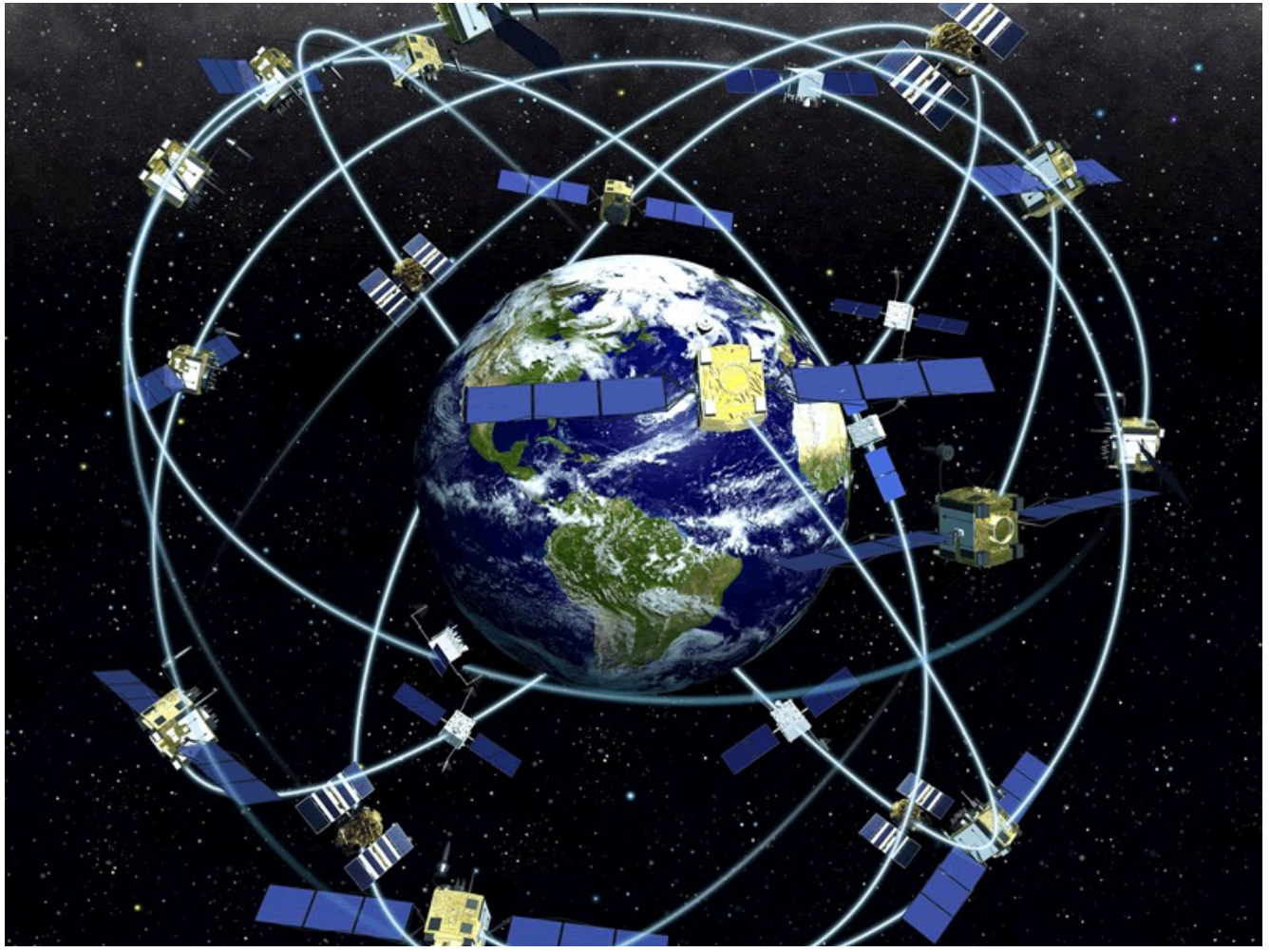
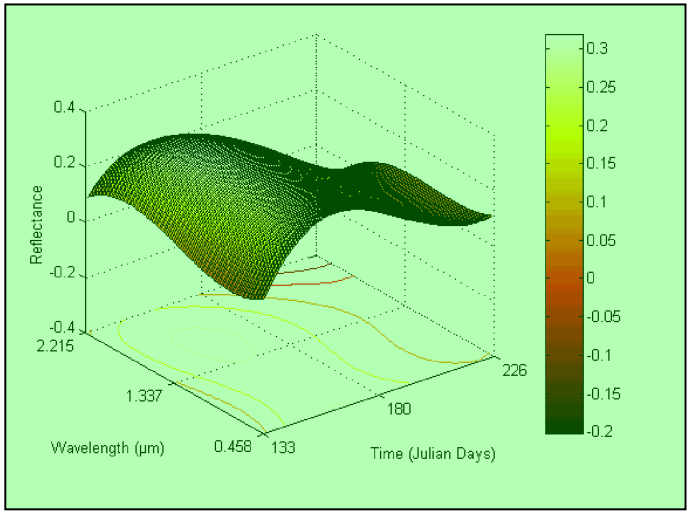
Classification of remote sensing images with GeoDMA - Geographic Data Mining Analyst

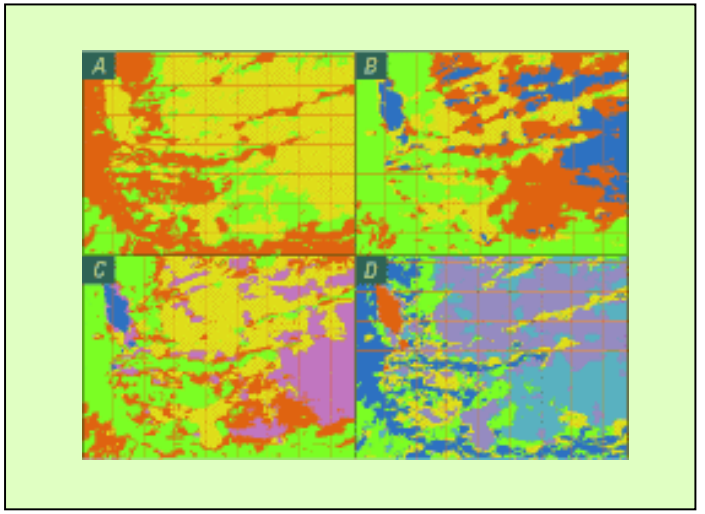
Thales Sehn Korting
tkorting@dpi.inpe.br

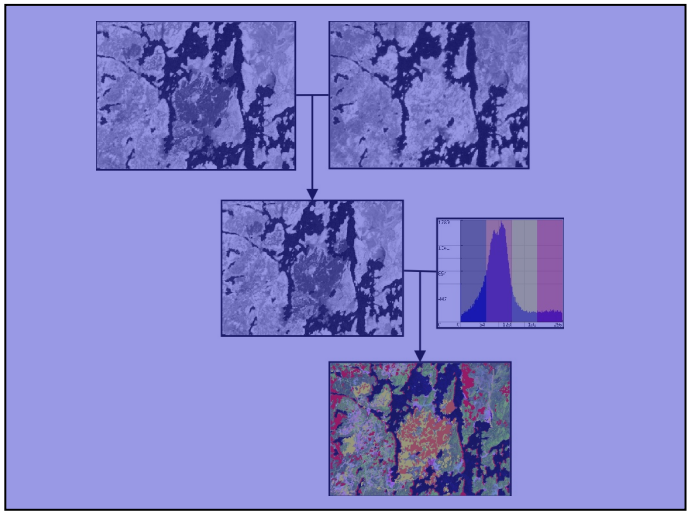
Starting with LANDSAT-1 in 1973, satellites provide a rich data set that helps us to follow changes in our planet.

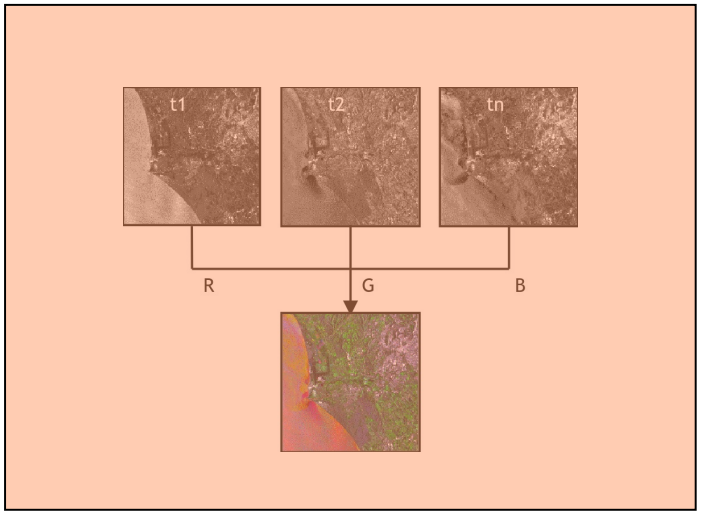


To extract information from images, we need proper image interpretation algorithms.

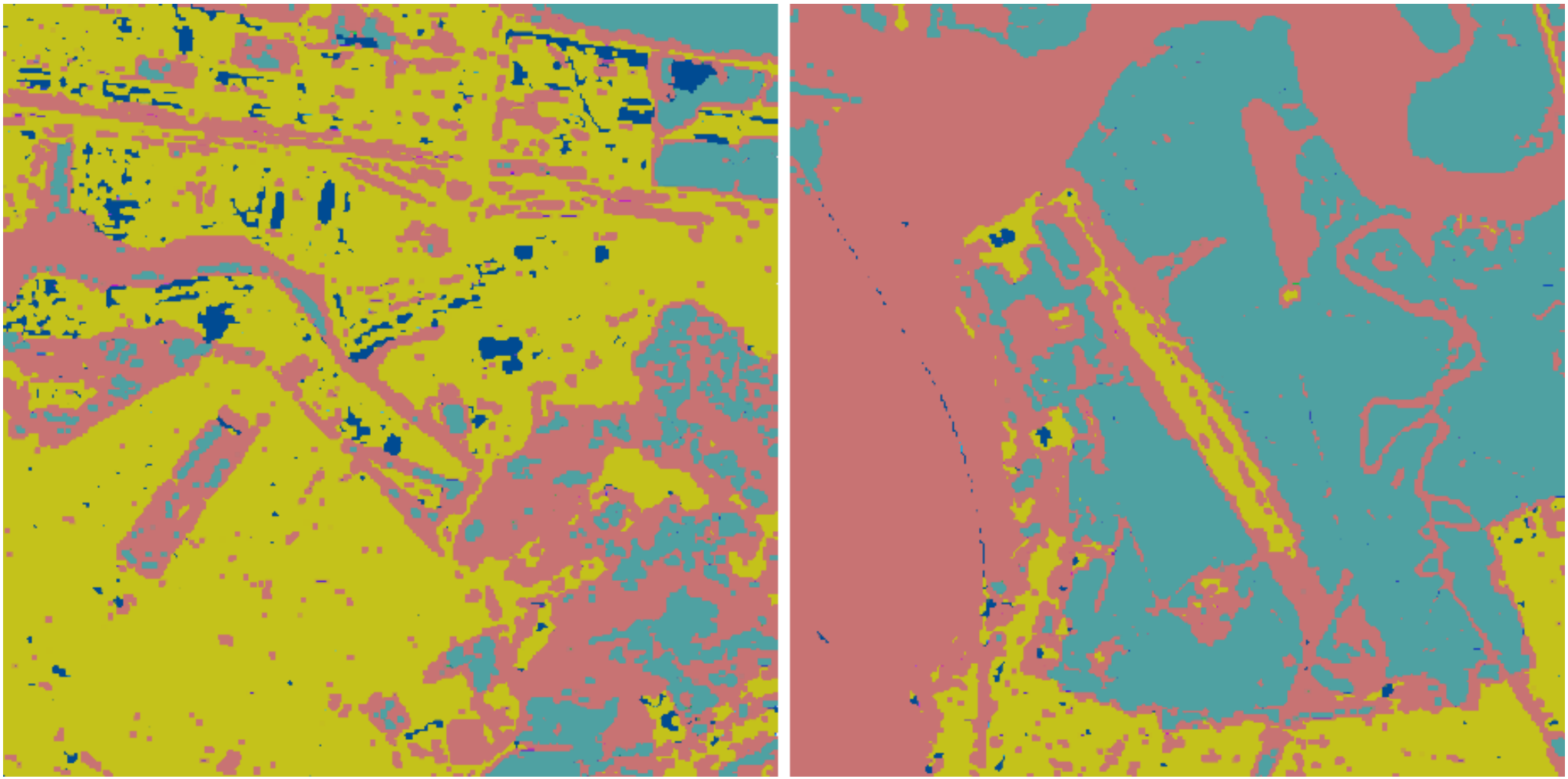








During the 1980s and 1990s, most image classification techniques used pixel-based statistical analysis.



Object-based classification took a long time to reach the mainstream users.

Haralick, R. and L. Shapiro, Image segmentation techniques. Applications of Artificial Intelligence II., **1985**. 548: p. 2-9.

Câmara, G., et al., Spring: Integrating remote sensing and GIS by object-oriented data modelling. Computers and Graphics, **1996**. 20(3): p. 395-403.

Available online at www.sciencedirect.com
ScienceDirect
Environmental Modelling & Software
Environmental Modelling & Software 23 (2008) 835–845
www.elsevier.com/locate/envsoft

Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application

A.M. Kalteh*, P. Hjorth, R. Berndtsson

Department of Water Resources Engineering, Lund University, P. O. Box 118, S-221 00 Lund, Sweden

Received 15 October 2006; received in revised form 3 October 2007; accepted 4 October 2007
Available online 19 November 2007

Abstract

The use of artificial neural networks (ANNs) in problems related to water resources has received steadily increasing interest over the last decade or so. The related method of the self-organizing map (SOM) is an unsupervised learning method to analyze, cluster, and model various types of large databases. There is, however, still a notable lack of comprehensive literature review for SOM along with training and data handling procedures, and potential applicability. Consequently, the present paper aims firstly to explain the algorithm and secondly, to review published applications with main emphasis on water resources problems in order to assess how well SOM can be used to solve a particular problem. It is concluded that SOM is a promising technique suitable to investigate, model, and control many types of water resources processes and systems. Unsupervised learning methods have not yet been tested fully in a comprehensive way within, for example water resources engineering. However, over the years, SOM has displayed a steady increase in the number of applications in water resources due to the robustness of the method.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Artificial neural networks; Self-organizing map; Review; Water resources

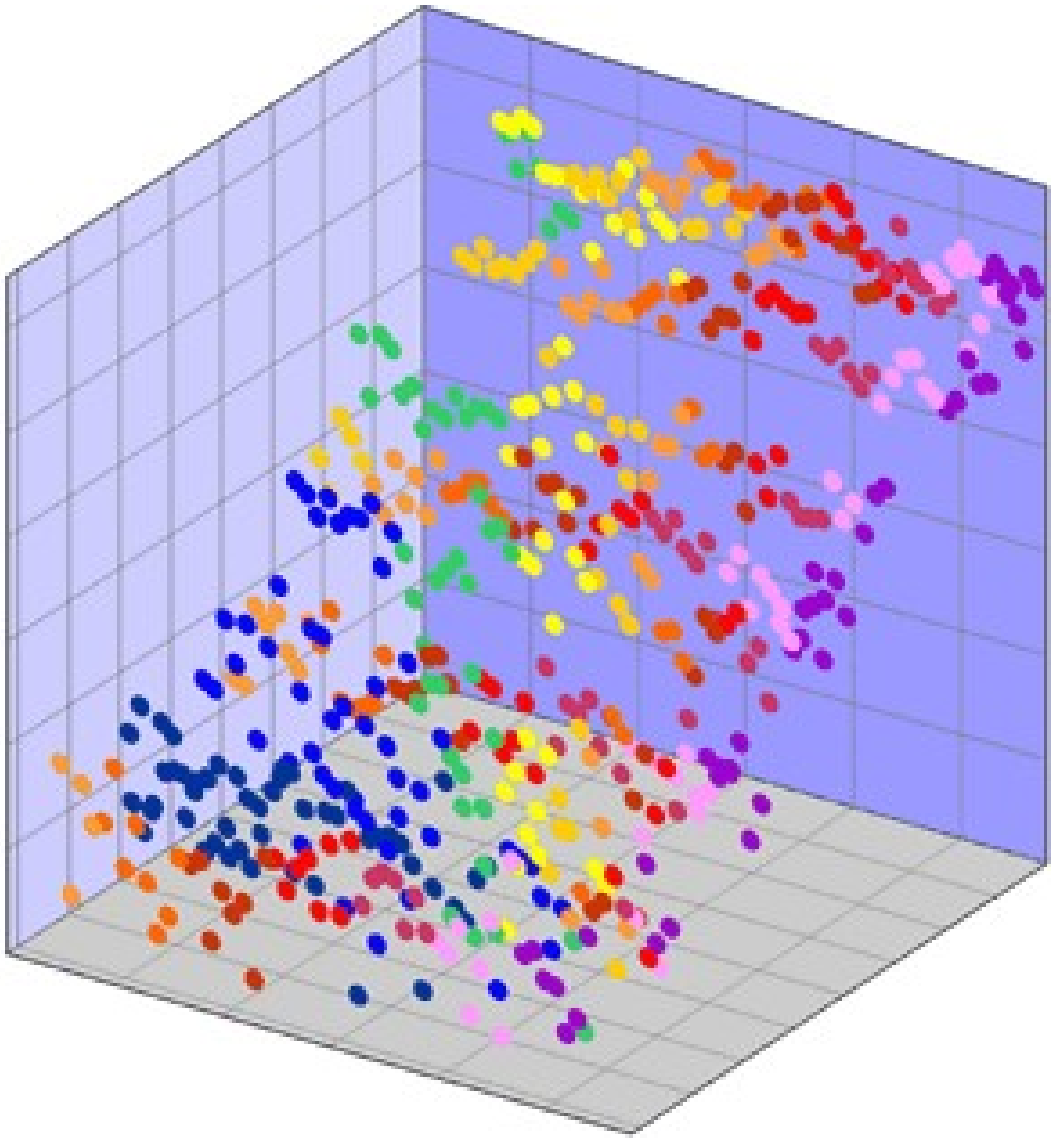
1. Introduction

Modelling of hydrological processes that are embedded with high complexity, dynamism, and non-linearity in both spatial and temporal scales is of prime importance for hydrologists and water resources engineers. In many cases, however, the lack of physical understanding of the complex processes involved creates problems to find efficient models. Over the last decades artificial neural networks (ANNs) have been subject to an increasing interest in water resources problems. This has led to a tremendous surge in research activities (ASCE, 2000b; Maier and Dandy, 2000; Dawson and Wilby, 2001; Alp and Cigizoglu, 2007; Darsono and Labadie, 2007; Iliadis and Maris, 2007; Rahaly et al., 2007). The increasing number of applications of ANNs in modelling of hydrological processes is related to their ability to relate input and output variables in complex systems without any requirement of a detailed understanding of the physics of the process involved (Dawson and Wilby, 2001). According to ASCE (2000a,b), an ANN is a massively parallel-distributed information processing system resembling biological neural networks of the human brain and capable of solving large-scale complex problems such as pattern recognition, non-linear modelling, classification, and control. The feed-forward multi-layer perceptron (MLP) is the most widely used ANN for prediction and forecasting of water resources variables (Maier and Dandy, 2000). Detailed reviews of ANNs along with assessments of their application in water resources and hydrology can be found in Maier and Dandy (2000), ASCE (2000a,b), and Dawson and Wilby (2001). The self-organizing map (SOM; also called Kohonen map or topology preserving feature map) is a kind of ANN method which is capable of clustering, classification, estimation, prediction, and data mining (Alhoniemi et al., 1999; Vesanto and Alhoniemi, 2000; Kohonen, 2001) in a wide-spread range of disciplines regarding signal recognition, organization of large collections

* Corresponding author. Present address: Department of Forestry, Faculty of Natural Resources, Gailan University, P.O. Box 1144, Sowmehe Sara, Gailan, Iran.
E-mail address: aman_mohammad.kalteh@vvl.lth.se (A.M. Kalteh).

1364-8152/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2007.10.001

How to choose suitable object features for a good classification?



GeoDMA
Geographic Data Mining Analyst

- Version 0.2.1, plugin for TerraView 4.0.0
- Linux/Windows
- Segmentation Algorithms (based on):
  - Bins et al. 1996.
  - Baatz, M.; Schäpe, A.2000.
  - Checkerboard
- Features
  - Shape
  - Texture
  - Landscape
- Classification
  - Decision Trees
  - Self-Organizing Maps
  - Neural Networks
- Validation
  - Kappa
  - Monte Carlo Simulatoion

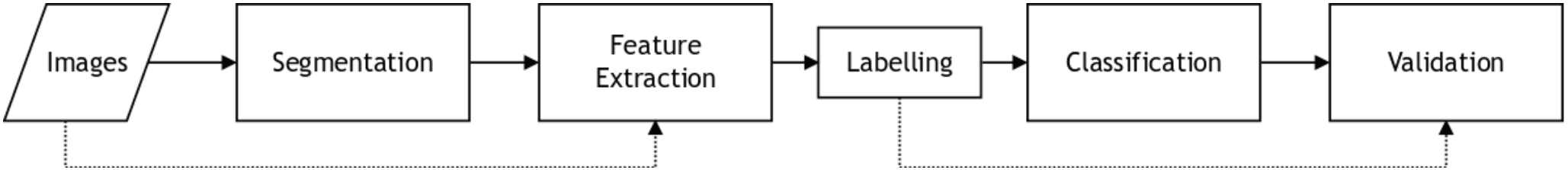
GeoDMA
Geographic Data Mining Analyst

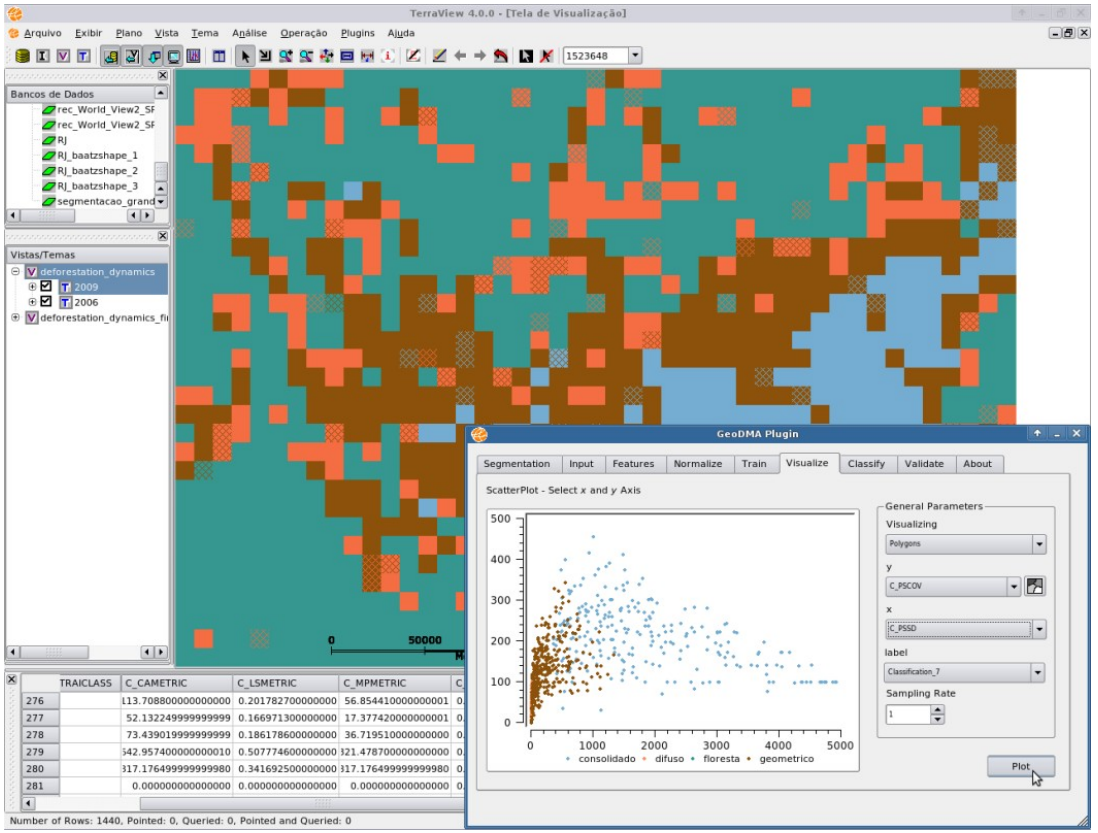
- Support and store different geographic data types in local or remote database.
- Calculate object properties including landscape ecology, spectral, and spatial features.
- Use data mining techniques to build automatically decision trees for object classification.

GeoDMA
Geographic Data Mining Analyst

- Provide a free software solution of data mining applied to geographic applications, running in different platforms.
- Integrate validation tools in a friendly Graphic User Interface (GUI), based on statistics and the users knowledge.

Outline for object-based image classification





Object's features available in GeoDMA

Patch metrics	**Landscape metrics**	**Spectral**
Angle	Area-Weighted Mean Patch Fractal Dimension	Amplitude
Area	Area-Weighted Mean Shape Index	Entropy
Bounding Box area	Class Area	Maximum
Contiguity	Edge Density	Mean
Elliptic Fit	Landscape Shape Index	Minimum
Fractal Dimension	Mean Patch Fractal Dimension	Mode
Gyration Radius	Mean Patch Size	Ratio
Length/Width	Mean Perimeter Area Ratio	Standard Deviation
Patch Density	Mean Shape Index	Sum
Perimeter	Number of Patches	
Perimeter-Area ratio	Patch Size Coefficient of Variation	
Rectangularity	Patch Size Standard Deviation	
	Percentage of Landscape	

Classification algorithms

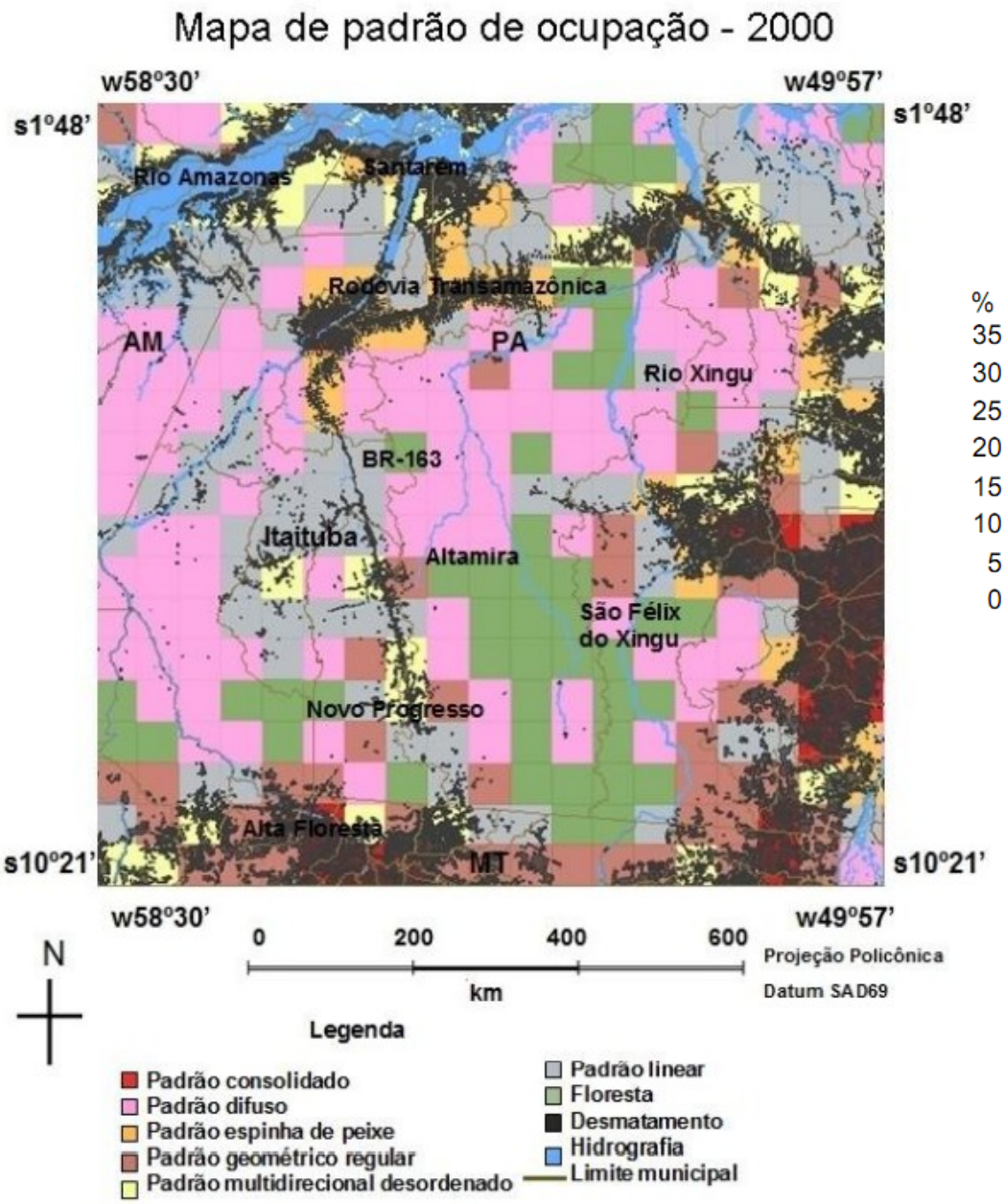
- The manual definition of an effective rule-set for classification is a time-consuming task.
- GeoDMA makes easy the knowledge discovery in the database, automatically creating decision trees, which find the best features for classification, and their relations to the classes of interest.

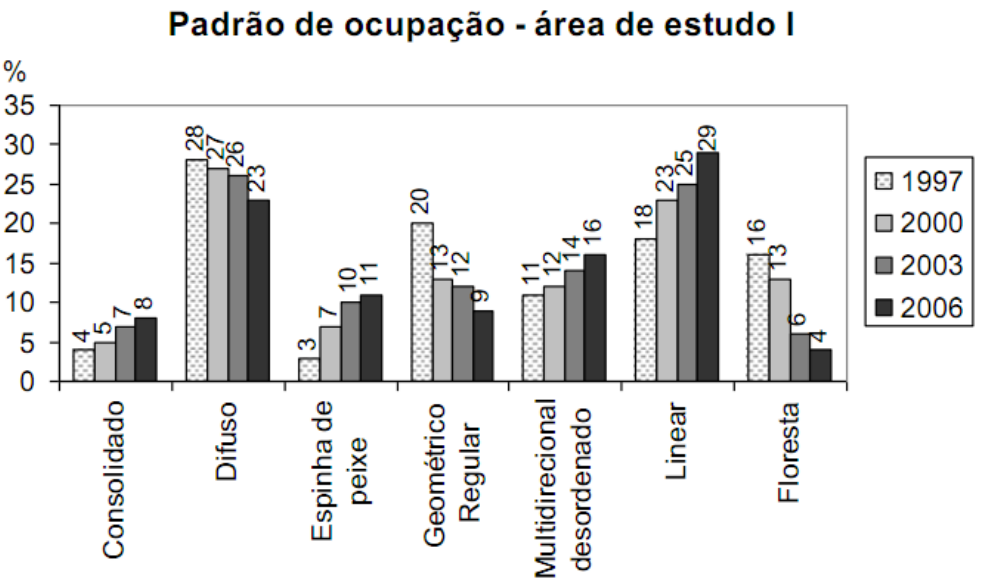
Applications – Urban classification

QuickBird scene, obtained in February, 2002, São Paulo, Brazil.

Applications – Urban classification

Applications – Deforestation Patterns





Future works

- New classification algorithms
- Feature selection user interface
- Time Series Analysis

Classification of remote sensing images with GeoDMA - Geographic Data Mining Analyst

Thales Sehn Korting
tkorting@dpi.inpe.br
http://geodma.sf.net/