


Support Vector Machine algorithm optimal parameterization for change detection mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil)

Sarah Martins¹ · Nariane Bernardo¹ · Igor Ogashawara² · Enner Alcantara¹ 

Received: 7 July 2016 / Accepted: 9 July 2016 / Published online: 21 July 2016
© Springer International Publishing Switzerland 2016

Abstract Change detection in Land Use and Land Cover (LULC) using Support Vector Machines (SVM) to mapping a geographic area is a way that has been studied and improved because of its advantages as low costs in terms of computing, field research and staff team. To use SVM, it is needed firstly to define the most efficient function to be used (linear, polynomial, and radial base function—RBF) and secondly to establish the most appropriate input parameters of them, based on the study area, which was the main challenge of using SVM algorithm. The main goal of this work was to test the performance of polynomial function and RBF, and to identify which input parameters combination are the best to use SVM algorithm for Funil Hydroelectric Reservoir (FHR) sub-watershed LULC mapping, using TM/Landsat-5 time-series images. After several tests and based on the obtained results, the RBF was identified as the best SVM's function, which was used to classify the time-series images. The referred SVM function has the following parameters to be defined: the error tolerance (ξ or C), the pyramid depths (P), the radial basis function parameter (γ), and the threshold. The most effective combination of input parameters to RBF was $C = 100$; $P = 2$, $\gamma = 0.1$, threshold = 0.05. LULC change detection analyses demonstrates that the obtained SVM parameterization made the algorithm able to differentiate large and continuous classes, lengthy and thin areas, as borders, and not continuous small areas located inside

wide classes, through the usage of effective, but small, training sample. The parameterization proposed for this work to FHR sub-watershed area resulted in great statistics classification with the overall's accuracy among 86 and 98 % over the time-series, the producer's accuracy of 90 %, the user's accuracy higher than 86 %, and the Kappa statistics ranged from 86 to 91 %.

Keywords SVM parameters · Change detection · TM/Landsat-5 · Automatic classification

Introduction

For planning and management the natural resources in a sustainable way, detailed knowledge about Land Use and Land Cover (LULC) and its changes detection is considered as a key parameter to evaluate the environment progress over time (Verburg et al. 2011; Usman et al. 2015). In this context, LULC encompass two separate terminologies which are often used interchangeably: land cover means the observed biophysical cover of the Earth's surface, and land use is the human use of land resources (Dimiyati et al. 1994; Briassoulis 1999). The LULC thematic mapping from remote sensing data is commonly based on image classification techniques and the choice of an appropriate method is considered crucial for reliable mapping (Lu and Weng 2007).

Classifying remote sensing images into thematic map still a challenge mainly due to selected image data, complexity of the landscape, image processing and classification approaches (Lu and Weng 2007). A numerous classification algorithms have been developed since the first TM/Landsat-5 image was acquired in 1972 (Lu and Weng 2007; Tewkesbury et al. 2015) and among the most

✉ Sarah Martins
sarahca.martins@gmail.com

¹ Department of Cartography, São Paulo State University, Presidente Prudente, SP, Brazil

² Department of Earth Sciences, Indiana University-Purdue University Indianapolis, Indianapolis, IN, USA

popular algorithms available we can highlight the maximum likelihood classifier—MLC (Huang et al. 2002). This classifier, MLC, is a parametric method that assumes a normal distribution for the data set, what can be invalid for some cases, as to classes consisting of several subclasses or to classes with different spectral features (Kavzoglu and Reis 2008). To overcome the referred kind of limitation, non-parametric classification techniques have been introduced, such as the Support Vector Machines (SVM). SVMs are based on the statistical learning theory proposed by Vapnik (1995).

SVMs are particularly suitable for remote sensing field because this method has the ability to handle successfully with small training datasets, producing higher classification accuracy than the traditional methods. SVMs minimize classification error on unseen data without prior assumptions made on the probability distribution of the data (Mantero et al. 2005). The main challenge in using the SVM method to classify satellite images is to select the most appropriate kernel function type (Radial Basis Function—RBF, Linear, Polynomial or Sigmoid) and its parameters, which influences SVM performance. The RBF and the polynomial kernel are commonly used in the literature for the remotely sensed images classification (Pal and Mather 2005). The SVMs need user-defined parameters and each parameter has different impact on kernels

function performance. Hence, the classification accuracy of SVMs is based upon the choice of the parameters and kernels (Ustuner et al. 2015).

In order to produce a better LULC mapping, the main goal of this work was to assess which combination of SVM's function and parameterization produce the most efficient classification, and consequently, represents the best combination to map LULC in the Funil Hydroelectric Reservoir (FHR) sub-watershed. The analysis ranged from 1995 to 2010 allowing assessing the change detection to provide useful information to explain the water quality in the reservoir.

Materials and methods

Study area

The FHR is part of the Funil Hydropower system, located in the middle reaches of Paraíba do Sul River, at the frontier between São Paulo and Rio de Janeiro states, Brazil (Fig. 1). The FHR has 6.2 billion m³, and 40 km² of flooded area in its maximum volume, employed for electric power production, generating average power about 216 MW, since 1969). Moreover, at upstream there is a large industrial complex from the region called “Vale do

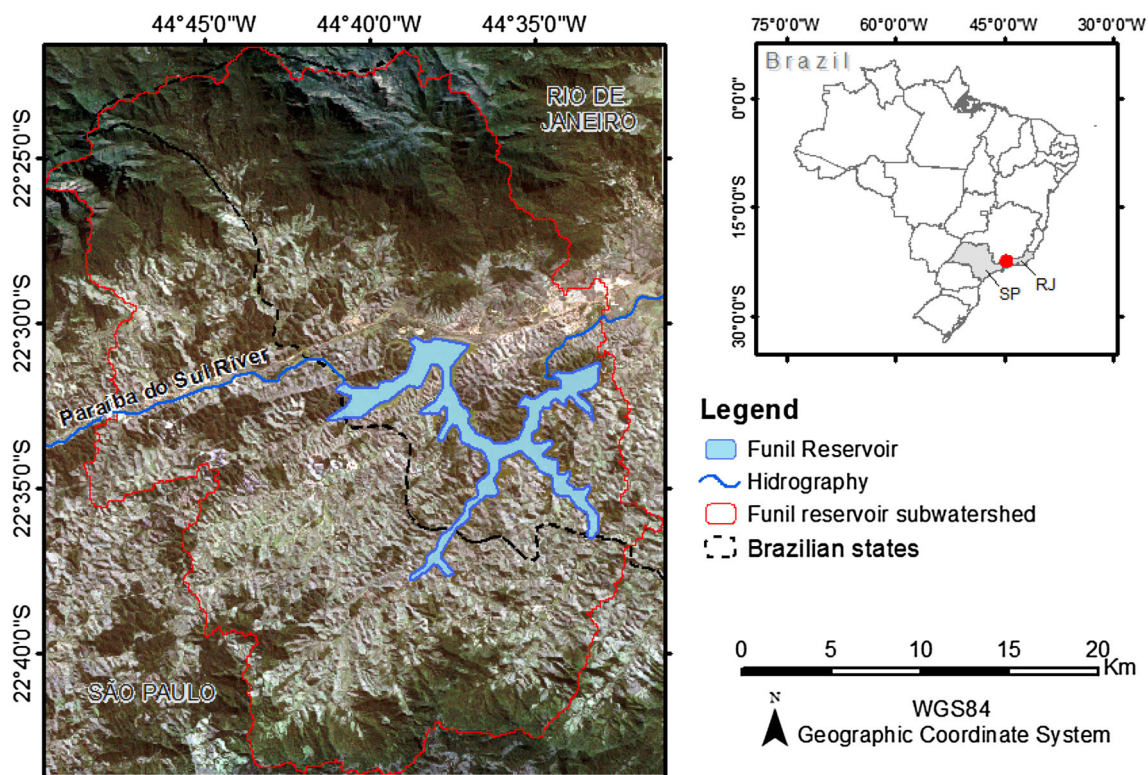


Fig. 1 Funil reservoir sub-watershed location. TM/Landsat-5 surface reflectance image, R3G2B1 color composite. Date: 08/01/2010. Path/row: 218/76

Paraíba”, at east of São Paulo Metropolitan Region (Branco et al. 2002).

Funil sub-watershed delimitation

The FHR sub-watershed delimitation was conducted using a 90 m digital elevation model (DEM) produced by Shuttle Radar Topography Mission—SRTM (RABUS et al. 2003) data, available from Consortium for Spatial Information (CGIAR-CSI 2008 - <http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp>). To process this data, the ArcHydro Tool available on ArcGIS-x was used.

Sub-watershed delimitation was based on filling DEM’s sinks to rectify the original data, and also used to define flow directions and accumulation which were used to estimate the drainage. The estimated drainage localization (streams) was defined based on slope and altitude characteristics. In addition to these data, relief characteristics (water divisor) and dam spot (bath point) were also computed to define two sub-watersheds: one for Funil reservoir and another at the reservoir immediately upstream (Walsh et al. 2015). Therefore, the FHR’ sub-watershed was clipped from the images (Fig. 2; Table 1).

Satellite images preprocessing

The TM/Landsat-5 images used in this work were acquired from the United States Geological Survey—USGS (<http://earthexplorer.usgs.gov/>). Images from

1995 to 2010, set in regular intervals of 5 years, were selected to apply SVM algorithm for LULC classification (see Table 2 for LULC class definition). TM/Landsat-5 was chosen because of its long and continuous time series of data with high quality and adequate spectral and spatial resolutions for the purposes of this work.

To ensure the high-quality of the surface information analyses it is necessary to apply atmospheric correction and normalization procedures to the images (Hadjimitsis and Clayton 2009). These procedures were made to minimize the effects on the signal by atmospheric scattering and absorption as well as differences attributed to the geometry acquisition and illumination conditions. Thus, it is essential to differentiate these interferences from the real radiometric signal from the targets. Therefore, radiometric calibration based on sensor parameters and atmospheric correction of the images can be used to obtain the surface reflectance (Schroeder et al. 2006). The atmospheric correction was conducted using Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) application; while the radiometric normalization process was made by the Iteratively Reweighted Multivariate Alteration Detection—IR-MAD (Canty et al. 2004).

FLAASH is an algorithm that uses the atmospheric compounds to establish regional models in proposing to compensate atmospheric effects in satellite spectral images by using MODerate spectral resolution atmospheric TRANsmittance version 4—MODTRAN4 (Adler-Golden et al. 1999). This algorithm was applied to convert the 1995 image to surface reflectance. This image was chosen as reference because it is the oldest in the image data set; therefore it has the closest characteristics from those originally defined for the sensor before its launching. In addition to sensor configuration parameters FLAASH set was: rural area model to estimate aerosol contribution, tropical as atmospheric model, and Kaufman over water as multispectral settings.

IR-MAD is an algorithm used to compute the radiometric normalization procedure, which has as main objective to find the pseudo invariant object in a bi-temporal images comparison. This procedure uses canonical components through sequences of MAD (Multivariate Alteration Detection) transformations until it reaches the stop criterion (Canty et al. 2004; Canty and Nielsen 2008; Canty 2009). IR-MAD was performed with surface reflectance image as a base data (1995 image), and relates it with an at-sensor reflectance data (2000, 2005 and 2020 images). The stop criteria were set as 30 interactions or no significant change in canonical correlations.

To conduct IR-MAD application, it is necessary to have a previous registration of all images to ensure that all of

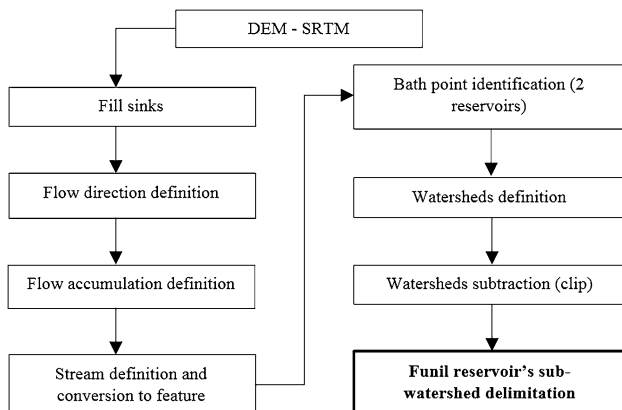


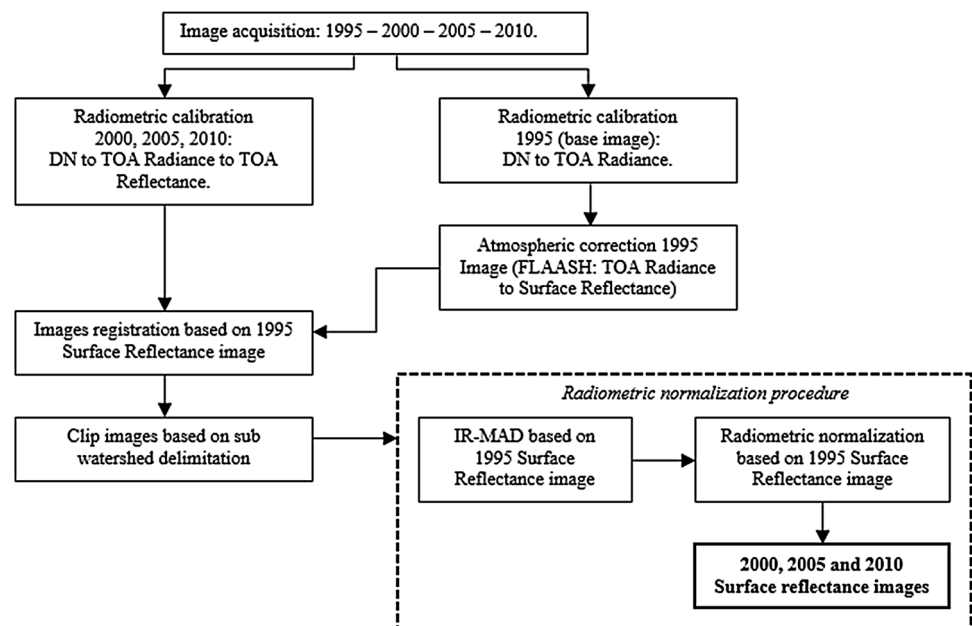
Fig. 2 FHR sub-watershed delimitation flowchart

Table 1 TM/Landsat-5 images characteristics

Image date	Path/row	Solar zenith angle	Solar azimuth angle
08/26/1995	218/076	56.019	57.195
08/21/2000	218/076	50.241	49.712
08/03/2005	218/076	53.031	43.287
08/01/2010	218/076	53.072	42.244

Table 2 Information classes used in the SVM classification, corresponding to the land cover identified in study area with their respective descriptions (adapted from UNEP/FAO 1994)

Information classes	Description
Natural and Semi-Natural Vegetation (NSNV)	Natural vegetated areas are defined as areas where the vegetative cover is in balance with the abiotic and biotic forces of its biotope. Semi-natural vegetation is defined as vegetation not planted by humans but influenced by human actions. These anthropogenic actions may result from grazing; possibly overgrazing the natural phytocenoses, or else from practices such as selective logging in a natural forest whereby the floristic composition has been changed. Previously cultivated areas which have been abandoned and where vegetation is regenerating are also included in this class
Terrestrial Primarily Non-Vegetated Areas (TPNVA)	The cover is influenced by the edaphic substratum (including bare soils)
Artificial Waterbodies (AW)	This class applies to areas those are covered by water due to the construction of artefacts such as reservoirs, canals, artificial lakes, etc. Thus without these artefacts, the area would not be covered by water
Cultivated and Managed Terrestrial Areas (CMTA)	This class refers to areas where the natural vegetation has been removed or modified and replaced by other types of vegetative cover of anthropogenic origin. This vegetation is artificial and requires human activities to maintain it in the long term. Between human activities, or before starting cropping, the surface can be temporarily without vegetative cover
Artificial Surfaces and Associated Areas (ASAA)	This class describes areas that have an artificially covered by human activities such as construction (cities, towns, transportation), extraction (open mines and quarries) or waste disposal

Fig. 3 Image preprocessing flowchart

them have a correspondence in their pixels localization, which will be essential for the change detection analyses. The maximum mean square error premised to ensure spatial correspondence among the image pixels for this registration process was 5 %. Finally, the time-series images radiometric normalization was processed based on 1995' surface reflectance image (FLAASH output), and statistical parameters (qui-square—IR-MAD output), generating surface reflectance images for all years as a result. Figure 3 presents a flowchart of the steps of the preprocessing used in this work.

Supervised classification

The SVM main framework is the structural risk minimization (SRM) which can be achieved by two pathways: (1) by fixing the empirical risk to obtain the smallest confidence interval to reduce the errors; (2) by finding an optimal hyperplane which maximizes the separation among the closest data of different classes. The data which must be considered the class boundary corresponds to the support vector, and could incorporate some error tolerance (ξ or C), like a soft margin for the information class (Huang

et al. 2002; Canty 2009; Kavzoglu and Colkesen 2009; Mountrakis et al. 2011).

SVM works considering each pixel as vector in a multidimensional system, where each class have a pixel as support vector to establish the class boundary. To perform the classification, it is necessary a training stage to allow the machine to learn the process by using examples which indicate the correspondence between a training sample and a class. The class has an unknown probability distribution which is compared to a training sample and, based on this comparison, the support vectors of each classes are defined. After the algorithm training stage, the machine is able to classify a non-training sample using the support vectors parameters. The training data are useful to define the support vector for each classes' boundaries, making statistical data unnecessary in this supervised classification (Huang et al. 2002; Canty 2009; Kavzoglu and Colkesen 2009; Mountrakis et al. 2011).

SVM was originally projected to perform separation between two linear classes and was adapted to work for multi-classes with nonlinear separation by the use of the kernelization method (Canty 2009; Mountrakis et al. 2011). For non-linear classes, there are four SVM's functions available: linear, quadratic, polynomial and radial basis function. According to Kavzoglu and Colkesen (2009) the last two functions are more efficient, presenting better performances than the others. However, it is still necessary to select one of them and the most suitable parameters for the chosen function to ensure a good performance of the algorithm. Because of that, we evaluated the best function and parameterization to be used on FHR sub-watershed.

Using the sample configuration described below, two tests were performed: (1) the first to decide what function must be used; and (2) the second to define the parameters to be used in the previously selected function. The 2010 surface reflectance image was chosen to perform the tests, because this is the image where the information classes are more fragmented. This fragmentation is ideal to evaluate the capability of the algorithm by observing if the algorithm could identify all of fragments. The information classes were based on the land cover classification system proposed by Food and Agriculture Organization—FAO(UNEP/FAO 1994; see Table 2).

To perform the tests, the information classes present in the study area were identified and one representative sample (area) of each one was taken in order to contain “pure” pixels. The training samples quantity, size, and localization were based on the premise stated by Mountrakis et al. (2011), who suggest that SVM algorithm is less

sensitive to sample size and could be able to obtain successful results using few samples.

The first test was based on Huang et al. (2002) findings which affirmed that the best results for SVM could be obtained by using the polynomial and radial basis kernel functions. These types of functions improve the results by defining complex decision boundaries among the information classes. It occurs because of the data transformation from non-linear to linear boundaries, which is useful to multi-dimensional spatial analysis (Huang et al. 2002). Additionally, these authors also presented the accuracy of the processing according to polynomial order (p) and Radial Basis Function (RBF) parameter (γ); and affirmed that the selection of parameters affects the final results.

To select the best parameters to be used in the RBF, a second evaluation was performed. For this type of kernel function, besides γ value, defined in the first test, the available parameters are: the soft margin penalization (C), corresponding to the tolerance to classification errors; the pyramid depths (P), that speed up the classification; and threshold value, which limits the perceptual of data that will be inserted in each information classes. The results of this second evaluation defined the parameters to be applied in all surface reflectance images in the multi-temporal series to obtain the LULC classification of study area using the SVM.

Land cover classification evaluation and change detection analysis

Accuracy and Kappa statistic were used to evaluate the SVM performance to land cover classification by a comparison between the classified data and regions of interest, obtained by visual analysis for each image in time-series. In the first test, one area was selected for each individual class. For the second one, it was selected a data set for each class, aiming to generate a robust output data.

Accuracy indicates the classification quality by relating the reference and the classified data in terms of the global value or in terms of its categories using the confusion matrix. Therefore, the classification result was evaluated by comparing the representativeness of the training pixels to the ones in each category (Lillesand and Kieffer 1999). While the global error is estimated by the overall accuracy; the categories errors are estimated by the producer's and the user's accuracies (Congalton 1991).

Producer's accuracy measures omissions errors, relating actual correctly classified pixels in a category with the amount of pixels classified in the same category on training

step. User’s accuracy indicates the probability of some classified pixels really correspond to the same category in the reference data (regions of interest), relating the actual correct classification pixels in a category with the same class amount of pixels classified on the actual classification (Congalton 1991).

Results and discussion

The results of the first test to find the best parameters combination function to classify the study area with the lowest error as possible are show in Fig. 4. The use of RBF function with γ values varying from 0.1, 0.3 and 0.5 showed no changes for accuracy (Fig. 4a). The use of polynomial kernel function with orders of 3, 6 and 9 showed that the best order was 6 (Fig. 4b).

Since there is no significant difference between these two functions, the lowest computational effort was chosen as the decision criteria to select the function to be used. Thus, the RBF kernel was selected to be performed in this work with $\gamma = 0.1$. The Table 3 shows the overall accuracy for the tested functions and its parameters.

The best parameters values for the chosen function (RBF) were defined in the second evaluation. Besides the γ value, the parameters available to change for this type of function were: margin penalization (C), pyramid depths (P), and threshold value. Combinations among these parameters in pairs were conducted and the best results, in terms of overall accuracy and Kappa statistic, were chosen to perform SVM classifications (Table 4).

The obtained results show that Kappa and overall accuracy are weakly influenced by variables C and the threshold. The C variable’s influence is probably related to a less restrictive classification with the increase of its value,

Table 3 Kappa and overall accuracy values obtained for both kernel functions according to their tests values

Radial basis kernel function		
γ	Kappa	Overall accuracy (%)
0.1	0.9589	97.6735
0.3	0.9589	97.6735
0.5	0.9589	97.6735
Polynomial kernel function		
p	Kappa	Overall accuracy (%)
3	0.9592	97.6886
6	0.9605	97.7639
9	0.9605	97.7639

Table 4 Kappa and accuracy values obtained in RBF parameters tests to define the values that will be used to apply SVM algorithm

γ	C	P	Threshold	Kappa	Accuracy (%)	
0.1	120	0	0.05	0.9727	98.0874	
		0	0.00	0.9727	98.0874	
		0	0.05	0.9727	98.0838	
	100	0	0.00	0.9727	98.0838	
		2	0.05	0.9844	98.9071	
		2	0.00	0.9873	99.1075	
	100 2 0.05				0.9873	99.1075
	120	2	0.00	0.9878	99.1403	

allowing a larger distance between the optimal hyperplane and the misclassified data. The threshold’s influence is presumably related to the largest percent of acceptance for the classes when its value decreases, resulting in a more restrictive classification for this value incensement. Based

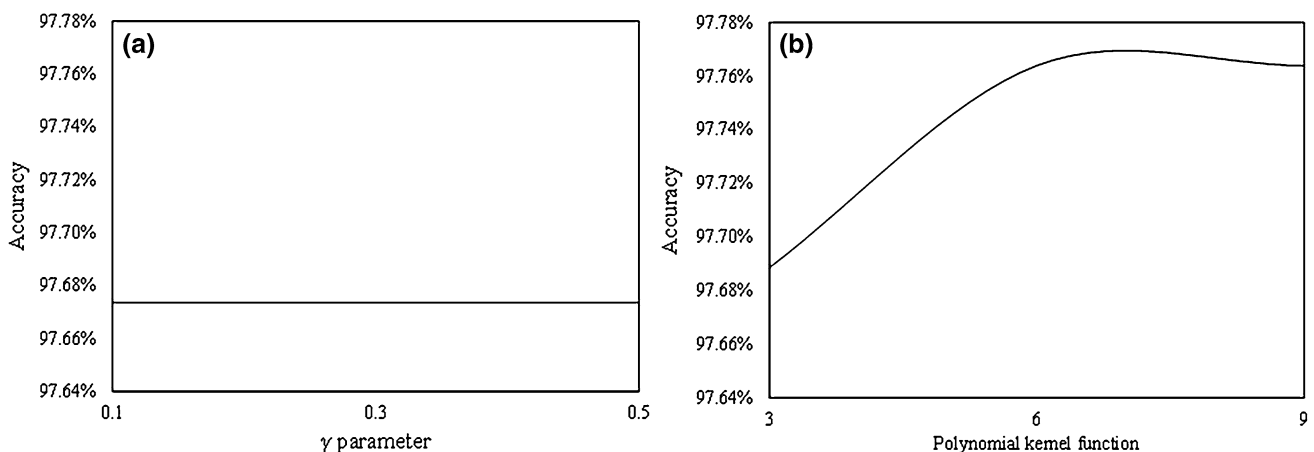


Fig. 4 Overall accuracy (%) resultant of tests between RBF function with γ values 0.1, 0.3 and 0.5 (a), and polynomial kernel function with orders 3, 6 and 9 (b)

Table 5 Kappa and overall accuracy values obtained for SVM classification per year

Year	Overall accuracy	Kappa
1995	0.98	0.96
2000	0.86	0.81
2005	0.98	0.97
2010	0.98	0.96

on these analyses, the variables values selected were the second higher Kappa and overall accuracy (bold line in Table 4). The selection for the second higher values was due to the fact that it is more restrictive than the highest one and the results were similar.

The results obtained indicated that the best BRF γ value was 0.1, because of the lack of significant improvement in the accuracy with changes in its value for the same function. Therefore, there was no relevant improvement in the Kappa and overall accuracy in a comparison between this RBF configuration and the best polynomial function parameter (Kappa difference = 0.0016 and accuracy difference = 0.0904 %). Therefore, the RBF with a γ value = 0.1 was chosen due to the lowest computational cost. The others parameters defined for the RBF were: C = 100, because it was the most restrictive value with highest accuracy; Threshold = 0.05 (acceptance = 95 %), because it improved Kappa and overall accuracy values, since it is more restrictive; and P = 2, to improve the computational performance.

Once defined the SVM parameters, the classification was performed and a confusion matrix was generated for each image classification. The overall accuracy (Table 5) showed a good performance of SVM classification, which was equal to 98 % for 1995, 2005 and 2010 images and equal to 86 % for 2000 image. These results indicate that the training sample had premised low errors in all classification categories relative to the image’s pixels amount.

Producer’s accuracy demonstrates that the SVM had a good performance to NSNV, CMTA and ASAA categories for all images, maintaining more than 90 % of accuracy for all images classification in the referred categories (Table 6).

The poorest performance was observed in the AW category for the 2000 year image with a 34.48 % of accuracy, and for TPNVA category in the 2010 image with a 76.96 % of accuracy. However, for the other images the same categories had satisfactory performances, with accuracy values higher than 80 % for AW and 90 % for TPNVA.

The poorest performance for AW category in the 2000 image classification is probably related to the low water level period in the entire system, river and reservoir, which made it difficult to collect samples from this category without any spectral interference of river-margin vegetation, margin soil, and areas with riverbed exposition, especially within the river’s channel. The classifications errors for TPNVA category in the 2010 image could be attributed to the presence of underbrush vegetation in contact to areas between TPNVA and NSNV, making it difficult to collect samples without spectral influences of the vegetation. For user’s accuracy, all results were satisfactory, showing for most of the classes more than 90 % of accuracy, the same was observed for the overall classification (Table 6).

The two, local and by category accuracies, corroborate to the overall accuracy results indicating the high consistency of SVM parameterization proposed. The complementary analysis of these classification results is the Kappa statistics that can variates from 1, representing total agreement, to 0 or negative numbers, meaning non agreement (Lillesand and Kieffer 1999). Kappa statistics results confirmed the high agreement between the training and actual classifications, indicating that some other random classification may obtain close results with this work classification. Kappa values were 0.81 for the 2000 image, 0.96 for the 1995 and 2010 images, and 0.97 for the 2005 image (Table 5).

For LULC mapping analysis, globally, it could be observed that the SVM parameterization made the algorithm sensible for variations not only large periods, as 1995 to 2010 (Table 7), also for small change, as year-by-year analysis (Tables 8 and 9).

As a SVM application result, the LULC changes between the time-series images can be accessed in Table 8. Through these results, it is possible to observe small

Table 6 Product and user’s accuracy of SVM classification based on confusion matrix

Year	Producer’s accuracy (%)					User’s accuracy (%)				
	NSNV	TPNVA	AW	CMTA	ASAA	NSNV	TPNVA	AW	CMTA	ASAA
1995	99.40	91.32	98.99	93.15	96.30	99.64	91.92	98.43	96.06	96.18
2000	91.93	95.47	34.48	99.44	96.84	93.26	76.85	99.49	85.96	96.81
2005	96.62	99.2	82.96	97.05	99.79	97.81	98.34	92.42	97.02	99.87
2010	99.98	76.96	99.08	93.67	95.89	99.93	86.66	97.77	96.89	97.51

Table 7 LULC change from first to last year (1995–2010)

	Class	1995 (%)				
		NSNTV	TPNVA	AW	CMTA	ASAA
2010 (%)	NSNTV	88.07	21.48	0.08	9.08	0.83
	TPNVA	2.03	58.46	0.01	0.71	0.38
	AW	0.02	0.01	93.79	0.55	12.29
	CMTA	9.89	20.02	5.42	89.31	57.11
	ASAA	0.01	0.03	0.70	0.36	29.39
Image difference		1.42	-25.50	11.89	3.70	-54.99

The words and values in bold represent the difference amount obtained between both images (1995 and 2010—the total period of analyse)

Table 8 LULC classification areas for each class

Year	LULC (km ²)				
	NSNV	TPNVA	AW	CMTA	ASAA
1995	328.68	49.44	28.20	433.16	12.80
2000	312.26	60.13	26.55	441.74	11.60
2005	291.52	52.95	26.32	470.20	11.28
2010	338.80	39.24	32.09	436.64	5.51

changes in LULC, around 3 % or lower (1995–2000: CMTA class = 2 %; 1995–2010: NSNTV = 1.42 and CMTA = 3.70 %, for example).

Observing the Fig. 5, is it possible to see that the obtained parameterization enables the SVM algorithm to differentiate the large and continuous classes, as well as lengthy and thin areas, as borders, and not continuous small areas located inside wide classes.

The conserved TPNVA area was around 58 %, 21.48 % of this area was converted into NSNTV, probably because of the regeneration in the borders. Around 20 % of this area was converted into CMTA. Water cover (AW), was the class with the lowest change, keeping 93.79 % of its original area; 5.42 % were changed to CMTA, probably because of the water level reduction in the river channel, exposing fertile areas which could be used for agricultural and livestock activities.

Table 9 LULC change year by year

Comparison period	LULC change (%)				
	NSNV (%)	TPNVA (%)	AW (%)	CMTA (%)	ASAA (%)
1995–2000	-5	22	-6	2	-9
2000–2005	-7	-12	-1	6	-3
2005–2010	16	-26	22	-7	-51

Agricultural and livestock class (CMTA) had maintained 89.31 % of its original area, and was mostly changed to natural areas (NSNTV—9.08 %). The anthropic areas (ASAA), kept only 29.39 % of its original area, becoming the lowest maintenance rate among all classes. The major change in ASAA class occurred in the conversion to CMTA, 57.11 %, indicating the reduction of exposed soil in rural areas, and 12.29 % of ASAA class were converted into water area (AW), probably because of the reservoir water level growth, recovering the margin area which was characterized by exposed soil. In general, all percentage lower than 1 % probably is associated to classification error real but unrepresentative land cover changes.

Conclusion

The tested SVM's parameters ensured the optimal parameterization of algorithm and allowed mapping the main changes in FHR, including the very small ones. Results demonstrated that the degree of polynomial function modifies the accuracy, and RBF showed best results when compared to polynomial function. Input parameters of RBF were also tested and shown that the highest regularization parameter ($C = 120$) did not retrieve the higher accuracy. For this study area, the higher accuracy was obtained using a lower C-value ($=100$), lower basis function parameter ($\gamma = 0.1$), to reduce the computational cost, and the pyramid depths ($=2$) to improve the performance of algorithm.

The change detection analyses demonstrates that the obtained SVM parameterization allows a great differentiation between the trained classes, supporting the premise of low sensitivity to sample size and high sensitivity to representative training sample selection. In addition to differentiation between the large and continuous classes, the parameterization obtained in this work made the algorithm able to differentiate lengthy and thin areas, as borders, and not continuous small areas located inside wide classes.

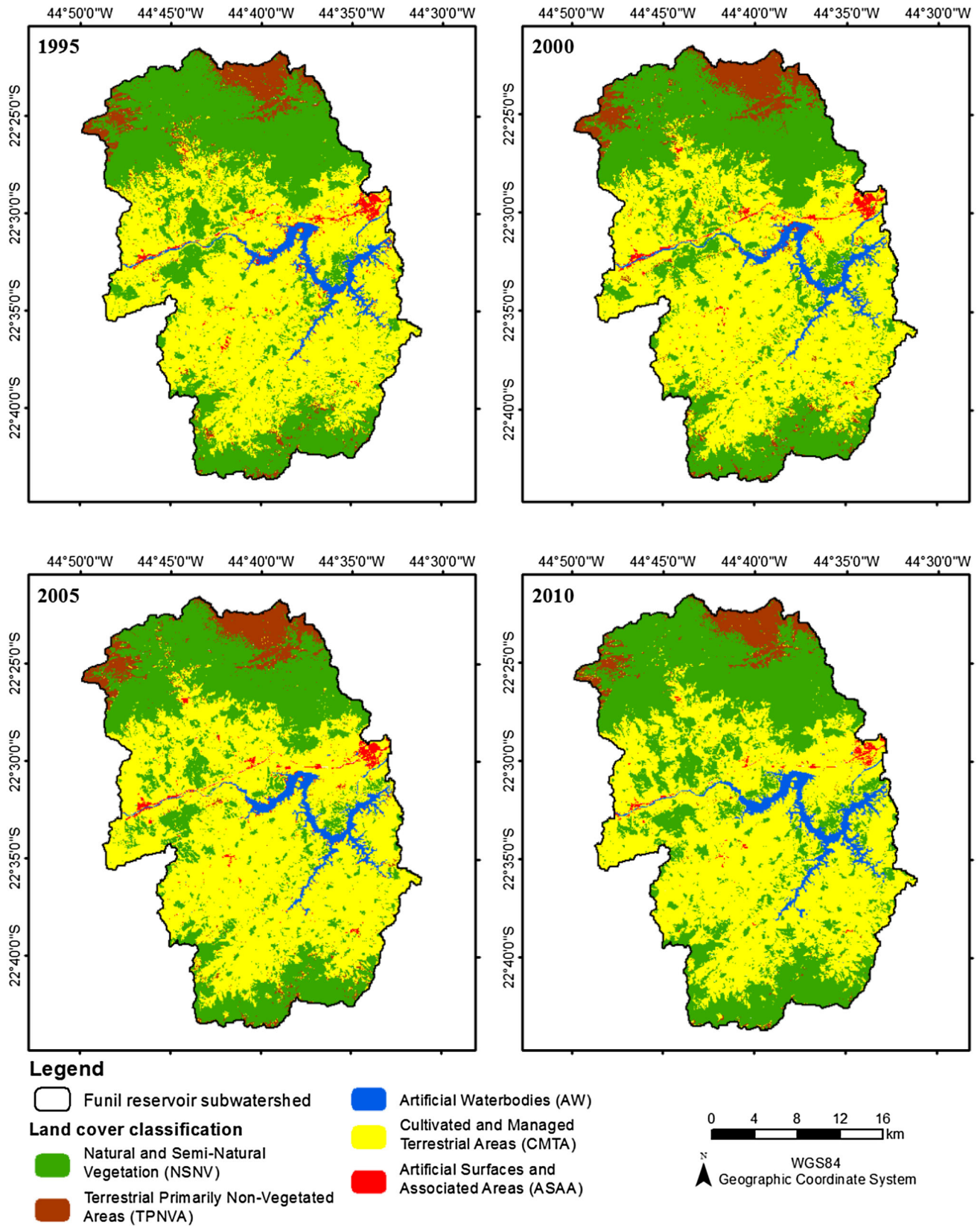


Fig. 5 LULC classification maps per year

The LULC mapping performed using the obtained SVM function and parameterization, for this multi-temporal series was efficient and provided trustful and real results, ensuring high values of accuracy, both overall (minimum 90 %), producer's (90 %) and user's (86 %), and Kappa statistics (86 % and 91 % over the time), allowing high level of confidence to LULC mapping obtained.

References

- Adler-Golden SM, Matthew MW, Bernstein LS, Levine RY, Berk A, Richtsmeier SC, Acharya PK, Anderson GP, Felde G, Gardner J, Hoke M, Jeong LS, Pukall B, Ratkowski A, Burke HH (1999) Atmospheric correction for shortwave spectral imagery based on MODTRAN4. *SPIE Proc Imaging Spectrom* 3753:61–69
- Branco CWC, Rocha M-I, Pinto GFS, Gômará GA, Filipo R (2002) Limnological features of Funil Reservoir (RJ, Brazil) and indicator properties of rotifers and cladocerans of the zooplankton community. *Lakes Reserv Res Manag* 7:87–92
- Briassoulis H (1999) Analysis of land use change: theoretical and modeling approaches. Regional Research Institute, West Virginia University, USA. Available on: <http://www.rri.wvu.edu/webbook/briassoulis/contents.htm>. Accessed Feb 2016
- Canty MJ (2009) Image analysis, classification, and change detection in remote sensing: with algorithms for ENVI/IDL, 2nd edn. CRC Press, Florida
- Canty MJ, Nielsen AA (2008) Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sens Environ* 112:1025–1036
- Canty MJ, Nielsen AA, Schmidt M (2004) Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sens Environ* 91:441–451
- CGIAR-CSI (2008) SRTM 90 m database home Available on: <http://srtm.csi.cgiar.org/SRTMdataProcessingMethodology.asp>. Accessed 15 Nov 2015
- Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens Environ* 37:35–46
- Dimiyati M, Mizuno K, Kitamura T (1994) An analysis of land use/cover change using the combination of MSS landsat and land use map: a case study in Yogyakarta, Indonesia. *Int J Remote Sens* 17:931–944
- FAO (1994) Land Cover Classification System. Available on: http://www.fao.org/docrep/003/x0596e/x0596e01f.htm#p381_40252. Accessed Feb 2016
- Hadjimitsis DG, Clayton C (2009) Assessment of temporal variations of water quality in inland water bodies using atmospheric corrected satellite remotely sensed image data. *Environ Monit Assess* 159:281–292
- Huang C, Davis LS, Townshend JRG (2002) An assessment of support vector machines for land cover classification. *Int J Remote Sens* 23:725–749
- Kavzoglu T, Colkesen I (2009) A kernel functions analysis for support vector machines for land cover classification. *Int J Appl Earth Obs Geoinf* 11:352–359
- Kavzoglu T, Reis S (2008) Performance analysis of maximum likelihood and artificial neural network classifiers for training sets with mixed pixels preview. *GISci Remote Sens* 45:330–342
- Lillesand TM, Kieffer RW (1999) Remote sensing and image interpretation, 4th edn. Wiley, USA
- Lu D, Weng Q (2007) A survey of image classification methods and techniques for improving classification performance. *Int J Remote Sens* 28:823–870
- Mantero P, Moser G, Serpico SB (2005) Partially supervised classification of remote sensing images through SVM-based probability density estimation. *IEEE Trans Geosci Remote Sens* 43:559–570
- Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. *ISPRS J Photogramm Remote Sens* 66:247–259
- Pal M, Mather PM (2005) Support vector machines for classification in remote sensing. *Int J Remote Sens* 26:1007–1011
- Rabus BM, Eineder AR, Bamler R (2003) The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *Photogramm Eng Remote Sens* 57:241–262
- Schroeder TA, Cohen WB, Song C, Canty MJ, Yang Z (2006) Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in west Oregon. *Remote Sens Environ* 103:16–26
- Tewkesbury AP, Comberb AJ, Tateb NJ, Lamba A, Fisher PF (2015) A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sens Environ* 160:1–14
- Usman M, Liedl R, Shahid MA, Abbas A (2015) Land use/land cover classification and its change detection using multi-temporal MODIS NDVI data. *J Geog Sci* 25:1479–1506
- Ustuner M, Sanli FB, Dixon B (2015) Application of support vector machines for landuse classification using high-resolution RapidEye images: a sensitivity analysis. *Eur J Remote Sens* 48:403–422
- Vapnik V (1995) The nature of statistical learning theory. Springer, New York
- Verburg PH, Neumann K, Nol L (2011) Challenges in using land use and land cover data for global change studies. *Glob Change Biol* 17:974–989
- Walsh SJ, Page PH, McKnight SA, Yao X, Morrissey TP (2015) A reservoir siting tool for North Carolina: system design & operations for screening and evaluation. *Appl Geogr* 60:139–149