



QUALIDADE DE DADOS GEOESPACIAIS: análise da consistência topológica no mapeamento de cobertura da terra do município de Curitiba - PR

GEOSPATIAL DATA QUALITY: consistency a topological analysis in the coverage land mapping of Curitiba city – PR

CALIDAD DE DATOS GEOESPACIALES: análisis de la consistencia topológica en el mapeamento de cobertura de la tierra del municipio de Curitiba – PR

Manoella Barros Pedreira Ferreira  
Universidade Federal do Paraná – UFPR
E-mail: <manoella.bpf@gmail.com>.

Otacilio Lopes de Souza da Paz  
Universidade Federal do Paraná – UFPR
E-mail: <otacilio.paz@gmail.com>.

RESUMO

Devido ao constante aprimoramento das técnicas de tratamento de dados geoespaciais são inúmeras as pesquisas e aplicações dessas técnicas nas mais variadas áreas do conhecimento. Porém, a necessidade de validação desses dados tem recebido pouca atenção, principalmente no meio acadêmico, pondo em risco a precisão e confiabilidade de muitos estudos. Dentre os parâmetros utilizados para o controle da qualidade de dados geográficos, o presente trabalho teve por objetivo a verificação da consistência topológica de dados referentes ao mapeamento de cobertura da terra do município de Curitiba/PR, na escala original de 1:35.000. Este parâmetro se distinguiu dos demais pela necessidade de quantificação da área de cada classe mapeada. Constatou-se que apesar desse mapeamento ter sido de uma pequena extensão, em uma escala considerada de detalhe, os dados apresentaram inconsistência em sua topologia. Os erros de topologia apontam para possíveis divergências na quantificação das áreas se comparados ao seu tamanho real, especialmente por conta de pequenos polígonos não classificados ou polígonos que extrapolam a área de classificação. Cabe ressaltar que, como as informações espaciais apresentadas em um trabalho científico podem servir de base a outros, é fundamental ter atenção na confiabilidade dos dados, para que possíveis erros não sejam replicados.

Palavras-chave: SIG. Cartografia. Topologia.

ABSTRACT

Due to constant improvement of geographic data processing techniques there are many research and applications of these techniques in various areas of knowledge. However, the need for validation of these data has received little attention, especially in academia, endangering the accuracy and trustworthiness of studies. Among the parameters used for the quality control of geographic data, this study aimed to verify the topological consistency of data related to land cover mapping of city Curitiba, in Paraná State, Brazil, in the original scale of 1: 35,000. This parameter is distinguished from the other by the need to quantify the area of each mapped class. It was found that despite this mapping have been performed on a scale considered detail, the data showed topological inconsistency. Even if these errors can be insignificant, they point to possible differences in the quantification of the areas compared to its actual size, especially for petty unclassified polygons or polygons that exceed the sorting area. Finally, it is worth noting that as the spatial information presented in a scientific work can be the basis for others, it is critical to have attention on the reliability of the data so that potential errors are not replicated.

Keywords: GIS. Cartography. Topology.

RESUMEN

Debido al constante perfeccionamiento de las técnicas de tratamiento de datos geográficos son innumerables las investigaciones y aplicaciones de esas técnicas en las más variadas áreas del conocimiento. Sin embargo, la necesidad de validación de estos datos ha recibido poca atención, principalmente en el medio académico, poniendo en riesgo la precisión y confiabilidad de muchos estudios. Entre los parámetros utilizados para el control de la calidad de datos geográficos, el presente trabajo tuvo por objetivo la verificación de la consistencia topológica de datos referentes al mapeamiento de cobertura de la tierra del municipio de Curitiba / PR, en la escala original de 1: 35.000. Este parámetro se distinguió de los demás por la necesidad de cuantificación del área de cada clase mapeada. Se constató que a pesar de ese mapeamiento haber sido de una pequeña extensión, en una escala considerada de detalle, los datos presentaron inconsistencia en su topología. Aunque estos errores pueden ser irrisorios, apuntan a posibles divergencias en la cuantificación de las áreas si se compara con su tamaño real, especialmente por pequeños polígonos no clasificados o polígonos que extrapolan el área de clasificación. Cabe resaltar que, como las informaciones espaciales presentadas en un trabajo científico pueden servir de base a otros, es fundamental tener atención en la confiabilidad de los datos, para que posibles errores no sean replicados.

Palabras clave: SIG. Cartografía. Topología.

INTRODUÇÃO

Em meio à necessidade de compreensão do espaço geográfico e de suas representações, as técnicas para manipulação de dados geoespaciais têm sido bastante aprimoradas ao longo dos últimos anos. Impulsionadas, sobretudo, pela necessidade de criação e aperfeiçoamento de mecanismos que auxiliassem as tomadas de decisão sobre o território (CÂMARA *et al.*, 2001; ROCHA, 2003; ROSA, 2005; LONGLEY, 2009).

Dado geoespacial pode ser entendido como qualquer tipo de dado (conjunto de informações) que descreva fenômenos aos quais esteja associada uma dimensão espacial, em um determinado tempo (ROSA, 2005; LONGLEY, 2009). Veregin (1999) defende que as bases de dados geoespaciais não devem ser espelhos da realidade e que todas as informações acerca do processo de produção dos dados geoespaciais devem ser relatados e comunicados aos usuários dos dados. Isso para que os mesmos avaliem se determinada base de dados é adequada a aplicação pretendida.

A automatização do processamento de dados geoespaciais, com técnicas computacionais aliadas às técnicas matemáticas para tratamento de processos que ocorrem no espaço geográfico, são fornecidas pelos Sistemas de Informações Geográficas (SIG). Os SIGs, são descritos por Câmara *et al.* (2001) como ferramentas computacionais que representam o aparato tecnológico que possibilita a realização de análises complexas. Segundo o autor, integram dados de diversas fontes e permitem criar bancos de dados geográficos, o que possibilita a automatização da produção de documentos cartográficos.

Os métodos de análise espacial e estatística são importantes (ROGERSON, 2012). A qualidade da pesquisa científica, quando da manipulação de dados geográficos, demanda o tratamento desses dados a fim de que esses sejam validados e os erros corrigidos ou, ao menos, reduzidos.

Em decorrência do constante aprimoramento das técnicas de tratamento de dados geográficos, tem-se hoje um grande número de trabalhos que utilizam essas técnicas nas mais variadas áreas do conhecimento. Porém, a necessidade validação desses dados tem recebido pouca atenção, pondo em risco a precisão e confiabilidade de muitos estudos acadêmicos.

Os dados geoespaciais são uma generalização ou aproximação da realidade, muitas vezes podem ser incertos e imprecisos. Por esse motivo, “os estudos das incertezas dos

dados geográficos, suas medidas, modelamento e a análise da propagação dos erros são sem dúvida horizontes de pesquisa bastante atuais dentro de um ambiente de SIG” (SILVA, 2003, p. 33).

A importância do controle de qualidade nos SIGs é discutida por Barros e Carneiro (2013), apresentando como foco a certificação de imóveis rurais e processos de vistoria. Na prática, os dados geoespaciais utilizados na pesquisa desenvolvida pelos autores podem servir para regularização de determinada propriedade, utilizando, por exemplo, os limites espaciais da hidrografia para verificar se a propriedade está dentro dele, e se existe a mata necessária na intersecção entre a propriedade e a hidrografia.

Sem a validação dos dados geoespaciais, se estes estiverem incorretos “os dados vetoriais podem parecer visualmente perfeitos, mas o polígono pode ser armazenado no sistema de forma incorreta, por exemplo, como uma poligonal aberta, sem área associada” (BARROS e CARNEIRO, 2013, p.275). Santos Jr. e Ribeiro (2012) apontam que na validação dos dados geoespaciais e qualidade do processo de produção cartográfica a qualidade posicional tem sido a principal preocupação.

De acordo com Veregin (1999), a preocupação com a qualidade dos dados geoespaciais aumentou nos últimos anos por três motivos: 1) aumento da participação do setor privado na produção de dados, não sendo obrigados a seguir padrões de qualidade; 2) aumento do uso de SIG no apoio à tomada de decisão; 3) aumento na confiança na fonte de dados secundários, influenciado pela redução no acesso de dados em função dos custos.

Os dados geoespaciais possuem propriedades geométricas e topológicas. Não é objetivo deste trabalho discuti-los, apenas destaca-se que dentro de tais propriedades alguns parâmetros devem ser considerados para o controle da qualidade dos dados geográficos.

Os metadados (informações que descrevem os dados) de qualidade de maior evidência compreendem a “linhagem, a acurácia, a consistência lógica e a completeza” (WEBER et al., 1999, p. 10). Cabe destacar, segundo Santos Jr. e Ribeiro (2012), que a norma ISO 19.113 (2000) define como elementos para a avaliação da qualidade de dados geoespaciais a acurácia posicional ou geométrica, metadados, consistência lógica, acurácia temporal, qualidade temática e completude ().

Segundo Davis (2001, p. 95), a topologia é o conceito mais usual de estrutura de dados em SIG. Este autor afirma que as definições para o termo topologia são variadas,

porém, de maneira simples pode-se considerar em associação duas perspectivas. Na primeira delas, o termo é definido como a matemática da conectividade e adjacência para características espaciais. Em segundo lugar, a topologia pode ser vista como a programação que fornece relações espaciais entre locais de conexão e junção de dados, entre redes e características geográficas. Em um SIG, isto significa que a topologia é uma estrutura especial que estabelece ligações entre os nós e as redes a fim de reconhecer as relações espaciais entre as características geográficas.

A topologia, em ambiente SIG, foi desenvolvida para fornecer um método automático rigoroso para verificação dos dados e limpeza dos erros de entrada (THEOBALD, 2001). De acordo com a ESRI (empresa desenvolvedora de *softwares* de SIG), os erros em dados geoespaciais significam necessariamente que as informações, produtos ou decisões obtidas desses dados são imprecisos. Em muitos casos, trabalhar com dados geoespaciais não confiáveis pode acarretar sérias consequências financeiras, ecológicas ou, na pior das hipóteses, risco de vida.

De acordo com Theobald (2001), a topologia garante uma solução automatizada, baseada em regras, para a manutenção da integridade espacial do dado em SIG. O autor ressalta que os polígonos no formato *shapefile* (nativo do sistema ArcGIS) podem conter uma ou mais partes desunidas e/ou sobrepostas.

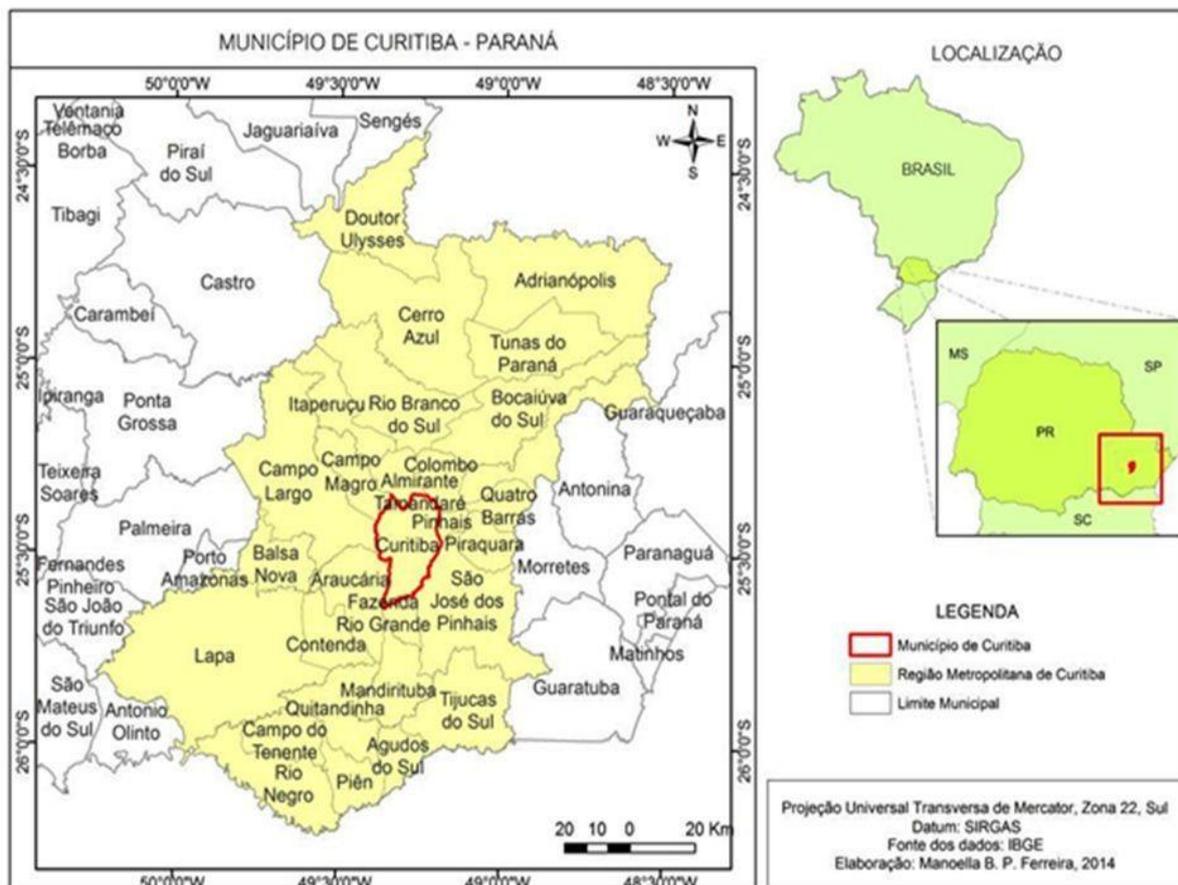
São exemplos de problemas na topologia dos dados geoespaciais: sobreposição de polígonos, vazios entre polígonos, cruzamento de linhas em nós, entre outros. O sucesso no processamento em SIG é condicionado pela análise da consistência topológica dos dados. Veregin (2005) defende que a análise de consistência topológica deve ser feita antes da liberação dos dados ao usuário.

Nesse contexto, o presente trabalho teve por objetivo a verificação da consistência topológica de dados geoespaciais referentes à classificação de cobertura da terra do município de Curitiba/PR, apresentados como resultado da pesquisa de Ferreira (2015). Optou-se pela análise de topologia devido à necessidade de quantificação da área das classes mapeadas.

ÁREA DE ESTUDO

O município de Curitiba, capital do Estado do Paraná, está situado ao centro-sul do primeiro planalto paranaense, com uma altitude de 945 m e clima subtropical, é a principal cidade da Região Metropolitana de Curitiba, formada por 29 municípios (Figura 1). O município possui uma área de 434,967 km², e uma população total de 1.746.898 habitantes, sendo a 8ª cidade mais populosa do Brasil (IBGE, 2010), encontrando-se dividido em 75 bairros. Administrativamente o município está subdividido em nove regionais, espécies de subprefeituras encarregadas dos bairros de cada região, são elas: Bairro Novo, Boa Vista, Boqueirão, Cajuru, CIC, Fazendinha/Portão, Matriz, Pinheirinho e Santa Felicidade.

FIGURA 1 - Mapa de localização do município de Curitiba/PR

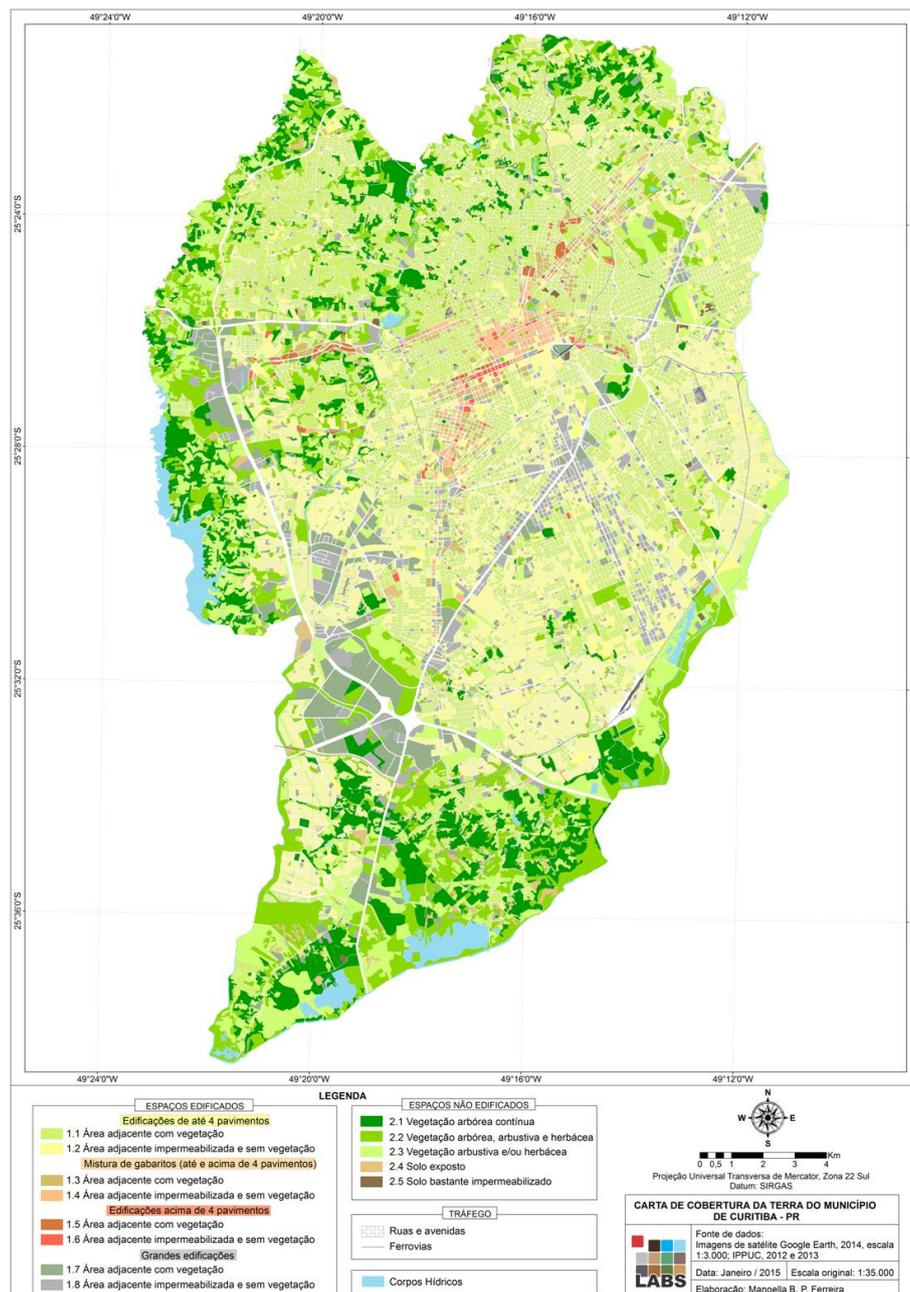


FONTE: IPPUC (2014). Org.: Autores (2017).

PROCEDIMENTOS METODOLÓGICOS

A análise topológica proposta no presente trabalho foi realizada em dados de classificação da cobertura da terra do município de Curitiba-PR apresentados em Ferreira (2015) (Figura 2).

FIGURA 2 - Carta da cobertura da terra do município de Curitiba, Paraná



FONTE: Ferreira (2015).

A classificação foi gerada de forma não automática (manual), em um *software* SIG, em dados de arruamento (limites das quadras) do município de Curitiba disponibilizados pelo Instituto de Pesquisa e Planejamento Urbano de Curitiba (IPPUC), atualizados nos anos de 2012 e 2013. Como base visual para a interpretação da paisagem foram utilizadas imagens de satélite do *Google Earth*, na escala aproximada de 1:3.000. Foi definida como sistema de projeção dos dados finais do mapeamento a Projeção Transversa de Mercator, Fuso 22, sul, com o *Datum* SIRGAS 2000.

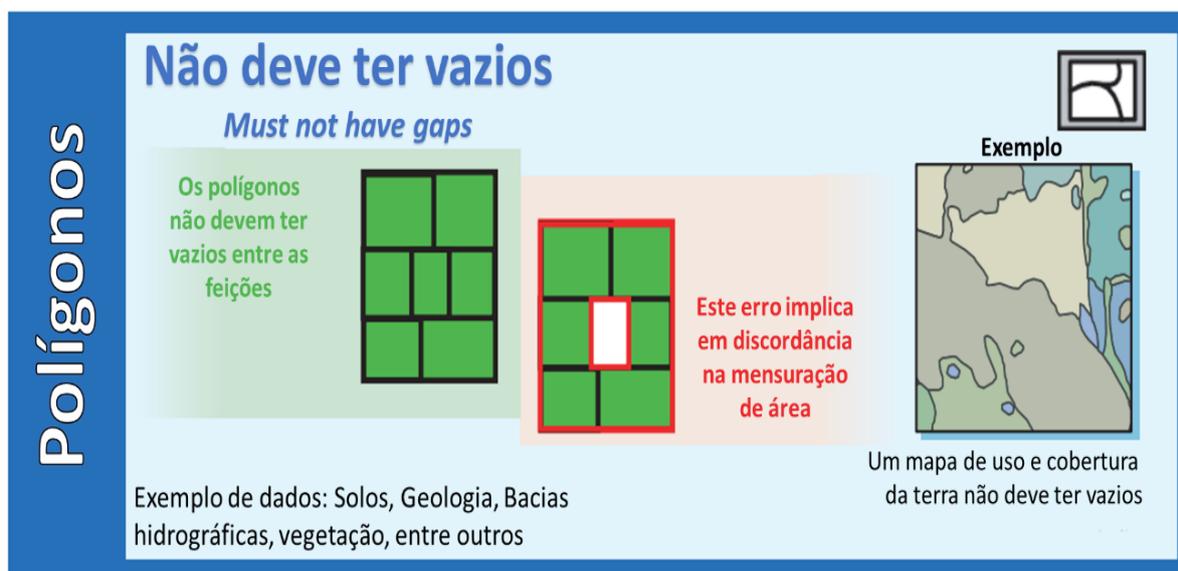
A análise de topologia foi realizada por meio de um SIG, o *software* ArcGIS 10.1, nível ArcInfo. Esse *software* disponibiliza a ferramenta para a análise topológica, a qual por meio de regras matemáticas de topologia permite a modelagem da relação espacial entre os polígonos, de acordo com a necessidade do usuário. Para a análise proposta no presente trabalho, foram destacadas as análises topológicas de sobreposição – *overlap* (Figura 3) e de vazios – *gaps* (Figura 4).

FIGURA 3 – Análise topológica no ArcGIS – análise de sobreposição (*overlap*)



FONTE: ArcGIS. Org.: Autores (2017).

FIGURA 4 – Análise topológica no ArcGIS – análise de vazios (*gaps*)



FONTE: ArcGIS. Org.: Autores (2017).

Sobre o dado de entrada, de classificação de cobertura da terra, executou-se a análise topológica pela *regra Must Not Overlap*, a qual é utilizada para mapeamento de todos os polígonos que possam estar sobrepostos.

Posteriormente, realizou-se a análise topológica pela regra *Must Not Have Gaps* para mapeamento de todas as áreas não classificadas, vazias. Após a validação das regras de topologia executadas na análise, gerou-se um novo dado, uma camada contendo os erros apresentados pelo dado de entrada. Esses erros topológicos são representados, automaticamente, por polígonos, linhas e pontos.

Utilizando esses resultados, foram gerados dois mapas onde as falhas topológicas foram sobrepostas ao *shapefile* onde estão especializadas as quadras do município de Curitiba para melhor visualização dos erros. Os erros foram corrigidos por meio da ferramenta de edição vetorial do ArcGIS.

RESULTADOS E DISCUSSÃO

A análise topológica pelas regras *Must Not Overlap* e *Must Not Have Gaps*, permitiu a identificação de vários polígonos com erros de consistência. No total, foram mapeados 614 erros (Quadro 1). Não foram encontrados erros de polígonos duplicados.

QUADRO 1 – Total de erros apresentados na aplicação de regras de topologia

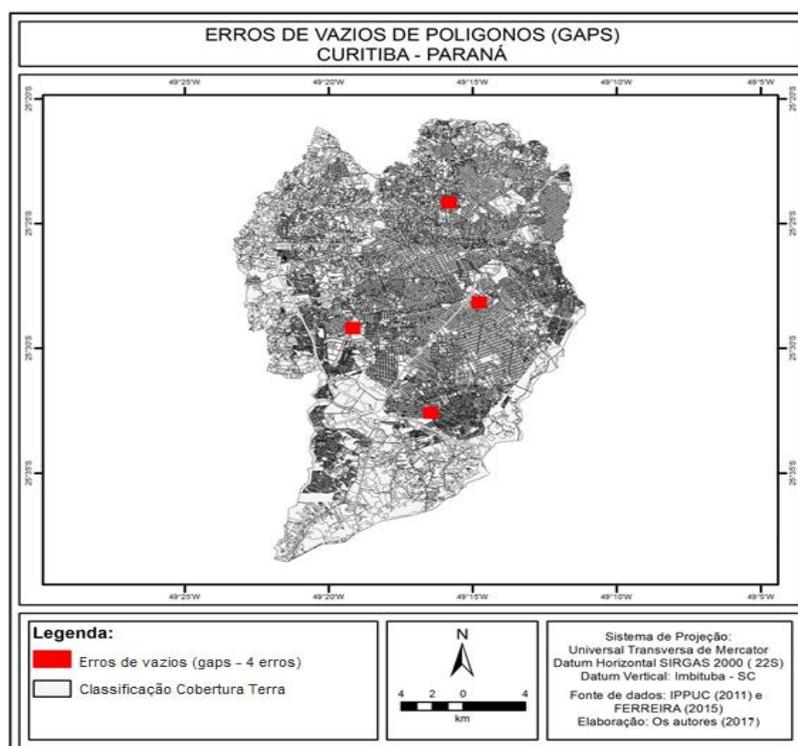
Regras de topologia	Total de erros
<i>Must not overlap</i>	4
<i>Must not have gaps</i>	612
Total	614

FONTE: Os autores (2017).

Esses erros são equivalentes à classificação da cobertura da terra urbana, em uma escala considerada de detalhe. Os dados geospaciais que expressam escalas grosseiras, mesmo a partir das escalas a nível municipal, carecem de análises mais aprofundadas que garantam a qualidade da informação como um todo.

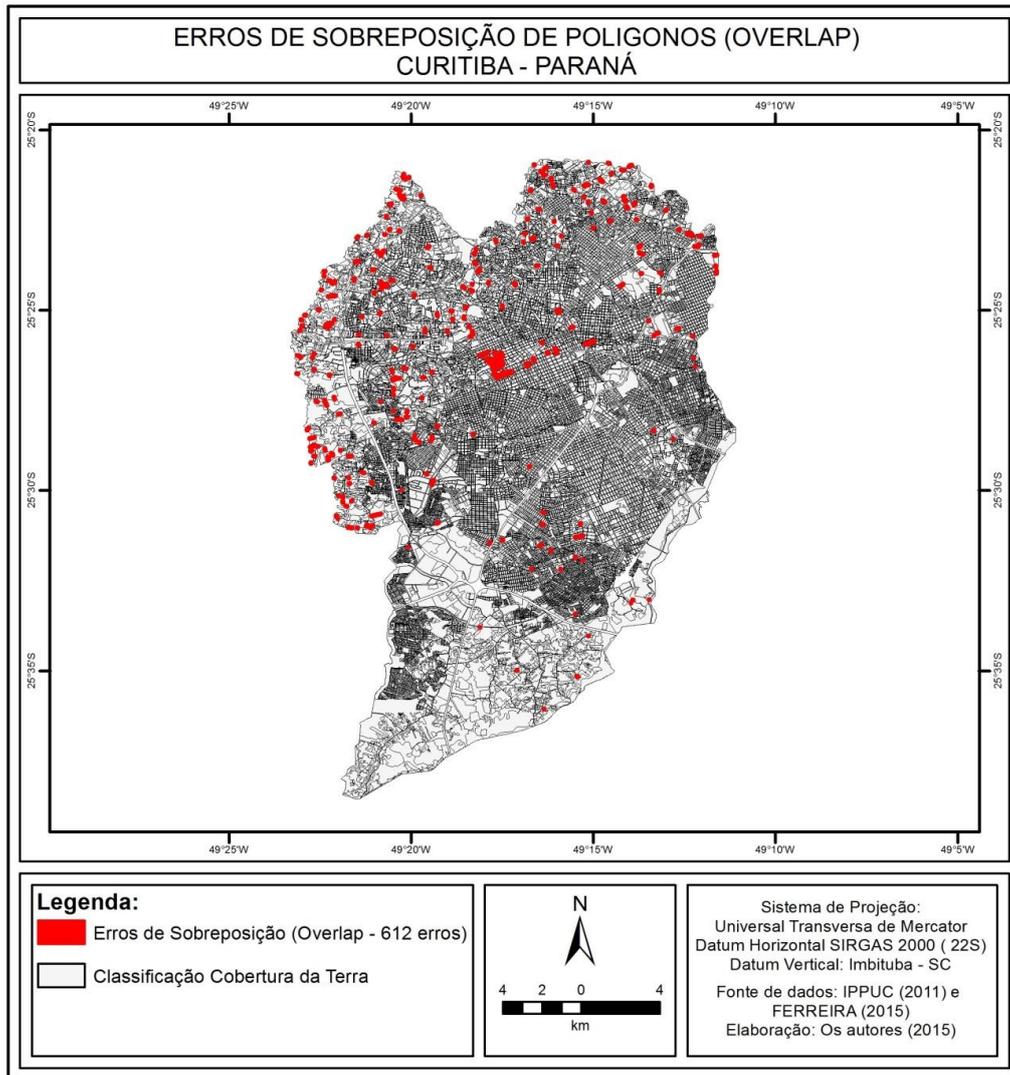
A espacialização dos erros expostos no Quadro 1 está representada nos mapas de análise topológica (Figuras 5 e 6). Por meio desses mapas é possível a visualização da distribuição dos erros ao longo de toda a área classificada, o que ressalta a importância da topologia dos dados espaciais.

FIGURA 5 - Mapa da análise topológica *must not have gaps* - vazios entre os polígonos



Fonte: Autores (2017).

FIGURA 6 - Mapa da análise topológica *must not overlap* – sobreposição entre os polígonos



Fonte: Autores (2017).

É possível verificar que os erros de topologia estão concentrados majoritariamente ao norte do município, com destaque à sobreposição de polígonos, principalmente nos vértices que fazem parte de limites com outros polígonos. Já os erros de vazios, ou áreas não classificadas, não foram expressivos, contando apenas com quatro ocorrências em toda a área mapeada. A quantificação da área de cada tema mapeada está representada nos Quadros 1 e 2. Optou-se por apresentar os valores das áreas antes e depois da verificação de topologia.

QUADRO 2 – Quantificação de área total de cada classe antes e após a correção topológica

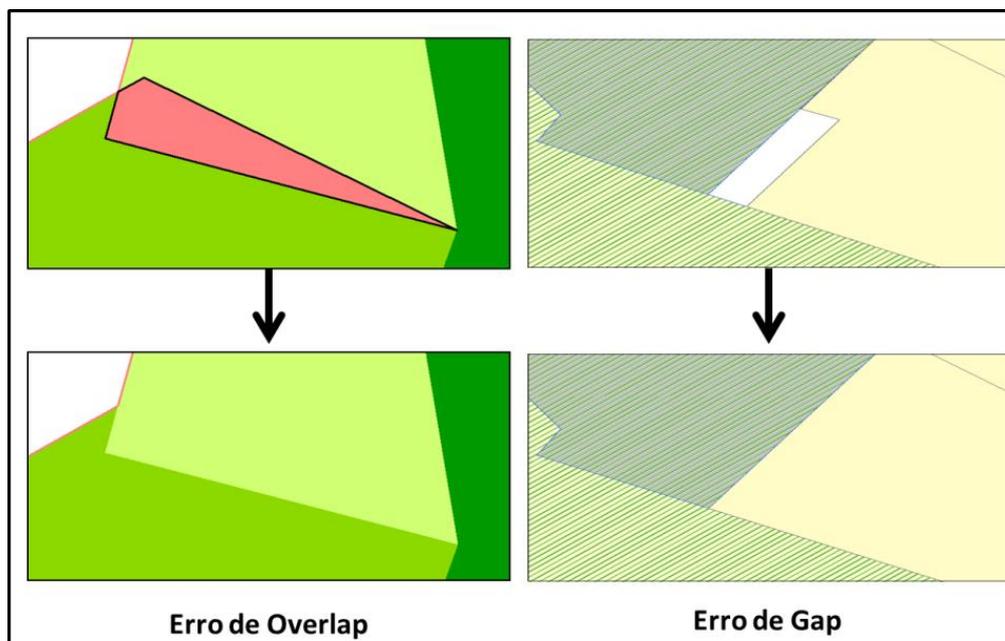
Antes da correção topológica			Depois da correção topológica		
Tema	Área km ²	Porcentagem %	Tema	Área km ²	Porcentagem %
1.1	104,90	24,12	1.1	102,26	23,51
1.2	60,62	13,94	1.2	60,76	13,97
1.3	4,34	1,00	1.3	4,42	1,02
1.4	3,71	0,85	1.4	3,69	0,85
1.5	1,09	0,25	1.5	1,08	0,25
1.6	1,01	0,23	1.6	0,98	0,23
1.7	19,16	4,41	1.7	19,16	4,41
1.8	18,59	4,27	1.8	18,58	4,27
2.1	41,48	9,54	2.1	40,97	9,42
2.2	65,93	15,16	2.2	51,00	11,73
2.3	44,58	10,25	2.3	43,43	9,98
2.4	3,94	0,91	2.4	3,93	0,90
2.5	0,89	0,20	2.5	0,87	0,20
Corpos d'água	6,34	1,45	Corpos d'água	6,34	1,46
Sistema viário	58,38	13,42	Sistema viário	77,49	17,80
Total	434,96	100,00	Total	434,96	100,00

FONTE: Autores (2017).

A quantificação de área total de cada classe antes e após a correção topológica, apresentada no quadro acima, demonstra que algumas classes apresentaram alterações em sua área total. Esses erros de topologia podem ser melhor visualizados nas Figuras 5 e 6, onde estão espacializados em mapas.

Vale ressaltar que, para a análise de topologia realizada no presente trabalho, foram utilizados dados de classificação de paisagem urbana, especificamente cobertura do solo. Sobre esse dado já foram realizadas diversas análises, como a quantificação de cada classe mapeada. Maiores detalhes acerca do mapeamento das unidades de paisagem do município de Curitiba-PR podem ser consultados em Ferreira (2015). Exemplos de erros presentes nessa classificação podem ser visualizados na Figura 7.

FIGURA 7 - Erros de sobreposição e de vazios encontrados no mapeamento da cobertura da terra de Curitiba, Paraná



Fonte: Os autores (2017).

Nos estudos de paisagem, o mapeamento do uso e/ou cobertura da terra urbana e a consistência topológica são relevantes, pois o dimensionamento de cada área mapeada constitui-se, por vezes, em um dos pressupostos da análise em questão. Erros de topologia podem significar que haverá grande discrepância na quantificação de cada unidade classificada, unicamente pelo fato de a área mapeada ser de grande extensão.

CONCLUSÕES

Na análise topológica da classificação da cobertura da terra do município de Curitiba/PR constatou-se que apesar dessa classificação ter sido realizada em escala de detalhe, os dados apresentaram várias inconsistências em sua topologia, o que influenciou diretamente na quantificação de algumas das classes mapeadas. Apesar dessas inconsistências, a quantificação dessas áreas não apresentou grande alteração. Todavia, dependendo da finalidade para a qual os dados forem utilizados, poderão não ser muito confiáveis.

A confiabilidade dos dados será mais intensamente afetada de acordo com a dimensão da área de estudo. Principalmente em escalas grosseiras, representativas de grandes áreas, qualquer erro apresentado pode implicar em significativas discrepâncias, se comparado a suas características reais na superfície terrestre.

Por meio da análise da consistência topológica é possível avaliar se os dados geoespaciais e determinados estudos, que deles se utilizam, são confiáveis ou não. Os erros topológicos apresentados por dados geográficos, de qualquer que seja a pesquisa, podem acarretar a não confiabilidade da mesma, ou mesmo sua invalidação. Considerando que as informações espaciais apresentadas em pesquisas acadêmicas podem servir de base a outros trabalhos, é fundamental primar pela confiabilidade dos dados, para que possíveis erros não sejam replicados.

REFERÊNCIAS

BARROS, E. R. O.; CARNEIRO, A. F. T. Uma Proposta de Controle de Qualidade de Informações Cadastrais de Imóveis Rurais. **Revista Brasileira de Cartografia**, n 65/2, p. 265-581, 2013.

CÂMARA, G.; DAVIS, C.; MONTEIRO, A. M. V.; MEDEIROS, J. S. DE. Introdução à ciência da geoinformação. **Introdução à Ciência da Geoinformação**, p. 345, 2001.

Disponível em:

<<http://www.dpi.inpe.br/gilberto/livro/introd/%7B%25%7D5Cnwww.dpi.inpe.br/gilberto/livro>>. Acesso em: 20/03/2017.

DAVIS, B. E. **GIS: a visual approach**. 2ª ed. New York: Thomson Learning, 2001.

ESRI. **Geodatabase topology rules and topology error fixes**. Disponível em:

<<http://resources.arcgis.com/en/home/>>. Acesso em: 22/03/2017.

LONGLEY, P. A.; GOODCHILD, M. F.; MAGUIRE, D. J.; RHIND, D. W. **Sistemas e Ciência da Informação Geográfica**. Bookman, 2009.

ROCHA, M. P. C. Desenvolvimento de referencial teórico para um sistema de informações gerenciais (SIG) para parlamentares e assessores na Câmara Legislativa do Distrito Federal: em busca de um modelo conceitual. **Ciência da Informação**, v. 32, p. 80–88, 2003.

ROGERSON, P. **Métodos estatísticos para geografia: um guia para o estudante**. 3ª Ed. Porto Alegre: Bookman, 2012, 348 p.

FERREIRA, M. B. P.; PAZ, O. L. S.

Qualidade de dados geoespaciais: análise da consistência topológica no mapeamento de cobertura da terra do município de Curitiba - PR

ROSA, R. Geotecnologias na Geografia aplicada. **Revista do Departamento de Geografia**, v. 16, p. 81–90, 2005. Disponível em:

<<http://citrus.uspnet.usp.br/rdg/ojs/index.php/rdg/article/view/55>>. Acesso em: 13/03/2017.

SANTOS JUNIOR, W.M.; RIBEIRO, G.P. Qualidade dos dados geográficos disponibilizados em ambiente de sistema de informação geográfica na internet. In: SIMPÓSIO BRASILEIRO DE CIÊNCIAS GEODÉSICAS E TECNOLOGIAS DA GEOINFORMAÇÃO, 4., 2012, Recife.. **Anais...** Recife: SIMGEO, 2012 p. 01-08.

SILVA, A. B. **Sistemas de Informações Geo-referenciadas: conceitos e fundamentos**. Campinas: Editora da Unicamp, 2003.

THEOBALD, D. M. **Understanding Topology and Shapefiles**. ArcUser Online, 2001. Disponível em: <<http://www.esri.com/news/arcuser/0401/topo.html>>. Acesso em: 09/04/2017.

VEREGIN, Howard. Data quality parameters. **Geographical information systems**, v. 1, p. 177-189, 1999.

WEBER, E.; ANZOLCH, R.; LISBOA FILHO, J.; COSTA, A. C.; IOCHPE, C. **Qualidade de dados geoespaciais**. Porto Alegre: URGS – Instituto de Informática, 1999. 37 p. Relatório Técnico.

Autores

Manoella Barros Pedreira Ferreira – Possui Graduação pela Universidade Federal do Pará (UFPA) e Mestrado em Geografia pela Universidade Federal do Paraná (UFPR). Atualmente é Professora do Colégio Adventista de São José dos Pinhais (CASJP).

Otacílio Lopes de Souza da Paz – Possui Graduação em Geografia pela Universidade Federal do Paraná (UFPR). Atualmente é Mestrando pelo Programa de Pós-Graduação em Geografia da Universidade Federal do Paraná (UFPR).

Artigo recebido em: 07 de outubro de 2017.

Artigo aceito em: 26 de abril de 2018.

Artigo publicado em: 30 de junho de 2018.