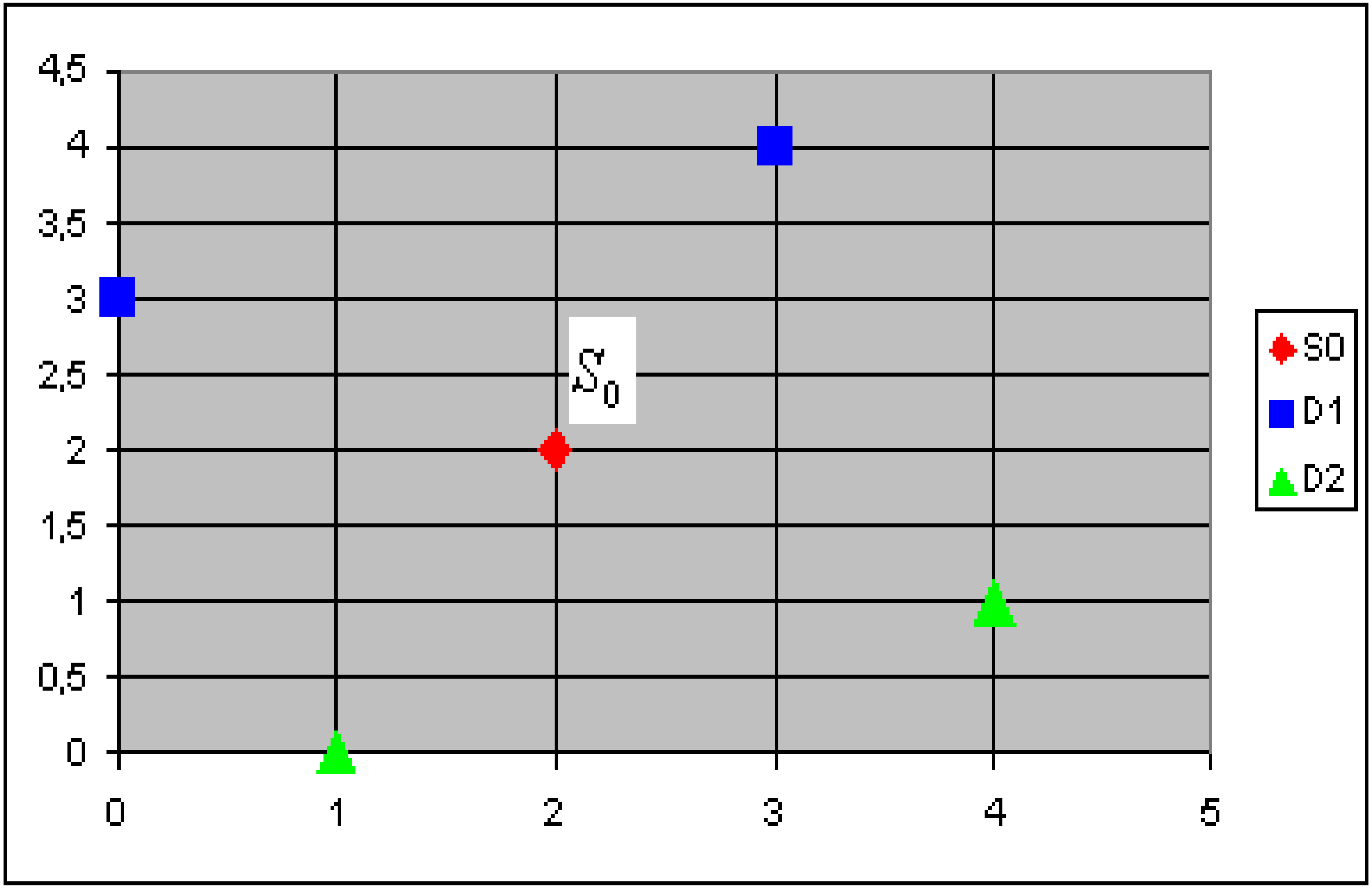
LAND COVER GEOSTATISTICAL CLASSIFICATION FOR REMOTE SENSING

- **Kęstutis Dučinskas, Lijana Stabingiene and Giedrius Stabingis**
- Department of Statistics, Klaipeda University, H. Manto 84, Klaipeda LT-92294, Lithuania

European Forum for Geostatistics Conference, Lisbon, October 12-14 ,2011

Intoduction



Relevance and the main objective

Territory is covered with the spatially correlated smoke.

Image from NASA MODIS satellite.

- Problem - non geostatistical classification methods are not efficient when territory is covered by the smoke of fire.
- Objective – to increase accuracy of land cover classification from remotely sensed imagery
- Proposed geostatistical classification technique is realised in the examples with images corrupted by spatially correlated Gaussian noise.

Introduction 1

- Switzer [1] was the first to treat classification of spatial data.
- For features based on Gaussian Markov model, the influence of texture rotation to image classification is considered by Deng and Clausi [2].
- Spatial contextual classification problems arising in geospatial domain is considered by Shekhar [3].
- Atkinson, Naser [4], [5] incorporated geostatistical information of features into plug-in versions of classifiers based on the marginal distribution of the observation to be classified.

Introduction 2

- Image classification problem is to divide an observed image into several homogeneous regions by labeling pixels, based on feature information and on information about spatial adjacency relationships with training sample.
- The stationary Gaussian Random Field (GRF) model for features and MRF model for class labels are considered.
- In the case of partial parametric uncertainty, the plug-in BDF is proposed.
- This is the generalization of the discriminant function derived in the case of training sample with fixed training sample and fixed prior probabilities for labels [6].

Introduction 3

- Geostatistical techniques that utilise spatial information in classification can be split into two distinct groups.(Atkinson,2000)
- In the first, spatial information is used to provide data on texture. It is implicit in such approaches that texture varies spatially across the image, and particularly between the classes of interest, so that data on texture can be used to inform classification (pixel-by pixel classification) .
- In the second group, spatial information is used to smooth the classified image. The rationale for smoothing is that inaccuracies that arise from simple spectral classification applied on a pixel-by-pixel basis can be reduced using the spatial dependence between neighbouring pixels. Proximate pixels are likely to be similar and this dependence can be formalised and utilised to increase classification accuracy. The goal is to choose a smoothing function based upon this spatial dependence.

Methods 1

- The marginal model of the observation Z(s) in class is

$$Z(s) = \mu_l + \varepsilon(s)$$

the error term is generated by zero-mean stationary GRF

$\{\varepsilon(s) : s \in \mathcal{D}\}$ with covariance function:

$$\operatorname{cov}(\varepsilon(s), \varepsilon(u)) = \sigma^2 r(s-u), \quad s,u \in \mathcal{D},$$

$r(s-u)$ - spatial correlation function,

σ^2 - scale parameter.

$L = \{1, 2\}$ - label set,

$S_n = \{s_i \in \mathcal{D}; i = 1, \dots, n\}$ - set of training pixels.

Methods 2

$T' = \ Z', Y'$ - training sample.

$Y = (Y(s_1), \ldots, Y(s_n))'$ - labels vector.

$Z = (Z(s_1), \ldots, Z(s_n))'$ - features vector

Assume that the *Z* model for given *Y*=*y* is

$$Z = X_y \mu + E.$$

Assumption. The conditional distribution of $Y(s_0)$ given $T=t$ depends only on $Y=y$, i.e.

$$\pi_l(y) = P(Y(s_0) = l \mid T = t), l = 1, 2.$$

Methods 3

Dependent BDF classification

BDF for the classification of Z_0 given $T=t$ (with $Z=z$, $Y=y$) under the Assumption is:

$$W_t\left(Z_0\right) = \left[Z_0 - \left(\mu_{1t}^0 + \mu_{2t}^0\right)/2\right]'\left(\mu_{1t}^0 - \mu_{2t}^0\right)/\sigma_{0t}^2 + \gamma(y)$$

where

$$\gamma(y) = \ln\left[\pi_1(y)/\pi_2(y)\right].$$

Methods 3

If Z_0 is assumed to be independent to T,

then BDF (BDFI) is implemented by Atkinson(2004,2010).

It has a following form

$$W \;\; Z_0 \;\; = \;\; Z_0 - \;\; \mu_1(s_0) + \mu_2(s_0) \;\; /2 \;\;' \;\; \mu_1(s_0) - \mu_2(s_0) \;\; /\sigma^2 + \gamma(y)$$

Methods 4

Bayes error rate for BDF

$$P_0(W_t(Z_0)) = \sum_{l=1}^{2} \pi_l(y)\Phi\;\; -\Delta_{0n}/2 + \;\; {-1}^{\,l}\gamma(y)/\Delta_{0n}$$

Bayes error rate for BDFI

$$P_0(W(Z_0)) = \sum_{l=1}^{2} \pi_l(y)\Phi\;\; -\Delta_{0}/2 + \;\; {-1}^{\,l}\gamma(y)/\Delta_{0}$$

Methods 4

PBDF to the classification problem based on BDF

$$W_t \;\; Z_0;\hat{\Psi} \;\; = \;\; Z_0 - \;\; \hat{\mu}_{1t}^{0} + \hat{\mu}_{2t}^{0} \;\; \big/ 2 \;\;\; \hat{\mu}_{1t}^{0} - \hat{\mu}_{2t}^{0} \;\; \big/ \hat{\sigma}_{0t}^{2} + \gamma(y)$$

PBDFI to the classification problem based on BDFI

(If *Z*0 is assumed to be independent to *T)*

$$W \;\; Z_0;\hat{\Psi} \;\; = \;\; Z_0 - \;\; \hat{\mu}_1 + \hat{\mu}_2 \;\; \big/ 2 \;\;\; \hat{\mu}_1 - \hat{\mu}_2 \;\; \big/ \hat{\sigma}^2 + \gamma(y)$$

Methods 5

The actual error rate for PBDF

$$P\ \ W_t\ \ Z_0,\hat{\Psi}\ \ =\sum_{l=1}^{2}\pi_l(y)\Phi\ \ \hat{Q}_l(t)$$

where

$$\hat{Q}_l(t)=\ \ -1\ ^{l}\ \ \mu_{lt}^{0}-\ \ \hat{\mu}_{1t}^{0}+\hat{\mu}_{2t}^{0}\ \ \Big/2\ \cdot\ \ \hat{\mu}_{1t}^{0}-\hat{\mu}_{2t}^{0}\ \ /\,\hat{\sigma}_{0t}^{2}+$$

$$+\gamma(y)/\sqrt{\ \ \hat{\mu}_{1t}^{0}-\hat{\mu}_{2t}^{0}\ \ ^{2}\sigma_{0t}^{2}\Big/\ \ \hat{\sigma}_{0t}^{2}\ \ ^{2}},\ l=1,2$$

Methods 5

The actual error rate for PBDFI

$$P \; W \;\; Z_0, \hat{\Psi} \quad = \sum\nolimits_{l=1}^{2} \pi_l(y) \Phi \;\; \hat{Q}_l^*(t)$$

$$\hat{Q}_l^*(t) = \; -1^{\;l} \;\; \mu_l - \;\; \hat{\mu}_1 + \hat{\mu}_2 \;\; \big/ 2 \;\; \cdot \;\; \hat{\mu}_1 - \hat{\mu}_2 \;\; / \, \hat{\sigma}^2 +$$

$$+\gamma(y) / \sqrt{\; \hat{\mu}_1 - \hat{\mu}_2 \;^{2} \, \sigma^2 \Big/ \;\; \hat{\sigma}^2 \;\;^{2}}, \quad l = 1, 2$$

Land cover classification example

•The example of land cover classification from remotely sensed image into two classes (Woodland:black; Grassland:white).

•Additive stationary GRF with isotropic exponential covariance is applied to the image.

•*Such situation can occur during fire or fog.*

•Our proposed classification methods, using different PBDF and PBDFI are applied.

Land cover classification example

- *N*0=*NN*(16) neighborhood model is used.

- Supervised classification methods are compared with unsupervised classification method using co-occurrence matrices (GLCM) by using the empirical error rates.

Forest territory classification example

Real satellite image crop

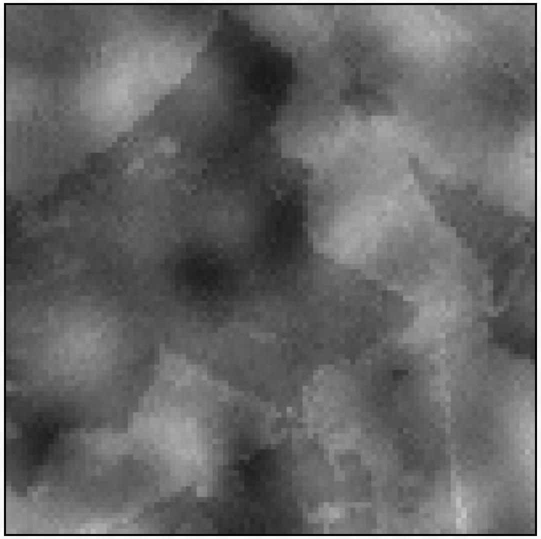
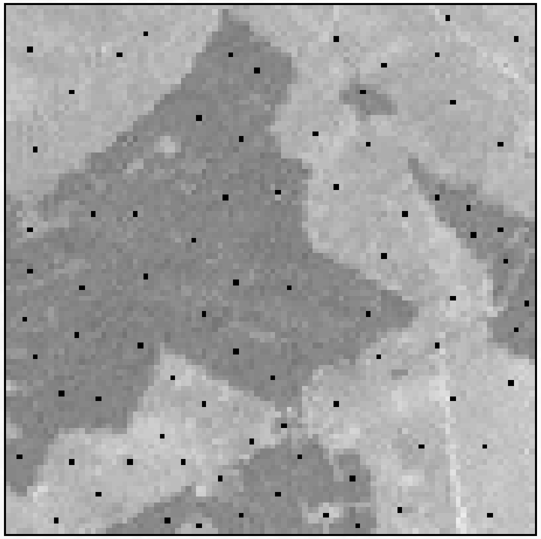
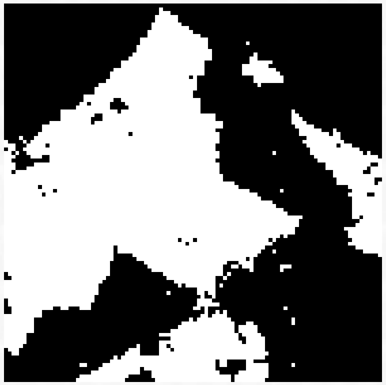


Image with spatially correlated GRF

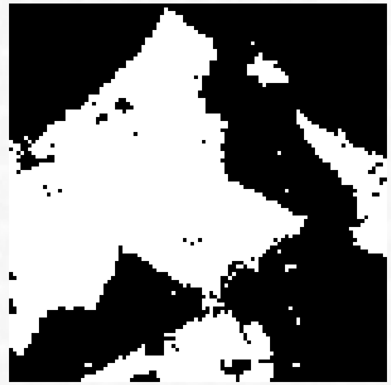


Training sample

Classification results of real image



Classification with PBDF



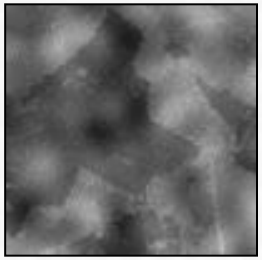
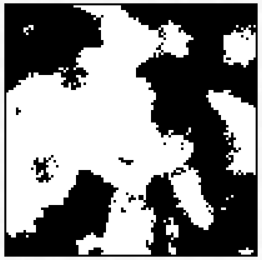
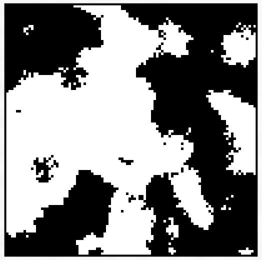
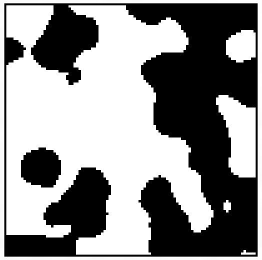
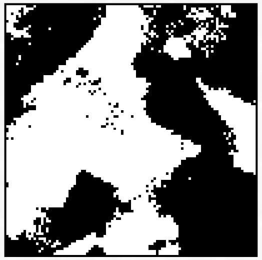
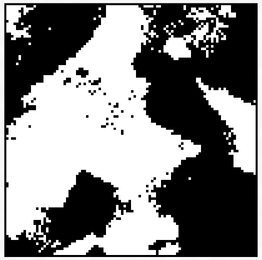
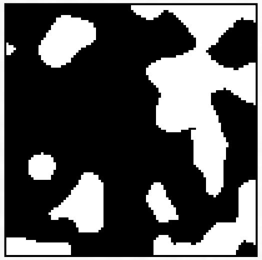
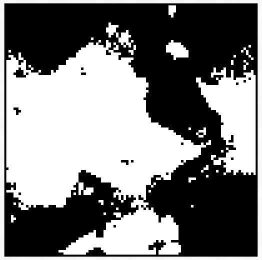
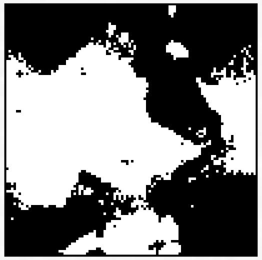
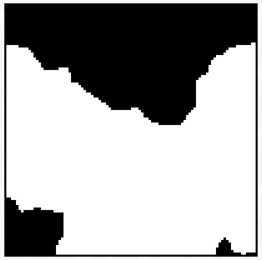
Classification with PBDFI

Classification with GLCM average

The empirically calculated errors.

DF	PBDF	PBDFI	GLCM
$\hat{P}\ 1\|2$	0.0363	0.0369	0.0502
$\hat{P}\ 2\|1$	0.0151	0.0153	0.0352

Classification results of corrupted images

α	Image for classification	Results with PBDF	Results with PBDFI	Results with GLCM average
10				
30				
50				

Classification results of corrupted images

α	Empirical errors	PBDF	PBDFI	GLCM average
10	$\hat{P}(1\|2)$	0.2174	0.2222	0.3236
	$\hat{P}(2\|1)$	0.0733	0.0770	0.1424
30	$\hat{P}(1\|2)$	0.1496	0.1540	0.9201
	$\hat{P}(2\|1)$	0.0283	0.0302	0.5865
50	$\hat{P}(1\|2)$	0.1502	0.1542	0.4771
	$\hat{P}(2\|1)$	0.0656	0.0678	0.2651

Conclusions

- The results of performed calculations give us the strong argument to encourage the users do not ignore the spatial dependence in image classification and reconstruction.
- The advantage of BDF against BDFI and the advantage of PBDF against PBDFI are shown numerically and visually in image restoration and in remotely sensed image classification examples.

References

- [1] Switzer P. Extensions of linear discriminant analysis for statistical classification of remotely sensed satellite imagery. Math. Geol 1980; 12:367–376.
- [2] Deng H, Clausi DA. Gaussian MRF rotation invariant features for image classification. IEEE Trans Pattern Anal Machine Intell 2004; 26(7):951-955.
- [3] Shekhar S, Schrater PR, Vatsavai RR, Wu W, Chawla S. Spatial contextual classification and prediction models for mining geospatial data. IEEE Trans on Multimedia 2002; 4(2):174-188.
- [4] Atkinson, P. M. (2004). Spatially weighted supervised classification for remote sensing. *International journal of applied earth observation and geoinformation*, *5* , 277–291.
- [5] Atkinson, P. M., & Naser, D. K. (2010). A geostatistically weighted k-nn classifier for remotely sensed imagery. *Geographical analysis*, *42* , 204–225.
- [6] Dučinskas K. Approximation of the expected error rate in classificattion of the gaussian random field observations. Statistics and Probability Letters 2009, 79:138–144.

Thank you for your attention.