Geostatistical stochastic simulation for spatial accuracy assessment of land cover maps derived from remotely sensed data

João D. Carneiro and Maria J. Pereira

Abstract The application of remote sensing image classification to derive land covers maps is widely used, because it is a simple and fast procedure. However, these maps are many times disregarded for land use planning and management due to the difficulty to assess accuracy, as well as the lack of reference methods to tackle the problem. Presently land cover classification accuracy assessments are based solely on the used of the confusion matrix, which is a simple cross-tabulation of the mapped class against that observed in the reference data at a set of validation pixels providing a summary of commission (type I) errors and omission (type II) errors.

Geostatistics framework is appropriate to model spatial variation of the classification uncertainty. Previous works proposed the use of indicator kriging to local varying means and sequential indicator simulation with prediction via collocated indicator cokriging. However, two main problems remain unsolved: the incorporation of distinct spatial error patterns for each thematic class due to its radiometric features and previous methodologies do not take into account patch sizes contribution to uncertainty. In the present work, these two issues are address through the use of patch size weighted spatial covariance estimation in conjunction within the framework of Direct Sequential Simulation algorithm. Early tests of the methodology applied to a segment of portuguese landscape shown promising results.

João D. Carneiro
CERENA/IST, Instituto Superior Técnico Av. Rovisco Pais 1049-001 Lisboa Portugal, e-mail: joaotrigocarneiro@gmail.com

Maria J. Pereira
CERENA/IST, Instituto Superior Técnico Av. Rovisco Pais 1049-001 Lisboa Portugal e-mail: maria.pereira@ist.utl.pt

Ninth International Geostatistics Congress, Oslo, Norway, June 11. – 15., 2012

1 Introduction

The classification of satellite images for producing land cover maps is one of the most common applications of remote sensing. As any classification model it contains errors of different types which may derive from several origins such as misregistration problems, mixed pixels, sensors properties and ground conditions and, last but not less important, the spatial stationarity assumptions on radiometric data between the training sites and the application sites. Although classification accuracy assessment is now widely accepted as a fundamental component of thematic mapping investigations it is not uncommon for map accuracy to be inadequately quantified. Presently land cover classification accuracy assessments are based on error or confusion matrix, which is a simple cross-tabulation of the mapped class against that observed in the reference data at a set of validation pixels [20, 2, 17, 18] providing a summary of commission (type I) errors and omission (type II) errors.

The widely used target accuracy of 85 % of global error can be inappropriate and that the approach to accuracy assessment commonly adopted in remote sensing is pessimistically biased [4]. The errors of classification derived from satellite images are not distributed evenly by all land cover classes: some land cover classes are easier to discriminate than others and very often the occurrence of mixed pixels of certain thematic classes difficult the labeling of those classes. Also, a certain real class tends to be misclassified in a small sub-set of land cover classes which have spectral features that are similar to the real class, and not indiscriminately in any class [21, 19]. Moreover, fragmented landscapes tend to be more difficult to classify, i.e., for a certain thematic class small patches usually tend to produce more uncertain classifications classes than more continuous patches. Consequently, the spatial distribution of classification errors is not typically but the confusion matrix and other metrics derived from it are location- independent measures, thus they dont provide any information about the spatial patterns of the error. Spatial characterization of classification errors is of prime importance [3] for the further use of classified thematic maps such as characterization of spatial uncertainty areas, evaluation of the classified themes which can be considered reliable or with the need of local field samples.

1.1 Current Geostatistical approaches to spatial accuracy assessment of land cover maps

Geostatistics framework is appropriate to model spatial variation of the classification uncertainty. [7] proposed to map local indices of classification quality derived from local (posterior) conditional distributions functions of thematic classes obtained by indicator kriging with locally varying means [12]. This method integrates in a unique step reference data (hard data) and the image classifiers posterior probability vectors (soft data). The same authors proposed the use of Sequential Indicator

Simulation (SIS) and the multi-phase classification [13, 14, 15] to evaluate the propagation of classification uncertainty on ecological model predictions, using as local mean values the users accuracy [3], which means that a constant local mean for each class was used in the conditioning process. This may be too simplistic because often there is a distinct pattern of the thematic errors over the region (classified area) due to sensors properties and/or the ground conditions. [1] and [9] used SIS with prediction via collocated indicator cokriging, which is a co-simulation method to generate stochastic realizations of thematic classes maps for updating cover type maps and that are suited for the estimation of the spatial distribution of prediction errors. The use of collocated indicator cokriging for data integration explicitly takes into account with cross-correlation between hard and soft data [5]. Consequently, the collocated cokriging estimates are potentially less influenced by sharp local contrasts in the soft data, then SIS with local varying means. [9] states that this is an advantage, but their results clearly shown that some of the thematic classes tend to disappear which in our opinion is a strong disadvantage for this type of applications. For example, the water bodies are usually very easy to classify in remote sensing images when compared with other land cover types and produce very sharp contrasts with other thematic classes, but water courses (linear) will tend to disappear when using this method for updating the land cover map, despite of the success of the classifier in discriminating them. Moreover, different thematic classes may produce distinct patterns of error and taking into account the spatial cross-correlation between hard and soft data may blur this pattern. None of the previously described methods take into account the influence of the uncertainty derived from the size of the patches.

1.2 Problem Statement

Given the shortcomings referred above (1.1), it is proposed in the current work to develop methodologies for mapping the spatial distribution of classification errors and updating land cover maps based on geostatistical stochastic simulation that takes into account the spatial continuity of each land cover class, the differences between spatial patterns of errors and the influence of uncertainty derived from the size of patches.

This is achieved mainly adopting three main ideas:

- the use of a weighted variogram (or spacial covariance) [10] in order to filter out the influence of the noise/uncertainty of patches sizes and get appropriate spatial structures of the errors
- Incorporation of the spatial patterns and patch sizes of each thematic class into the model using Poisson kriging [11, 6, 10]. This kriging method is also adapted to the use of local varying means in order to include the local LULC class (Poisson kriging with local varying means)
- The modification of indicator simple kriging (probability kriging [7]) in order to directly include a measurement error directly related to patch size in a procedure similar to the one used in block kriging in order to include block uncertainty [8].

- The incorporation of the referred kriging methods within the framework of Direct Sequential Simulation [16].

Details of the stated methods are reviewed and applied to estimate probability fields and uncertainty assessment for each thematic class in the remaining sections.

2 Used Models

2.1 Poisson Kriging Framework

The process of classifying pixel by pixel a remote sensed image can be viewed as whole as a collection of Bernoulli trials where the probability p of an error occurring is low and being the number of total image pixels typically very high. That being, the process of error occurrence in the classification procedure can be globally modeled by a Poisson process.This reasoning can also be extended in terms of random fields. It can be assumed that error occurrence for each pixel is also modeled by a Poisson distribution for each pixel. Thus being, we define for each pixel a random function of error occurrence with a Poisson distribution:

$$E\left(u\right) \sim Poi\left(\lambda\right) \tag{1}$$

where u is a spatial location within the study field D. From the relation between the Binomial and Poisson distributions, the parameter λ equals $n \times p$ where n represents the number of trials and p its probability of success (in our case, the probability of a classification error occurring). The classification error probability is assumed spatially dependent. Hence it defines its own random field:

$$p \sim P\left(u\right) \tag{2}$$

The parameter n, which represents the number of trials, can be assumed to be closely related to the number of pixels classified for a given patch belonging to a given class. Hence for each class C_i constituted by K morphologically separated patches Pa_j there will be K na_j classified pixels:

$$C_i = \left\{Pa_1 \sim na_1, Pa_2 \sim na_2, ..., Pa_j \sim na_j, ..., Pa_k \sim na_k\right\} \tag{3}$$

where the i, j indexes represent the ith class and its corresponding jth patch pixel number. The na_j values are thus used as an approximation to the true value of n for each pixel. This leads to the definition for each pixel of the following random field:

$$E\left(u\right) \sim Poi\left(\, na_{i,j} \times P\left(u\right)\right) \tag{4}$$

An error rate random variable can also be defined:

$$Z(u) = \frac{E(u)}{na_{i,j}} \tag{5}$$

Measured errors of remote sensed are usually represented by and indicator variable for each u belonging to the Domain D (the classified remote sensed image). In the present case measurement errors can be represented by indicator variables representing extreme values of the random field Z. Thus in the present case the random field Z is defined for the hard data available as:

$$Z(u_\alpha) \sim I(u_\alpha) \tag{6}$$

where $I(u)$ assumes value 1 at hard data locations where a misclassification is found, zero otherwise and is undefined elsewhere.

The used procedure then follows from this point the Poisson kriging framework sequence [6, 10], including the modeling of weighted variograms using the $na_{i,j}$ weights.

2.2 Poisson kriging with local varying means

The derivation of the Poisson kriging system [10] includes the usual kriging unbiased constrain $\sum_{\alpha=1}^{N} \lambda_\alpha = 1$. If this unbiasedness constrain is removed, a kriging system similar to simple kriging can be derived:

$$\sum_{\beta=1}^{N} \lambda_\beta C\left(u_\alpha, u_\beta\right) + \frac{\lambda_\alpha m^*}{na_\alpha} = C(u_\alpha, u_0) \tag{7}$$

Where the covariances are estimated according to the procedure outlined in [6, 10], using weighted variograms and m^* is the weighted mean [6] of the rate $Z(u)$. The estimator of $P(u_0)$ at location u_0 is thus modified as follows:

$$P(u_0) = \sum_{\alpha=1}^{N} \lambda_\alpha \left(\frac{Z(u_\alpha)}{na(u_\alpha)} - m(u_\alpha) \right) + m(u_\alpha) \tag{8}$$

where $m(u_\alpha)$ is the expected value of $P(u_0)$. In this way, the soft data provided by the classified remote sensed image can be incorporated in the Poisson Kriging framework.

2.3 Simple Kriging with hard data error incorporation

It is possible to include in the simple kriging system an error assigned to hard data. The reasoning follows the one used in incorporating block measuring error in the block kriging procedure [8]. Using simple kriging in this way. The $P(u)$ error occurrence is directly estimated together with an error term $R(u)$. The error term follows the usual conditions for block error in block kriging [8].

2.4 Accuracy Assessment: Direct Sequential Simulation

The referred kriging variants where all used in simulation studies for accuracy assessment integrated in the Direct Sequential Simulation framework [16].

3 Case sudy

The exposed metrologies are now compared with a pratical case study of a thematic classification procedure executed on satellite image of the portuguese eastern countryside. Of the total number of classes descrimitaded only four have ground truth available and of these, three remained with enough information to make a viable use of the methods under study. The relevant classes are shown with its corresponding ground reference data on figure 1 . Also needed is the total pixel population in each patch (figure 2) . Also needed in the estimation procedures are the local means calculated on soft data. That is performed using a small moving window (3x3) is used in order to reduce the weight of borderline pixels (figure 2) .

3.1 Spatial covariance modelling

Given the scarcity of available data, weighted variograms with approximately half the range of the study field, few distance lags and large angular tolerances are used. The resulting anisotropy ellipsoids are modeled using always a spherical variogram model. In figure 3 the calculated weighted variograms for the main direction are shown with their corresponding anisotropy correlogram tables used in all algorithms applied. All the algorithms are scaled to use internally (whenever required) scaling with Sill values computed in the variogram modeling stage. These are for each class 0.15, 0.30 and 0.15 (from class 1 to 3).

The weighted variogram modeling framework is part of the Poisson kriging framework.In spite of that fact, there is no reason to not using the variograms modeled in this fashion in the non-Poisson algorithms, hence they are the only ones used in the present work.

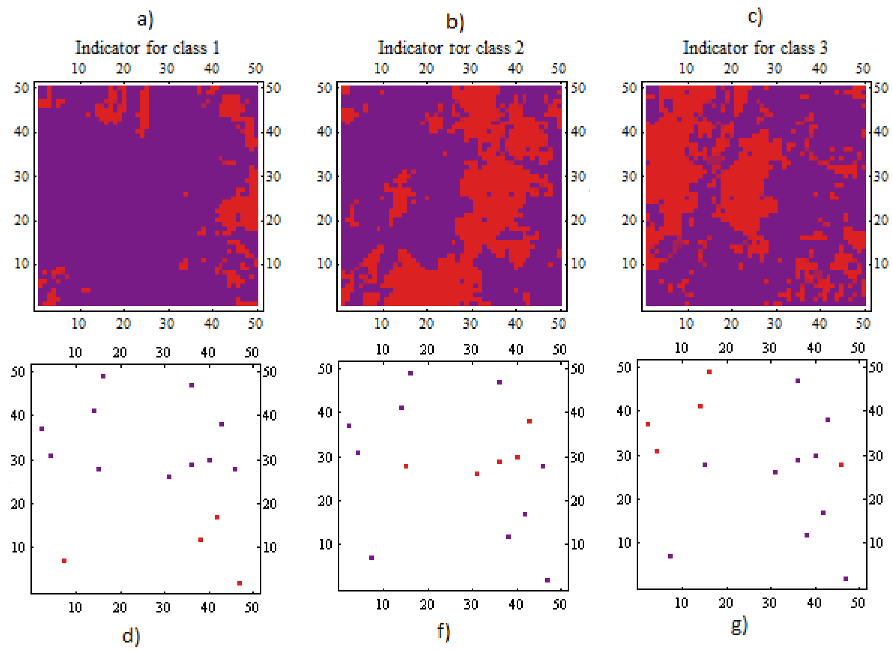

Fig. 1 Indicator coding of thematic classes derived from satellite images (top) and the corresponding ground reference data. Class 2 is used in all estimation and simulation procedures in the present work.

3.2 Estimation of class probability fields and stochastic simulation

The estimation of error probabilities yields the results the results presented is only for class 2. The three used kriging algorithms (all of them using local means derived from soft data) generally reproduce the same error patterns and very similar probability values for a given pixel. The notable exception in the present case is (figure 4) concerns the lower left corner of the estimated images. The three methods have a considerable variation in that neighborhood. That happens due to the presence of a hard data value assigned to a small area. The Poisson and weighted simple kriging algorithms take this data value into account a lot less than the simple kriging algorithm.

Contrary to the estimation images, the ones of standardized kriging variance show marked differences between them. The influence of the weighting scheme used can clearly be noticed. As such, it is expectable that the use of Direct Sequential Simulation with these methods will show significant differences among them. When inspecting the results for 40 simulation runs for each of the kriging methods (figure 5), that is indeed what happens. The methods which take into account the pixel population associated with hard data clear that influence. Noticing again the lower left corner, that influence is clearly felt: the values of the mean images for the

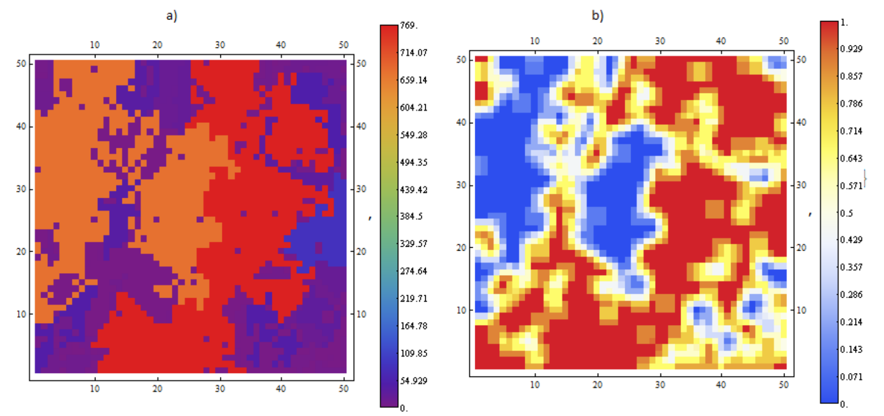

Fig. 2 Left: the values of pixel population for each identified thematic patch are shown. Zero values correspond to patches where no hard data is available. Right: A moving average of error occurrence in a 3x3 window is shown. It is used to lower the influence of border pixels in estimation procedures (presented mean is derived from class 2).

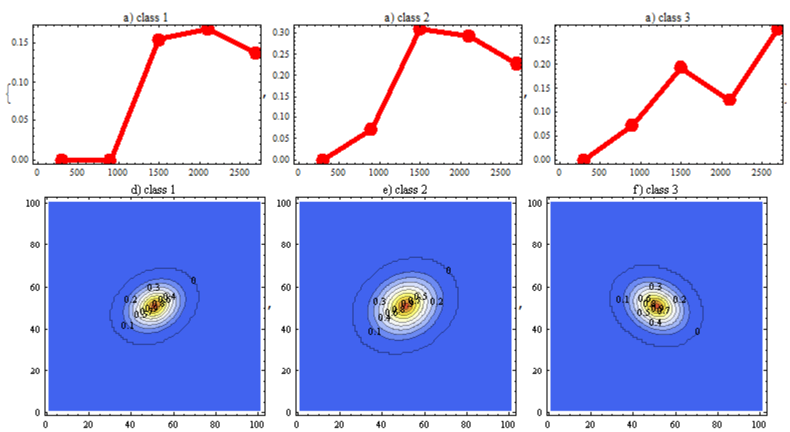

Fig. 3 Top: The experimental weighted variograms obtained for the main continuity direction for each thematic class are shown. Bottom: Correlogram tables showing the anisotropy ellipses for each class. A spherical variogram model is used.

weighted methods indicate that the error occurrence is higher and simultaneously the simulated values varied more in this area in spite of the presence of a hard data point.

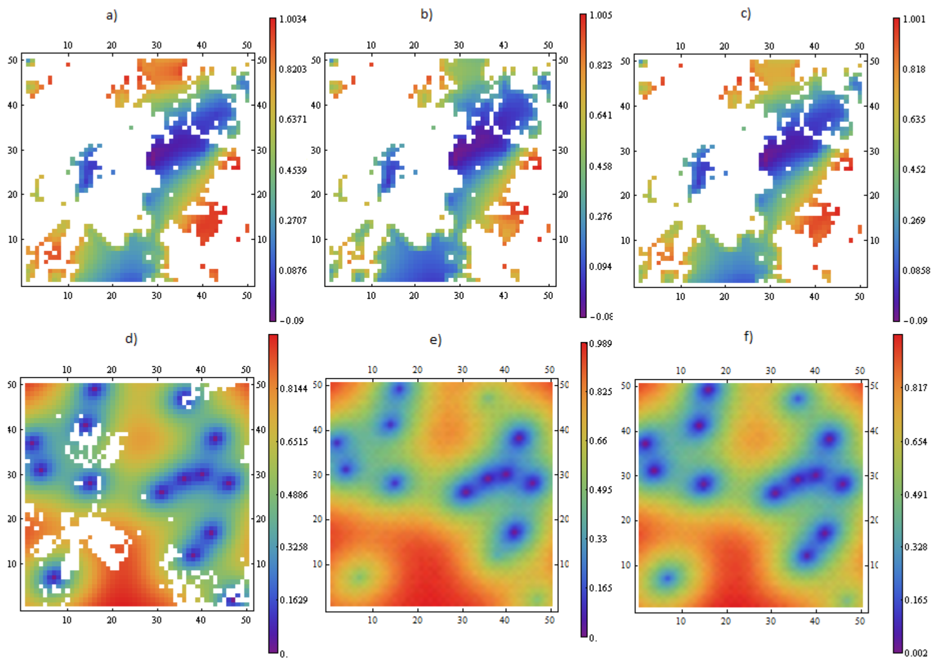

Fig. 4 kriging results and associated standard kriging variances. Simple Kriging (a and d), Weighted Simple Kriging (b and e) and Poisson kriging with local varying means (c and f). The kriging variances of patch size influenced methods show a non zero error on some hard data points.

4 Conclusions and final remarks

The proposed modifications to the kriging algorithms tested in the present work seem to be achieved to a good degree. The incorporation of patch size influence as an error source is clearly shown especially when considering the stochastic simulation results. More testing is however needed, either using synthetic or real data in order to clearly demonstrate the advantages of the proposed methodologies. The shown results are however already promising.

Future avenues are being considered as for instance the modification of Sequential Indicator Simulation in order to update class patch forms and the use of the weighted variogram modeling applied to multiphasic structures.

Also of note is the software development undertaken for this work. All software used was built from scratch and based on Wolfram's Mathematica 8.0 software. The use of a high level language is justified by the speed in which algorithm modifications can be modified continuing that fact a very acceptable trade-off with lower execution speeds. The software is freely available online at `https://sites.google.com/site/geostatmathematica/`.

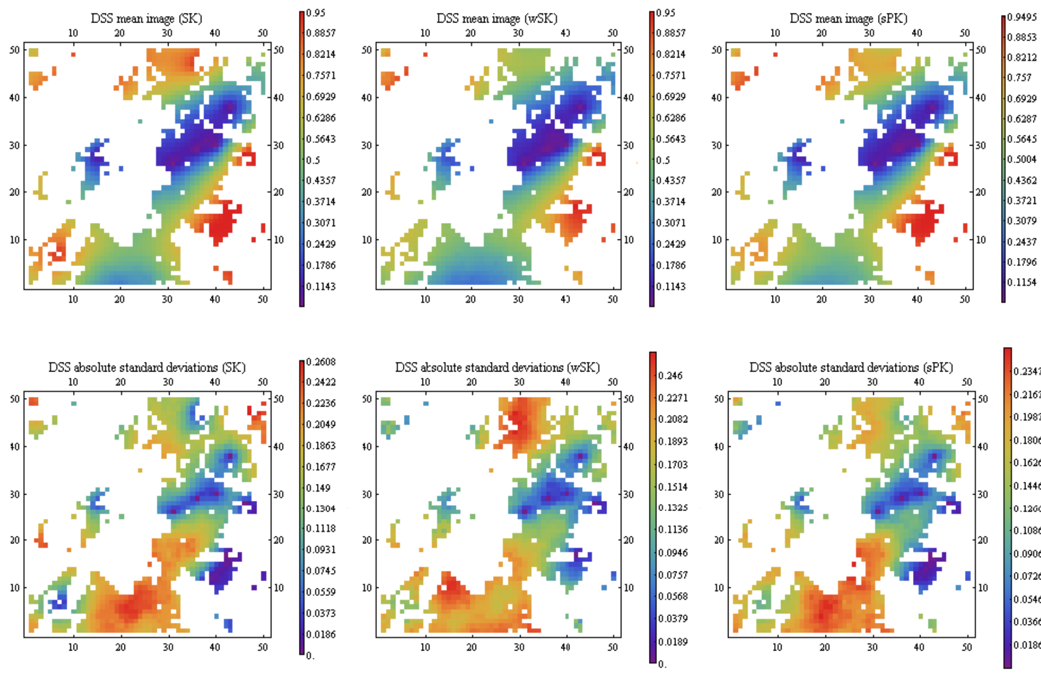

Fig. 5 Results obtained with 40 runs of Direct Sequential Simulation with Simple kriging with local varying means (SK), error weighted Simple kriging with local varying means (wSK) and Poisson kriging with local varying means.

References

1. de Bruin, S.: Predicting the Areal Extent of Land-Cover Types Using Classified Imagery and Geostatistics. Remote Sensing of Environment **74**(3), 387–396 (2000). DOI 10.1016/S0034-4257(00)00132-2. URL `http://linkinghub.elsevier.com/retrieve/pii/S0034425700001322`
2. Congalton, R.: A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment **37**(1), 35–46 (1991). DOI 10.1016/0034-4257(91)90048-B. URL `http://linkinghub.elsevier.com/retrieve/pii/003442579190048B`
3. Foody, G.: Status of land cover classification accuracy assessment. Remote Sensing of Environment **80**(1), 185–201 (2002). DOI 10.1016/S0034-4257(01)00295-4. URL `http://linkinghub.elsevier.com/retrieve/pii/S0034425701002954`
4. Foody, G.: Harshness in image classification accuracy assessment. International Journal of Remote Sensing **29**(11), 3137–3158 (2008). DOI 10.1080/01431160701442120. URL `http://www.informaworld.com/openurl?genre=article\&doi=10.1080/01431160701442120\&magic=crossref||D404A21C5BB053405B1A640AFFD44AE3`
5. Goovaerts, P.: Comparison of CoIK, IK and MIK performances for Modeling Conditional Probabilities of Categorical Variables. In: Dimitrakopoukos (ed.) Geostatistics for the Next Century, pp. 18–29. Kluwer, Dordrecht (1994)

6. Goovaerts, P.: Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. International journal of health geographics **4**, 31 (2005). DOI 10.1186/1476-072X-4-31. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1360096\&tool=pmcentrez\&rendertype=abstract`
7. Kyriakidis, P., Dungan, J.: A geostatistical approach for mapping thematic classification accuracy and evaluating the impact of inaccurate spatial data on ecological model predictions. Environmental and Ecological Statistics pp. 311–330 (2001)
8. Liu, Y.: Geostatistical Integration of Linear Coarse Scale and Fine Scale Data. phd thesis, Stanford University (2007)
9. Magnussen, S., Bruin, S.: Updating cover type maps using sequential indicator simulation. Remote Sensing of Environment **87**(2-3), 161–170 (2003). DOI 10.1016/S0034-4257(03)00138-X. URL `http://linkinghub.elsevier.com/retrieve/pii/S003442570300138X`
10. Monestiez, P., Dubroca, L., Bonnin, E., Durbec, J., Guinet, C.: Geostatistical modelling of spatial distribution of Balaenoptera physalus in the Northwestern Mediterranean Sea from sparse count data and heterogeneous observation efforts. Ecological Modelling **193**(3-4), 615–628 (2006). DOI 10.1016/j.ecolmodel.2005.08.042. URL `http://linkinghub.elsevier.com/retrieve/pii/S0304380005004436`
11. Oliver, M.a., Webster, R., Lajaunie, C., Muir, K.R., Parkes, S.E., Cameron, a.H., Stevens, M.C., Mann, J.R.: Binomial cokriging for estimating and mapping the risk of childhood cancer. IMA journal of mathematics applied in medicine and biology **15**(3), 279–97 (1998). URL `http://www.ncbi.nlm.nih.gov/pubmed/9773520`
12. Pereira, M.J., Soares, A., Rosário, L.: Characterization of forest resources with satellite SPOT images by using local models of co-regionalization. In: Geostatistics 2000 Cape Town, pp. 581–590. W.J. Kleingeld and D. Krige (eds.), South Africa, Cape Town (2000)
13. Soares, A.: Geostatistical Estimation of Orebody Geometry: Morphological Kriging. Mathematical Geology (1990)
14. Soares, A.: Geostatistical Estimation of Multi-Phase Structures. Mathematical Geology (1992)
15. Soares, A.: Short Note Sequential Indicator Simulation with Correction for Local Probabilities. Mathematical Geology **30**(6), 761–765 (1998)
16. Soares, A.: Direct Sequential Simulation and Cosimulation. Mathematical Geology **33**(8), 911–926 (2001)
17. Stehman, S.: Selecting and interpreting measures of thematic classification accuracy. Remote Sensing of Environment **62**(1), 77–89 (1997). DOI 10.1016/S0034-4257(97)00083-7. URL `http://linkinghub.elsevier.com/retrieve/pii/S0034425797000837`
18. Stehman, S., Czaplewski, R.: Design and Analysis for Thematic Map Accuracy AssessmentFundamental Principles. Remote Sensing of Environment **64**(3), 331–344 (1998). DOI 10.1016/S0034-4257(98)00010-8. URL `http://linkinghub.elsevier.com/retrieve/pii/S0034425798000108`
19. Stehman, S., Wickhamb, J., Smithc, J., Yang, L.: Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: Statistical methodology and regional results. Remote Sensing of Environment **86**(4), 500–516 (2003). DOI 10.1016/S0034-4257(03)00128-7. URL `http://linkinghub.elsevier.com/retrieve/pii/S0034425703001287`
20. Story, M., Congalton, R.G.: Accuracy Assessment : A User ' s Perspective. Society **52**(3), 397–399 (1986)
21. Zhu, Z., Yang, U., Stehman, S.V., Czaplewski, R.L.: Accuracy Assessment for the U . S . Geological Survey Regional Land-Cover Mapping Program : New York and New Jersey Region. Photogrammetric Engineering & Remote Sensing **66**(12), 1425–1435 (2000)