

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA

DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

**ACURÁCIA DAS MÉTRICAS DE VALIDAÇÃO DA CLASSIFICAÇÃO DE
IMAGENS**

Josiane Aparecida Cardoso de Souza

Belo Horizonte

2020

JOSIANE APARECIDA CARDOSO DE SOUZA

**ACURÁCIA DAS MÉTRICAS DE VALIDAÇÃO DA CLASSIFICAÇÃO DE
IMAGENS**

Dissertação de mestrado apresentada ao programa de pós-graduação em Engenharia de Produção da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Engenharia de Produção.

Orientador: Prof. Britaldo Silveira Soares Filho

Belo Horizonte
2020

S729a

Souza, Josiane Aparecida Cardoso de.

Acurácia das métricas de validação da classificação de imagens [recurso eletrônico] / Josiane Aparecida Cardoso de Souza. - 2020. 1 recurso online (39 f. : il., color.) : pdf.

Orientador: Britaldo Silveira Soares Filho.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f.36-39.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia de Produção - Teses. 2. Confiabilidade (Engenharia) - Teses. 3. Mapas – Teses. I. Soares Filho, Britaldo Silveira.
II. Universidade Federal de Minas Gerais. Escola de Engenharia.
III. Título.

CDU: 658.5(043)

Ficha catalográfica: Biblioteca Profº Mário Werneck, Escola de Engenharia da UFMG



UNIVERSIDADE FEDERAL DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

UFMG

FOLHA DE APROVAÇÃO

ACURÁCIA DAS MÉTRICAS DE VALIDAÇÃO DA CLASSIFICAÇÃO DE IMAGENS


JOSIANE APARECIDA CARDOSO DE SOUZA

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA DE PRODUÇÃO, como requisito para obtenção do grau de Mestre em ENGENHARIA DE PRODUÇÃO, área de concentração PESQUISA OPERACIONAL E INTERVENÇÃO EM SISTEMAS SOCIOTÉCNICOS, linha de pesquisa Modelagem Estocástica e Simulação.

Aprovada em 17 de fevereiro de 2020, pela banca constituída pelos membros:


Prof(a). Britaldo Silveira Soares Filho - Orientador
UFMG


Prof(a). Marcelo Azevedo Costa
DEP/UFMG


Prof(a). Ubirajara de Oliveira
UFMG

Belo Horizonte, 17 de fevereiro de 2020.

AGRADECIMENTOS

Primeiramente, a Deus e a Nossa Senhora Aparecida e São Francisco de Assis por terem me concedido essa graça e por nunca me desampararem.

Aos meus pais, Jorge e Silvana, ao meu irmão, Bruno, que são minha base, por me apoiarem e me incentivarem.

Aos professores, Britaldo e Marcelo pelas valiosas contribuições, orientação e atenção que me foi dirigida.

Ao pessoal do Centro de Sensoriamento Remoto (CSR) e do Laboratório de Apoio à Decisão e Confiabilidade (LADEC) por toda atenção e ajuda.

A todos, meu muito obrigado.

“Senhor, dai-me força para mudar o que pode ser mudado...
Resignação para aceitar o que não pode ser mudado...
E sabedoria para distinguir uma coisa da outra.”

São Francisco de Assis

RESUMO

O Sensoriamento Remoto é uma importante ferramenta na aquisição de dados relacionados à Terra para a realização de diversos estudos e, conseqüentemente nas tomadas de decisões. Devendo esses dados ser acurados para que não causem conseqüências irreversíveis. No uso e cobertura do solo, a acurácia temática avalia as concordâncias entre a classificação e a verdade terrestre, geralmente representada por uma matriz de confusão, que por sua vez são aplicados índices de acurácia como, por exemplo: acurácia total, Kappa, entre outros. Considerando que os dados são acurados, estes oferecem confiabilidade na análise de acuracidade dos mapas? Este trabalho tem como objetivo analisar a confiabilidade de cinco índices de acurácia: acurácia total, Kappa, Pi de Scott, Tau e Pabak, de maneira em que o impacto das amostragens podem influenciar no valor de acurácia. Para isso, foram criados mapas de referência com quatro, cinco e seis classes, mapas de classificação com acurácias atribuídas de 50%, 70%, 85% e 95%. Em seguida foi realizada a validação considerando que os mapas são reais, para que após o cálculo dos índices de acurácia fossem comparados com os índices de acurácia atribuídas na criação dos mapas de classificação. Para a validação foram utilizadas janelas de tamanhos 5x5, 20x20 e 25x25 *pixels* em amostragens sistemática e aleatórias realizado no *Dinamica EGO 5*, onde foi realizado simulação de Monte Carlo, através de varreduras para analisar o comportamento dos dados das acurácias calculadas. As análises consistem em medidas de dispersão e tendência central dos dados, além da construção de histogramas e análise de regressão que foram realizados no *Rstudio 3.6.2*. Os resultados mostraram que os índices não sofrem variações em função do tipo de amostragem, mas podem ser influenciados pelo número de classes bem como pelo tipo de mapa, além dos valores de acurácias mais baixos como 0,50 e 0,70 tendem a sofrer maiores variações do que acurácias mais altas independentemente do tipo de índice.

Palavras-chave: índices de acurácia, matriz de confusão, confiabilidade, mapas.

ABSTRACT

Remote Sensing is an important tool on acquisition of information related to Earth and to accomplish many studies and consequently decision making. The data must be accurate to avoid irreversible consequences. On the usage and coverage of soil, the thematic accuracy evaluates the concordance between classification and true terrestrial, usually represented by a confusion matrix, then, accuracy index are applied, such as: total accuracy, Kappa, and others. Assuming the data is accurate, does those index offer reliability on the accuracy analysis of the maps? This paper has the objective to analyse the reliability of five accuracy index: total accuracy, Kappa, Scott's Pi, Tau and Pabak. To analyse it, was created maps of reference with four, five and six classes and maps of classification with attributed accuracy of 50%, 70%, 85% and 95%. After that, the validation was made considering that the maps are real. It was done with the purpose of compare the calculated accuracy index with the accuracy index attributed on the classification map. To validated it, were utilized windows of size 5x5, 20x20 and 25x25 pixels on random and systematic sampling on software Dinamica EGO 5, the maps were sweep to analyse the behavior of the calculated accuracy. The analysis consists on dispersion measure and central tendency of the data, histograms and regression analysis. The results shown that the most reliable index not vary on the type of sampling, but can be influenced by the number of class as well as by the type of map, in addition to the lower accuracy values such 0,50 e 0,70 tend to suffer greater variations than higher accuracy regardless of the type of index.

Keywords: *accuracy index, confusion matrix, reliability, maps*

LISTA DE FIGURAS

Figura 1 – etapas para execução do trabalho	13
Figura 2 – estrutura de agrupamento do mapa com 4 classes	15

SUMÁRIO

1. INTRODUÇÃO	9
1.1 Contextualização do tema	9
1.2 Objetivos	11
1.2.1 Objetivo geral	11
1.2.2 Objetivos específicos	11
1.3 Escopo do trabalho	11
2. METODOLOGIA	13
2.2.1 Construção dos mapas	14
2.3 Matriz de confusão	16
2.3.1 Acurácia total	18
2.3.2 Kappa de Cohen	19
2.3.3 Pi de Scott	20
2.3.4 Índice Tau	20
2.3.5 Pabak	21
3. RESULTADOS	22
3.1 Amostragem sistemática	22
3.1.1 Acurácia total	22
3.1.2 Kappa	25
3.1.3 Pi de Scott	26
3.1.4 Tau	27
3.1.5 Pabak	28
3.2 Amostragem aleatória	28
3.3 Análise de regressão	32
3.3.1 Amostragem sistemática	32
3.3.2 Amostragem aleatória	34
4 CONCLUSÃO	35
Referências	36

1. INTRODUÇÃO

Este trabalho objetiva analisar a confiabilidade dos índices de acurácia baseadas na matriz de confusão na matriz de confusão, de maneira em que o impacto das amostragens pode influenciar no valor de acurácia. Este capítulo contextualiza o tema, bem como define os objetivos e o escopo do trabalho.

1.1 Contextualização do tema

O Sensoriamento Remoto é uma técnica de observação terrestre que permite a obtenção de imagens de um objeto sem contato direto, através de sensores em aeronaves ou satélites (Richards, 2013). Ou seja, consiste em coletar e interpretar informações sobre um item a partir de uma determinada distância (Congalton e Green, 2019), sendo conhecida como a ciência da observação à distância (Barret e Curtis, 1999).

É uma ferramenta importante no fornecimento de dados de uso e cobertura da Terra (Anderson et al, 1976; Turner, 2005; Whitehead, Hugenholtz, 2014; Haack, Mahabir, Kerkering, 2015; Costa et al, 2018). Estes dados são fontes de informações imprescindíveis para a tomada de decisões nas áreas científicas, atividades de manejo e conservação do solo (Lyons et al, 2018) e em campos envolvendo detecção de mudança e classificação (Ma et al, 2019). Alguns exemplos da aplicação dos dados do uso de sensoriamento remoto envolvem estimar e mapear a área basal de povoamentos de eucaliptos (Reis et. al. , 2018); avaliar quantitativamente a suscetibilidade de deslizamento de terra (Liu e Duan, 2018); analisar o nível de infestação de ervas daninhas em uma produção agrícola (Ottofen et al, 2019) e mediar emissões de gases de efeito estufa (Li et al, 2017), dentre outros.

Existem diversas maneiras de criar mapas a partir do sensoriamento remoto, sendo imprescindível analisar a sua acurácia, isto é, a sua capacidade de representar fidedignamente os dados reais. Para isso, a classificação de imagens é a abordagem mais comum para mensurar a acurácia (Lyons, 2018). A classificação de imagem é geralmente referida ao mapeamento temático, a rotulação de objetos e características de determinados grupos baseados em seus atributos (Sokal, 1974; Jensen, 2015). A avaliação da acurácia indica a qualidade do mapa que foi criado através dos dados de sensoriamento remoto, sendo dividida em acurácia posicional e acurácia temática (Congalton e Green, 2009.p. 02).

A acurácia posicional corresponde à diferença da concordância locacional entre dados gerados por Sensoriamento Remoto e pontos terrestres conhecidos ou outra fonte considerada mais precisa que o mapa (Bolstad, 2005; Congalton, 2009.p.12). Seu cálculo é

feito através da Raiz do Erro Quadrático Médio (REQM) (Bolstad, 2005). A acurácia temática avalia a concordância entre a classificação ou atributos dos recursos de um mapa gerados e a verdade em campo, geralmente representada através de uma matriz de confusão (Congalton, 2009.p. 32). Em outras palavras, a matriz de confusão é o resultado de uma tabulação cruzada entre dados classificados (através de análises por sensoriamento remoto por sensoriamento) e os dados de referência (dados válidos) (Comber et al, 2019).

Através da matriz de confusão podem ser aplicadas várias medidas para avaliar a acurácia dos dados, dentre elas tem-se a Acurácia Total, a estatística Kappa de Cohen, o Índice de Tau, dentre outras. Também é possível analisar a acurácia de cada classe individual, através de medidas como acurácia do produtor e do usuário, Kappa condicional etc. (Couto, 2003). Sendo o uso e cobertura do solo, variáveis ambientais muito importantes, é indispensável que as informações sejam atualizadas e de alta qualidade, uma vez que sem uma alta qualidade há o risco de má interpretação e compreensão (Stehman e Foody, 2019).

Tendo em vista estes aspectos, de que os dados de uso e cobertura do solo são atualizados e possuem alta qualidade, cabe aplicar as medidas de acurácia. Várias são as literaturas que atribuem um valor de acurácia mínimo aceitável de 85% tais como Anderson (1971), Fitzpatrick-Lins (1981), Senseman, Bagley e Tweddale (1995). Já Aronoff (1985) sugere um intervalo de confiança para acurácia entre 85% e 95%. Jensen (1986; 2015) discute equações aceitáveis para determinar um número mínimo de pixels de acordo com o nível de acurácia desejado.

Além disso, diversos são os estudos que comparam medidas de acurácia tal como acurácia total, Kappa, Tau, etc (Couto, 2013; Fitzgerald e Lees, 1994; South, Qi e Lusch, 2004; Xia, Cheng e Li, 2019), mas sem deixar claro qual é a melhor dentre estas e se são realmente confiáveis.

O objetivo desse trabalho é analisar a confiabilidade de cinco índices de acurácias baseadas na matriz de confusão: acurácia total, Kappa, Pi de Scott, Tau e Pabak. Para alcançar tal objetivo, foram criados mapas de referências com 4, 5 e 6 classes, e mapas de classificação atribuindo acurácias de 50%, 70%, 85% e 95%. Em seguida foi realizada a validação considerando os mapas como fossem reais para posteriormente realizar a análise entre as acurácias que foram atribuídas e as novas acurácias que foram encontradas nos cálculos dos cinco índices. Na validação foram realizadas simulações de Monte Carlo, através de varreduras das janelas de tamanhos 5x5, 20x20 e 25x25 *pixels* para amostragens sistemáticas e aleatórias. O objetivo é avaliar como os dados de acurácias calculados se

comportam mediante aos números de classes, tamanhos das janelas, tipo de amostragem. As análises consistem em medidas de dispersão, centralidade e histogramas e análise de regressão linear. Os processos de construção dos mapas, validação e cálculo dos índices foram realizados no *Dinamica EGO 5* e a construção de tabelas e histogramas realizados no *Rstudio 3.6.2*.

1.2 Objetivos

1.2.1 Objetivo geral

Analisar a confiabilidade de cinco índices de acurácia baseados na matriz de confusão, de maneira em que o impacto das amostragens pode influenciar no valor de acurácia.

1.2.2 Objetivos específicos

- Gerar mapas de referências;
- Gerar mapas de classificação atribuindo diferentes valores de acurácia;
- Realizar validação através criação de janelas de tamanhos 5x5, 20x20 e 25x25 *pixels* para as amostragens sistemáticas e aleatórias;
- Aplicar simulação de Monte Carlo através de varreduras das janelas;
- Calcular os índices de acurácia através da matriz de confusão gerada: Acurácia Total, Kappa, Pi de Scott, Tau e Pabak;
- Realizar análises estatísticas através das medidas de dispersão e centralidade, comparando quais índices possui maior proximidade entre as acurácias atribuídas e calculadas.
- Realizar análise de regressão linear.

1.3 Escopo do trabalho

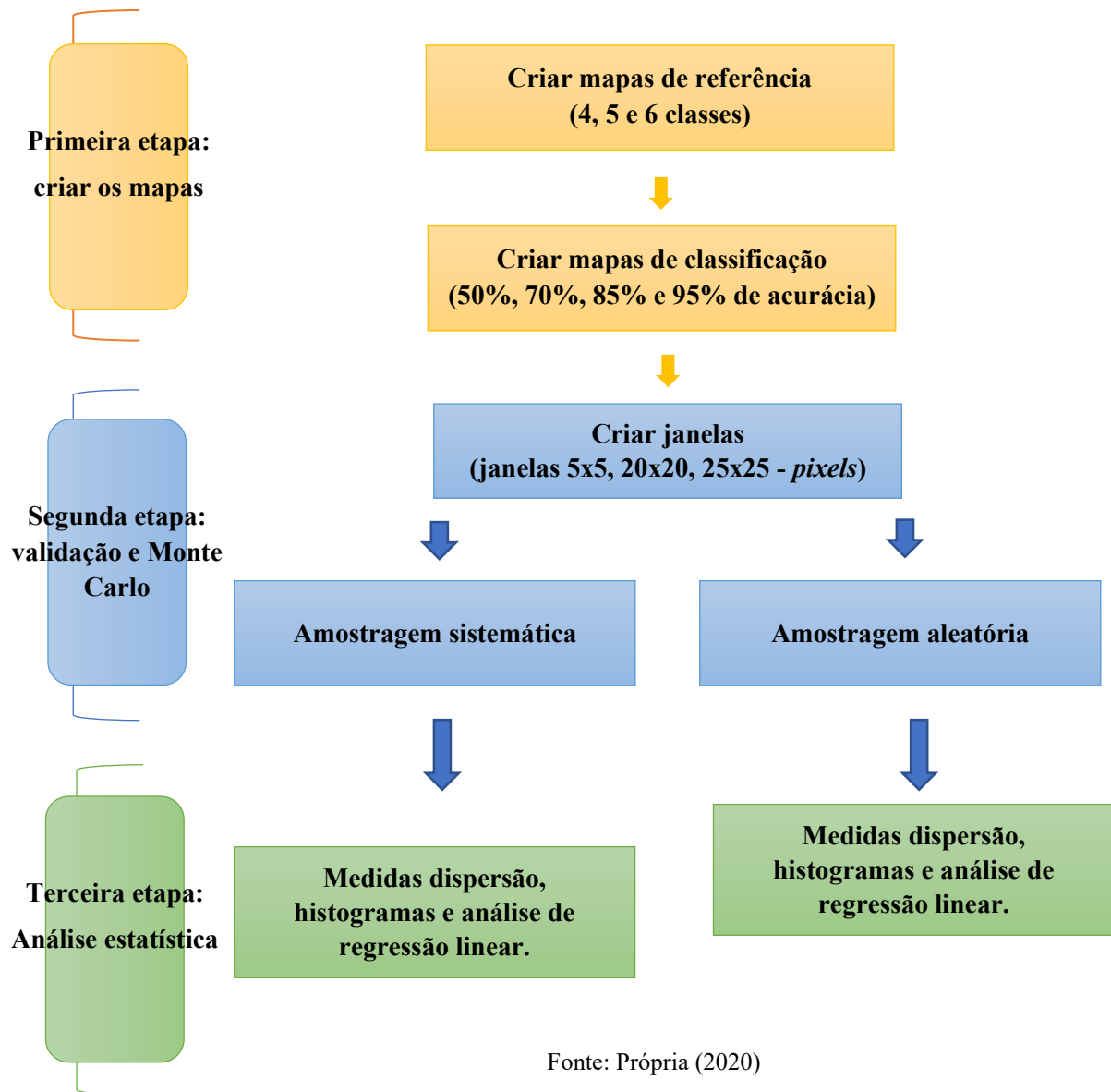
O presente trabalho pode ser dividido em três etapas, sendo ilustradas na figura 1:

- Primeira etapa: criação dos mapas de referências e de classificação. Inicialmente foram criados os mapas de referências contendo 4, 5 e 6 classes. No total foram criados 9 mapas de referências, sendo divididos entre as três classes, ou seja, três mapas para 4, 5 e 6 classes, e cada um deles possuindo diferentes configurações. Em seguida foram criados quatro mapas de classificação com níveis de acurácia iguais a 50%, 70%, 85% e 95% para cada mapa de referência. Maiores detalhes

podem ser encontrados no subtópico 2.2.1 e 2.2.2. Ambos os tipos de mapas foram criados no *Dinamica EGO 5*.

- Segunda etapa: Validação dos mapas. Para a validação dos mapas foram criadas janelas de tamanho iguais a 5x5, 20x20 e 25x25 *pixels*. Essas janelas foram criadas para dois tipos de amostragens: sistemática e aleatória. A amostragem sistemática contém distâncias iguais entre as janelas (35 *pixels*) e a amostragem aleatória contém distâncias aleatórias. Foram realizadas simulações Monte Carlo através de varreduras das janelas para avaliar o comportamento dos dados das acurácias calculadas. As simulações foram realizadas de acordo com os tamanhos das janelas, uma vez que os mapas possuem 300 linhas por 300 colunas. Assim, para a janela de tamanho 5x5 *pixels* foram realizadas 60 varreduras, janela 20x20 *pixels* foram realizadas 15 varreduras e 25x25 *pixels* foram realizadas 12 varreduras. Essa etapa também foi criada no *Dinamica EGO 5*.
- Terceira etapa: análise estatística. Com a varredura das janelas para cada amostragem, os dados foram analisados através da construção de histogramas, medidas de tendência central e dispersão, e análise de regressão. Essa etapa foi realizada no *RStudio 3.6.2*.

Figura 1 – etapas para execução do trabalho



2. METODOLOGIA

Com o objetivo de analisar se as medidas de acurácia temática baseadas na matriz de confusão realmente são confiáveis foram realizadas as etapas citadas na introdução (figura 1) que estão detalhadas aqui. As etapas de criação dos mapas, de validação e simulação Monte Carlo foram realizadas no *Dinamica EGO 5* e as análises estatísticas no *RStudio 3.6.2*. O tópico 4.1 detalha sobre o programa *Dinamica EGO 5* e seus subtópicos 4.1.1, 4.1.2 e 4.1.3 explicam a construção dos modelos para geração dos mapas referência, classificação e validação respectivamente.

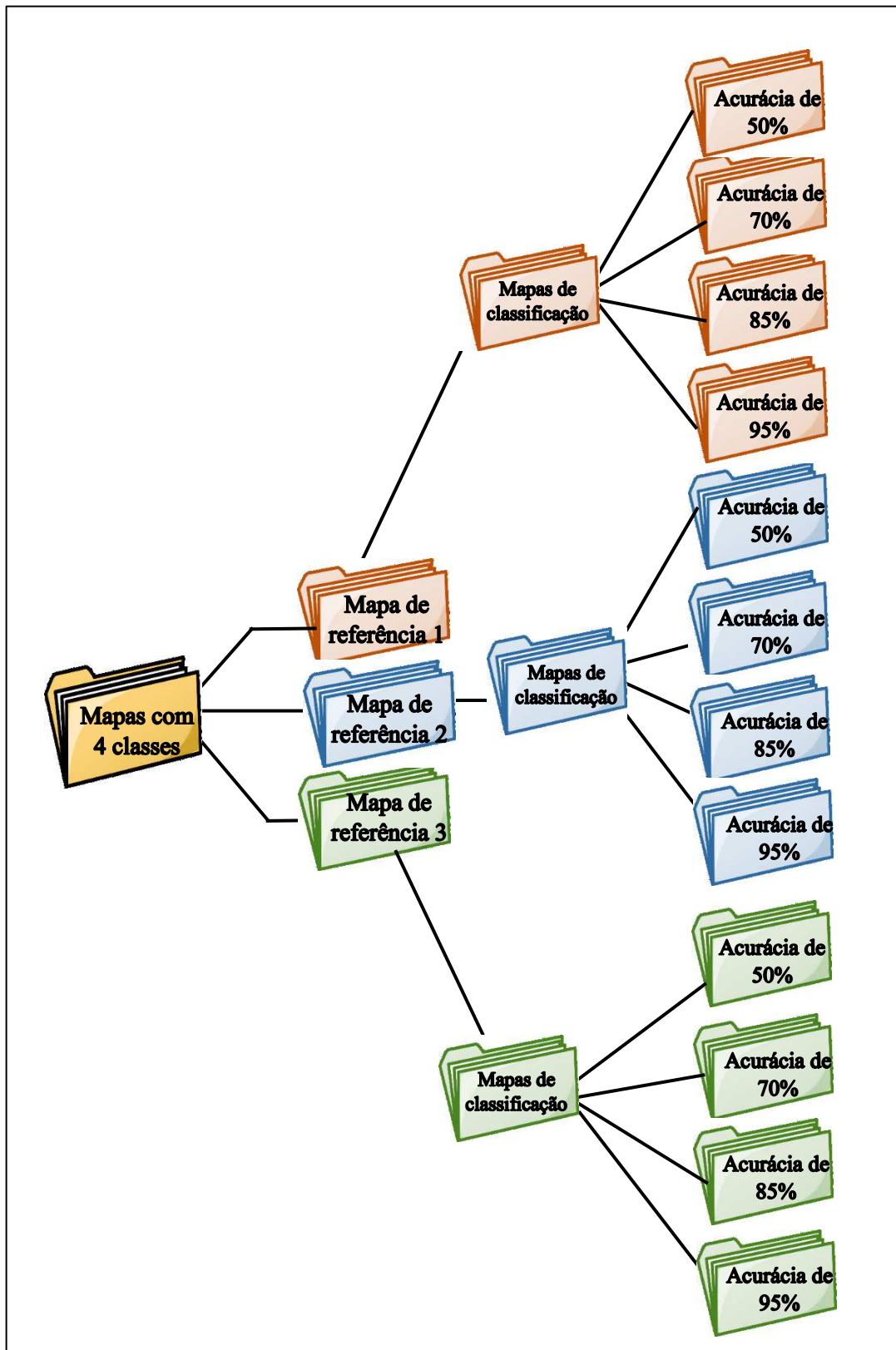
2.2.1 Construção dos mapas

Para a construção dos mapas foi estabelecido uma base com 300 linhas por 300 colunas, compondo assim uma base com 90000 *pixels* ou 90000 hectares, já que foi definido que cada *pixel* possua 1 hectare (100x100 m²). Foram gerados grupos de mapas no *Dinamica EGO 5* com 4, 5 e 6 classes, e para cada grupo há três subgrupos de mapas de referência – com diferentes configurações cada um. Para cada mapa de referência foram simulados mapas de classificação com quatro níveis de acuracidade: 50%, 70%, 85% e 95%. Isso é ilustrado na figura 15 para os mapas com quatro classes, sendo iguais para os outros mapas com cinco e seis classes – as cores servem apenas para diferenciar os subgrupos de mapas de referência. Assim, foram gerados 9 mapas de referências e 36 mapas de classificação, totalizando 45 mapas para as três classes.

Um único modelo foi construído no *Dinamica EGO 5*, porém para maiores detalhes sobre a criação este foi dividido em três submodelos:

1º – Submodelo para criação do mapa de referência: tais mapas foram criados tendo como base 13 mapas de Belo Horizonte e 8 da Amazônia extraídos aleatoriamente do MapBiomas, com número de classes iguais a 4, 5 e 6. O número de cada classe corresponde a classe dos mapas de Belo Horizonte ou da Amazônia de acordo com o MapBiomas, e as cores foram escolhidas aleatoriamente. As classes são: 3 (Formação Florestal) na cor verde, 11 (Área Úmida Natural não Florestal) na cor azul, 15 (Pastagem) na cor amarela, 21 (Mosaico de Agricultura e Pastagem) na cor bege, 24 (Infraestrutura Urbana) na cor cinza e 25 (Outra Área não Vegetada) na cor marrom. Foram escolhidos três mapas de cada classe e utilizadas as respectivas porcentagens de áreas de cada classe para a criação dos mapas de referência

Figura 2 – estrutura de agrupamento do mapa com 4 classes



Fonte: Própria (2020)

2º – Submodelo para criação do mapa de classificação: os mapas de referência foram utilizados como base para criar os mapas de classificação. Foram atribuídos quatro valores de acuracidade: 50%, 70%, 85% e 95% para cada mapa de classificação. Cada classe sofreu o mesmo número de transição no mapa de classificação, por exemplo, para acurácia de 85% em um mapa com 4 classes tem-se 15% de transição entre elas, assim cada classe sofreu 1,25% de transição.

3º – Submodelo para validação dos mapas de referência e classificação: foram utilizadas amostragens sistemáticas e aleatórias, para cada uma delas foram criados três tamanhos de janelas, 5x5 pixels, 20x20 pixels e 25x25 pixels. Para a amostragem sistemática foram utilizadas distâncias iguais entre as janelas, sendo definidos 35 pixels, já a amostragem aleatória, a distância entre as janelas foram geradas aleatoriamente, mas de forma que não houvesse sobreposições de janelas. Com a criação das janelas para cada tipo de amostragem, foram realizadas simulações de Monte Carlo através de varreduras. O número de varreduras variou de acordo com o tamanho da janela uma vez que o mapa possui dimensão 300x300 *pixels*. Desta forma, para a janela de tamanho 5x5 foram realizadas 60 varreduras, janela 20x20 foram feitas 15 varreduras e a janela 25x25 foram realizadas 12 varreduras.

2.3 Matriz de confusão

A matriz de confusão ou matriz de erro (Congalton, 1983; Stehman, Wickham, 2011; Khatami, Mountraki, Stehman, 2017) é uma ferramenta com arranjo quadrado que apresenta o grau de similaridade entre dados de referência (geralmente dispostos em colunas, derivados de fotos ou vídeos aéreos, medição de solo, etc) e dados de classificação (geralmente dispostos em linhas, gerados a partir de mapas ou dos dados de sensoriamento remoto) (Congalton, 1991). A representação da matriz de confusão é apresentada na Tabela 1.

Assumindo que n amostras estão dispostas dentre k^2 células, em que cada amostra é atribuída a uma das k classes no mapa, distribuídas de maneira independente a uma das mesmas k classes no conjunto de dados de referência. Deste modo, n_{ij} corresponde ao número de amostras distribuídas na classe i ($i = 1, 2, \dots, k$) no conjunto de dados de classificação e classe j ($j = 1, 2, \dots, k$) no conjunto de dados de referência. Já n_{i+} é o somatório do número de amostras denominadas como classe i do mapa e n_{+j} é o somatório do número de amostras da classe j nos dados de referência, respectivamente mostrados pelas equações 1 e 2.

$$n_{i+} = \sum_{j=1}^k n_{ij} \quad (1)$$

$$n_{+j} = \sum_{i=1}^k n_{ij} \quad (2)$$

A equação 1 corresponde à soma marginal de cada linha dos dados de classificação e, a equação 2 corresponde à soma marginal de cada coluna dos dados de referência.

A diagonal principal da matriz de confusão representa a concordância entre os dados de classificação e os dados de referência sendo representada por n_{ii} , onde $i = j$ ($i = 1, 2, \dots, k$) (Congalton, 2009). Portanto, a classificação correta é representada pelas amostras da diagonal principal da matriz de confusão e a classificação incorreta pelas amostras fora da diagonal principal (Couto, 2003; Stehman, 2006).

		Dados de referência (j = colunas)			Total das linhas n_{i+}
		1	2	k	
Dados de classificação (i = linhas)	1	n_{11}	n_{12}	n_{1k}	n_{1+}
	2	n_{21}	n_{22}	n_{2k}	n_{2+}
	k	n_{31}	n_{32}	n_{kk}	n_{k+}
Total das colunas n_{+j}		n_{+1}	n_{+2}	n_{+k}	n

Tabela 01: Adaptado de Congalton (2009, p.60).

Fonte: Congalton (2009)

Denotando p_{ij} como a proporção de amostras correspondente a i, j -ésima célula (n_{ij}), a equação 3 mostra a proporção de amostras que cada célula possui de uma combinação particular de classes de cobertura de terra classificada e referenciada (Salk et al, 2018).

$$p_{ij} = \frac{n_{ij}}{n} \quad (3)$$

Quando $i = j$ tem-se a proporção da diagonal principal da matriz:

$$p_{ii} = \frac{n_{ii}}{n} \quad (4)$$

Então a proporção marginal das linhas (p_{i+}) e colunas (p_{+j}) são definidas como:

$$p_{i+} = \sum_{j=1}^k p_{ij} = \frac{n_{i+}}{n} \quad (5)$$

$$p_{+j} = \sum_{i=1}^k p_{ij} = \frac{n_{+j}}{n} \quad (6)$$

com p_{i+} e p_{+j} definidas previamente como “concordância ao acaso”.

Além de exibir resultados empíricos convenientes, a matriz de confusão facilita a análise estatística (Card, 1982), sendo utilizados diversos métodos de análises de acurácia destacando-se acurácia total, Kappa e Tau.

Os métodos ou medidas de análise de acurácia foram escolhidos de acordo com seu nível de uso, ou seja, frequência com que são utilizados no Sensoriamento Remoto. Foi analisada a quantidade de artigos bibliográficos publicados que possuem em revistas ou jornais de sensoriamento remoto. A busca foi realizada na base *Scopus* com palavras chave para cada método conforme mostrado:

- overall accuracy: acurácia total representam 62,7% dos artigos pulicados;
- cohen's kappa: o Kappa de Cohen representa 29,5% desses artigos;
- weighted cohen's kappa: o Kappa de Cohen Ponderado possui 1,07%;
- tau index: índice de Tau possui 1,13% do total;
- Scott's pi: Pi de Scott possui 0,7%;
- Ground truth index: Índice Ground Truth 0,08%.

Os índices Acurácia Total e Kappa foram escolhidos por serem os mais utilizados. Kappa ponderado não foi escolhido devido à complexidade de atribuir pesos a classes, sendo escolhidos os índices Tau e Pi de Scott. Por fim, o índice Pabak também foi utilizado, sendo dentre estes o índice mais recente e portanto pouco difundido na análise de acurácias de uso e cobertura do solo.

2.3.1 Acurácia total

A acurácia total ou acurácia global apresentada por Story e Congalton (1986), tem como objetivo analisar a concordância geral entre os dados de classificação e os de referência através da soma da diagonal principal dividida pelo total de amostras, representada por p_o (Eq. 7).

$$p_o = \frac{\sum_{i=1}^k n_{ii}}{n} \quad (7)$$

p_o também é chamada de porcentagem de concordância por autores como Story, Congalton (1986), Naeset (1996), Ma e Redmond (1995), e de proporção observada por Koukolas, Blackburn (2001), Congalton (2009) e representada pelo somatório à proporção das amostras classificadas corretamente (p_{ii}) (Eq.8).

$$p_o = \sum_{i=1}^k p_{ii} \quad (8)$$

2.3.2 Kappa de Cohen

A análise Kappa foi desenvolvida por Cohen (1960) e é uma técnica multivariada discreta utilizada em avaliação de acurácia (Bishop et al, 1975), mede a proporção de concordância, eliminando-se a chance ao acaso (Cohen, 1960). O resultado gerado pela análise Kappa é a estatística KHAT (\hat{K}) que mede o quão bem os dados de classificação concordam com os dados de referência. Considerando duas matrizes, a estatística permite inferir diferenças entre as matrizes de dois avaliadores, duas datas diferentes, etc. (Congalton, 2009.p.107)

Para o estimador \hat{K} é necessário determinar a proporção da concordância atual ou observada p_o (Eq. 7) e a proporção de concordância casual p_c (Eq.9), ao qual corresponde ao produto das proporções marginais das linhas e colunas dos dados de referência e de classificação. p_c também é chamada de proporção esperada.

$$p_c = \sum_{i=1, j=1}^k p_{i+} p_{+j} \quad (9)$$

Assumindo um modelo de amostragem multinominal, a estimativa de máxima verossimilhança de \hat{K} é dada por:

$$\hat{K} = \frac{p_o - p_c}{1 - p_c} \quad (10)$$

Vale ressaltar que o numerador dessa equação é a diferença entre os dados observados e os esperados (dados de classificação), sendo esta estatística similar à análise da família Qui-quadrado (χ^2).

Os resultados de \hat{K} variam entre -1 e +1, onde -1 corresponde a nenhuma concordância, 0 relaciona-se a concordâncias iguais entre os dados e +1 corresponde a uma perfeita. Autores como Landis e Koch (1977) caracterizaram o nível de concordância \hat{K} em

três intervalos: valor maior que 0,80 representa forte concordância, entre 0,40 e 0,80 representa concordância moderada e menor que 0,40 representa uma concordância fraca.

2.3.3 Pi de Scott

O índice Pi de Scott (Scott's π) desenvolvido por Scott (1955) corrige a concordância esperada através do número de classes e a frequência com que são utilizadas (Scott, 1955). Sua equação é similar ao Kappa (Eq. 29), diferenciando-se no cálculo da concordância esperada obtida através das proporções de cada classe em toda a amostra de unidades do conteúdo (Craig, 1981; Riff, Lacy e Fico, 2014, p.132).

$$\pi = \frac{p_o - p_s}{1 - p_s} \quad (11)$$

A concordância esperada depende do número de classes e da frequência com que cada um deles é utilizado em um dado teste (Scott, 1955; Riffe, Lacy e Fico, 2014, pag. 130). Assumindo que as proporções de concordância são iguais entre as classificações e as unidades nas classes de toda população são conhecidas (Craig, 1981). O valor de p_s é dado por:

$$p_s = \sum_{i=1}^k \left(\frac{p_{i+} + p_{+i}}{2} \right)^2 \quad (12)$$

O índice pode variar de 0 a 1, independentemente do número de classes, permitindo uma comparação com a acurácia total (Craig, 1981). Quanto mais próximo de um, maior é a concordância dos classificadores. A concordância esperada mínima ou incerteza máxima acontece quando todas as classes são utilizadas com igual frequência, e aumenta quando há qualquer desvio de distribuição uniforme de frequências entre as classes (Scott, 1955).

2.3.4 Índice Tau

O coeficiente *Tau* foi desenvolvido por Klecka (1980) e introduzido no sensoriamento remoto por Ma e Redmond (1995). Similar ao Kappa, seu objetivo é analisar o nível de concordância entre os dados da matriz de confusão. (Couto, 2003) afirma que *Tau* mede o progresso da acurácia dos dados de classificação em função do ajuste aleatório de amostras para as classes.

Ma e Redmond (1995) relatam que a diferença crítica entre os *Kappa* e *Tau* está nas probabilidades, enquanto o *Kappa* é baseado na probabilidade a *posteriori*, *Tau* usa a probabilidade a *priori*, ou seja, o *Kappa* depende da matriz de confusão e o *Tau* pode ser calculado antes mesmo da construção da matriz de confusão.

De acordo com Klecka (1980) o coeficiente de *Tau* pode ser calculado através da seguinte equação:

$$T = \frac{p_o - p_r}{1 - p_r} \quad (13)$$

Onde $p_r=1/k$, assim para probabilidades a *priori* de associação iguais entre os grupos:

$$T_e = \bar{\tau} = \frac{p_o - (1/k)}{1 - (1/k)} \quad (14)$$

T_e representa o coeficiente Tau para casos onde as classificações são baseadas em probabilidades iguais de pertencer à classe. Ma e Redmond (1995) afirmam que T_e aproxima da acurácia total (P_o) à medida que o número de classes aumenta.

2.3.5 Pabak

Pabak (*Prevalence – adjusted bias – adjusted kappa*) foi desenvolvido por Byrt, Bishop e Carlin (1993) e ajusta o índice de Kappa para as diferenças entre a proporção total de avaliações erros e acertos e para o viés entre os mapas de referência e classificação. Sendo assim, o índice apresenta resultado mais apurado na classificação da acurácia temática (Avelino, 2018). A equação para seu cálculo é:

$$PABAK = \frac{2 \left(\frac{\sum_{i=1}^k n_{ii}}{n} \right) - p_e}{1 - p_e} \quad (15)$$

Em que $p_e = 0,50$, obtido através das médias das diagonais (Byrt, Bishop e Carlin, 1993). Substituindo na equação (15), tem-se a nova equação,

$$PABAK = \frac{2 \left(\frac{\sum_{i=1}^k n_{ii}}{n} \right) - 0,5}{1 - 0,5} = 2p_o - 1 \quad (16)$$

De acordo com os autores Byrt, Bishop e Carlin (1993), esse índice “é um processo natural de ajustamento para viés e prevalência levando a um índice linearmente relacionado à concordância observada, p_o ”.

3. RESULTADOS

Para as análises estatísticas, os índices de acurácia foram agrupados em amostragens sistemática e aleatória. Foram construídos histogramas de cada acurácia com os três tamanhos de janela (5x5, 20x20 e 25x25 *pixels*) para cada modelo de mapa e número de classes – se encontram ao final de cada tópico – além de tabelas com medidas de média, mediana e desvio padrão – se encontram nos apêndices. Primeiro foi realizado uma análise estatística mapa a mapa e após uma avaliação geral sobre o respectivo índice.

Foram analisadas as proximidades entre a acurácia calculada (durante a validação) e a acurácia atribuída (durante a construção do mapa de classificação). Também foi observada a dispersão dos dados em relação à média.

3.1 Amostragem sistemática

3.1.1 Acurácia total

- Primeiro mapa com quatro classes: o segundo tamanho de janela (20x20) apresentou melhor confiabilidade dentre os três, tendo a melhor proximidade entre as acurácias atribuídas e suas médias, além das médias e das medianas que ficaram próximas entre si, tendo uma diferença menor que 1%. Com o aumento das janelas 5x5 para 20x20, as médias das acurácias atribuídas 0,50 e 0,60 sofreram variações entre as janelas de aproximadamente 5%, em que a primeira reduziu de 0,53978 para 0,57398 e a segunda aumentou de 0,63928 para 0,68412. Já as acurácias 0,85 e 0,95 sofreram menores variações em torno de 3%, respectivamente de 0,88085 para 0,86131 e de 0,93425 para 0,96871. Com o aumento das janelas 20x20 para 25x25, as médias da acurácias reduziram entre 1% a 10% – exceto 0,80 que aumentou em 1% – e os dados ficam mais dispersos da média, fazendo com que as acurácias calculadas ficassem mais distantes das acurácias atribuídas. O aumento do tamanho das janelas não indica um aumento do valor de acurácia, uma vez que aumentando o tamanho da janela está aumentando a área do mapa a ser amostrada e conseqüentemente aumenta-se a chance de captar não somente áreas acuradas mas também não acuradas. O que se percebe é

que valores mais baixos de acurácias como 0,50 e 0,70 tendem a sofrer maiores variações com o aumento do tamanho de janelas.

- Segundo mapa com quatro classes: o primeiro tamanho de janela (5x5) apresentou melhor confiabilidade, tendo a melhor proximidade entre as acurácias atribuídas e suas médias. Entretanto os três tamanhos de janelas apresentaram valores de médias próximas entre esses. Com o aumento das janelas 5x5 para 20x20 quanto de 20x20 para 25x25 as médias não sofreram muitas variações, em que essas variaram entre 0,16% e 2,51%, bem como os desvios padrões que também não sofreram grandes variações com os aumentos das janelas. Contudo, as acurácias 0,50 e 0,70 tendem a sofrer mais variações quando comparadas a 0,85 e 0,95.
- Terceiro mapa com quatro classes: diferentemente do segundo mapa, a acurácia total sofre maiores alterações em suas médias tanto com o primeiro aumento (de 5x5 para 20x20) quanto com o segundo (20x20 para 25x25), em que as médias das acurácias atribuídas aumentaram entre 0,22% e 6,84% no primeiro e reduziram entre 2,12% e 4,41%. Entretanto, na segunda janela houve menores dispersões dos dados em relação à média.
- Primeiro mapa com cinco classes: as proximidades entre as médias e as acurácias atribuídas ocorreram na segunda janela, sendo que a diferença entre as janelas não ultrapassaram 4%. Com o primeiro aumento das janelas, as médias das acurácias 0,50 e 0,70 variaram (respectivamente aumentou 5% e reduziu 14%), enquanto que as médias das acurácias 0,85 e 0,95 mantiveram-se muito próximas entre as janelas. No segundo aumento, as médias das acurácias 0,50 e 0,95 mantiveram-se enquanto que as outras duas aumentaram em torno de 3%.
- Segundo mapa com cinco classes: as acurácias atribuídas 0,50 e 0,70 apresentaram maiores desvios padrões em relação as acurácias 0,85 e 0,95. Com o primeiro aumento de janelas, apenas 0,70 teve sua acurácia média elevada, sendo que as outras mantiveram muito próximas entre as duas janelas. Porém, com o segundo aumento de janelas, todas as acurácias tiveram variações nas médias, 0,50 e 0,70 reduziram e 0,85 e 0,95 aumentaram.
- Terceiro mapa com cinco classes: os dados das acurácias apresentaram maior desvio padrão dentre todos os mapas acima, incluindo os mapas com quatro classes, mesmo que o desvio padrão tenha reduzido com o aumento das janelas, estes ainda apresentam mais distantes da média dentre todos os mapas apresentados até agora.

Com o primeiro aumento das janelas, todas as médias das acurácias aumentaram diferentemente do aumento das janelas 20x20 para 25x25 em que apenas as acurácia 0,50 e 0,85 tiveram as médias elevadas e as outras duas (0,70 e 0,95) reduzidas.

- Primeiro mapa com seis classes: as médias não sofreram grandes variações com o aumento das janelas 5x5 para 20x20, chegando a uma variação máxima de 2% entre as médias dessas janelas. O aumento das janelas 20x20 para 25x25, fez com que todas as acurácias tivessem suas médias elevadas. A dispersão dos dados em relação à média não foi alta, sendo 0,50 a maior dispersão dos dados em relação às outras acurácias – isso para todos três tamanhos de janelas.
- Segundo mapa com seis classes: as acurácias atribuídas 0,50 e 0,70 apresentaram maiores desvio padrão em relação a 0,85 e 0,95. Dentre os três tamanhos de janelas, a primeira (5x5) apresentou melhor aproximação entre as acurácias atribuídas e as médias das acurácias calculadas – exceto 0,50 que ficou distante e, não somente na primeira janela como nas outras duas também. Com o primeiro aumento de janelas, a dispersão dos dados reduziram, porém apenas 0,50 mantiveram os valores médios próximos entre a primeira e a segunda janela. No segundo aumento de janelas, a dispersão dos dados continuaram maiores em 0,50 e 0,70.
- Terceiro mapa com seis classes: nas três janelas as acurácias médias foram próximas entre 0,50 e 0,70 (exceto na última janela) e também entre 0,85 e 0,95. As dispersões reduzem com o aumento das janelas, contudo as acurácias que apresentam médias próximas entre si foram 0,70 e 0,85 na terceira janela e 0,50 e 0,95 na segunda.

Em resumo, a acurácia total não é fortemente influenciada pelo tamanho das janelas, ou seja, pela quantidade de amostras, uma vez que o aumento do tamanho das janelas não ocasionam variações maiores que 5% nos valores de acurácia real (nesse caso, atribuídas). Apesar dos valores de acurácia média sempre apresentarem próximos da acurácia real, esses em grande parte sempre apresentam abaixo da acurácia real e nunca acima (mostrando que aqueles valores que ficaram um pouco maiores na criação dos mapas de classificação não interferiram nas acurácias). Entretanto, este índice é influenciado pelo tipo de mapa, quanto mais as classes são dispersas – ou predominância de uma classe – menor são os valores das acurácias, chegando a quase 25% abaixo da acurácia atribuída. As acurácias 0,50 e 0,70 tendem a sofrer maiores variações quando comparadas com as acurácias 0,75 e 0,85 na variação das janelas e independentemente do tipo de mapa e quantidade de classes.

3.1.2 Kappa

- Primeiro mapa com quatro classes: As médias e medianas foram próximas, variando entre 0,07% a 2,50% para os três tamanhos de janela. Quando há aumento dos tamanhos das janelas em ambos os casos, as médias das acurácias do índice de Kappa sofreram os mesmos tipos de variação que no índice de acurácia total, entretanto com valores diferentes, uma vez as acurácias médias e seus desvios são diferentes entre os dois índices. As maiores proximidades entre médias e as acurácias aconteceram nas acurácias atribuídas de 0,85 e 0,95 para os três tamanhos de janela.
- Segundo mapa com quatro classes: com o aumento das janelas 5x5 para 20x20, as médias das acurácias 0,50 e 0,70 diminuíram, da acurácia 0,85 reduziu e 0,95 permaneceram em torno de 0,91. Com o aumento de 20x20 para 25x25, apenas a média da acurácia 0,95 reduziu enquanto as outras permaneceram muito próximas entre as duas janelas. Apenas 0,95 teve sua média próxima nas três janelas, contudo as melhores proximidades aconteceram para o primeiro e segundo tamanho de janela. Os dados ficam mais dispersos na terceira janela – com exceção de 0,95 – indicando menor confiabilidade entre as acurácias calculadas e atribuídas.
- Terceiro mapa com quatro classes: o aumento do tamanho das janelas 5x5 para 20x20 fez com que todas as médias aumentassem, e com o segundo aumento (de 5x5 para 20x20) os desvios dos dados reduziram para todas as acurácias, apenas 0,50 e 0,70 tiveram suas médias elevadas, 0,85 manteve seu valor (aproximadamente 33%) e 0,95 teve a média reduzida.
- Primeiro mapa com cinco classes: o aumento das janelas 5x5 para 20x20 fez com que a média da acurácia atribuída 0,50 reduzisse e 0,70 aumentasse, e as outras duas acurácias permaneceram suas médias praticamente iguais aos da janela anterior. No segundo aumento (de 20x20 para 25x25), as médias das acurácias 0,50 e 0,95 reduziram e 0,70 e 0,85 aumentaram.
- Segundo mapa com cinco classes: com o primeiro aumento as acurácias médias 0,70 e 0,85 aumentaram, 0,95 reduziu e 0,50 manteve seu valor praticamente o mesmo. No segundo aumento, as médias das acurácias 0,50 e 0,70 reduziram e as outras duas aumentaram em relação à janela anterior. A dispersão dos dados foi mais alta na segunda janela.
- Terceiro mapa com cinco classes: a dispersão dos dados foi alta nos três tamanhos de janelas quando comparada aos outros dois mapas anteriores. O primeiro aumento das

janelas fez com que apenas a média da acurácia 0,95 aumentasse, já no segundo aumento as médias 0,50 e 0,85 elevaram e as outras duas reduziram.

- Primeiro mapa com seis classes: com o aumento das janelas 5x5 para 20x20 as médias das acurácias de 0,85 e 0,95 reduziram, de 0,70 aumentou e de 0,50 permaneceu próxima à média da janela anterior. O aumento das janelas 20x20 para 25x25, fez com que todas as acurácias tivessem suas médias elevadas. A dispersão dos dados não foi alta, sendo 0,50 a maior dispersão dos dados em relação às outras acurácias – isso para todos três tamanhos de janelas.
- Segundo mapa com seis classes: dentre os três tamanhos de janelas, a primeira (5x5) apresentou melhor aproximação entre as acurácias atribuídas e as médias das acurácias calculadas – exceto 0,50 que ficou distante e, não somente na primeira janela como nas outras duas também. Com o primeiro aumento de janelas, a dispersão dos dados reduziram, porém apenas 0,50 mantiveram os valores médios próximos entre a primeira e a segunda janela. No segundo aumento de janelas, a dispersão dos dados continuaram maiores em 0,50 e 0,70.
- Terceiro mapa com seis classes: os desvios padrões reduzem com o aumento das janelas, contudo as acurácias atribuídas que apresentam médias próximas de si foram 0,70 e 0,85 na terceira janela e 0,50 e 0,95 na segunda. Contudo, estes desvios foram mais altos quando comparados aos outros mapas. A acurácia atribuída 0,95 foi a que teve a média mais próxima de si, sendo que ficaram 10% abaixo.

O índice de Kappa sofre variações com tamanho de janelas, sendo a segunda janela melhor dentre as três. Esse índice também sofre variação em função do tipo de mapa, onde a predominância de uma e de duas classes fazem com que os valores de acurácia fiquem muito abaixo da acurácia atribuída (no mínimo 20%). Os melhores valores de acurácia são 0,85 e 0,95, tendo seus valores de acurácia mais próximas da acurácia atribuída.

3.1.3 Pi de Scott

Este índice apresentou características apresentadas anteriormente semelhantes e até mesmo iguais aos da acurácia total e do Kappa em determinados mapas, como pode ser visto a seguir.

Tanto o primeiro e o terceiro mapa com quatro classes apresentaram características iguais ao do índice Kappa – inclusive os valores de média, mediana e desvio padrão foram

muito próximos entre estes dois índices. Já o segundo mapa com quatro classes, as acurácias atribuídas 0,85 e 0,95 tiveram suas médias mais próximas de si, e o primeiro aumento das janelas ocasionou redução da média da acurácia 0,70, aumentou a média da acurácia 0,85 e 0,95 e de 0,50 permaneceu a mesma entre as duas janelas. No aumento de 20x20 para 25x25, todas permaneceram sem aumentos maiores que 1% , exceto 0,95 que reduziu.

O Pi de Scott também teve suas características similares ao Kappa ao primeiro, segundo e terceiro mapas com cinco classes, todavia apenas no primeiro mapa os valores de média, mediana e desvio padrão foram muito próximos entre estes dois.

Os índices Kappa e Pi de Scott também apresentaram similaridades em todos os três mapas com seis classes, mas todos os valores de média, mediana e desvio padrão diferentes entre estes dois.

Esse índice, assim como o Kappa, sofre variações de acurácias com tamanho de janela, tipo de mapa e número de classes, sendo também a janela 20x20 *pixels* o melhor tamanho de janela. Os melhores valores de acurácia são 0,85 e 0,95, tendo seus valores de acurácia mais próximas da acurácia atribuída.

3.1.4 Tau

O índice Tau apresentou similaridades com os índices de Kappa no primeiro e segundo mapa com quatro classes, sendo que no segundo a média da acurácia atribuída 0,50 não mudou com aumento das janelas 5x5 para 20x20, sendo os valores de média, mediana e desvio padrão foram diferentes entre estes dois. O terceiro mapa com quatro classes possui maior dispersão dos dados em relação a média para as acurácias de 0,50 e 0,70 nas três janelas, sendo a maior proximidade entre as médias e as acurácias atribuídas de 0,85 e de 0,95. O primeiro aumento de tamanho de janela fez com que todas as médias aumentassem – exceto 0,70 que manteve sem alterações – e no segundo aumento do tamanho de janela, todos os índices sofreram redução.

As similaridades entre Tau e Kappa permanecem nos três mapas com cinco classes, sendo que os valores de média, mediana e desvio padrão foram diferentes entre estes dois.

O mesmo acontece com os três mapas com seis classes, ou seja, as características entre Kappa e Tau são similares, com a exceção de que no aumento dos tamanhos de janelas 5x5 para 20x20 no segundo mapa, a média da acurácia atribuída 0,50 permaneceu sem alteração entre as janelas. Os valores de média, mediana e desvio padrão foram diferentes entre estes dois.

O Tau também apresenta variações parecidas com do índice Kappa, onde sofre variações das acurácias em função do tamanho de janela, tipo de mapa e número de classes, sendo também a janela 20x20 *pixels* o melhor tamanho de janela. Os melhores valores de acurácia são 0,85 e 0,95, tendo seus valores de acurácia mais próximas da acurácia atribuída. Vale ressaltar que as variações entre a primeira e segunda janela são maiores para os valores de acurácia atribuída 0,50 e 0,70 (podendo chegar a 15%) e menores para os valores das acurácias atribuídas 0,85 e 0,95 (podendo chegar a 7%), isso para os índices Kappa, Pi de Scott e Tau.

3.1.5 Pabak

Pabak apresentou similaridade com a acurácia total no primeiro mapa com quatro classes, e com Tau no terceiro mapa com quatro classes. Já no segundo mapa com quatro classes, o aumento das janelas 5x5 para 20x20 fez com todas as acurácias tivessem as médias elevadas, exceto 0,50 que teve sua média reduzida, e no aumento de 20x20 para 25x25, as mudanças foram similares aos de Kappa.

Nos mapas com cinco classes, houve similaridades entre os índices de Kappa e Pabak nos primeiro e terceiro mapas, e o segundo a similaridade de Pabak ocorreu com Pi de Scott.

Por fim, os três mapas com seis classes apresentaram similaridades com Kappa, exceto o segundo mapa que com o aumento de janelas de 5x5 para 20x20 teve a média da acurácia atribuída 0,50 elevada.

Vale ressaltar que os valores de média, mediana e desvio padrão foram diferente entre todos os índices de Pabak, Tau, Pi de Scott e Kappa.

Pabak também possui similaridade nas variações das acurácias assim como Kappa, Pi de Scott e Tau, entretanto é o índice com maior sensibilidade dentre os cinco, onde seus valores de acurácia chegam a ser negativos quando valores de acurácia atribuídas 0,50. A acurácia mais próxima da atribuída é 0,95, e em alguns casos 0,85 (mapas com predominância de uma classe).

3.2 Amostragem aleatória

3.2.1 Acurácia total

- Primeiro mapa com quatro classes: Os dados apresentaram maiores dispersão nos dados das acurácias 0,50 e 0,70 para os três tamanhos de janelas. O aumento das janelas 5x5 para 20x20 fez com que médias das acurácias 0,70 e 0,85 permanecessem muito próximas entre as janelas, 0,50 diminuísse e 0,95 aumentasse. O segundo aumento fez com que todas as médias das acurácias aumentassem, exceto 0,50 que permaneceu próximo. A dispersão dos dados diminuiu com o tamanho das janelas, apenas 0,70 na janela 25 que aumentou. As médias de 0,85 e 0,95 foram as que ficaram mais próximas nas três janelas. Apenas nas janelas de tamanho 20x20 e 25x25 que 0,50 ficaram próximo. Média e mediana ficaram próximas em no máximo 3% para 0,50 e 0,70 enquanto os outros valores ficaram próximos em um máximo 1%. Idem para Kappa
- Segundo mapa com quatro classes: o tamanho de janela 20x20 apresentou menor desvio padrão dos dados dentre as três. No aumento das janelas de 5x5 para 20x20, a média da acurácia atribuída 0,50 reduziu, 0,70 e 0,85 aumentou e de 0,95 manteve-se. No segundo aumento, apenas 0,70 teve sua média elevada, enquanto as outras se mantiveram próximas entre as janelas.
- Terceiro mapa com quatro classes: no aumento das janelas de 5x5 para 20x20 aumentaram as médias 0,85 e 0,95, reduziu a média de 0,70 e manteve-se de 0,50. No segundo aumento, todos reduziram exceto 0,70 que se manteve.
- Primeiro mapa com cinco classes: no aumento das janelas de 5x5 para 20x20 as médias 0,50 e 0,95 elevaram e as outras duas reduziram. No segundo aumento, ocorreu o contrário, onde 0,50 e 0,95 reduziram e 0,70 e 0,85 aumentaram (0,85 não mais que 2%). A acurácia que mais variou em torno da média foi 0,70 em todas as janelas. Proximidades entre acurácias ficaram até 4% nas três janelas. Melhor proximidade entre as acurácias foram 0,85 (-5%) e 0,95 (-10%)
- Segundo mapa com cinco classes: os maiores desvios padrão foram em 0,50 e 0,70 nas três janelas. O primeiro aumento das janelas fez com que média 0,50 permanecesse, 0,70 e 0,85 aumentassem e 0,95 diminuísse. O segundo aumento fez com que as médias de 0,95 tivesse um leve aumento (menor que 1%) e as outras reduzissem. Com exceção de 0,50, todas as outras acurácias apresentaram proximidades entre as médias e acurácias, nas três janelas.
- Terceiro mapa com cinco classes: O primeiro aumento das janelas fez com que média de todas as acurácias aumentassem. O segundo aumento fez com que as médias de

0,50 e 0,75 aumentassem e as outras duas reduzissem. Melhor aproximação entre média e acurácia foi na segunda janela. Desvio padrão mais alto que dos outros dois mapas com cinco classes.

- Primeiro mapa com seis classes: O primeiro aumento das janelas fez com que média de 0,50 mantivesse e as outras acurácias aumentassem. O segundo aumento fez com que as médias de 0,50 e 0,70 mantiveram e outras duas aumentassem.
- Segundo mapa com seis classes: os menores desvios padrão foram em 0,85 e 0,95. O primeiro aumento das janelas fez com que média de 0,50 e 0,75; 0,85 aumentasse e 0,95 reduzisse. O segundo aumento fez com que as médias de 0,50 mantivesse, 0,70 e 0,85 que reduziu e 0,95 aumentou.
- Terceiro mapa com seis classes: tanto com o primeiro quanto com o segundo aumento dos tamanhos de janelas, todas as médias das acurácias aumentaram. a terceira janela que ficou mais próxima entre media e acurácia 0,85 e 0,95 (10%)

3.2.2 Kappa

- Primeiro mapa com quatro classes: mesmas características que a acurácia total, entretanto com valores de média, desvio padrão, mediana e valores de mudanças diferentes.
- Segundo mapa com quatro classes: o tamanho de janela 20x20 apresentou menor dispersão dos dados dentre as três. No aumento das janelas de 5x5 para 20x20, a média da acurácia atribuída 0,50 reduziu, 0,70 e 0,85 aumentou e de 0,95 manteve-se. No segundo aumento, apenas 0,50 teve sua média reduzida enquanto as outras se mantiveram próximas entre as janelas.
- Terceiro mapa com quatro classes: no aumento das janelas de 5x5 para 20x20 todas as médias, exceto 0,70 que se manteve, elevaram. No segundo aumento, todos reduziram exceto 0,70 que aumentou – no máximo 2%. Todas as acurácias em todas as janelas ficaram muito abaixo da acurácia atribuída tendo um valor de 40% abaixo.
- Primeiro mapa com cinco classes: idem acurácia total
- Segundo mapa com cinco classes: o menor desvio padrão foi em 0,50 nas três janelas. O primeiro aumento das janelas fez com que média 0,50 e 0,95 reduzissem, 0,70 permanecesse e 0,85 aumentasse. O segundo aumento fez com que as médias de 0,50 aumentasse as outras reduzissem. (Melhor aproximação entre média e acurácia foi 0,95 – 10%)

- Terceiro mapa com cinco classes: idem acurácia total. 0,85 mais próxima entre acurácia atribuída e média, porém com 30% de diferença
- Primeiro mapa com seis classes: O primeiro aumento das janelas fez com que média 0,50 e 0,70 mantivessem, 0,85 e 0,95 aumentassem. O segundo aumento fez com que as medias de 0,50 aumentasse as outras reduzissem (0,85 e 0,95 mais próximos das acurácias e das médias porém com 10% diferença.
- Segundo mapa com seis classes: idem acurácia total
- Terceiro mapa com seis classes: idem ao anterior sendo melhor proximidade entre média e acurácia somente 0,95 (20% a menos).

3.2.3 Pi de Scott

Pi de Scott apresentou similaridades com Kappa nos segundo e terceiro mapas com quatro classes, já no primeiro apresentou similaridades com a acurácia total, a única diferença foi no aumento das janelas 5x5 para 20x20 que manteve a média da acurácia atribuída 0,95 e aumentou a 0,70. Nos mapas com cinco classes, as similaridades aconteceram entre Pi de Scott e acurácia total no primeiro mapa, e entre Pi de Scott e Kappa nos segundo e terceiro mapas, contudo o aumento de 20x25 no terceiro mapa fez com que a média da acurácia 0,50 mantivesse. Nos mapas com seis classes as similaridades aconteceram entre Pabak e Tau no primeiro e terceiro mapa, e o segundo apresentou similaridades estatísticas com Pi de Scott.

Em todos eles, os valores de média, desvio padrão, mediana e valores de mudanças diferentes entre os índices de acurácia.

3.2.4 Tau

Tau apresentou similaridades com Pi de Scott nos primeiro mapa com quatro classes, mas no aumento das janelas 5x5 para 20x20 a média da acurácia 0,95 aumentou. O segundo mapa com quatro classes foi idem acurácia total. E o terceiro mapa com quatro classes: no aumento das janelas de 5x5 para 20x20 as médias, de 0,50 e 0,70 reduziram, e de 0,85 e 0,95 elevaram. No segundo aumento, todos reduziram exceto 0,70 que aumentou – no máximo 2%. Melhor aproximação entre média e mediana foi em 0,85 e 0,95 – aproximadamente 10%.

Nos mapas com cinco classes, as similaridades aconteceram entre Tau e acurácia total no primeiro e terceiro mapa. No segundo mapa com cinco classes: no aumento das janelas de 5x5 para 20x20 as médias de 0,50 e 0,95 reduziram, de 0,70 e 0,85 elevaram. No segundo

aumento, todos reduziram exceto 0,95 que aumentou – no máximo 2%. Aproximação entre média e mediana variaram, em que tais aproximações ficam distantes.

Terceiro mapa com cinco classes: idem acurácia total e Kappa.

Por fim, nos mapas com seis classes, Tau apresentou similaridade da acurácia total no primeiro e o terceiro mapas e no segundo Tau foi similar ao Pi de Scott.

Em todos eles, os valores de média, desvio padrão, mediana e valores de mudanças diferentes entre os índices de acurácia.

3.2.5 Pabak

O índice de Pabak apresentou as mesmas características que a acurácia total para todos os mapas com quatro classes, entretanto os valores de média, mediana e desvio padrão foram diferente entre eles. Nos mapas com cinco classes as similaridades aconteceram entre Pabak e acurácia total no primeiro e terceiro mapa, e o segundo apresentou similaridades estatísticas com Tau. Nos mapas com seis classes as similaridades aconteceram entre Pabak e Tau no primeiro e terceiro mapa, e o segundo apresentou similaridades estatísticas com Pi de Scott. Em todos eles, os valores de média, desvio padrão, mediana e valores de mudanças diferentes entre os índices de acurácia.

3.3 Análise de regressão

Inicialmente foi realizada a análise de regressão para cada índice de acurácia desconsiderando o valor da acurácia atribuída, ou seja, foi avaliada como os índices podem ser influenciados pela quantidade de classes, tipo de mapa e tamanho de janela independentemente do valor da acurácia atribuída (real). As comparações são sempre realizadas em função do modelo de mapa 1, com quatro classes e tamanho de janela 5x5 *pixels*. Após foram realizadas o mesmo tipo de análise, porém de acordo com cada valor de acurácia atribuída.

3.3.1 Amostragem sistemática

Analisando separadamente os índices, desconsiderando o valor da acurácia atribuída – separadamente – o índice de acurácia total com quatro classes, modelo de mapa 1 com

tamanho de janela 5x5, a média foi de 76,94% ($0,7694 \pm 0,0092$). Tanto o número de classes 5 quanto 6, a acurácia total tende a reduzir respectivamente em 3,538% ($0,0353 \pm 0,0096$) e em 4,868% ($0,0486 \pm 0,0096$). O modelo de mapa também tende a reduzir o valor desse índice em quase 2%, para os modelos de mapas 2 e 3. Já o tamanho da janela aumenta o valor da acurácia, sendo esse valor de aumento muito próximo entre as janelas 20x20 e 25x25, em que a primeira aumenta em 2,305% ($0,0230 \pm 0,0106$) e a segunda aumenta em 1,999% ($0,01999 \pm 0,0116$), sendo a diferença entre elas apenas de 0,3139%.

O índice Kappa apresentou média de 60,96% ($0,6095 \pm 0,011$) para modelo de mapa 1 com quatro classes e tamanho de janela 5x5. Diferente da acurácia total, o número de classes aumenta o valor de Kappa em 4,957% ($0,0496 \pm 0,0117$) para mapas com 5 classes e em 5,444% ($0,0544 \pm 0,0117$) para mapas com 6 classes. Kappa reduz com o modelo de mapa, sendo a redução maior no modelo de mapa 3 (28,27%) do que no mapa 2 (4,44%). O tamanho da janela tende a aumentar o valor de Kappa em 2,52% para tamanho de janela 20x20 e 3,75% para tamanho de janela 25x25.

A média do índice Tau para modelo de mapa 1 com quatro classes e tamanho de janela 5x5 foi de 69,01% ($0,6901 \pm 0,0114$). Esse índice sofre reduções tanto com o número de classes quanto com o modelo de mapa, sendo que o número de classes reduz a acurácia em torno de 2,5% e o modelo de mapa 3 reduz o valor da acurácia em 11,81% e o modelo de mapa 2 reduz em 1,93%. O tamanho da janela tende a aumentar o valor da acurácia para ambos os tamanhos (20x20 e 25x25), sendo o aumento maior na janela 20x20.

Pi de Scott apresentou acurácia média de 60,12% ($0,6012 \pm 0,0133$). Tanto o número de classes quanto o tamanho das janelas aumentam o valor de acurácia desse índice em aproximadamente 4,5% pra cada uma dessas variáveis, entretanto os modelos de mapa tende a reduzir o valor de acurácia em 5,17% para o modelo de mapa 2 e em 36,59% para o modelo de mapa 3.

Por fim, a acurácia média de Pabak foi de 53,90% ($0,53897 \pm 0,0184$). Tal índice tende a reduzir sua acurácia tanto com o número de classes quanto com os modelos de mapa, sendo a maior redução causada para modelo de mapa 3 e para mapas com 6 classes. Contudo, os tamanhos das janelas aumentam o valor de acurácia desse índice, sendo a segunda janela (20x20) com aumento um pouco maior em relação a janela 25x25.

Analisando os índices separadamente, agora com valores de acurácia isolados, foi observado que de maneira geral a acurácia total mantém maior proximidade entre a acurácia atribuída e a acurácia calculada do que os outros índices, todos os seus valores de acurácia tendem a sofrer maiores variações com número de classes (todas as quatro acurácias – 0,50;

0,70; 0,85 e 0,95 – sofreram redução em relação ao modelo de mapa 1 com quatro classes e janelas 5x5). Já Kappa, sofre reduções dos valores de acurácia em função do modelo de mapa, sofrendo reduções em todos os valores de acurácia. Pi de Scott, assim como Kappa sofre variações de redução com número de classes quanto com modelo de mapa. Tau sofre reduções tanto com número de classes quanto com modelo de mapa. Pabak tende a sofrer reduções em função do número de classes. Em todos os cinco índices, o mapa 3 ocasionou maior redução dos valores de acurácia, as reduções ocasionaram maior impacto no valor da acurácia do que os aumentos, uma vez que os aumentos não ultrapassaram 3% - exceto para Pabak que sofreu maior aumento dentre os índices.

3.3.2 Amostragem aleatória

Analisando separadamente os índices, desconsiderando o valor da acurácia atribuída – separadamente – a média do índice de acurácia total é próxima ao da amostragem sistemática, 76,34% ($0,7634 \pm 0,0095$) sendo a diferença de -0,6% em relação à amostragem sistemática. Nesse tipo de amostragem, o número de classes e o modelo de mapa também tendem a reduzir o valor da acurácia total, em que essas reduções são 1% abaixo em relação à amostragem sistemática. Os tamanhos das janelas aumentam o valor da acurácia total, tendo também o valor de aumento muito próximo ao da amostragem sistemática.

O valor da acurácia média para Kappa foi de 60,0% ($0,60087 \pm 0,0112$), sendo muito próximo ao da amostragem sistemática. Além disso, as variações são as mesmas que da amostragem sistemática, sendo os valores dessas variações 1% menor.

Tau também apresentou similaridades com a primeira amostragem, a média foi próxima sendo menor 0,819% apenas – 68,19% ($0,68191 \pm 0,118$). As variações foram as mesmas que a amostragem sistemática e os valores da redução da acurácia para modelos de mapas 2 e 3 foram muito próximos, contudo as variações das classes 5 e 6 e do tamanho de janelas foram em torno de 1% maiores quando comparadas com a amostragem sistemática.

Pi de Scott apresentou acurácia média de 59,18% ($0,5918 \pm 0,0139$), também bem próxima da amostragem sistemática. As variações foram as mesmas, exceto pelo fato dessas variações serem menores para todas as variáveis quando comparadas com a amostragem sistemática.

Por fim, a acurácia média de Pabak foi de 52,68% ($0,52684 \pm 0,0191$). Valo próximo ao da amostragem sistemática. As variações também foram iguais, onde a maior redução causada para modelo de mapa 3 e para mapas com 6 classes. Contudo, os tamanhos das janelas

aumentam o valor de acurácia desse índice, sendo a terceira janela (25x25) com aumento um pouco maior em relação à janela 20x20.

Analisando os índices separadamente, agora com valores de acurácia isolados, tem-se a mesma conclusão que da amostragem sistemática, diferenciando apenas os valores de acurácias calculadas e variações.

4 CONCLUSÃO

Com o objetivo de analisar a confiabilidade dos índices de acurácias baseados na matriz de confusão, a forma em que o impacto das amostragens pode influenciar no valor de acurácia foi criado mapas de referência e classificação. Na criação dos mapas de classificação foram atribuídos valores de acurácia que ao final foram comparados com os novos valores de acurácia calculados, após varreduras realizadas de acordo com o tamanho das janelas, e que foram aplicados em dois tipos de amostragens: sistemática e aleatória.

Primeiramente vale ressaltar que todas as acurácias de todos os índices apresentaram médias e medianas quase sempre abaixo das acurácias atribuídas, sendo que ao apresentar tais valores acima, estes não ultrapassavam 2% em relação a acurácia atribuída.

O aumento do tamanho das janelas, ou seja, o aumento do número de amostras não indica um aumento do valor de acurácia, uma vez que aumentando o tamanho da janela está aumentando a área do mapa a ser amostrada e conseqüentemente há uma maior chance de captar não somente áreas acuradas, mas também não acuradas. O que se percebe é que valores mais baixos de acurácias como 0,50 e 0,70 tendem a sofrer maiores variações com o aumento do tamanho de janelas. Isto para todos os índices de acurácias e em todos os mapas com 4, 5 e 6 classes.

Em relação ao tipo de mapa, os índices de acurácias podem sofrer variações em seus valores de acurácia real, uma vez que foi observado que a predominância de uma ou duas classes no mapa fazem com que esses valores caiam em no mínimo 15% para acurácias mais altas como 0,85 e 0,95 e 25% para acurácias mais baixas como 0,50 e 0,75. Esses valores dependem do tipo de índice. O número de classes também pode influenciar nas acurácias, em que foi observado que todos os índices tendem a se comportar de maneira mais homogênea com maior número de classes, nesse caso nos mapas de 6 classes.

De maneira geral, os índices não são influenciados diretamente pelo tipo de amostragem, ou seja, entre a amostragem sistemática e a amostragem aleatória com janelas de mesmo tamanho, os índices ficam com uma diferença de até 3% entre as duas amostragens.

Analisando as amostragens separadamente, a amostragem sistemática se mostra levemente melhor quando o tamanho de janela é igual a 20x20 *pixel* e amostragem aleatória se mostra melhor para janelas de tamanho 5x5 *pixels*.

A acurácia total mantém seus valores de acurácias sempre próximos da acurácia real (aqui acurácia atribuída) em que esta sofre maior variação quando o mapa apresenta predominância de uma classe (terceiro mapa com 4, 5 e 6 classes), essa variação é menor nas acurácias mais altas (0,85 e 0,95).

O Kappa não somente sofre variações em mapas com predominância de uma classe, com também quando há predominância de duas classes (primeiro mapa com seis classes e segundo mapa com cinco classes). E independentemente do tipo de mapa e classes, o melhor tamanho de janela para esse índice é 20x20 *pixels*. Os melhores valores de acurácia são 0,85 e 0,95, tendo seus valores de médias mais próximas da acurácia verdadeira.

Os outros três índices: Pi de Scott, Tau e Pabak apresentaram similaridades ora com a acurácia total ora com Kappa, o que se notou é que com o aumento do número de classes essas similaridades tendem a ser ainda maior, exceto pelo fato dos valores de média, mediana e desvio padrão serem diferentes.

Em trabalhos futuros podem ser analisados os outros tipos de amostragens, tamanhos de janela, quantidade de classes, bem como outros tipos de índices de acurácia que utilizam regressão logística, por exemplo.

Referências

ABER, James S.; MARZOLFF, Irene; RIES, Johannes. **Small-Format Aerial Photography: Principles, Techniques and Geoscience Applications**. Amsterdam: Elsevier Science, 2010. 268 p.

ANDERSON, J.R. et al. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. 1. Ed. Washington: United States Government Printing Office. **Geological Survey Professional Paper 964**. 1976. 28 p.

ANDERSON, James R. Land-Use Classification Schemes. **Photogrammetric Engineering**, [s.l.], v. 37, n. 4, p.379-387, abr. 1971.

ARONOFF, Stan. The Minimum Accuracy Value as an Index of Classification Accuracy. **Photogrammetric Engineering & Remote Sensing**, Berkeley, v. 51, n. 1, p.99-111, jan. 1985.

ARONOFF, Stanley. **Remote Sensing for GIS Managers**. Redlands: Esri Press, 2005. 487 p.

- BARRETT, Eric Charles; CURTIS, Leonard Frank. **Introduction to Environmental Remote Sensing**. 4. ed. London: Stanley Thornes, 1999. 457 p.
- BOLSTAD, P. **GIS Fundamentals: A First Text on Geographic Information Systems**. 5th Edition, United States: XanEdu Publishing Inc, 2016, 770 p.
- BRIGGS, Asa; BURKE, Peter. **Uma história social da mídia: De Gutenberg à Internet**. 3. ed. Rio de Janeiro: 2016, Zahar, 423.
- CAMPBELL, J. B.; WYNNE, R.H. **Introduction to Remote Sensing**. 5th Edition, New York: The Guildford Press, 2011, 667 p.
- CHEN, J.; MORA, O. E.; CLARKE, K. C.. Assessing the accuracy and precision of imperfect point clouds for 3d indoor mapping and modeling. **Isprs Annals Of Photogrammetry, Remote Sensing And Spatial Information Sciences**, [s.l.], v. -4/6, p.3-10, 12 set. 2018. Copernicus GmbH. <http://dx.doi.org/10.5194/isprs-annals-iv-4-w6-3-2018>.
- CHUVIECO, Emilio. **Fundamentals of satellite remote sensing: An Environmental Approach**. 2. ed. Boca Raton: Crc Press, 2016. 486 p.
- CONGALTON, Russell G.; GREEN, Kass. **Assessing the Accuracy of Remotely Sensed Data: Principles and Practices**. 3. ed. Boca Raton: Crc Press, 2019. 346 p.
- CONGALTON, Russell G.; GREEN, Kass. **Assessing the Accuracy of Remotely Sensed Data: Principles and Practices**. 2. ed. Boca Raton: Crc Press, 2009. 183 p.
- COSGROVE, Denis; FOX, William L. **Photography and Flight**. London: Reaktion Books, 2010. 148 p.
- COSTA, Hugo et al. Land Cover Mapping from Remotely Sensed and Auxiliary Data for Harmonized Official Statistics. **Isprs International Journal Of Geo-information**, [s.l.], v. 7, n. 4, p.157-178, 21 abr. 2018. MDPI AG. <http://dx.doi.org/10.3390/ijgi7040157>.
- COUTO, Paula. Assessing the accuracy of spatial simulation models. **Ecological Modelling**, [s.l.], v. 167, n. 1-2, p.181-198, set. 2003. Elsevier BV. [http://dx.doi.org/10.1016/s0304-3800\(03\)00176-5](http://dx.doi.org/10.1016/s0304-3800(03)00176-5).
- COUTO, Paula. Assessing the accuracy of spatial simulation models. **Ecological Modelling**, [s.l.], v. 167, n. 1-2, p.181-198, set. 2003. Elsevier BV. [http://dx.doi.org/10.1016/s0304-3800\(03\)00176-5](http://dx.doi.org/10.1016/s0304-3800(03)00176-5).
- FITZGERALD, R.w.; LEES, B.g.. Assessing the classification accuracy of multisource remote sensing data. **Remote Sensing Of Environment**, [s.l.], v. 47, n. 3, p.362-368, mar. 1994. Elsevier BV. [http://dx.doi.org/10.1016/0034-4257\(94\)90103-1](http://dx.doi.org/10.1016/0034-4257(94)90103-1).
- FITZPATRICK-LINS, Katherine. Comparison of sampling procedures and data analysis for a land-use and land-cover map. **Photogrammetric Engineering & Remote Sensing**, Reston, v. 3, n. 47, p.343-351, mar. 1981.
- GIRARD, Colette M.; GIRARD, Michel-claude. **Processing of Remote Sensing Data**. Paris: Crc Press, 2003. 508 p.
- HAACK, Barry; MAHABIR, Ron; KERKERING, John. Remote sensing-derived national land cover land use maps: a comparison for Malawi. **Geocarto International**, [s.l.], v. 30, n. 3, p.270-292, 4 set. 2014. Informa UK Limited. <http://dx.doi.org/10.1080/10106049.2014.952355>.
- HIRSCH, Robert. **Seizing the Light: A Soical History of Photography**. 2. ed. New York: Mcgraw-hill Higher Education, 2008. 496 p.

JENSEN, John R. **Introductory Digital Image Processing: A Remote Sensing Perspective**. 4. ed. Glenview: Pearson Education, 2015. 544 p.

JENSEN, John R. **Introductory Digital Image Processing: A Remote Sensing Perspective**. 1. ed. Glenview: Pearson Education, 1986. 379 p.

KHORRAM, Siamak et al. **Remote Sensing**. New York: Springer, 2012. 134 p.

LI, Deren; SHEN, Xin; WANG, Le. Connected Geomatics in the big data era. **International Journal Of Digital Earth**, [s.l.], v. 11, n. 2, p.139-153, 11 abr. 2017. Informa UK Limited. <http://dx.doi.org/10.1080/17538947.2017.1311953>.

LI, Wei et al. Land-use and land-cover change carbon emissions between 1901 and 2012 constrained by biomass observations. **Biogeosciences**, [s.l.], v. 14, n. 22, p.5053-5067, 14 nov. 2017. Copernicus GmbH. <http://dx.doi.org/10.5194/bg-14-5053-2017>.

LIU, Jie; DUAN, Zhao. Quantitative Assessment of Landslide Susceptibility Comparing Statistical Index, Index of Entropy, and Weights of Evidence in the Shangnan Area, China. **Entropy**, [s.l.], v. 20, n. 11, p.868-890, 10 nov. 2018. MDPI AG. <http://dx.doi.org/10.3390/e20110868>.

LYONS, Mitchell B. et al. A comparison of resampling methods for remote sensing classification and accuracy assessment. **Remote Sensing Of Environment**, [s.l.], v. 208, p.145-153, abr. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2018.02.026>.

MA, Lei et al. Deep learning in remote sensing applications: A meta-analysis and review. **Isprs Journal Of Photogrammetry And Remote Sensing**, [s.l.], v. 152, p.166-177, jun. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.isprsjprs.2019.04.015>.

MASRON, Tarmiji; OSMAN, Mohd Azam. **Spatial Information in Local Knowledg**. Malaysia: Penerbit Usm, 2015. 95 p.

OMBER, Alexis et al. Spatial analysis of remote sensing image classification accuracy. **Remote Sensing Of Environment**, [s.l.], v. 127, p.237-246, dez. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2012.09.005>.

OTTOSEN, Thor-bjørn; LOMMEN, Suzanne T.e.; SKJØTH, Carsten Ambelas. Remote sensing of cropping practice in Northern Italy using time-series from Sentinel-2. **Computers And Electronics In Agriculture**, [s.l.], v. 157, p.232-238, fev. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.compag.2018.12.031>.

POWELL, R.I et al. Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon. **Remote Sensing Of Environment**, [s.l.], v. 90, n. 2, p.221-234, mar. 2004. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2003.12.007>.

REES, W. G.. **Physical Principles of Remote Sensing**. 3. ed. Boston: Cambridge University Press, 2013. 492 p.

REIS, Aliny Aparecida dos et al. Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods. **New Zealand Journal Of Forestry Science**, [s.l.], v. 48, n. 1, p.1-17, 16 jan. 2018. Scion. <http://dx.doi.org/10.1186/s40490-017-0108-0>.

RICHARDS, John A. **Remote Sensing Digital Image Analysis: An Introduction**. Berlin: Springer Science, 2013. 281 p.

SENSEMAN, Gary M; BAGLEY, Calvin F; A TWEDDALE, Scott. Accuracy Assessment of the Discrete Classification of Remotely-Sensed Digital Data for Landcover Mapping. **Usacerl Technical Report En-95/04**, [s.l.], v. s/ n., p.1-27, abr. 1995.

- SOKAL, R. R.. Classification: Purposes, Principles, Progress, Prospects. **Science**, [s.l.], v. 185, n. 4157, p.1115-1123, 27 set. 1974. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.185.4157.1115>.
- SOUTH, Stephen; QI, Jiaguo; LUSCH, David P. Optimal classification methods for mapping agricultural tillage practices. **Remote Sensing Of Environment**, [s.l.], v. 91, n. 1, p.90-97, maio 2004. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2004.03.001>.
- STEHMAN, Stephen V.; FOODY, Giles M. Key issues in rigorous accuracy assessment of land cover products. **Remote Sensing Of Environment**, [s.l.], v. 231, p.1-23, set. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.rse.2019.05.018>.
- THENKABAIL, Prasad S.. **Remotely Sensed Data Characterization, Classification, and Accuracies**. Boca Raton: Crc Press, 2015. 712 p.
- TURNER, Monica G. et al. Landscape Ecology: What Is the State of the Science?. **Annual Review Of Ecology, Evolution, And Systematics**, [s.l.], v. 36, n. 1, p.319-344, dez. 2005. Annual Reviews. <http://dx.doi.org/10.1146/annurev.ecolsys.36.102003.152614>.
- VÁZQUEZ, M.; PALLÉ, E.; RODRÍGUEZ, P. Montañas. **The Earth as a Distant Planet: A Rosetta Stone for the Search of Earth-Like Worlds**. New York: Springer, 2010. 422 p.
- WHITEHEAD, Ken; HUGENHOLTZ, Chris H.. Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: a review of progress and challenges. **Journal Of Unmanned Vehicle Systems**, [s.l.], v. 02, n. 03, p.69-85, set. 2014. Canadian Science Publishing. <http://dx.doi.org/10.1139/juvs-2014-0006>.
- XIA, Nan; CHENG, Liang; LI, Manchun. Mapping Urban Areas Using a Combination of Remote Sensing and Geolocation Data. **Remote Sensing**, [s.l.], v. 11, n. 12, p.1470-1493, 21 jun. 2019. MDPI AG. <http://dx.doi.org/10.3390/rs11121470>.