777.409	9024.182	8734.227	7153.167	171.985	SWAMP	2
8780.429	8144.643	7277.643	16234.786	11516.714	7954.786	222.571	Y.PALM	13

The table exemplifies the field information classes (AG as Agriculture, YPALM as Young Palm, MPALM as Mature Palm, FOREST, WATER, SWAMP, OPENFOR as Opening in the forests, DEVELOP as Development activities, OPENLAND as open land) are attached to the polygon segments. The table illustrated based on selected 8 polygon segments. The 9 field information classes will be modeled as the response variables.

Preparation of data for Modeling

The predictor variables and response variables were rearranged as follow in order to assign the response variables into indicator variables (0, 1) for the spatially joined 15045 polygon segments. The following table illustrates based on selected 9 polygon segments as an example out of 15045 polygon segments. The blue font variables are response variables which are designed as the indicator variables which represent land use and land covers to be predicted. The red font variables are predictor variables which represent **mean** spectral response of each field information class.

RecordID	B1MEAN	B2MEAN	B3MEAN	B4MEAN	B5MEAN	B6MEAN	B7MEAN	AG	DEVELOP	FOREST	MPALM	OPENFOR	OPENLAND	SWAMP	WATER	YPALM
1	8665.615	7897.577	6859.385	17108.885	9962.038	6990.192	230.115	0	0	1	0	0	0	0	0	0
4843	9488.522	9085.435	8820.652	16537.13	14801.435	10594.043	210.348	0	0	0	0	0	1	0	0	0
4334	9065.176	8401.412	7881.412	13948.529	12274.471	9359.882	205.118	0	0	0	0	0	0	1	0	0
4415	8733.375	7909.813	7048.063	16866.25	10521.125	7363.375	227.438	0	0	0	1	0	0	0	0	0
4703	8997.5	8319.95	8180.2	14844.25	13354.2	9292.3	207.85	0	0	0	0	0	0	0	0	1
4792	10229.375	9672.875	9882.75	13958.375	16433.5	13821.75	188.625	0	0	0	0	1	0	0	0	0
4892	9133	8335.778	8163.889	12954.333	14051.889	10246.333	197.667	0	1	0	0	0	0	0	0	0
4993	9778.357	9333.357	9634.714	13563.643	14351.214	11103.857	188.5	1	0	0	0	0	0	0	0	0
14381	9138.192	8352.385	8026.038	10166.154	7580.192	6307.154	178.385	0	0	0	0	0	0	0	1	0

MODELING in R

The modal was implemented in R (R Development Core Team 2009) using the VGAN (Thomas W. Yee, 2008). The following line of code is important for modeling probability of land cover classes of 15045 segments.

“mysegment.vglm = vglm (cbind (AG, DEVELOP, FOREST, MPALM, OPENFOR, OPENLAND, SWAMP, WATER, YPALM) ~ B1MEAN + B2MEAN + B3MEAN + B4MEAN + B5MEAN + B6MEAN + B7MEAN, family=multinomial, data=mysegment)”

The following table illustrates the selected 16 polygon segments and the probability of classification to the particular land cover classes based on the predictor variables. In this project and paper, the land cover which has the highest probability is assigned as the true land cover to the particular segment polygon.

	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$
RecordID	AG	DEVELOP	FOREST	MPALM	OPENFOR	OPENLAND	SWAMP	WATER	YPALM
1	0.00	0.00	0.98	0.01	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.76	0.13	0.00	0.00	0.00	0.00	0.11
8	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.89	0.08	0.00	0.00	0.00	0.00	0.02
48	0.00	0.00	0.71	0.05	0.00	0.00	0.00	0.00	0.24
50	0.00	0.00	0.88	0.01	0.00	0.00	0.00	0.00	0.10
51	0.00	0.00	0.51	0.03	0.00	0.00	0.03	0.00	0.43
4482	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.98
4732	0.02	0.00	0.00	0.03	0.00	0.03	0.01	0.00	0.90
9334	0.00	0.19	0.00	0.00	0.00	0.00	0.50	0.31	0.00
9761	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
14663	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9334	0.00	0.19	0.00	0.00	0.00	0.00	0.50	0.31	0.00
14378	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00
14663	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

In this project and paper, the land cover which has the highest probability is assigned as the true land cover to the particular segment polygon.

	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$	$P(Y_i = J)$
RecordID	AG	DEVELOP	FOREST	MPALM	OPENFOR	OPENLANI	SWAMP	WATER	YPALM
1	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
48	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
50	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
51	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
4482	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
4732	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
9334	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
9761	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
14663	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9334	0.00	0.19	0.00	0.00	0.00	0.00	1.00	0.00	0.00
14378	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
14663	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

SEGID	CLID	FieldLU
1	10	FOREST
2	9	FOREST
7	13	FOREST
8	9	FOREST
13	14	FOREST
48	11	FOREST
50	11	FOREST
51	13	FOREST
4482	20	YPALM
4732	19	YPALM
9334	1	SWAMP
9761	20	YPALM
14378	2	WATER
14663	8	AG

Then calculate the percentage of matching of field data (field information classes by the GPS data collector) and predicted information classes by the multiple logistic regression. In the study, the probability of prediction from model to field information class is 0.9683616. In another word, the agreement of field information classes and model predicted information classes are 96.84%. Then all the field polygon segments (15045 segments in this study) were classified individually as the land cover type of maximum probability. Therefore, spectral class of each polygon segment has the regression derived information class. The following tables illustrates the selected 14 polygon segments out of 15045 for example purpose, with spectral class (CLID) and multinomial logistics regression derived field data based on actual field data and mean spectral values of spectral bands. Actual field data is the data that the analyst (our understanding) perceive the land cover we collected in the field. Multinomial logistic regression derived field land use data is the data that the spectral channels perceive based on the spectral signatures and analyst perceptions to the land cover.

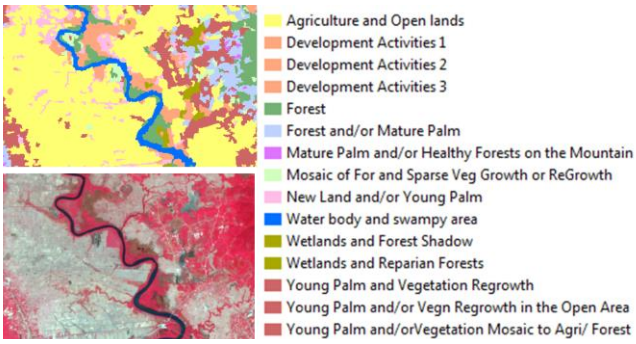
The above table (actually holds 15045 records in this project) could contain additional fields of categorical data such as stand structure or other additional exploratory variables in addition to land cover. However, in this project, it only contained spectral class ID and land cover. The probability of each land cover to spectral class will be derived for the labeling purpose using multinomial logistics regression.

The modal was implemented in R using the VGAN. The following line of code in R (as example) is important for modeling probability of nine land cover classes to 21 spectral classes.

“ltmlm<-vglm (FieldLU~ CLID, family=multinomial (), na.action=na.pass)” produce the following table that contains the spectral class (CLID) and probability of each land cover except the last column.

CLID	P_AG	P_DEVELOP	P_FOREST	P_MPALM	P_OPENFOR	P_OPENLAND	P_SWAMP	P_WATER	P_YPALM	NewLU
1	0.00	0.00	0.00	0.00	0.00	0.00	0.86	0.14	0.00	Water body and swampy area
2	0.03	0.02	0.60	0.00	0.00	0.00	0.30	0.06	0.00	Wetlands and Forest Shadow
3	0.02	0.00	0.91	0.00	0.00	0.00	0.07	0.00	0.00	Wetlands and Reparian Forests
4	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Forest
5	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Forest
6	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Forest
7	0.09	0.06	0.64	0.00	0.00	0.00	0.20	0.00	0.01	Development Activities 1
8	0.27	0.29	0.02	0.00	0.07	0.00	0.35	0.00	0.00	Development Activities 2
9	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Forest
10	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	Forest
11	0.00	0.00	0.98	0.02	0.00	0.00	0.00	0.00	0.00	Forest
12	0.18	0.14	0.03	0.00	0.02	0.00	0.34	0.00	0.28	Development Activities 3
13	0.00	0.00	0.76	0.03	0.00	0.00	0.05	0.00	0.16	Mosaic of For and Sparse Veg Growth or ReGrowth
14	0.00	0.00	0.91	0.08	0.00	0.00	0.00	0.00	0.01	Forest and/or Mature Palm
15	0.01	0.00	0.15	0.00	0.00	0.00	0.10	0.00	0.74	Young Palm and Vegetation Regrowth
16	0.41	0.03	0.00	0.00	0.08	0.00	0.21	0.00	0.26	New Land and/or Young Palm
17	0.01	0.00	0.33	0.15	0.00	0.00	0.00	0.00	0.51	Young Palm and/or Vegn Regrowth in the Open Area
18	0.00	0.00	0.55	0.44	0.00	0.00	0.00	0.00	0.01	Mature Palm and/or Healthy Forests on the Mountain
19	0.14	0.00	0.04	0.00	0.00	0.01	0.05	0.00	0.75	Young Palm and/orVegetation Mosaic to Agri/ Forest
20	0.40	0.00	0.00	0.00	0.08	0.21	0.00	0.00	0.30	Agriculture and Open lands
21	0.68	0.00	0.00	0.00	0.02	0.29	0.00	0.00	0.00	Agriculture and Open lands

The land cover labels could be developed based on the probability of different land cover to each spectral class with visual image analyses. The land cover labels were illustrated as the last column. The definition of land cover classes are labelled to the original vector segments file which contain spectral classes (21 classes in this study) and 38978 polygon segments in order to create the final land use and land cover map. The following graphics provide a small part of the study area which is classified using multinomial logistics regression in R. The result is visualized in GIS.



Future work is being carried out to add the additional layer of digital elevation models in order to differentiate forests and palms using the same procedures as describe in this paper. In the future work, vegetation structure is also considered addition to land cover classes as the response variables.

RESULT AND DISCUSSION

“Good in Good out” truly apply in the use of Geo-Informatics. The field or ground truth data quality is important for classification of remotely sensed data. The analysts recognize land cover objects by looking at them *horizontally* in the field and noted. The remote sensing sensor registers the spectra reflectance by scanning the land cover *vertically*. It is subtle but important that the perceptions of the analysts to the land cover and signatures of land cover received by the sensor is different because they are looking at spatial objects from different angles. It is important to reconcile the perceptions on the field data from the analyst point of view to image sensor point of view instead of using the field data directly as the training samples. The first part of the multinomial logistics regression provides the field data reconciliation with the probability from the analyst point of view to sensor’s spectral signature point of view based on the 15045 polygon segments. The land cover which has the highest probability is assigned as the true land cover in the field to the particular segment polygon as the remote sensing sensor recognizes. In the first part of modeling, field land cover classes and mean spectral reflectance are the response variables and predictor variables. In this paper generally 96.84% of land cover could be recognized by the sensors based on the input of the analysts. If the percentage is very low, it cautions that the quality of field data may be required to achieve the good output.

The second part of modeling is labeling of the spectral classes to the land cover classes by the multinomial logistic regression. The spectral classes (categorical valued classes) are predictor variables and land covers classes (categorical classes) as the response variables. Based on the predicted probability, visualization of image with polygon segments and interpretation capability of analysts, the land cover classes are derived for each spectral class. In this stage, the spectral discrimination capability could be evaluated and decisions to add additional layers of information such as digital elevation model and or radar data, and additional field sample data collections could be made and rerun the multinomial logistics regression models. All the spectral classes of segmented polygons are assigned in GIS based on the labels derived from the second part of the modeling.

Object-based classification considers not only the individual pixels but also neighboring pixels for grouping individual pixels into information classes. With this approach, spatially contiguous pixels are usually first grouped into spectrally homogeneous objects, and then classification is conducted with objects as the minimum processing units (Yu et.al 2006). This paper attempted to implement the object-based classification using multinomial logistics regression.

The study proved that the multinomial logistic regression could be successfully applied for object based classifications using R statistics software and VGAN package.

Additional spatial information such as digital elevation models and microwave data is recommended as the predictor variables and vegetation structure data as the additional response variables for better classification and stratification of land cover using multinomial logistics regression.

REFERENCES

"A comparison of algorithms for maximum entropy parameter estimation". Sixth Conf. on Natural Language Learning (CoNLL). 2002. pp. 49–55.

Friedman, Jerome; Hastie, Trevor; Tibshirani, Rob 2010. "Regularization paths for generalized linear models via coordinate descent". *Journal of Statistical Software* **33** (1)

Greene, William H., 1993.*Econometric Analysis*, fifth edition, Prentice Hall, pp 720-723.

Lillesand M.T, Kiefer W. R, 2004. Remote Sensing and Image Interpretation (Fifth Edition), Wiley, pp 574-575

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Thomas W. Yee, 2008. VGAM: Vector Generalized Linear and Additive Models. R package version 0.7-7. http://www.stat.auckland.ac.nz/~yee/VGAM

Yee, T. W. and Hastie, T. J., 2003. Reduced-rank vector generalized linear models. *Statistical Modelling*, **3**, 15–41.

Yee, T. W. and Wild, C. J., 1996. Vector generalized additive models. *Journal of the Royal Statistical Society, Series B, Methodological*, **58**, pp 481–493.

Yu Q, Gong P, Clinton N et. al., 2006. Object-based detail vegetation classification with airborne high spatial resolution remote sensing imagery. Photogrammetric Engineering and Remote Sensing 72:pp799-811

Thomas W. Yee, 2008. VGAM: Vector Generalized Linear and Additive Models. R package version 0.7-7. http://www.stat.auckland.ac.nz/~yee/VGAM