MAP SIMILARITY MEASUREMENT AND ITS APPLICATION TO THE SADO ESTUARY

Sandra Caeiro [1]

Sandra Sousa [2]

Marco Painho [3]

Abstract – In the past thirty years GIS technology has progressed from computer mapping to spatial database management, and more recently, to quantitative map analysis and modeling. However, most applications still rely on visual analysis for determining similarity within and among maps. The aim of this study is to compare maps of *homogenous* areas computed from estuarine sediment characterization indicators, using different approaches. These maps were defined using three different interpolation methods. Different Kappa statistics, visual map overlays or components of agreement and disagreement due to chance, quantity and location were used for single cell and/or neighborhood (hard and soft) map comparison. Although the three methods were computed with different statistical techniques, their results are similar, supporting the choice of any of the methods as equivalent and thus of equal value to be used as management units of the estuary. Hence the significance of choosing one of the methods is reduced.

Key words : Categorical maps, binary comparison, Kappa statistics, neighborhood statistics, estuarine management.

Resumo – Medidas de semelhança de mapas e a sua aplicação no estuário do Sado. Nos últimos trinta anos, a tecnologia dos *Sistemas de Informação Geográfica* tem progredido desde a elaboração automática de mapas à gestão de base de dados espaciais e, mais recentemente, à análise quantitativa de mapas e modelação. No entanto, a maioria das aplicações baseia-se ainda na análise visual para determinar a similaridade entre mapas. Neste trabalho, pretende-se efectuar uma comparação entre mapas de áreas *homogéneas* de caracterização sedimentar do estuário do Sado, resultante de três métodos de interpolação por geostatística multivariada, de modo a avaliar as diferenças existentes entre eles. Foram utili-

---

[1] Assistente no Departamento de Ciências Exactas e Tecnológicas. Universidade Aberta. Rua Escola Politécnica, n.º 141, 1269-001 Lisboa. Tel: +351 213 916 340, Fax: +351 213 969 293. E-mail: scaeiro@univ-ab.pt

[2] Investigadora. Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa (ISEGI/UNL), Campus de Campolide, 1070-312 Lisboa. Tel: +351 213 870 413, Fax: +351 213 872 140. E-mail: ssousa@ isegi.unl.pt

[3] Professor catedrático no ISEGI/UNL. E-mail: painho@isegi.unl.pt

zadas três abordagens de comparação de mapas: 1.ª – comparação com uma única célula – sobreposição visual de mapas e várias estatísticas Kappa (Kstandard, Kquantidade e Klocalização); 2.ª – comparação *bruta* em contiguidade de células – sobreposição visual de mapas e estatísticas Kstandard e Klocalização; 3.ª – comparação *fuzzy* em contiguidade de células – balanço das componentes de concordância e discordância em termos de quantidade e localização. As três abordagens, embora com vantagens e desvantagens, podem ser complementares. As estatísticas Kappa ainda que não avaliem similaridades intermédias e atribuam classificações correctas devido ao acaso, dão, na realidade, resultados rápidos e de fácil determinação. A comparação por contiguidade de células apresenta claras vantagens, quando comparada com uma única célula, especialmente se é aplicada a ecossistemas dinâmicos como são os estuários. A comparação *fuzzy* em células contíguas permite avaliar as componentes de concordância e discordância em conexão com a quantidade e localização. Este tipo de comparação contém mais informação e permite uma interpretação mais objectiva dos dados, quando confrontado com a comparação *bruta*, dado que esta última modifica os mapas. As três abordagens de comparação de mapas demonstraram que os três mapas das áreas homogéneas são semelhantes. As diferenças registadas são essencialmente devidas a discrepâncias na localização no caso da comparação do Mapa 1 com os Mapas 2 e 3. Este facto levou à conclusão de que a utilização de técnicas diferentes de geostatística multivariada produziu essencialmente diferenças apenas dos padrões espaciais. Por último, embora os três métodos tenham sido estabelecidos com fundamento em técnicas desiguais, os resultados são análogos, considerando os métodos com igual valor para serem aplicados em unidades de gestão ambiental do estuário em estudo, tendo-se optado por seleccionar o método 1, como o mais apropriado para o objectivo em questão.

Palavras-chave : Mapas temáticos, comparação binária, estatística Kappa, estatísticas de vizinhança, unidades de gestão estuarina.

Resumé – Les mesures de similarité cartographique et leur application à l'estuaire du Sado. Depuis trente ans, la technique SIG a beaucoup progressé, en passant de l'élaboration automatique des cartes et de la gestion des bases de données spatiales à l'analyse quantitative des cartes et à la modélisation. Pourtant, la plupart des applications s'appuient encore sur l'analyse visuelle des similarités entre cartes. On tente ici une comparaison entre des cartes d'aires sédimentologiquement homogènes dans l'estuaire du Sado, par l'application de trois techniques d'interpolation par géostatistique multivariée: 1. la comparaison d'une seule cellule (superposition visuelle des cartes et statistiques Kappa (Kstandard, Kquantité et Klocalisation)); 2. la comparaison *brute*, par contiguïté de cellules (superposition visuelle et statistiques Kstandard et Klocalisation); 3. la comparaison *fuzzy*, par contiguïté de cellules (bilan des composantes de concordance ou de discordance, en quantité et en localisation). Les trois procédés peuvent se compléter. La première technique (statistiques Kappa) fournit, malgré ses limitations, des résultats rapides et de détermination facile. La deuxième technique s'applique particulièrement bien aux écosystèmes dynamiques comme c'est le cas de l'estuaire du Sado. La troisième technique contient davantage d'information et permet une interprétation plus concrète des données. Bien que les trois méthodes aient des fondements techniques différents, leurs résultats demeurent semblables.

Mots-clés : Statistiques Kappa, statistiques de contiguïté, unité de gestion de l'environnement de l'estuaire.

I.   INTRODUCTION

In the different Geographical Information Systems (GIS) applications, the environment in general and coastal areas in particular, comparing or detecting different categorical maps is an essential issue. The accuracy of a comparison procedure based on a more reliable and robust approach could lead to a marked improvement in the ability to detect a map change.

The simplest way to compare two maps is to compute the correlation between the mapped variables. But with this method the locations of the points are not considered. This reflects a major drawback on the method in an overall comparison, because a given correlation may reflect the degree of correspondence over the entire map area, or may be the result of a large deviation in a small region of the map (DAVIS, 1986). Map spatial comparison procedures are then, amongst others, important for the validation and calibration of spatial models (HAGEN, 2002b). These procedures can express the similarity between two maps by looking at simple proportions of areas or by measuring it numerically. This numerical similarity can be assessed by categorical representation of overlay results as a contingency table. Statistical analyses are then made with various integral measures of association, log-linear models, among others (ZASLAVSKI, 1995). When the assessment consists of a number of pair-wise comparisons, based on a cell-by-cell agreement using the confusion matrix, the Kappa statistic (COHEN, 1960) can be a suitable approach (CARLETTA, 1996). The result of a map comparison can be an overall value for similarity (e.g. a value between 0 and 1) or a map in its own right, which means that the result of a comparison of two maps is a third map indicating per location how strong the similarity is (HAGEN, 2002b). However the confusion matrix fails to distinguish between a near-miss and a far-off map. 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; PONTIUS and SUEDMEYER, 2004).

There is wide disagreement about the usefulness of Kappa statistics to assess raster agreement. The model for chance agreement is statistical independence, which is not the expected basis within or between coverages (CHRISMAN, 1997). Kappa statistics should not be viewed as the unequivocal standard or default way to quantify agreement and alternatives to make an informed choice should be considered. Nevertheless, the ease in calculation and its appropriateness for testing whether agreement exceeds chance levels for binary and nominal ratings (UEBERSAX, 2003). Also new variants of Kappa were recently introduced to consider similarity of location (PONTIUS, 2000) and quantity (HAGEN, 2002b).

When the comparison is to be done in non-static environments, like coastal environments it is difficult to define sampling grids in exact positions and therefore a single cell-by-cell analysis comparison is less representative. If a specific cell fails to have the correct category, then it is counted as complete error, even when the correct category is found in a neighbouring cell. Cell-by-cell analysis

can fail to indicate general agreement of pattern because it fails to consider spatial proximity to agreement (Pontius and Suedmeyer, 2004). Therefore, in this case, a neighbourhood cell comparison is more appropriate. Using the neighbourhood to compare categorical maps could be computed using a hard or soft classification.

Neighbourhood hard statistics using a mode function are very similar to frequency filters. These filters count the frequency of the attribute values occurring in a chosen window with a fixed size. The majority filter selects the value with the highest frequency. Majority filters are very useful for smoothing irregular edges between adjacent areas and they eliminate rare attribute values from a raster (Molenaar, 1998). A wide variety of functions can be used for hard neighbourhood calculations depending on the goal of the work, but Mode is the more accurate for categorical data (Murteira and Black, 1983). Hard neighbourhood operations summarize the attributes occurring in the vicinity of each location. It creates a new map where the value assigned to a location is computed as a function of independent values surrounding that location. This group of operations can be conceptualised as 'moving windows' sliding throughout the mapped area.

However, 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 soft classification (also here referred to as 'fuzzy') to the comparison of categorical maps it is possible to obtain a spatial and gradual analysis of the similarity of two maps at different multi-resolution. In addition, in this soft comparison, it would be helpful to have an analytical technique that allocates the sources of agreement and disagreement indicating in the ways in which the comparison map is strong and weak (Pontius and Suedmeyer, 2004).

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

The team (Caeiro *et al*., 2002) has been working on the development of an environmental data management system through sediment quality assessment for the Sado Estuary (EMMSado) in Southeast Portugal (fig. 1). The units of this management system are spatially contiguous regions of sediment structure (*homogenous* areas). To delineate these *homogenous* areas, three maps were computed using different multivariate geostatistical approaches. Strong agreement of similarities will further support the choice of any of the methods as equivalent, hence, the reduced significance of choosing one of the methods.

The aim of this work is to assess the difference between the three maps using different similarity assessment approaches: first approach – single cell comparison using binary comparison and Kappa standard and its new variants Klocation to evaluate agreement due location and Khisto to evaluate agree-
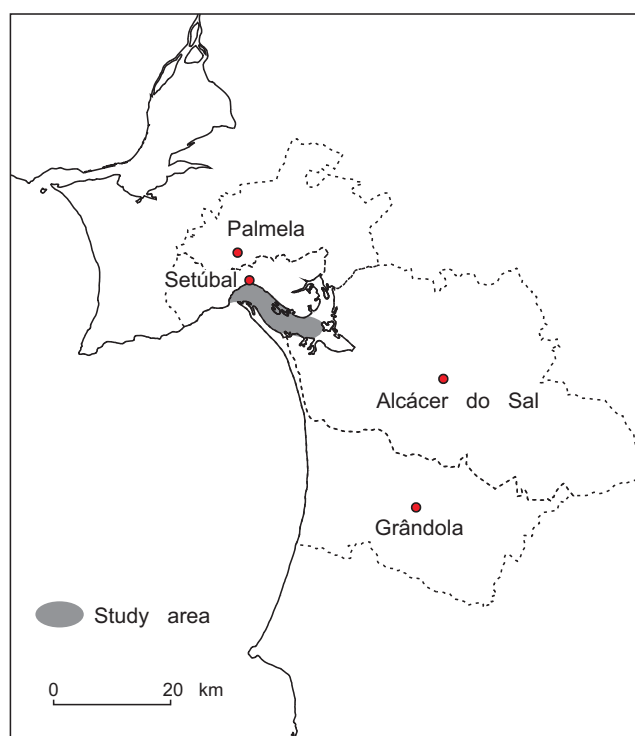
Fig. 1 – Study area.
Fig. 1 – Localização da área em estudo.

ment due to quantity; second approach – hard neighbourhood comparison, using binary comparison and Kappa statistic; and the third approach soft neighbourhood comparison using components of agreement and disagreement due to change, quantity and location. The advantages and disadvantages of using these different approaches are also compared and discussed.

II.   METHODS

1. Previous work

The three maps of *homogenous* areas were computed based on three sediment characterization indicators (fine fraction, organic matter and redox potential), using the following methods (fig. 2):

Method 1: Cluster analysis of dissimilarity matrix that accounts for distances in both the attribute and geographical spaces, followed by indicator kriging of the classification. Cluster classification yielded four clusters. For each cluster indicator semivariogram

was computed and model-fitted for indicator kriging (100×100m grid) (Map 1);

Method 2: Block kriging (discretisation by four points) of the three sediment indicators, followed by a discriminated analysis of K-means clustering predicted values (Map 2);

Method 3: A hybrid approach that incorporates the probabilities of occurrence at non-sample locations of indicator kriging used in Method 1 into discriminated analysis of Method 2 (Map 3).
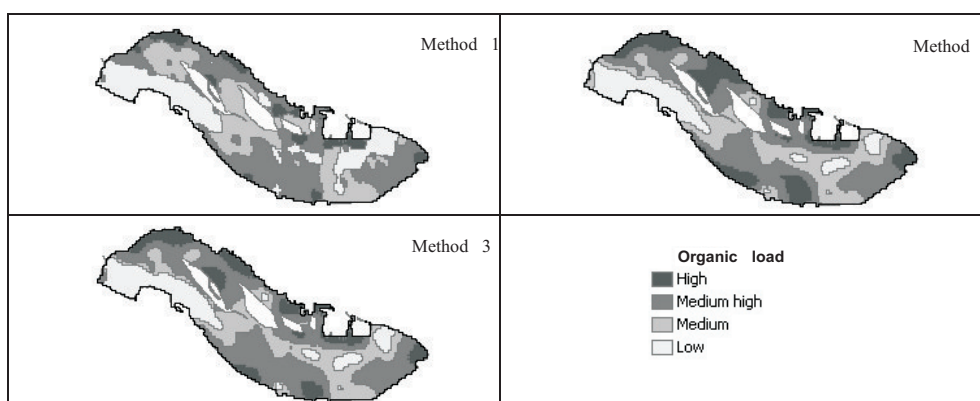

Fig. 2 – Maps representing the three methods for Sado estuary management area delineation.

Fig. 2 – *Mapas representativos das três metodologias para definição das áreas homogéneas do Estuário do Sado (1cm no mapa: cerca de 6km)*.

A detailed explanation of these methods, including their methodological advantages and drawbacks, is available in Caeiro *et al.* (2003a). These interpolation surfaces (100×100m grid) were based on a systematic unaligned campaign in 153 points spread over a final sampling grid cell of 750×500m, using prior information on the spatial variation in the estuarine sediments (Caeiro *et al.*, 2003b). In each of these categorical maps, four organic matter contents categories were computed: 1 - high, 2 - medium high, 3 - medium, and 4 - low organic load. Results of Map 1 seem to be in better agreement with estuary behaviour, assessment of contamination sources and previous work conducted at this site (Caeiro *et al.*, 2003a). For that reason, Map 1 was considered the reference for the comparison between Map 1 and Map 2 and Map 1 and 3. 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.

2. Single cell comparison

This first comparison approach involved the overlay of the original maps on a cell-by-cell basis, to produce a map and attribute table of site-specific

differences (fig. 3). Simple map algebra operations (union) were used. Then, based on the contingency table (or confusion matrix), Kappa and variants were calculated. Klocation, as discussed by PONTIUS (2002), was used to access the similarity of location, and Khisto, as developed by HAGEN (2002b), was used to access the similarity of quantity. These Kappa calculations are explained in a previous work (SOUSA *et al.*, 2002). Binary classification was computed using raster calculator and reclassified within ArcGIS™ Spatial Analyst, which states for each cell whether or not the maps are identical at that location.

3. Hard neighbourhood comparison

In this approach each location is a function of the input cells of different neighbourhood sizes. The approach was applied to square windows of 3, 5, 7, 9, 11, 13, 15 and 29 cells, at the finest resolution, each cell being 100×100m. The last neighbourhood (29) was only used to evaluate the sensitivity of this approach. The use of the different neighbourhood sizes, or map resolutions, permits for the gauging of the map results' sensitivity to scale variation and finding key map resolutions in the case of map behaviour changes. 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 neighbourhood size of 3 means that a 3×3 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 consti-tute the coarse cell. The GridStat software application was developed for this purpose. It results from a refinement of the majority function of 'Neighbourhood Statistics' in ArcGIS™ Spatial Analyst. The majority function in 'Neighbourhood Statistics' is very similar to the mode. Majority computes the value that occurs most often in the neighbourhood but has a flaw: when a tie occurs, the processing cell is classified as 'No Data'. GridStat Mode will classify the same cell either with its original value or with the closest category. Map algebra and contingency tables were then used to obtain the difference between each of the two maps and create a classification of their differences. For quantification of map comparison approaches, Kstandard and Klocation and binary classifica-tion were used (fig. 3). Khisto was not calculated in this approach since it can be calculated through Kstandard and Klocation.

4. Soft neighbourhood comparison

For computing fuzzy map comparison the module VALIDATE in Idrisi Kili-manjaro® 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
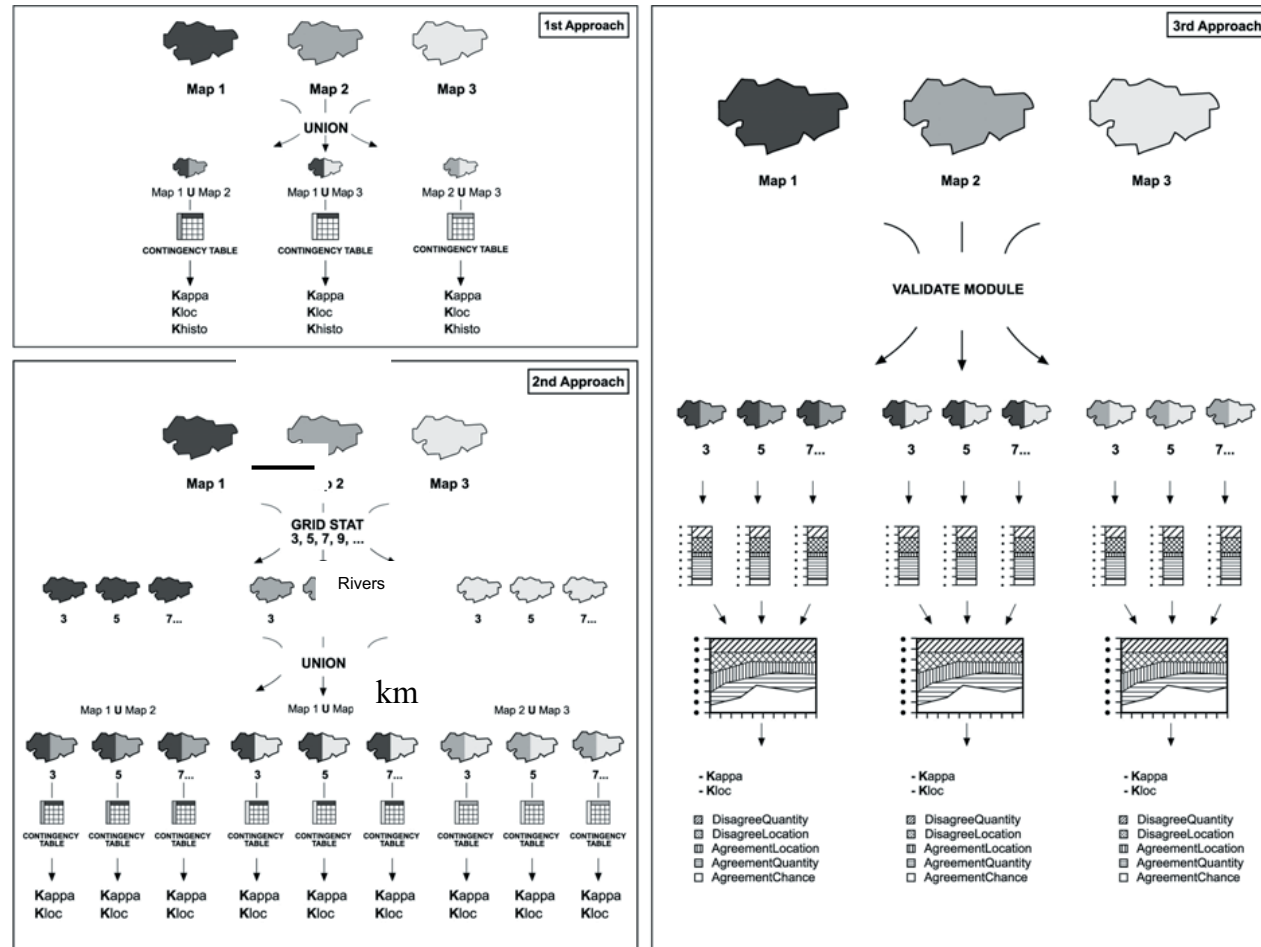
Fig. 3 – Approaches used for map comparison.
Fig. 3 – Métodos utilizados para a comparação dos mapas.

raw data (finest resolution) to a very coarse resolution. An arithmetic sequence was used to create the aggregating neighbouring 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 approach.

For maps with one single stratum/sub-region, VALIDATE computes five especially important numbers for each resolution that constitute the basis for the components of agreement and disagreement between the reference map and other maps that have increasingly accurate information, from none to medium and perfect information: i) correct due to change, ii) correct due to quantity, iii) correct due to location, iv) error due to location, and v) error due to quantity. Each cell has partial membership in all of the categories, and the agreement for category j in cell n is to be the minimum of proportion of category j in grid cell n of Map 1 and proportion of category j in grid cell n of Map 2. VALIDATE module also computes the Kstandard and Klocation.

For a more detail and understanding of all VALIDATE calculation, see Pontius (2000, 2002) and Pontius and Suedmeyer (2004) or, for specific application into this study, see Caeiro *et al.* (2003c).

III. RESULTS AND DISCUSSION

1. Single cell comparison

Analysis of the Kappa values of the three map comparison shows almost perfect agreement between Maps 2 and 3 (Kstandard=0.85), confirmed not only for quantity (Khisto=0.89) but also for location similarity (Klocation=0.95). This result was expected since Method 3 is a refinement of Method 2. Method 3 is moderately similar to Method 1 (Kstandard=0.55). These maps are computed using different multivariate geostatistics but Method 3 uses results from Method 1. Maps 1 and 2 are the ones with less agreement (Kstandard=0.42) since the computed *homogenous* areas use independent interpolation techniques. Looking at the Klocation value of Maps 1 and 2 (0.51) the differences between these two maps are mainly due to spatial location rather to quantitative dissimilarities (Khisto=0.83). Binary comparison between Maps 1 and 2 in figure 4 also confirmed this local difference due to fewer areas of agreement classification. This major location difference can also be true for the Map 1 and 3 comparison since the Klocation value (0.55) is further away from the maximum value than Khisto (0.87). On the other hand, the small difference between Maps 2 and 3 may be due to the quantity category values since the Kloc value is close to maximum similarity (Table 1). The refinement of Map 3, i.e. the use of probabilities of the indicator kriging in the Discriminate analysis, seems to compute mainly small differences in quantity, compared to Map 2. Nevertheless, all Klocation values show agreement substantially greater than the agreement expected due to chance (Sousa *et al.*, 2002).

2. Hard neighbourhood cell comparison

The Kappa values (Kstandard and Klocation) do not vary significantly along an increase of neighbourhood cells as confirmed by observation of Table I. On average, Kappa values are almost constant up to a neighbourhood cell size of 9. This cell size (900×900m) corresponds to less than double the grid sampling size (500×750m). For cell sizes larger than 9, the Kappa values tend to decrease, in other words, the differences between the maps increase. This is due to a greater number of neighbourhood cells and, thus, the inclusion of cells of homogenous areas with a different organic content category (categories 1 to 4, fig. 2). Only for neighbourhood values that are too high (29 cells, or 2900m) does the agreement between methods increase more considerably (fig. 4, Table I), since the smaller areas almost disappear. This window size of neighbourhood cells (29) is disproportional since it is almost four times larger than the size of the grid sampling cells.

Table I – *Kstandard* and *Klocation* for the different resolutions
and according to hard and soft classification (maximum similarity = 1)
(adapted from Caeiro *et al.*, 2003c).

Quadro I – Kstandard *e* Klocalização *para as diferentes resoluções
e de acordo com as classificações rígida e flexível (máxima similaridade = 1)
(adaptado de* Caeiro *et al., 2003c).*

| 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
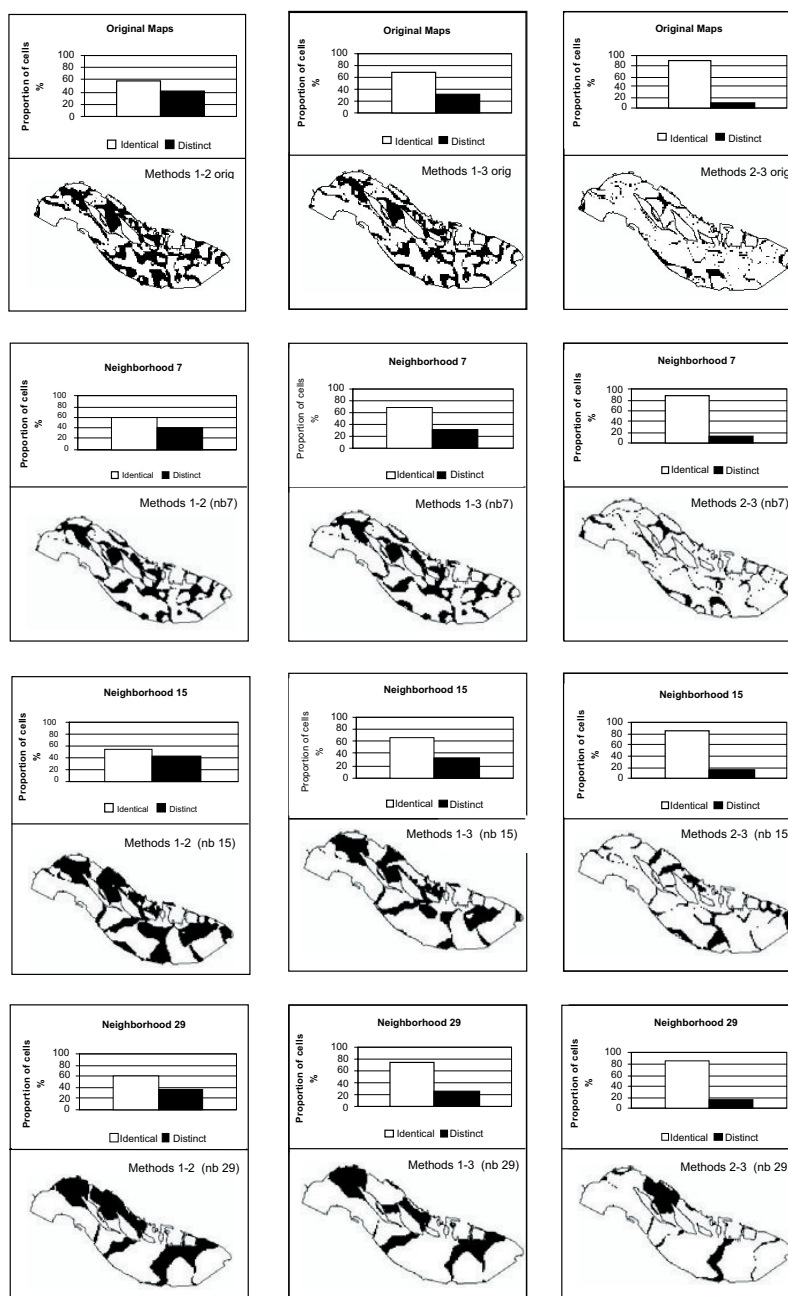
Fig. 4 – Map comparison for Maps 1, 2 and 3 using binary classification
for single cell and hard neighborhood cells 7, 15 and 29.

*Fig. 4 – Mapas de classificação binária de comparação dos Mapas 1, 2 e 3, utilizando
uma única célula ou contiguidade de 7, 15 e 29 células (1cm no mapa: cerca de 9km)*

The map comparison taking into account the hard neighbourhood cells emphasises the distinctions between methods, such as areas with more variation within the estuary. These areas are in the opposite direction to water flow (30º), in locations with less hydrodynamics subject to different sources of pollution, like on the north side of the estuary (Caeiro, 2003a) (i.e., the map comparison between methods 1-3 [n.b. 15] in figure 4).

3. Soft neighbourhood cell comparison

At the finest resolution (original maps) the overall correct proportion is 58%, 68% and 89%, for comparison between Maps 1 and 2, 1 and 3 and 2 and 3, respectively. A large percent correct is not necessarily 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%). Comparing Map 1 with Maps 2 and 3, Map 3 is in more agreement, not only in terms of quantity but also location (figs. 5a and 6a and Table 1), for the same reasons explained earlier.

In contrast, among the maps (2 and 3) with more similarities, the small differences are due to quantity (8%), compared to only 3% due to location disagreement (see fig. 7a and Table 1). These results coincide with Kappa values obtained in the previous approaches.

Figures 5b, 6b and 7b show how percent agreement increases as resolution becomes coarser from 1 to 29 grid cells per side of each coarse grid cell for all
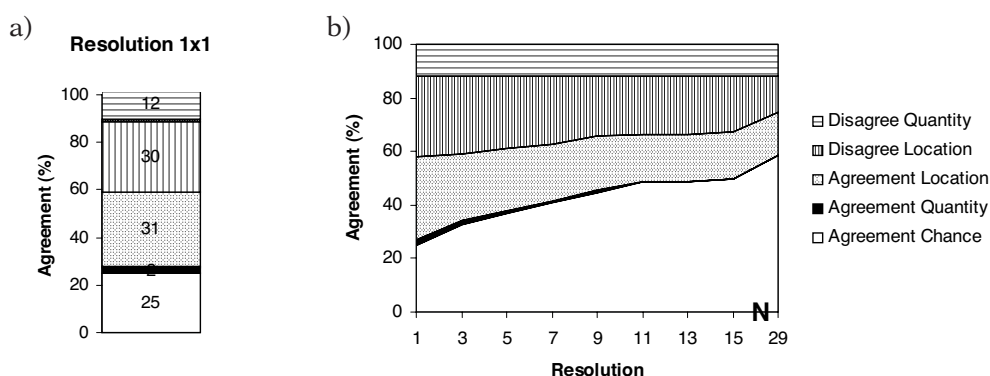

Fig. 5 – a) Cumulative percent of agreement at fine resolution and b) classification versus resolution for the agreement, between maps of Methods 1 and 2 (adapted from Caeiro *et al.*, 2003c).

Fig. 5 – a) Percentagens cumulativas de concordância nos mapas originais e b) classificação versus resolução para a concordância, obtido pelos mapas dos métodos 1 e 2 (adaptado de Caeiro et al., 2003c).
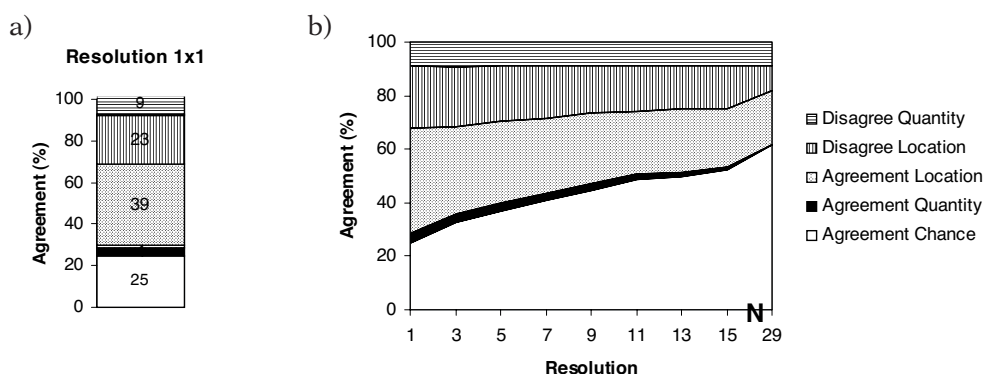
a)

b)

Fig. 6 – a) Cumulative percent of agreement at fine resolution and b) classification versus resolution for the agreement, between Maps of method 1 and 3. (adapted from CAEIRO *et al.*, 2'003c).

Fig. 6 – a) Percentagens cumulativas de concordância nos mapas originais e b) classificação versus resolução para a concordância, entre os Mapas dos métodos 1 e 3 (adapted from CAEIRO et al., 2003c).

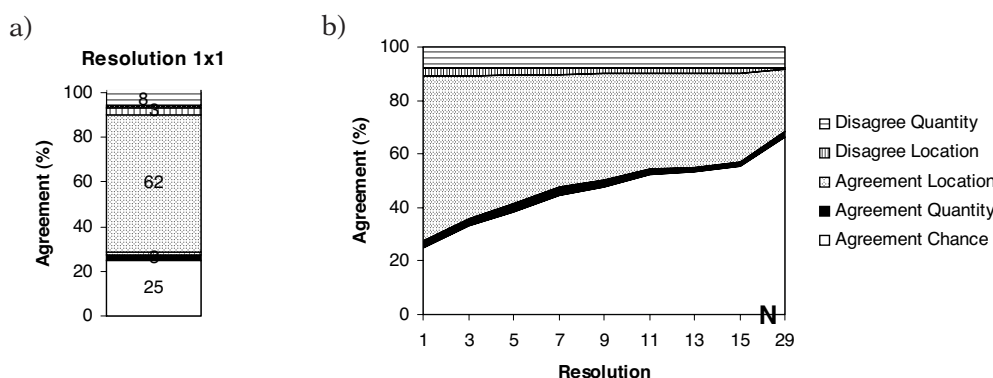a)

b)

Fig. 7 – a) Cumulative percent of agreement at fine resolution and b) classification versus resolution for the agreement, between Maps of method 2 and 3. (adapted from CAEIRO *et al.*, 2'003c).

Fig. 7 – a) Percentagens cumulativas de concordância nos mapas originais e b) classificação versus resolução para a concordância, entre os Mapas dos métodos 2 e 3 (adapted from CAEIRO et al., 2003c).

method comparisons. 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 does not change substantially (or tends towards zero when comparing Maps 1 and 2 and Maps 1 and 3), and disagree-

ment due to location decreases. Disagreement due to quantity remains constant since changing the resolution does not change the quantity when the soft (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.

The percent agreement between Maps 1 and 2 increases from 58 to 75% moving from the finest resolution to the coarsest resolution. On those maps at the finest resolution, 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. 5b and 8 and Table 1).

Percent agreement between Maps 1 and 3 increases from 68 to 81.7% as resolution becomes coarser. As resolution becomes coarser, Kstandard slightly decreases until grid cell size 15 and Klocation slightly decreases until grid cell size 7, and increases in the coarser grid cells having a higher value (0.68) (figs. 6b and 8 and Table 1).
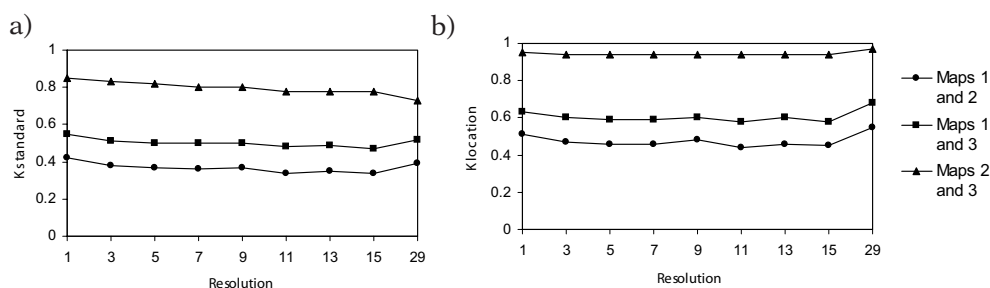

Fig. 8 – a) Kstandard and b) Klocation for the different resolutions, calculated using fuzzy classification (adapted from Caeiro *et al.,* 2003c).

Fig. 8 – a) Kstandard e b) Klocalização para as diferentes resoluções, calculadas usando a classificação fuzzy (adaptado de Caeiro et al.*, 2003c).*

For both comparisons of Maps 1 and 2 and Maps 1 and 3, the disagreement due to location at resolution 7 is about 90% of the disagreement due to location at resolution 1, indicating that only 10% of the disagreement due to location happens over distances less than 700m. This grid cell size is similar to the sediment sampling's grid used for computing the maps (750×500). The sampling grid was calculated according to the principle that there are not important differences in sediment characteristic at distances smaller than the sampling grid (Caeiro *et al.,* 2003b).The percent of agreement between Maps 2 and 3 increases from 89 to 91.6%, moving 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 (fig. 7b).

In hard comparison, only with the coarser resolution (29 cells, i.e. 2900m) does the agreement between methods also increase more significantly (see

fig. 8 and Table 1), with the exception of Kstandard of map comparison between Map 2 and 3.

Comparing hard with soft comparison, comparison between Map 1 and 2 and Map 1 and 3 values of Kstandard calculated through hard comparison classification show higher variation than the ones calculated through soft classification. This is especially visible at cell size 9 (see Table 1) as previously stressed in point III.2. The influence of the hardening step is likely to be the source of this more pronounced variation. Similarly, values of Klocation obtained with hard comparison classification are slightly higher then the ones computed through soft classification. The maps look more similar in terms of location using the hard classification compared to the fuzzy classification since hard classification reclassifies each grid cell at map resolution according to Mode function.

IV. CONCLUSIONS

In this paper, we have used different map comparison approaches to compare spatial models of *homogenous* estuarine sediment areas. All of the approaches can complement each other. Since the 1960s, the map comparison technique has been assessed using the Kappa index of agreement. However, Kstandard fails to penalise, to not give intermediate similarities and to attribute correct classification to chance. It also fails to penalise for large quantitative error and to reward sufficiently for small quantitative error. In addition, it is unsuccessful in clearly distinguishing between quantification error and location error (PONTIUS, 2000). Nevertheless, in this work, Kappa gave good, fast and easy information that can be added to other methods, in particular to the new Kappa variants to quantify location and quantitative errors. The binary classification also gave a quick and easy visual comparison. The classification schemes that attempt to accurately specify both quantity and location are better to evaluate the marginal distributions in a spatial model. Here we presented a new method (third approach) of accuracy assessment to budget the component of agreement and disagreement in terms of quantity and location between any two maps that show a categorical variable. This assessment can be done not only at raw map resolution but also at multiple resolutions by using fuzzy classification (PONTIUS and SUEDMEYER, 2004). By using the minimum proportion of each category, in each grid cell, the results are the proportion of a category that is within the same pixel as another category. This allows fuzzy agreement maps containing more information and giving an easier and more realistic interpretation of the dataset, when compared with hard comparison. Despite this, the results, though easily omputed in the software and represented in a simple graphical form, can be tricky to interpret. Similar conclusions were also obtained about the maps' differences using either Kappa indexes or the component of agreement.

The hard classification can lead to changes in the quantity of each category, generating misleading results. Even so, hard neighbourhood comparison may be useful for an easy visualization of map overlays (e.g. fig. 4) as a criterion for

defining a reduced number of these *homogenous* areas for a future extended management programme of this estuary. Coastal zone management represents a dynamic process that develops and implements a co-ordinated strategy to allocate resources to achieve the conservation and sustainable varied use of coastal zones (FRENCH, 1997). A reduced number of *homogenous* areas (units of management) provides a model of estuary management that is more appropriate, easier to manage and less expensive to monitor.

The different comparison approaches demonstrated that, when using either single cell, neighbourhood hard or soft comparison (though giving different interpretations), the three *homogenous* estuarine area maps are still similar. The differences are mainly caused by location disagreement in comparing Map 1 with Maps 2 and 3. This helped to conclude that the use of different geostatistical multivariate methods mainly computed differences in spatial patterns. These different comparison approaches also helped to consolidate the lack of benefit in using the indicator kriging probabilities of the Method 1 into Method 2, resulting in Method 3 as stated in CAEIRO *et al.* (2003a). Finally, it also confirmed that the choice of Method 1 seems to better represent the spatial pattern of the four categories of organic load.

All of the results reinforce the robustness of *homogenous* area calculation. Moreover, the results support the choice of any of the methods as equivalent and, thus, of equal value for environmental management. Hence, maps resulting from Method 1 are less likely to be a bad choice.

ACKNOWLEDGMENTS

REEFERENCES

CAEIRO, S.; COSTA, M. H.; PAINHO, M. and RAMOS, T. B. (2002) – Sado Estuary environmental management: A GIS approach. *Proceedings of Euroworkshop ECO-GEOWATER GI and Water Resources Assessment*, 9-13 July, Oxford. http://www.gisig.it/eco-geowater/VirtualPConference: 1-13.

CAEIRO, S.; GOOVAERTS, P.; PAINHO, M. AND COSTA, M.H. (2003a) - Delineation of estuarine management areas using multivariate geostatistics: the case of Sado estuary. *Journal of Environmental Science and Technology*, 37(18): 4052-4059.

CAEIRO, S.; GOOVAERTS, P.; PAINHO, M.; COSTA, M.H. and SOUSA, S. (2003b) – Spatial sampling design for sediment quality assessment in estuaries. *Journal of Environmental Modelling and Software*, 18(10): 853-859.

Caeiro, S.; Sousa, S.; Pontius Jr., R.G. and Painho, M. (2003c) – Sado Estuary management areas: Hard versus soft classification map comparison. *Proceedings of CoastalGIS 2003, Fifth International Symposium on GIS and Computer Cartography for Coastal Zone Management*, 16-18 October, Genova: 1–9.

Carletta, J. (1996) – Assessing agreement on classification tasks: the Kappa statistic. *Computational linguistics*, 22: 249-254.

Chrisman, N. (1997) – *Exploring geographic information systems*. John Wiley & Sons, Inc., 298 p.

Cohen, J. (1960) – A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20: 37-46.

Davis, J. (1986) – *Statistics and data analysis in Geology*. Second Edition. John Wiley & Sons, Inc. 646 p.

French, P.W. (1997) **–** *Coastal and estuarine management***.** Environmental Management Series, Routledge, London, 251 p.

Hagen, A. (2002a) – Approaching human judgement in the automated comparison of categorical maps. *Proceedings of International Conference on recent advantages in soft computing 2002*, Nottingham: 131-132.

Hagen, A. (2002b) – *Technical Report: Comparison of maps containing nominal data*. National Institute for Public Health and the Environment, Maastricht.

Hagen, A. (2002c) – Multi-method assessment of map similarity. *In* Ruiz, M., Gould, M. and Ramon, J. (Eds.), *Proceedings of the 5th AGILE Conference on Geographic Information Science*. Palma de Majorca.

Molenaar, M. (1998) **–** *An introduction to the theory of spatial object modelling for GIS*. Wiley & Francis, London.

Murteira, B. J. F. and Black, G. H. J. (1983) – *Estatística Descritiva*. McGraw-Hill, Lisboa, 285 p.

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

Pontius, R. G. Jr. (2002) - Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering & Remote Sensing*, 68: 1041-1049.

Pontius, R. G. Jr. and Schneider, L. C. (2001) – Land-cover change model validation by an ROC method for the Ipswich watershed. Massachusetts, USA. *Agriculture, Ecosystems and Environment*, 85(1-3): 239-248.

Pontius, R.G. Jr. and Suedmeyer, B. (2004; forthcoming) – Components of Agreement between Categorical Maps at Multiple Resolution. *In* Lunetta, R. S. and Lyon, J. G. (eds.) – *Remote Sensing and GIS Accuracy Assessment*, CRC Press, Boca Raton.

Sousa, S.; Caeiro, S. and Painho, M. (2002) – Assessment of map similarity of categorical maps using Kappa statistics, the case of Sado Estuary. *In* Sousa, A. *et al*. (eds.) – *Proceedings of ESIG 2002*, 13-15 November, Lisbon: 1-6.

Uebersax, J. S. (2003) – Kappa Coefficients: A Critical Appraisal. http://ourworld.compuserve.com/homepages/jsuebersax/kappa.htm.

Zaslavsky, I. (1995) – Analysis of association between categorical maps in multi-layer GIS. *Proceedings of GIS/LIS 95*, Vol. 2, 14 - 16 November: 1066-1074.