SADO ESTUARY MANAGEMENT AREAS: HARD VERSUS SOFT CLASSIFICATION MAPS COMPARISON

(1) Sousa, S., (2) Caeiro, S., (3) Pontius Jr, R. G. (2) Painho, M.

(1) IMAR, Depart. of Exact and Technological Sciences of the Portuguese Distant Learning University, Lisbon, (PT)
(2) ISEGI/CEGI, Institute for Statistics and Information Management - New University of Lisbon, (PT)
(3) Department of International Development, Community and Environment, Graduate School of Geography, Clark University, Worcester, (USA)

Abstract

The aim of this work is to assess the difference between three categorical maps of spatially contiguous regions of sediment structure for Sado Estuary in Portugal. These maps were computed for the same purpose but with different spatial statistics. For the map comparison fuzzy classification at different resolutions are used and compared with cell-by-cell and neighborhood hard comparison. These comparison approaches demonstrate that using either single cell, neighborhood hard or soft comparison the three estuarine management areas maps are still similar. Their major differences are mainly due to location disagreement. Advantages of using fuzzy map comparison and evaluation of agreement and disagreement components are discussed.

Introduction

In the different Geographical Information Systems (GIS) applications, and particularly in coastal zone management, compare different maps is an essential issue. The accuracy of a comparison procedure based on a more reliable and robust approach could have a marked improvement in the ability to detect a map change. Costal hydrodynamics makes it difficult to define sampling grids in exact positions and therefore a single cell-by-cell analysis comparison is less representative. Also in the cell-by-cell agreement between the two maps each cell is crisply classified, since the confusion matrix contains information about only cell-by-cell agreement. The confusion matrix fails to distinguish between a near miss and a far miss. In other words, the confusion matrix records zero agreement when a cell is not classified correctly, even when the correct category is found in the neighbouring cell, or even when the correct category is found nowhere near the cell (Pontius, 2002 and Pontius and Suedmeyer, 2003).

Therefore, a neighbourhood cell comparison is more appropriate. Using the neighbourhood to compare categorical maps could be computed using a hard or fuzzy classification. Hard classification has the disadvantage of modifying the maps before the comparison. After hardening, there could be a substantial change in the quantity of each category, leading to errors and misleading results. By applying fuzzy classification for the comparison of categorical maps it is possible to obtain a special and gradual analysis of the similarity of two maps. Also, it would be helpful to have on that soft comparison an analytical technique that allocates the sources of agreement and disagreement indicating in what respects the comparison map is strong and weak.

Within GIS usually the map comparison statistics are used mainly for measuring the goodness-of-fit of simulation land-change models (e.g. Pontius, 2000, Hagen, 2002, Pontius, 2002 and Pontius and Schneider, 2001) and not to evaluate differences between spatial patterns models of regions with very dynamic characteristics like estuaries.

The team has been working on the development of an environmental data management system through sediment quality assessment for the Sado Estuary (EMSado) in the south of Portugal (Caeiro *et al*., 2002). The units of this management system are spatially contiguous and homogenous regions (management areas). To delineate these management areas three maps were computed using multivariate geostatistical tools. A great agreement of similarities will further support the choice of any of the methods as appropriate for environmental management, and hence the less significance of choosing one of the methods. The aim of this work is to assess the difference between the three maps in which the cells are fuzzy classified, and to separate sources of agreement due to quantity and location. This article comes in the sequence of two other where cell-by-cell comparison and hard neighbourhood classification were computed and results discussed (Sousa *et al.* 2002 and Sousa *et al.,* unpublished) . In this work we want also to compare the fuzzy map comparison with this earlier comparison methods.

Methods

Previous work

In order to divide the Sado Estuary in homogenous areas for future environmental management of this ecosystem, geostatistical multivariate techniques were used. Three maps of final homogenous areas were computed from three sediment characterization indicators, using: Map 1) cluster analysis of dissimilarity matrix function of geographical separation followed by indicator kriging of the cluster data, Map 2) discriminant analysis of kriged values of the three sediment attributes, Map 3) combination of methods 1 and 2. (fig. 1). In each of these categorical maps four organic matter contents categories were computed: 1- for high organic load, 2- for medium high, 3- for medium, and 4- for low organic load. Results of Map 1 seem to be in better agreement with estuary behavior, assessment of contamination sources and previous work conducted at this site (Caeiro *et al.*, 2003). For that reason, Map 1 was considered the reference for the comparison between Map 1 and Map 2. For comparison between Map 2 and 3, Map 2 was considered the reference since Map 3 results are from a refinement of Map 2 using data from Map 1. For these same reasons, Map 1 is considered the reference in the comparison between Map 1 and 3.

Visual map overlays were used either for single cell, or neighborhood sizes (3, 5, 7, 9, 11, 13, 15 and 29) using hard data map comparison. The last neighborhood (29) was used to evaluate the sensitivity of this approach. At the finest resolution, each cell is 100-by-100 meters. A two-step process converts the fine-resolution cells to coarse hard-classified cells. For the first step, the size of the coarse cells is determined by aggregating several fine resolution cells. The resolution of the coarse cell is expressed as a multiple of the length of the side of a fine resolution cell. For example, a neighborhood size of 3 means that a 3-by-3 block of fine resolution pixels are aggregated to form one coarse cell. For the second step, a single category is assigned to the coarse cell, based on the majority category among the fine-resolution cells that constitute the coarse cell. Using this neighborhood hard comparison each location is a mode function of the input cells of different neighborhood sizes, instead of a single input cell-by-cell comparison. In both comparisons map algebra and contingency tables were used to obtain the difference between each of the two maps and create a classification of their differences. For quantification of map comparison approaches, Kappa statistics (*kstandard, Klocation* to evaluate location errors and *kh*isto to evaluate quantity errors) and agreement space were used (Sousa *et al.* 2002 and Sousa *et al.,* unpublished).

Fuzzy comparison

For computing fuzzy map comparison the module VALIDATE in Idrisi Kilimanjaro® GIS software was used. The module computes statistics for different resolutions (i.e. length of a fine grid cell size) starting from the resolution of the raw data (finest resolution) to a very coarse resolution. An arithmetic sequence was used to create the aggregating neighboring cells into an increasing coarse grid (from 3 to 29 grid cells). We computed until the grid-cell size of 29 to allow comparing with the previous work.

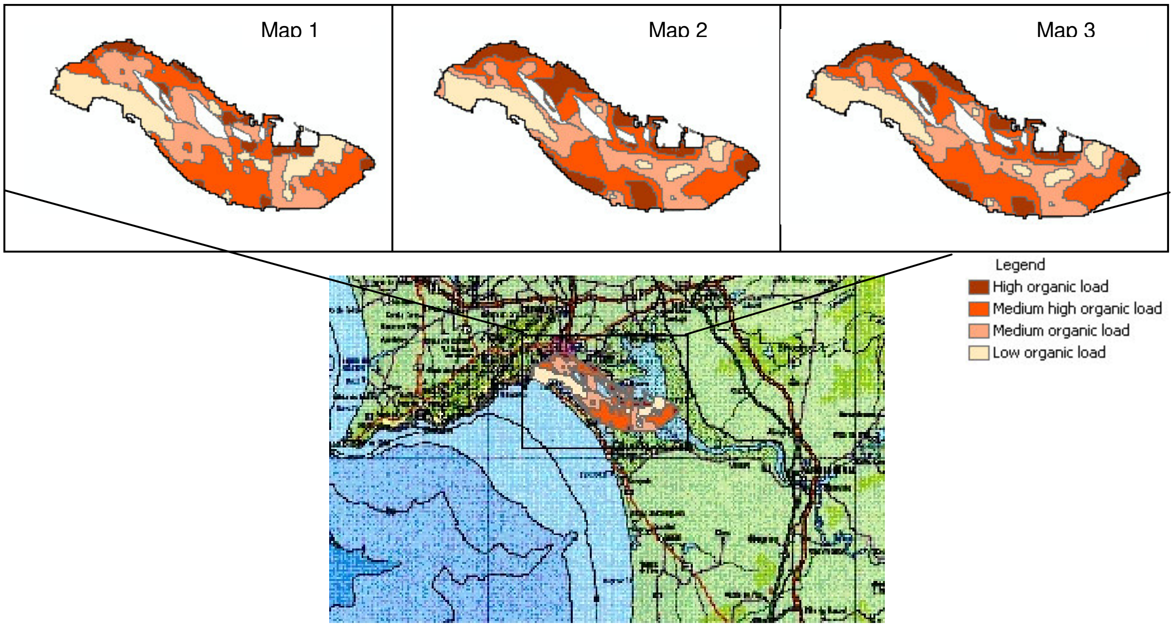

Figure 1 – Study area and maps representing the 3 methods for Sado estuary management area delineation (Adapted from Sousa *et al.*, 2002).

For maps with one single strata/sub-region VALIDATE computes five especially important numbers that constitute the basis for the components of agreement and disagreement between the reference map and other maps that have increasingly accurate information (from no (n), to medium (m) and perfect information(p)). They are denoted as *N(n)*, *N(m)*, *M(m)* (components of agreement) and , *P(m)* and *P(p)* (components of disagreement). VALIDATE computes these statistics for each resolution. Each cell have partial membership in any of the categories, and the agreement for category *j* in cell *n* is to be minimum of proportion of category *j* in grid cell *n* of Map *M* (*Mn, j*) and proportion of category *j* in grid cell *n* of Map M' (*Mn, j*). Figure 2 gives the mathematical definition for each expression. For *N(n)*, each cell of the other map is the same and has a membership in each category equal to 1/J. For *N(m)*, each cell of the other map is the same and has a membership in each category equal to the proportion of that category in the comparison map. *M(m)* denotes the proportion correct between the reference map and the comparison map. For *P(m)*, the other map is the comparison map with the locations of the grid swapped anywhere within the map, so as to have the maximum possible agreement with the reference map. For *P(p)*, the agreement between the reference map and the other map that has perfect information of quantity and perfect information of location, therefore the agreement is perfect.

Since the study area is not perfectly square, the aggregation technique will produce coarse resolution cells that are made up of different numbers of fine resolution cells. Therefore, it is important to weigh (*Wn*) each cell according to its importance in the analysis, being *Wn* the number of fine resolution cells that constitute a coarse cell, *n*.

For each resolution the components of agreement are separated into:

1. proportion agreement due to chance = *MIN*[*N(n), N(m), M(m)*];
2. proportion agreement due to quantity = if MIN[*N(n), N(m), M(m)*] = *N(n)*, then MIN[*N(m)-N(n)*, *M(m)-N(n)*], else 0;
3. proportion agreement due to location = MAX[*M(m) – N(m)*, *0*];
4. proportion disagreement due to location = *P(m) – M(m);*
5. proportion disagreement to quantity = *P(p)-P(m).*

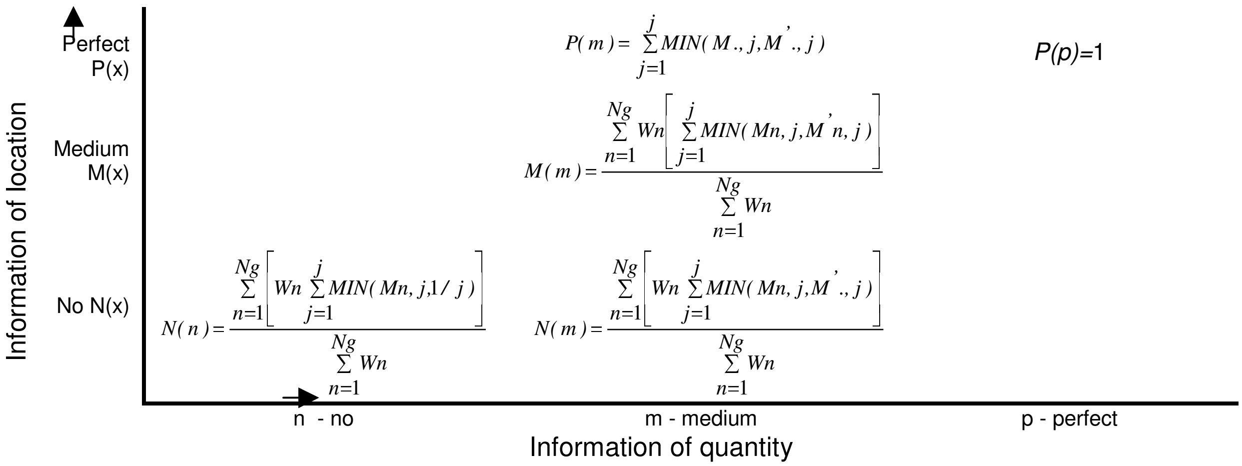

Fig. 2 – Mathematical expressions computed by VALIDADE module for map comparison. The expression in the middle column and middle row gives the agreement between reference Map *M* and comparison Map *M'* at resolution *g*. The other four expressions are idealized agreement between *M* and *M'* maps, based on the combination of information available concerning quantity and location. *n* = grid cell index; *j* = category 1 to 4; J = number of categories (4 in our study), *Ng* = number of grid cells in the map at resolution g (from 1 to 29 in our study), *Wn* the number of fine resolution cells that constitute a coarse cell. When a subscript is a dot (.), it means that the term is summed over that subscript (Adapted from Pontius, 2002).

VALIDATE module also computes the Kappa index of agreement and its variants (Pontius, 2000): *Kstandard* and for location (*Klocation*), calculated through the following equations:

$$Kstandard = \frac{M(m) - N(m)}{P(p) - N(m)} \quad (1)$$

$$Klocation = \frac{M(m) - N(m)}{P(m) - N(m)} \quad (2)$$

For a more detail and understanding of all these statistics, see (Pontius, 2000, Pontius, 2002 and Pontius and Suedmeyer, 2003).

Results and discussion

Previous results

Analysis of the three map comparison using only cell-by-cell comparison shows a good agreement between Maps 2 and 3 (*Kstandard* = 0.85). Maps 1 and 2 are the ones with less agreement since the homogenous areas were computed using independent interpolation techniques. The differences between Maps 1 and 2 and 1 and 3 are mainly due to spatial location (*Klocation* = 0.51, for comparison between Maps 1 and 2, and *Klocation* = 0.63 for comparison between Maps 1 and 3) rather to quantity dissimilarities (*Khisto* = 0.83, for comparison between Maps 1 and 2, and *Khisto* = 0.87 for comparison between Maps 1 and 3). On the other hand, the small difference between Maps 2 and 3 may be due to the quantity category values, since the *Klocation* value is close to maximum similarity. Nevertheless all *Klocation* values show agreement substantially greater than the agreement expected due to chance (Sousa *et al.*, 2002) (see also table 1).

Using the hard neighborhood map comparison, the kappa values (*Kstandard, Klocation or Khisto*) do not vary substantially as cells become coarser, although for grid cell size values higher then 9 the kappa values tend to decrease. Only for neighborhood values that are very high (29 cells, i.e. 2900 m) does the agreement between methods increase substantially (table 1) (Sousa *et al.*, Unpublished).

Fuzzy comparison

Finest resolution

At the finest resolution the overall proportion correct is 58 %, 68 % and 89 %, for comparison between Maps 1 and 2, 1 and 3 and 2 and 3, respectively. These results are in accordance with Kappa values obtained in the previous studies (see table 1). A large percent correct is not necessary an important criterion to judge classification schemes because a large portion of percent correct can be attributable to chance (Pontius, 2000). In the case of comparison between Maps 1 and 2, the proportion of disagreement is mainly due to location errors (30 %) and only 12 % is due to quantity disagreement. Also the differences between Maps 1 and 3 are mainly due to location disagreement (23 %) when compared to quantity disagreement (9 %). So, comparing Map 1, with Maps 2 and 3, Map 3 is in more agreement, not only as quantity but also as location (Fig. 3a), 4a), and Table 1).

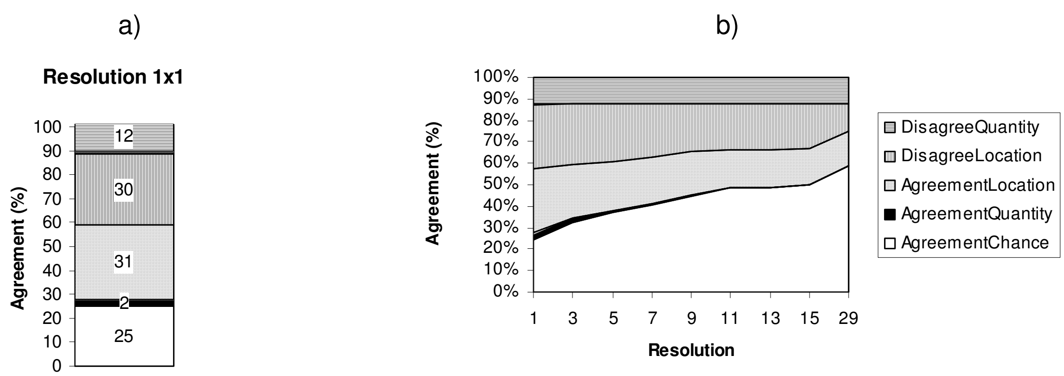

Fig. 3 – a) Percent agreement at fine resolution and b) classification versus resolution for the agreement, between Maps of method 1 and 2.

In contrast, in the more similar Maps (2 and 3) the small differences are due to quantity (8 %), compared to only 3 % due to location disagreement (see Fig. 5a) and table 1). The refinement of Map 3 (i.e., use of probabilities of Map 1 indicator kriging in discriminate analysis of Map 2) seems to compute mainly small differences in quantity, compared to Map 2.

Multiple resolutions

Figures 3a), 4b) and 5b) show how percent agreement increases as resolution becomes coarser from 1 to 29 grid cells per side of each coarse grid cell, for all method comparison. At the finest resolution, percent correct due to chance is 25, in all the figures, since there are four categories. As resolution becomes coarser, agreement due to chance tends to increase, agreement due to location decreases, agreement due to quantity doesn’t change substantially (or tend to zero in comparison Maps 1 and 2, and 1 and 3), and disagreement due to location decreases. Disagreement due to quantity remains constant since changing the resolution does not change the quantity when the fuzzy aggregation method is used. Both disagreement and agreement due to location decrease as resolution becomes coarser, because location is less important at coarser resolutions.

a) b)

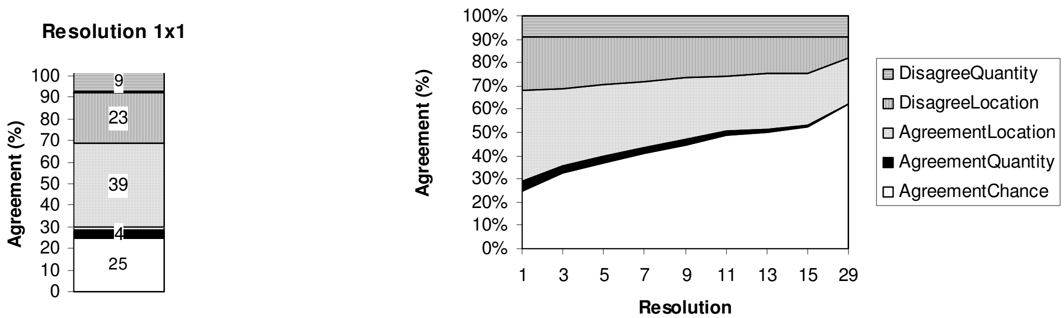

Fig. 4 – a) Cumulative percent agreement at fine resolution and b) classification versus resolution for the agreement, between Maps of method 1 and 3.

a) b)

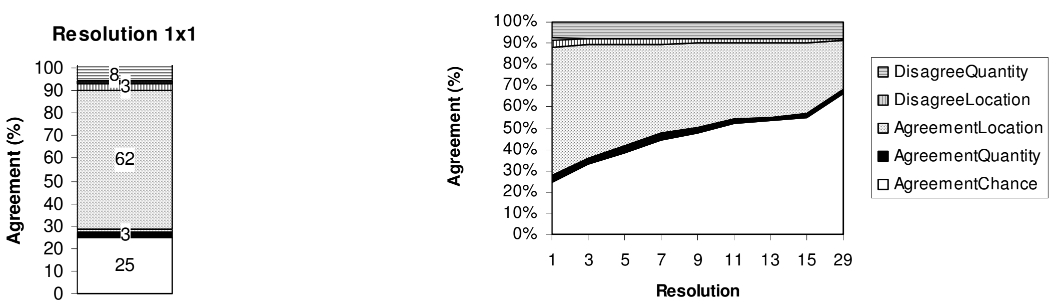

Fig. 5 – a) Percent agreement at fine resolution and b) classification versus resolution for the agreement, between Maps of method 2 and 3.

The percent agreement between Maps 1 and 2 increases from 58 to 75 % as one moves from the finest resolution to the coarsest resolution. On those maps at finest resolution the *Kstandard* is 0.42, and *Klocation* is 0.51, as resolution became coarser the *Kstandard* decreases until the grid cell size reaches 15 and *Klocation* slightly decreases until grid cell size 7, and increase in the following grid cell (9) and on the coarser one (Figs. 3a), 6 and table 1).

As resolution becomes coarser, percent agreement between Maps 1 and 3 increases from 68 to 81.7 %, also at the finest resolution the *Kstandard* is 0.55, and *Klocation* is 0.63. As resolution becomes coarser *Kstandard* slightly decreases until grid cell size 29 and *Klocation* slightly decreases until grid cell size 7, and increases in the coarser grid cells having is higher value (0.68) (Figs. 4a), 6 and table 1).

For both comparisons of Maps 1 and 2 and Maps 1 and 3 the disagreement due to location at resolution 1 is about 90 % the disagreement due to location at resolution 7, indicating that 10 % of the disagreement due to location happens over distances less than 700 m. This grid cell size is similar to the sediment sampling's grid used for computing the maps (750 x 500). This sampling grid was calculated with the principle that there are not important differences in sediment characteristic at distances smaller than the sampling grid (Caeiro, *et al., in press*).

Table 1 – *Kstandard* and *Klocation* for the different resolutions and according to hard and soft classification (maximum similarity = 1).

Maps comparison			**1 and 2**	**1 and 3**	**2 and 3**	**1 and 2**	**1 and 3**	**2 and 3**
Kappa			***Kstandard***			***Klocation***		
Resolution	1	Hard or Soft	0.42	0.55	0.85	0.51	0.63	0.95
	3	Hard	0.42	0.55	0.85	0.5	0.63	0.95
		Soft	0.38	0.51	0.83	0.47	0.6	0.94
	5	Hard	0.42	0.55	0.83	0.51	0.64	0.95
		Soft	0.37	0.50	0.82	0.46	0.59	0.94
	7	Hard	0.42	0.56	0.83	0.51	0.64	0.95
		Soft	0.36	0.5	0.8	0.46	0.59	0.94
	9	Hard	0.34	0.47	0.77	0.52	0.58	0.95
		Soft	0.37	0.5	0.8	0.48	0.6	0.94
	11	Hard	0.4	0.54	0.81	0.5	0.63	0.95
		Soft	0.34	0.48	0.78	0.44	0.58	0.94
	13	Hard	0.38	0.53	0.8	0.49	0.62	0.95
		Soft	0.35	0.49	0.78	0.46	0.60	0.94
	15	Hard	0.37	0.53	0.78	0.48	0.64	0.95
		Soft	0.34	0.47	0.77	0.45	0.58	0.94
	29	Hard	0.44	0.56	0.78	0.52	0.74	0.98
		Soft	0.39	0.52	0.73	0.55	0.68	0.97

Percent agreement between Maps 2 and 3 increases from 89 to 91.6 % as one move from the finest resolution to the coarsest resolution. The *Kstandard* decreases as resolution becomes coarser and *Klocation* is almost constant, only slightly increasing at the coarser resolution.

As well as in hard comparison only for the coarser resolution (29 cells, i.e. 2900 m) does the agreement between methods increase more significantly (see Figure 6 and Table 1), with the exception of map comparison between Map 2 and 3. This could be due to the disappearance of smaller management areas at that resolution.

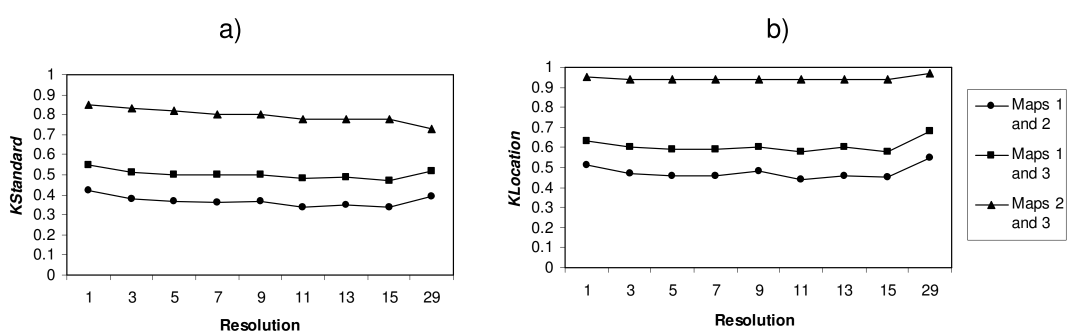

Fig. 6 – a) *Kstandard* and b)*Klocation* for the different resolutions, calculated using fuzzy classification.

Hard versus soft comparison

Values of *KStandard* calculated through hard comparison classification show higher variation than the ones calculated through soft classification. This is specially noticed at cell size 9 (see Table 1). As already explained in previous works (Sousa *et al.*, unpublished), this number of grid cells includes cells of homogenous areas belonging to different organic matter content categories (categories 1 to 4 see Fig. 1) The influence of the hardening step is likely to be the source of this pronounced variation. Similarly, values of *Klocation* obtained with hard comparison classification are slightly higher then the ones computed through fuzzy classification, because the maps look more similar in terms of location using the hard classification compared to the fuzzy classification.

Conclusions

In earlier times the map comparison technique was assessed using Kappa index of agreement. However, *KStandard* fails to penalize for large quantification error and fails to reward sufficiently for small quantification error. When the quantities of each category in Map *M* is similar to another Map *M'*, then Kappa should indicate so, but *Kstandard* attributes those correct classifications to chance. Also *Kstandard* fails to distinguish clearly between quantification error and location error. The classification schemes that attempt to specify accurately both quantity and location are better to evaluate the marginal distributions in spatial models (Pontius, 2000). The new methods presented here of accuracy assessment allows to budget the component of agreement and disagreement between any two maps that show a categorical variable, not only at raw map resolution but also at multiple resolutions using fuzzy classification (Pontius and Suedmeyer, 2003). These techniques compare the maps in terms of quantity and location.

In this work we have shown a complementary application of these comparison techniques, which are usually used for remote sensing, simulation modeling and land change analysis. Our application has been to evaluate the differences between three spatial models of estuarine sediment management areas. The different comparison approaches demonstrated that using either single cell, neighborhood hard or soft comparison the three estuarine management areas maps are still similar, being the differences mainly due to location disagreement. All the results reinforce the robustness of the method for management area calculation. Moreover, support the choice of any of the methods as appropriate for environmental management, and hence moderate the significance of choosing the map resulting from method 1.

Nevertheless there are advantages in using fuzzy classification and budget assessment of component of agreement and disagreement. Fuzzy agreement maps compared with earlier works of hard comparison contains more information and gives a more easy and realistic interpretation of the dataset. As explained in the introduction the hard classification changes the maps leading to misleading results.

Acknowledgments

The research was funded by the Portuguese Science and Technology Foundation and POCTI (Research Project nº POCTI/BSE 35137/99) and financed by FEDER. Sandra Caeiro holds a PRODEP grant.

References

Caeiro, S., Costa, M. H., Painho, M. and Ramos, T. B. Sado Estuary Environmental Management: A GIS Approach. *Proceedings* of Euroworkshop ECO-GEOWATER "GI and Water Resources Assessment", 9 – 13 Julho, Oxford, Inglaterra. http://www.gisig.it/eco-geowater/VirtualPConference. 1 – 13. 2002.

Caeiro, S., Goovaerts, P., Painho, M. and Costa, M. H.. Delineation of Estuarine management areas using multivariate geostatistics: the case of Sado estuary. *Journal of Environmental Science and Technology*. 2003. In press.

Caeiro, S., Goovaerts, P., Painho, M., Costa, M. H. and Sousa, S. Spatial sampling design for sediment quality assessment in estuaries. *Journal of Environmental Modeling and Software.* In press.

Hagen, A. Multi-method assessment of map similary. (Ed.) Ruiz, M., Gould, M., and Ramon, J. 5th AGILE Conference on Geographic Information Science. Palma de Maiorca. Spain. 2002.

Pontius, Jr. R. G. Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering & Remote Sensing.* 66, 1011-1016. 2000.

Pontius, Jr. R. G. Statistical Methods to Partition Effects of Quantity and Location During Comparison of Categorical Maps at Multiple Resolutions. *Photogrammetric Engineering & Remote Sensing.* 68,:1041-1049. 2002.

Pontius, R. G. Jr. and. Schneider. L. C. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment.* 1778. 1-10. 2001

Pontius, R. G. J. and Suedmeyer, B. Components of Agreement between Categorical Maps at Multiple Resolution. *in Remote Sensing and GIS Accuracy Assessment*, CRC Press, Boca Raton FL. 2003 In press.

Sousa, S., Caeiro, S. and Painho, M.. Assessment of map similarity of categorical maps using kappa statistics, the case of Sado Estuary. *Proceedings of ESIG 2002* . 1 – 6 Lisbon, Portugal. 2002.