# DIGITAL SOIL MAPPING USING MULTIPLE LOGISTIC REGRESSION ON TERRAIN PARAMETERS IN SOUTHERN BRAZIL

Elvio Giasson[1]*; Robin Thomas Clarke[2]; Alberto Vasconcellos Inda Junior[1]; Gustavo Henrique Merten[2]; Carlos Gustavo Tornquist[3]

[1]UFRGS - Faculdade de Agronomia - Depto. de Solos, C.P. 15100 - 90001-970 - Porto Alegre, RS - Brasil.
[2]UFRGS - Instituto de Pesquisas Hidráulicas, C.P. 15029 - 91501-970 - Porto Alegre, RS - Brasil.
[3]UFRGS - Programa de Pós-Graduação em Ciência do Solo.
*Corresponding author <giasson@ufrgs.br>

ABSTRACT: Soil surveys are necessary sources of information for land use planning, but they are not always available. This study proposes the use of multiple logistic regressions on the prediction of occurrence of soil types based on reference areas. From a digitalized soil map and terrain parameters derived from the digital elevation model in ArcView environment, several sets of multiple logistic regressions were defined using statistical software Minitab, establishing relationship between explanatory terrain variables and soil types, using either the original legend or a simplified legend, and using or not stratification of the study area by drainage classes. Terrain parameters, such as elevation, distance to stream, flow accumulation, and topographic wetness index, were the variables that best explained soil distribution. Stratification by drainage classes did not have significant effect. Simplification of the original legend increased the accuracy of the method on predicting soil distribution.
Key words: GIS, DEM, soil survey, terrain analysis

# MAPEAMENTO DIGITAL DE SOLOS UTILIZANDO REGRESSÕES LOGÍSTICAS MÚLTIPLAS E PARÂMETROS DO TERRENO NO SUL DO BRASIL

RESUMO: Os levantamentos de solos são fontes de informação necessárias para o planejamento de uso das terras, entretanto eles nem sempre estão disponíveis. Este estudo propõe o uso de regressões logísticas múltiplas na predição de ocorrência de classes de solos a partir de áreas de referência. Baseado no mapa original de solos em formato digital e parâmetros do terreno derivados do modelo numérico do terreno em ambiente ArcView, vários conjuntos de regressões logísticas múltiplas foram definidas usando o programa estatístico Minitab, estabelecendo relações entre as variáveis do terreno independentes e tipos de solos, usando tanto a legenda original como uma legenda simplificada, e usando ou não estratificação da área de estudo por classes de drenagem. Os parâmetros do terreno como elevação, distância dos rios, acúmulo de fluxo e índice de umidade topográfica foram as variáveis que melhor explicaram a distribuição das classes de solos. A estratificação por classes de drenagem não teve efeito significativo. A simplificação da legenda aumentou a precisão do método na predição da distribuição dos solos.
Palavras-chave: levantamento de solos, SIG, MNT, análise do terreno

## INTRODUCTION

Soil surveys are recognized as important sources of information for land use planning and management. In Brazil, soil surveys for most of the country are available only at small scale (1:750,000), and just a small portion of the Brazilian territory has semi-detailed or detailed soil surveys because of funding limitations.

A more cost-effective approach to traditional, large scale soil surveys would be to map soils of representative areas within homogeneous regions and use the soil-landscape relationships to predict soil distribution on non-surveyed areas (Schneider & Klamt, 1996). This approach is similar to the reference area method (Lagacherie et al., 2001), which is based on the hypothesis that it is possible to sample a reference area including most of the soil classes of a region. Based on this area, the prediction of soil distribution on other areas may be facilitated if the landscape is modeled by digital terrain analysis (Hengl & Rossiter, 2003), and if relationships between soils and landscape are modeled.

Multiple logistic regression have been used successfully in soil science and related fields, as to predict landslide hazard (Ohlmacher & Davis, 2003) or probability of occurrence of soil drainage classes (Campling et al., 2002; Kravchenko et al., 2002), or to relate the presence of a non-calcareous clay–loam horizon to terrain attributes (King et al., 1999). As soil map units are categorical variables, multiple logistic regression may be suitable for predicting occurrence of soil classes from landscape variables, with the advantage of allow to associate the prediction with probabilities of occurrence of soil mapping units.

Although previous works used logistic regression to estimate the occurrence of specific soil characteristics instead of soil taxonomic classes or mapping units, they suggest that logistic regressions may have potential for producing soil maps from terrain parameters based on relationships between these parameters and soil occurrence. This study evaluates a new method of extrapolation on landscape parameters by testing how well multiple logistic regressions can reproduce a soil map of a reference area.

## MATERIAL AND METHODS

The study area was the Sentinela do Sul county, located in southeastern Rio Grande do Sul, Brazil (UTM 22S, Easting from 434,000 m to 453,000 m, and Northing from 6,596,000 m to 6,626,000 m), comprising an area of 253 km$^2$ with minimum elevation at sea level and maximum elevation of 525 m. The regional climate is subtropical (Köppen classification: Cfa II 1 d), with mean annual temperature 17.6ºC and mean annual precipitation 1600 mm.

The relief is nearly flat in the alluvial plains, gently sloping or strongly sloping in the colluvial deposits, and very steep in the highlands, the landform that dominates the landscape. There are two different types of parent material: a granite-gneiss mixture with associated migmatite, and Cenozoic sediments, including sediments of gravitational and colluvial sources (Silveira, 1984). Recent and old fluvial deposits complete the transition of sedimentation between the complexes of granites and the coastal sedimentation on the eastern side of the county, making the parent material distribution complex.

A detailed reconnaissance soil survey of the county at scale 1:50,000 (Klamt et al., 1996), produced according to usual soil survey procedures that included extensive field work and photointerpretation, established eight mapping units (MU1, MU2, ... and MU8) (Table 1). Most MUs are combinations (undifferentiated group, complex, or association) of two or more of the nine different soil taxonomic units found in the area because of the difficulty in separate individual soil classes as consociations (i.e., mapping units with only one dominant soil class) given the complexity of soil-landscape relationships, where each distinguishable landform usually had several classes of soils (Klamt et al., 1996).

The digital elevation model (DEM) used had a resolution of 3 arc sec, corresponding to a pixel size of approximately 92 m, and was obtained from USGS SDTS - SRTM (Shuttle Radar Topography Mission)

Table 1 - Soil mapping units in the 1:50,000 detailed reconnaissance soil survey map of Sentinela do Sul, RS, Brazil (Klamt et al., 1996).

| Mapping unit | | Taxonomic units | | Drainage Class | Area |
|---|---|---|---|---|---|
| ID | Type | Soil Taxonomy | Brazilian System | | |
| | | | | | % |
| MU1 | Undiffe rentiated group | Typic Hapludult and Humic Hapludult | Argissolo Amarelo e Argissolo Vermelho | Well drained | 19.7 |
| MU2 | Complex | Humic Hapludult and Umbric Dystrochrept | Argissolo Vermelho-Cambissolo | Well drained | 21.5 |
| MU3 | Complex | Humic Hapludult, Umbric Dystrochrept and Udarent | Argissolo Amarelo-Cambissolo-Neossolo Litólico | Well drained | 24.8 |
| MU4 | Consociation | Udarent | Neossolo Litólico | Well drained | 2.4 |
| MU5 | Association | Humic Hapludult, Plintohumult, and Typic Endoaquult | Argissolo Vermelho-Plintossolo-Planossolo | Moderately well drained | 6.8 |
| MU6 | Association | Plintohumult, Typic Endoaquult, and Typic Endoaquent | Plintossolo-Planossolo-Gleissolo Háplico | Poorly drained | 16.0 |
| MU7 | Association | Typic Endoaquent and Entisol | Gleissolo Háplico-Neossolo Aluvial | Poorly drained | 4.3 |
| MU8 | Consociation | Fluvaquentic Humaquept | Gleissolo Melânico | Poorly drained | 4.4 |

(Rabus et al., 2003). The DEM was used, directly or as a component, to calculate other nine soil prediction variables for each pixel: slope gradient, profile curvature, planar curvature, curvature (combination of planar and profile curvature), flow direction, flow accumulation, flow length, stream power index (SPI), and topographic wetness index (TWI) (Wolock & McCabe, 1995). Each of these landform parameters was selected to be tested as explanatory variable because they were expected to represent changes on soil-forming factors and, therefore, are believed to be informative on the occurrence of soil mapping units.

For the establishment of relationships between these explanatory variables and soil distribution, logistic regressions were used because they allow to predict probabilities of occurrence of soil mapping units based on terrain variables, an advantage when compared to other prediction techniques. The multiple logistic regression model is a non-linear transformation of the linear regression, which allows to estimate the probability of occurrence of any number of classes of a dependent variable (in this case, soil mapping units) based on explanatory variables (Hosmer & Lemeshow, 1989).

Data sampling for training points consisted of 7,500 random map observations (one observation per each 3.5 ha) consisting of DEM, DEM derived parameters, and soil classes, as classified in the soil survey. The option for not using the entire area as training points intended to allow to test the results of the prediction in a data set different that the training set of points, which in this case was defined as the entire study area. Hengl & Rossiter (2003) used operator-selected representative sampling points, the option for using random points intended to eliminate subjectivity and to allow simple reproducibility. The data sampled in ArcView 3.2 environment (ESRI, 1999) was exported as tables and analyzed statistically using software Minitab version 11 (Minitab Inc., 1996).

Best sets of logistic regressions explaining the soil distribution were selected based upon two prediction methods: 1) stratified prediction, by modeling the entire study area without stratification of areas by drainage classes; and 2) unstratified prediction, by modeling the entire study area by stratification of areas by drainage classes. This separation in drainage classes was based on TWI values and used a threshold value of TWI = 7.7 (well drained < 7.7 < poorly drained), after applying a 3 pixel by 3 pixel mean filter to the computed TWI map. The threshold value was determined by testing filtered TWI classification looking for classes that would better reproduce drainage classes groups as represented on the original soil map.

Parameters for multiple nominal logistic regressions were calculated in Minitab environment using soil mapping unit as response or dependent variable, classified in eighth nominal classes (UM1 to UM8). A step-by-step procedure was used to obtain the best fit set of logistic regressions, starting with a larger number of variables and excluding the variables considered less related to variations on the response variable. For each prediction method, a set of best fit regressions was selected based on criteria as goodness to fit tests (Pearson and deviance), log-likelihood, odds ratio, and Z test (Hosmer & Lemeshow, 1989). A set of logistic regressions was determined for the unstratified prediction method, and two sets of logistic regressions were selected for the stratified prediction method, respectively for well drained areas and for poorly drained areas. These sets of logistic regressions were used to estimate the probabilities of occurrence of each soil mapping unit. These sets of equations were organized as scripts in ArcView 3.2 environment, assigning a probability value to each pixel for the entire study area and creating eight maps, each of them representing for the entire study area the probability of occurrence of each soil mapping unit. Using the function Map Query of ArcView Spatial Analyst Extension on the probability maps, two soil maps were estimated (one for stratified and one for not stratified procedures) by assigning to every single pixel the denomination of the soil mapping unit that had the larger probability of occurrence on that pixel.

The accuracy of the estimated soil maps was determined by using error matrices (Congalton, 1991) comparing all pixels of the estimated maps to the original soil map, by this way using a larger dataset than in the training points dataset, and intending to evaluate what would be the effects of using logistic regressions for extrapolating prediction of soil occurrence to other areas than only the training points. The four map accuracy indicators were used: 1) overall accuracy, calculated by dividing the number of correctly-classified pixels by the total number of pixels; 2) producer accuracy, which measures how well an area has been classified, or what the proportion of each mapping unit was mapped according to the original soil map; 3) user accuracy, which measures the reliability of the map, or how well the estimated map is reproducing the original map in a specific point; and 4) Cohen's Kappa statistic (Cohen, 1960), which corrects for agreements that would happen by chance between the original soil map (with and without legend simplification) and the estimated soil maps.

## RESULTS AND DISCUSSION

The unstratified prediction of soil mapping units with multiple logistic regressions selected vari-

ables elevation, distance to streams, and TWI to explain the occurrence of soil mapping units; these variables related to water accumulation and water table depth. For the stratified prediction, variables selected to explain the occurrence of soil types in well drained areas were elevation, distance to streams, and slope, showing that soil occurrence in these areas is more related to factors that affect water movement and soil erosion processes. For poorly drained areas the selected variables were elevation, distance to streams, and flow accumulation, indicating that soil distribution in these areas is more related to factors that affect soil drainage and water table depth (Table 2). These sets of best predictors are related to long-term known relationships between soil forming factors, landform, and soil distribution, which associates soil distribution to erosion processes in steep, well drained areas, and with water dynamics, as water table depth (variable elevation) and water movement and accumulation, in poorly drained areas.

Table 2 presents three sets of equations, respectively for unstratified area and for area stratified by drainage classes. Each set of equations allows calculate the probability of occurrence of soil mapping units. For example, from the first column of Table 2 we may extract the following equation: log [p1/p8] = 10.073 - (0.9359 * TWI) - (0.030455 * elevation) + (0.0044564 * distance to streams).

For evaluating the reproducibility of the original soil map, the concordance between the original soil map and the two newly generated maps (stratified and unstratified by drainage classes) (Figure 1) was evaluated using error matrices (Table 3). Considering the major classification errors, without stratification MU1, MU3, MU5, and MU6 were underclassified, while MU4 and MU7 were overclassified. When the study area was stratified, MU2, MU4, MU7, MU7, and MU8 were overclassified and MU3 and MU6 were underclassified.

Attempts to classify the entire landscape never achieved better results than a 48% overall accuracy (Kappa = 36%), obtained without drainage class separation, while the map estimated with separation of drainage classes had an overall accuracy of 45% (Kappa = 31%). Overall accuracy and Kappa indices were considered unsatisfactory in both cases, although they are in the same magnitude that values found by Hengl & Rossiter (2003). Best user accuracy was obtained for MU3 (61.5%), MU5 (54.5%), and MU7 (51.0%) when drainage classes were not separated, and MU8 (86.9%), MU4 (85.4%), and MU6 (67.7%), when drainage classes were separated (Table 3).

For most soil mapping units, higher producer and user accuracy was achieved when the study area

Table 2 - Multiple logistic regressions for estimation of probabilities of occurrence of soil mapping units.

| Predictor | pUM1/pUM8 | pUM2/pUM8 | pUM3/pUM8 | pUM4/pUM8 | pUM5/pUM8 | pUM6/pUM8 | pUM7/pUM8 |
|---|---|---|---|---|---|---|---|
| NO STRATIFICATION - ENTIRE AREA[1] | | | | | | | |
| Intercept | 10.073**[4] | 13.581** | 22.208** | 46.30** | -4.213** | 1.704 | -10.099** |
| TWI | -0.9359** | -1.8560** | -3.3408** | -7.2133** | 1.1683** | 0.1556 | 1.2945** |
| Elevation (m) | -0.030455** | 0.021819** | 0.040814** | 0.037082** | -0.123897** | -0.033091** | -0.0022896** |
| Distance (m) | 0.0044564** | 0.0018649** | 0.0021843** | 0.0031854** | 0.0030610** | 0.0027811** | 0.0047173** |

| Predictor | p1/p4 | | p2/p4 | | p3/p4 | | |
|---|---|---|---|---|---|---|---|
| STRATIFICATION BY DRAINAGE CLASSES - WELL DRAINED AREAS[2] | | | | | | | |
| Intercept | 13.1865** | | 9.2824** | | 6.2110** | | |
| Elevation (m) | -0.087643** | | -0.029264** | | -0.006844* | | |
| Slope (degrees) | -0.37344** | | -0.31386** | | -0.23154** | | |
| Distance (m) | 0.0011651* | | -0.0017638** | | -0.0013055** | | |

| Predictor | p5/p8 | | p6/p8 | | p7/p8 | | |
|---|---|---|---|---|---|---|---|
| STRATIFICATION BY DRAINAGE CLASSES - POORLY DRAINED AREAS[3] | | | | | | | |
| Intercept | 5.038** | | 9.284** | | -0.679* | | |
| Elevation (m) | -0.149197** | | -0.044773** | | -0.030200** | | |
| TWI | 0.3454* | | -0.5519** | | 0.4083** | | |
| Flow accumulation | -1.437 E-04** | | -1.399 E-0.4** | | -2.893 E-0.4** | | |

[1]Log-likelihood = -9007.9; test that all slopes are zero: G = 8661.2, DF = 21, $P < 0.001$. [2]Log-likelihood = -3534.8; test that all slopes are zero: G = 2135.6, DF = 9, $P < 0.001$. [3]Log-likelihood = -1970.8; test that all slopes are zero: G = 691.8, DF = 9, $P < 0.001$. *indicates $P < 0.05$ and **indicates $P < 0.01$.

Table 3 - Error matrices comparing the field surveyed soil map with the original legend with soil maps estimated either with stratification or without stratification by drainage classes.

| | | ORIGINAL SOIL MAP | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | *Unstratified by drainage classes* | | | | | | | | |
| | | MU1 | MU2 | MU3 | MU4 | MU5 | MU6 | MU7 | MU8 | Total | UA |
| ESTIMATED MAP | MU1 | 11.9 | 4.9 | 0.6 | 0.0 | 0.2 | 5.1 | 1.2 | 1.0 | 25.0 | 47.6 |
| | MU2 | 2.9 | 8.0 | 6.0 | 0.1 | 0.0 | 1.9 | 0.1 | 1.7 | 20.7 | 38.8 |
| | MU3 | 1.0 | 8.2 | 18.1 | 1.5 | 0.0 | 0.4 | 0.0 | 0.2 | 29.4 | 61.5 |
| | MU4 | 0.0 | 0.0 | 0.2 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.3 | 46.6 |
| | MU5 | 0.1 | 0.0 | 0.0 | 0.0 | 2.5 | 1.1 | 0.2 | 0.6 | 4.5 | 54.5 |
| | MU6 | 3.4 | 0.9 | 0.1 | 0.0 | 1.4 | 6.8 | 4.2 | 0.9 | 17.8 | 38.2 |
| | MU7 | 0.1 | 0.0 | 0.0 | 0.0 | 0.3 | 0.6 | 1.1 | 0.1 | 2.2 | 51.0 |
| | MU8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 |
| | Total | 19.5 | 22.1 | 24.9 | 1.7 | 4.4 | 16.0 | 6.8 | 4.5 | 100.0 | |
| | PA | 60.9 | 36.4 | 72.4 | 8.4 | 55.4 | 42.6 | 16.4 | 0.0 | | |
| | OA = 48% | | | | | K = 36% | | | | | |
| | | *Stratified by drainage classes* | | | | | | | | |
| | | MU1 | MU2 | MU3 | MU4 | MU5 | MU6 | MU7 | MU8 | Total | UA |
| ESTIMATED MAP | MU1 | 7.0 | 3.8 | 0.5 | 0.0 | 0.1 | 1.9 | 0.3 | 0.7 | 14.2 | 48.8 |
| | MU2 | 2.3 | 5.3 | 4.7 | 0.0 | 0.0 | 1.1 | 0.1 | 0.7 | 14.3 | 37.3 |
| | MU3 | 1.0 | 8.9 | 18.3 | 1.6 | 0.0 | 0.3 | 0.0 | 0.3 | 30.5 | 60.1 |
| | MU4 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 14.6 |
| | MU5 | 0.1 | 0.0 | 0.0 | 0.0 | 2.5 | 1.2 | 0.3 | 0.6 | 4.6 | 53.5 |
| | MU6 | 8.8 | 3.3 | 1.0 | 0.0 | 1.8 | 10.6 | 5.5 | 1.9 | 32.9 | 32.3 |
| | MU7 | 0.2 | 0.0 | 0.0 | 0.0 | 0.2 | 0.7 | 0.6 | 0.1 | 1.8 | 34.3 |
| | MU8 | 0.2 | 0.7 | 0.2 | 0.0 | 0.0 | 0.3 | 0.0 | 0.2 | 1.6 | 13.1 |
| | Total | 19.6 | 22.0 | 24.9 | 1.7 | 4.5 | 16.0 | 6.8 | 4.5 | 100.0 | |
| | PA | 35.5 | 24.2 | 73.5 | 1.2 | 54.8 | 66.3 | 8.9 | 4.8 | | |
| | OA = 45% | | | | | K = 31% | | | | | |

UA = user accuracy, PA = producer accuracy, OA = overall accuracy, K = Kappa index
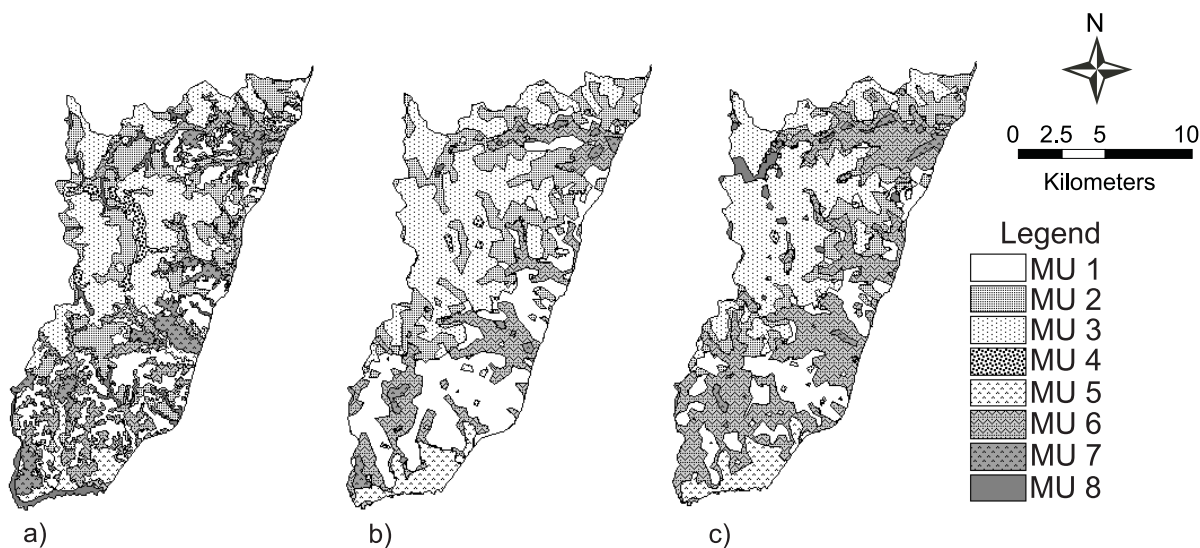


Figure 1 - Field surveyed soil map with original legend (a), and maps estimated with similar legend, respectively without stratification by drainage classes (b) and with stratification by drainage classes (c).

was not stratified. Although it would be expected higher accuracy when separating areas by drainage classes, this was not observed. Given the characteristics of the SRTM DEM, its precision may not have reproduced small variations of elevation, and its resolution may have not showed the actual terrain variations at short distance. Significant differences with stratification of a study area in hill land and plains were obtained by Hengl & Rossiter (2003), who affirm that although stratification has advantages, "it would be more practical to develop a single data set and predictive map of the entire area at once".

To improve the capacity to reproduce the original soil map, the legend was simplified by grouping similar mapping units based on their higher taxonomic categorical level, aiming to verify if a simplified categorical legend, which would usually correspond to a cartographic scale reduction, could be better predicted. The simplified legend joined mapping units MU2, MU3, and MU4 (reclassified as MU24) and mapping units MU6, MU7, and MU8 (reclassified as MU68). Thus, the simplified legend was formed by mapping units MU1, MU24, MU5, and MU68. Same procedures with and without stratification of the area by drainage classes were used for estimating multiple logistic regressions and evaluating map accuracy.

For the maps estimated using the simplified legend (Figure 2), the procedure without stratification by drainage classes had an overall accuracy of 71% (Kappa = 54%), which is an increase of 48% in relation to the same procedure without stratification using the original legend. The map estimated with stratification of the area by drainage classes had overall accuracy of 68% (Kappa = 51%), an increase of 51% in comparison to when the original legend was used. Overall percent correct and Kappa were similar for situations with or without stratification of the area by drainage classes. Higher user accuracy was obtained with stratification by drainage classes in most of the mapping units, with values of 83.7% for MU24 and 70.5% for MU68. Higher producer accuracies were obtained for the procedure without stratification by drainage classes for soil mapping units MU24 (86.6%) (Table 4), and major errors were underclassification of MU1 and overclassification of MU68 when not stratifying by drainage classes. When stratification was used, major errors were overestimation of MU1 and underestimation of occurrence of MU68.

Although the use of a simplified legend makes the soil map lose precision (more soil classes included in a map unit), it makes the soil map gain accuracy, i.e., the capacity to reproduce a reference soil map, either using an original field surveyed or a map with simplified legend.

Table 4 - Error matrices comparing the field surveyed soil map with simplified legend with soil maps estimated either with stratification or without stratification by drainage classes.

| | | ORIGINAL SOIL MAP | | | | | |
|---|---|---|---|---|---|---|---|
| | | *Unstratified by drainage classes* | | | | | |
| | | MU1 | MU24 | MU5 | MU68 | Total | UA |
| | MU1 | 11.9 | 5.5 | 0.2 | 7.3 | 24.9 | 47.8 |
| | MU24 | 3.9 | 42.2 | 0.0 | 4.3 | 50.4 | 83.7 |
| | MU5 | 0.1 | 0.0 | 1.7 | 1.9 | 3.7 | 45.8 |
| ESTIMATED MAP | MU68 | 3.5 | 1.0 | 1.7 | 14.8 | 21.0 | 70.5 |
| | Total | 19.4 | 48.7 | 3.6 | 28.3 | 100.0 | |
| | PA | 61.3 | 86.6 | 47.2 | 52.3 | | |
| | OA = 71% | | | | K = 54% | | |
| | | *Stratified by drainage classes* | | | | | |
| | | MU1 | MU24 | MU5 | MU68 | Total | UA |
| | MU1 | 7.0 | 4.4 | 0.1 | 2.9 | 14.3 | 48.8 |
| | MU24 | 3.3 | 39.0 | 0.0 | 2.6 | 44.9 | 86.8 |
| | MU5 | 0.1 | 0.0 | 2.5 | 2.0 | 4.6 | 53.5 |
| ESTIMATED MAP | MU68 | 9.2 | 5.3 | 1.9 | 19.8 | 36.3 | 54.7 |
| | Total | 19.6 | 48.7 | 4.5 | 27.3 | 100.0 | |
| | PA | 35.5 | 80.1 | 54.8 | 72.6 | | |
| | OA = 68% | | | | K = 51% | | |

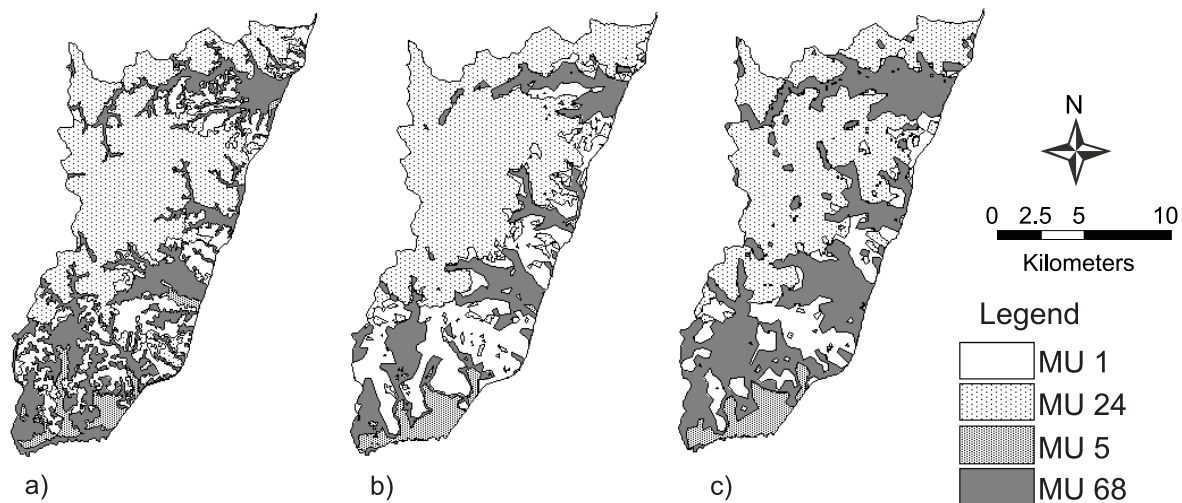UA = user accuracy, PA = producer accuracy, OA = overall accuracy, K = Kappa index

Figure 2 - Field surveyed soil map with simplified legend (a), and maps produced with similar legend, respectively without stratification by drainage classes (b) and with stratification by drainage classes (c).

## REFERENCES

CAMPLING, P.; GOBIN, A.; FEYEN, J. Logistic modeling to spatially predict the probability of soil drainage classes, **Soil Science Society of America Journal**, v.66, p.1390-1401, 2002.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, v.20, p.37-46, 1960.

CONGALTON, R.G. A review of assessing the accuracy of classification of remotely sensed data. **Remote Sensing of Environment**, v.37, p.35-46, 1991.

ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE - ESRI. **ArcView 3.2.** Redland, California, 1999.

HENGL, T.; ROSSITER, D.G. Supervised landform classification to enhance and replace photo-interpretation in semi-detailed soil survey. **Soil Science Society of America Journal**, v.67, p.1810-1822, 2003.

HOSMER, D.W.; LEMESHOW, S. **Applied logistic regression**. New York: John Wiley & Sons, 1989. 307p.

KLAMT, E.; SCHNEIDER, P.; KÄMPF, N.; GIASSON, E.; BASTOS, C.A.B.; LIMA, V.S.; GREHS, S.A.; HESSELN, N.E. **Levantamento de reconhecimento de alta intensidade dos solos do município de Sentinela do Sul, RS**. Porto Alegre: Departamento de Solos, UFRGS, 1996. (Relatório Técnico).

KING, D.; BOURENNANE, H.; ISAMBERT, M.; MACAIRE, J. Relationship of the presence of a non-calcareous clay-loam horizon to DEM attributes in a gently sloping area. **Geoderma**, v.89, p.95-111, 1999.

KRAVCHENKO, A.N.; BOLLERO, G.A.; OMONODE, R.A.; BULLOCK, D. Quantitative mapping of soil drainage classes using topographical data and soil electrical conductivity. **Soil Science Society of America**, v.66, p.235-243, 2002.

LAGACHERIE, P.; ROBBEZ-MASSON, J.; NGUYEN-THE, N.; BARTHS, J. Mapping of reference are a representativity using a mathematical soilscape distance. **Geoderma**, v.101, p.105-118, 2001.

MINITAB INC. **Minitab User's Guide**. Release 11 for Windows. State College: Minitab Inc., 1996.

OHLMACHER, G.C.; DAVIS, J.C. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. **Engineering Geology**, v.69, p.331-343, 2003.

RABUS, B.; EINEDER, M.; ROTH, A.; BAMLER, R. The shuttle radar topography mission - a new class of digital elevation models acquired by spaceborne radar. ISPRS. **Journal of Photogrammetric and Remote Sensing**, v.57, p.241-262, 2003.

SCHNEIDER, P.; KLAMT, E. Necessidades e perspectivas em levantamentos de solos no Rio Grande do Sul.. In: SIMPÓSIO BRASILEIRO SOBRE ENSINO DE SOLOS, 2., Santa Maria, 1995. **Proceedings**. Santa Maria: SBCS, UFSM, 1996.

SILVEIRA, R.J.C. da. Variabilidade das características dos solos e relações solo-superfícies geomórficas na Encosta do Sudeste do Rio Grande do Sul. 139f. Porto Alegre:UFRGS, 1984. 139p. (Dissertation - M.Sc.).

WOLOCK, D.M.; McCABE, G.J. Comparison of single and multiple flow-direction algorithms for computing topographic parameters in TOPMODEL: **Water Resources Research**, v.31, p.1315-1324, 1995.