

ANÁLISE DA QUALIDADE DE DADOS DE ENTRADA PARA MANIPULAÇÃO EM AMBIENTE DE SISTEMA DE INFORMAÇÃO GEOGRÁFICA

Bárbara Costa Lima¹

Heloísa Alencar²

1. Mestranda do Programa de Pós-graduação Profissional Biodiversidade em Unidades de Conservação – Escola Nacional de Botânica Tropical (ENBT/JBRJ) (barbaracosta@jbrj.gov.br)
2. Ecóloga - Universidade Federal da Paraíba (locusgisbr@gmail.com)

ABSTRACT

Geographic Information Systems (GIS) refers to the tools computational methods of Geoprocessing, which enable the execution of complex analyzes, from the aggregation of data from different sources and when producing georeferenced databases. This study aimed to report a critical analysis regarding the quality of geographic data provided by public institutions. Among the types of spatial data quality elements defined by ISO 19157: 2013, logical consistency is a little discussed issue. Using data available in the online database of the environmental agency of the state of Rio de Janeiro, we found many topology errors, a subelement of data quality. The GIS used to correct the errors was QGIS, a free and multiplatform software. For the topological verification, the Topology Checker plugin was used. After identifying the errors, the data gaps were corrected using the GRASS v.clean and Correct geometries tools coupled with QGIS. Overlapping errors were manually corrected in the vector data, as only two overlaps were identified. Considering that the analyzed Geospatial Database is widely used by the agency's analysts, professionals in the environmental area, as well as by the scientific community, we understand that it is necessary to carry out a review regarding the topological consistency of all cartographic material online, as well as include more information in the catalog metadata.

Keywords: geospatial data; data quality; metadata

INTRODUÇÃO

Após a metade do século XX, os dados geográficos começaram a serem tratados como conjunto de técnicas matemáticas e computacionais, denominadas Geoprocessamento. Para Câmara *et al.* (2005), esta ciência teria como objetivo estudo e implementação de formas distintas de representação computacional do espaço geográfico”. Trata-se de um tema controverso, pois há quem o considere como a automatização de processos de tratamento e manipulação de dados geográficos que antes eram feitos manualmente. Estes dados podem ser utilizados por exemplo, para realização de estudos ambientais, demográficos, climáticos, hidrográficos, entre outros; representando um objeto ou fenômeno para o qual a localização geográfica é uma característica inerente à informação e indispensável para analisá-la (RODRIGUES, 2018).

Em Geoprocessamento, a topologia expressa a relação espacial entre as camadas (geometrias de ponto, linha ou área) que se conectam ou são adjacentes. Segundo Santos (2015), sua validação objetiva estabelecer o controle sobre os dados produzidos durante o processo de edição, o qual deve obedecer a regras específicas para todos os elementos vetoriais, em especial para as geometrias derivadas (linhas e polígonos).

O presente estudo objetiva relatar uma análise crítica quanto à qualidade de dados geográficos de entrada disponibilizados em banco de dados geoespaciais por órgãos e/ou instituições públicas.

METODOLOGIA

Os elementos de qualidade de dados espaciais são componentes que descrevem aspectos de qualidade do dado geoespacial. A Norma ISO 19157:2013 aborda seis elementos de qualidade de dados espaciais: completude; acurácia posicional; acurácia temática; acurácia temporal; consistência lógica; usabilidade; e seus respectivos subelementos, vide Tabela 1.

TABELA 1: ELEMENTOS DE QUALIDADE DE DADOS ESPACIAIS.

Elemento	Subelemento	Código
Completude	Comissão	11
	Omissão	12
Consistência lógica	Consistência conceitual	21
	Consistência de domínio	22
	Consistência de formato	23
	Consistência topológica	24
Acurácia posicional	Acurácia posicional absoluta	31
	Acurácia posicional relativa	32
	Acurácia posicional dos dados em grade	33
Acurácia temática	Correção da classificação	41
	Correção dos atributos não quantitativos	42
	Acurácia dos atributos quantitativos	43
Acurácia temporal	Acurácia de uma medida temporal	51
	Consistência temporal	52
	Validade temporal	53
Usabilidade		6

Fonte: IBGE (2017) *apud* International Organization for Standardization (2013).

Neste estudo, abordaremos somente os procedimentos realizados para alcançar a correção topológica dos dados obtidos pelo Portal GEOINEA.

O SIG utilizado para corrigir os dados obtidos pelo Portal GEOINEA em diferentes projeções foi o QGIS, um *software* livre, multiplataforma. Nele é possível criar, editar e visualizar dados disponibilizados em um banco de dados geoespacial público, a partir da criação de uma conexão entre o *software* e o banco de dados.

A Tabela 2, abaixo, são descritas as anomalias topológicas que podem ser encontradas em arquivos geoespaciais no decorrer do processo de validação.

TABELA 2: VALIDAÇÃO TOPOLÓGICA NA MESMA CLASSE: ANOMALIAS DE GEOMETRIA

Anomalia	Descrição	Correção
Geometria vazia (<i>Empty geometry</i>)	Feições com a geometria nula ou vazias.	Automática
Geometria desconhecida (<i>Unknown geometry</i>)	Feições cuja geometria não é reconhecida no ambiente SIG.	
Geometria inválida (<i>Invalid geometry</i>)	O tipo de geometria não corresponde a geometria vetorizada.	
Poucos vértices (<i>Too few vertices</i>)	O tipo de geometria possui poucos vértices para representar sua geometria primitiva (ponto, linha ou área).	Automática
Buracos não contidos (<i>Uncontained holes</i>)	Área de feição com limites internos, buracos, que não estão contidos total ou parcialmente dentro do limite externo.	Manual
Áreas não fechadas (<i>Unclosed areas</i>)	Área de feição cujo limite não está fechado, isto é, o primeiro e o último vértice não tem o mesmo valor de coordenada.	Automática
Sobreposição de buracos (<i>Overlapping holes</i>)	Área de feição com limites internos, buracos, se sobrepõem um ao outro.	Manual
Linhas com comprimento zero (<i>Zero-length lines</i>)	Comprimento de feição que não possui nenhuma extensão, isto é, os vértices estão sobre um mesmo ponto.	Automática
Áreas com cobertura zero (<i>Zero-coverage areas</i>)	Cobertura de feição que não possui nenhuma área, isto é, os vértices são todos colineares.	Automática
Coordenadas inválidas (<i>Invalid coordinates</i>)	Feições cujo um valor de coordenadas (X, Y ou Z) é superior a um determinado valor absoluto.	
Laço em área (<i>Area loop</i>)	Áreas de feição com um laço em qualquer um dos limites exteriores ou interiores.	Automática
Pontos (vértices) duplicados (<i>Kickback/Duplicate point</i>)	Área de feição ou feição linear cuja geometria possui pontos duplicados num mesmo vértice.	Automática
Componente de geometria inválido (<i>Invalid geometry component</i>)	Feições onde a geometria é sintaticamente correta, mas cuja especificação não define um componente de geometria válida. Por exemplo: arco inválido, raio inválido, limite inválido, descontinuidade.	
Ponta repentina (<i>Kink / Spike</i>)	Feições lineares ou área onde há divergência repentina na vetorização dos vértices.	Automática
Laço em linha (<i>Loop in line</i>)	Feições lineares cuja geometria cria um laço e a área é maior do que uma determinada tolerância especificada.	Automática
Vetor curto (<i>Short vector</i>)	Feições lineares ou área com dois vértices sequenciais com a distância inferior a uma determinada tolerância especificada. A tolerância a ser considerada pode ser o erro gráfico na escala do projeto (0.2 mm na escala do projeto).	Automática
Geometria nula (<i>Null geometry</i>)	Feições com geometria nula.	Automática
Componente de geometria desnecessário (<i>Superfluous geometry component</i>)	Feições onde a geometria contém um componente que não é necessária para descrever a sua propriedade geométrica.	Automática
Geometria fragmentada (<i>Fragmented geometry</i>)	Feições onde o tipo de geometria é uma coleção e as partes desta feição não precisam ser contíguas.	Automática / Manual
Feição duplicada (<i>Duplicate feature</i>)	Feições cujos atributos combinam e as geometrias são idênticas, ou com tamanho, forma e localização semelhantes, dentro de uma tolerância especificada.	Automática

Fonte: IBGE (2017)

Os dados trabalhados estão atrelados às temáticas de biologia da conservação, áreas protegidas e Unidades de Conservação (UC), disponíveis na aba principal “Biodiversidade e Áreas Protegidas”, conforme disponibilizado na Figura 1.

Com o intuito de quantificar informações referente às Áreas de Preservação Permanente (APP's), procedeu-se o download dos arquivos em formato *shapefile* das UC e APP's de interesse, sendo posteriormente analisados visualmente no *software* QGIS.



Figura 1. Base de dados geoespaciais do INEA.

Os primeiros procedimentos a realizar em qualquer dado geoespacial vetorial, após o seu *download*, são a leitura do metadados e posteriormente a verificação topológica. Para o caso em tela, foi utilizado o *plugin* Verificador de Topologia que pode ser instalado por meio do menu Complementos localizado na barra superior do *software*. Ao clicar na opção Gerenciar e Instalar Complementos será aberta uma janela onde é possível realizar a instalação. O *plugin* possui o painel que pode ser habilitado ao clicar em “Verificador de Topologia”, em seguida, basta clicar em um ícone de uma chave de boca que se encontra nele. Desta forma, é aberta uma janela de configurações de regras de topologia (Figura 2) e por meio dela é possível analisar se determinado vetor apresenta erros topológicos. Outra forma de obter a localização do erro topológico é através da ferramenta “Verificar a validade” que se encontra no menu Vetor, opção Geometrias (Figura 3). Essa ferramenta gera o dado virtual de pontos que permite ao analista encontrar a feição que apresenta o erro.

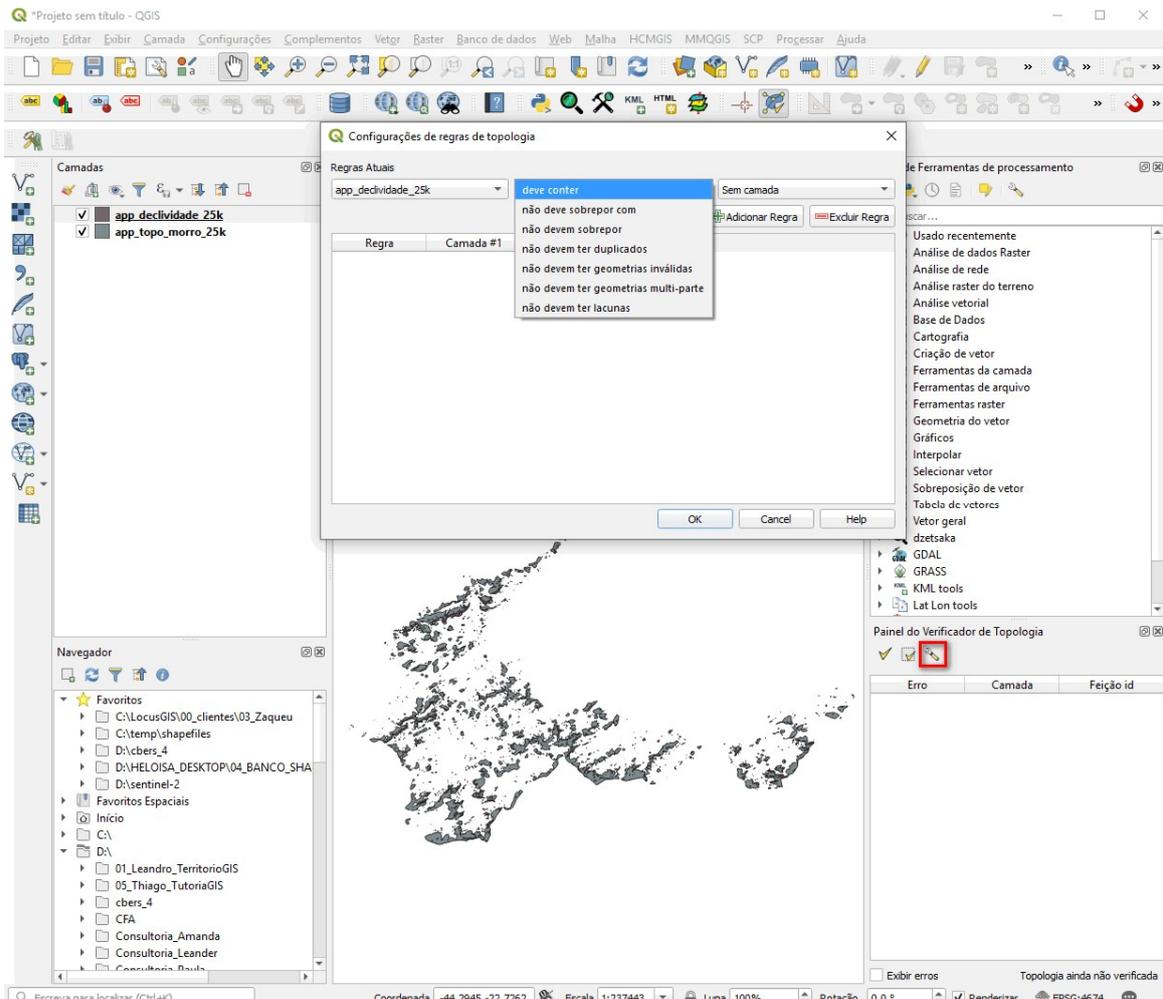


Figura 2. Janela do plugin Verificador de Topologia responsável pela aplicação de regras topológicas nos dados vetoriais.

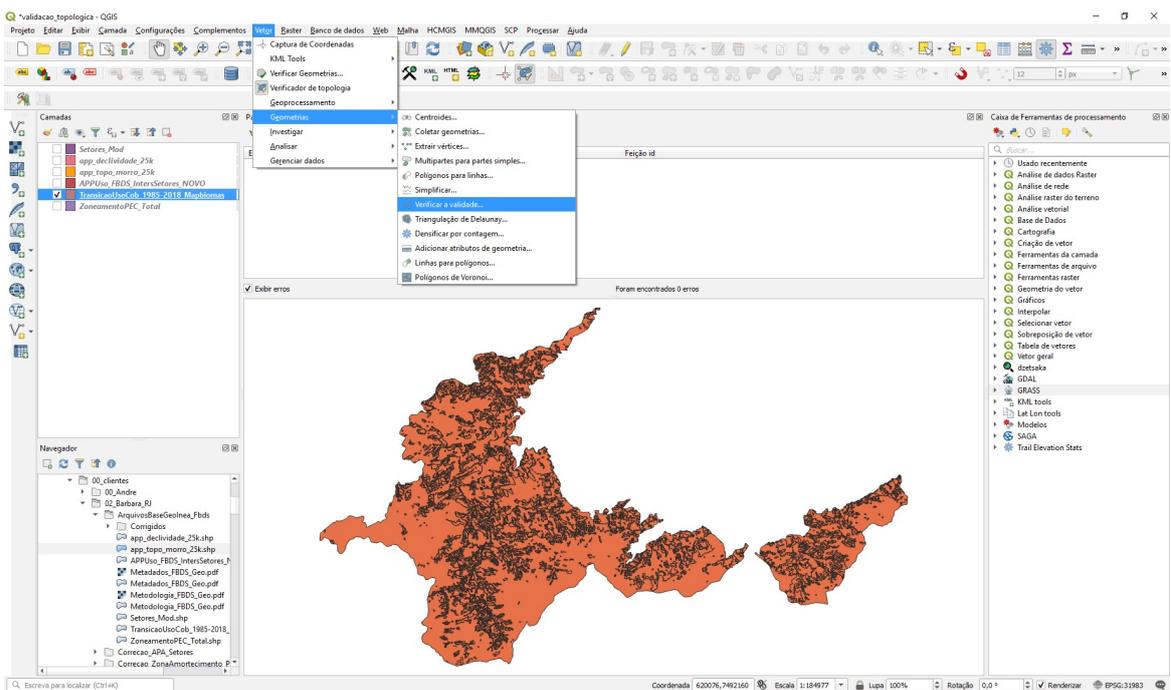


Figura 3. Localização da ferramenta do menu Geometria denominada “Verificar validade”.

Para cada *shapefile*, neste caso de Áreas de Preservação Permanente de Topo de Morro e de Declividade do estado do Rio de Janeiro, foram aplicadas todas as regras do Verificador e posteriormente executadas as devidas correções por meio dos algoritmos respectivos a cada erro.

Para obter a correção de lacunas ou fendas dos dados foi utilizado o *v.clean* do GRASS acoplado ao QGIS, ferramenta de limpeza *snap*, que pode ser acessada na caixa de ferramentas de processamentos. Outra ferramenta utilizada foi a Corrigir geometrias, encontrada também nessa mesma caixa de ferramentas que serve para corrigir geometrias inválidas. Erros de sobreposição encontrados foram corrigidos de forma manual com edição do dado vetorial, por se tratar apenas de duas sobreposições.

RESULTADOS

Após aplicação de todas as ferramentas mencionadas, foram encontrados nos dois *shapefile* de Áreas de Preservação Permanente de Topo de Morro e de Declividade de uma Unidade de Conservação municipal, mais de 1.200 erros topológicos de diferentes formas, dentre lacunas e feições multiparte, conforme pode ser visualizado nas Figuras 4 e 5 a seguir.

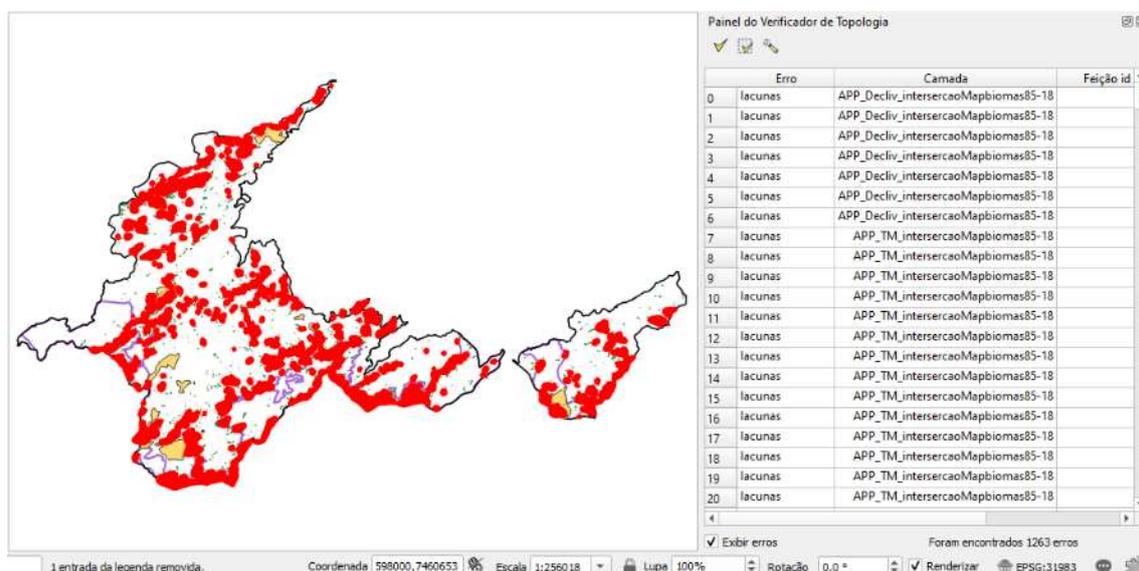


Figura 4. Identificação de erros topológicos pelo painel do “Verificador de Topologia”.

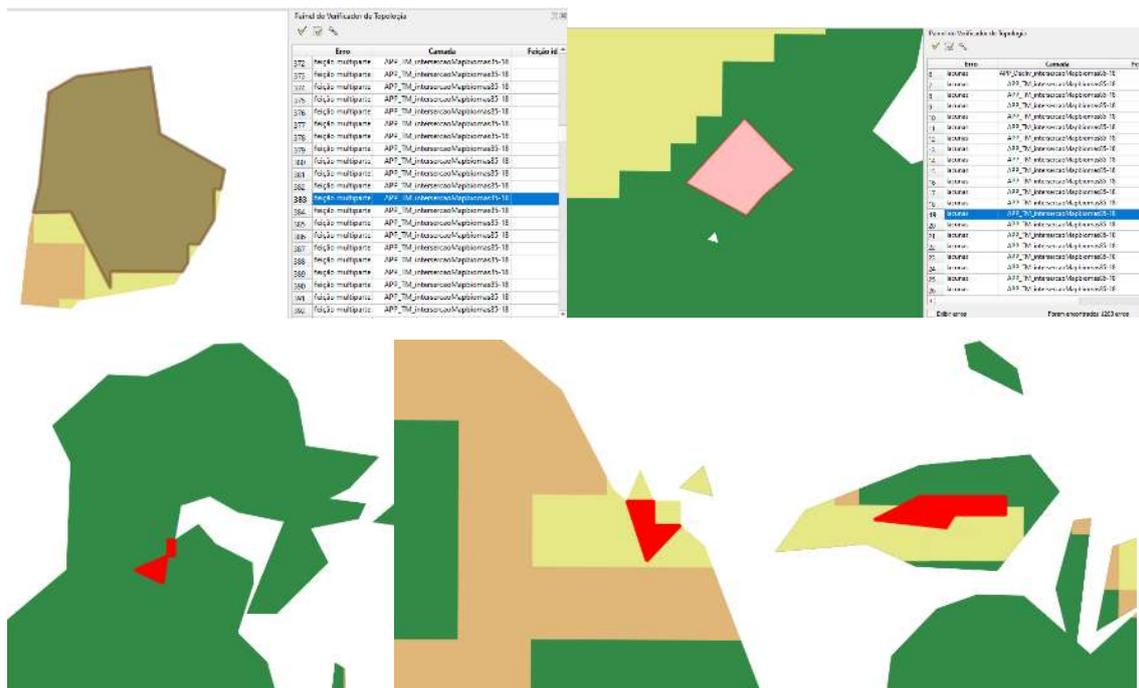


Figura 5. Detalhe de algumas das lacunas identificadas pelo painel do “Verificador de Topologia”.

Devido a considerável quantia de erros topológicos encontrados em dois dos arquivos disponibilizados no Portal de Base de Dados Geoespaciais (GEOINEA) aqui analisados, apesar dos mesmos informarem que os metadados estão atrelados ao Catálogo de Metadados do Portal de dados geoespaciais da Infraestrutura Nacional de Dados Espaciais (INDE), não evidenciamos informações relevantes que pudessem assegurar a consistência topológica dos dados do referido portal.

CONCLUSÕES E CONSIDERAÇÕES

Considerando que a Base de Dados Geoespaciais analisada é referência para nortear as análises de licenciamento ambiental do órgão ambiental do estado do Rio de Janeiro, incluindo anuência de gestores de Unidades de Conservação e implantação/operação de empreendimentos em áreas protegidas pela Lei 12.651/2012, e que ainda, possui acesso público e é amplamente utilizado pela comunidade científica nas mais variadas análises de distintas áreas do conhecimento, entendemos que é necessário haver uma revisão quanto à consistência topológica de todo material cartográfico que compõe a base de dados geoespaciais online, bem como incluir informações no catálogo de metadados dos arquivos disponibilizados.

REFERÊNCIAS BIBLIOGRÁFICAS

CÂMARA, G.; DAVIS, C.; MONTEIRO, A. M.V. *Introdução à Ciência da Geoinformação*. Disponível em: <<http://www.dpi.inpe.br/livros.html>> Acesso em: agosto de 2020.

Instituto Brasileiro de Geografia e Estatística – IBGE. Manuais Técnicos em Geociências, n. 13. Avaliação da Qualidade de Dados Geoespaciais, 2017.

INSTITUTO ESTADUAL DO AMBIENTE – INEA. Base de Dados geoespaciais. Disponível em: <https://inea.maps.arcgis.com/apps/MapSeries/index.html?appid=00cc256c620a4393b3d04d2c34acd9ed> Acesso em: agosto de 2020.

RODRIGUES, M. A. da S. Armazenamento e Manipulação de Dados Espaciais no PostgreSQL/PostGIS. Trabalho de conclusão de curso. UFU, 2018. Disponível em: <<https://repositorio.ufu.br/handle/123456789/24785>> Acesso em: agosto de 2020.

SANTOS, J. Regra Topológicas para verificar a Validade da Geometria Disponível em: <http://www.processamentodigital.com.br/wp-content/uploads/2015/09/QGIS28_Regras_Topologicas.pdf> Acesso em: agosto de 2020.