

ADRIANO LUIS SCHÜNEMANN

GEOTECNOLOGIAS PARA MAPEAMENTO DIGITAL NA ANTÁRTICA MARÍTIMA

Trabalho de conclusão de curso apresentado ao Departamento de Geotecnologias da Universidade Federal de Viçosa - Campus Viçosa

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Doctor Scientiae*.

Orientador: Eládio Fernando Faria
Tese (Doutorado) - Universidade Federal de Viçosa
Instituição: Viçosa, Minas Gerais, Brasil

1. Geotecnologias 2. Mapeamento digital 3. Antártica 4. Universidade Federal de Viçosa 5. Departamento de Solos 6. Programa de Pós-graduação em Solos e Nutrição de Plantas 7. Título

CDD 22 ed. 624.151

VIÇOSA
MINAS GERAIS - BRASIL
2016

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

S392g Schönemann, Adriano Luis, 1975-
2016 Geotecnologias para mapeamento digital na Antártica
marítima / Adriano Luis Schönemann. – Viçosa, MG, 2016.
ix, 105f. : il. (algumas color.) ; 29 cm.

Orientador: Elpídio Inácio Fernandes Filho.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Geomorfologia. 2. Engenharia geotécnica.
3. Mapeamento digital. 4. Antártida. I. Universidade Federal de
Viçosa. Departamento de Solos. Programa de Pós-graduação em
Solos e Nutrição de Plantas. II. Título.

CDD 22 ed. 624.151

ADRIANO LUIS SCHÜNEMANN

GEOTECNOLOGIAS PARA MAPEAMENTO DIGITAL NA ANTÁRTICA MARÍTIMA

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Doctor Scientiae*.

APROVADA: 03 de novembro de 2016.


Márcio Rocha Francelino
(Coorientador)


Carlos Ernesto Gonçalves
Reynaud Schaefer


Jorge Arigony Neto


Gérson Rodrigues dos Santos


Elpídio Inácio Fernandes Filho
(Orientador)

AGRADECIMENTOS

Ao colega André Thomazini, pelo companheirismo e auxílio nas atividades desenvolvidas nas Operações Antárticas XXXIII e XXXIV, além do processamento de amostras e ensinamentos para escrita científica.

Ao bolsista de iniciação científica, Pedro Henrique Araújo Almeida, pelo auxílio no levantamento da Península Keller, coleta de solos e processamento de dados.

Ao professor Antônio Batista Pereira pelo incentivo incansável para a realização do doutorado e auxílio em todas as atividades de campo realizadas na Península Keller.

A meus co-orientadores, Carlos Schaefer e Márcio Francelino e orientador Elpídio Filho, pelo conhecimento e oportunidades em diferentes trabalhos científicos e atividades a campo.

Aos colegas do Laboratório de Geoprocessamento pelo companheirismo e aprendizado nesta área do conhecimento.

Ao Francisco Fasana, pelo amizade e compartilhamento de conhecimento técnico na área de informática.

Aos professores do Departamento de Solos, da Universidade Federal de Viçosa, pelo conhecimento e aprendizado.

Aos bolsistas Thamiris, Silvia, Júlia, Adalberto, Daniel e a colega Marcele, pelo auxílio nas atividades laboratoriais.

À minha esposa, pela compreensão e esforço, em função minha ausência.

À Universidade Federal do Pampa, por permitir meu afastamento para a realização do doutorado.

Ao Programa Antártico Brasileiro e Secretaria do Comitê Interministerial para os Recursos do Mar, pela realização das Operações Antárticas Brasileiras XXXIII e XXXIV.

À Marinha do Brasil e Força Aérea Brasileira, pelo apoio a logística nas Operações Antárticas Brasileiras XXXIII e XXXIV.

À Emilia, pelos dados da base GPS situados na Península Keller.

SUMÁRIO

Resumo.....	vii
Abstract.....	ix
Introdução geral	1
Referências bibliográficas	3
CAPÍTULO 1	4
Abstract.....	4
1. Introduction.....	5
2. Materials and Methods	6
2.1 Study area	6
2.2 High Resolution Topography Survey.....	7
2.3 Data Processing	8
2.4 DTM validation	9
3. Results	11
3.1 Cloud adjustment	11
3.2 DTMs Validation	11
3.3 Digital Terrain Model (DTM)	13
3.4 Morphometric variables comparison.....	15
4. Discussion	20
5. Conclusions.....	21
References	22
CAPÍTULO 2	24
Resumo.....	24
Abstract.....	25
1. Introdução	26
2. Materiais e Métodos	28
2.1 Área de estudo	28
2.2 Levantamento de elevada resolução topográfica (ERT)	29
2.3 Processamento dos dados de campo	31

2.4 Obtenção das variáveis morfométricas	32
2.6 Comparação entre as variáveis morfométricas	35
3. Resultados	36
3.1 Processamento dos dados de campo	36
3.3 Comparação entre as variáveis morfométricas	36
4. Discussão.....	44
5. Conclusões.....	46
Referências	47
CAPÍTULO 3	50
Resumo.....	50
Abstract.....	51
1. Introdução	52
2. Materiais e Métodos	54
2.1 Áreas de estudo	54
2.2 Levantamento de Elevada Resolução Topográfica (ERT).....	54
2.1 Processamento	55
2.1.1 Laser Scanner Terrestre.....	55
2.1.2 Grades Amostrais.....	56
2.1.3 Interpoladores	59
2.1.4 Geração de curvas de nível.....	60
2.1.5 Tempo de processamento.....	60
3. Resultados	61
3.1 Interpoladores	61
3.2 Validação interna versus externa	69
3.3 Análise qualitativa de resultados	71
3.4 Tempo de processamento.....	74
4. Discussão.....	75
5. Conclusões.....	77
Referências	78
CAPÍTULO 4	80
Resumo.....	80
Abstract.....	81

1. Introdução	82
2. Materiais e Métodos	83
2.1 Área de estudo	83
2.2 Análise de solos	85
2.3 Densificação da amostragem	86
2.4 Interpoladores preditivos	87
2.5 Processamento dos interpoladores	91
2.6 Comparação entre preditores	93
3. Resultados	93
3.1 Densificação da amostragem	93
3.2 Processamento dos interpoladores	96
3.3 Comparação entre predições	99
4. Discussão	101
5. Conclusões	102
Referências	103
Conclusões gerais	106

Resumo

SCHÜNEMANN, Adriano Luis, D.Sc., Universidade Federal de Viçosa, novembro de 2016. **Geotecnologia para mapeamento digital na Antártica marítima**. Orientador: Elpídio Inácio Fernandes Filho. Coorientadores: Márcio Rocha Francelino.

O mapeamento da geomorfologia, permafrost, atributos químicos, comunidades vegetais, fluxo de CO₂, entre outros, em áreas livres de gelo da Antártica é de suma importância para o entendimento dos efeitos das mudanças globais. Essa região é muito sensível a pequenas alterações, sendo, portanto, um eficiente indicador de tais mudanças. Para isso torna-se necessário levantar as condições atuais da paisagem dessas áreas para servir de parâmetro para futuras análises de mudanças. Novas técnicas de levantamento e obtenção de dados, que envolvem o uso de equipamentos de elevada acurácia, permitem mapear elementos indicadores de mudanças globais em grandes áreas. O laser scanner terrestre (TLS) permite o mapeamento de áreas livres de gelo a partir de levantamentos terrestres in loco. A nuvem tridimensional gerada por esta tecnologia permite obter dados de elevação a partir de uma malha muito densa de pontos. A partir desta é possível obter uma série de covariáveis morfométricas (CM) e modelos. Os interpoladores preditivos Random Forest, Krigagem, IDW e Spline permitem prever diferentes fenômenos em locais não amostrados. Já o interpolador Topo to Raster, só permite prever elevações. Neste sentido o presente trabalho teve os seguintes objetivos: i) elaborar um modelo digital de elevação para a Península Keller com elevada resolução espacial; ii) verificar a influência da escala nas covariáveis morfométricas oriundas dos modelos digitais de elevação; iii) verificar e comparar o poder preditivo dos interpoladores IDW, Spline, Topo to Raster, Krigagem ordinária e simples, para as elevações em duas áreas da Antártica Marítima e iv) verificar o poder dos interpoladores preditivos Random Forest e Krigagem Simples, para a predição de atributos pedogequímicos selecionados. Para tanto realizou-se um levantamento com o uso do TLS nas penínsulas Keller e Coppermine, nos anos de 2015 e 2016, durante o verão austral. Os dados brutos foram processados e filtrados. No

programa ArcGIS 10.1 os mesmos foram convertidos em rasters com diferentes tamanhos de células. Estes formaram a base para os quatro trabalhos. Para o trabalho (i), tratou da produção de MDT com elevada resolução topográfica, possuindo tamanho de célula de 0,20 m, servindo de referência para verificação de mudanças futuras na paisagem local. Para o trabalho (ii) foram gerados modelos digitais de terreno (MDTs) em diferentes resoluções e a partir dos mesmos, foram obtidas as covariáveis morfométricas. As covariáveis obtidas nas diferentes resoluções foram comparadas de diferentes formas. Já para o trabalho (iii), foram extraídos os valores de elevação para MDTs de tamanho de célula de 1 m, em diferentes densidades de amostragem e três formatos de grades amostrais. Com os pontos extraídos, foram preditos os valores de elevação em locais não extraídos e os erros nos valores obtidos, comparados aos valores existentes nestes locais. Por fim, para o trabalho (iv), a partir do MDT e análise química de amostras foram preditos os valores de atributos químicos pelos interpoladores preditivos Random Forest e Krigagem e seus indicadores foram comparados

Palavras chave: interpoladores, Criosolos, Geomorfologia, modelos digitais de terreno, covariáveis morfométricas, amostragem, grades regulares.

Abstract

SCHÜNEMANN, Adriano Luis, D.Sc., Universidade Federal de Viçosa, November, 2016. **Geotechnologies for digital mapping in maritime Antarctica**. Adviser: Elpídio Inácio Fernandes Filho. Co-advisers: Márcio Rocha Francelino.

Mapping geomorphology, permafrost, soil chemical properties, vegetation communities, CO₂ dynamics in ice-free areas of Maritime Antarctica are fundamental for the understanding of global change effects. Periglacial and paraglacial areas are highly sensitive to little disturbances, being indicator of climate changes. Hence, it is necessary to know the current survey conditions of Antarctica landscape to use it as a parameter to monitor landscape modifications. New techniques for landscape survey and data gathering, involving the use of high-accuracy equipments, allow mapping of elements that contribute to this aim in large areas. In this context, the terrestrial laser scanner (TLS) allows mapping of ice-free areas. The three-dimensional cloud generated by TLS technology generate elevation data from a dense cloud of points. From this cloud, is possible to access a series of morphometric covariables (CM) and survey models. Predictive interpolators such as Random Forest, Kriging, Inverse of Distance Weighted and Spline can predict different phenomena in non-sampled locations. On the other hand, Top to Raster interpolator only predict heights from sample data. Thus, this study aimed: (i) to develop a digital elevation model for Keller Peninsula with a high spatial resolution; (ii) to check the influence of scale in CM derived from digital elevation models; (iii) to check and compare the predictive interpolators IDW, Spline, Topo to Raster and simple kriging for elevations in two areas of the Maritime Antarctica. In the chapter four, the goal was to identify the best sampling size and patten in those areas and (iv) to check the power of predictive interpolators (Random Forest and Simple Kriging) in order to predict soil chemical properties. A survey with TLS was carried out in Keller and Coppermine Peninsulas in 2015 and 2016, during the austral summer. The raw data were processed and filtered. By using ArcGIS software 10.1, the filtered data were interpolated to rasters, with different cell sizes, being the basis of data

processing. For the first chapter a digital terrain models (DTMs) in different resolutions was generated, which was used to obtain the CM. The CM obtained with different resolutions were compared. For the second chapter, by using the DTM and chemical analysis of soil samples, non-sampled values were predicted with Random Forest and Kriging techniques. The aim of the third chapter was to produce a DTM with high resolution topographic techniques. The final map had a 0.20 m cell size resolution. This will be a map reference to monitor landscape changes in this area. In the last chapter, elevation values from 1 m cell size DTMs were extracted using eight sample densities and three sampling shape grids. With the extracted points, elevation values and his errors were predicted, computed and compared.

Key words: interpolators, cryosols, geomorphology, digital terrain models (DTMs), morphometric covariables, sampling, regular grids.

Introdução geral

O mapeamento de fenômenos em áreas livres de gelo, é de suma importância para o entendimento das mudanças globais. Somente 0,36% das áreas da antártica são livres de gelo. Estas áreas estão dispersas pela periferia do continente, ilhas e pelas montanhas Transantárticas(BOCKHEIM, 2015). O mesmo autor cita como um dos desafios para o estudo dos solos antárticos, o mapeamento dos mesmos. As porções da chamada Antártica Marítima, são muito sensíveis a pequenas alterações, sendo então indicadores de mudanças globais.

As Penínsulas Keller e Coppermine estão situadas na chamada Antártica Marítima, mais especificamente no arquipélago das Schetlands do Sul. A primeira situada na Ilha Rei George entre as latitudes 62°03'00" e 62°05'30" sul e longitudes 58°28'00" e 58°22'00" oeste. Já a segunda está localizada na Ilha Robert, entre as latitudes 62°21'00" e 62°24'00" sul e longitudes de 59°45'00" e 59°37'00" oeste.

Enquanto que o clima de algumas regiões do manto antártico, como em McMurdo Dry Valleys, possuem precipitações inferiores a 10 mm ano⁻¹, existem regiões com registros de mais de 600 mm ano⁻¹, como por exemplo, na Península Antártica (BOCKHEIM, 2015), já na Península Keller, os valores ficam em aproximadamente 360 mm ano⁻¹(FRANCELINO et al., 2011). Os solos desta península são basicamente Criossolos, Fluvisolos e Leptossolos (WRB, 2015), possuindo área de cobertura de aproximadamente 17% da mesma (FRANCELINO et al., 2004). O restante das áreas possuem geleiras, bancos de neve, afloramentos rochosos e sedimentos inconsolidados (FRANCELINO et al., 2011).

Novas técnicas que envolvem o uso de equipamentos de elevada acurácia, permitem mapear grandes áreas. O laser scanner terrestre (TLS) permite o mapeamento de áreas livres de gelo a partir de levantamentos terrestres in loco. A nuvem de pontos gerada por esta tecnologia permite obter as elevações a partir de uma malha muito densa de pontos. A partir da nuvem de pontos é possível obter uma série de covariáveis morfométricas (CM), modelos digitais de superfície e de terreno em formatos matriciais. Estas covariáveis podem ser usadas para prever

atributos do solo, quando correlacionadas aos mesmos. Alguns dos interpoladores que permitem essa predição são o Random Forest (BREIMAN, 2001) e a krigagem (GOOVAERTS, 1999). O primeiro utiliza-se de covariáveis para prever fenômenos em locais não amostrados enquanto que o segundo, utiliza-se somente de valores amostrados da variável e determina o padrão estatístico a uma certa distância, para prever esta variável.

Entre os desafios para utilização ótima de preditores, está a amostragem, que pode ser realizada com diferentes padrões e densidades. Os padrões mais utilizados são o hexagonal, quadrado e randômico.

Neste sentido o presente trabalho teve os seguintes objetivos: i) verificar a influência da escala nas covariáveis morfométricas oriundas dos modelos digitais de elevação; ii) verificar o poder dos interpoladores preditivos Random Forest e Krigagem Simples, para a predição de atributos pedogeoquímicos selecionados; iii) elaborar um modelo digital de elevação para a Península Keller com elevada resolução especial, e iv) verificar e comparar o poder preditivo dos interpoladores IDW, Spline, Topo to Raster, Krigagem ordinária e simples para as elevações em duas áreas da Antártica Marítima.

Referências bibliográficas

BOCKHEIM, J. G. (ED.). **The Soils of Antarctica**. Cham: Springer International Publishing, 2015.

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

FRANCELINO, M. R. et al. GEOFORMAS DA PENÍNSULA KELLER, ANTÁRTICA MARÍTIMA: SUBSÍDIOS AO MONITORAMENTO AMBIENTAL. In: M.R., C. E. S.; FILHO, F. F. N. B. S. M. R. A. (Eds.). **Ecosistemas Costeiros e Monitoramento Ambiental da Antártica Marítima: Baía do Almirantado, Ilha Rei George - Rede 2.1**. ed. Viçosa-MG: NEPUT, 2004. p. 7–13.

FRANCELINO, M. R. et al. Geomorphology and soils distribution under paraglacial conditions in an ice-free area of Admiralty Bay, King George Island, Antarctica. **Catena**, v. 85, n. 3, p. 194–204, 2011.

GOOVAERTS, P. Geostatistics in soil science: State-of-the-art and perspectives. **Geoderma**, v. 89, n. 1–2, p. 1–45, 1999.

WRB. World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps. **World Soil Resources Reports No. 106**. Rome: [s.n.].

CAPÍTULO 1

High resolution topography for Digital Terrain Model (DTM) in Keller Peninsula, Maritime Antarctica

Adriano Luis Schünemann ¹, Pedro Henrique Araújo Almeida ², André Thomazini ², Elpídio Inácio Fernandes Filho ², Márcio Rocha Francelino ², Carlos Ernesto Gonçalves Reynaud Schaefer ² e Antonio Batista Pereira ¹

¹ Universidade Federal do Pampa;

² Universidade Federal de Viçosa.

Abstract

High resolution topography (HRT) surveys is an important tool to model landscapes, especially in zones subjected to strong environmental changes, such as Antarctica, where landform is highly influenced by cryoclasty and permafrost melting. The aim of this work was to obtain a high accurate DTM for Keller Peninsula, Maritime Antarctica. The survey study was assessed in the 2014/2015 and 2015/2016 during the austral summer, by using Terrestrial Laser Scanner (TLS). In order to cover the entire Peninsula, the TLS equipment was installed in 81 different points. Results of the DTM and the terrain variables Aspect, Slope and Hillshade obtained were compared with previous models generated by aero photographic (APG) survey. The RMSE for the DTM and APG method was 0.726 and 2.397 m, respectively. Spatial resolution of the DTM was 0.20 m. Morphometric variables obtained from the two methods presented visual differences on the thematic maps. Generalization was the main process, whereas interpolation occurred for the HRT survey, being the process of choice for the APG method. A large number of points can be obtained by using laser scanner, providing a dense cloud of points, spatially well-distributed, enabling the generalization process to obtain surface models with high performance.

Keywords: terrestrial laser scanner; remote sensing; geomorphology

1. Introduction

The ability to predict the transfer of water, sediments and nutrients, map catastrophic natural events, identify the radiation balance, biophysical responses of the environment that influence the landscape and its functions, all have great importance to science and society. Several studies are focused on Earth's surface and its natural resources and changes (PASSALACQUA et al., 2015), especially at the present, when mankind face severe climate changes, an widespread modifications on landscape (MICHEL et al., 2007).

Passalacqua et al. (2015), defined surveys with high resolution topography (HRT) as any set of data in its raw form, consisted of locations measurements (x, y) and elevations (z), which together make up a cloud of points having spatial resolution greater than or equal to one point per square meter. One of the advantages of using surveys with HRT is that it enables quantitative study of phenomena (TAROLLI, 2014). This is important in zones subjected to strong environmental changes, such as Antarctica, where relief dynamic is highly influenced by physical weathering and permafrost melting.

The Digital Elevation Model is a representation of a continuous surface elevation using points with well-known x, y and z coordinates. When the elevation represents the earth surface, it is so-called Digital Terrain Model (DTM) (LI; ZHU; GOLD, 2004). The main interpolators to model the landscape or a phenomena, may be listed in three basic categories as indicated by (WASKLEWICZ et al., 2013): i) local neighborhood (e.g., nearest neighbor, inverse distance weighted, triangulation); ii) geostatistics methods (e.g. kriging), and iii) spline methods.

DTMs can be generated in many ways as follow: conventional topographic surveys, kinematic GNSS (Global Navigation Satellite System) surveys, analogue and digital photogrammetric approaches, radar techniques, laser surveys, shipboard echo sounding, airborne optical sensing, satellite radar altimetry, soil augering and geological boring, three-dimensional (3D) seismic survey, airborne ice-penetrating radar techniques, radio-echo sounding surveys and digitizing of contours (FLORINSKY, 2011). The obtained data can be represented as regular grids, square

grids, profiles, contouring, regular triangles, hexagons, breaks and feature lines and random forms (LI; ZHU; GOLD, 2004).

Another issue to be studied is focused on the use of processing data. Frequently, the processed data is converted to raster format and later submitted to analyze. However, conversion to the matrix format causes a reduction of detail in the data set, which is avoided by using point cloud, evidenced in the study of Lague, Brodeur & Leroux (2013). Also, these authors used temporal techniques to evaluate erosion using the point cloud. Despite this previous literature, there is a lack of HRT studies in Polar environments, especially in Maritime Antarctica.

In the past, the topographic representation of the earth's surface was limited to spatial resolutions of up to 10 m, approximately (HUTCHINSON; GALLANT, 1999). An example that demonstrates the improvement over the spatial resolution, is the case of SRTM program (Shuttle Radar Topography Mission). Initially, SRTM program had spatial resolution of 90 m and was later made available a model with spatial resolution of 30 m (FARR et al., 2007). Although occurred an advance in data obtained by this program, these resolutions would still be insufficient to identify small changes in the landscape that have great importance for numerous environmental phenomena studies. For this reason, the aim of this work was to obtain a high accurate DTM (0.04 m² square cell) for Keller Peninsula, Maritime Antarctica. In a near future, the high accurate DTM will be used to map temporal morphometric processes and identify landform changes in the local scale.

2. Materials and Methods

2.1 Study area

The study was carried out at Keller Peninsula (Admiralty Bay, King George Island), Maritime Antarctica, part of the South Shetlands (Fig. 1). The area comprises about 500 ha, being approximately 4 km of length, considering the north-south direction, and 2 km of width, considering east-west direction (FRANCELLINO et al., 2011). The altitude reach up 380 m above sea level (FRANCELLINO, 2004), with an

annual average air temperature and rainfall of $-1,8^{\circ}\text{C}$ and 360 mm, respectively (FRANCELINO et al., 2011). The Brazilian Station is placed in this peninsula, which began operating in 1984 (SANTOS et al., 2005). Periglacial Landform are influenced by physical weathering processes an cryoclasty associated with freezing and thawing cycles, besides strong winds and snow permafrost melting during the summer (FRANCELINO, 2004).

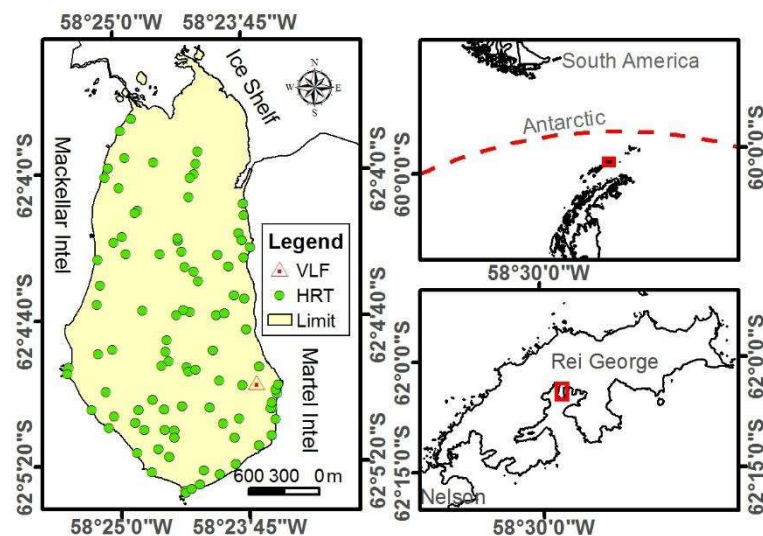


Fig. 1 – Illustrative location map of the study area, indicating the stations points (HRT) and the GNSS base (VLF) for post-processing coordinates.

2.2 High Resolution Topography Survey

The survey study was assessed in the 2014/2015 and 2015/2016 during the austral summer, by using Terrestrial Laser Scanner (TLS) (RIEGL VZ-1000 model), with a nominal accuracy and precision of 8 and 5 mm, respectively. A GNSS equipment (Leica Model GS80), was used in order to georeference the stations. A base located near the Comandante Ferraz Antarctic Station (Fig. 1) ($62^{\circ} 04' \text{S}$ 58.12 south and $58^{\circ} 23' 39.31$ west, and 57.53 m of ellipsoidal height), was utilized for differential correction of data. The GNSS scans were performed during 10 minutes at each point, where the TLS was also installed, and fixed solutions for all points were obtained.

Due the methodology adopted for scanning survey, sequential surveys were necessary to obtain a minimum overlap of 50%. With this, the concatenation of all

measurements generated the point cloud in post-processing. To cover the entire Peninsula, the TLS equipment was installed in 81 different points as shown in Fig. 1. The scanning survey strategy was designed to cover the largest possible area with an overlap between scanned zones. In areas with high snow cover, the high sensitivity mode of the software was used, to obtain a reasonable density of dots per square meter.

Results of the DTM obtained were compared with the DTM generated by Júnior et al. (2012). The terrain variables Aspect, Slope, Contour and Hillshade were generated with the two DTMs and their results were also compared.

2.3 Data Processing

The software Leica Geoffice was used to process GNSS data. All data were referenced with the WGS84 *Datum*. Also, altitudes and elevation of the DTM were referenced accordingly to the ellipsoid and local geoid, respectively. After obtained the coordinates of all scanned stations, the processing was realized comprising 81 scans, using the software Riscan Pro 6.1-64 bits version. The first stage of processing consisted of sequential scans turning for concatenation. This processing was developed visually. Positions of each scan, do not change on this processing. Later, an adjustment between consecutive and overlapped SCANS was realized with the tool "Multi Station Adjustment" in the software Riscan Pro. When the standard deviation of error (σ_e) obtained from the adjustment process was lower or equal to 0.040 m, the adjustment process was finished. However, in some cases, this accurate adjustment could not be obtained, but did not committed the quality of point cloud. By adjusting the SCANS, a new cloud was created. The raw cloud has points of objects that are not part of the natural terrain, such as birds and buildings. Also, airborne particles are another issue to be removed from the cloud. Usually, this may be caused by the high-sensitivity mode, as shown in Fig. 2. All these points were eliminated from the cloud, in order to obtain the filtered cloud of the terrain.

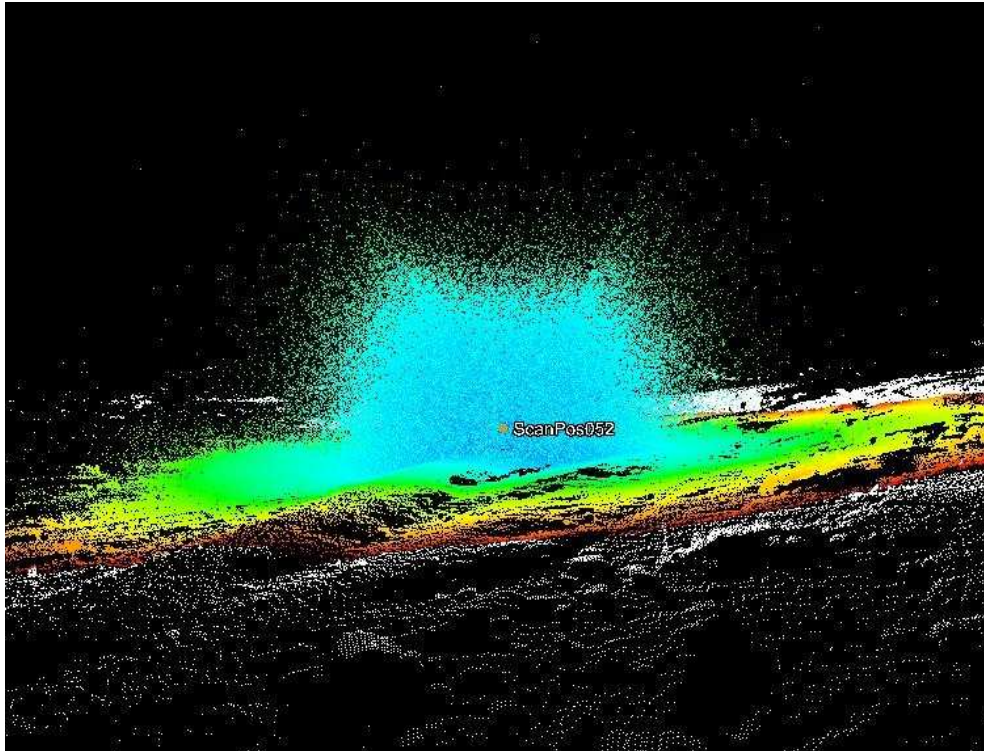


Fig. 2 – Cloud with some points (airborne particles registered with high sensibility mode), to be extracted and filtered from the current terrain in the study area.

The entire area was divided into 8 sub-areas, in order to filter the cloud through the visual filter as well as with automatic filter so-called "vegetation filter" available in Riscan Pro software. All airborne particles, buildings, shelters, non-terrain features were removed from the raw cloud. With this procedure, a new cloud was produced from the merge of the 8 filtered sub-areas, hereafter called "filtered cloud" (FC). The FC comprises 54 points per square meter, which was exported in the ".las" format. After, a "Lasdataset" was created with the ArcGIS 10.1 software. The DTM was performed by the generalization of the FC by using "LAS dataset to raster", utilizing average cell recording and "Natural Neighbor" void fill method. The size of the generated cells were 0.20 m. The elevation of reference was altered to the local geoid, using the online Geoid Height Calculator (UNAVCO, 2016) tool. This process uses the potential coefficient model EDM96 and a set of spherical harmonic coefficients. Finally, the mean Geoid Undulation (\bar{N}) was obtained for the 81 points, and DTM values were subtracted from the mean Geoid Undulation.

2.4 DTM validation

An independent GNSS survey was performed to assess the DTM quality. To this end, 57 georeferenced points distributed across the peninsula were obtained. Each point was compared with the DTM generated by using both TLS (hereafter, HRT-DTM), and APG survey (hereafter, APG-DTM) obtained by Júnior et al. (2012). Errors were considered as outliers when values exceeded three times the standard deviation (99.7%). These errors were desconsidered on the validation process. The Root Mean Square Error (RMSE) was calculated comparing DTM elevations (HRT-DTM and APG-DTM) with validation points values (57 points), by using the equation below (WILLMOTT; MATSUURA, 2005):

$$RMSE = [n^{-1} \sum_{i=1}^n |e_i|^2]^{1/2},$$

where

$RMSE$ = root mean square error (m);

e_i = each difference between DTM elevations and validation points, at an i position (m).

According to the outlier criterion utilized, 53 validation points were extracted from the HRT-DEM. Also, 56 validation points were used to extract the elevation values from the APG-DTM. The extractions were performed using R software.

2.5 Morphometric variables

In this study, slope, aspect and hillshade were obtained from the HRT-DTM and APG-DTM. Morphometric variables were generated in the 10.1 ArcGIS software version. With this, all morphometric variables were subtracted between HRT-DTM and APG-DTM, and then compared. Hence, these subtractions were considered as errors, which were used to obtain RMSE value.

3. Results

3.1 Cloud adjustment

The maximum and minimum standard deviation error (σ_e) obtained from the adjustments was 0.0848 and 0.0081 m, respectively. The mean standard deviation error ($\bar{\sigma}_e$) obtained was 0.0245 m. This value is 4.9 times higher than the nominal accuracy of the equipment. The mean error for latitude, longitude and ellipsoidal height (considering all coordinates) obtained from the GNSS survey in the HRT for TLS stations, were 0.90, 1.63 and 1.9 mm, respectively.

3.2 DTMs Validation

Mean error between the HRT-DTM and the validation point values was 0.718 m, with a standard deviation error of 1.599 m. The threshold to be considered as outliers was 4.796 m. RMSE obtained was 1.740 m utilizing 57 points. By using outlier criterion, the following points: 48 (error = 5.94 m), 55 (error = 6.06 m), 59 (error = 6.04 m) and 65 (error = 6.01 m) were excluded from the validation cloud in HRT-DTM (Fig. 3). With this criterion, RMSE decreased, reaching to 0.726 m.

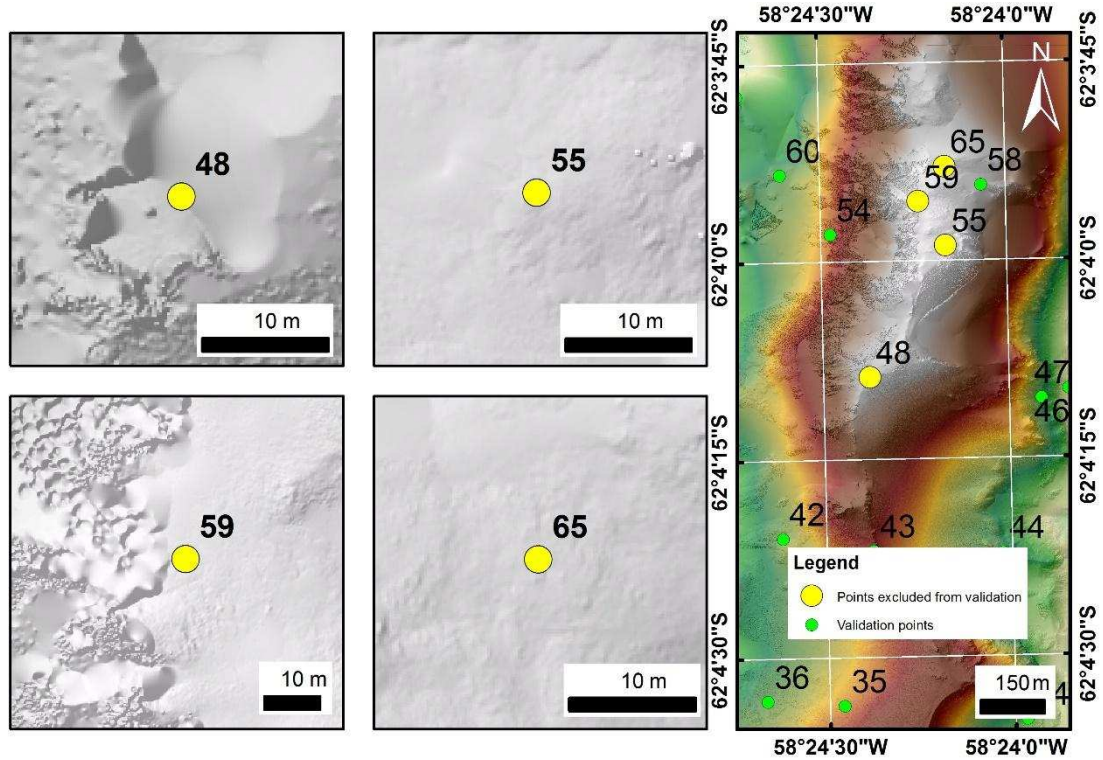


Fig. 3 – Points considered as outliers for HRT-DTM.

The APG-DTM validation showed mean error of 2.044 m, which is higher than HRT-DTM. Slightly elevated standard deviation error (4.458 m) was obtained for the APG-DTM validation when compared with HRT-DTM. In addition, RMSE value was 4.868 m, using 57 points. Using the same criterion to eliminate the outliers, only the point 47 (error = 32.08 m) (Fig. 4) was excluded from the validation data. With this, a value of 2.397 m was obtained, which is more than 3 times higher than HRT-DTM procedure considered.

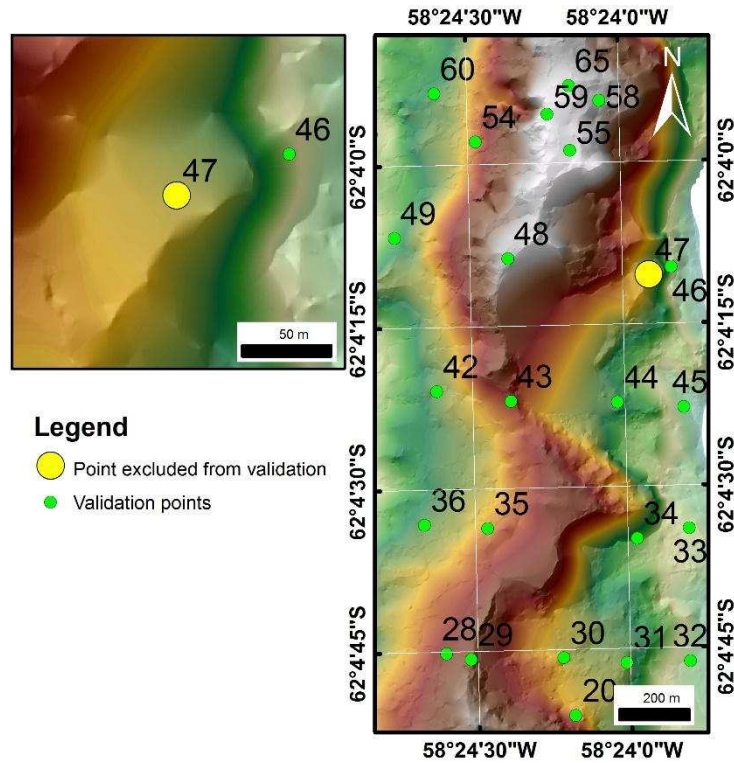


Fig. 4 – Point considered as outliers for APG-DTM.

3.3 Digital Terrain Model (DTM)

The DTM generated with the HRT survey is shown in Fig. 5. The maximum elevation value recorded was 331.503 m at the Mount Birkenmajer. Flat areas are located on locations closer to the sea level, especially on the west side.

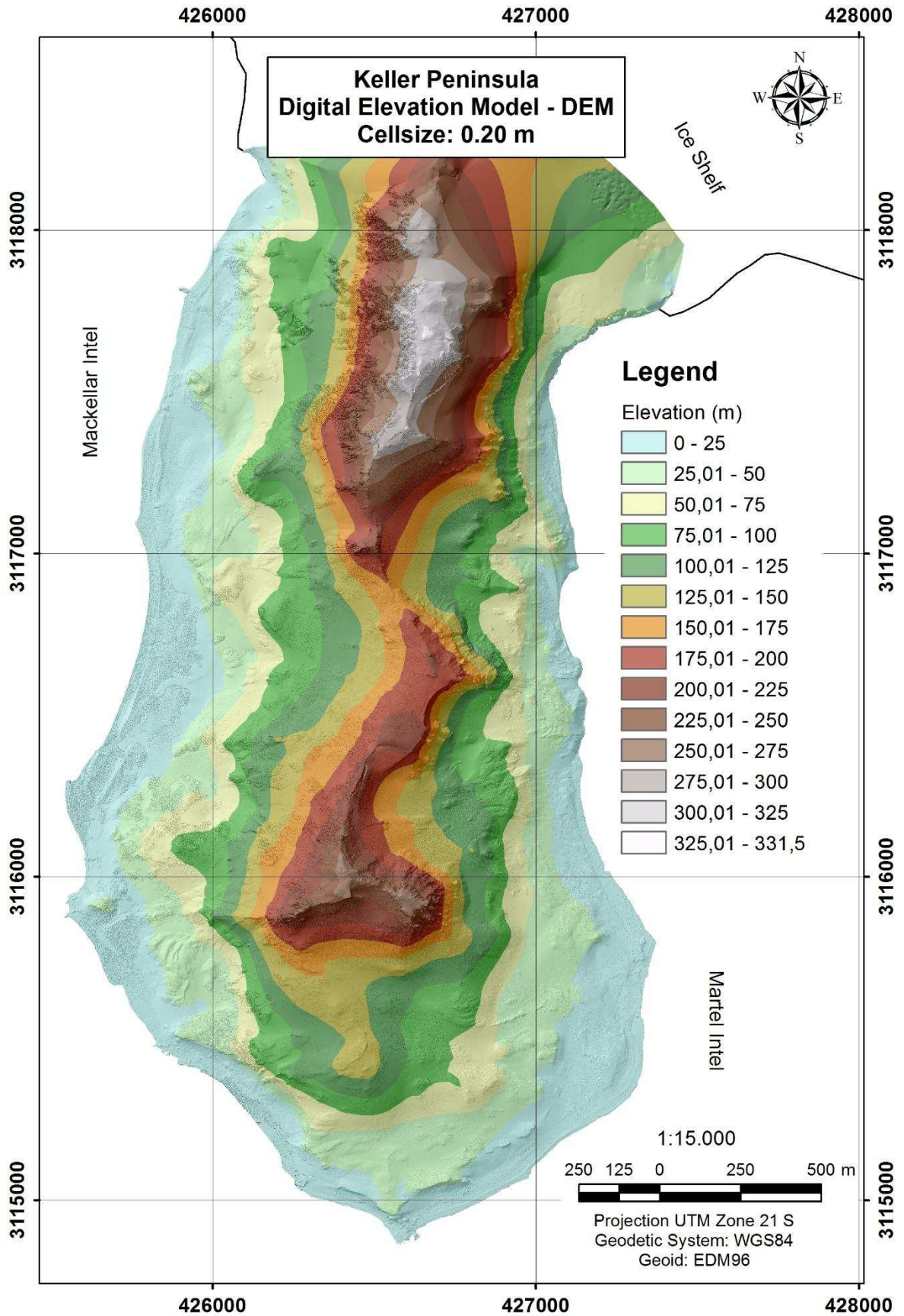


Fig. 5 – Digital Terrain Model of Keller Peninsula with 0.20 m cell size (HRT-DTM).

3.4 Morphometric variables comparison

Clearly, morphometric variables obtained from the two methods present differences on the thematic maps in Fig. 6, 7 and 8. According to the RMSE, a contrast exist between the two surveys. An amount of 82.184 (with the same units of the raster) was recorded for the Aspect. Visually, Aspect in the HRT-DTM showed a better representation when compared with the APG-DTM (Fig. 6).

For the Slope, a RMSE of 67.725 (percent rise) was obtained, comparing the two methods. Thematic maps of slope can be viewed in Fig. 7, where a marked difference on the accuracy in the HRT-DTM is observed. In addition, a difference of 48.796 between HRT-DTM and APG-DTM was recorded, for the generated hillshades. The visual comparison of the two positions is presented in Fig. 8.

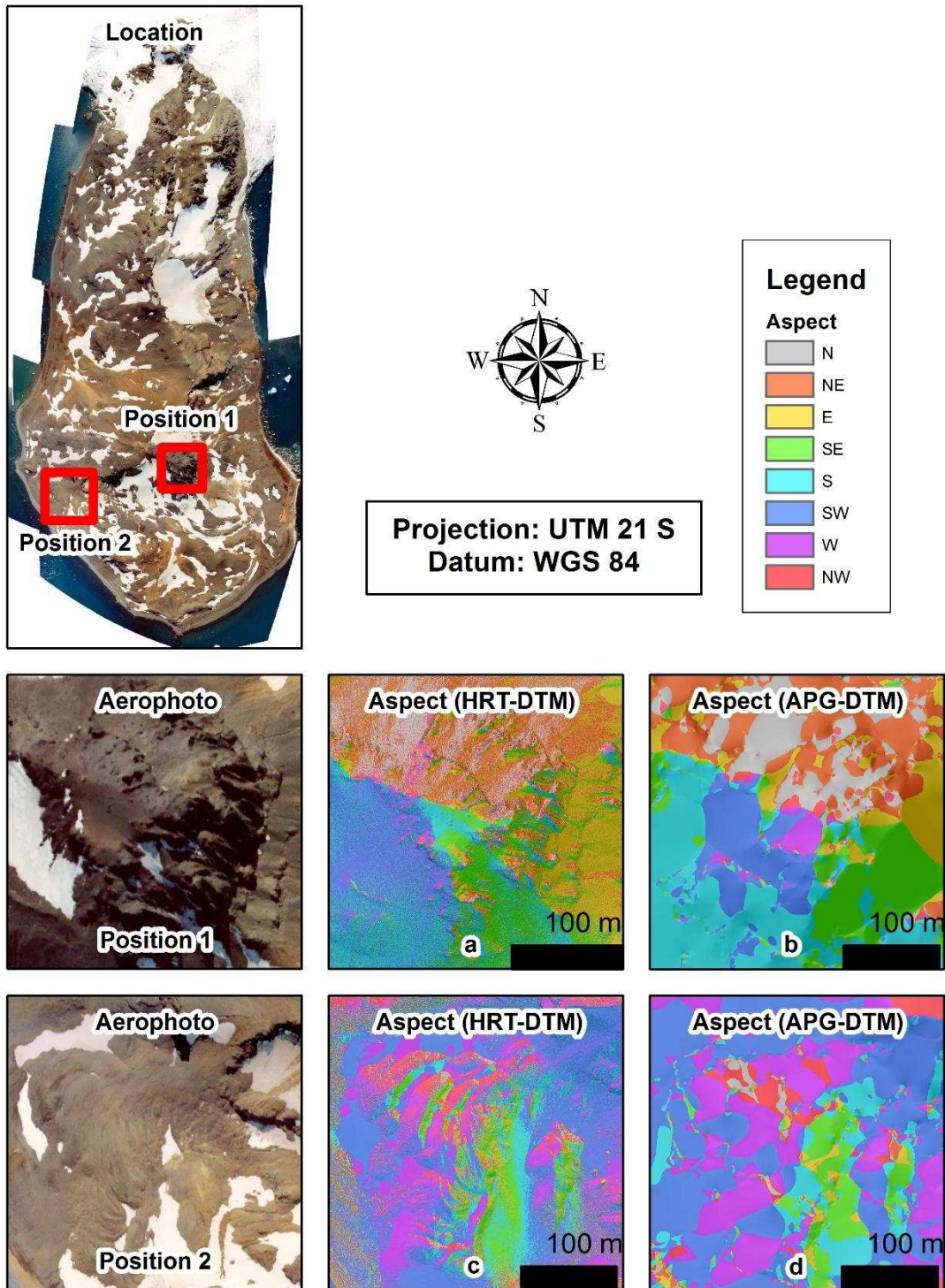


Fig. 6. Illustrative representation of the Aspect for the HRT-DTM (a and c) and the APG-DTM (b and d) in two different positions (Aerophotos adapted from(FRANCELINO, 2004)).

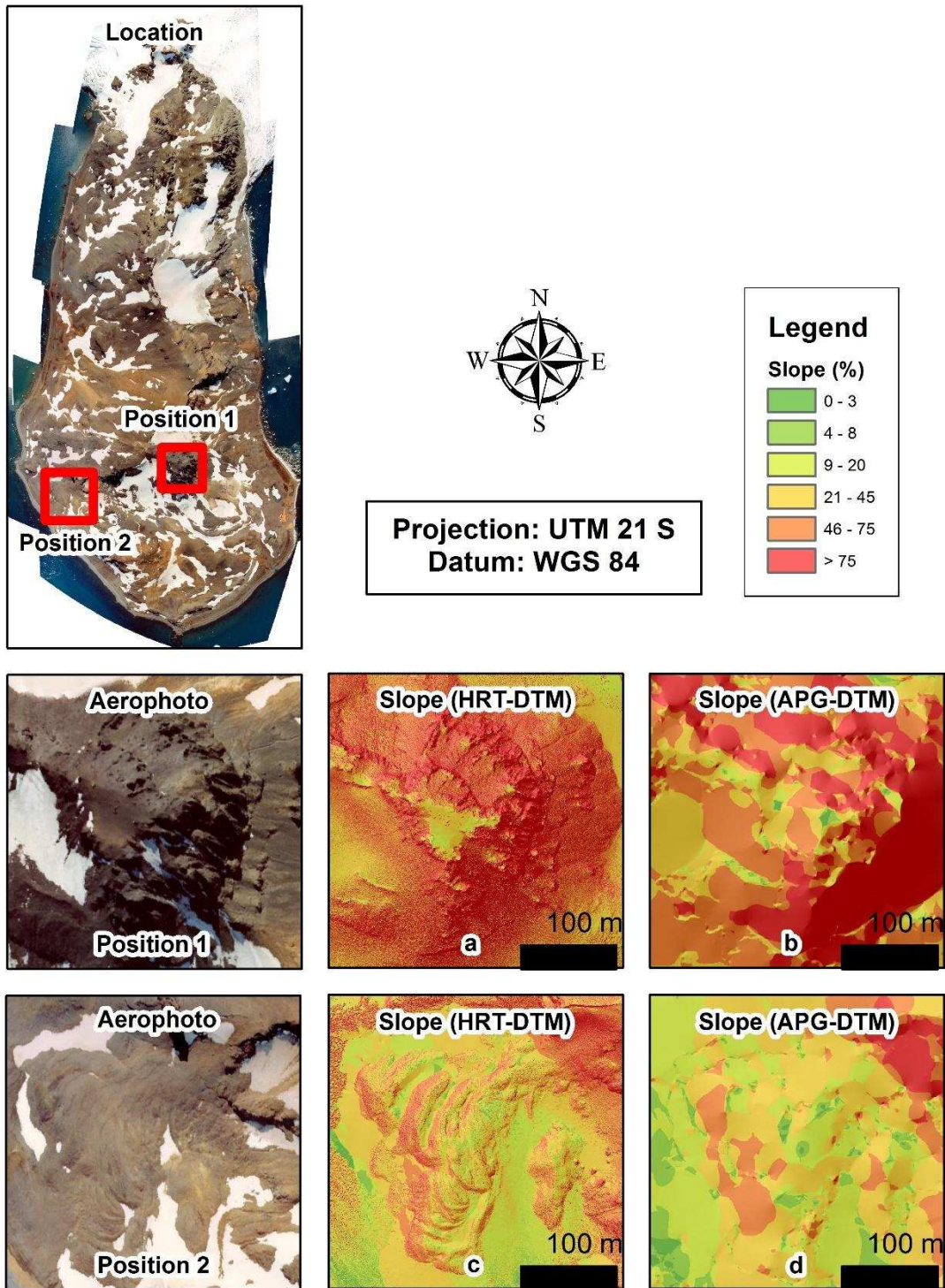


Fig. 7. Illustrative representation of the Slope for the HRT-DTM (a and c) and the APG-DTM (b and d) in two different positions (Aerophotos adapted from(FRANCELINO, 2004)).

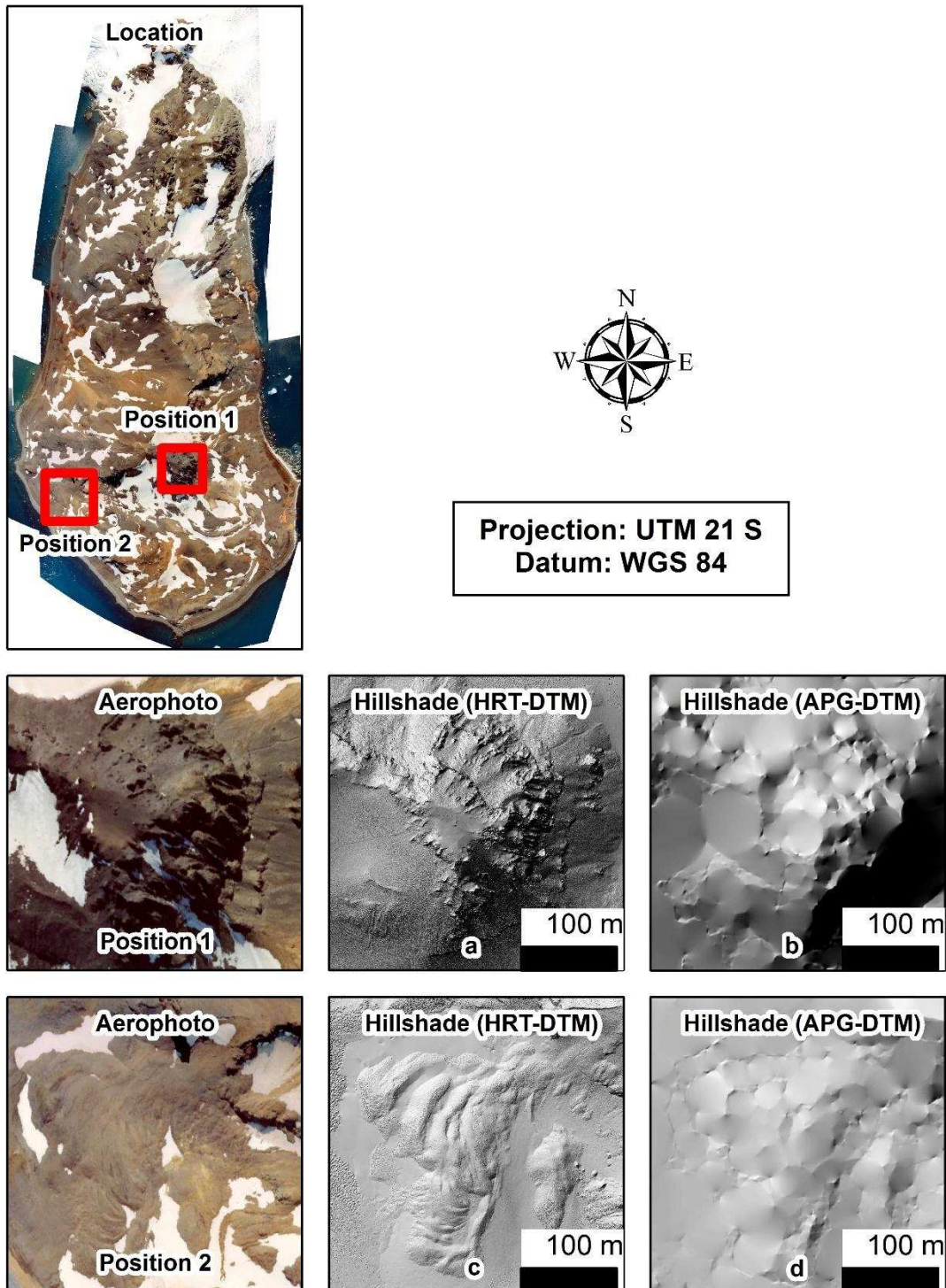


Fig. 8. Illustrative representation of the Hillshade for the HRT-DTM (a and c) and the APG-DTM (b and d) in two different positions (Aerophotos adapted from(FRANCELINO, 2004)).

Results indicated a RMSE of 10.510 m when comparing HRT-DTM and APG-DTM methods, indicating difference on terrestrial modeling performance.

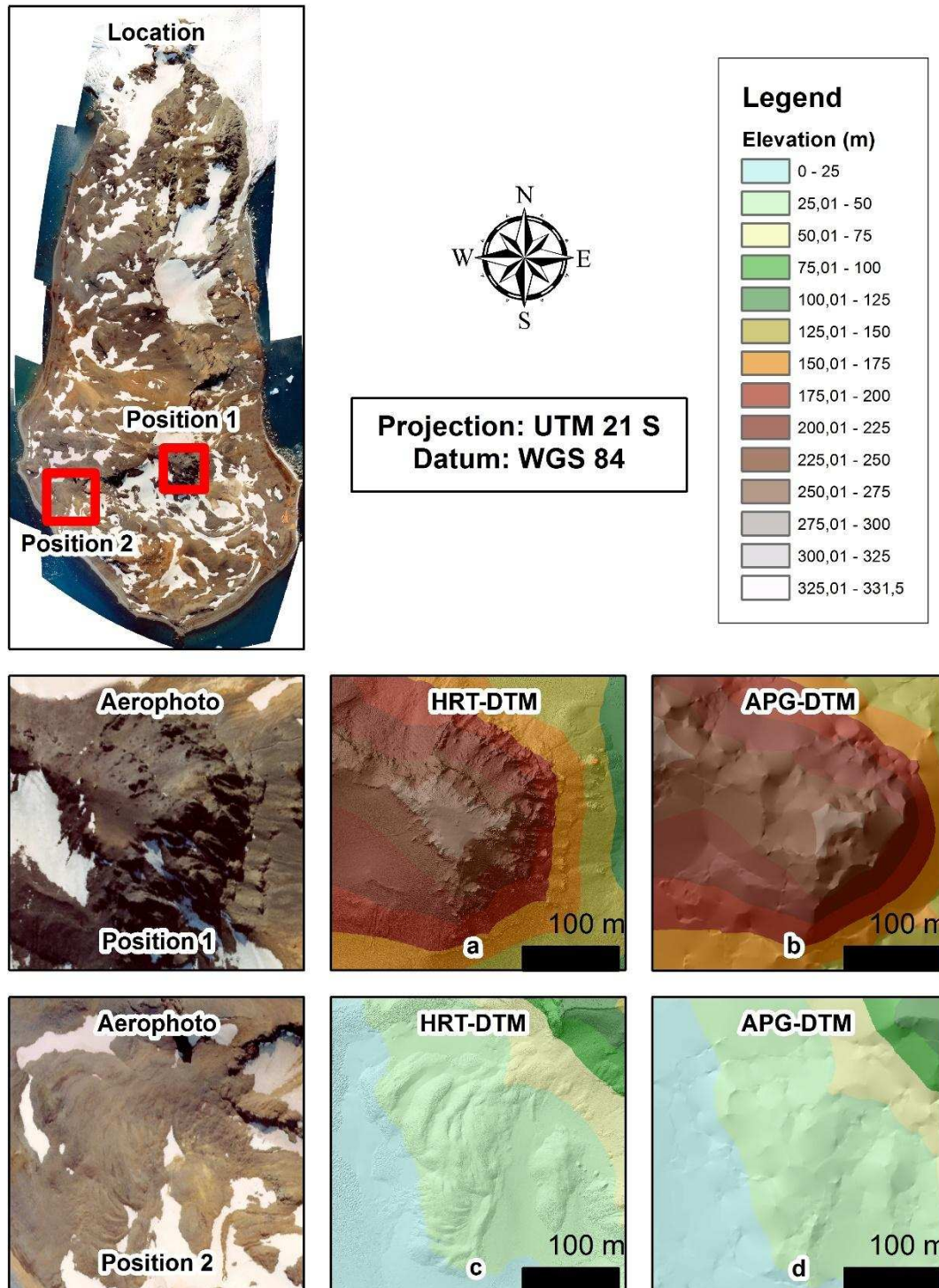


Fig. 9. Illustrative representation of the DTM generated by HRT-DTM (a and c) and the APG-DTM (b and d) methods, in two different positions (Aerophotos adapted from (FRANCELINO, 2004)).

4. Discussion

This is the first detailed survey using a high resolution topography equipment (laser scanner) in ice-free areas in Maritime Antarctica. Previously, studies of Francelino (2004) and Júnior et al. (2012) developed digital elevation models and morphometric variables and topographic maps based on APG and GNSS surveys. In these studies, the authors utilized interpolation as a main process to obtain digital elevation models, in order to generate terrain attributes. Júnior et al. (2012) utilized 70,787 points derived from the stereomodel, combined with the data from the EACF topographic plan to obtain the DTM. Despite being the first works in this region, the DTM do not have enough resolution to characterized small changes on relief. This important, to evaluate landscape modification in para and periglacial zones, which are in constant changes, especially in Maritime Antarctica, subjected to more humid and warm climate conditions. In a small temporal period of monitoring, these changes can be observed by using HRT method, such as laser scanner.

In this work, approximately 270 million of points (about 54 points per square meter) were used to obtain DTM in Keller Peninsula, representing the more detailed terrestrial modeling documented in Maritime Antarctica yet. Overall, the main process to elaborate the DTM was generalization. The interpolation method predicts values for cells in a raster from a limited number of sample data points, while generalization calculates values for cells from several set of data within the same cell. Cell value in this study was calculated based on 2.16 points, approximately, emphasizing the precision of the method chosen to model all of surface. Interpolation method was the main process used in the APG-DTM modeling, because only a few points per square meters (less than 1 point per square meter) were available in the studied area, explaining the better performance of the HRT-DTM.

Although generalization was the main process, in some places interpolation also occurred. This is due for the fact that laser scanner did not accessed surfaces

with snow, ice and lack of overlap during survey, resulting in empty zones in the cloud. Possibly, the regions of the points 48, 55, 59 and 65 (Fig. 3) were modeled by interpolation instead of generalization. Results also indicated difference of altitude value in the highest peak in Keller Peninsula. In this work, highest peak in Keller Peninsula presented 331.503 m when compared with 380 m above the sea level, reported by Francelino et al. (2011). This result is mainly related to the differences in the generalization and interpolation methods, indicating an overestimation when using interpolation as a main method to generate DTM.

During the HRT processing, the cloud adjustments did not suffer influence of the GNSS survey, since GNSS positions were considered fixed on the adjustment, varying only ΔRoll , ΔPit and Δyaw (adjustment parameters). Eliminating outliers by using a statistical criteria, RMSE tended to decrease in both surveys (HRT-DTM and APG-DTM), improving the quality of the modeling. RMSE is reported to be an efficient indicator of modeling performance. RMSE decreased up to 49% when eliminating outliers in the APG-DTM. RMSE values indicated differences among morphometric variables. However, in this case, the significance of the differences was not quantified. Qualitatively, visual performance was clearly identified in Fig. 6, 7 and 8 for morphometric variables and in Fig. 9 for DTM. With this, the visual quality of the thematic maps is more pronounced in the HRT-DTM maps, due the elevated number of points utilized to model the surface.

5. Conclusions

High resolution topography represents an efficient toll to map landscapes in ice-free areas in Maritime Antarctica. Digital terrain models obtained by generalization as a main process, present lower root mean square error, with elevated number of points per cell in comparison with those generated by interpolation. A large number of points can be obtained by using laser scanner, providing a dense cloud of points, well-spatially distributed, enabling the generalization process to obtain surface models with high performance. Hence, morphometric variables can be detailed mapped and studied, accurately representing the local relief. Also, this technique allows studies of landscape evolution at a micro

scales in time, in which geomorphological process can be assessed. The modeling of similar periglacial landscapes in Maritime Antarctica must be achieved, in order to better understand landscape changes in polar zones under current climate warming.

Acknowledgments: We acknowledge Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (556794/2009-5) and Ministério da Ciência, Tecnologia e Inovação (MCTI) for granting financial support. This work is a contribution of INCT-Criosfera TERRANTAR group. Thanks for the Project Geoespaço (<http://www.inpe.br/crs/pan/pesquisas/geoespaco.php>), process CNPq/PROANTAR: 556872/2009-6, for the GNSS base data. We thank Universidade Federal do Pampa, for support.

References

- PASSALACQUA, P et al. Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review. **Earth-Science Reviews**, 2015, Volume 148, p. 174-193, DOI: 10.1016/j.earscirev.2015.05.012.
- MICHEL, R. F. M. et al. Ornithogenic Gelisols (Cryosols) from Maritime Antarctica Soil **Science Society of America Journal**, 2007, Volume 70, No. 4, p.1370-1376, DOI: 10.2136/sssaj2005.0178.
- TAROLLI, P. High-resolution topography for understanding Earth surface processes: Opportunities and challenges. **Geomorphology**, 2014, Volume 216, p. 295-312, DOI: 10.1016/j.geomorph.2014.03.008.
- LI , Z.; ZHU , Q.; GOLD , C. **Digital Terrain Modeling: Principles and Methodology**. Publisher: CRC PRESS, Whashington, D.C., USA, 2004.
- WASKLEWICZ, T. et al. 3.6 Digital Terrain Modeling. In: **Treatise on Geomorphology**. 2013, p. 130–161.
- FLORINSKY, I. **Digital terrain analysis in soil science and geology**. 2011, Volume 31, p. 31–41, DOI: 10.1016/B978-0-123-85036-2.00003-1.
- LAGUE, D.; BRODU, N.; LEROUX, J. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). **ISPRS Journal of Photogrammetry and Remote Sensing**, Volume 82, p. 10–26, 2013, DOI: 10.1016/j.isprsjprs.2013.04.009.

HUTCHINSON, M. F.; GALLANT, J. C. **Representation of terrain Geographical information systems**. 1999. Disponível em: http://www.geos.ed.ac.uk/~gisteac/gis_book_abridged/files/ch09.pdf\nHutchinson1999GIS.pdf

FARR, T. G. et al. The shuttle radar topography mission. **Reviews of Geophysics**, Volume 45, n. 2, 2007.

FRANCELINO, M. R. et al. Geomorphology and soils distribution under paraglacial conditions in an ice-free area of Admiralty Bay, King George Island, Antarctica. **Catena**, Volume 85, n. 3, p. 194–204, 2011.

FRANCELINO, M. R. **GEOPROCESSAMENTO APLICADO AO MONITORAMENTO AMBIENTAL DA ANTÁRTICA MARÍTIMA: SOLOS, GEOMORFOLOGIA E COBERTURA VEGETAL DA PENÍNSULA KELLER**. Universidade Federal de Viçosa, 2004.

SANTOS, I. R. et al. Heavy metal contamination in coastal sediments and soils near the Brazilian Antarctic Station, King George Island. **Marine pollution bulletin**. Volume 50, n. 2, p. 185–94, 2005.

JÚNIOR, C. W. M. et al. A new topographic map for Keller Peninsula, King George Island, Antarctica. **Brazilian Antarctic Research**. Volume 5, p. 105–113, 2012.

UNAVCO. Available online: <https://www.unavco.org/software/geodetic-utilities/geoid-height-calculator/geoid-height-calculator.html> (accessed on 09/08/2016).

WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. **Climate Research**. Volume 30, n. 1, p. 79–82, 2005.

CAPÍTULO 2

Variáveis morfométricas como covariáveis preditivas para mapeamento digital em ambiente polar complexo e influência da resolução espacial

Resumo

Plataformas remotas têm possibilitado a obtenção de modelos digitais de elevação em locais de difícil acesso, sendo as derivadas destes usadas como covariáveis preditivas. O objetivo do presente trabalho foi verificar a influência da mudança do tamanho de célula, na produção de covariáveis morfométricas como preditores para mapeamento digital. Para tanto foi realizada a varredura da Península Keller, situada na Ilha Rei George, Antártica com o auxílio do Laser Scanner Terrestre (TLS). O ajuste das varreduras obteve erros variando entre 0,008 e 0,085 m com erro médio de 0,025 m. A partir da nuvem de pontos gerada, foram produzidos Modelos Digitais de Elevação (MDEs) nos tamanhos de célula de 0,2; 0,5; 1,0; 5,0; 10,0; 20,0 e 30,0 m. A partir desses MDEs foram derivados 39 temas morfométricos no programa SAGA GIS. Para comparação entre as variáveis morfométricas, foi gerada uma nuvem de pontos a partir do MDE de maior resolução (30 m) ao qual foram extraídos todos os valores das 280 matrizes geradas. Para verificar as concordâncias foi realizada a correlação de Pearson e Randomized Dependence Coefficient (RDC) (LOPEZ-PAZ, HENNIG, SCHÖLKOPF, 2013). A fim de separar as correlações em 3 grupos, sensíveis a resolução (SR), pouco sensíveis a resolução (PSR) e altamente sensíveis a resolução (ASR) foi realizada a análise de agrupamento das correlações. Dependendo do elemento morfométrico utilizado como covariável, a resolução espacial poderá melhorar o modelo de predição. A mesma variável, enquadrada neste trabalho no grupo das ASR, poderá ser utilizada em diferentes tamanhos de célula, carregando informações diferenciadas como preditora em MD. Para obtenção de variáveis PSR, é possível utilizar modelo digitais de elevação com resoluções espaciais de 30 m.

Palavras chave: Random Forest, interpoladores, sensoriamento remoto

Morphometric variables as predictive covariates for digital mapping in a complex polar environment and spatial resolution influence

Abstract

Remote platforms have made possible to obtain digital elevation models in places of arduous access, where the derivatives of these models are used as predictive covariates. The objective of the present work was to verify the influence of cell size change in the creation of morphometric covariates as digital mapping predictors. To accomplish this objective, the Keller Peninsula surface, located on King George Island, Antarctica was scanned with a Terrestrial Laser Scanner (TLS). The scanning adjustment obtained errors ranging from 0.008 to 0.085 meters with an average error of 0.025 meters. From the points cloud generated, Digital Elevation Models (DEMs) were produced with cell sizes of 0.2; 0.5; 1.0; 5.0; 10.0; 20.0 and 30.0 meters. From these DEMs, 39 morphometric themes were derived in the SAGA GIS software. To comparison among the morphometric variables, a points cloud was generated from the DEM with the highest resolution (30 m) of which all the 280 arrays values generated were extracted. Pearson correlation and Randomized Dependence Coefficient (RDC) were performed to verify the concordances (LOPEZ-PAZ, HENNIG, SCHÖLKOPF, 2013). In order to separate the correlations into 3 groups: resolution-sensitive (RS), low-resolution-sensitive (LRS) and high-resolution-sensitive (HRS), clustering analysis of the correlations were performed. Depending on the morphometric element used as covariate, the spatial resolution may improve the prediction model. The same variable, framed in the HRS group, can be used in different cell sizes, carrying differentiated information as a predictor in digital mapping. To obtain LRS variables, it is possible to use digital elevation models with spatial resolutions of 30 m.

Keywords: Random Forest, interpolators, remote sensing

1. Introdução

Predizer respostas biofísicas do ambiente que influenciam na paisagem e suas funções são de grande valia para estudos da superfície terrestre e seus recursos naturais, sendo o gradiente topográfico fator chave nesse processo, que inclui ainda transporte de massa e energia (PASSALACQUA et al., 2015).

No passado, a representação topográfica da superfície terrestre estava restrita ao esboço de linhas analógicas denominadas curvas de níveis, as quais possuíam aplicações limitadas no ambiente computacional mesmo quando digitalizada devido à sua estrutura vetorial, que a impedia de representar adequadamente um dado contínuo como o relevo. Paralelamente, o formato matricial (*raster*) é mais adequado para modelo contínuos, além de permitir uma série de processamentos utilizando álgebra de mapas.

Segundo Luiz, Santos e Brenner(2007), a representação *raster* de uma superfície é composta pelas coordenadas planimétricas (x, y) e um valor de intensidade da célula (z). Esse conjunto é comumente designado como MDE, que permitiu uma grande avanço na modelagem computacional de dados ambientais.

Atualmente é possível acessar gratuitamente vários desses modelos, como o gerado pelo programa SRTM, que inicialmente possuía resolução espacial de 90 m e posteriormente foi disponibilizado outro com resolução espacial de 30 m (FARR et al., 2007); o gerado através do uso de imagens do sistema Aster, também com 30 m (ABRAMS& HOOK, 2002) o processado através das imagens de radar do sensor Palsar (ALOS), com 12,5 m (BORNER et al., 2007), entre outros. Embora ocorresse um avanço nos dados obtidos por esse programa, estas resoluções ainda seriam insuficientes para a identificação de pequenas mudanças na paisagem, que são de grande importância para o estudo de diversos fenômenos ambientais.

Com o avanço das tecnologias de levantamentos geodésicos, como por exemplo os Sistemas de Posicionamento Global (GPS), que permitem a obtenção de coordenadas com erros milimétricos e identificação dos desvios (incertezas) associados às mesmas, em conjunto com técnicas de mapeamento a laser (e.g. LIDAR, TLS) tornou-se possível a obtenção de levantamentos com elevada

resolução topográfica (ERT), o que revolucionou a forma de modelar as variáveis da paisagem (PASSALACQUA et al., 2015; TAROLLI, 2014).

Levantamentos com ERT compreendem qualquer conjunto de dados que na sua forma bruta consistem em localizações (x, y) e elevações (z) medidas e que em conjunto compõe uma nuvem de pontos que possuem resolução espacial maior ou igual a um ponto por metro quadrado (PASSALACQUA et al., 2015).

Dados para levantamentos de ERT podem ser obtidos de forma remota (e.g. botes, aeronaves, veículos) ou de forma estática (e.g. tripés) com diferentes técnicas de registro dos mesmos, como por exemplo Synthetic Aperture Radar (SAR), Structure from Motion (SfM), Sound Navigation and Ranging (SONAR) (PASSALACQUA et al., 2015). Uma das vantagens do uso de levantamentos com ERT é que possibilita o estudo de fenômenos terrestres quantitativamente (TAROLLI, 2014).

Lague, Brodu e Leroux(2013), utilizaram técnicas temporais para estudo de erosão com o uso da nuvem de pontos, na Nova Zelândia. Entretanto, com grande frequência os dados processados são convertidos no formato *raster* e posteriormente realizadas as análises. Porém, essa conversão causa uma perda no detalhamento dos dados, o que seria evitado com o uso da nuvem de pontos, notoriamente evidenciado no estudo de áreas com elevada declividade. Os mesmos autores utilizaram técnicas temporais para estudo de erosão com o uso da nuvem de pontos, na Nova Zelândia, por exemplo.

Modelos digitais de elevação tem sido utilizados para descrever e estudar solos e suas interconexões com as paisagens, além de outros modelos preditivos para mapeamento digital de solos (MDS) (MINASNY; MCBRATNEY; SALVADOR-BLANES, 2008). Um dos fatores limitantes para o MDS é a escala de trabalho, a qual, segundo Cavazzi et al., (2013), apresentam dois problemas principais associados ao tamanho do pixel do MDE que são: i) atributos do terreno produzidos a uma resolução espacial muito fina carregam excesso de detalhes gerando muito ruído, o que pode prejudicar a acurácia da predição; ii) baixas resoluções espaciais podem generalizar as propriedades da superfície perdendo sua capacidade preditiva. Outro inconveniente no uso de MDEs de elevada resolução espacial é o grande incremento no volume de dados armazenado, necessitando de unidades de armazenamento muito robustas. Os mesmos autores salientam que com o aumento

do uso de equipamentos que possibilitam a obtenção de dados com ERT, que permitem obter MDEs com elevada resolução espacial, tem aumentado o uso de *rasters* com este nível de resolução.

Os atributos do terreno oriundos do processamento do MDE são amplamente utilizados como preditores em MDS em função de sua correlação com a formação do solo (BEHRENS et al., 2010). As variáveis morfométricas fornecem uma gama de informações locais quantitativamente derivadas do MDE, incluindo declividade, face de exposição, curvatura, além de diversos índices, como o topográfico de umidade, rugosidade do terreno, entre outros, os quais constituem-se em importantes dados para uma gama de modelos ambientais (GALLANT & HUTCHINSON, 1997).

O presente estudo teve como objetivo avaliar a influência da mudança do tamanho da célula de MDEs, na qualidade das variáveis morfométricas utilizadas como preditores para MDS. Esta influência será traduzida em medidas de sensibilidade a alteração no tamanho de células.

2. Materiais e Métodos

2.1 Área de estudo

A península Keller está localizada na Baía do Almirantado, Ilha Rei George, entre as enseadas Martel e Mackellar, na região denominada de Antártica Marítima ao sul das ilhas Shetlands do Sul (Figura 1). Ela possui aproximadamente 500 ha de área, com aproximadamente 4 km de extensão, considerando a direção norte-sul e com aproximadamente 2 km de largura, considerando a direção leste-oeste (FRANCELINO et al., 2011). Ainda segundo esses autores, o ponto mais elevado da península possui altitude de 380 m e possui temperatura média anual de $-1,8^{\circ}\text{C}$ com precipitações médias de 360 mm ano^{-1} . Nesta península localiza-se a Estação Antártica Comandante Ferraz, que começou a operar no ano de 1984 (SANTOS et al., 2005). As principais modificações ocorridas no relevo são em função dos processos de intemperismo físico e solifluxão, ocorridos através dos ciclos de congelamento e descongelamento, além dos ventos fortes que atingem a região

(FRANCELINO, 2004).

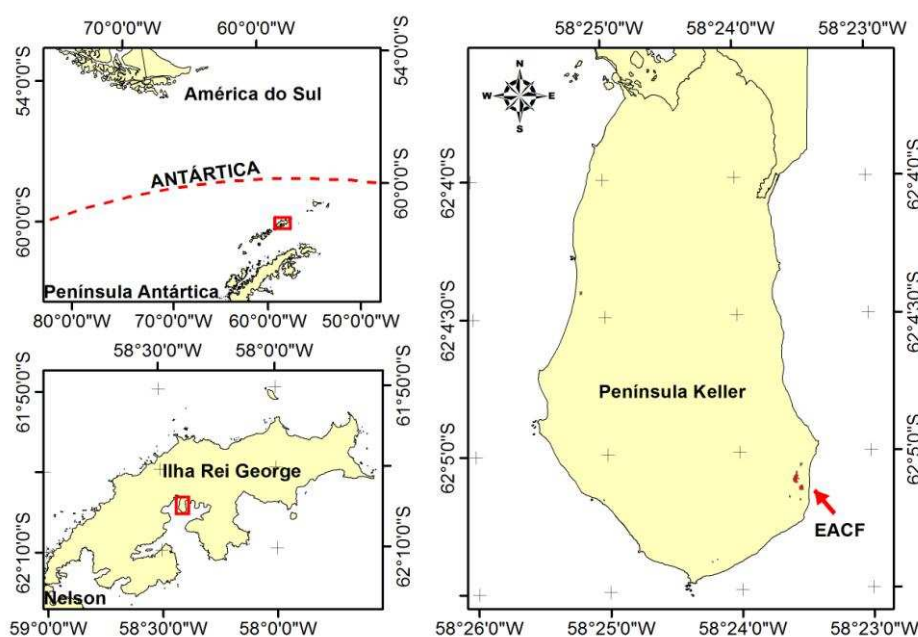


Figura 1 – Mapa de localização da área de estudo.

2.2 Levantamento de elevada resolução topográfica (ERT)

Para a obtenção dos MDEs a partir de uma nuvem tridimensional de pontos, foi realizado um levantamento de ERT nos verões Austrais 2014/2015 e 2015/2016, utilizando-se o equipamento Laser Scanner Terrestre (TLS) da fabricante RIEGL modelo VZ-1000 (Figura 2). Foi utilizado para levantamento das coordenadas das estações o receptor GNSS da Leica Modelo GS08 Plus. Os dados para determinação das coordenadas foram pós-processadas utilizando-se como base o receptor GNSS localizado nas proximidades da Estação Antártica Comandante Ferraz (EACF)¹. Os rastreamentos GNSS foram realizados por 10 minutos em cada ponto onde foi instalado o TLS, sendo obtidas soluções fixas para todos os pontos.

1 Projeto Geoespaço (<http://www.inpe.br/crs/pan/pesquisas/geoespaco.php>), processo CNPq/PROANTAR: 556872/2009-6, coordenada pela Dr^a Emilia Correa, da Universidade Presbiteriana Mackenzie.



Figura 2 – Fotografia do equipamento TLS utilizado, com antena GPS acoplada ao mesmo.

Para que toda a área da península fosse varrida pelo equipamento TLS, foi necessário que se instalasse o mesmo, em 81 pontos diferentes (Figura 3). Essa configuração e número de pontos instalado, foi suficiente para a varredura de toda a área, ocorrendo pequenas omissões no terreno.

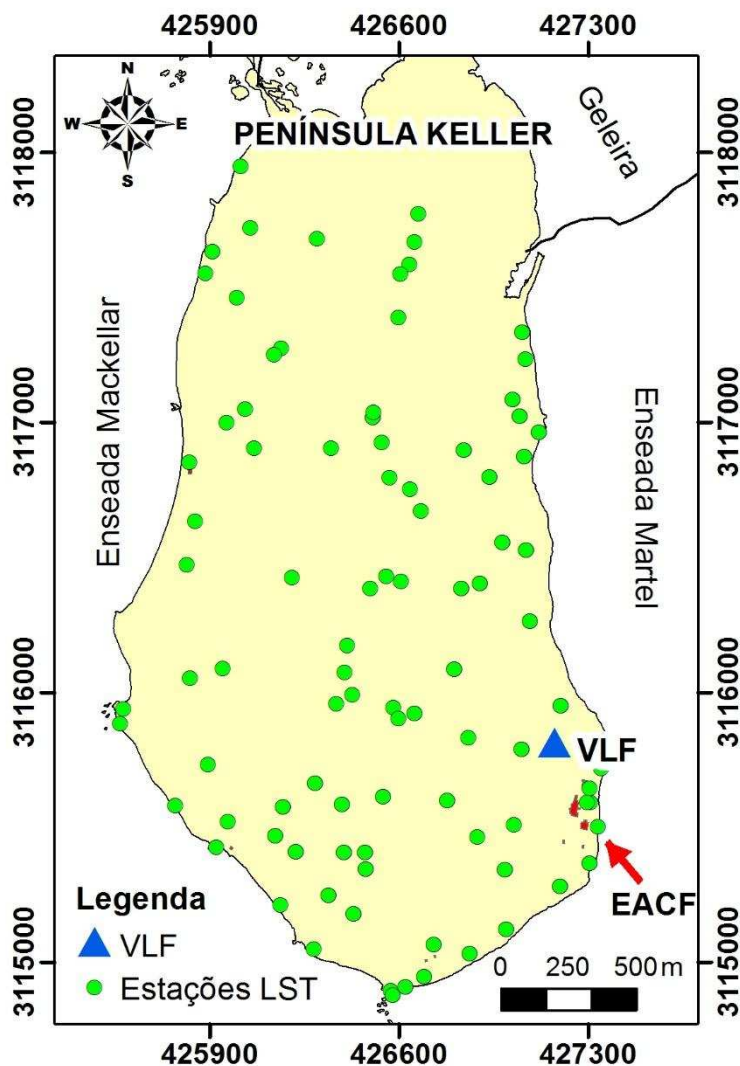


Figura 3 – Localização dos 81 pontos (estações) onde foram realizadas as varreduras.

2.3 Processamento dos dados de campo

Para o processamento dos dados GNSS foi utilizado o programa computacional Leica Geoffice, utilizando como base a antena situada no local denominado VLF, com latitude de $62^{\circ} 04' 58,11696$ S, longitude de $58^{\circ} 23' 39,30959$ W e altura elipsoidal de 57,53 m (Figura 3). Todos os dados foram referenciados ao *Datum* WGS84 com suas altitudes referenciadas a altura elipsoidal.

Obtidas as coordenadas de todos os pontos rastreados, foi realizado o processamento de todos as 81 varreduras realizadas com o TLS. Para tanto utilizou-

se o programa computacional Riscan Pro, versão 2.1 64 bits. A nuvem bruta de pontos foi processada removendo elementos que não correspondiam a superfície estudada, como edificações, aves, etc. A nuvem final de pontos ficou composta de aproximadamente 270 milhões de pontos, o que correspondeu a aproximadamente 54 pontos por metro quadrado. No programa ArcGIS 10.1 (ESRI, 2012) os pontos foram interpolados através da ferramenta *LAS dataset to raster*, com registro de células utilizando a média (av) e método de preenchimento de espaços vazios com o interpolador *Natural Neighbor* (nn), gerando MDEs com resoluções espaciais de 30,0; 20,0; 10,0; 5,0; 1,0; 0,5 e 0,2 metros.

A partir desses MDEs, foram derivados 39 temas morfométricos (Tabelas 1, 2 e 3) para comparação entre os mesmos para as diferentes resoluções espaciais, com o objetivo de verificar sua influência nos resultados gerados.

2.4 Obtenção das variáveis morfométricas

Para comparar a influência da resolução espacial nas diferentes variáveis morfométricas passíveis de serem obtidas a partir dos MDEs, foram gerados 39 derivadas dele (Tabelas 1, 2 e 3).

Tabela 1 – variáveis morfométricas do grupo das curvaturas

Derivadas da Curvatura		
Nº	VARIÁVEIS MORFOMÉTRICA	Descrição Breve
1	Curvatura de Seção	
2	Curvatura da linha de fluxo	
3	Curvatura Geral	
4	Curvatura Longitudinal	
5	Curvatura Máximal	Na seção normal local
6	Curvatura Minimal	Referente a seção normal local
7	Curvatura Plana	
8	Curvatura de Perfil	Descreve o segundo mecanismo de acumulação
9	Curvatura Tangencial	Descreve o primeiro mecanismo de acumulação
10	Classificação da Curvatura	Divide as curvaturas em 8 classes (discreto)
11	Curvatura Total	TOTAC = MEANC2 – DIFFC2 (Curvatura média menos Diferença de Curvatura)

Tabela 2 – variáveis morfométricas do grupo da radiação solar local

Derivadas da Radiação Solar		
Nº	VARIÁVEIS MORFOMÉTRICA	Descrição Breve
12	Solrad_diffuse1	Radiação solar difusa incidente no mês de janeiro
13	Solrad_diffuse2	Radiação solar difusa incidente no mês de junho
14	Solrad_direct1	Radiação solar direta incidente no mês de janeiro
15	Solrad_direct2	Radiação solar direta incidente no mês de junho
16	Solrad_dur1	Duração da radiação solar no mês de janeiro
17	Solrad_dur2	Duração da radiação solar no mês de junho
18	Solrad_total1	Radiação solar total (direta mais difusa) incidente no mês de janeiro
19	Solrad_total2	Radiação solar total (direta mais difusa) incidente no mês de junho

Tabela 3 – outras variáveis utilizadas

Outras covariáveis		
Nº	VARIÁVEIS MORFOMÉTRICAS	Descrição Breve
20	Face de Exposição do Relevo	Orientação em relação aos pontos cardiais de cada face do relevo;
21	Índice de Convergência	Calcula um índice de convergência / divergência em relação ao escoamento superficial;
22	Diferença	Diferença no gradiente de declividade
23	Aquecimento Anisotrópico Diurno	Representa a assimetria no aquecimento comparando as faces leste e oeste locais, baseado em horas de radiação;
24	Gradiente	Corresponde ao gradiente hidráulico
25	Índice de Balanço de Massa	Índice que representa o balanço de massa em cada pixel
26	Modelo Digital de Elevação	Representa a elevação em cada célula do modelo;
27	Posição de Meia Encosta	Representa a distância em relação ao topo e vale, variando entre 0 e 1. Cobre as regiões mais quentes da declividade
28	Índice de multiresolução da planicidade do topo dos vales	Utiliza a planicidade e outras características dos fundos de vale. O método demonstra uma idéia baseada na lógica Fuzzi de Análise de Paisagens. Varia entre 0 e 1
29	Elevação normalizada	É uma medida de altura relativa de um ponto ao invés de seu valor elevação propriamente dito. Atribui o valor 1 para o ponto mais alto e o valor 0 para a posição mais baixa dentro de uma respectiva área de busca
30	Área real da superfície	Não considera a superfície como projetada, e sim a superfície real
31	Declividade	Representa a declividade local que pode ser angular ou em porcentagem;
32	Elevação da declividade	
33	Elevação Padronizada	É o produto da altura normalizada multiplicada pela altura absoluta
34	Superfície de pontos específicos	Indica diferenças entre pontos específicos de mudança da superfície, resultando em valores discretos positivos e negativos
35	Índice de Rugosidade Terrestre	Calcula a diferença dos valores de elevação a partir de uma célula central e suas oito células vizinhas
36	Convexidade da superfície terrestre	É calculado como a razão entre o número de células que têm curvatura positiva (células convexas) para o número de todas as células válidas dentro de um raio de pesquisa específico
37	Textura da superfície terrestre	Divide a textura da superfície em 8, 12 ou 16 classes
38	Profundidade dos vales	Inverte a elevação, deriva as redes de drenagem e calcula a
39	Vetor de medidas de rugosidade	Mede a variação na rugosidade do terreno, variando de
40	Índice topográfico de umidade	Descreve a tendência de cada célula em acumular água em função do relevo

Para a obtenção destas variáveis elaborou-se um *script* no programa computacional R (R CORE TEAM, 2015), que de forma automatizada gerou, a partir de cada MDE nas sete diferentes resoluções espaciais, as variáveis listadas nas Tabelas 1, 2 e 3. Utilizou-se para tanto os pacotes “RSAGA” (BRENNING, 2008) e

“*raster*” (ROBERT, 2015), sendo que o primeiro possibilita o uso das ferramentas computacionais do programa SAGA GIS e o segundo permite operações com matrizes. Deste processamento resultaram então as 39 variáveis para cada um dos sete MDEs nas diferentes resoluções, obtendo portanto 280 matrizes.

2.6 Comparação entre as variáveis morfométricas

Para a comparação das variáveis morfométricas, inicialmente foi gerada uma nuvem de pontos composta por 5.397 pontos a partir do MDE de resolução de 30m, que serviu como base para a extração dos valores das variáveis morfométricas para cada resolução estudada. De posse destes pontos, foram extraídos os valores de cada variável morfométrica para todas as resoluções e comparadas entre si. Por fim foram realizadas as correlações (Pearson e RDC) entre as variáveis morfométricas originadas a partir dos sete MDEs, tomando como referência os modelos com tamanho de célula de 0,2; 1,0 e 30,0 metros. Algumas correlações foram plotadas em gráficos para cada variável morfométrica estudada, para exemplificar os comportamentos padrão. Para a realização específica de todo este processamento, foi elaborado um *script* no programa computacional R.

Para determinar a sensibilidade das variáveis morfométricas em relação a resolução espacial, foi realizada a análise de agrupamento das correlações, separando as mesmas em 3 grupos (clusters): sensíveis a resolução espacial (SR), pouco sensíveis a resolução espacial (PSR) e altamente sensíveis e resolução espacial (ASR). À título de comparação, este agrupamento foi realizado tomando como referência, em cada momento, as resoluções espaciais de 0,2; 1,0 e 30 m. Este agrupamento foi realizado usando a ferramenta *cluster*, onde se define o número de grupos em que deseja separar o conjunto (3) de dados e são identificados padrões nos valores estudados.

Para verificar a correlação, ajuste linear e não linear das distribuições, foram selecionadas como exemplo duas variáveis morfométricas de cada grupo das 3 sensibilidades sugeridas, tendo como critério de escolha que a variável deveria estar dentro do mesmo grupo, quando tomadas as 3 resoluções de referência. Foram então escolhidas entre as PSR as variáveis Face de Exposição e Declividade, entre as ASR Gradiente e Índice de Balanço de Massa e por fim, as SR Diferença e

Profundidade de Vale.

3. Resultados

3.1 Processamento dos dados de campo

O desvio padrão do erro, máximo obtido dos ajustamentos do levantamento do TLS foi de 0,0848 m e o valor mínimo de 0,0081 m. O desvio padrão do erro médio obtido, considerando todos os ajustamentos foi de 0,0245 m, o que representa valor 4,9 vezes maior que a precisão nominal do equipamento, sendo então considerado um ajustamento com grau de precisão satisfatório, quando considerados os levantamentos tradicionais para tal finalidade. Foi observado que a variância média, para latitude, longitude e altitude elipsoidal, considerando a tomada de todas as coordenadas das estações (medidas por GPS), foi de 0,90 mm; 1,63 mm e 1,9 mm, respectivamente. Porém, cabe salientar, que estes desvios, não podem ser considerados nos erros de ajustamento das varreduras, pois as coordenadas de cada estação foram consideradas fixas no ajustamento, variando apenas ΔRoll , ΔPit e ΔYaw , que representam os ajustes vertical, horizontal e azimute.

3.3 Comparação entre as variáveis morfométricas

O agrupamento das variáveis morfométricas em três grupos (ASR, PSR e SR) em função das resoluções é observado na Figura 4.

Variável Morfométrica	Referência 20 cm (Pearson)	Referência 1 m (Pearson)	Referência 30 m (Pearson)	Referência 20 cm (RDC)	Referência 1 m (RDC)	Referência 30 m (RDC)
Face de Exposição do Relevo	SR	PSR	SR	SR	PSR	PSR
Índice de Convergência	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura de Seção Transversal	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura da linha de fluxo	ASR	ASR	ASR	SR	SR	SR
Curvatura Geral	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura Longitudinal	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura Maximal	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura Minimal	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura Plana	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura de Perfil	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura Tangencial	ASR	ASR	ASR	ASR	ASR	ASR
Curvatura Total	ASR	ASR	ASR	SR	SR	SR
Classificação da Curvatura	ASR	ASR	ASR	ASR	ASR	ASR
Diferença	ASR	SR	ASR	ASR	ASR	ASR
Aquecimento Anisotrópico Diurno	PSR	PSR	PSR	SR	PSR	PSR
Gradiente	ASR	ASR	ASR	SR	SR	SR
Índice de Balanço de Massa	ASR	ASR	ASR	ASR	ASR	ASR
Modelo Digital de Elevação	PSR	PSR	PSR	PSR	PSR	PSR
Posição de Meia Encosta	ASR	SR	ASR	ASR	ASR	SR
Índice de multiresolução da planicidade do topo dos vales	PSR	PSR	PSR	PSR	PSR	PSR
Elevação normalizada	SR	PSR	SR	ASR	SR	SR
Área real da superfície	SR	PSR	SR	PSR	PSR	PSR
Declividade	PSR	PSR	PSR	PSR	PSR	PSR
Elevação da declividade	SR	PSR	SR	SR	PSR	PSR
Radiação solar difusa 1	SR	PSR	SR	SR	PSR	PSR
Radiação solar difusa 2	SR	PSR	SR	SR	PSR	PSR
Radiação solar direta 1	SR	PSR	SR	SR	PSR	PSR
Radiação solar direta 2	PSR	PSR	PSR	SR	PSR	PSR
Duração da radiação solar 1	SR	PSR	SR	SR	PSR	PSR
Duração da radiação solar 2	SR	PSR	SR	SR	PSR	PSR
Radiação solar total 1	SR	PSR	SR	SR	PSR	PSR
Radiação solar total 2	PSR	PSR	PSR	SR	PSR	PSR
Elevação Padronizada	PSR	PSR	PSR	PSR	PSR	PSR
Superfície de pontos específicos	ASR	ASR	ASR	ASR	ASR	ASR
Índice de Rugosidade Terrestre	SR	PSR	SR	PSR	PSR	PSR
Convexidade da superfície terrestre	ASR	SR	ASR	ASR	ASR	ASR
Textura da superfície terrestre	ASR	SR	ASR	ASR	ASR	ASR
Profundidade dos vales	SR	SR	SR	ASR	SR	SR
Vetor de medidas de rugosidade	SR	PSR	PSR	PSR	PSR	PSR
Índice topográfico de umidade	PSR	PSR	SR	SR	PSR	PSR

Figura 4 – Agrupamento das covariáveis conforme sensibilidade e resolução de referência, em função das correlações de Pearson e RDC. Legenda: ASR = altamente sensível a resolução; PSR = pouco sensível a resolução; SR = sensível a resolução.

Entre as variáveis consideradas como SR (Figura 4), observa-se que ao usar outra resolução de referência, em alguns casos as variáveis podem se enquadrar em outra classe de sensibilidade, sendo que a maioria não alterou a classe com essa mudança. Esse tipo de informação é particularmente importante na aplicação de ferramentas de MDS, pois poderá aumentar o poder de predição do algoritmo.

As Figuras 5 e 6 representam a correlação de Pearson, ajuste linear e não linear das variáveis SR Diferença e Profundidade de Vale. É possível verificar que quando comparadas as diferentes resoluções, os gráficos se mostraram diferentes, ao alterar a resolução. As correlações também se alteram, segundo o que reforça a hipótese de que as variáveis inclusas neste grupo possuem sensibilidade a alteração na resolução espacial.

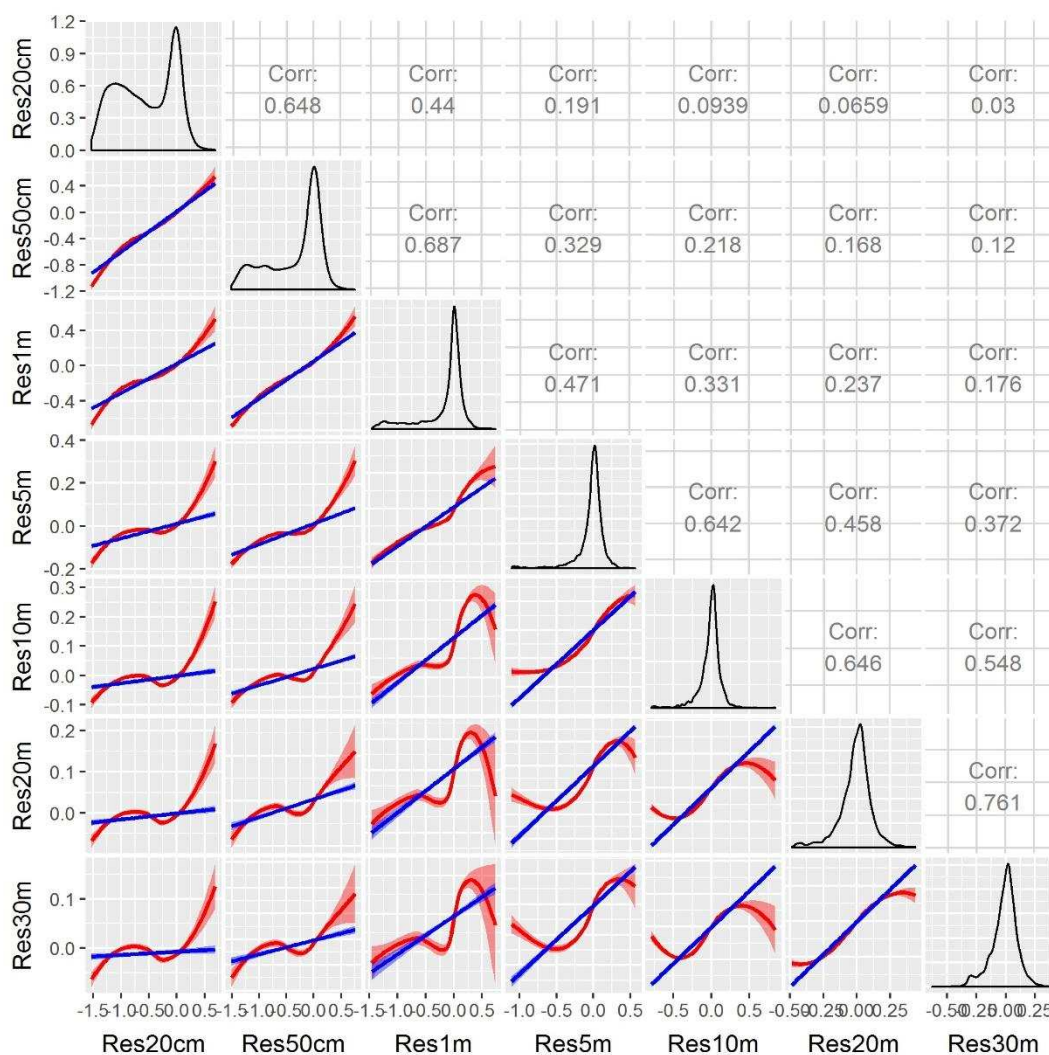


Figura 5– Correlação de Pearson, ajuste linear e não linear da variável SR Diferença.

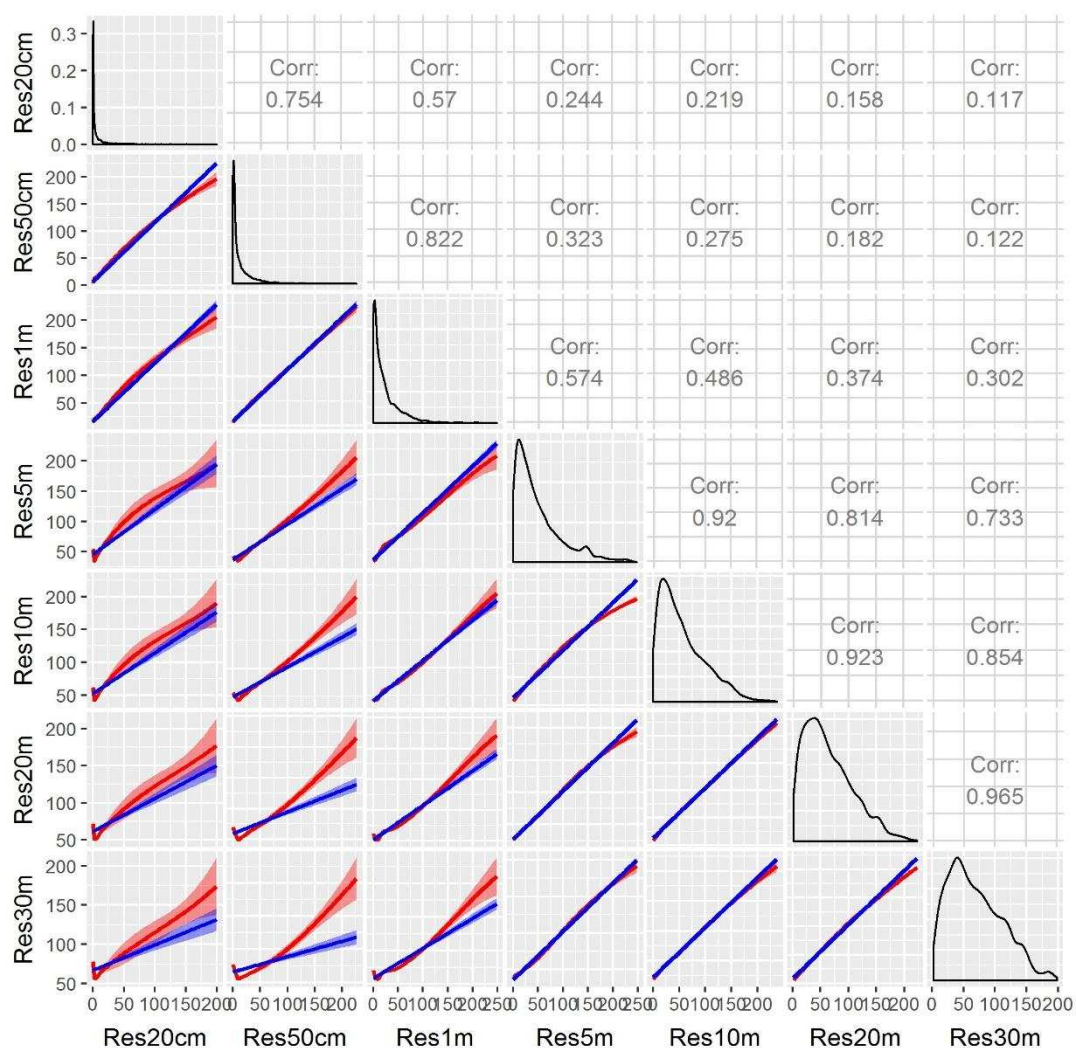


Figura 6– Correlação de Pearson, ajuste linear e não linear da variável SRProfundidade de Vale.

Em relação às variáveis ASR, apenas as variáveis Índice de convergência, curvatura maximal, posição da declividade e textura da superfície mudaram de grupo, as demais foram consideradas como altamente sensíveis à resolução espacial. As Figuras 7 e 8 indicam correlação, ajuste linear e não linear da variável ASR Gradiente e Índice de Balanço de Massa respectivamente. É possível observar a mudança nos coeficientes de correlação entre as mesmas, o que reforça a hipótese de sensibilidade alta, das variáveis supracitadas tomadas como exemplo. Quando considerada a correlação entre as diferentes resoluções, observa-se valores elevados da mesma juntamente com valores baixos e mantendo-se estáveis, reforçando mais uma vez a sensibilidade a resolução espacial.

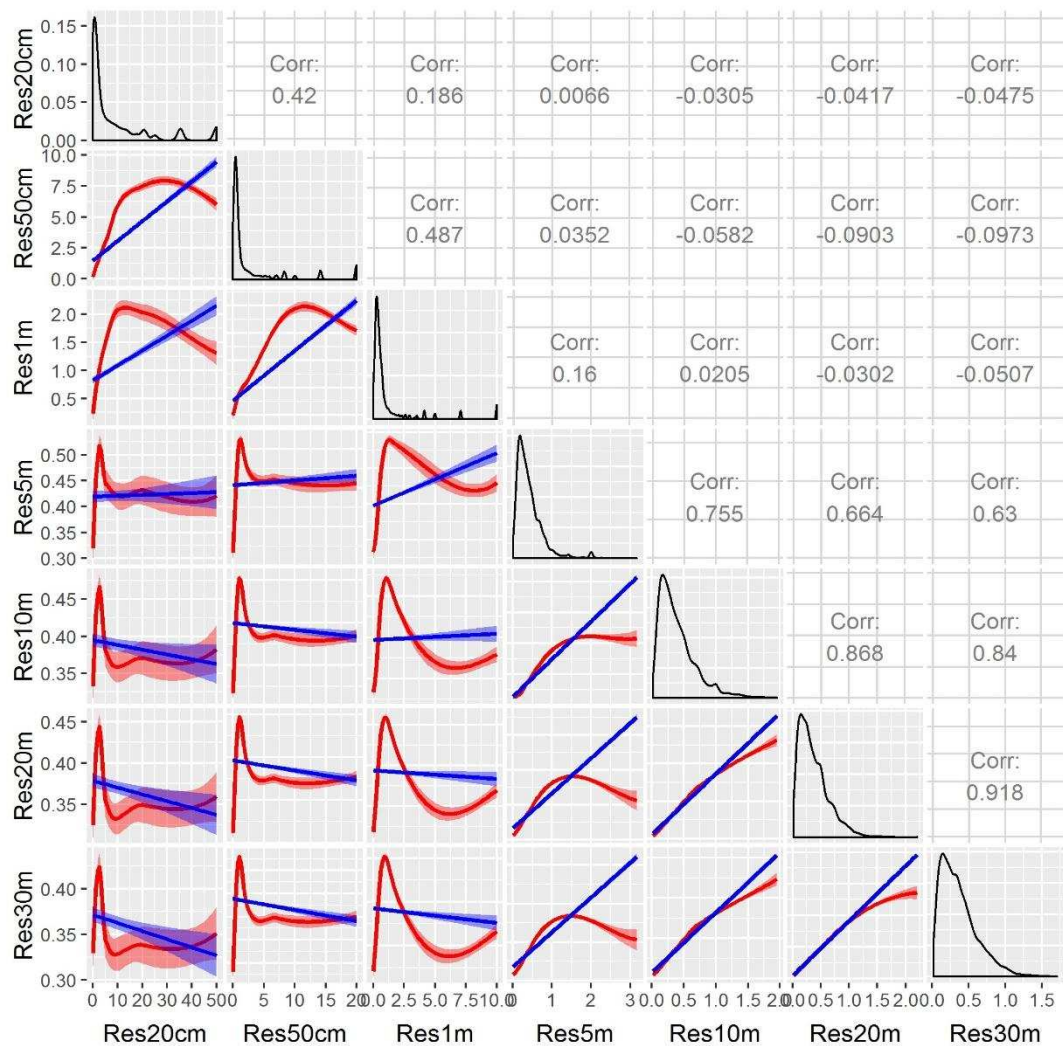


Figura 7 – Correlação de Pearson, ajuste linear e não linear da variável ASR Gradiente.

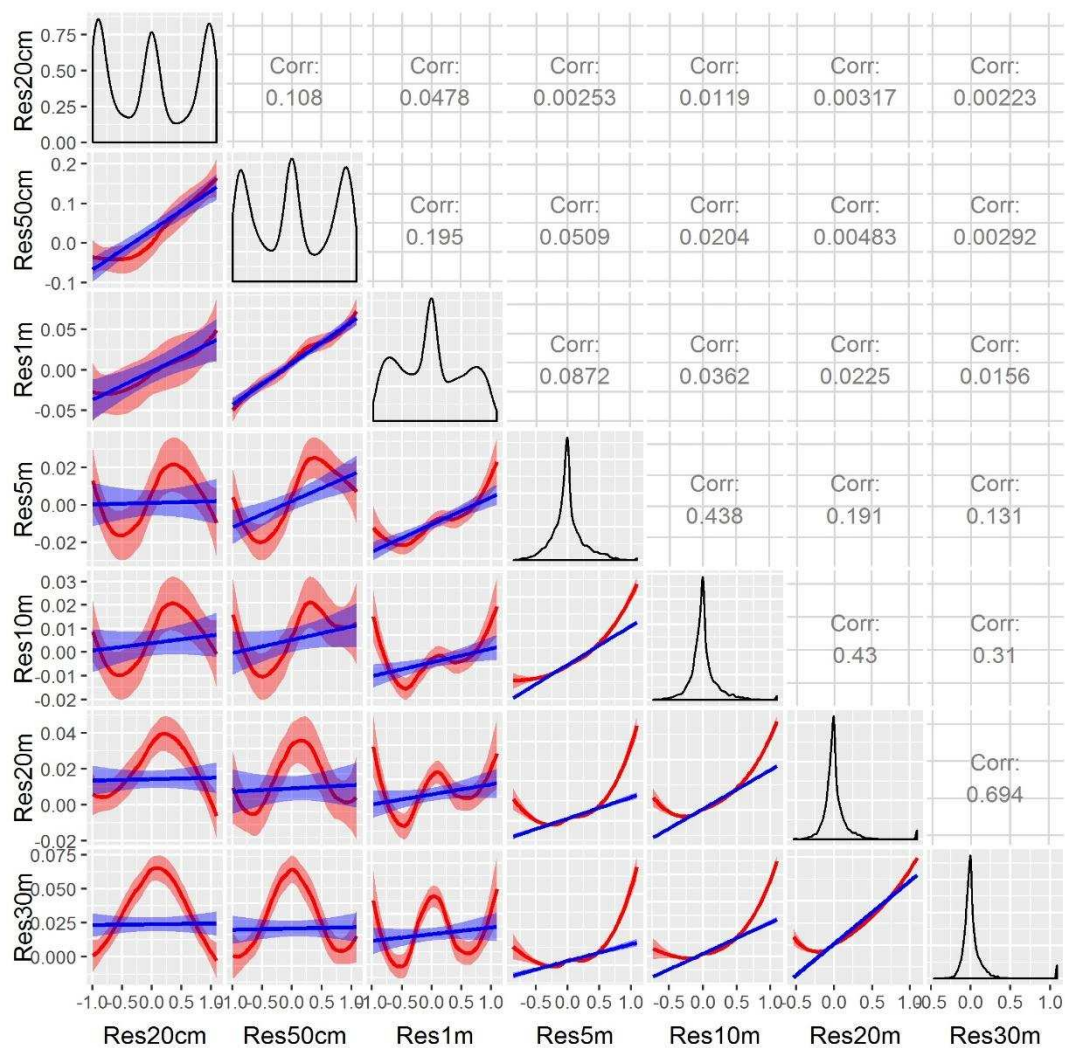


Figura 8 – Correlação de Pearson, ajuste linear e não linear da variável ASR Índice de Balanço de Massa.

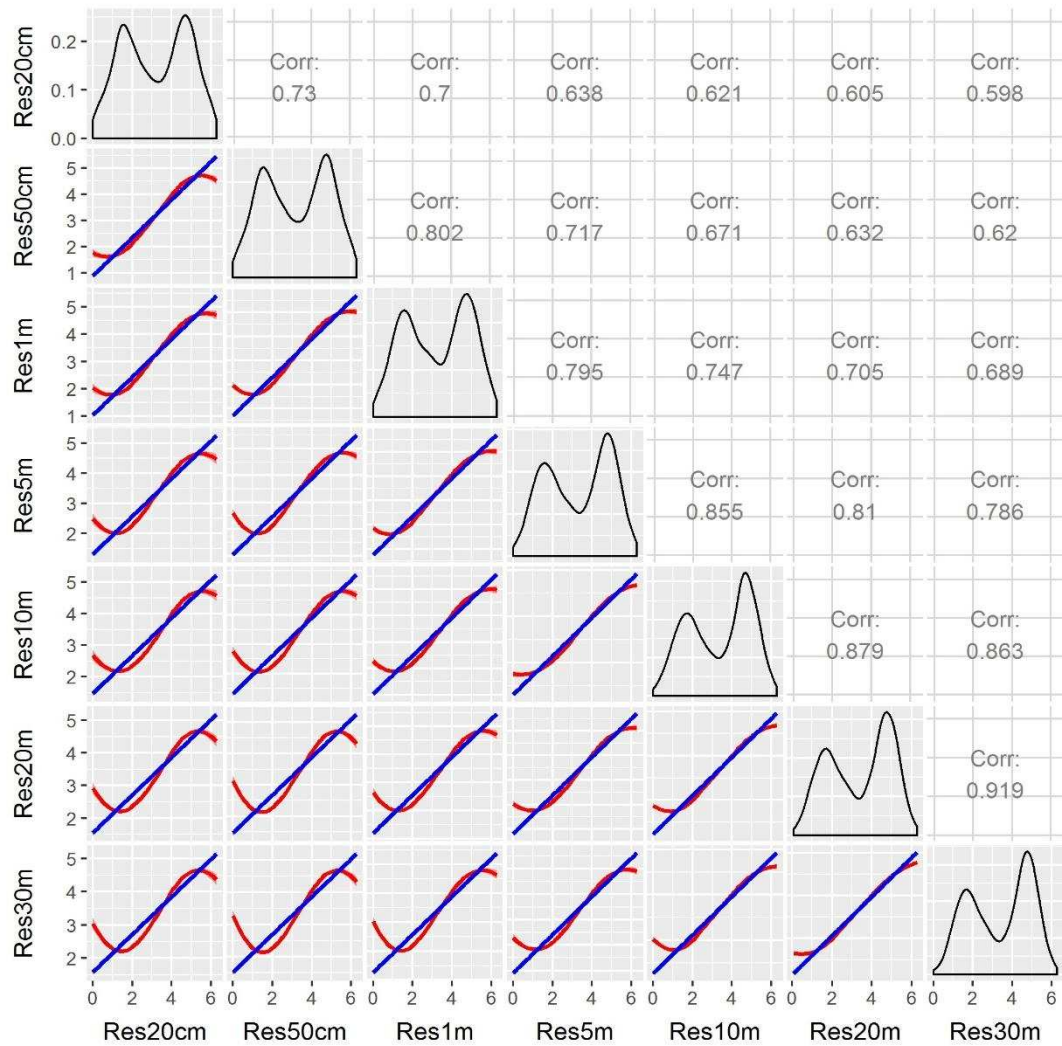


Figura 9–Correlação de Pearson, ajuste linear e não linear da variável PSR Face de exposição.

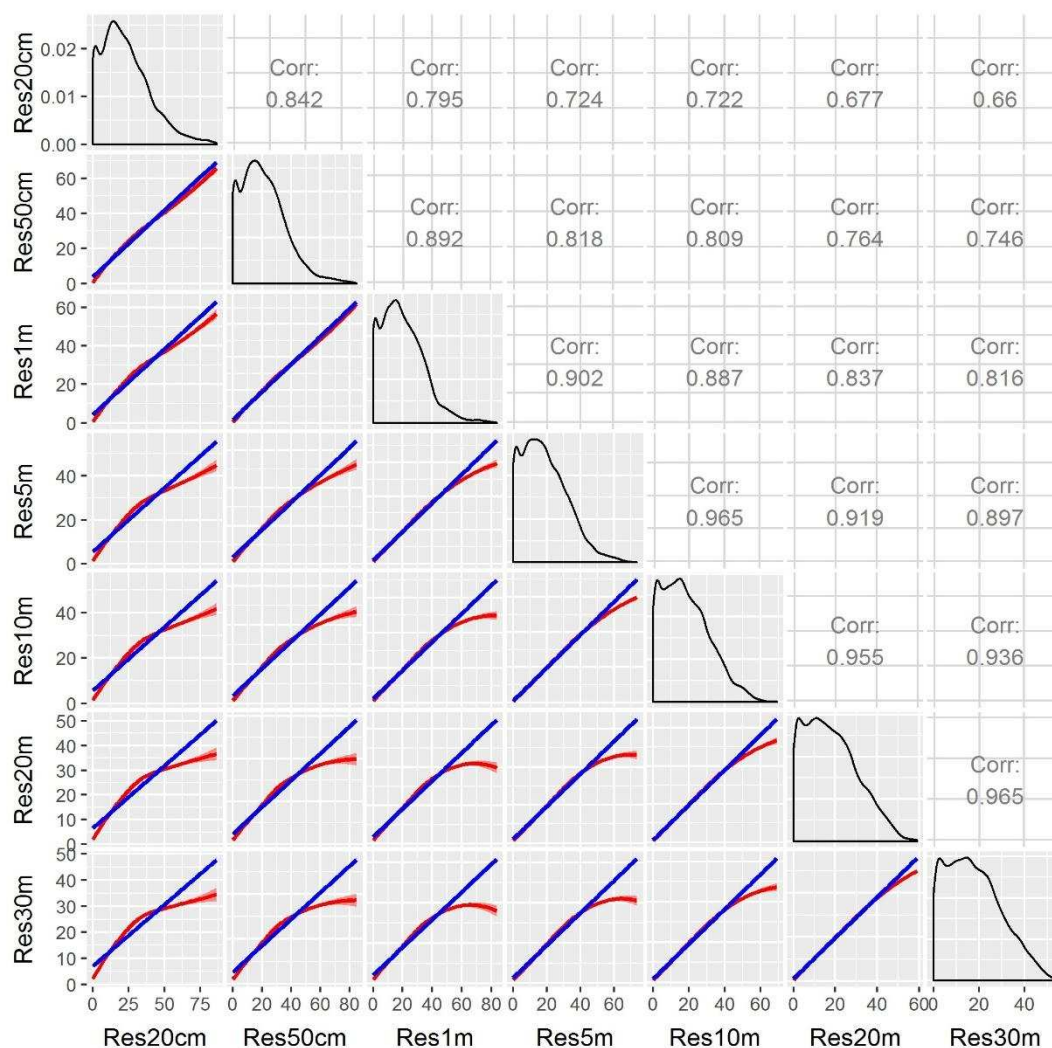
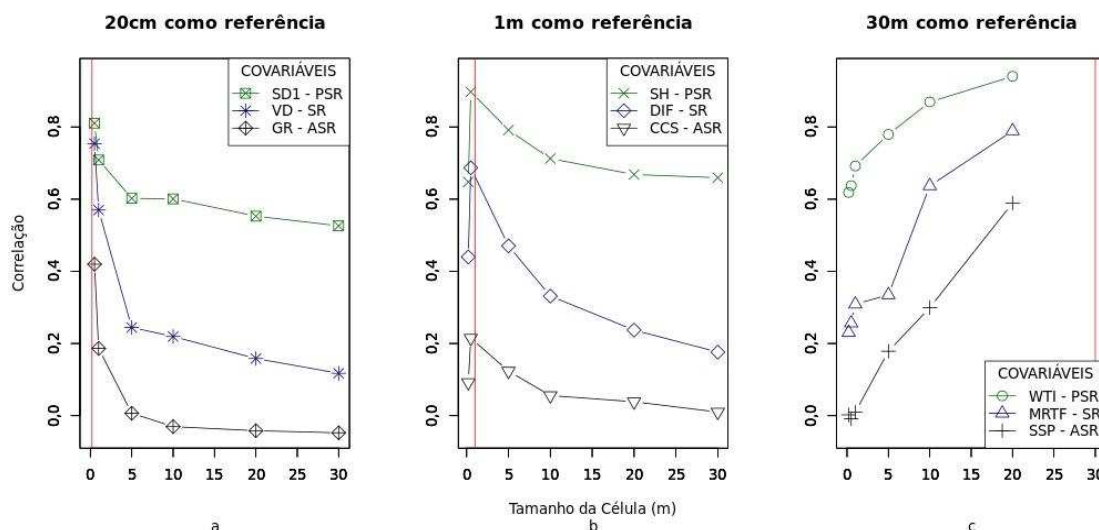


Figura 10 – Correlação de Pearson, ajuste linear e não linear da variável PSR Declividade.

As Figuras 9 e 10, representam a correlação, ajuste linear e não linear da variável PSR Face de exposição e Declividade, respectivamente. Observa-se que mesmo alterando a resolução da célula do *raster*, ocorrem poucas alterações entre os gráficos, correlações e modelagem. Neste caso as correlações se mantiveram estáveis, com a mudança da resolução, permanecendo com valores elevados de correlação, considerando a correlação de Pearson.

A Figura 11 demonstra os gráficos com os exemplos de padrões encontrados de correlação de Pearson para algumas variáveis utilizadas.



Legenda: SD1 = Radiação solar direta 1; VD = Profundidade de vale; GR = Gradiente; SH = Elevação da declividade; DIF = Diferença; CCS = Curvatura secante cruzada; WTI = índice topográfico de umidade; MRTF = MRRTF; SSP = Superfície específica de pontos; PSR = Pouco Sensível a Resolução; SR = Sensível à Resolução; ASR = Altamente Sensível a Resolução.

Figura 11 - Exemplos de padrões encontrados para os gráficos de correlação de Pearson das variáveis enquadradas como PSR, SR e ASR, tomando os tamanhos de célula de a) 0,20, b) 1,00 e c) 30,00 m como referências para as correlações (linhas verticais em vermelho).

De forma geral, quando utilizada a resolução de 30 m, como referência, a análise gráfica indica variações quando comparada a outras resoluções, porém a análise de agrupamento indicou poucas mudanças entre as variáveis nos grupos (SR, PSR e ASR).

4. Discussão

Ao analisar os resultados, é possível afirmar que a mesma variável, quando usada em diferentes resoluções, como variável em modelos preditivos para MDS, poderá carregar informações relevantes e não redundantes, podendo ser usadas em conjunto no modelo preditivo. Isto pode ser afirmado devido ao fato de que com a mudança no tamanho da célula, algumas variáveis, vão diminuindo suas correlações, de um tamanho maior para um tamanho menor na mesma variável. Portanto, a mesma variável, obtida a partir de um tamanho diferente de célula, poderá carregar

informação não redundante.

No caso das variáveis PSR, segundo os resultados obtidos neste trabalho, é possível utilizar somente a variável com baixa resolução espacial, sem perder a informação da mesma. Existem poucas diferenças entre a mesma variável tomada em diferentes resoluções espaciais, considerando os valores numéricos obtidos das correlações. Se tomada essa decisão, é possível diminuir o processamento computacional, o que agilizará o seu processamento.

Considerando o comportamento da correlação, entre as diferentes resoluções com as resoluções de referência, observa-se que não ocorrem grandes alterações quando se compara os gráficos gerados (Figura 11- a, b e c). Isto significa que ao mudar o tamanho da célula de referência, não ocorrem grandes alterações na classificação gerada pelos agrupamentos (*cluster* = ASR, SR e PSR) das diferentes variáveis estudadas. Embora a diferenciação utilizando os agrupamentos como referência não se alterasse muito, visualmente é possível distinguir uma mudança de comportamento, notoriamente evidenciado entre a Figura 11-a e as Figuras 11-b e c, sendo que as duas últimas se mantiveram muito semelhantes, o que poderia ser explicado em parte pela menor distância entre as duas referências.

As variáveis que mantiveram seus gráficos estáveis e com altas correlações indicam as que são pouco sensíveis a mudança do tamanho da célula como no exemplo de Figuras 11-a, b e c em suas linhas superiores (SD1, SH e WTI), ou seja, não ocorre incremento na qualidade das variáveis, utilizando-se células com tamanho menor. Desta forma é possível lançar mão do uso das variáveis PSR com o maior tamanho de célula (30m), em detrimento da agilidade no processamento computacional.

Aquelas variáveis que possuíam gráficos que demonstrassem maior estabilidade, porém com baixas correlações, representam variáveis altamente sensíveis a mudança do tamanho do pixel (ASR) como no exemplo da Figura 11-a e b, em suas linhas inferiores (GR e CCS). Neste caso, ao utilizar a mesma variável com tamanho de pixel de 30,0 m e 0,2 m, por exemplo, é possível que cada uma delas carregue consigo uma informação diferenciada, quando usadas como preditores em MD. Portanto é possível incluir a mesma variável ASR, com diferentes resoluções, no processo de seleção de variáveis preditoras em MD.

Por fim, aquelas que obtiveram gráficos com um comportamento ascendente

ou descendente, representam as, aqui chamadas, variáveis sensíveis a mudança no tamanho da célula, como no exemplo da Figura 11-a, b e c, em suas linhas intermediárias (VD, DIF e MRTF). Este caso representa tanto altas quanto baixas correlações, quando se altera o tamanho da célula de uma mesma variável. A variável MRRTF, por exemplo, quando realizada a sua correlação dos diferentes tamanhos de célula, com o tamanho de 30m, demonstra uma alta correlação com os tamanhos mais próximos aos 30m, e a medida que diminui-se o tamanho da célula, a correlação também diminui. Neste caso, ao adotar o critério de utilizar diferentes tamanhos de células para a mesma variável como variáveis preditivas para MD, recomenda-se que seja utilizada a mesma variável com pequeno valor do tamanho de célula somando-se o uso da mesma com alto valor de tamanho de célula. A diferença com relação às ASR é que para as variáveis SR, deve ser adotado o critério de utilização de valores extremos de tamanho de célula de uma mesma variável. Já para ASR, é possível utilizar diversos valores de tamanho de célula de uma mesma variável, com todas as outras variáveis para a seleção de preditores em MDS.

Os resultados apresentados indicam de forma geral que existem variáveis que poderão ser utilizadas como variáveis em diferentes resoluções espaciais. Essas variáveis podem carregar informações diferenciadas, o que auxiliaria na predição para MD. Observa-se também que existem variáveis que independente da resolução espacial, poderão carregar a mesma informação para predição. Neste caso se optaria pelo uso daquela que possua a pior resolução espacial, em detrimento do processamento e menor redundância de dados.

5. Conclusões

Dependendo do elemento morfométrico utilizado como variável, como por exemplo, Curvatura secante cruzada, Curvatura da linha de fluxo, Curvatura geral, Curvatura minimal, Curvatura maximal, Curvatura plana, Curvatura de perfil, Curvatura tangencial, Classificação da curvatura, Índice de balanço de massa e Superfície de pontos específicos, a resolução espacial poderá melhorar o modelo de predição. Neste sentido a mesma variável, enquadrada neste trabalho no grupo das

altamente sensíveis a resolução, como elencado acima, poderá ser utilizada em diferentes tamanhos de célula, carregando informações diferenciadas como preditoras em MD.

A tomada de uma resolução como referência para estudar o comportamento, não altera de forma geral, a tomada de decisão. A maioria das variáveis quando enquadradas em um dos grupos aqui propostos, se mantiveram no mesmo grupo, quando alterada a sua referência.

Para obtenção de variáveis pouco sensíveis, tais como Aquecimento anisotrópico diurno, Área real de superfície, Índice de multiresolução da planicidade do topo dos vales, Declividade, Elevação padronizada e Vetor de medida de rugosidade é possível utilizar modelos digitais de elevação com resoluções espaciais de 30 m, para a derivação destas variáveis, disponibilizados gratuitamente na internet, com a vantagem da abrangência de grandes áreas.

A comparação entre as correlações de Pearson e Randomized Dependence Coefficient (RDC) indica as variáveis que possuem correlações lineares ou não. Quando os valores são concordantes e elevados, entre as duas correlações, é possível afirmar que o comportamento é linear. Quando os valores das correlações forem baixos em Pearson e altos em RDC, é possível afirmar que existe uma correlação e a mesma é não linear.

Referências

ABRAMS, M.; HOOK, S. ASTER User Handbook Version 2. **Jet Propulsion**, v. 2003, n. 23/09/2003, p. 135, 2002.

BEHRENS, T. et al. Multi-scale digital terrain analysis and feature selection for digital soil mapping. **Geoderma**, v. 155, n. 3–4, p. 175–185, 2010.

BORNER, T. et al. ALOS PALSAR products verification. **International Geoscience and Remote Sensing Symposium (IGARSS)**. Anais, 2007.

BRENNING, A. Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. In: **SAGA--Seconds Out (Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, vol. 19)**. v. 19, p. 23–32, 2008.

CAVAZZI, S. et al. Are fine resolution digital elevation models always the best choice in digital soil mapping? **Geoderma**, v. 195–196, p. 111–121, 2013.

ESRI 2012. ArcGIS Desktop: Release 10.1. Redlands, CA: Environmental Systems Research Institute.

FARR, T. G. et al. The shuttle radar topography mission. **Reviews of Geophysics**, v. 45, n. 2, 2007.

FRANCELINO, M. R. **GEOPROCESSAMENTO APLICADO AO MONITORAMENTO AMBIENTAL DA ANTÁRTICA MARÍTIMA: SOLOS, GEOMORFOLOGIA E COBERTURA VEGETAL DA PENÍNSULA KELLER**. Universidade Federal de Viçosa, 2004.

FRANCELINO, M. R. et al. Geomorphology and soils distribution under paraglacial conditions in an ice-free area of Admiralty Bay, King George Island, Antarctica. **Catena**, v. 85, n. 3, p. 194–204, 2011.

GALLANT, J. C.; HUTCHINSON, M. F. Scale dependence in terrain analysis. **Mathematics and Computers in Simulation**, v. 43, n. 3–6, p. 313–321, 1997.

LAGUE, D.; BRODU, N.; LEROUX, J. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 82, p. 10–26, 2013.

LOPEZ-PAZ, D.; HENNIG, P.; SCHÖLKOPF, B. The Randomized Dependence Coefficient. **Advances in neural** , p. 1–9, 2013.

LUIZ, S.; SANTOS, A. R. S.; BRENNER, T. L. Geração de modelo digital de elevação a partir de imagens reference stereo do satélite IKONOS. **Anais do XIII Simpósio Brasileiro de Sensoriamento Remoto**, v. 2, n. Figura 1, p. 581–587, 2007.

MINASNY, B.; MCBRATNEY, A. B.; SALVADOR-BLANES, S. Quantitative models for pedogenesis - A review. **Geoderma**, 2008.

PASSALACQUA, P. et al. Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review. **Earth-Science Reviews**, 2015.

R Core Team. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. ISBN 3-900051-07-0, URL <https://www.R-project.org/>. 2015).

Robert J. H. raster: **Geographic Data Analysis and Modeling**. R package version

2.4-20. <http://CRAN.R-project.org/package=raster>, 2015.

SANTOS, I. R. et al. Heavy metal contamination in coastal sediments and soils near the Brazilian Antarctic Station, King George Island. **Marine pollution bulletin**, v. 50, n. 2, p. 185–94, 2005.

TAROLLI, P. High-resolution topography for understanding Earth surface processes: Opportunities and challenges. **Geomorphology**, 2014.

CAPÍTULO 3

Grades amostrais e interpoladores para elaboração de Modelos Digitais de Elevação – MDE

Resumo

A determinação do tamanho da amostragem para mapear fenômenos como atributos do solo, elevação, entre outros é de suma importância para interpolação de dados e geração de superfícies preditivas. O presente trabalho teve o objetivo principal verificar a melhor forma, densidade e o melhor interpolador para a predição de elevação em duas áreas com diferentes rugosidades. Este trabalho teve também o objetivo de comparar a validação interna do interpolador da krigagem com a validação externa. O trabalho foi desenvolvido nas Penínsulas Keller (PK) e Coppermine (PC), situadas na Antártica Marítima, onde foi realizado um levantamento de elevada resolução topográfica (ERT) com o uso do Laser Scanner Terrestre. Para testar as densidades e padrões das grades amostrais, foram geradas 60 grades amostrais para 8 (0,30 a 4,56 pt ha⁻¹) densidades de pontos por hectare. Em cada ponto amostrado foi extraído o valor de elevação do MDE de 1m de célula gerado a partir do levantamento de ERT. A partir das grades amostradas, foram interpolados os valores e gerados MDEs. As superfícies preditas foram comparadas por validação cruzada e validação e seus RMSEs foram computados. Para a PC o RMSE decresceu de forma assintótica até o menor valor encontrado, para a grade mais densa. Os resultados dos interpoladores mostraram diferentes comportamentos entre as duas áreas estudadas. Para a PK o melhor interpolador foi o Spline enquanto que para a PC, foi a krigagem simples. A comparação entre a validação interna e externa demonstra que a krigagem obtém resultados melhores do que sua validação interna demonstra, em relação ao RMSE.

Palavras chave: amostragem, interpoladores, geoestatística, grades amostrais

Sample grids and interpolators for Digital Elevation Models – DEM elaboration

Abstract

The sample size determination to map phenomena such as soil attributes, elevation, among others is important for data interpolation and predictive surface generation. The present work had as main objective verify the best shape, density and the best interpolator for the elevation prediction in two areas with different roughness. This work also had the objective to compare internal validation of the kriging interpolator with external validation. The work was developed in the Keller (KP) and Coppermine (CP) Peninsulas, located in the Maritime Antarctic, where a high-resolution topographic survey (HRT) was performed with the use of the Terrestrial Laser Scanner. To test the densities and patterns of the sampling grids, 60 of them were generated for 8 (0.30 to 4.56 ptha⁻¹) point densities per hectare. At each point sampled, the elevation value generated from the HRT survey of the DEM with one meter cell was extracted. From the grids sampled, the values were interpolated and DEMs generated. The predicted surfaces were compared by cross-validation and validation and their RMSEs were computed. For CP the RMSE decreased asymptotically to the lowest value found for the densest grid. The results of the interpolators showed different behaviors between the two studied areas. For KP, the Spline was the best interpolator whereas for the CP, it was the simple kriging. The comparison between internal and external validation demonstrates that kriging obtains better results than its internal validation in relation to RMSE.

Keywords: sampling, interpolators, geostatistics, sample grids

1. Introdução

Brus & Heuvelink (2007) citam dois estágios principais que podem ser distinguidos para mapear variáveis ambientais sendo o primeiro a amostragem e o segundo a predição. Neste sentido a determinação do tamanho da amostragem para mapear fenômenos é de suma importância para interpolação de dados e geração de superfícies preditivas contínuas. A determinação do tamanho amostral pode muitas vezes ser superestimado o que normalmente causa uma perda de tempo e desperdício financeiro (DE GRUIJTER et al., 2008).

A grade amostral para mapear solos, atributos, relevo, etc, tanto sua densidade, quanto sua regularidade e forma pode influenciar nos resultados de predição destes fenômenos. As grades regulares normalmente são usadas com forma hexagonal ou quadrada. A grade não regular utilizada é a aleatória, onde são amostrados locais de forma aleatória na área que se deseja mapear (HENGL, 2009). O mesmo autor descreve que a amostragem regular é aquela que cobre sistematicamente a área de interesse. Em contrapartida este método tem a desvantagem de que nas distâncias menores que o tamanho da grade, existirá uma baixa representatividade do fenômeno estudado. A amostragem aleatória possui a vantagem de representar diferentes distâncias entre pontos (pt), com a desvantagem de que a dispersão dos pontos na área de interesse é menor (HENGL, 2009). Minasny et al.(2013), exemplificam que para mapeamento digital de carbono do solo, as escalas atualmente utilizadas incluem a regional, nacional e continental com densidades de amostragem de 0,2 à 110.000 amostras por hectare. Para krigagem, densidades inferiores a 0,3 pt ha⁻¹ tendem a formar pequena quantidade de pares nos primeiros pontos do semivariograma. Isto pode prejudicar a predição pelo interpolador. Destacam-se dentre os interpoladores mais utilizados para geração de MDEs, o Inverse Distance Weighted (IDW), Multiquadratic Radial Basis Function, Thin Plate Spline, Nearest Neighbor, Radial Basis Functions, Krigagem, Spline (SPLINE), Topo to Raster (TTR), entre outros (ARUN, 2013; CHAPLOT et al., 2006; ERDOGAN, 2009; YUE et al., 2007; ZIMMERMAN et al., 1999). Erdogan (2009) indica que não há regra definida para definir o método de interpolação mais

adequado, sendo necessário realizar diversos estudos que analisem a precisão de cada modelo. Este aspecto assume grande importância quando se deseja prever superfícies de diferentes áreas, com diferentes rugosidades.

A validação ou *Hold-out* (ARLOT & CELISSE, 2010; DEVROYE & WAGNER, 1979) ou validação simples, utiliza-se da separação de uma parte do conjunto de dados para verificar a qualidade dos valores finais preditos, a partir deles. A validação cruzada (CV) tem o objetivo de verificar a performance estatística de algoritmos sobre conjuntos de dados (ARLOT & CELISSE, 2010). O método chamado de *leave-one-out* é um caso especial do método de estimação do erro CV (KIM et al., 2012). Este método é realizado separando-se apenas um ponto com o objetivo de estimar o valor do mesmo, a partir do conjunto restante. Este processo é feito recorrentemente com todos os valores e a média geral é estimada. Minasny et al., (2013) descrevem explicitamente que metade dos estudos, ao qual os autores citaram em seu artigo, para mapeamento de carbono do solo, não possuem nenhuma validação. Os mesmos autores salientam que a outra metade dos trabalhos em sua maioria utilizam a validação cruzada e validação interna para determinar a qualidade da predição.

O presente trabalho teve o objetivo principal verificar as melhores formas (quadrada, hexagonal e aleatória), densidade (menos densa a mais densa) e o melhor interpolador (5) para a predição de elevação em duas áreas com diferentes complexidade de relevo. Uma das áreas possui rugosidade elevada, o que se traduziu neste trabalho em maior complexidade de relevo e a outra com rugosidade menor (menor complexidade), situadas na Antártica Marítima. A melhor forma diz respeito a três padrões, sendo dois regulares (hexagonal e quadrado) e um aleatório. Outro objetivo do presente trabalho foi verificar a qualidade do indicador Root Mean Square Error (RMSE), realizado internamente pela krigagem (validação cruzada – *leave-one-out*) comparado ao mesmo indicador, utilizando todos os pontos de elevação já existentes para a área, doravante chamada de validação externa.

2. Materiais e Métodos

2.1 Áreas de estudo

Para o presente estudo foram selecionadas duas áreas que compreendem porções da Antártica Marítima. As duas áreas possuem tamanhos e rugosidades diferentes. A rugosidade diz respeito a irregularidades que a superfície apresenta. Considerando um modelo digital de elevação, a rugosidade será tanto maior, quanto maior for a diferença dos valores entre células vizinhas. As áreas estudadas compreendem a Península Coppermine (PC) (Figura 1 – A1) e a Península Keller (PK) (Figura 1 – A2).

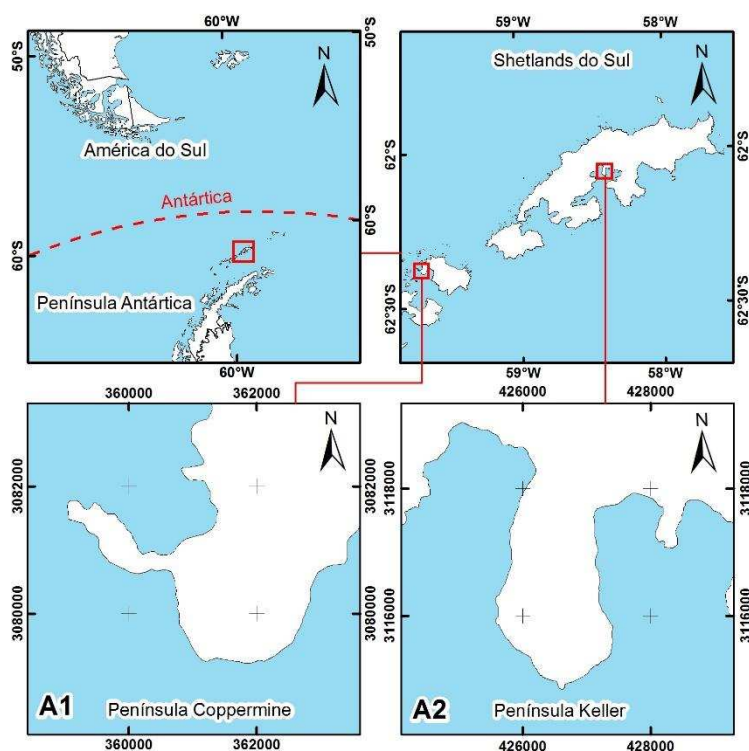


Figura 1 - Localização da Península Coppermine (A1) e Península Keller (A2).

Os locais A1 e A2 possuem áreas de 329,1470 e 452,0771 ha, respectivamente.

2.2 Levantamento de Elevada Resolução Topográfica (ERT)

Para o levantamento das áreas de estudo foi utilizado um equipamento Laser Scanner Terrestre da Riegl Modelo VZ-1000, que possui precisão e acurácia de 0,005 e 0,008 mm, respectivamente, conforme manual do fabricante. Todos os locais onde o equipamento foi instalado foram georreferenciados utilizando o equipamento GNSS da Leica GS80. A técnica de posicionamento utilizada foi a diferencial.

2.1 Processamento

2.1.1 Laser Scanner Terrestre

As nuvens de pontos brutas, obtidas pelos escaneamento das áreas de estudo foram submetidas a filtragem de acordo com a recomendação do fabricante do Laser Scanner Terrestre, no programa computacional Riscan Pro 2.1 64bits. Os principais filtros utilizados foram o filtro automático de vegetação e a remoção visual de artefatos e elementos que não correspondiam a superfície estudada. Salienta-se que para estas áreas a vegetação existente é rasteira, sendo portanto o filtro de vegetação, uma adaptação para remover artefatos, e não a vegetação propriamente dita.

A partir das nuvens de dados filtradas das duas áreas de estudo, foram gerados MDEs (Figura 2) com tamanho de célula de um (1) metro no programa computacional ArcGIS 10.1. Estes MDEs no contexto deste trabalho, são denominados MDEs de referência. Para a generalização e geração dos mesmos foram utilizadas a média e assinatura de célula com o interpolador *Natural Neighbors*.

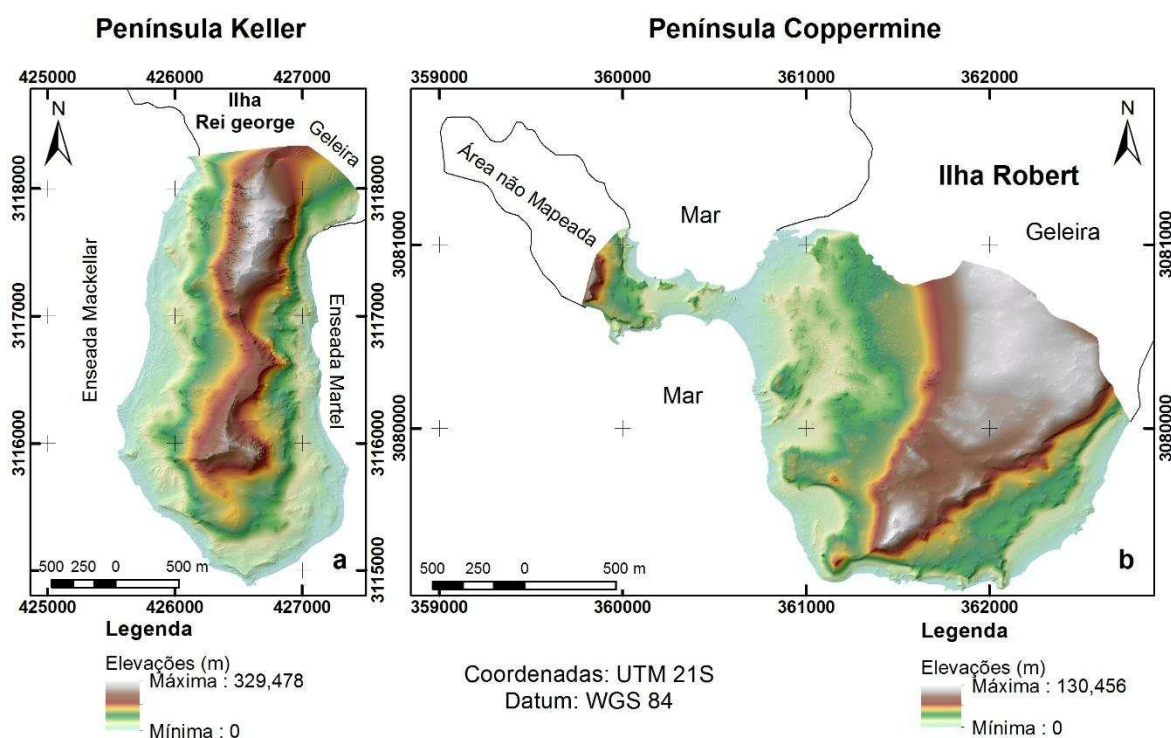


Figura 2 - MDEs das Penínsulas Keller (a) e Coppermine (b) com tamanho de célula de um (1) m.

2.1.2 Grades Amostrais

Para a verificação da exatidão de cada interpolador foram geradas oito grades amostrais, nos padrões hexagonal, quadrado e aleatório com densidades de 0,304; 0,608; 0,911; 1,215; 1,519; 2,279; 3,038 e 4,557 pontos por hectare (pt ha⁻¹). Para cada grade foram geradas 10 repetições, sendo que para as grades regulares foi aleatorizada a coordenada do ponto de origem da grade (Figuras 3 e 4). Este processo garantiu que cada repetição das grades regulares, obtivesse valores diferentes de elevação, independentes. Os valores de elevação foram extraídos dos MDEs de cada uma das áreas estudadas para cada ponto da grade amostral gerada. Para este fim um script foi criado no programa computacional R, tanto para a geração das grades quanto para extração dos valores de altitude, nos pontos amostrais, em todas as áreas de estudo.

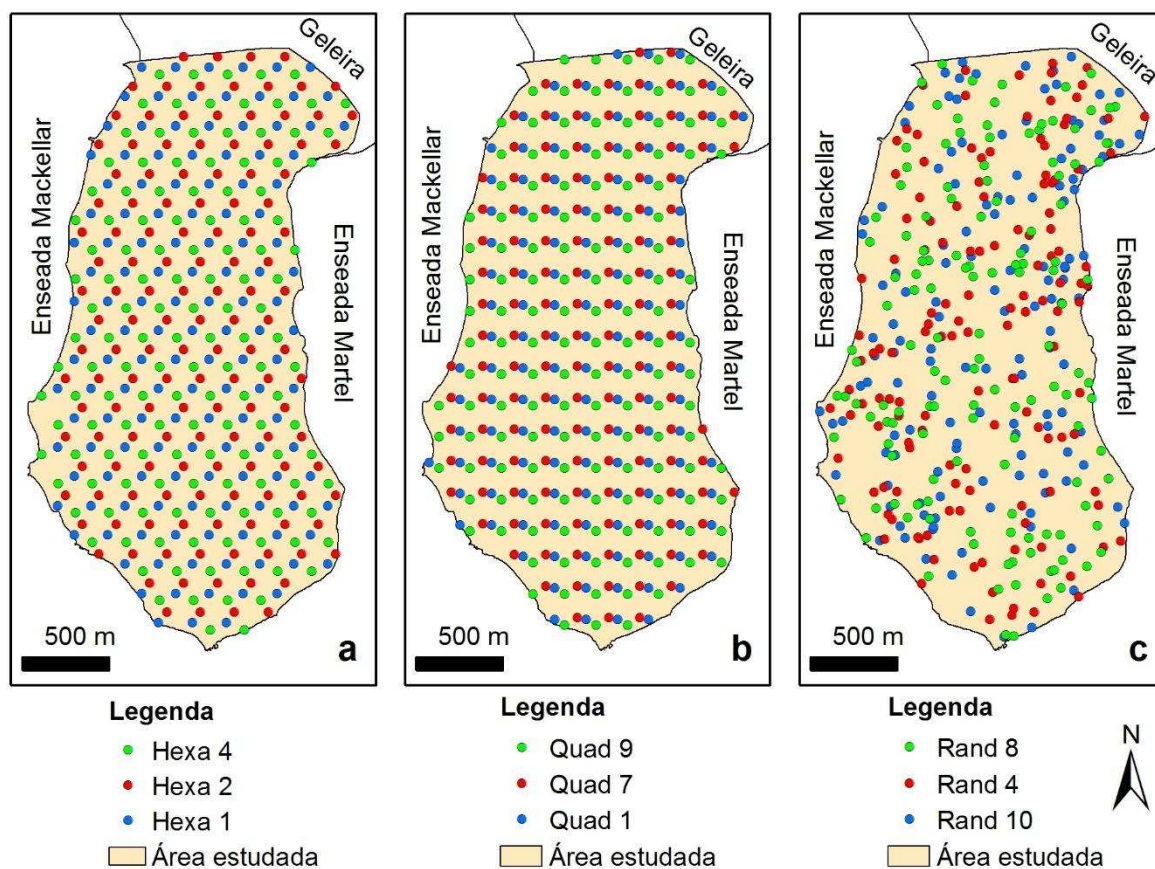


Figura 3 - Representação das grades de amostragem hexagonal (a), quadrada (b) e aleatória (c) exemplificando a diferença na posição de três repetições com densidade amostral de $0,304 \text{ pt ha}^{-1}$, para a PK.

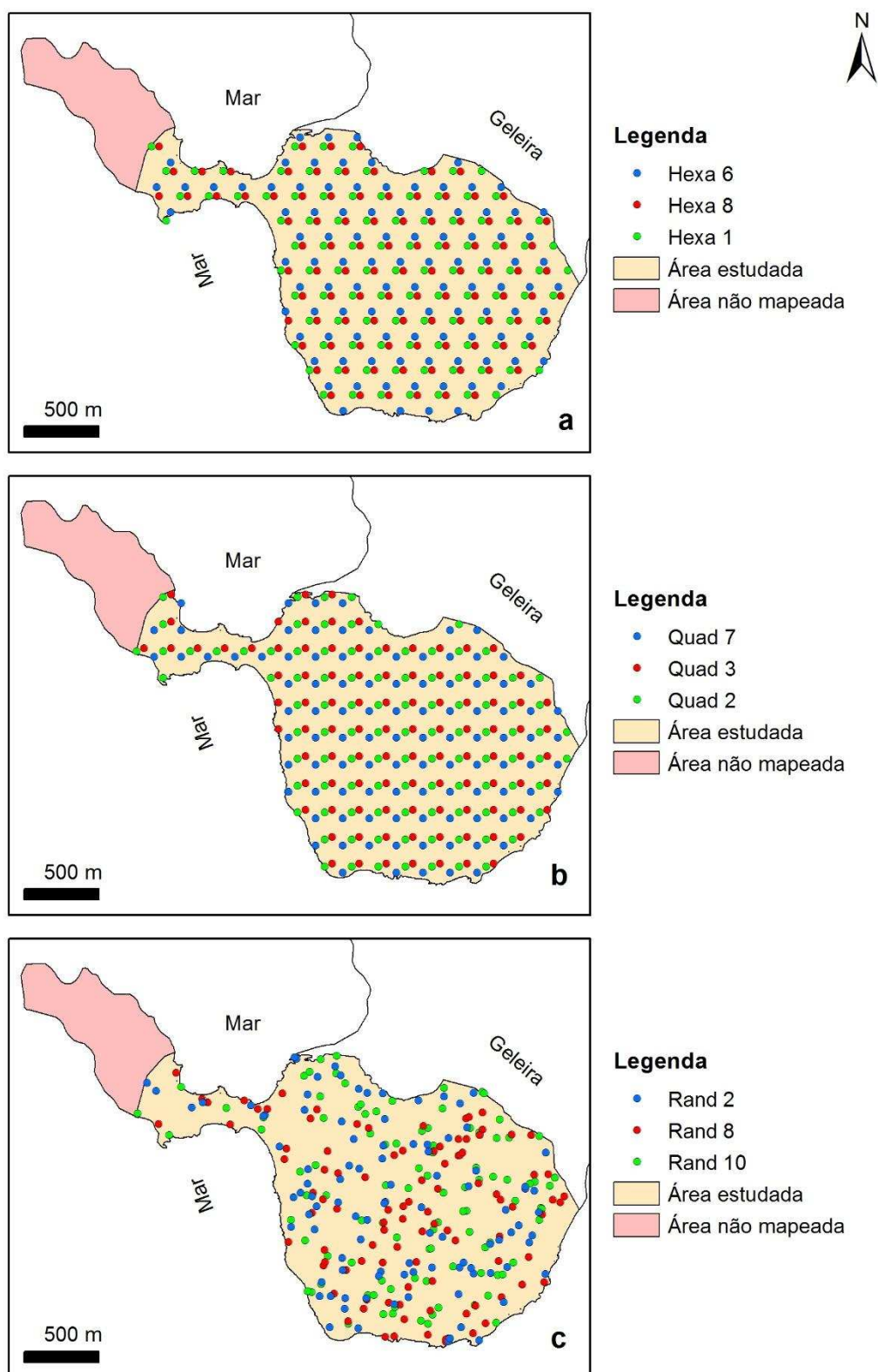


Figura 4 - Representação das grades de amostragem hexagonal (a), quadrada (b) e

aleatória (c) exemplificando a diferença na posição de três repetições com densidade amostral de 0,304 pt ha⁻¹, para a PC.

2.1.3 Interpoladores

No presente estudo foram utilizados os interpoladores IDW (inverso da distância ponderada), Spline (SPLINE), Topo to Raster (TTR), Krigagem Universal (KU) e Krigagem Simples (KS) para as duas áreas de estudo. Os interpoladores IDW, Spline e Topo to Raster foram processados no programa computacional ArcGIS 10.1 a partir de modelos automatizados (*model builder*) para as duas áreas e todas as grades geradas. As krigagens foram realizadas a partir de um script elaborado para o programa computacional R, utilizando como base o pacote *automap* (HIEMSTRA et al., 2009). Todos os MDEs preditos, foram comparados célula a célula com o MDE de referência. A partir destas comparações foram determinados os erros, com relação a referência. A partir dos erros foi determinado o RMSE. Foram calculadas as médias das 10 repetições para o parâmetro anteriormente citado. O RMSE foi obtido de duas formas para as áreas de estudo, tendo como base a equação (WILLMOTT& MATSUURA, 2005) :

$$RMSE = \left[n^{-1} \sum_{i=1}^n |e_i|^2 \right]^{1/2}, \quad (1)$$

onde

$RMSE$ = raiz quadrada do erro médio quadrático, em m;

e_i = cada diferença entre os valores de elevação preditos e os valores originais do MDE de referência, em m.

Para os tratamentos Krigagem Universal e Krigagem Simples foi feita a determinação do RMSE. Foram calculados dois valores de RMSE, a saber: modelo ou validação interna (RMSE.mod) e RMSE da validação ou validação externa (RMSE.val). O RMSE.mod, foi obtido pelo algoritmo de krigagem, através da validação cruzada (leave-one-out). Já o RMSE.val, foi obtido comparando o MDE original e o MDE gerado pelas krigagens, comparando todas as células dos mesmos.

Os valores dos RMSE, aqui denominados de RMSE.mod e RMSE.val foram comparados, para verificar o quão confiável é o RMSE.mod, como indicador para a seleção dos melhores modelos e para avaliar a qualidade dos valores preditos. Neste caso, considerou-se o RMSE.val, como valor base para a comparação.

Para os interpoladores IDW, Spline e TTR, o RMSE foi obtido apenas comparando o modelo predito e o modelo original, para todas as células de cada uma das áreas de trabalho. A Tabela 1 apresenta um resumo dos processamentos realizados.

Tabela 1 - Resumo dos processamentos realizados

MDE	Tamanho da Célula	Número de Células	Padrão de amostragem	Rep.	Dens.	Interpol.	Total
Keller	1 m	4.520.768	Hexagonal	10	8	5	400
			Quadrado	10	8	5	400
			Randômico	10	8	5	400
Coppermine	1 m	3.289.265	Hexagonal	10	8	5	400
			Quadrado	10	8	5	400
			Randômico	10	8	5	400
SOMA				60	48	30	2400

Legenda: Rep. = repetições (nº), Dens. = densidades (nº), Interpol. = interpoladores (nº).

Como critério, para a determinação da melhor densidade amostral, utilizou-se como linha de corte o valor de RMSE.val correspondente a 95% (OLEA, 1999) da diferença entre o valor máximo e mínimo deste parâmetro, obtido para cada área.

2.1.4 Geração de curvas de nível

Para a realização de uma análise qualitativa das superfícies produzidas, foram selecionadas duas grades amostrais no padrão hexagonal, para as duas áreas de estudo. Uma das grades foi a de menor e a outra a de maior densidade amostral. A partir dos rasters preditivos obtidos por estas grades, para cada interpolador, foram geradas as curvas de nível, com equidistância de 10 m.

2.1.5 Tempo de processamento

Com o objetivo de verificar a velocidade de processamento, o tempo em que

cada interpolador levou para realizar a predição do fenômeno estudado foi computado.

3. Resultados

As quantidades de pontos, resultante da adequação das densidades amostrais para cada área e as distâncias mínimas para cada padrão amostral são apresentadas na Tabela 2:

Tabela 2 - Menores distâncias entre os pontos em função da densidade de pontos

Local	Medidas	Padrão Amostral	Densidades (pontos ha ⁻¹)							
			0,30	0,61	0,91	1,22	1,52	2,28	3,04	4,56
Keller	Pontos		137	275	412	549	687	1030	1373	2060
	Menor Distância (m)	Hexagonal	194,86	137,49	112,42	97,43	87,14	71,17	61,64	50,33
		Quadrada	181,63	128,27	104,77	90,75	81,10	66,25	57,38	46,84
Coppermine	Menor Distância (m)	Aleatória	9,38	7,00	4,77	12,00	4,34	5,83	8,09	2,89
		Pontos	100	200	300	400	500	750	1000	1500
		Hexagonal	193,62	137,56	112,50	97,33	87,12	71,14	61,64	50,33
		Quadrada	181,41	128,27	104,74	90,71	81,13	66,24	57,37	46,84
		Aleatória	13,96	11,41	8,37	6,00	4,47	7,26	5,03	4,10

Observa-se que o padrão que obteve a menor distância entre os pontos foi o aleatório, justamente pela ausência da regularidade. Estas pequenas distâncias são independentes da densidade da grade amostral.

3.1 Interpoladores

Os resultados dos interpoladores mostraram diferentes comportamentos entre as duas áreas estudadas.

Os semivariogramas obtidos pelas krigagens, para cada área não se diferenciaram significativamente com o aumento das densidades, como pode ser exemplificado para o caso da KS, nas Figuras 5 e 6.

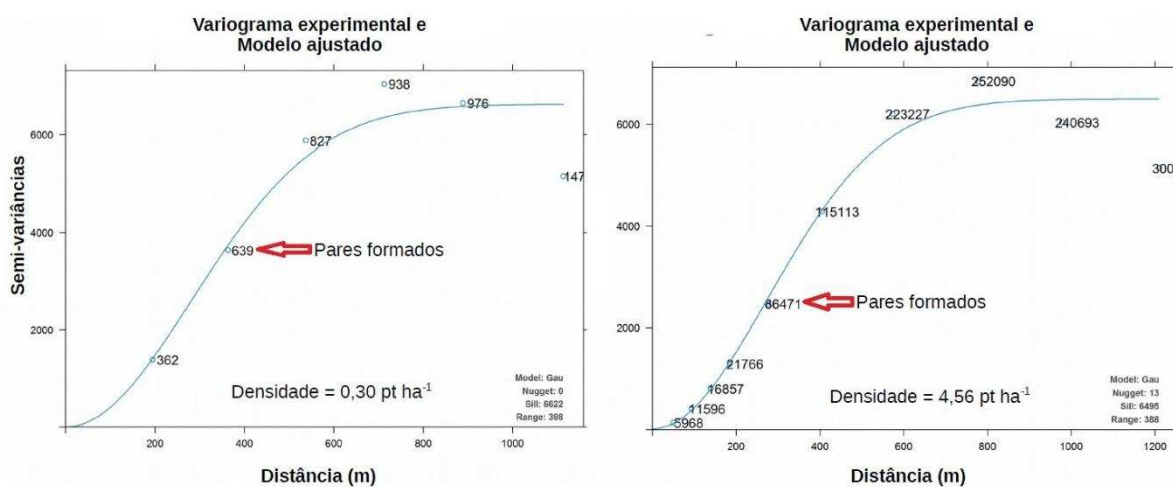


Figura 5 - Semivariogramas obtidos para maior e menor densidade para a PK.

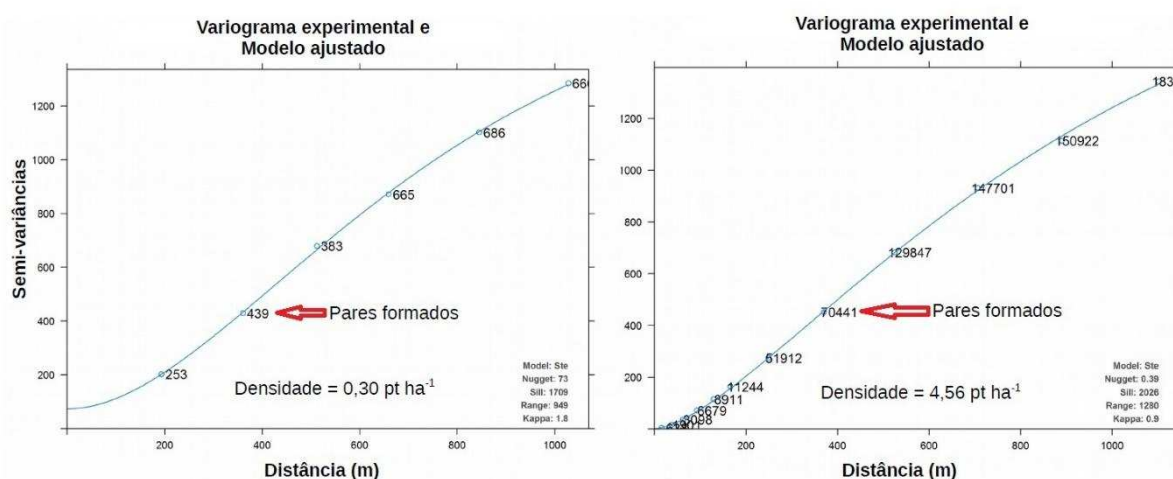


Figura 6 - Semivariogramas obtidos para maior e menor densidade para a PC.

Para a PK a linha de corte que representa 95% de mudança nos valores de RMSE.val foi de 4,31 m, e para a PC foi de 2,32.

Observou-se que para a PK ocorreram problemas nos modelos preditos pelas krigagens em algumas repetições de diferentes densidades e tipos de grade. Este comportamento gerou valores anômalos, notoriamente observados nas bordas da área da península, como pode ser observado na Figura 7 com o padrão hexagonal com 412 pontos. Em outra amostragem, com 275 pontos e no padrão quadrado, na mesma figura, os valores anômalos não se restringiram apenas as bordas da península. Na PC, em nenhuma das krigagens esse comportamento foi observado.

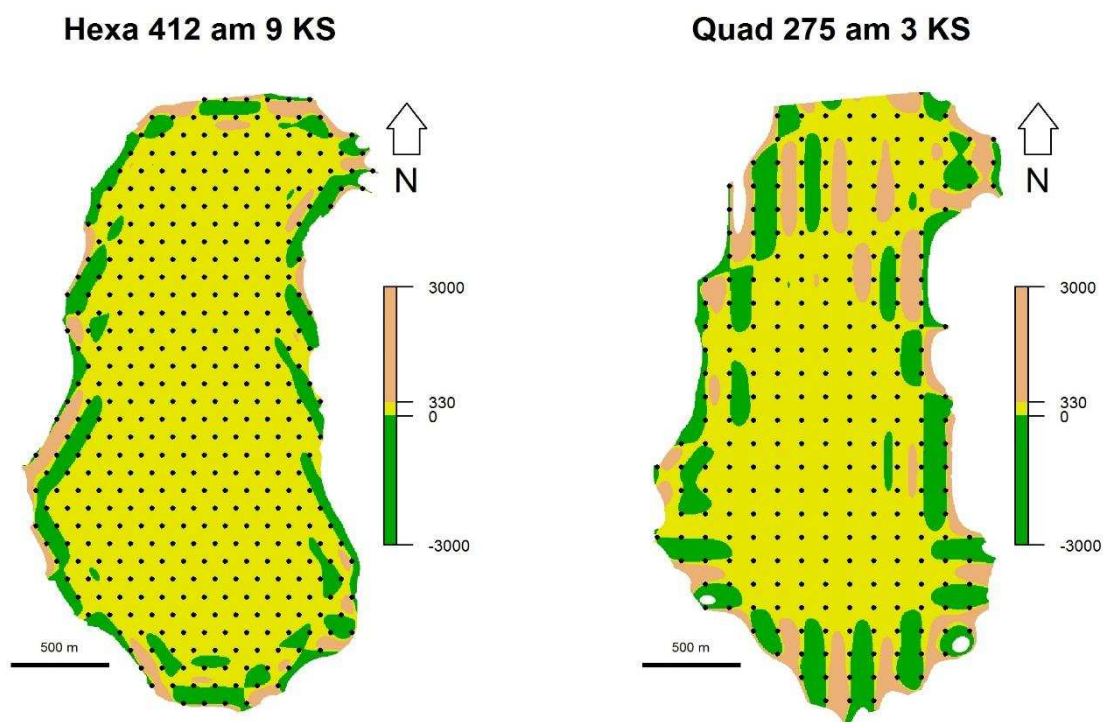


Figura 7 - Mapas representando as elevações anômalas previstas pela KS para uma grade hexagonal (Hexa) com 412 pontos (densidade $0,91 \text{ pt ha}^{-1}$) e para uma grade quadrada (Quad) com 275 pontos (densidade $0,61 \text{ pt ha}^{-1}$).

A Figura 7, exemplifica os valores anômalos representados nos intervalos maior que 330 m e menor que 0 m. Estes valores representam o valor máximo aproximado de elevação obtido a partir do MDE original e mínimo do mesmo, respectivamente. Observa-se que os valores previstos fora deste intervalo, em alguns casos são muitas vezes superiores aos máximos e mínimos, existentes para esta área. Este comportamento de valores anômalos, se repetiu para a krigagem universal, para esta área e nas mesmas grades e densidades.

Foi necessária a eliminação das repetições e amostras que continham valores espúrios, para PK, para a análise dos dados e confecção dos gráficos. As grades, cujos valores de RMSE.val ficaram acima de 50, foram eliminados e considerados como valores anômalos. Entre as amostras eliminadas estiveram 7 com densidade de $0,30 \text{ pt ha}^{-1}$ com o padrão aleatório, 26 com densidade de $0,61 \text{ pt ha}^{-1}$ sendo destas 14 no padrão hexagonal, sete no padrão quadrado e 5 no padrão aleatório. A terceira densidade em que foi necessário remover os valores anômalos foi a de $0,91$

pt ha⁻¹ com seis repetições eliminadas, sendo cinco no padrão hexagonal e uma no padrão quadrado. A quarta densidade em que foram eliminados estes valores foi a de 1,22 pt ha⁻¹ com cinco repetições eliminadas, dentre elas, três com padrão hexagonal e duas com padrão quadrado. Por fim, a última densidade que apresentou valores de RMSE acima da linha de corte determinada foi a de densidade 1,22 pt ha⁻¹ com sete amostras eliminadas, sendo cinco no padrão quadrado e duas no padrão hexagonal.

O RMSE.val para a PK obtido considerando-se as grades estudadas para todos os interpoladores é apresentado na Figura 8.

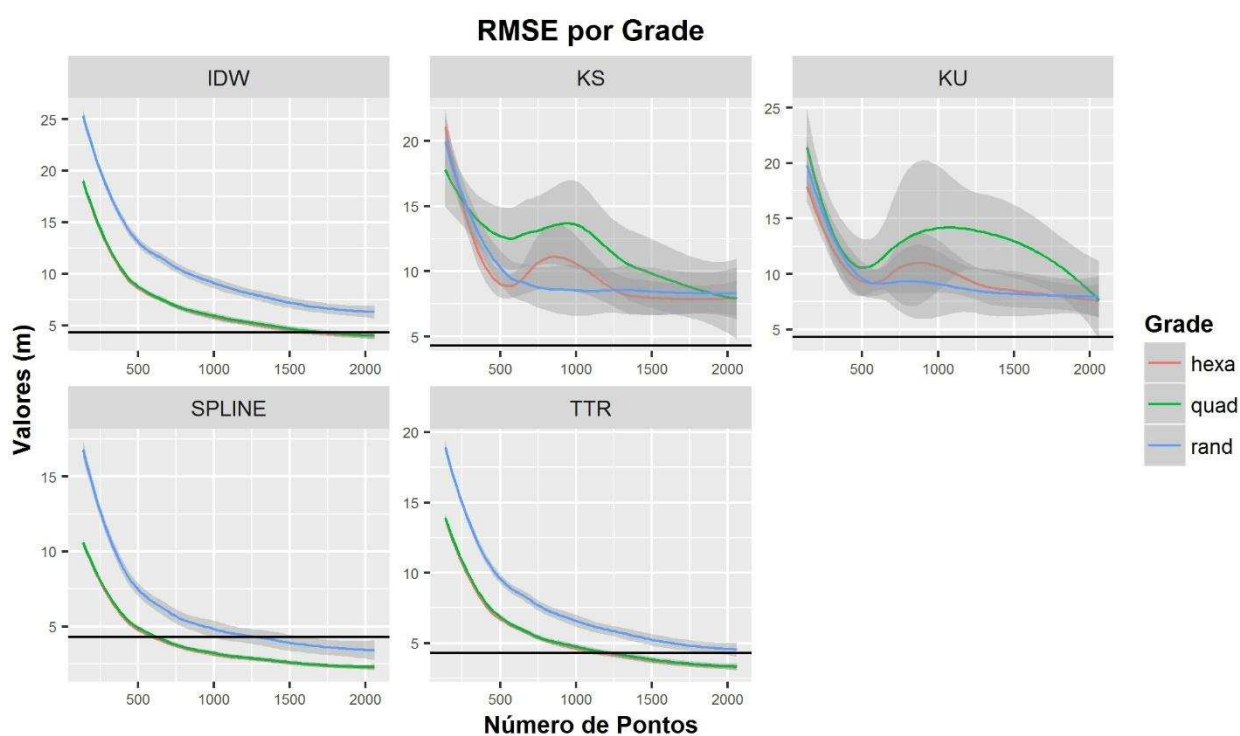


Figura 8 - Gráficos representando os resultados do RMSE.val para a PK por grade para cada interpolador utilizado. Legenda: hexa = hexagonal; quad = quadrado; rand = aleatório; IDW = Inverso da distância ponderado; KS = krigagem simples; KU = krigagem universal; TTR = topo to raster; Linha preta = 95% do Δ RMSE.

O comportamento dos interpoladores não geoestatísticos ocorreu de forma semelhante. O padrão amostral, para estes interpoladores, com melhores resultados de validação, foram sempre melhores para as grades regulares (Figura 8 – IDW, Spline e TTR). Já, para o caso da PK, os interpoladores geoestatísticos, tiveram

seus melhores resultados com o padrão de amostragem aleatório.

Outra representação permite visualizar a relação entre os interpoladores, como apresentado na Figura 9.

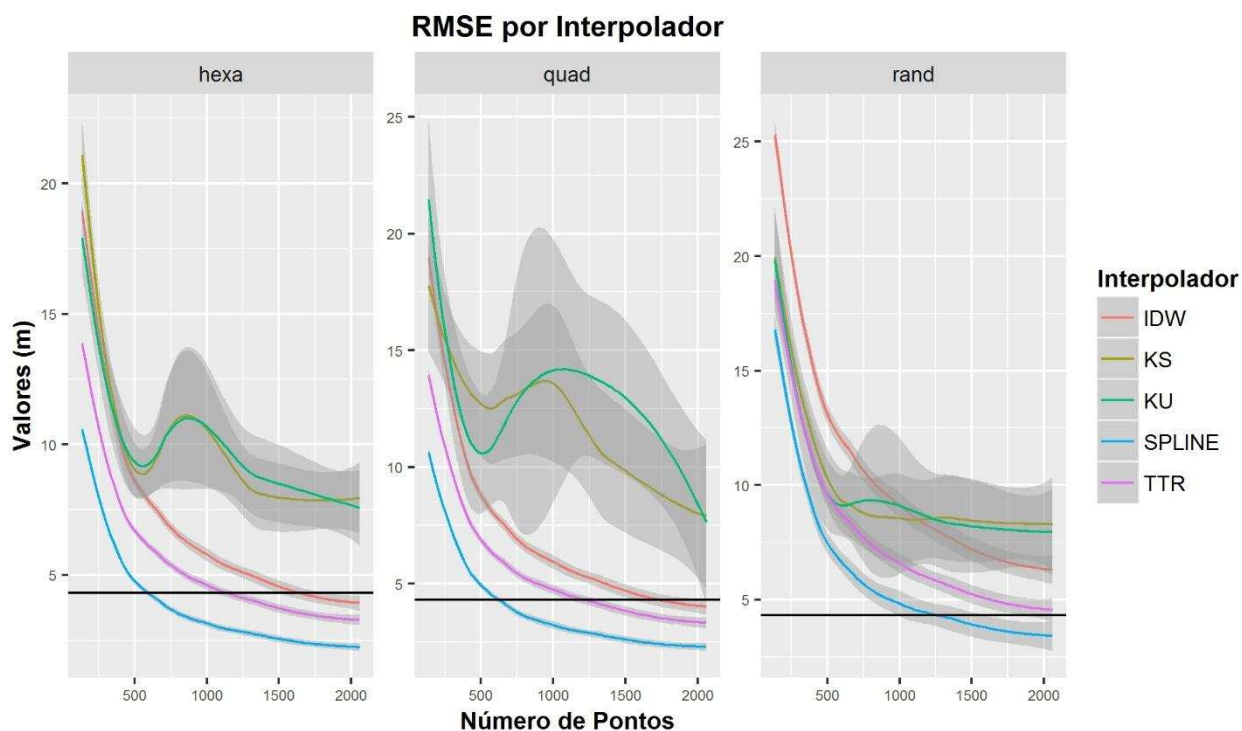


Figura 9 - Gráficos representando os resultados do RMSE.val para a PK por interpolador para cada grade testada. Legenda: hexa = hexagonal; quad = quadrado; rand = aleatório; IDW = Inverso da distância ponderado; KS = krigagem simples; KU = krigagem universal; TTR = topo to raster; Linha preta = 95% do Δ RMSE.

As krigagens não mostraram diferenças visuais entre si, o que pode ser observado nos gráficos da Figura 9, tendo resultados muito semelhantes, para esta área.

A representação geral do RMSE.val para a PK é observada na Figura 10.

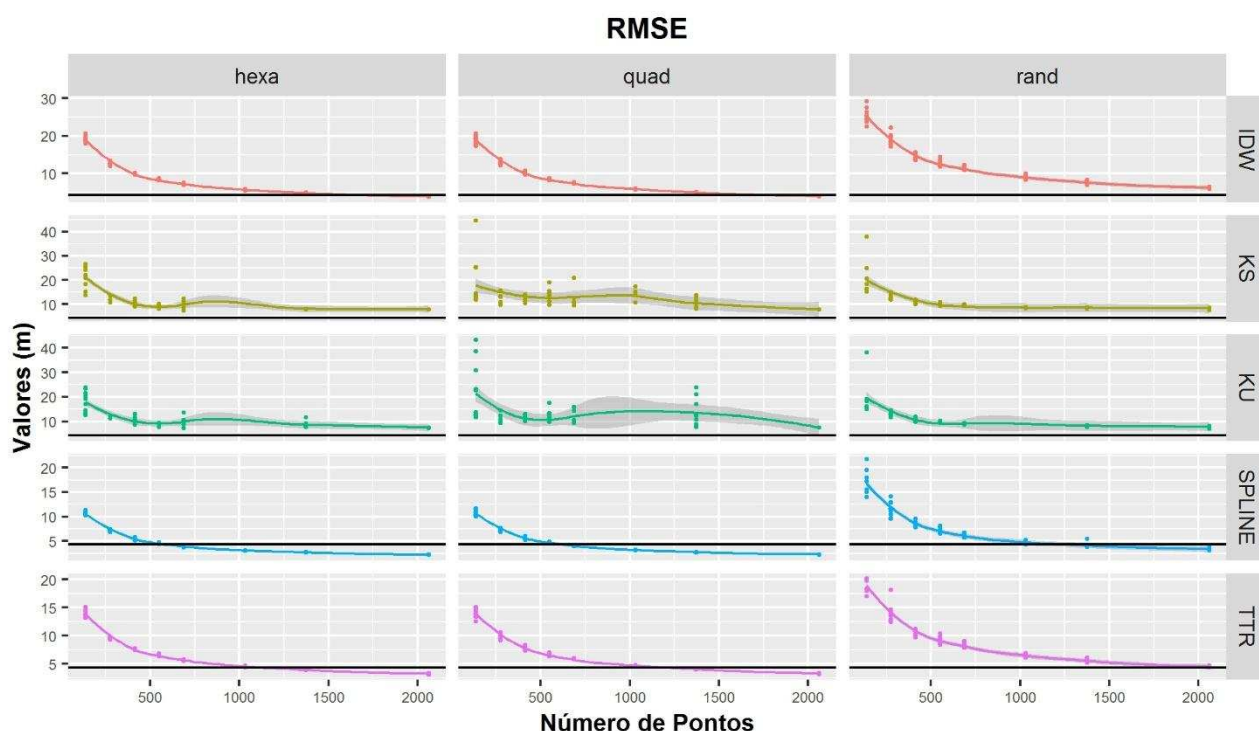


Figura 10 - Gráficos representando os resultados do RMSE.val para a PK com os pontos amostrados. IDW = inverso da distância ponderado; Legenda: hexa = hexagonal; quad = quadrado; rand = aleatório; IDW = Inverso da distância ponderado; KS = krigagem simples; KU = krigagem universal; TTR = topo to raster; Linha preta = 95% do Δ RMSE.

A diferença nos resultados entre o RMSE.val por grade para as duas áreas estudadas é notável, quando compara-se os comportamentos dos gráficos nas Figuras 8 e 11. Enquanto que para a área da PK o interpolador que obteve o melhor RMSE.val com menor densidade de amostragem foi o SPLINE (Figura 8), para a PC o interpolador que obteve o melhor resultado, considerando este requisito foi a KS com uma densidade amostral de aproximadamente $2,43 \text{ pt ha}^{-1}$, comparado aos $1,11 \text{ pt ha}^{-1}$ para a PK.

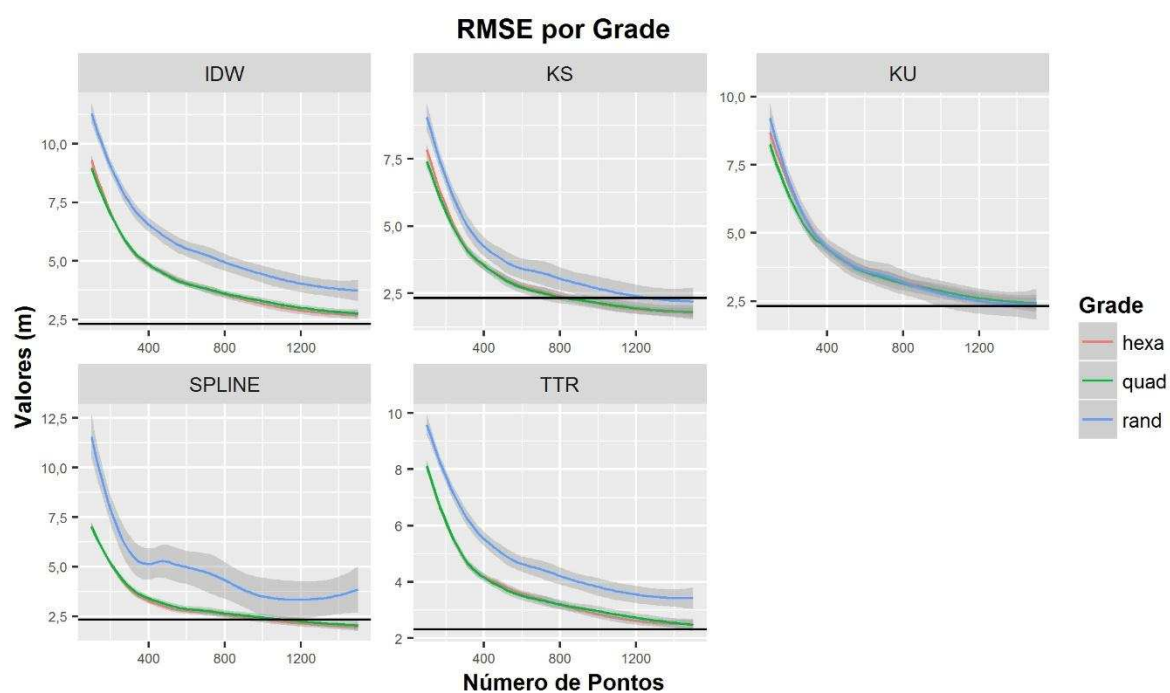


Figura 11 - Gráficos representando os resultados do RMSE.val para a PC por grade para cada interpolador utilizado. Legenda: hexa = hexagonal; quad = quadrado; rand = aleatório; IDW = Inverso da distância ponderado; KS = krigagem simples; KU = krigagem universal; TTR = topo to raster; Linha preta = 95% do Δ RMSE.

Da mesma forma as diferenças são grandes, principalmente quando comparadas as krigagens, considerando o RMSE.val por interpoladores, representados nas Figuras 9 e 12.

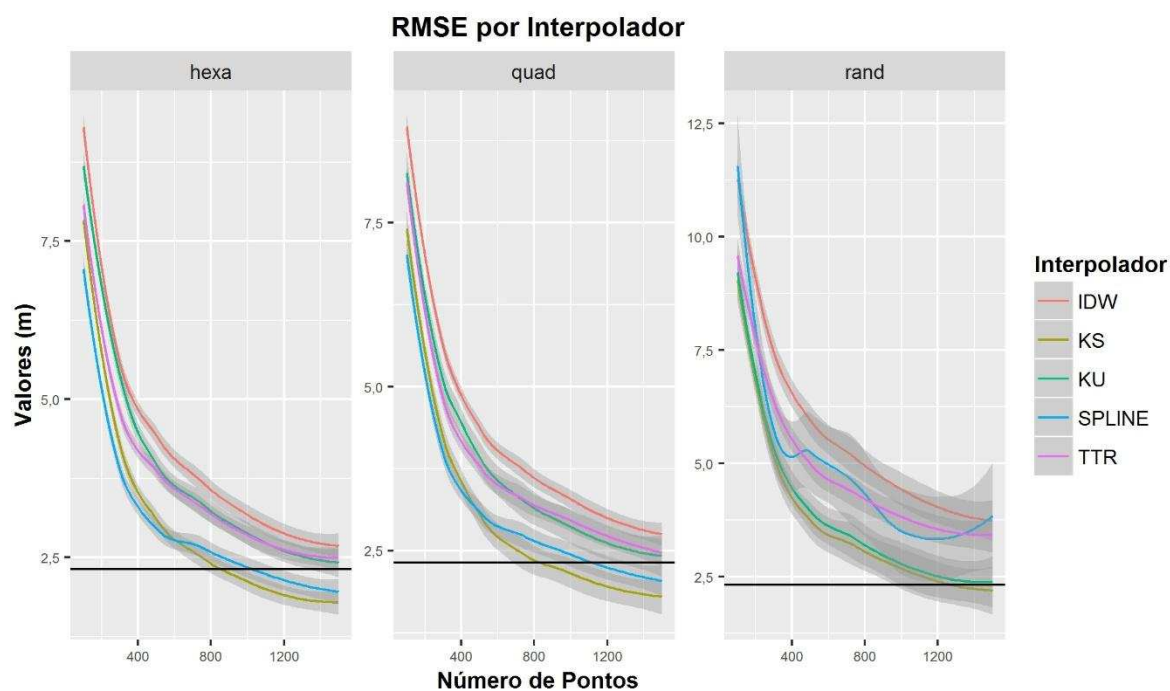


Figura 12 - Gráficos representando os resultados do RMSE.val para a PC por interpolador para cada grade testada. Legenda: hexa = hexagonal; quad = quadrado; rand = aleatório; IDW = Inverso da distância ponderado; KS = krigagem simples; KU = krigagem universal; TTR = topo to raster; Linha preta = 95% do Δ RMSE.

Uma visão geral das diferenças entre os RMSE.val das duas áreas em estudo pode ser obtida comparando as Figuras 10 e 13.

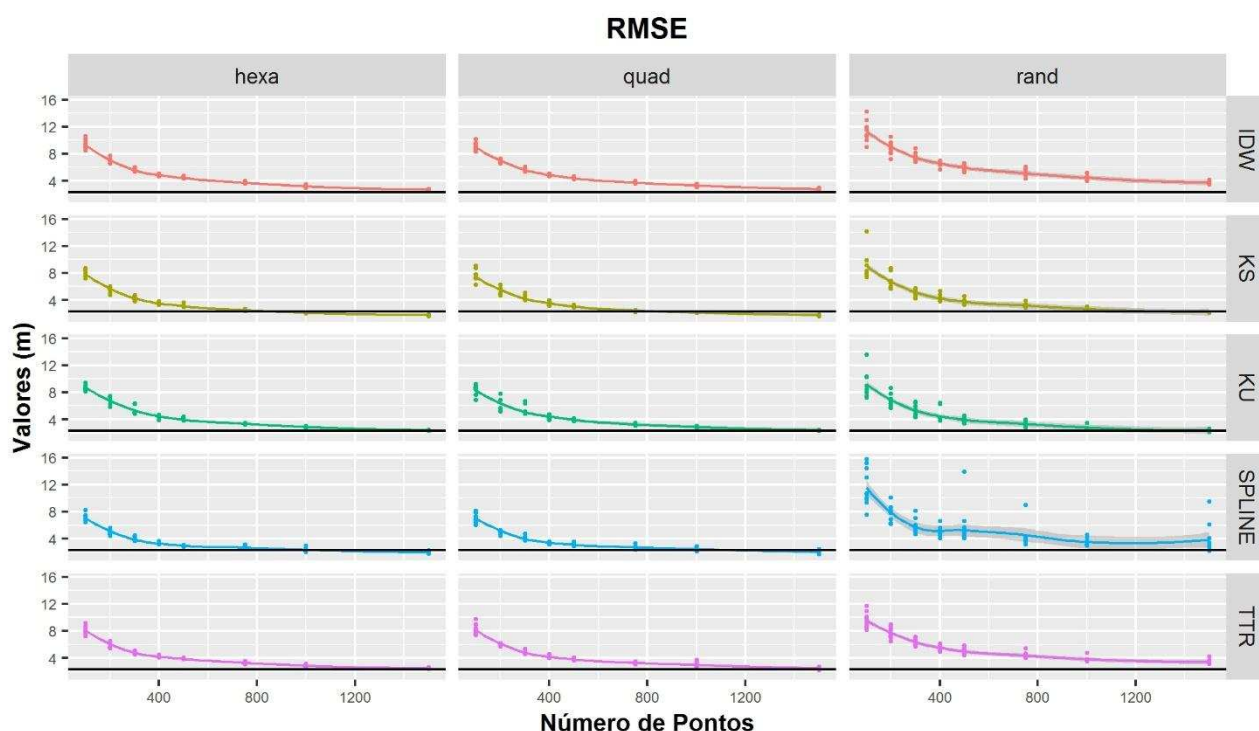


Figura 13 - Gráficos representando os resultados do RMSE.val para a PC. IDW = inverso da distância ponderado; Legenda: hexa = hexagonal; quad = quadrado; rand = aleatório; Linha preta = 95% do Δ RMSE.

De uma forma geral, observou-se que a grade com padrão aleatório, obteve valores de RMSE.val superiores quando comparados os interpoladores. No caso da KS para a PC os gráficos se mantiveram semelhantes, quando considerados os três padrões de amostragem nos padrões regulares (Figura 13 – hexa e quad).

3.2 Validação interna versus externa

A comparação entre o RMSE.mod e o RMSE.val, permite verificar se a validação obtida a partir da krigagem, é otimista ou pessimista em relação aos valores reais da população do fenômeno estudado.

Os gráficos das Figura 14 e 15, permitem visualizar as diferenças obtidas considerando este indicador, para as duas krigagens estudadas e as três grades nas duas áreas estudadas.

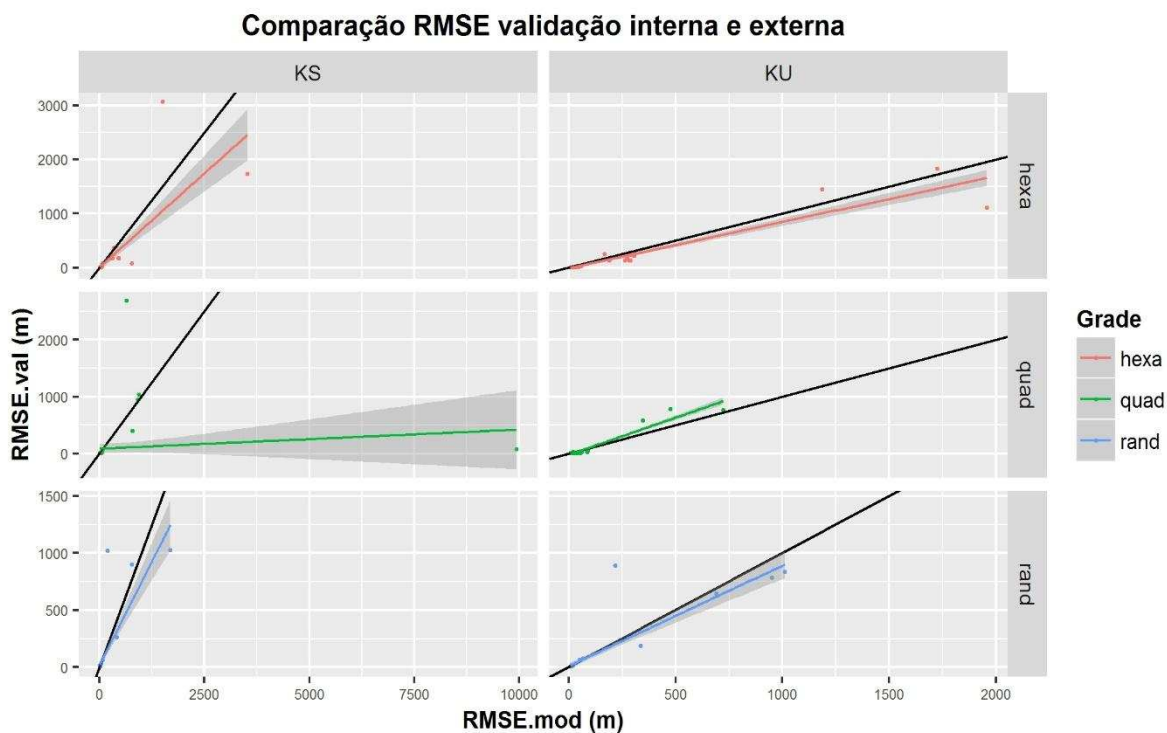


Figura 14 - Gráficos representando os resultados da comparação entre os RMSEs da validação interna e externa para a PK. KS = krigagem simples; KU = krigagem universal; hexa = hexagonal; quad = quadrado; rand = aleatório; Linha preta = reta de intercepto zero e ângulo de inclinação de 45°.

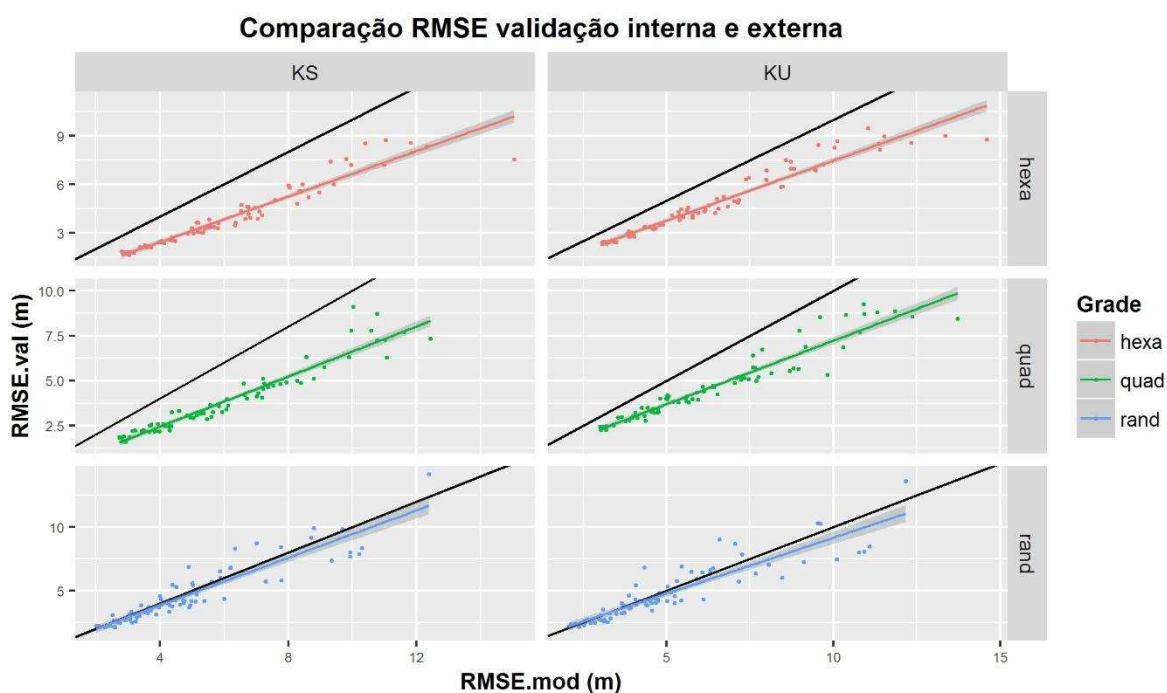


Figura 15 - Gráficos representando os resultados da comparação entre os RMSEs da validação interna e externa para a PC. KS = krigagem simples; KU = krigagem universal; hexa = hexagonal; quad = quadrado; rand = aleatório; Linha preta = reta de intercepto zero e ângulo de inclinação de 45°.

Os resultados da comparação entre a validação interna (RMSE.mod) e validação externa (RMSE.val), mostram de uma forma geral, que as grades amostrais com padrão aleatório, possuem menores diferenças entre as mesmas. Isto é observado pela equação linear para este padrão estar muito próxima a reta de intercepto zero e ângulo de inclinação de 45° (linha preta) (Figuras 14 e 15 - rand). Outro detalhe observado é o de que para grades regulares as krigagens são pessimistas ao indicar a qualidade do RMSE através de sua validação interna. Isto é observado nas Figura 14 e 15, onde os ajustes lineares sempre ficaram abaixo da linha com intercepto zero e ângulo de inclinação de 45°.

3.3 Análise qualitativa de resultados

As Figuras 16 e 17 exemplificam a qualidade das superfícies preditas para duas grades. Uma das grades é a de menor densidade (0,30 pt ha⁻¹) e a outra é a de maior densidade (4,56 pt ha⁻¹).

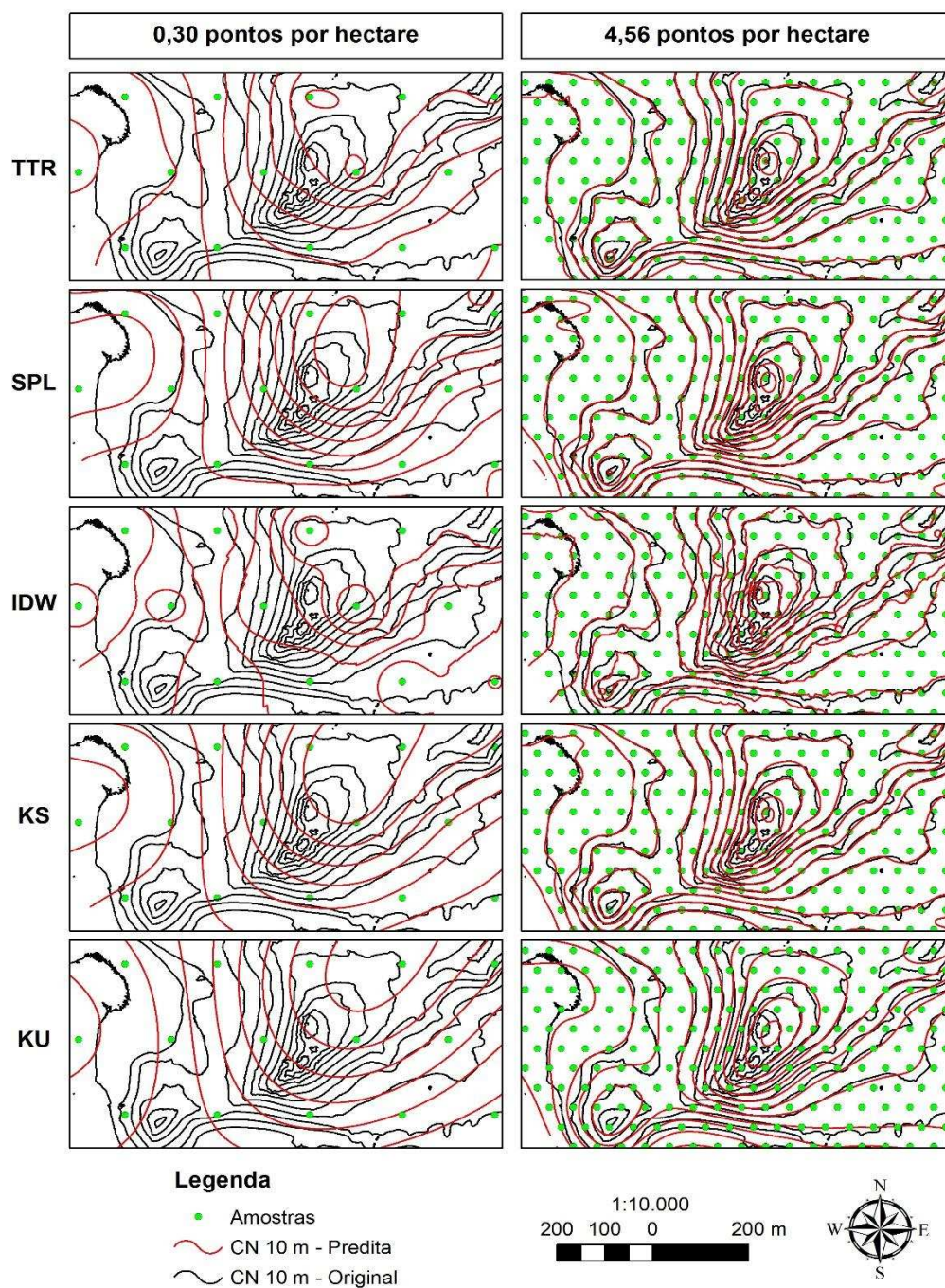


Figura 16 - Curvas de nível com equidistância de 10 m geradas a partir das superfícies preditas, nas maiores e menores densidades para os 5 interpoladores em uma das repetições no padrão hexagonal para a PC. Legenda: CN = curva de nível; SPL = Spline.

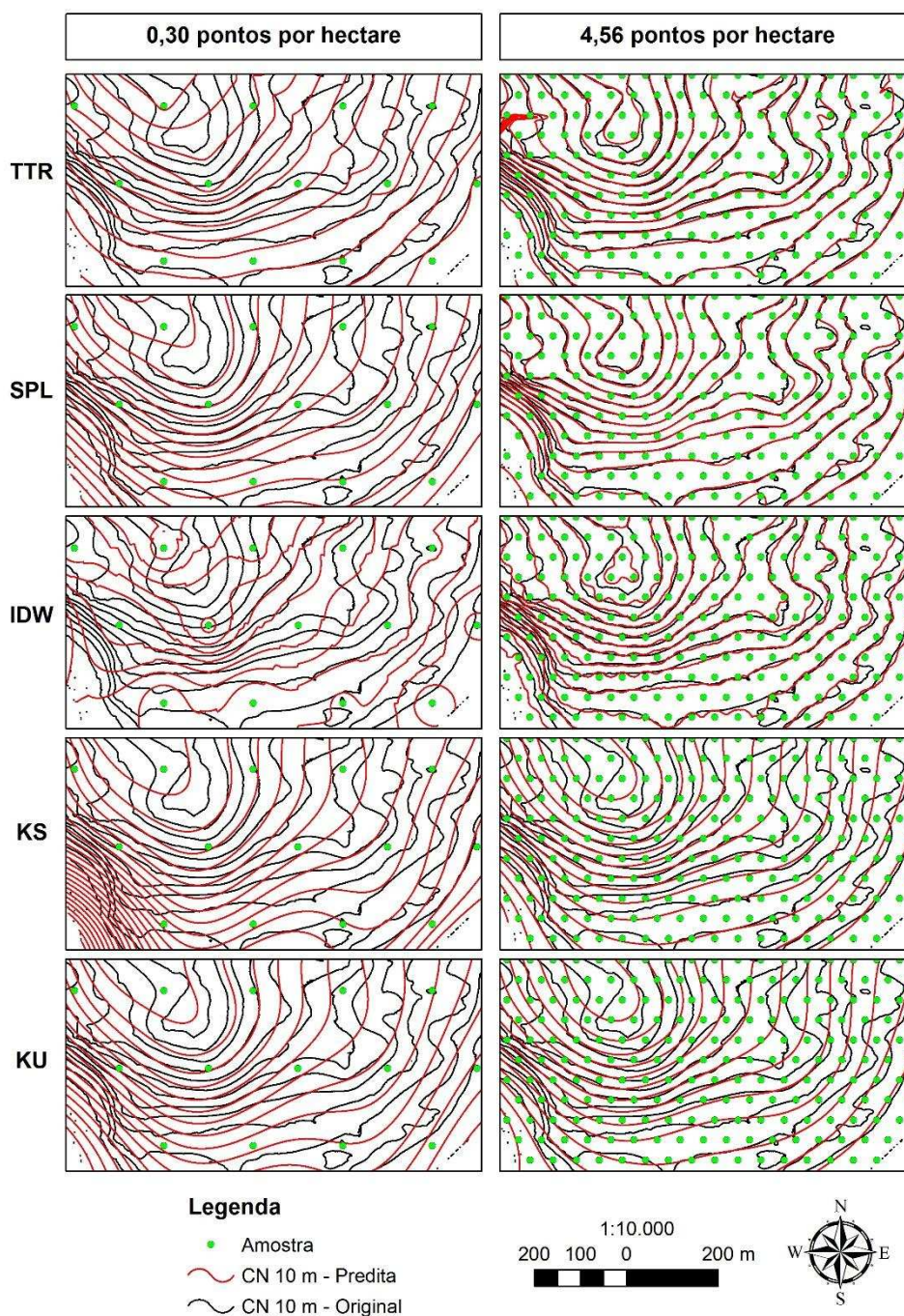


Figura 17 - Curvas de nível com equidistância de 10 m geradas a partir das superfícies preditas, nas maiores e menores densidades para os 5 interpoladores em uma das repetições no padrão hexagonal para a PK. Legenda: CN = curva de nível; SPL = Spline.

A partir das duas figuras anteriores é possível verificar de forma qualitativa os

resultados das superfícies preditas pelos interpoladores nas grades selecionadas. Observa-se com clareza que as grades amostrais de menor densidade produzem superfícies com baixa representatividade das feições do terreno. A análise visual confirma os resultados dos melhores interpoladores obtidos a partir da análise numérica. Desta forma o interpolador que demonstrou melhor concordância das curvas de nível produzidas pelo MDE predito, quando comparada as curvas produzidas pelo MDE original, foi a KS para a PC, além dos interpoladores SPLINE e TTR. Já para a PK a análise qualitativa demonstra as melhores representatividades para os interpoladores SPLINE e TTR.

3.4 Tempo de processamento

Os gráficos com os tempos de processamento das krigagens, para as duas áreas são apresentados na Figura 18. Os tempos de processamento para os interpoladores não geoestatísticos não se alterou em função da densidade amostral, nas duas áreas estudadas. O tempo médio para a PK ficou com aproximadamente 0,25 minutos para o interpolador IDW, 0,50 minutos para o SPLINE e de aproximadamente 3,00 minutos para o TTR. Já para a PC os valores aproximados obtidos foram de 0,20 minutos para IDW, e os outros interpoladores ficaram com os mesmos tempos que os obtidos para a PK.

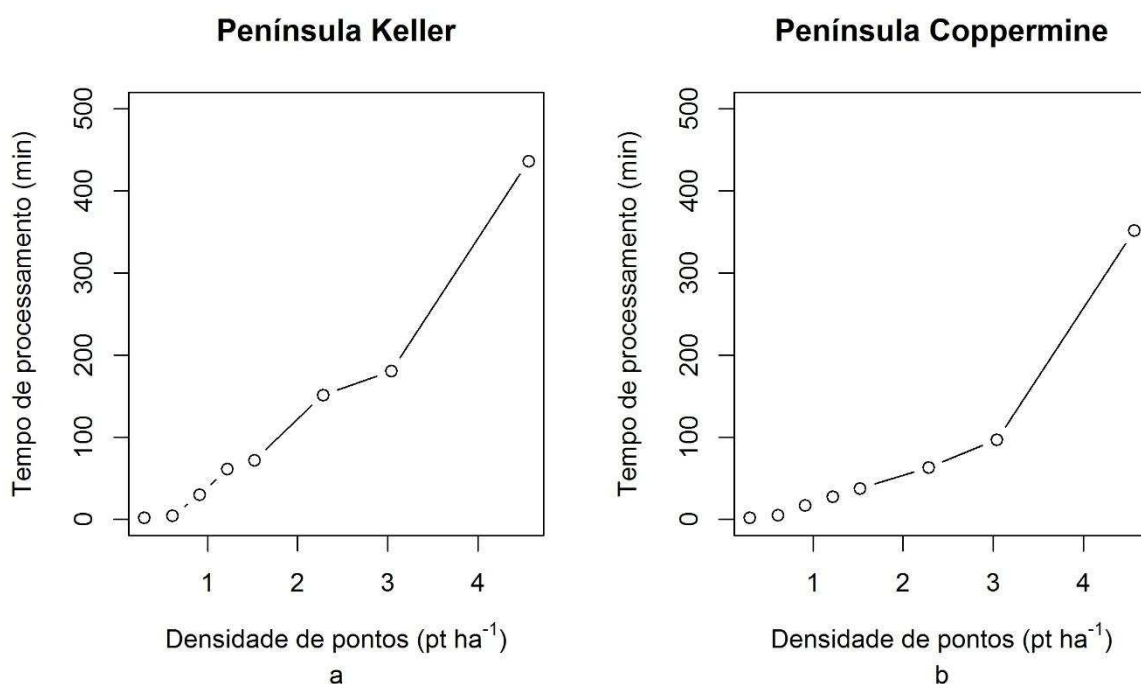


Figura 18 – Tempo de processamento de cada krigagem em função do aumento da densidade.

4. Discussão

A distância entre pontos para a interpolação das krigagens é importante na medida em que grandes distâncias entre pontos tendem a formar poucos pares nos primeiros pontos dos semi-variogramas, o que pode ser um problema para a caracterização dos indicadores da mesma. O indicador pepita por exemplo, pode ser pouco representado, quando se obtém poucos pares nos primeiros pontos. Isso pode alterar a escolha do modelo ajustado, para o fenômeno em estudo.

Existiram muitas diferenças nos resultados obtidos pelas duas áreas de estudo. Estas diferenças indicam que é necessário obter um indicador de complexidade do fenômeno, para determinar a melhor grade e densidade de amostragem. As krigagens, obtiveram piores valores de RMSE.val na área com maior complexidade, quando comparados aos interpoladores não geoestatísticos. Este valor diminuiu à medida que se aumentou a densidade amostral. Para a PC os

interpoladores geostatísticos obtiveram os melhores resultados, independente do padrão amostral e com valores de RMSE.val muito inferiores aos demais interpoladores. Isso reforça a necessidade de análise criteriosa para escolha do padrão e densidade amostral em função da complexidade do fenômeno em estudo. A reamostragem interativa do fenômeno, seria uma solução para diminuir diferenças na complexidade. Porém, este processo, envolveria o aumento dos custos, para levantamento e análise do mesmo. Esta situação é ainda mais importante quando a área de estudo se situa em locais remotos, como no caso em tela, em regiões antárticas. Muitas vezes é impossível reamostrar em um mesmo local em curtos períodos de tempo. Desta forma com reamostragem em intervalos de tempo maiores, o ambiente poderia estar sujeito a alterações, o que prejudicaria a correta caracterização e mapeamento destes locais. Embora a diferença da densidade para os melhores interpoladores, quando considerada a densidade amostral por grade (2,43 pt ha⁻¹ - PC e 1,11 pt ha⁻¹ - PK) indique uma menor densidade amostral para a PK com o interpolador SPLINE, o RMSE.val que foi utilizado como linha de corte para determinar a melhor densidade, foi muito inferior para a PC. Isto indica melhores resultados quando se considera os valores absolutos de RMSE.val para esta área. Este comportamento pode ser explicado pela menor complexidade do relevo da PC, em comparação a da PK. Para a PK as elevações variaram entre 0 e aproximadamente 330 m (Figura 2 - a), em distâncias de aproximadamente 1000 m. Já para a PC as elevações variaram entre 0 e aproximadamente 130 m (Figura 2 - b).

A Figura 11 demonstra a melhor qualidade de grades regulares para a amostragem, quando considera-se a menor amostragem possível. Para a KS, no caso de PC, seria necessário incrementar o número de pontos amostrais de aproximadamente 800 (2,43 pt ha⁻¹) pontos para aproximadamente 1200 (3,65 pt ha⁻¹) pontos para se obter o mesmo valor de RMSE.val. Isto indica um incremento de aproximadamente 50% na densidade amostral, se a decisão for por utilizar uma grade aleatória. Comportamento semelhante pode ser observado no melhor interpolador para a PK. Neste caso o incremento da densidade amostral teria de ser de aproximadamente 99% (1,11 para 2,21 pt ha⁻¹) que corresponderia ao incremento de 500 pontos amostrais para 1000 pontos amostrais, para esta área (Figura 8 -

SPLINE). É possível afirmar para este caso, que grades com padrão regular seriam a melhor opção para áreas com reduzida complexidade do relevo.

A comparação entre as validações interna e externa (Figura 14 e 15), mostra a alta qualidade dos resultados preditos pelas krigagens. Porém, a ocorrência de valores anômalos na área de maior complexidade (PK), demonstra a necessidade da utilização de outros indicadores, para acessar a qualidade dos modelos preditos por interpoladores geoestatísticos. Em situações que não se conhece a população dos valores do fenômeno estudado, a validação interna pode ser um bom indicador, sem superestimar a qualidade das superfícies preditas.

A análise qualitativa indicou que os interpoladores geoestatísticos devem ser melhor estudados para a predição de fenômenos, como o caso em tela, para diferentes complexidades do mesmo. Isto foi evidenciado pelos resultados apresentados para a PK, onde a KS não representou bem a realidade, quando consideradas as curvas de nível (Figura 17 – KS). Esta baixa representatividade ocorreu também na grade com maior densidade amostral. Já para a PC, que possui uma menor complexidade do relevo, os interpoladores geoestatísticos, tiveram uma boa representatividade. A mesma análise, também indicou a baixa representatividade qualitativa, em grades amostrais com baixa densidade ($0,30 \text{ pt ha}^{-1}$), tanto na área mais complexa, quanto na menos complexa.

5. Conclusões

As grades com formas regulares, hexagonal e quadrática, apresentaram os melhores resultados de RMSE para todos os interpoladores.

As melhores densidades amostrais para o mapeamento de elevação, considerando o critério de corte de 95% de mudança nos valores de RMSE da validação externa são a de $1,11 \text{ pt ha}^{-1}$, para áreas com maior complexidade (Península Keller), e de $2,43 \text{ pt ha}^{-1}$, para a área com menor complexidade (Península Coppermine). Porém, o critério de corte, mostra melhores valores do mesmo indicador para áreas de menor complexidade (Península Coppermine).

Considerando o indicador RMSE, as krigagens são pessimistas em relação a predição de elevações. As superfícies preditas pelos mesmos possuem melhor qualidade do que a observada pelo validador interno do método. Ou seja, a comparação com os valores verdadeiros conhecidos e omitidos, antes de realizar a interpolação, são melhores representados, do que a validação interna indica.

Para as krigagens, as grades amostrais com padrões aleatórios, tendem a diminuir os problemas com a ocorrência de valores anômalos. Para os outros interpoladores a grade aleatória apresentou sempre um RMSE mais elevado do que para as grades regulares. Dependendo da grade amostral, padrão geométrico e densidade, podem ocorrer valores anômalos a partir das krigagens em áreas com elevada complexidade. Isto indica que nestas áreas é necessário realizar mais estudos com relação ao tipo de amostragem e valores preditos pelo interpolador.

Referências

ARLOT, S.; CELISSE, A. A survey of cross-validation procedures for model selection *. **Statistics Surveys**, v. 4, p. 40–79, 2010.

ARUN, P. V. A comparative analysis of different DEM interpolation methods. **The Egyptian Journal of Remote Sensing and Space Science**, v. 16, n. 2, p. 133–139, 2013.

BRUS, D. J.; HEUVELINK, G. B. M. Optimization of sample patterns for universal kriging of environmental variables. **Geoderma**, v. 138, n. 1-2, p. 86–95, 2007.

CHAPLOT, V.; DARBOUX, F.; BOURENNANE, H.; LEGUÉDOIS, S.; SILVERA, N.; PHACHOMPON, K. Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density. **Geomorphology**, v. 77, n. 1-2, p. 126–141, 2006.

DE GRUIJTER, J.; BRUS, D.; MARTIN, M. B.; Knotters, M. **Sampling for Natural Resource Monitoring**. v. 103, p. 889-890, 2008, DOI: 10.1198/jasa.2008.s240.

DEVROYE, L. P.; WAGNER, T. J. Distribution-Free Performance Bounds for

Potential Function Rules. **IEEE Transactions on Information Theory**, v. 25, n. 5, p. 601–604, 1979.

ERDOGAN, S. A comparison of interpolation methods for producing digital elevation models at the field scale. **Earth Surface Processes and Landforms**, v. 34, n. 3, p. 366–376, 2009.

HENGL, T. **A Practical guide to Geostatistical Mapping**. P. 271, 2009, DOI: 10.1016/0277-9390(86)90082-8.

HIEMSTRA, P. H.; PEBESMA, E. J.; TWENHÖFEL, C. J. W.; HEUVELINK, G. B. M. Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. **Computers and Geosciences**, v. 35, n. 8, p. 1711–1721, 2009.

KIM, D. R.; ALI, M.; SUR, D.; KHATIB, A.; WIERZBA, T. F. Determining optimal neighborhood size for ecological studies using leave-one-out cross validation. **International Journal of Health Geographics**, 2012.

MINASNY, B.; MCBRATNEY, A. B.; MALONE, B. P.; WHEELER, I. Digital Mapping of Soil Carbon. **Advances in Agronomy**, v. 118, p. 1–47, 2013.

OLEA, R. A. **Geostatistics for Engineers and Earth Scientists**. Boston, MA: Springer US, 1999.

WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. **Climate Research**, v. 30, n. 1, p. 79–82, 2005.

YUE, T.-X.; DU, Z.-P.; SONG, D.-J.; GONG, Y. A new method of surface modeling and its application to DEM construction. **Geomorphology**, v. 91, n. 1-2, p. 161–172, 2007.

ZIMMERMAN, D.; PAVLIK, C.; RUGGLES, A.; ARMSTRONG, M. P. An experimental comparison of ordinary and universal kriging and inverse distance weighting. **Mathematical Geology**, v. 31, n. 4, p. 375–390, 1999.

CAPÍTULO 4

Predição de atributos pedogeoquímicos selecionados através de interpoladores em áreas livres de gelo da Antártica

Resumo

O conhecimento de valores de variáveis ambientais, em locais não amostrados a partir de um conjunto mínimo de amostras da mesma, pode ser realizado através de uma série de interpoladores. A krigagem e o Random Forest são alguns exemplos de preditores com esta capacidade. Este trabalho teve o objetivo de verificar a melhora na predição dos atributos selecionados, com o uso de covariáveis, em solos com áreas livres de gelo comparada a modelos geoestatísticos que não utilizam covariáveis. Para tanto foram coletadas 58 amostras de solo na camada de 0 a 0,10 m de profundidade na Península Keller, Ilha Rei George, Antártica, sendo esta malha densificada com levantamento pré-existente. Foi feita a análise química do solo pelo método gravimétrico e determinados os valores de potássio, cálcio e magnésio, para cada ponto amostrado. De posse do conjunto de pontos amostrais realizou-se a krigagem simples (KS). Com o mesmo conjunto de dados amostrais, além das covariáveis foi utilizado o interpolador Random Forest, para prever os valores dos atributos estudados. Os MDEs de 1, 5, 10, 20 e 30 m de tamanho de célula, foram obtidos a partir de um levantamento com o Laser Scanner Terrestre (TLS) na área de estudo e a partir da nuvem de pontos gerada. O uso de covariáveis, não melhorou a predição de bases para a área estudada. Os mapas finais preditos não possuíam grandes diferenças em seus RMSEs, porém com diferenças visuais. O uso de mapas com melhor resolução, não melhoram a predição usando o Random Forest, para Ca^{2+} . Áreas com elevada complexidade, merecem mais estudos em seus atributos químicos, considerando a produção de mapas preditivos.

Palavras chave: Covariáveis preditivas, interpolação, mapeamento digital.

Pedogeochemical attributes prediction selected throw interpolators in Antarctic ice free areas

Abstract

The knowledge of environmental variables values, in places not sampled from a minimum set of samples of the same, can be made through a series of interpolators. Kriging and Random Forest are some examples of predictors with this capability. The objective of this work was to verify the improvement in the prediction of the selected attributes, with the use of covariates, in soils with ice - free areas compared to geostatistical models that do not use covariates. For this, 58 soil samples were collected in the layer of 0 to 0.10 meters deep on the Keller Peninsula, King George Island, Antarctica, and the mesh densified with a pre-existent survey. The chemical analysis of the soil was performed by gravimetric method and the values of potassium, calcium and magnesium were determined for each point sampled. From the sample points set, simple kriging (SK) was performed. With the same sample dataset, besides the covariates, the Random Forest interpolator was used to predict the values of the studied attributes. The DEMs of 1, 5, 10, 20 and 30 meters of cell size were obtained from a survey with the Terrestrial Laser Scanner (TLS) in the study area and from the points cloud generated. The use of covariates did not improve the prediction of bases for the studied area. The predicted final maps did not have great differences in their RMSEs, but had visual differences. The use of maps with better resolution does not improve the prediction using Random Forest, for Ca^{2+} . Areas with high complexity deserve more studies on their chemical attributes, considering the production of predictive maps.

Keywords: Predictive covariates, interpolation, digital mapping

1. Introdução

A variabilidade espacial de propriedades do solo é uma preocupação antiga dentro da Ciência do Solo (VIEIRA, 2000). A estatística clássica, pressupõe a independência entre os dados, enquanto que a estatística espacial utiliza como pressuposto a primeira lei da Geografia, “*Tudo está relacionado com todo o resto, mas coisas próximas estão mais relacionadas do que coisas distantes*” (TOBLER, 1970), ou seja, a geoestatística pressupõe a dependência entre os dados a uma certa distância. A krigagem foi elaborada para tentar ajudar a resolver um grande número de problemas nas áreas de mineração, engenharia de petróleo, controle de poluição e saúde pública (WEBSTER& OLIVER, 2008). Os métodos de krigagem podem ser lineares e não-lineares; estacionários e não-estacionários. Os métodos de krigagem encontrados na literatura são a ordinária, simples, disjuntiva, indicativa, co-krigagem e universal entre outras (FERREIRA; SANTOS; RODRIGUES, 2013; SARMA, 2009). A krigagem possui uma série de vantagens em relação a outros métodos de interpolação, como por exemplo, fornecer em média estimativas sem viés e com variância mínima, além de ser possível conhecer a variância da estimativa (FERREIRA; SANTOS; RODRIGUES, 2013; VIEIRA, 2000). Os tipos mais utilizados são a krigagem simples (KS) e ordinária, sendo a primeira utilizada quando a média é assumida como conhecida em toda a área estudada e a segunda considera a média flutuante ou móvel (FERREIRA; SANTOS; RODRIGUES, 2013).

Esses métodos não utilizam nenhum mecanismo de seleção por árvores de decisão.

Classificadores que usam árvores de decisão possuem limitações quando o modelo estatístico se ajusta em demasiado com dos dados de treinamento (HO, 1995).

Os resíduos de processos de modelos que utilizam a autocorrelação são úteis para a interpretação e processamento para interpolação de dados não amostrados, pois indicam o grau em que as amostras estão sujeitas a processos atípicos (KIRKWOOD et al., 2016).

Kirkwood et al. (2016) salientam que variáveis auxiliares (covariáveis) podem ser introduzidas para predição de atributos geoquímicos, como a gamaespectrometria e dados a partir de outros levantamentos. Entre os métodos de obtenção de covariáveis, salienta-se os dados obtidos por levantamentos LIDAR (*Light Detection And Ranging*), aerotransportados e ou laser Scanner terrestre (TLS), que obtém pontos com a informação espacial tridimensional e, em alguns casos, é possível obter a reflectância e intensidade do sinal refletido em cada ponto registrados pelo equipamento.

Já Breiman & Breiman (2001) conceituam a técnica conhecida como Random Forest (RF), como sendo sucessivas árvores de decisão que se desenvolvem pela introdução de elementos randômicos em sua construção. Esta técnica pode ser utilizada para predizer tanto dados categóricos quanto contínuos.

Neste contexto o presente trabalho tem o objetivo de verificar a existência de incremento (melhor qualidade) na predição de Ca^{2+} , Mg^{2+} e K^+ , com o uso de covariáveis, em superfícies restritas de áreas livres de gelo comparada a modelos geoestatísticos que utilizam apenas um conjunto de amostras da variável para predizer o fenômeno.

2. Materiais e Métodos

2.1 Área de estudo

O presente trabalho foi realizado na Península Keller, Ilha Rei George situada no arquipélago das Schetlands do Sul, Antártica Marítima (Figura 1).

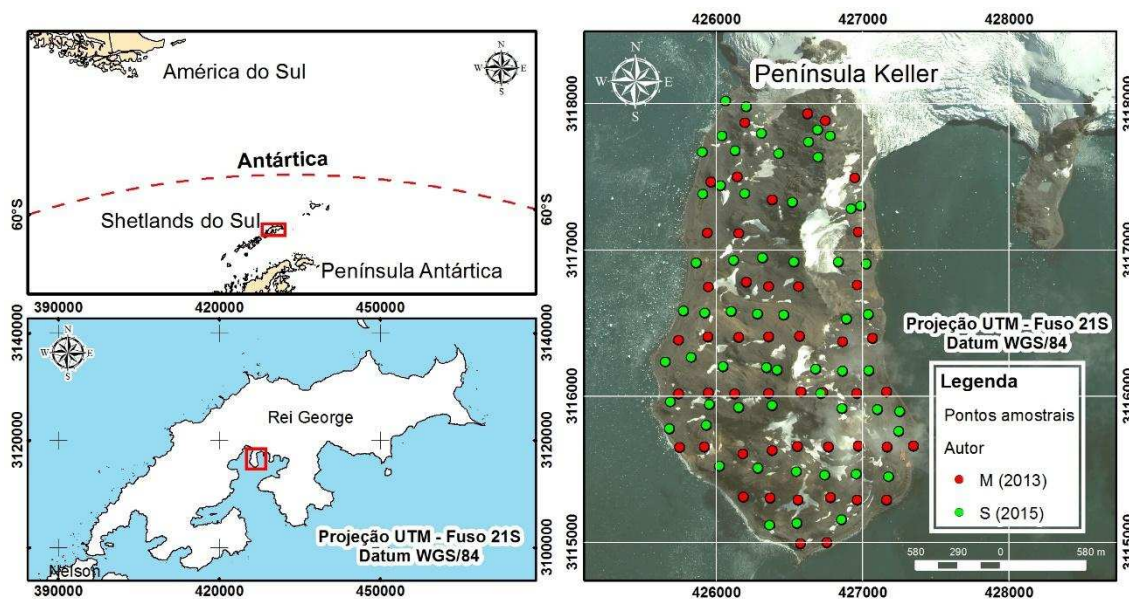


Figura 1 – Localização das Schetlands do Sul, Ilha Rei George e Península Keller, com os pontos de coleta de amostras realizados por Schünemann (2015) (S) e Moraes(2013) (M).

A península fica entre as enseadas Martel e Mackellar, em local denominado de Baía do Almirantado. A mesma possui aproximadamente 500 ha de área, com um comprimento aproximado de 4km na direção norte-sul e aproximadamente 2km na direção leste-oeste. As temperaturas anuais ficam na média de $-1,8^{\circ}\text{C}$ com precipitações médias de 360mm por ano (FRANCELINO et al., 2011). O local apresenta ventos com velocidade média anual de 21km.h^{-1} , com rajadas de 176km.h^{-1} (MORAES, 2013). Em partes da península ocorrem afloramentos de andesitos piritizados os quais originam sedimentos sulfatados ácidos (FRANCELINO et al., 2011; SIMAS et al., 2006) e solos que contrastam com material basáltico pela sua coloração amarelada (FRANCELINO et al., 2011). As unidades de solos da península, utilizando o sistema da FAO (WRB, 2015), e adaptações realizadas por Francelino et al. (2004), com suas áreas e porcentagens de cobertura na península estão sumarizadas na Tabela 1.

Tabela 1 – Unidades de solo da Península Keller, área e porcentagem de cobertura.

Unidades de solos	Área (m ²)	% Cobertura
Criossolo Lítico + Leptossolo Gélico CRli (Lpge)	77582,00	1,50
Criossolo Lítico + Regossolo Esqueletic-gélico (CRli + RGsk)	82635,50	1,60
Criossolo Lítico + Regossolo Esqueletic-gélico + Leptossolo Ornitogenic-gélico (CRli + RGsk + LPoge)	43640,00	0,90
Criossolo Paralítico + Cambissolo Leptic-gélico (CRlip + CMle)	89380,10	1,80
Criossolo Esquelético + Cambissolo Esqueletic-gélico (CRsk + CMsk)	143112,90	2,80
Criossolo Esquelético + Leptossolo Gélico (Crsk + LPge)	41233,10	0,80
Criossolo Vitric-léptico + Leptossolo Orthithionigélico + Regossolo Gelistágnico (CRvi + LPoge + RGge)	46111,90	0,90
Criossolo Vitric-léptico + Regossolo Gelistágnico (CRvi + RGgt)	57336,30	1,10
Criossolo Vitric-léptico + Regossolo Gelistágnico + Leptossolo Gélico (CRvi + RGgt + Lpge)	46733,10	0,90
Fluissolo Gélico (FLge)	3542,90	0,10
Fluissolo Vitric-gélico (FLvi,ge)	7588,30	0,10
Leptossolo Ornitogenic-gélico + Regossolo Paralítico (LPge + RG)	122822,80	2,40
Leptossolo Ornitogenic-gélico (LPoge)	86820,00	1,70
Leptossolo Vitric-gélico + Regossolo Gelistágnico (LPvi + RGgt)	14361,30	0,30
Total	862900,2	16,9

Fonte: Adaptado de Francelino et al. (2004).

Conforme verificado, apenas cerca de 17% da península é coberta por algum tipo de solo, as demais áreas estão com geleiras, banco de neve, afloramento rochoso e sedimentos inconsolidados (FRANCELINO et al., 2011). Estas áreas de solos são restritas a bolsões, não sendo em muitos casos grandes extensões contínuas. Segundo Simas et al. (2008), o teor de Ca^{2+} varia entre 4,7 e 24,6 $\text{cmol}_c.\text{dm}^{-3}$, na Baía do Almirantado. Já Albuquerque-Filho (2005), encontrou valores variando entre 5 a 47 $\text{cmol}_c.\text{dm}^{-3}$ na própria Península Keller.

2.2 Análise de solos

Para a análise dos atributos pedogeoquímicos foram coletadas 58 amostras de solo em duas profundidades: 0 a 10 cm (A) e 10 a 30 cm (B), distribuídos espacialmente dispersos, em forma de uma grade semi-regular (Figura 1 - S). A coleta foi realizada durante a XXXIII Operação Antártica no verão austral 2014-2015.

Todas as amostras foram secas ao ar e passada por peneira de diâmetro de 2 mm. Depois foram pesadas, sendo quantificadas as porcentagens da fração grosseira (CASC) que ficaram retidas na peneira e fração fina denominada aqui de

terra fina seca ao ar (TFSA) que passou pela peneira.

Os atributos físicos analisados foram: areia grossa (AG), areia fina (AF), silte (SIL), argila (ARG), cascalho (CASC) e TFSA. Já os atributos químicos analisados foram pH em água, pH KCl, fósforo (P), potássio (K), cálcio (Ca^{2+}), magnésio (Mg^{2+}), alumínio (Al^{3+}), acidez potencial (H + Al), soma de bases trocáveis (SB), capacidade de troca catiônica efetiva (t), capacidade de troca catiônica a pH 7,0 (T), índice de saturação por bases (V), índice de saturação por alumínio (m), matéria orgânica (MO), fósforo remanescente (P-Rem) e enxofre (S).

Para as determinações foram pesadas as amostras, utilizando o método gravimétrico para determinação da quantidade de amostras a ser analisada.

Todas as amostras foram submetidas ao laboratório de análise de solos da Universidade Federal de Viçosa para as determinações elencadas.

2.3 Densificação da amostragem

Para a densificação da amostragem foram utilizadas as análises realizadas por Moraes (2013), cuja distribuição dos 48 pontos de coletas na península está ilustrada na Figura 1 (M). Essas foram somadas ao conjunto com 58 amostras realizadas no presente trabalho. Moraes (2013) coletou suas amostras nas profundidades de 0 a 5, 5 a 10 e 10 a 30cm. Para utilizar os dados coletados pelo mesmo, foram observados os valores das médias entre as duas camadas superiores do solo para que os valores pudessem ser utilizados em conjunto com os outros 58 pontos amostrais (S) obtidos na camada de 0 a 10cm. Um ajuste nos valores das análises dos 58 pontos (S) foi realizado para utilizar a mesma unidade e metodologia de análise, sendo apresentado como um fator de conversão. Para tanto as 58 amostras nas duas profundidades foram pesadas e seus resultados convertidos de gravimétrico para volumétrico. O fator de conversão foi obtido da seguinte forma:

1º) Foram pesadas cada amostra (S) com um cilindro (Cx) de 10 cm^{-3} e obtendo a densidade das mesmas, o que resultou numa densidade em $\text{g} \cdot 10\text{cm}^{-3}$.

2º) Para as determinações das análises de S foram utilizadas 10g, portanto:

$$Cx_{(g)} \rightarrow 10\text{cm}^{-3};$$

$10_{(g)} \rightarrow V_{(\text{cm}^3)}$, sendo V , o volume correspondente as 10g (S) dado em cm^{-3} , através

da Equação 1:

$$V_{(cm^{-3})} = 100 / Cx_{(g)}, \quad (1)$$

sendo:

Cx = valor obtido da pesagem do volume de 10 cm^{-3} do cilindro em g;

V = volume da amostra de 10 g dada em cm^{-3} .

3º) Como a análise de M foi obtida em 10 cm^{-3} e não 10 g, multiplicou-se o resultado de S por 10 (cm^{-3}) e dividir por V , ficando a Equação 2 assim:

$$Rot_{(M)} = \frac{Rot_{(S)} * 10}{V}, \quad (2)$$

sendo:

$Rot_{(M)}$ = resultado da análise da rotina para Moraes (2013), ou fator de conversão;

$Rot_{(S)}$ = resultado da análise de rotina nesse trabalho (S);

V = volume equivalente da amostra de 10 g deste trabalho (S);

De posse desses valores foi possível juntar as duas bases de análises o que tornou a malha de pontos mais densa, como pode ser observado na Figura 1, onde a distribuição das amostras dos dois autores são representadas em conjunto.

2.4 Interpoladores preditivos

Para comparar o poder de predição de atributos pedogeoquímicos para a área de estudo foram utilizados os interpoladores RF com tamanhos de célula de 1, 5, 10, 20 e 30 m para o Ca^{2+} e KS com o tamanho de célula de 30 m. Para cada processamento foi elaborado um *script* no programa computacional R (R Core Team, 2015).

Para a krigagem não foi necessária a utilização de covariáveis, sendo portanto utilizadas somente as variáveis amostrais para Ca^{2+} , Mg^{2+} e K^{+} , obtidas das análises e junção das duas bases amostrais, conforme descrito na densificação da amostragem. O total de amostras obtido foi de 102. Para a obtenção do mapa preditivo destas variáveis, foi necessário obter uma grade local, onde foram assinalados os valores preditos pelo interpolador. Esta grade foi confeccionada a partir de modelos digitais de elevação (MDEs) obtidos a partir de uma nuvem de

pontos, originada a partir de um levantamento realizado com Laser Scanner Terrestre (TLS) da RIEGL, modelo VZ-1000, realizado durante a XXXIII Operação Antártica Brasileira, no ano de 2015. As grades foram produzidas a partir da nuvem contendo aproximadamente 270 milhões de pontos, gerando MDEs com 1, 5, 10, 20 e 30 m de resolução geométrica.

Já o RF, utiliza como base os valores das variáveis amostradas em conjunto com covariáveis amostradas em pontos intermediários às primeira, que representam melhor o universo dos solos. Para o presente trabalho, as covariáveis utilizadas foram os mapas de litologia (FRANCELINO et al., 2004), geomorfologia (FRANCELINO et al., 2011), distância euclidiana dos ninhos de aves presentes na península Keller, gerada no programa ArcGIS 10.1 a partir dos dados obtido de (CARNEIRO et al., 2010), três bandas do visível e uma do infravermelho próximo da imagem de satélite Quick Bird (2005), o MDE gerado pelo TLS e as covariáveis morfométricas oriundas deste, conforme a Tabela 3, somando um total de 50 covariáveis.

Para a obtenção destas covariáveis elaborou-se um *script* no programa computacional R, que de forma automatizada gerou, a partir de cada MDE nas 5 diferentes resoluções espaciais, os diferentes temas morfométricos. Utilizou-se para tanto os pacotes “RSAGA” (BRENNING, 2008) e “*raster*” (HIJMANS, 2015), sendo que o primeiro possibilitou o uso das ferramentas computacionais do programa SAGA GIS e o segundo permitiu operações com matrizes (*rasters*).

As covariáveis obtidas a partir do MDE utilizadas são apresentadas nas Tabelas 2, 3 e 4.

Tabela 2 – covariáveis morfométricas do grupo das curvaturas

Derivadas da Curvatura		
Nº	VARIÁVEIS MORFOMÉTRICA	Descrição Breve
1	Curvatura de Seção	
2	Curvatura da linha de fluxo	
3	Curvatura Geral	
4	Curvatura Longitudinal	
5	Curvatura Máximal	Na seção normal local
6	Curvatura Minimal	Referente a seção normal local
7	Curvatura Plana	
8	Curvatura de Perfil	Descreve o segundo mecanismo de acumulação
9	Curvatura Tangencial	Descreve o primeiro mecanismo de acumulação
10	Classificação da Curvatura	Divide as curvaturas em 8 classes (discreto)
11	Curvatura Total	TOTAC = MEANC2 – DIFFC2 (Curvatura média menos Diferença de Curvatura)

Tabela 3 – covariáveis morfométricas do grupo da radiação solar local

Derivadas da Radiação Solar		
Nº	VARIÁVEIS MORFOMÉTRICA	Descrição Breve
12	Solrad_diffuse1	Radiação solar difusa incidente no mês de janeiro
13	Solrad_diffuse2	Radiação solar difusa incidente no mês de junho
14	Solrad_direct1	Radiação solar direta incidente no mês de janeiro
15	Solrad_direct2	Radiação solar direta incidente no mês de junho
16	Solrad_dur1	Duração da radiação solar no mês de janeiro
17	Solrad_dur2	Duração da radiação solar no mês de junho
18	Solrad_total1	Radiação solar total (direta mais difusa) incidente no mês de janeiro
19	Solrad_total2	Radiação solar total (direta mais difusa) incidente no mês de junho

Tabela 4 – outro grupo decovariável utilizadas

Outras covariáveis		
Nº	VARIÁVEIS MORFOMÉTRICAS	Descrição Breve
20	Face de Exposição do Relevo	Orientação em relação aos pontos cardiais de cada face do relevo;
21	Índice de Convergência	Calcula um índice de convergência / divergência em relação ao escoamento superficial;
22	Diferença	Diferença no gradiente de declividade
23	Aquecimento Anisotrópico Diurno	Representa a assimetria no aquecimento comparando as faces leste e oeste locais, baseado em horas de radiação;
24	Gradiente	Corresponde ao gradiente hidráulico
25	Índice de Balanço de Massa	Índice que representa o balanço de massa em cada pixel
26	Modelo Digital de Elevação	Representa a elevação em cada célula do modelo;
27	Posição de Meia Encosta	Representa a distância em relação ao topo e vale, variando entre 0 e 1. Cobre as regiões mais quentes da declividade
28	Índice de multiresolução da planicidade do topo dos vales	Utiliza a planicidade e outras características dos fundos de vale. O método demonstra uma idéia baseada na lógica Fuzzi de Análise de Paisagens. Varia entre 0 e 1
29	Elevação normalizada	É uma medida de altura relativa de um ponto ao invés de seu valor elevação propriamente dito. Atribui o valor 1 para o ponto mais alto e o valor 0 para a posição mais baixa dentro de uma respectiva área de busca
30	Área real da superfície	Não considera a superfície como projetada, e sim a superfície real
31	Declividade	Representa a declividade local que pode ser angular ou em porcentagem;
32	Elevação da declividade	
33	Elevação Padronizada	É o produto da altura normalizada multiplicada pela altura absoluta
34	Superfície de pontos específicos	Indica diferenças entre pontos específicos de mudança da superfície, resultando em valores discretos positivos e negativos
35	Índice de Rugosidade Terrestre	Calcula a diferença dos valores de elevação a partir de uma célula central e suas oito células vizinhas
36	Convexidade da superfície terrestre	É calculado como a razão entre o número de células que têm curvatura positiva (células convexas) para o número de todas as células válidas dentro de um raio de pesquisa específico
37	Textura da superfície terrestre	Divide a textura da superfície em 8, 12 ou 16 classes
38	Profundidade dos vales	Inverte a elevação, deriva as redes de drenagem e calcula a
39	Vetor de medidas de rugosidade	Mede a variação na rugosidade do terreno, variando de
40	Índice topográfico de umidade	Descreve a tendência de cada célula em acumular água em função do relevo

2.5 Processamento dos interpoladores

Todo o processo da krigagem foi realizado a partir de um *script* construído para a linguagem do programa computacional R, sendo utilizados os pacotes geoR (JR.& DIGGLE, 2006), moments (NOVOMESTKY& KOMSTA, 2015), scatterplot3d (LIGGES& MARTIN, 2003), tcltk2 (GROSJEAN, 2015), sp (PEBESMA& BIVAND, 2005), rgdal (BIVAND; KEITT; ROWLINGSON, 2016) e RSAGA.

Para o processamento da krigagem, primeiramente foram carregados os dados amostrais para o ambiente do programa R, onde foi realizada a estatística descritiva e teste de normalidade, utilizando-se o teste de Shapiro e Wilk (1965).

O primeiro critério para a preparação do conjunto dos dados, é a determinação da máxima distância que foi considerada para o pareamento entre os dados e construção do semivariograma experimental, que é realizada interativamente de forma omnidirecional alterando os valores de máxima distância e verificando o semivariograma gerado.

Após esse primeiro processamento foi realizada a análise da anisotropia, plotando os semivariogramas nas 4 direções, 0°, 45°, 90° e 135°. A semelhança entre os 4 variogramas permitiu verificar visualmente a existência de anisotropia na área de estudo. A próxima etapa foi realizado o ajuste do modelo teórico visualmente e interativamente, onde foram determinados os melhores valores de patamar (C), alcance (a), efeito pepita (C_0) e o modelo matemático teórico. O ajuste dos dados foi realizado pelo método dos mínimos quadrados (MMQ) OLS (*Ordinary Least Squares*) (Figura 3, a, b e c). Os métodos mais utilizados são o OLS, REML (*Restricted Maximum Likelihood*), WLS (*Weighted Least Squares*), e ML (*Maximum Likelihood*). Para fins de identificação da qualidade do modelo utilizado (OLS), foi realizada uma autovalidação, tendo os indicadores média dos erros padronizados (MEP) e desvio dos erros padronizados (DEP). A MEP é melhor quanto mais próxima do valor zero, já DEP é tão melhor quanto mais próxima do valor 1 (VIEIRA, 2000). Outro indicador de qualidade é a variância dos erros, que pode ser interpretada como a variância da krigagem (WEBSTER& OLIVER, 2008). Ainda

testando o modelo, Vieira (2000) sugere que seja realizada uma regressão linear entre os valores observados e preditos pelo modelo, sendo os valores ideais de β_0 como o valor zero e β_1 como valor um.

Realizado o processamento independente do grid preditivo, realizou-se a KS preditiva, para a grade de 30 m e os mapas preditos foram exportados no formato ascii.

Tanto para a KS quanto para o random forest, foi realizado o RMSE através da validação cruzada conhecida como “leave-one-out”, tendo como base a equação a seguir (WILLMOTT& MATSUURA, 2005):

$$RMSE = \left[n^{-1} \sum_{i=1}^n |e_i|^2 \right]^{1/2}, \quad (3)$$

onde

$RMSE$ = raiz quadrada do erro médio quadrático;

e_i = erro de cada ponto a partir da diferença “leave-one-out”.

O processamento do RF, da mesma forma que a krigagem, foi realizado de forma automatizada e interativa a partir de um *script* desenvolvido para o programa R. Para este *script* foram utilizados os pacotes "gstat" (PEBESMA, 2004), "randomForest" (LIAW& WIENER, 2002), "raster", "sp", "rgdal" e "SDMTools" (VANDERWAL et al., 2014).

Para o início do processo foram definidas as covariáveis a serem selecionadas, como descrito anteriormente, e carregadas no ambiente computacional, para cada tamanho de célula utilizado e para cada variável estudada. Para o treinamento do preditor foram extraídos os valores de cada covariável nos locais correspondentes aos pontos amostrais e armazenados em uma matriz de treinamento. De posse desta matriz, em conjunto com o vetor dos valores amostrados, realizou-se a primeira rodada do interpolador. Após o processamento anterior o interpolador fornece um valor de ordem de importância de cada covariável, sendo que quanto maior o valor, maior a importância na predição da variável. Um dos indicadores desse processo é a porcentagem da variável que é explicada com todas as variáveis sendo utilizadas. Ao ordenar os valores de importância e plotar em um gráfico, verificou-se quais covariáveis foram as mais importantes, até a descendência dos pontos no gráfico, selecionando as mesmas para uma nova

rodada do interpolador, até este ponto. Logo após foi realizado um segundo processamento dos dados, porém somente com as variáveis selecionadas na análise visual do gráfico e verificado o incremento na porcentagem da variável explicada pelas covariáveis. Por fim realizou-se a predição pelo interpolador e o raster exportado para o formato ascii ESRI.

2.6 Comparação entre preditores

Para comparar o grau de conformidade entre os mapas preditos, realizou-se o RMSE entre todos os mapas gerados e variáveis estudadas. Para tanto utilizou-se o grid de 30 m como referência para gerar pontos, no caso do Ca^{2+} , ao qual foram extraídos os valores de todos os *rasters* gerados. Com os vetores produzidos da extração e excluídas as linhas que possuíam valores faltantes, foi realizado o cálculo do grau de concordância entre cada par de mapas, sendo considerada tanto maior a concordância, quanto mais próximo de zero fosse o valor.

Outra forma de verificar a concordância entre os mapas produzidos é através da correlação. Para eliminar a possibilidade de não detecção de correlações não lineares entre os mapas, utilizou-se o coeficiente de dependência randomizada, descrito por Lopez-Paz, Hennig, Schölkopf (2013), a partir de sua função adaptada e otimizada, a qual permitiu que todos os rasters fossem comparados entre si e seus valores representados através de um gráfico de matrizes.

Uma terceira forma de verificação de concordância realizada foi através do coeficiente de determinação (R^2), entre todas as matrizes produzidas.

Por fim foram verificadas as porcentagens de diferença, considerando os indicadores anteriormente elencados.

3. Resultados

3.1 Densificação da amostragem

A partir da densificação da amostragem e transformação dos valores obteve-

se a Tabela 5 com os valores de Ca^{2+} , Mg^{2+} e K^+ , para os 102 pontos amostrados.

A amplitude entre os valores observados para Ca^{2+} , a partir da densificação da amostragem dos valores de cálcio variando entre 42,09 e 0,45 $\text{cmol}_c \text{ dm}^{-3}$. Quando comparados os menores valores encontrados, com os valores observados por Simas et al. (2008) e Albuquerque-Filho (2005), este valor é relativamente inferior, sendo aproximadamente 91% inferior aos mínimos encontrados pelos autores anteriormente citados. Quando considerado o maior valor encontrado a partir da densificação dos dados, o mesmo ficou muito próximo ao valor observado por Albuquerque-Filho (2005) sendo somente 10% inferior ao mesmo.

Tabela 5 – Valores de Ca^{2+} , Mg^{2+} e K^+ , pontos densificados e identificação do autor

Dados Moraes (2013)						Dados Schünemann (2016)					
Nº	Long. E (m)	Lat. N (m)	Ca ²⁺ (cmol _c dm ⁻³)	Mg ²⁺ (cmol _c dm ⁻³)	K ⁺ (mg dm ⁻³)	Nº	Long. E (m)	Lat. N (m)	Ca ²⁺ (cmol _c dm ⁻³)	Mg ²⁺ (cmol _c dm ⁻³)	K ⁺ (mg dm ⁻³)
1	426181,2	3115307,1	7,05	1,75	42,17	48	426360,1	3115115,9	14,46	5,58	91,18
2	426363,0	3115300,7	2,20	1,65	24,48	49	426548,9	3115127,1	10,46	8,85	133,43
3	426554,9	3115286,7	12,65	1,80	38,38	50	426854,0	3115152,4	7,19	4,95	107,74
4	426776,4	3115304,0	19,85	0,85	39,64	51	426019,4	3115518,4	8,09	7,64	113,53
5	426961,4	3115286,7	6,15	5,90	126,83	52	426280,8	3115504,0	12,82	6,94	113,80
6	427161,6	3115288,2	5,15	2,85	53,54	53	426544,3	3115479,5	0,73	0,35	38,40
7	425744,3	3115648,5	3,70	2,05	52,28	54	426737,9	3115458,4	29,68	5,19	91,08
8	425913,6	3115652,5	1,35	5,95	32,06	55	427175,8	3115447,8	17,32	2,26	94,25
9	426178,8	3115601,6	11,60	10,70	77,55	56	425675,4	3115774,5	3,04	2,68	82,30
10	426379,4	3115625,8	17,55	7,90	99,03	57	425930,0	3115800,4	30,12	3,07	45,11
11	426552,6	3115654,7	28,05	1,45	43,43	58	427243,2	3115759,3	9,27	9,77	141,31
12	426761,8	3115650,0	19,75	0,85	35,85	59	427250,9	3115893,1	7,19	2,88	99,78
13	426966,6	3115657,1	14,15	1,95	51,01	60	427100,0	3115908,2	11,95	5,23	133,42
14	427162,8	3115651,0	8,30	5,35	54,80	61	426854,7	3115913,3	18,56	4,76	142,74
15	427342,9	3115657,4	0,80	1,95	90,18	62	426711,5	3116018,6	16,34	1,82	96,43
16	425739,6	3116014,7	13,20	1,95	32,06	63	426378,2	3115935,9	42,09	5,43	58,25
17	425943,7	3116021,2	13,30	4,60	45,96	64	426150,0	3115922,4	23,16	3,29	74,39
18	426123,8	3116013,8	23,90	3,55	23,21	65	425950,2	3115942,3	27,30	3,74	46,85
19	426354,2	3116017,1	25,10	1,00	35,85	66	425682,2	3115959,4	11,01	6,45	88,77
20	426574,5	3116029,4	23,35	0,40	28,27	67	425648,3	3116231,6	6,63	3,96	82,00
21	426724,0	3116024,5	17,25	0,95	30,79	68	425824,8	3116265,1	7,93	4,78	88,54
22	426957,4	3116017,1	17,40	2,35	38,38	69	426043,2	3116202,3	17,14	2,83	48,53
23	427161,7	3116027,9	9,90	5,75	62,38	70	426341,3	3116195,1	35,18	4,50	75,04
24	425738,7	3116381,4	4,15	4,25	55,35	71	426412,4	3116178,3	35,85	3,41	75,11
25	425937,5	3116405,3	0,45	1,00	38,59	72	426677,2	3116184,2	25,47	2,06	58,60
26	426149,4	3116402,3	12,75	6,05	82,66	73	426858,2	3116169,8	32,39	3,44	103,65
27	426353,1	3116405,2	13,20	8,25	62,83	74	427039,5	3116175,1	30,57	3,67	59,79
28	426566,1	3116407,9	30,30	3,70	56,22	75	427036,0	3116556,6	14,83	9,47	149,69
29	426860,9	3116370,3	21,05	3,75	67,24	76	426887,5	3116526,1	19,08	2,61	60,08
30	427063,4	3116395,1	13,10	8,30	67,24	77	426458,9	3116553,6	19,78	2,74	50,59
31	425942,8	3116745,8	15,65	3,00	69,44	78	426278,1	3116562,6	1,83	0,03	85,82
32	426201,8	3116780,4	22,15	3,00	6,90	79	426096,5	3116576,7	16,90	2,68	56,93
33	426355,3	3116746,5	18,95	3,40	1,62	80	425918,8	3116566,4	9,14	5,23	77,90
34	426558,2	3116749,3	19,15	3,90	-0,14	81	425773,4	3116582,0	8,75	4,10	70,82
35	426959,5	3116757,0	19,35	5,90	3,98	82	425858,0	3116911,6	16,87	1,13	69,11
36	425936,9	3117115,0	6,90	2,55	20,94	83	426114,6	3116927,1	11,12	3,22	74,39
37	426149,8	3117110,1	9,65	3,60	16,61	84	426313,3	3116944,9	19,38	5,09	142,58
38	426967,2	3117120,5	0,75	2,30	0,00	85	426526,0	3116917,6	29,49	3,62	58,50
39	425960,1	3117463,6	12,50	2,20	27,34	86	426830,3	3116916,3	6,74	2,24	20,75
40	426138,4	3117498,6	8,10	2,15	33,74	87	427020,4	3116902,8	16,45	4,38	104,92
41	426378,7	3117344,1	17,20	3,55	40,13	88	426983,2	3117303,1	13,62	4,60	150,58
42	426940,9	3117493,9	9,65	3,60	24,14	89	426919,7	3117279,9	9,76	2,61	123,58
43	426190,5	3117868,8	4,80	2,65	49,73	90	426516,4	3117325,1	25,19	1,55	37,22
44	426621,8	3117932,8	8,70	2,25	38,53	91	426191,4	3117383,2	15,86	1,60	106,40
45	426740,3	3117883,1	11,80	4,50	27,18	92	426025,2	3117438,9	14,34	1,67	76,31
46	426571,8	3114989,7	4,10	8,80	86,50	93	425904,9	3117382,1	13,17	1,42	93,17
47	426750,9	3114996,8	2,05	1,70	33,74	94	425902,2	3117669,6	6,61	2,27	60,06
						95	426125,7	3117680,6	10,72	1,67	101,05
						96	426423,2	3117658,7	15,98	0,81	59,68
						97	426693,4	3117634,9	8,65	2,80	107,06
						98	426776,8	3117778,0	9,96	5,16	68,91
						99	426628,3	3117738,4	12,70	1,29	93,79
						100	426304,4	3117796,4	13,08	1,72	186,96
						101	426037,3	3117780,7	9,30	1,48	32,86
						102	426690,3	3117820,3	7,02	3,38	102,24

Fonte: adaptado de Moraes (2013).

Legenda: Lat. = latitude; Long. = longitude.

A distribuição em quartis representa a forma com que os valores densificados ficaram distribuídos na península, para as três variáveis estudadas (Figura 2):

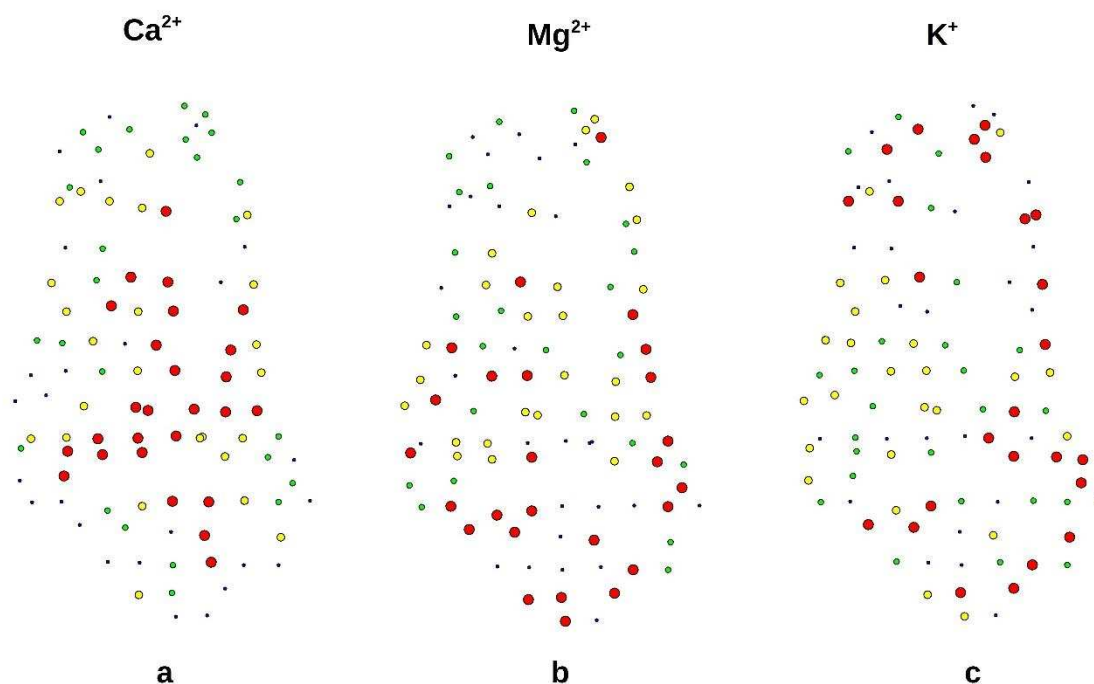


Figura 2 – Distribuição dos quartis para Ca^{2+} (a), Mg^{2+} (b) e K^+ (c), na área de estudo. Legenda: ● = 25% dos maiores valores; ● = 25% dos menores valores;

Observa-se que os maiores valores de Ca^{2+} são encontrados na parte central da península, evidenciado pelo quartil representado pelos pontos de coloração vermelha (4º quartil). Já os menores valores de quartis, são observados em sua maioria nas regiões costeiras, aqui representados pelos pontos azuis escuro (1º quartil). Isso evidencia que pode ocorrer uma influência do relevo e suas covariáveis no fenômeno estudado. Os valores intermediários compreendidos entre o segundo e terceiro quartis, estão distribuídos de uma maneira geral, em toda a área da península. Este comportamento não é observado para as outras variáveis estudadas, o que refletiu na utilização das covariáveis do MDE para a predição dos mesmos pelo RF.

3.2 Processamento dos interpoladores

Neste trabalho, optou-se por não realizar nenhum tratamento para valores discrepantes, em nenhum dos dois interpoladores utilizados.

A Figura 3, permite a visualização dos semivariogramas nas 4 direções principais para as três variáveis estudadas. Como observado na mesma, existe uma tendência a aproximação dos gráficos nas 4 direções para Ca^{2+} (a) e Mg^{2+} (b), o que indica a possível isotropia em todas as direções. Este comportamento aparentemente não se repetiu, para K^+ (c). Desta forma o variograma omnidirecional (considerando as 4 direções), representa com grande aproximação o comportamento dos dados observados.

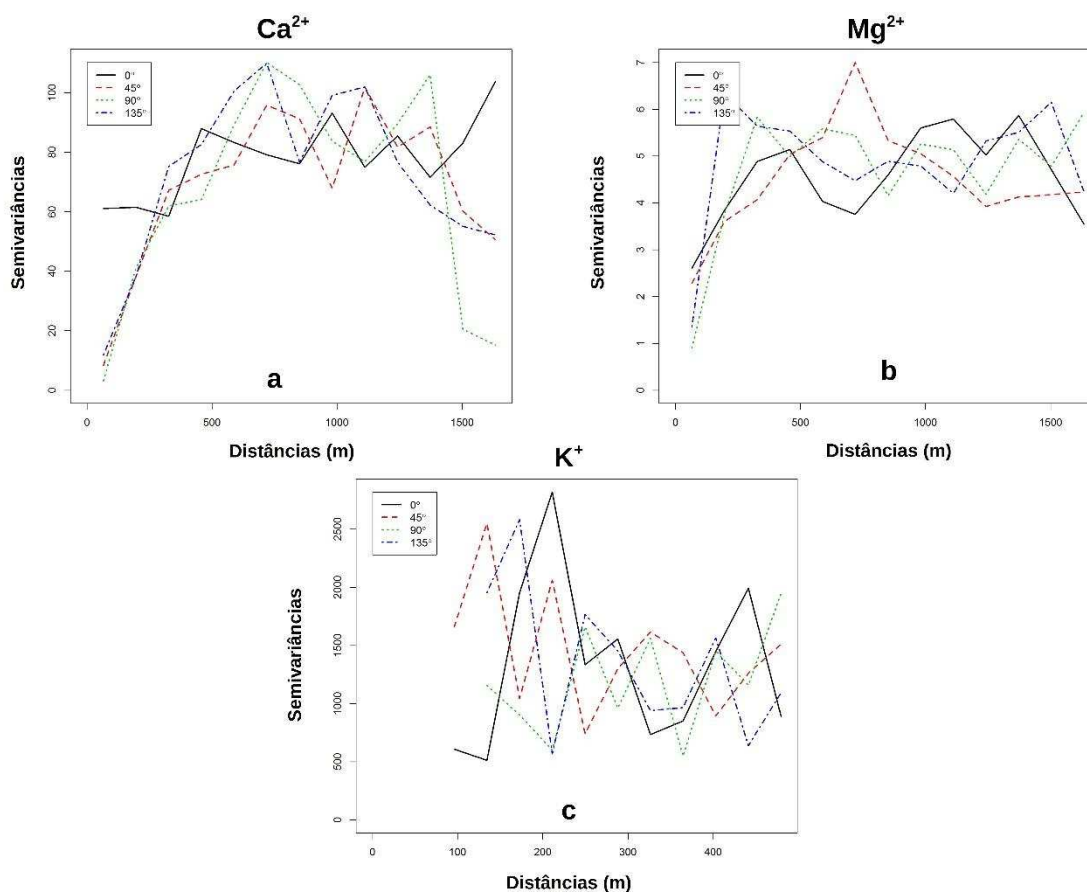


Figura 3 – Semivariogramas nas quatro direções principais para a variável estudada.

A realização da KS, com o método de ajuste OLS, para as variáveis estudadas, podem ser observados na Figura 4.

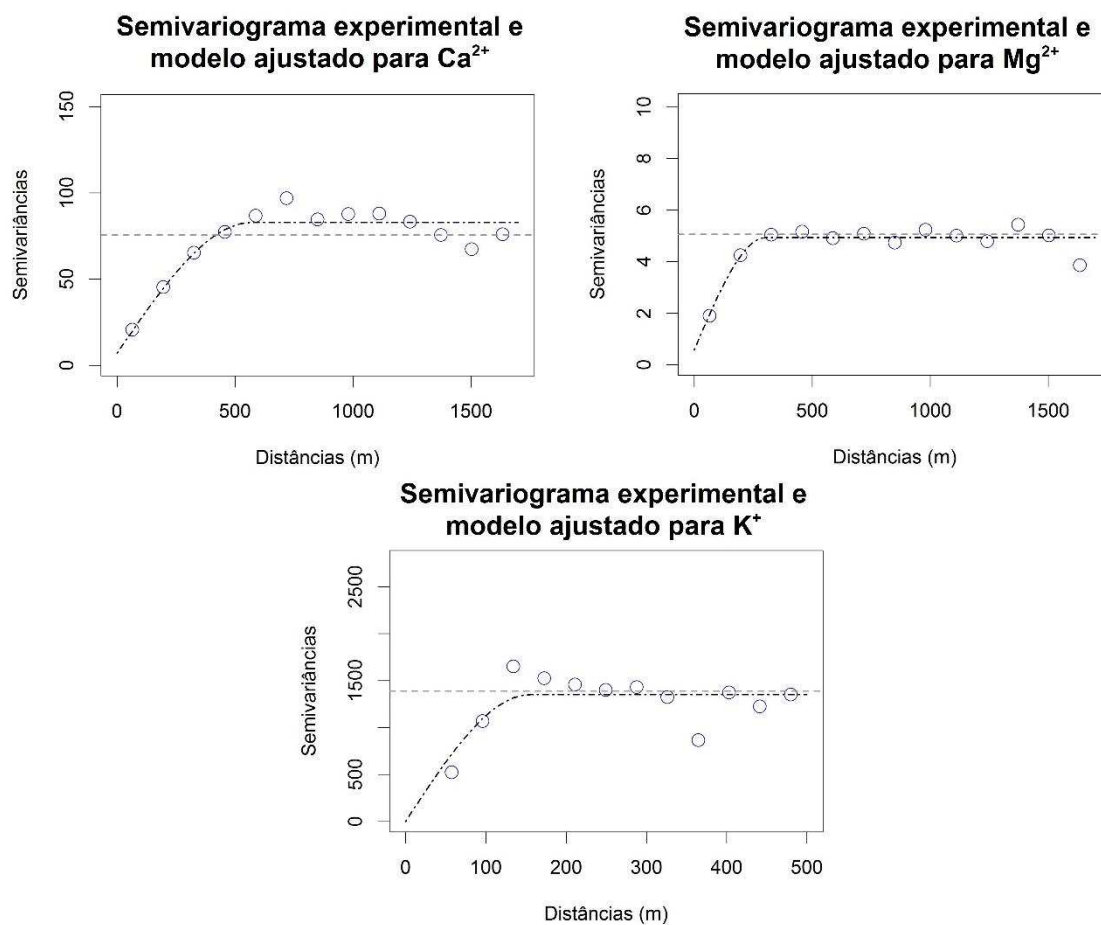


Figura 4 - Gráficos de semivariogramas experimentais (), ajuste por OLS (linha traço e ponto) e variância dos dados (linha tracejada).

Os indicadores obtidos das krigagens para as 3 variáveis estudadas são apresentados na Tabela 6.

Tabela 6 – indicadores da KS, e RMSE obtido das validações cruzadas para todos os interpoladores.

Base	Interp.	RMSE	Unidade	P1 (nº)	P2 (nº)	P3 (nº)	MEP	DEP	Beta 0	Beta 1	AP	Pep
Ca^{2+}	KS	6,686	($\text{cmol}_c \text{ dm}^{-3}$)	17	196	246	-0,004	1,08	1,26	0,91	569,457	7,00
Ca^{2+}	RF	7,569	($\text{cmol}_c \text{ dm}^{-3}$)	-	-	-	-	-	-	-	-	-
Mg^{2+}	KS	2,134	($\text{cmol}_c \text{ dm}^{-3}$)	17	196	246	-0,004	1,06	0,80	0,77	303,096	0,57
Mg^{2+}	RF	2,234	($\text{cmol}_c \text{ dm}^{-3}$)	-	-	-	-	-	-	-	-	-
K^{+}	KS	37,893	(mg dm^{-3})	3	9	25	0,000	1,12	83,26	-0,22	154,979	0,00
K^{+}	RF	37,269	(mg dm^{-3})	-	-	-	-	-	-	-	-	-

Legenda: Interp. = interpolador; P1 = número de pares no 1º ponto; P2 = número de pares no segundo ponto; P3 = número de pares no 3º ponto; MEP = média dos erros

padronizados; DEP = desvio dos erros padronizados; AP = alcance prático; Pep = pepita.

Os resultados apresentados de RMSE, quando comparados os interpoladores (KS e RF) possuem valores muito semelhantes.

Quando realizada a interpolação por RF, com tamanhos de células diferentes, para Ca^{2+} , o valor do RMSE, entre os diferentes tamanhos, se manteve numericamente muito semelhante, conforme observado na Tabela 7.

Tabela 7 – Valores de RMSE, alterando a resolução das covariáveis para a variável Ca^{2+} .

Resolução (m)	Interpolador	RMSE ($\text{cmol}_c \text{ dm}^{-3}$)
1	RF	7,781
5	RF	7,863
10	RF	7,792
20	RF	7,905
30	RF	7,569
30	KS	6,686

As covariáveis MDE, Índice de Umidade Topográfica e Elevação Padronizada, apareceram como as que mais explicam a variação da variável em todos os tamanhos de célula, para Ca^{2+} . Para Mg^{2+} , as covariáveis que melhor explicaram o fenômeno e foram selecionadas foram Índice de posição topográfica, perfil de curvatura, índice de convergência, convexidade de superfície, posição da meia encosta, índice de multiresolução de planicidade do topo dos vales e face de exposição. Já para K^+ , as covariáveis selecionadas pelo RF foram aquecimento anisotrópico diurno, curvatura longitudinal, perfil da curvatura e convexidade de superfície.

3.3 Comparação entre predições

Os mapas produzidos para Ca^{2+} , com diferentes tamanhos de célula, pelo RF são apresentados nas Figuras 5 e 6:

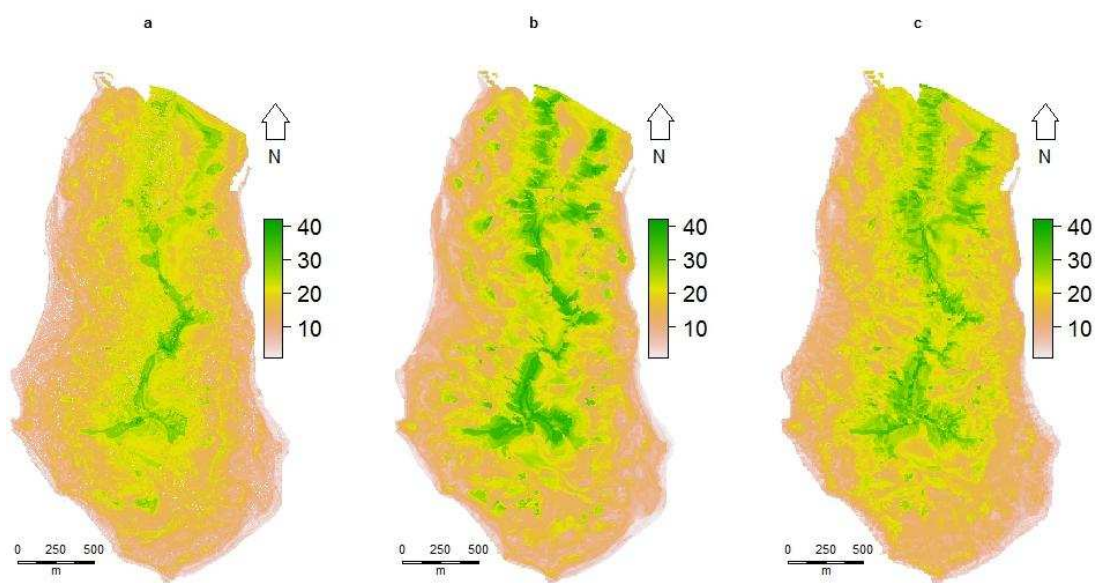


Figura 5 - Mapa preditos de Ca^{2+} ($\text{cmol}_c \text{ dm}^{-3}$) por RF a) 1m de tamanho de célula; b) 5m de tamanho de célula; c) 10m de tamanho de célula.

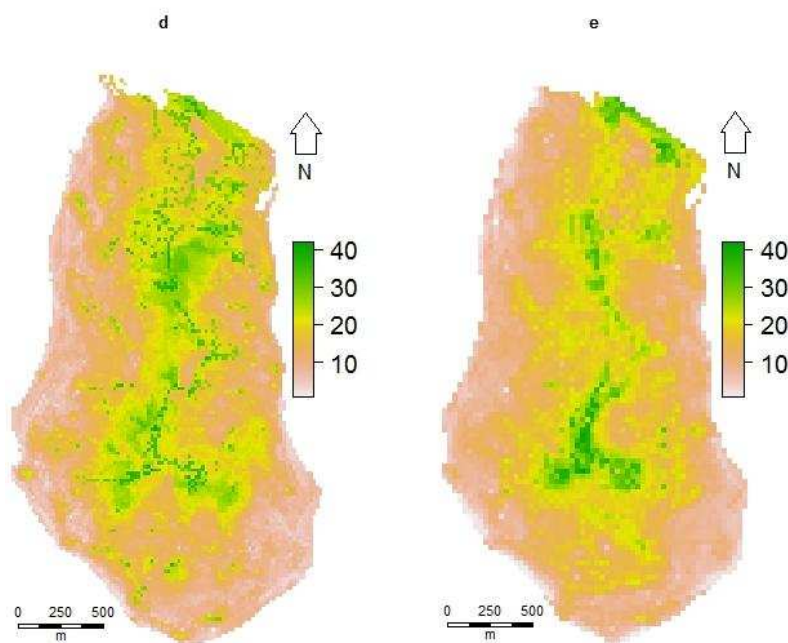


Figura 6 - Mapa preditos de Ca^{2+} ($\text{cmol}_c \text{ dm}^{-3}$) por RF d) 20m de tamanho de célula; e) 30m de tamanho de célula.

Ao realizar a análise visual dos mapas realizados por RF, observa-se uma diferença entre os mapas gerados. Quando comparados estes mapas com a KS, utilizando a análise visual, observa-se uma grande diferença nos mapas preditos.

A Figura 7 representa os mapas preditos para os 3 elementos estudados pelos diferentes interpoladores.

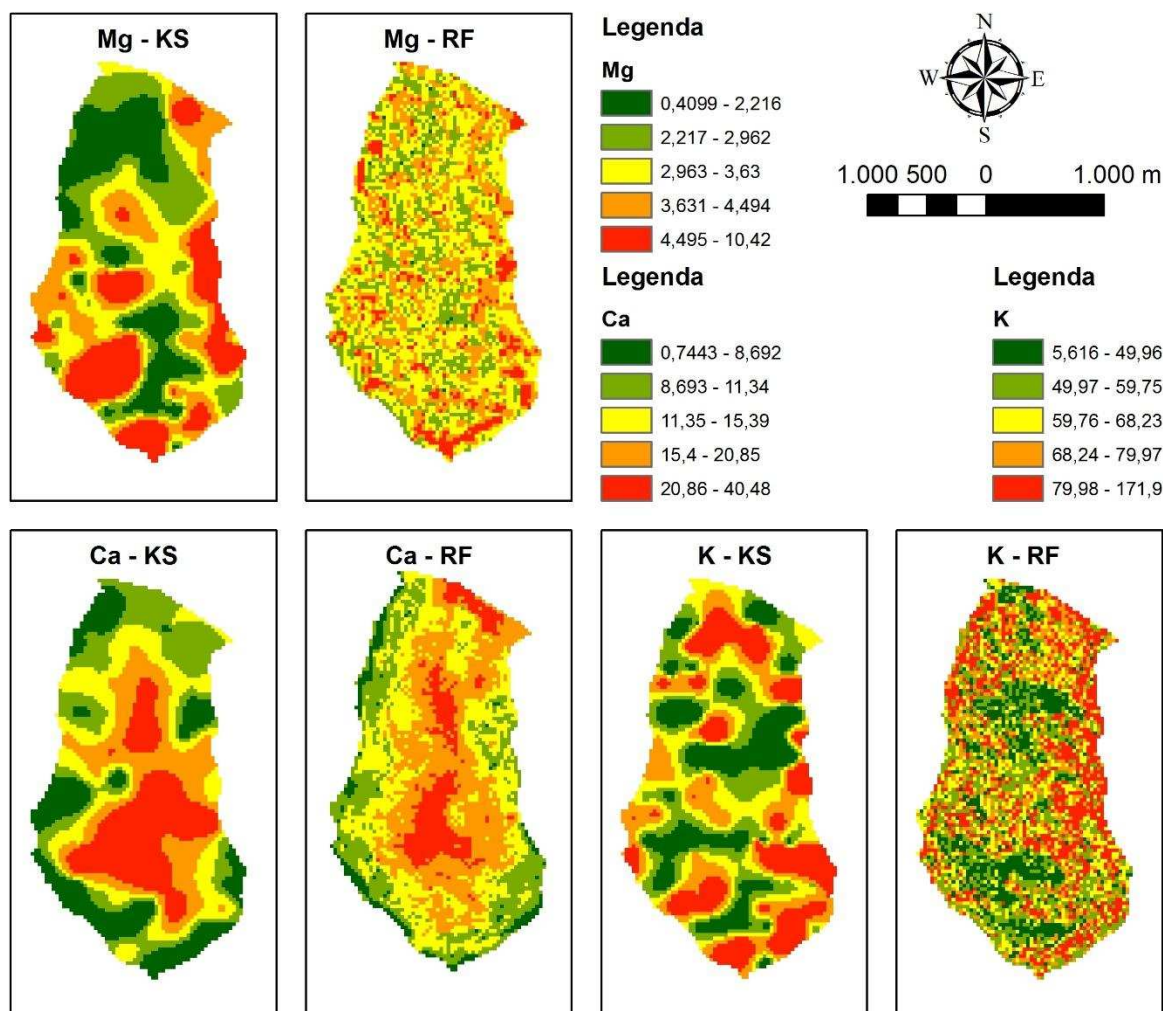


Figura 7 – Mapas preditos de Ca^{2+} ($\text{cmol}_c \text{ dm}^{-3}$), Mg^{2+} ($\text{cmol}_c \text{ dm}^{-3}$) e K^+ (mg dm^{-3}) pelos interpoladores KS e RF com tamanho de célula de 30m.

4. Discussão

A baixa representatividade dos valores preditos em relação aos valores observados, para K^+ , pode ser explicado pelo padrão observado no semivariograma experimental (Figura 4) para este elemento. Este semivariograma além de possuir poucos pares nos três primeiros pontos, apresenta um padrão semelhante ao chamado pepita puro. Se forem eliminados os pontos com menos de 30 pares

(Tabela 6) para este elemento, observa-se que a contribuição é quase nula, o que reforça essa hipótese. Embora esse comportamento seja observado, os valores de RMSE, quando comparados KS e RF, para K^+ , são muito semelhantes, numericamente.

Os resultados semelhantes, considerando o RMSE, para os dois interpoladores e variáveis estudadas, são um indicativo de que a utilização de covariáveis, não melhora a predição destes fenômenos. Mesmo no caso dos valores obtidos para Ca^{2+} , onde ocorreu uma diferença de 0,88 unidades, a KS obteve o melhor resultado numérico. Este fato chama a atenção, já que os quartis (Figura 2), indicaram um viés da elevação nesta variável o que demonstra a influência do pedoclima na distribuição deste atributo pedogeoquímico.

Esperava-se que as covariáveis Geologia e Geomorfologia contribuíssem para a explicação das variáveis estudadas, o que não foi observado em nenhuma das rodadas realizadas para o RF. Isto pode ser atribuído ao possível baixo nível de detalhamento (escala) destes mapas.

A análise visual demonstra grandes diferenças nos mapas preditos pelos dois interpoladores, para os três elementos estudados (Figura 7). Isto demonstra que mesmo que o RMSE demonstre valores muito semelhantes, podem ocorrer diferenças grandes nos mapas preditos. A análise qualitativa também demonstra uma maior continuidade nos mapas preditos para a KS quando comparados ao RF. Aparentemente o RF, representa melhor estes fenômenos, pois a complexidade do relevo, sugere esta descontinuidade. Este comportamento pode ser explicado pela natureza fragmentada das manchas solo na região, não possuindo grandes continuidades.

Analisando a mesma figura, o elemento Ca^{2+} , aparentemente teve maior concordância entre os mapas preditos por KS e RF, embora os valores de RMSE, tivessem uma pequena diferença.

5. Conclusões

Observou-se no presente trabalho interpoladores krigagem simples e random forest, não possuem grandes diferenças, quando consideradas suas superfícies

preditas, para as três variáveis estudadas. Salieta-se que o uso de covariáveis, incrementa o processamento computacional, da mesma forma que o uso de grids e ou tamanhos de célula menores, sem o aparente aumento na qualidade da predição.

As covariáveis Modelo Digital de Elevação, Índice de Umidade Topográfica e Elevação Padronizada são importantes na predição de Ca^{2+} , com o uso do interpolador random forest, independente do tamanho de célula utilizado, porém com baixa explicação da variabilidade desta variável (~24%).

Na ausência de covariáveis preditivas, a krigagem simples pode ser usada com resultados razoáveis para as variáveis estudadas, considerando os indicadores aqui apresentados, a partir de uma amostragem relativamente pequena, em áreas livres de gelo e de elevada complexidade em seus atributos químicos.

Para mapeamento de K^+ , a dependência espacial aparentemente exige um grid amostral mais denso. Isso pressupõe a diminuição da distância entre as amostras, o que pode encarecer o mapeamento deste fenômeno, em áreas livres de gelo. Para lançar mão do random forest, será necessário obter covariáveis que se correlacionem melhor com esta variável.

Para mapear Ca^{2+} , em áreas de elevada complexidade, como em regiões polares, é possível lançar mão da krigagem, em detrimento do processamento computacional e dificuldade de obtenção de covariáveis.

O random forest não melhorou a predição das variáveis aqui estudadas, sendo necessário aprofundar os estudos, para a predição de atributos químicos, em áreas de elevada complexidade em regiões polares.

Referências

ALBUQUERQUE-FILHO, M. R. **Geoquímica de solos da Península Keller, Ilha Rei George, Antártica, como subsídio ao monitoramento ambiental.** [s.l.] Universidade Federal de Viçosa, 2005.

BIVAND, R.; KEITT, T.; ROWLINGSON, B.. rgdal: Bindings for the Geospatial Data Abstraction Library. **R package version 1.1-9.** <https://CRAN.R-project.org/package=rgdal>, 2016.

BREIMAN, L.; BREIMAN, L. Statistical Modeling: The Two Cultures. **Statistical Science**, v. 16, n. 3, p. 199–215, 2001.

BRENNING, A. Statistical geocomputing combining R and SAGA: The example of landslide susceptibility analysis with generalized additive models. In: **SAGA--Seconds Out (Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, vol. 19)**. [s.l.: s.n.]. v. 19p. 23–32.

CARNEIRO, A. P. B. et al. Abundance and spatial distribution of sympatrically breeding *Catharacta* spp. (skuas) in Admiralty Bay, King George Island, Antarctica. **Polar Biology**, v. 33, n. 5, p. 673–682, 2010.

FERREIRA, Í. O.; SANTOS, G. R. DOS; RODRIGUES, D. D. Estudo sobre a utilização adequada da krigagem na representação computacional de superfícies batimétricas. **Revista Brasileira de Cartografia**, v. 5, n. 65/5, p. 831–842, 2013.

FRANCELINO, M. R. et al. Geomorphology and soils distribution under paraglacial conditions in an ice-free area of Admiralty Bay, King George Island, Antarctica. **Catena**, v. 85, n. 3, p. 194–204, 2011.

FRANCELINO, MÁRCIO ROCHA; SCHAEFER, CARLOS ERNESTO GONÇALVES REYNAUD; SIMAS, FELIPE NOGUEIRA BELLO; COSTA, L. M. GEOFORMAS DA PENÍNSULA KELLER, ANTÁRTICA MARÍTIMA: SUBSÍDIOS AO MONITORAMENTO AMBIENTAL. In: M.R., C. E. S.; FILHO, F. F. N. B. S. M. R. A. (Eds.). **Ecosistemas Costeiros e Monitoramento Ambiental da Antártica Marítima: Baía do Almirantado, Ilha Rei George - Rede 2.1**. ed. Viçosa-MG: NEPUT, 2004. p. 7–13.

GROSJEAN, P. H. **SciViews: A GUI API for R**. UMONS, Mons, Belgium. URL <http://www.sciviews.org/SciViews-R>, 2015.

HIJMANS, R. J. raster: Geographic Data Analysis and Modeling. **R package version 2.4-20**. <http://CRAN.R-project.org/package=raster>, 2015.

JR., P. J. R.; DIGGLE, P. J. geoR: Package for Geostatistical Data Analysis An illustrative session. **Artificial Intelligence**, v. 1, n. 2, p. 1–24, 2006.

KIRKWOOD, C. et al. A machine learning approach to geochemical mapping. **Journal of Geochemical Exploration**, maio 2016.

KOMSTA, L.; NOVOMESTKY, F. moments: Moments, cumulants, skewness, kurtosis and related tests. **R package version 0.14**. <http://CRAN.R-project.org/package=moments>, 2015.

LIAW, A; WIENER, M. Classification and Regression by randomForest. **R news**, v. 2, n. December, p. 18–22, 2002.

LIGGES, U.; MARTIN, M. Scatterplot3d – an R package for Visualizing Multivariate Data. **Cran**, p. 1–36, 2003.

LOPEZ-PAZ, D.; HENNIG, P.; SCHÖLKOPF, B. The Randomized Dependence Coefficient. **Advances in neural** , p. 1–9, 2013.

MORAES, A. G. DE L. **Mapeamento Digital de Atributos Físicos e Químicos do Solo na Península de Keller- Ilha Rei George-Antártica Marítima**. [s.l.] Universidade Federal Rural do Rio de Janeiro, 2013.

PEBESMA, E. J. Multivariable geostatistics in S: The gstat package. **Computers and Geosciences**, v. 30, n. 7, p. 683–691, 2004.

PEBESMA, E. J.; BIVAND, R. S. Classes and methods for spatial data in R. **R News**, v. 5, n. 2, p. 9–13, 2005.

R Core Team. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. ISBN 3-900051-07-0, URL <https://www.R-project.org/>, 2015.

SARMA, D. D. **Geostatistics with applications in earth sciences**. [s.l: s.n.].

SHAPIRO, S. S.; WILK, M. B. An Analysis of Variance Test for Normality (Complete Samples). **Biometrika**, v. 52, n. 3-4, p. 591–611, 1965.

SIMAS, F. N. B. et al. Clay-sized minerals in permafrost-affected soils (Cryosols) from King George Island, Antarctica. **Clays and Clay Minerals**, v. 54, n. 6, p. 721–736, 2006.

SIMAS, F. N. B. et al. Genesis, properties and classification of Cryosols from Admiralty Bay, maritime Antarctica. **Geoderma**, v. 144, n. 1-2, p. 116–122, 2008.

TIN KAM HO. Random decision forests. **Proceedings of 3rd International Conference on Document Analysis and Recognition**, v. 1, p. 278–282, 1995.

TOBLER, W. R. A computer movie simulating urban growth in the Detroit region. **Economic Geography**, p. 234–240, 1970.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In: **Tópicos em Ciência do Solo**. [s.l: s.n.]. v. 1p. 1–54.

VanDerWal, J.; Falconi, L.; Januchowski, S.; Shoo, L.; Storlie, C. SDMTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises. **R package version 1.1-221**. <https://CRAN.R-project.org/package=SDMTools>, 2014.

WEBSTER, R.; OLIVER, M. A. **Geostatistics for Environmental Scientists: Second Edition**. [s.l: s.n.], 2008.

WRB. World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps. **World Soil Resources Reports No. 106**. Rome: [s.n.], 2015.

Conclusões gerais

O presente trabalho demonstra o potencial do uso de geotecnologias para o mapeamento de atributos químicos, como Ca^{2+} , Mg^{2+} e K^+ , elevações e obtenção de covariáveis para o mapeamento de diferentes fenômenos. Porém o uso da krigagem simples, não aparenta ser adequado para mapear o K^+ , nestas áreas.

O mapeamento digital, com elevado detalhamento, em áreas com elevada complexidade e de difícil acesso, pode ser realizado com o uso do laser scanner terrestre. Estes modelos poderão servir para identificar pequenas mudanças na paisagem local, a partir de novos mapeamentos detalhados.

A densidade de amostragem de pontos em áreas de difícil acesso pode ser equalizada para diminuir os custos de análises e evitar a revisita em locais de difícil acesso, como em regiões polares.

Áreas com diferentes complexidades, podem produzir diferentes resultados, quando do uso de interpoladores geoestatísticos, podendo ocorrer valores anômalos. É necessário aprofundar os estudos neste sentido, para evitar a ocorrência de valores anômalos e aumentar a confiabilidade nos métodos de amostragem.

É possível, em alguns casos, lançar mão de interpoladores que não pressupõe a dependência espacial. Neste caso cita-se o exemplo da área da Península Keller, que possui uma maior complexidade. O processamento com o interpolador Spline obteve melhores resultados, com boa predição de elevação. Uma grande vantagem, é o ganho no tempo de processamento com o mesmo, sendo o anterior muito inferior ao obtido pelos interpoladores geoestatísticos.