

Individualization of soil classes by disaggregation of physiographic map polygons

Abstract – The objective of this work was to disaggregate the polygons of physiographic map units in order to individualize the soil classes in each one, representing them as simple soil map units and generating a more detailed soil map than the original one, making these data more useful for future reference. A physiographic map, on a 1:25,000 scale, of the Tarumãzinho watershed, located in the municipality of Águas Frias, in the state of Santa Catarina, Brazil, was used. For disaggregation, three geomorphometric parameters were applied: slope and landforms, both derived from the digital terrain model; and an elevation map. The boundaries of the physiographic units and the elevation, slope, and landform maps were subjected to cross tabulation to identify the existing combinations between the soil classes of each physiographic unit. Based on these combinations, rules were established to select typical areas of occurrence of each soil type in order to train a decision tree model to predict the occurrence of soil classes. The model was trained using the Weka software and was validated with a set of georeferenced soil profiles. Disaggregation enables the individualization and spatialization of soil classes and is useful in producing detailed soil maps.

Index terms: decision trees, digital soil mapping, pedology, soil class prediction.

Individualização de classes de solos por desagregação de polígonos de mapa fisiográfico

Resumo – O objetivo deste trabalho foi desagregar os polígonos de mapas de unidades fisiográficas, de modo a individualizar as classes de solos ocorrentes em cada unidade, para representá-las como unidades de mapeamento simples de solos e gerar um mapa de solos com maior detalhe cartográfico que o mapa original, ampliando a utilidade desses dados em demandas futuras. Foi utilizado um mapa fisiográfico, em escala 1:25.000, da microbacia Córrego Tarumãzinho, localizada no Município de Águas Frias, no Estado de Santa Catarina. Para realizar a desagregação, foram utilizados três parâmetros geomorfométricos: declividade e formas do terreno, ambas derivadas do modelo digital do terreno; e mapa de elevação. Os limites das unidades fisiográficas e os mapas de elevação, declividade e formas do terreno foram submetidos à tabulação cruzada para identificar as combinações existentes entre as classes de solos que compõem cada unidade fisiográfica. A partir dessas combinações, foram elaboradas regras para selecionar áreas de ocorrência típica de cada tipo de solo, para treinar um modelo de árvores de decisão para predição da ocorrência das classes de solos. O treinamento do modelo foi realizado no programa Weka, e a sua validação foi feita com um conjunto de perfis de solos georreferenciados. A desagregação possibilita a individualização e a espacialização das classes de solos e é útil para a produção de mapas de solos detalhados.

Termos para indexação: árvores de decisão, mapeamento digital de solos, pedologia, predição de classes de solos.

José Janderson Ferreira Costa⁽¹⁾  ,
Elvio Giasson⁽¹⁾  ,
Elisângela Benedet da Silva⁽²⁾  ,
Alcinei Ribeiro Campos⁽¹⁾  ,
Israel Rosa Machado⁽¹⁾  ,
Benito Roberto Bonfatti⁽¹⁾  ,
Ivan Luiz Zilli Bacic⁽²⁾  

⁽¹⁾ Universidade Federal do Rio Grande do Sul, Faculdade de Agronomia, Departamento de Solos, Avenida Bento Gonçalves, nº 7.712, CEP 91540-000 Porto Alegre, RS, Brazil.
E-mail: janderson.ferreirac@gmail.com, giasson@ufrgs.br, alcineicampos@gmail.com, israel-mac@hotmail.com, brbonfatti@yahoo.com.br

⁽²⁾ Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina, Centro de Informações de Recursos Ambientais e de Hidrometeorologia de Santa Catarina, Rodovia Admar Gonzaga, nº 1.347, Itacorubi, Caixa Postal 502, CEP 88034-901 Florianópolis, SC, Brazil.
E-mail: elisbenedetsilva@gmail.com, ivan.zilli@gmail.com

 Corresponding author

Received
November 6, 2017

Accepted
May 31, 2019

How to cite
COSTA, J.J.F.; GIASSON, E.; SILVA, E.B. da;
CAMPOS, A.R.; MACHADO, I.R.; BONFATTI
B.R.; BACIC, I.L.Z. Individualization of soil
classes by disaggregation of physiographic
map polygons. *Pesquisa Agropecuária
Brasileira*, v.54, e00290, 2019. DOI: <https://doi.org/10.1590/S1678-3921.pab2019.v54.00290>.



This is an open-access article distributed under the
Creative Commons Attribution 4.0 International License

Pesq. agropec. bras., Brasília, v.54, e00290, 2019
DOI: 10.1590/S1678-3921.pab2019.v54.00290

Introduction

Detailed knowledge regarding the spatial distribution of soil is vital for soil management programs aimed at monitoring the environment and managing agricultural production. However, information on soil is often scarce or only available on a small scale, limiting its usefulness. Factors such as reduced investments in soil surveys and in the time spent on soil research have contributed to this scenario (Arruda et al., 2016; Regmi & Rasmussen, 2018; Vincent et al., 2018).

In this context, it is important to take advantage of existing information that explains the occurrence of soil in the landscape. Physiographic maps (PMs) are a source of soil data and consist of physiographic units (PUs), each exhibiting a degree of unique pedological homogeneity in relation to neighboring units. However, the lack of detail and the insufficient information provided by these maps limits their application, precluding their use in projects that require in-depth knowledge of soil distribution, such as the assessment of ecosystem services (Vincent et al., 2018). An example is the technical report of Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (Epagri, 2004), which describes the spatial distribution of the soils in the landscape, but does not represent graphically the soil classes on the produced map. Despite this, PMs may still be important starting points for producing more detailed soil maps. A possible approach is the spatial disaggregation of the polygons of the combined PUs to better represent the spatial distribution of soils by individualizing and locating soil types in the landscape (Häring et al., 2012; Li et al., 2012; Odgers et al., 2014; Sarmento et al., 2017; Machado et al., 2018; Vincent et al., 2018).

Disaggregation is a conceptual approach aimed at converting current data into formats compatible with modern needs and the pedological concepts of soil formation (Bui & Moran, 2001), in an attempt to update soil maps (Wei et al., 2010; Smith et al., 2012) and distinguish between soil classes within the boundaries of combined map units (MUs) in less detailed soil maps.

Soil-landscape relationships and quantitative prediction are frequently used in soil map disaggregation. They are important tools in characterizing the spatial variability of the soil resource and in understanding its distribution in the landscape, since the association between predictor variables and soil classes has

significant potential to improve predictions of the occurrence of soil types in the landscape. Häring et al. (2012) used soil-landscape relationships based on environmental covariates in the spatial disaggregation of complex soil map units. According to the authors, this method is based on the functional relationship between different soil types and their position in the landscape, in line with the concept of toposequences.

Decision trees (DTs) are a useful tool in the disaggregation of conventional soil maps (Wei et al., 2010; Häring et al., 2012; Li et al., 2012; Sarmento et al., 2017) and in geology (Bui & Moran, 2001) because of their ability to transcribe the complex relationship between soil and landscape (Walter et al., 2006), as well as to deal with qualitative and quantitative predictor variables (Scull et al., 2005). DTs are hierarchical sequential structures that partition data into subsets: each internal node represents a test of a covariate; each branch, a result of the test; and each leaf, a soil class (Moran & Bui, 2002; Grinand et al., 2008; Crivelenti et al., 2009). Vincent et al. (2018) used DTs and expert knowledge to spatially disaggregate soil MUs of the Brittany region in Northwest France. The authors retrieved information on soil-landscape relationships from an existing soil database and used it for disaggregation, combined with a set of covariates.

The objective of this work was to disaggregate the polygons of physiographic map units in order to individualize the soil classes in each one, representing them as soil map units and generating a more detailed soil map than the original one, making these data more useful for future reference.

Materials and Methods

The study was conducted in the Tarumãzinho watershed, located in the municipality of Águas Frias, in the west of the state of Santa Catarina, Brazil (Figure 1), in an area of approximately 27.7 km², containing five main reliefs: flat, slightly undulating, undulating, strongly undulating, and mountainous (Potter et al., 2004). The climate in the region is classified as humid subtropical, Cfa according to Köppen's classification, with an average annual rainfall of 1,725 mm and an average annual temperature of 18.7°C. The parent material is basalt from the Serra Geral formation. The PU, on a 1:25,000 scale, and the technical report on the watershed describe the predominant soil types in the PUs and their associations (Epagri, 2004). Although

the map does not depict the spatial distribution of the soil, this information is provided in the report.

Through photo interpretation and a field survey of the study area conducted by Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (Epagri, 2004), six physiographic units were identified (Figure 1), taking into account their depositional or erosional origin: erosional summit (ES), erosional slope (ESL), erosional colluvial slope (ECS), colluvial erosional slope (CES), plateau slope (PS), alluvial colluvial valley bottom (ACVB).

The ESs are slightly undulating, subrounded, and flat forms with varying widths, at high altitudes (500 to 655 m), with a wide elongated base. This PU is located at a steeper gradient subject to constant soil loss mainly in deforested areas. Erosive processes (material losses) predominated over the depositional ones on ECSs and determined the current features of this unit, which is uniquely shaped and located immediately below ESLs and/or PSs. CESs are formed after ECSs, where the dominant process is the gradual deposition of sediments via rain and sheet erosion, which accumulate downhill

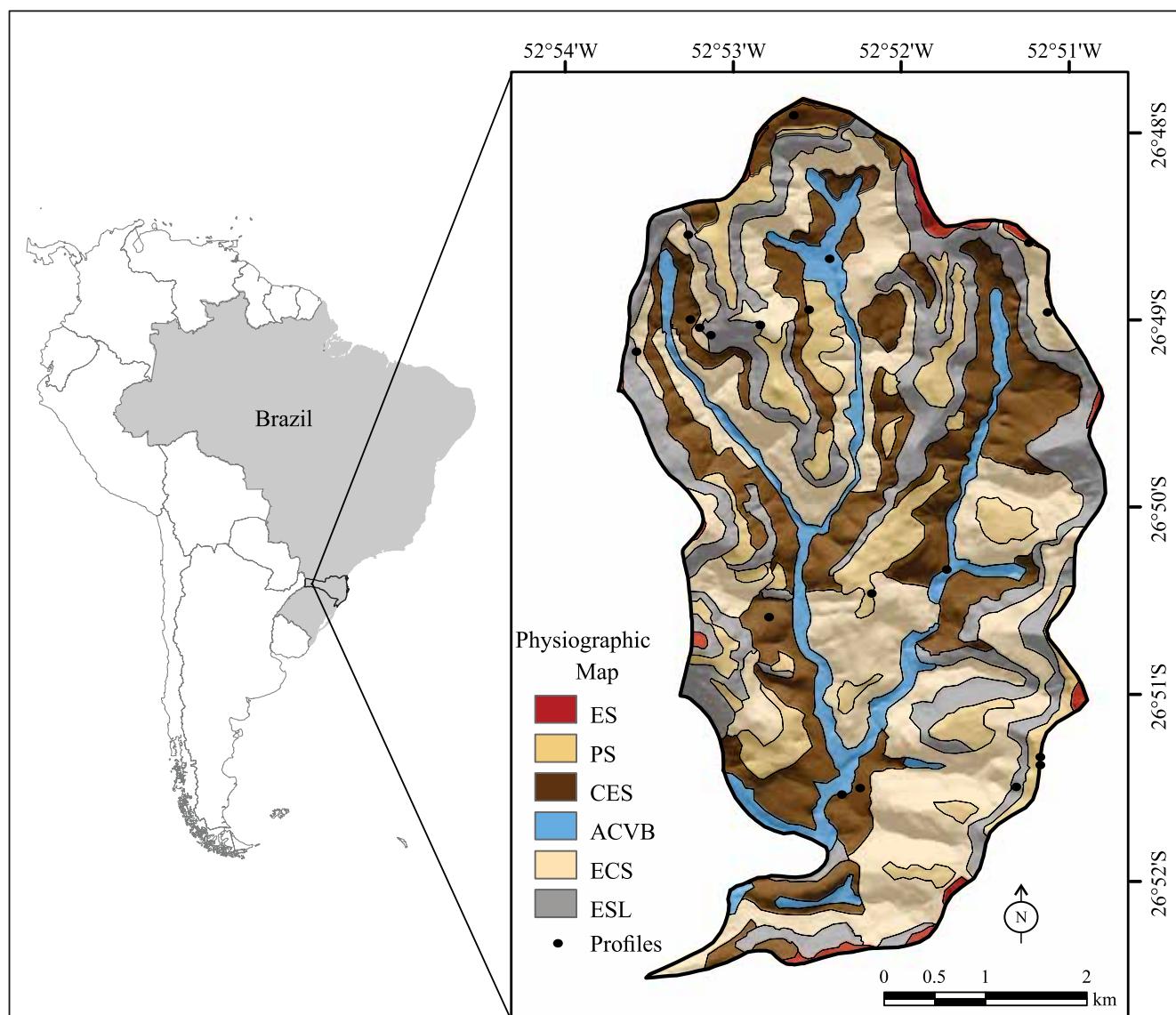


Figure 1. Physiographic map of the Tarumãzinho watershed, located in the municipality of Águas Frias, in the state of Santa Catarina, Brazil. Components: ES, erosional summit; PS, plateau slope; CES, colluvial erosional slope; ACVB, alluvial colluvial valley bottom; ECS, erosional colluvial slope; and ESL, erosional slope.

through the force of gravity and give rise to this PU. The PSS result from the differentiation of basalt flows and their internal variation: the upper section of basalt is more easily weathered in relation to the intermediate zone, generating a topographic rebound. These alternating factors lead to the formation of the steps that characterize the plateaus. The ACVB consists of the lowest altitude and lowest gradient areas of the watershed in the form of flat-floored valleys influenced by the Tarumãzinho stream.

The adopted methodology proposes creating disaggregation rules based on formal knowledge about soil-landscape relationships, as described by Häring et al. (2012), Odgers et al. (2014), and Sarmento et al. (2017), among others. However, adjustments were made to overcome the lack of a soil structure database to aid in the construction of rules, since the aforementioned studies all used a less detailed soil map to be disaggregated.

The first phase in the procedure involved analyzing the PM and the technical report, in order to identify the total number of PU polygons on the PM, the number of PUs related to one or more soil classes, references to possible inclusions of other soils, the main characteristics and properties of each PU, as well as the location of representative profiles and their analytical and morphological data. Data analysis took into account the reported distribution of soil classes in the landscape according to altitude, slope, and the other characteristics described in the technical report.

Based on this information, it was found that the PM consisted of 87 polygons representing the spatial

distribution of six PUs, four of which were single and two combined MUs. One of the combined MUs contained two soil classes and the other three, all in the form of soil associations, but the soils were not mapped individually due to scale limitations. The area occupied by the PUs with more than one soil class corresponds to 25.7% of the total area of the watershed (Table 1).

A digital terrain model (DTM) with a spatial resolution of 1 m was used to derive the geomorphometric parameters. The DTM was obtained from the aerophotogrammetric survey for the state of Santa Catarina, conducted by Secretaria de Estado do Desenvolvimento Econômico Sustentável (Santa Catarina, 2012), using resampling with bilinear interpolation techniques and Geotiff images. It was generated based on the digital elevation model (DEM), according to criteria such as visual analysis of existing features, creation of filtering masks, and softening and adjusting points from the original DEM within the created masks, among others. Based on the DTM, the ArcGIS, version 10.4, software (ESRI, 2015) was used to generate a slope class map representing the types of relief (Manual..., 2015), a landform class map (MacMillan, 2003), and an elevation map to identify the altitude at which soil classes occur.

The PU boundaries and the slope class, elevation, and landform maps were subjected to cross tabulation to generate a table with a combination of these covariates. Cross tabulation, also known as a contingency table, simultaneously describes two or more predictive covariates and reflects their joint

Table 1. Description of the physiographic units (PUs) and soils in the Tarumãzinho watershed, located in the municipality of Águas Frias, in the state of Santa Catarina, Brazil.

PU ⁽¹⁾	Brazilian system ⁽²⁾	Soil taxonomy ⁽³⁾	Inclusions	Area	
				(%)	(ha)
ES	Nitossolo Háplico (NX) + Nitossolo Vermelho (NV)	Ultisol (NX) + Ultisol (NV)	Inceptisol	1.4	38
CES	Cambissolo Háplico (CX) + Nitossolo Háplico (NX)+ Nitossolo Vermelho (NV)	Inceptisol (CX) + Ultisol (NX) + Ultisol (NV)	-	24.3	669
ESL	Neossolo Litólico	Entisol	Inceptisol	21.1	580
ECS	Cambissolo Háplico	Inceptisol	Ultisol	32.7	900
PS	Cambissolo Háplico	Inceptisol	-	12.3	339
ACVB	Cambissolo Háplico	Inceptisol	Aquent	8.2	226
Total				100.0	2,752

⁽¹⁾ES, erosional summit; CES, colluvial erosional slope; ESL, erosional slope; ECS, erosional colluvial slope; PS, plateau slope; and ACVB, alluvial colluvial valley bottom. Source: (Epagri, 2004). ⁽²⁾According to Santos et al. (2013). ⁽³⁾According to Soil Taxonomy (Soil Survey Staff, 2014).

distribution. This analysis aimed to individualize the covariate combinations in the PUs where each soil class occurs. The most common combinations among the PUs and the slope, elevation, and landform maps were then selected in cross tabulation. Based on these combinations, rules were created to individualize the typical areas of occurrence of each soil type. For example, according to the technical report, in CES, Cambissolos Háplicos (Inceptisols) are expected to occur on convergent back slopes (CBS) and on back slopes (BSL) with a 9 to 12% gradient (Figure 2 and Table 2); therefore, several individual rules were made for each soil component (Table 2).

The set of rules was applied to all PUs because the used report refers to associated soils as inclusions in single PUs. The rules were translated into logical expressions to select areas that correspond to PU boundaries, slopes,

landforms, and elevation, representing the locations (pixels) that depict each soil type.

These typical areas were then converted into vector polygons to facilitate the creation of sampling points used to train the predictive model. Random sampling stratified by the number of PUs was used, and a point density value was established and applied to all polygons to produce a similar point density, allowing a better distribution (Odgers et al., 2014).

The predictive covariates and PUs containing spatial distribution information on the soil classes were sampled at 7,085 points on the map, which is equivalent to a sampling density of 2.5 points per hectare. Values for all the predictive covariates were sampled at these points. The data were compatibilized with the Weka, version 3.6.6, data mining software (Hall et al., 2009) and used to train the J48 algorithm (Quinlan, 1993) of

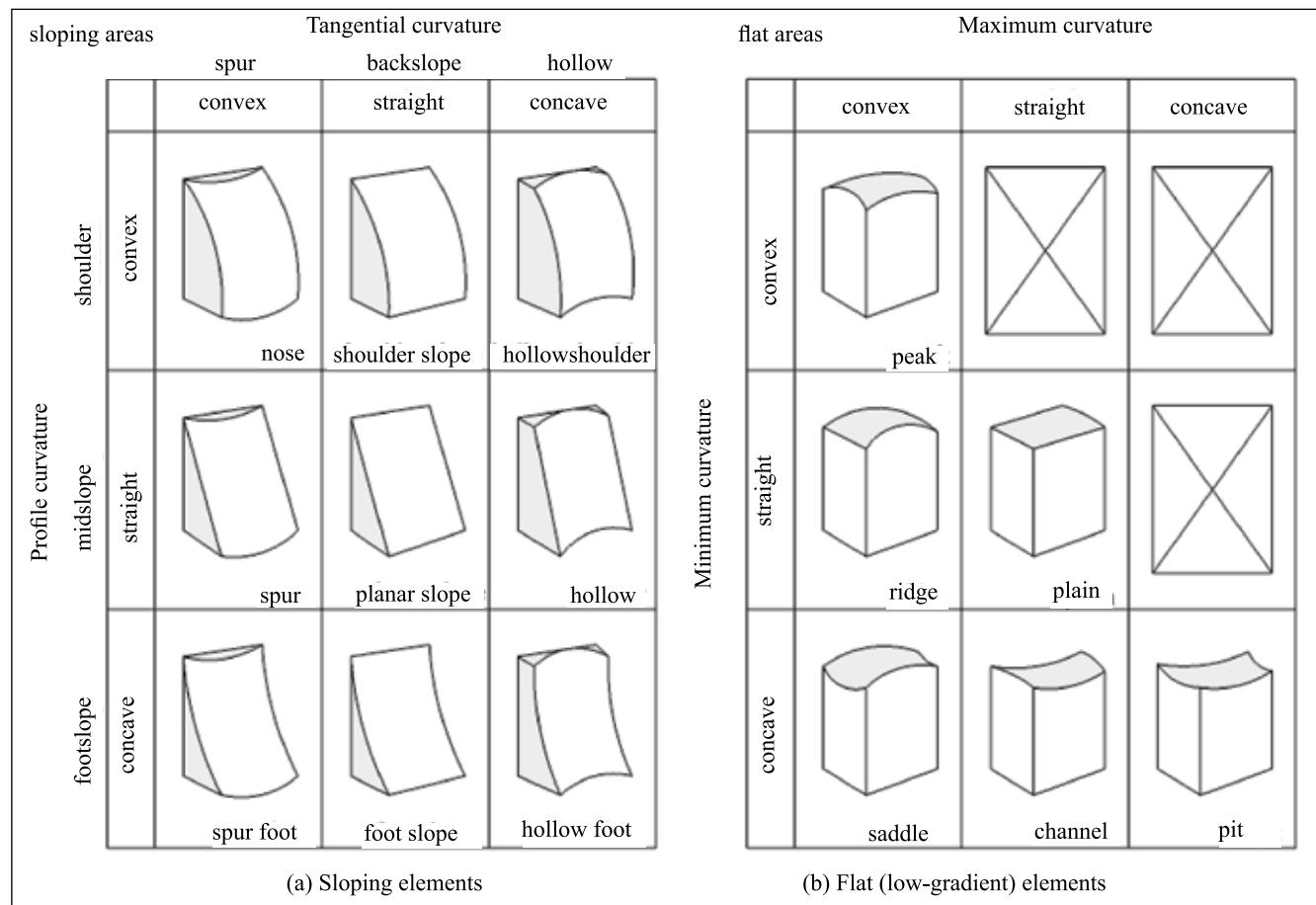


Figure 2. Fifteen landforms (fundamental elements) based on slope and curvature classification. Source: adapted from Schmidt & Hewitt (2004).

the DT, which was chosen because it allowed obtaining better results than the other algorithms used in digital soil mapping (Moran & Bui, 2002; Ten Caten et al., 2013; Dias et al., 2016). In order to implement the DT in an ArcGIS environment, the classification rules were adjusted for direct application in the geographic information system to predict the spatial distribution of soil classes in the disaggregated PM.

The model was validated based on a set of 20 georeferenced soil profiles (Figure 1), not used in the DT construction, considering the spatial agreement between the georeferenced profiles and the location of individual pixels on the disaggregated map with a 1 (5 m) and 2 pixel (10 m) search radii (buffers). In this case, the prediction was deemed correct when the soil class predicted in the search radius was the same as that observed in the reference profile (Smith et al., 2012; Nauman & Thompson, 2014). The total size of the soil classes of the disaggregated PUs on the PM was compared with that of the PUs from the reference map because the size of map units varies significantly and can result in varying effects. The calculation considered

the absolute error, the relative error between the fraction of the reference map and the disaggregated map, and a visual comparison of the disaggregated polygons (Li et al., 2012; Sarmento et al., 2017).

Results and Discussion

The reference and soil maps produced by disaggregating the polygons of combined PUs from the PM showed good agreement, evident by visual comparison (Figure 3 B and C). Although the proportion of individual soil classes predicted per unit on the disaggregated map was relatively similar to that of the research report (Table 3), with a mean relative error of 0.09 ha, it was underestimated, with a mean absolute error of 20.37 ha. This may be because some soil classes can occur as inclusions that were not mapped on the original map due to scale limitations. In the adopted methodology, overestimating a soil class in a given PU results in the underestimation of the other classes (Sarmento et al., 2017). Using a similar technique, Li et al. (2012) obtained a mean relative error

Table 2. Soil class individualization rules for the polygons of combined physiographic units (PUs), based on slope, elevation, and landform (LF) class maps and the boundaries of polygons from the physiographic map.

PU ⁽¹⁾	Slope (%)	Elevation (m)	LF ⁽²⁾	Soil taxonomy ⁽³⁾	Legend ⁽⁴⁾	Brazilian system ⁽⁴⁾
ES	<5	655	DSH	Ultisols	NX	Nitossolo Háplico Eutrófico típico
	<3	603	FSL	Ultisols	NV	Nitossolo Vermelho Distrófico típico
	0–3	509	DSH	Ultisols	NV	Nitossolo Vermelho Eutrófico típico
CES	>12	480	CBS	Inceptisols	CX	Cambissolo Háplico Alumínico típico
	9–12	320	BSL	Inceptisols	CX	Cambissolo Háplico Ta Eutrófico típico
	10–20	387	DBS	Ultisols	NX	Nitossolo Háplico Eutrófico típico
	<9	380	DBS	Ultisols	NV	Nitossolo Vermelho Eutrófico latossólico
ESL	<27	516	DBS	Entisols	RL	Neossolo Litólico Distrófico típico
	>46	496	DSH	Entisols	RL	Neossolo Litólico Eutrófico típico
ECS	8–20	437	CBS	Inceptisols	CX	Cambissolo Háplico Eutrófico típico
	>30	557	DSH	Inceptisols	CX	Cambissolo Háplico Ta Eutrófico típico
	<23	483	DBS	Inceptisols	CX	Cambissolo Háplico Eutrófico típico
PS	>6	575	DSH	Inceptisols	CX	Cambissolo Háplico Ta Eutrófico típico
	>11	530	CBS	Inceptisols	CX	Cambissolo Háplico Ta Eutrófico típico
	<6	478	DSH	Inceptisols	CX	Cambissolo Háplico Ta Eutrófico típico
ACVB	0–3	300	BSL	Inceptisols	CX	Cambissolo Háplico Ta Distrófico léptico
	0–4	450	LSM	Inceptisols	CX	Cambissolo Háplico Ta Eutrófico típico

⁽¹⁾ES, erosional summit; CES, colluvial erosional slope; ESL, erosional slope; ECS, erosional colluvial slope; PS, plateau slope; and ACVB, alluvial colluvial valley bottom. Source: (Epagri, 2004). ⁽²⁾DSH, divergent shoulder; FSL, foot slope; CBS, convergent back slope; BSL, back slope; DBS, divergent back slope; and LSM, lower slope mound. Source: MacMillan (2003). ⁽³⁾According to Soil Taxonomy (Soil Survey Staff, 2014). ⁽⁴⁾According to Santos et al. (2013).

between 0.49 and 0.67 for the relationship between the total areas of disaggregated classes and those recorded in the Canada Land Inventory. However, the authors only took into account the areas of the dominant classes in each MU, whereas the present study considered correct predictions in relation to all the soil classes occurring in the PUs.

The lowest relative errors were recorded in the ECS, ESL, and CES (Table 3). Although the ECS unit contains only one soil class, namely Cambissolo Háplico (Inceptisol), it occupies a 900 ha area on the original map, around 32.7% of the total area of the watershed. According to the technical report (Epagri, 2004), the ESL covers 580 ha, i.e., about 21.1% of the total area, consisting of a Neossolo Litólico (Entisol).

On the original map, the total area of the CES is 669 ha – approximately 24.3% of the total area; however, after disaggregation, the sum of the Cambissolo Háplico (Inceptisol) + Nitossolo Vermelho (Ultisol) + Nitossolo Háplico (Ultisol) soil classes was 670.9 ha, that is, an absolute error of only 1.9 ha. These results corroborate those of Vincent et al. (2018), who spatially disaggregated a conventional soil map and found that large areas were overestimated and small ones were underestimated.

The highest relative error was recorded for the sum of the soil classes in the ES (Table 3). On the original map, this PU covered a total area of 38 ha, and, after disaggregation, the sum of the Nitossolo Háplico (Ultisol) + Nitossolo Vermelho (Ultisol) soil classes

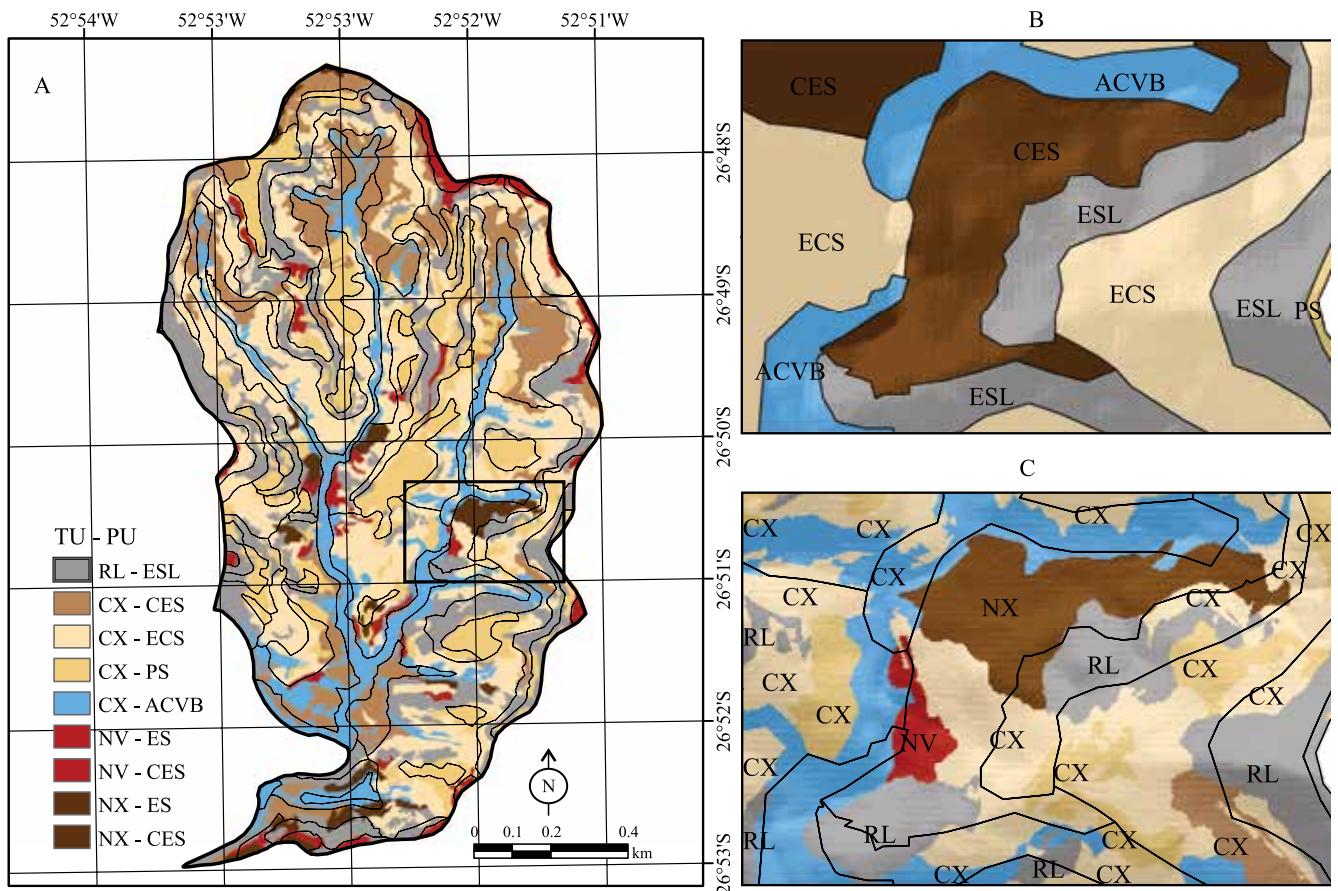


Figure 3. Map of the taxonomic units (TUs) and corresponding physiographic units (PUs) represented by the outlined areas (A); detail of a section of the original physiographic map (B); and detail of a section of the soil map generated by the disaggregation of the polygons of combined PUs on the physiographic map (C). TU components: RL, Neossolo Litólico, an Entisol; CX, Cambissolo Háplico, an Inceptisol; NV, Nitossolo Vermelho, an Ultisol; and NX, Nitossolo Háplico, an Ultisol. PU components: ESL, erosional slope; CES, colluvial erosional slope; ECS, erosional colluvial slope; PS, plateau slope; ACVB, alluvial colluvial valley bottom; and ES, erosional summit.

was 28.7 ha. The observed error may be due to the fact that only one georeferenced point was available for these classes, which hampered accuracy assessment. In addition, the soil classes with a larger number of available georeferenced points showed better agreement with the disaggregated map. Therefore, the most significant errors occurred in the smallest areas, reducing the total error of the disaggregated map.

In relation to the area occupied by soil classes after disaggregation (Figure 4), Cambissolos Háplicos (Inceptisols) predominate in the watershed and occur primarily in the ECS and PS (total area of 1,895 ha), followed by Neossolos Litólicos (Entisols) in the ESL (611 ha), at a 27 to 46% gradient (Epagri, 2004).

Two combined PUs provide an approach that identifies how disaggregation can determine the size of the area occupied by each soil class. The first is the CES (Table 4), which, on the original map, is composed of associated Cambissolo Háplico (Inceptisol), Nitossolo Háplico (Ultisol), and Nitossolo Vermelho (Ultisol), according to the technical report (Epagri, 2004). The dominant soil class in the map generated by the disaggregation of the PM polygons (Table 4) is Cambissolo Háplico (Inceptisol), which occupies around 67.75% of the CES, followed by Nitossolo Vermelho (Ultisol), with 17% of the area, and Nitossolo Háplico (Ultisol), with 15.25%. It should

be noted that Inceptisols are located in areas closest to the valley bottoms and have drainage problems at depths below 1 m, when situated near floodplains. Ultisols are deep, well drained, and found in the areas closest to the toeslopes, with dark-red coloring in the subsurface soil horizons (Epagri, 2004).

The second combined PU is the ES, which covers approximately 1.4% of the total watershed area on the original map and contains associated Nitossolo Vermelho (Ultisol) and Nitossolo Háplico (Ultisol), in alignment with the technical report (Epagri, 2004). Table 4 indicates that Nitossolo Vermelho (Ultisol) is the dominant soil class on the disaggregated map, occupying about 69.36% of the surface of this PU, followed by Nitossolo Háplico (Ultisol), with 30.64%. The Ultisols in this unit are largely located at high altitudes (500 to 655 m) in slightly undulating areas, in the drainage divides of the watershed.

The overall accuracy results of the PM disaggregation are listed in Table 5. The spatial agreement between the georeferenced soil profiles and the location of individual pixels on the disaggregated map showed an average overall accuracy of 44.4% for simple agreement. The overall accuracy increased when correct predictions for a 5- and 10-m search radius were considered, reaching 52.8 and 72.2%, respectively. When validating the model with georeferenced soil profiles in a disaggregation study, Vincent et al. (2018) found 41 to 45% accuracy for simple agreement, which

Table 3. Absolute and relative error values between the areas on the original map (OA) and the total area of the soil classes in the physiographic units (PUs) on the disaggregated physiographic map (DM).

Soil class ⁽¹⁾ (PU) ⁽²⁾	Area of the OA	Area of the DM	Absolute error	Relative error
----- (ha) -----				
NV+NX (ES)	38.0	28.7	9.30	0.24
NV+NX+CX (CES)	669.0	670.9	1.90	0.00
RL (ESL)	580.0	611.1	31.10	0.05
CX (ECS)	900.0	927.9	27.90	0.03
CX (PS)	339.0	311.2	27.80	0.08
CX (ACVB)	226.0	201.8	24.20	0.11
Mean		20.37		0.09

⁽¹⁾NV, Nitossolo Vermelho, an Ultisol; NX, Nitossolo Háplico, an Ultisol; CX, Cambissolo Háplico, an Inceptisol; and RL, Neossolo Litólico, an Entisol. Source: Santos et al. (2013). ⁽²⁾ES, erosional summit; CES, colluvial erosional slope; ESL, erosional slope; ECS, erosional colluvial slope; PS, plateau slope; and ACVB, alluvial colluvial valley bottom. Source: Epagri (2004).

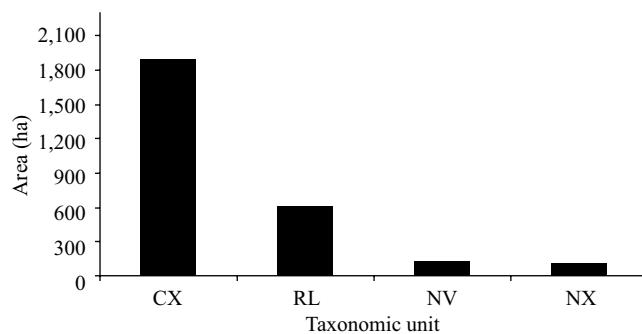


Figure 4. Distribution of the area in relation to the soil classes represented in the simple soil map units on the disaggregated map of the Tarumãzinho watershed, located in the municipality of Águas Frias, in the state of Santa Catarina, Brazil. Soil class: CX, Cambissolo Háplico, an Inceptisol; RL, Neossolo Litólico, an Entisol; NV, Nitossolo Vermelho, an Ultisol; and NX, Nitossolo Háplico, an Ultisol. Source: Santos et al. (2013).

increased to 65 to 72% when a larger search radius was considered. Nauman & Thompson (2014) assessed disaggregation methods in two soil surveys using DTs based on the Gini index, obtaining an overall accuracy of 56.4 to 65.1% for pixel classification on the conventional map in relation to its disaggregated counterpart. This was considered a good result since all the map units contained 10 to 25% of other soils or inclusions that could influence accuracy.

The accuracy values obtained showed that it was possible to generate a soil map that accurately depicted the soil classes in MUs by disaggregating the polygons of combined PUs on a PM. This is because the typical areas of occurrence for each soil type selected to train

the predictive model accurately represented soil spatial distribution in the field. The results showed that the relative errors of the disaggregated fraction in relation to the fraction described in the report are lower in classes covering larger areas, indicating that these soils are distributed as in the report.

Some of the differences between the disaggregated and original maps are not necessarily errors, but rather deviations from one estimate to another in an attempt to extract as much as possible from the existing information. This is done because the values provided in the report are often the only available references on the spatial distribution of soil classes (Odgers et al., 2014).

Table 4. Area occupied by the original physiographic map in relation to that occupied by the soil classes occurring in the physiographic units (PUs) of the disaggregated physiographic map.

Soil class ⁽¹⁾ (PU) ⁽²⁾	Physiographic map (ha)	Legend ⁽¹⁾	Soil class ⁽³⁾	Disaggregated map (ha)	Area / unit (%)
NV+NX (ES)	38	NV	Ultisol	8.8	30.64
		NX	Ultisol	19.9	69.36
		NV	Ultisol	102.3	15.25
NV+NX+CX (CES)	669	NX	Ultisol	114.1	17.00
		CX	Inceptisol	454.5	67.75
CX (PS)	339	CX	Inceptisol	311.2	100.0
CX (ECS)	900	CX	Inceptisol	927.9	100.0
RL (ESL)	580	RL	Entisol	611.1	100.0
CX (ACVB)	226	CX	Inceptisol	201.8	100.0
Total	2,752			2,751.54	

⁽¹⁾NV, Nitossolo Vermelho, an Ultisol; NX, Nitossolo Háplico, an Ultisol; CX, Cambissolo Háplico, an Inceptisol; and RL, Neossolo Litólico, an Entisol. Source: Santos et al. (2013). ⁽²⁾ES, erosional summit; CES, colluvial erosional slope; PS, plateau slope; ECS, erosional colluvial slope; ESL, erosional slope; and ACVB, alluvial colluvial valley bottom. Source: Epagri (2004). ⁽³⁾According to Soil Taxonomy (Soil Survey Staff, 2014).

Table 5. Accuracy values for the agreement between georeferenced soil profiles and the location of individual pixels on the soil map generated by the disaggregation of the polygons of combined physiographic units on the physiographic map of the Tarumãzinho watershed, located in the municipality of Águas Frias, in the state of Santa Catarina, Brazil.

Forms of validation	Disaggregated map
	Overall accuracy (%)
Simple agreement	44.4
1 pixel radius (5 m)	52.8
2 pixel radius (10 m)	72.2

Conclusions

1. Disaggregation enables the individualization and spatialization of soil classes and is useful in producing detailed soil maps.

2. The validation of the soil map with georeferenced soil profiles reveals that the larger the search radius, the higher the overall accuracy values, with 44.4% in simple agreement between soil map pixels and soil profiles, and 52.8 to 72.2% when larger search radius are used.

3. Disaggregation is most successful for colluvial erosional slopes, erosional slopes, and erosional

colluvial slopes, which showed the lowest relative errors in relation to the other physiographic units.

Acknowledgments

To Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), for scholarships awarded to the first, fourth, and fifth authors, and for the research productivity grant awarded to the second author; to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), for scholarships awarded to the third and sixth authors; and to Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (Epagri), for providing some of the material used in the study.

References

- ARRUDA, G.P. de; DEMATTÉ, J.A.M.; CHAGAS, C. da S.; FIORIO, P.R.; SOUZA, A.B. e; FONGARO, C.T. Digital soil mapping using reference area and artificial neural networks. *Scientia Agricola*, v.73, p.266-273, 2016. DOI: <https://doi.org/10.1590/0103-9016-2015-0131>.
- BUI, E.N.; MORAN, C.J. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. *Geoderma*, v.103, p.79-94, 2001. DOI: [https://doi.org/10.1016/S0016-7061\(01\)00070-2](https://doi.org/10.1016/S0016-7061(01)00070-2).
- CRIVELENTI, R.C.; COELHO, R.M.; ADAMI, S.F.; OLIVEIRA, S.R. de M. Mineração de dados para inferência de relações solo-paisagem em mapeamentos digitais de solo. *Pesquisa Agropecuária Brasileira*, v.44, p.1707-715, 2009. DOI: <https://doi.org/10.1590/S0100-204X2009001200021>.
- DIAS, L.M. da S.; COELHO, R.M.; VALLADARES, G.S.; ASSIS, A.C.C. de; FERREIRA, E.P.; SILVA, R.C. da. Predição de classes de solo por mineração de dados em área da bacia sedimentar do São Francisco. *Pesquisa Agropecuária Brasileira*, v.51, p.1396-1404, 2016. DOI: <https://doi.org/10.1590/s0100-204x2016000900038>.
- EPAGRI. Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. **Inventário de terras**: “Microbacia Córrego Tarumãzinho”: Município de Águas Frias, SC. Florianópolis, 2004. 68p. CD-ROM.
- ESRI. Environmental Systems Research Institute. **ArcGIS Desktop**. Version 10.4. Redlands: Environmental Systems Research Institute, 2015. Computer program.
- GRINAND, C.; ARROUAYS, D.; LAROCHE, B.; MARTIN, M.P. Extrapolating regional soil landscapes from an existing soil map: sampling intensity, validation procedures, and integration of spatial context. *Geoderma*, v.143, p.180-190, 2008. DOI: <https://doi.org/10.1016/j.geoderma.2007.11.004>.
- HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I.H. The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*, v.11, p.10-18, 2009. DOI: <https://doi.org/10.1145/1656274.1656278>.
- HÄRING, T.; DIETZ, E.; OSSENSTETTER, S.; KOSCHITZKI, T.; SCHRÖDER, B. Spatial disaggregation of complex soil map units: a decision-tree based approach in Bavarian forest soils. *Geoderma*, v.185-186, p.37-47, 2012. DOI: <https://doi.org/10.1016/j.geoderma.2012.04.001>.
- LI, Z.; HUFFMAN, T.; ZHANG, A.; ZHOU, F.; MCCONKEY, B. Spatially locating soil classes within complex soil polygons – mapping soil capability for agriculture in Saskatchewan Canada. *Agriculture, Ecosystems and Environment*, v.152, p.59-67, 2012. DOI: <https://doi.org/10.1016/j.agee.2012.02.007>.
- MACHADO, I.R.; GIASSON, E.; CAMPOS, A.R.; COSTA, J.J.F.; SILVA, E.B. da; BONFATTI, B.R. Spatial disaggregation of multi-component soil map units using legacy data and a tree-based algorithm in Southern Brazil. *Revista Brasileira de Ciência do Solo*, v.42, e0170193, 2018. DOI: <https://doi.org/10.1590/18069657rbcs20170193>.
- MACMILLAN, R.A. **LandMapR Software Toolkit-C++ version (2003)**: users manual. Edmonton: LandMapper Environmental Solutions, 2003. 110p.
- MANUAL técnico de pedologia. 3.ed. Rio de Janeiro: IBGE, 2015. 430p. (Manuais técnicos em geociências, n.4).
- MORAN, C.J.; BUI, E. Spatial data mining for enhanced soil map modelling. *International Journal of Geographical Information Science*, v.16, p.533-549, 2002. DOI: <https://doi.org/10.1080/13658810210138715>.
- NAUMAN, T.W.; THOMPSON, J.A. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification trees. *Geoderma*, v.213, p.385-399, 2014. DOI: <https://doi.org/10.1016/j.geoderma.2013.08.024>.
- ODGERS, N.P.; SUN, W.; MCBRATNEY, A.B.; MINASNY, B.; CLIFFORD, D. Disaggregating and harmonising soil map units through resampled classification trees. *Geoderma*, v.214-215, p.91-100, 2014. DOI: <https://doi.org/10.1016/j.geoderma.2013.09.024>.
- POTTER, R.O.; CARVALHO, A.P. de; FLORES, C.A.; BOGNOLA, I. **Solos do Estado de Santa Catarina**. Rio de Janeiro: Embrapa Solos, 2004. CD-ROM. Mapa color. (Embrapa Solos. Boletim de pesquisa e desenvolvimento, 46). Anexo Mapa Levantamento de Reconhecimento dos Solos do Estado de Santa Catarina. Escala: 1:250.000.
- QUINLAN, J.R. **C4.5**: programs for machine learning. San Francisco: Morgan Kaufmann, 1993. 302p.
- REGMI, N.R.; RASMUSSEN, C. Predictive mapping of soil-landscape relationships in the arid Southwest United States. *Catena*, v.165, p.473-486, 2018. DOI: <https://doi.org/10.1016/j.catena.2018.02.031>.
- SANTA CATARINA. Secretaria de Estado do Desenvolvimento Econômico Sustentável. **Relatório de concorrência pública nº 0010/2009, para contratação de serviços de aerolevantamento na escala 1:10.000 de aproximadamente 97.037 km², referentes**

- ao Estado de Santa Catarina.** Florianópolis, 2012. Disponível em HD.
- SANTOS, H.G. dos; JACOMINE, P.K.T.; ANJOS, L.H.C. dos; OLIVEIRA, V.A. de; LUMBRERAS, J.F.; COELHO, M.R.; ALMEIDA, J.A. de; CUNHA, T.J.F.; OLIVEIRA, J.B. de. **Sistema brasileiro de classificação de solos.** 3.ed. rev. e ampl. Brasília: Embrapa, 2013. 353p.
- SARMENTO, E.C.; GIASSON, E.; WEBER, E.J.; FLORES, C.A.; HASENACK, H. Disaggregating conventional soil maps with limited descriptive data: a knowledge-based approach in Serra Gaúcha, Brazil. **Geoderma Regional**, v.8, p.12-23, 2017. DOI: <https://doi.org/10.1016/j.geodrs.2016.12.004>.
- SCHMIDT, J.; HEWITT, A. Fuzzy land element classification from DTMs based on geometry and terrain position. **Geoderma**, v.121, p.243-256, 2004. DOI: <https://doi.org/10.1016/j.geoderma.2003.10.008>.
- SCULL, P.; FRANKLIN, J.; CHADWICK, O.A. The application of classification tree analysis to soil type prediction in a desert landscape. **Ecological Modelling**, v.181, p.1-15, 2005. DOI: <https://doi.org/10.1016/j.ecolmodel.2004.06.036>.
- SMITH, C.A.S.; DANESHFAR, B.; FRANK, G.; FLAGER, E.; BULMER, C. Use of weights of evidence statistics to define inference rules to disaggregate soil survey maps. In: MINASNY, B.; MALONE, B.P.; MCBRATNEY, A.B. (Ed.). **Digital soil assessments and beyond.** London: Taylor & Francis, 2012. p.215-220.
- SOIL SURVEY STAFF. **Keys to soil taxonomy.** 12th ed. Washington: USDA, 2014. 372p.
- TEN CATEN, A.; DALMOLIN, R.S.D.; PEDRON, F. de A.; RUIZ, L.F.C.; SILVA, C.A. da. An appropriate data set size for digital soil mapping in Erechim, Rio Grande do Sul, Brazil. **Revista Brasileira de Ciência do Solo**, v.37, p.359-366, 2013. DOI: <https://doi.org/10.1590/S0100-06832013000200007>.
- VINCENT, S.; LEMERCIER, B.; BERTHIER, L.; WALTER, C. Spatial disaggregation of complex soil map units at the regional scale based on soil-landscape relationships. **Geoderma**, v.311, p.130-142, 2018. DOI: <https://doi.org/10.1016/j.geoderma.2016.06.006>.
- WALTER, C.; LAGACHERIE, P.; FOLLAIN, S. Integrating pedological knowledge into digital soil mapping. In: LAGACHERIE, P.; MCBRATNEY, A.B.; VOLTZ, M. (Ed.). **Digital soil mapping: an introductory perspective.** Amsterdam: Elsevier, 2006. p.289-310. (Developments in Soil Science, v.31).
- WEI, S.; MCBRATNEY, A.; HEMPEL, J.; MINASNY, B.; MALONE, B.; AVELLO, T.D.; BURRAS, L.; THOMPSON, J. Digital harmonisation of adjacent analogue soil survey areas – 4 Iowa Counties. In: WORLD CONGRESS OF SOIL SCIENCE, 19, 2010, Brisbane. **Soil solutions for a changing world: proceedings.** Brisbane: International Union of Soil Science: ASSSI, 2010. 1 CD-ROM. Editado por R.J. Gikes e N. Prakongkep.
-