

Análise de Consistência Lógica conforme padrões de qualidade da norma ISO 19133

Vanessa Costa Maranhão*
Andrea F.T. Carneiro**

Recibido el 01 de marzo de 2016; aceptado el 8 de septiembre de 2016

Resumen

Sobre la importancia de la organización de las infraestructuras de datos espaciales nacionales, la calidad de los datos ha logrado prominencia en la producción cartográfica. La familia de normas ISO 19000 presenta los indicadores utilizados para el control de calidad de productos y procesos cartográficos. En este artículo se presentan los resultados de la verificación de la consistencia lógica, según los subelementos definidos en la ISO 19113 y previstos en la Infraestructura Nacional de Datos Espaciales de Brasil (INDE), en que se evalúa la calidad de los datos bajo los aspectos conceptuales, de dominio, formato y topología.

Palabras clave: *Información Geográfica, Calidad de Datos Espaciales, Consistencia Lógica, ISO 19133.*

Abstract

With the importance of structuring the national spatial data infrastructures, data quality has gained prominence in the cartographic production. The family of ISO 19000 standards shows the indicators used for quality control of products and cartographic processes. In this paper, it is presented the results of the logical consistency verification according to the sub-elements defined in the ISO 19113 and required for the Brazilian National Spatial Data Infrastructure (INDE). It takes in consideration the data quality from the spectrum of conceptual aspects, domain, format and topology.

Key words: *Geographic Information, Quality of Spatial Data, Consistency Logical, ISO 19133.*

* Mestre em Ciências Geodésicas, Brasil, e-mail: vanessa.maranhao@oi.com.br

** Professora do Departamento de Engenharia Cartográfica da Universidade Federal de Pernambuco, Brasil, e-mail: aftc@ufpe.br

Resumo

Com a importância da estruturação das infraestruturas de dados espaciais nacionais, a qualidade dos dados ganhou destaque na produção cartográfica. A família de normas ISO 19000 apresenta os indicadores utilizados para o controle de qualidade de produtos e processos cartográficos. Neste artigo, são apresentados os resultados da verificação da consistência lógica, conforme os subelementos definidos na ISO 19113 e previstos na Infraestrutura Nacional de Dados Espaciais do Brasil (INDE), avaliando a qualidade dos dados quanto aos aspectos conceituais, de domínio, de formato e de topologia.

Palavras-chave: Informação Geográfica, Qualidade de Dados Espaciais, Consistência Lógica, ISO 19133.

Introdução

A crescente difusão da Informação Geográfica (IG) pressupõe a necessidade de conhecer a origem dos dados (linhagem), sua acurácia, sua aplicabilidade, entre outros. A funcionalidade cartográfica só pode ser afirmada a partir do conhecimento da qualidade de seus dados. Mas, como avaliar os erros contidos na informação?, Quais são os parâmetros capazes de identificar a qualidade desses dados?, Como corrigir?

As vantagens e facilidades nas novas técnicas de aquisição de dados permitem cada vez mais a completude das informações geográficas, melhores resultados quanto à acurácia posicional, entre outros benefícios. Diante disso, surge a necessidade de implementações de técnicas e normativas que indiquem até que ponto determinada ferramenta espacial pode ser potencialmente aplicada.

Um termo bastante debatido atualmente pelos estudiosos da geomática é Infraestruturas de dados Espaciais (IDEs), cuja definição apresentada pela Associação Global de Infraestruturas de Dados (sigla em inglês *GSDI*) considera as IDEs como o conjunto de tecnologias, políticas e mecanismos institucionais que facilita a disponibilidade de acesso aos dados espaciais. No Brasil, o Decreto 6.666/08 instituiu a criação da Infraestrutura Nacional de Dados Espaciais (INDE) e, desde então, têm sido elaboradas normas e padrões para a integração e disponibilização de dados espaciais produzidos por instituições públicas.

A ISO/TC211, Comitê Técnico 211 da Organização Internacional de Normalização, tem entre outras funções, o objetivo de desenvolver um conjunto abrangente de padrões de cunho geográfico, promovendo o desenvolvimento, implantação e normalização da IG. Nesse contexto, o comitê desenvolveu uma série de normas ISO 19000, que fornece visões gerais de qualidade cartográfica para as categorias de padronização de Infraestrutura Espacial, modelagem de dados de geoespaciais, gestão da IG, serviços de IG, codificação da IG e temas específicos.

Na família ISO 19000, a norma ISO 19113 (ISO, 2002), revisada pela ISO 19157 (2013), rege os princípios de qualidade, apresentando como elementos de qualidade: Completude - presença e ausência de características, seus atributos e relacionamentos; Consistência lógica - grau de aderência às regras lógicas de estruturas, atribuições e relacionamentos dos dados; Exatidão posicional - acurácia da posição das feições; Exatidão temporal - acurácia dos atributos temporais e relações de características temporais; Exatidão temática - acurácia de atributos quantitativos, correção de atributos não quantitativos, classificação das características e seus relacionamentos.

O Plano de Ação para Implantação da Infraestrutura Nacional de Dados Espaciais do Brasil, elaborado pela Comissão Nacional de Cartografia (CONCAR), órgão responsável pelas diretrizes e bases da cartografia brasileira, prevê a elaboração de normas e padrões para aspectos variados das informações espaciais (metadados, dados vetoriais, qualidade, mapeamento cadastral). Algumas dessas especificações já foram publicadas, outras encontram-se em fase de elaboração.

Destacam-se, dentro do Plano de Ação para implantação da INDE, as especificações técnicas que vem sendo elaboradas para garantir os padrões almejados numa IDE. São elas: Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV), Especificação Técnica para a Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV), Especificação Técnica de Produtos de Conjuntos de Dados Geoespaciais (ET-PCDG), Especificação Técnica para o Controle de Qualidade dos Produtos de Conjuntos de Dados Geoespaciais (ET-CQDG) e a Especificação para a Representação de Dados Geoespaciais (ET-RDG).

A ET-EDGV apresenta um padrão a ser utilizado por todos os que contratam, produzem, fiscalizam ou apenas são usuários de dados vetoriais associados a bancos de dados, contemplando ainda um modelo conceitual a ser utilizado na cartografia nacional, em que o espaço geográfico brasileiro é representado por feições geográficas agregadas em treze categorias de informações: Relevo; Vegetação; Hidrografia; Sistema de Transporte; Energia e Comunicações; Abastecimento de Água e Saneamento Básico; Educação e Cultura; Estrutura Econômica; Localidades; Pontos de Referência; Limites; Administração Pública; Saúde e Serviço Social. A ET-EDGV é concebida por diagramas de classe objetos correspondentes a cada uma das categorias, explicando relacionamentos entre as feições, os seus atributos e seus campos de domínio.

A ET-CQDG foi publicada em 2016 com o objetivo de padronizar a avaliação da qualidade dos produtos de conjuntos de dados geoespaciais integrantes do Sistema Cartográfico Nacional (SCN) do Brasil. Trata-se de uma especificação em conformidade com a ISO 19157:2013, posto que os procedimentos de avaliação seguem os passos descritos na norma ISO, e também descreve como reportar a qualidade dos produtos usando procedimentos padronizados (DSG, 2016).

Neste artigo, são apresentados os resultados da verificação da consistência lógica, conforme os subelementos definidos na ISO 19113 e previstos na INDE, avaliando a qualidade dos dados quanto aos aspectos conceituais, de domínio, de formato e de topologia. Além disso, busca compatibilizar os dados ao modelo conceitual previsto na INDE, através da ET-EDGV.

Os dados utilizados para análise de consistência lógica e contabilização de padrões conceituais foram dados vetoriais integrantes do CACI – Cadastro Comprometido com Intervenções. O CACI é uma base cartográfica e temática, implementada por meio da coleta, tratamento, lançamento e registro descritivo e cartográfico de programas, planos, projetos (de instituições federais, estaduais e municipais), de forma sistemática, e articulada com os órgãos com poderes de intervir no espaço metropolitano, de modo que todas as áreas objeto de intervenção sejam analisadas à luz das normas e parâmetros de ocupação (Fidem, 1981). É a principal ferramenta para auxílio nas tomadas de decisões quanto ao uso e ocupação do solo no estado de Pernambuco, pela Agência Estadual de Planejamento e Pesquisa de Pernambuco (CONDEPE-FIDEM).

As novas tecnologias e o controle de qualidade de dados espaciais

Segundo Delavar e Devillers (2010), a importância da qualidade dos dados espaciais aumenta à medida em que testemunhamos o surgimento e o uso generalizado de sensores terrestres, aerotransportados e espaciais (como imagens de sensoriamento remoto de alta resolução, imagens de radar interferométrico, imagens hiperespectrais), sistemas baseados em posicionamento por satélite. A estes pode-se acrescentar a disseminação do uso de imagens de VANTs (veículos aéreos não tripulados). A necessidade de integrar uma quantidade crescente de dados heterogêneos resulta num desafio significativo no que diz respeito à medição, descrição e comunicação da qualidade da informação geográfica.

Wan *et. al.* (2015) ampliam essa visão, afirmando que os aspectos da qualidade de dados espaciais tem um papel fundamental na evolução das técnicas cartográficas e no aperfeiçoamento de aplicações práticas. Por isso, questões relacionadas à qualidade tem atraído cada vez mais a atenção de pesquisadores, empresas e instituições governamentais, como resultado de exigências de interoperabilidade e integração de dados de diferentes fontes.

Guptill e Morrison (1995) especificam cinco componentes que se reportam à qualidade do dado espacial relativos à fidelidade em aspectos semânticos e temporais, sendo os mesmos aceitos pela Comissão de Qualidade de Dados Espaciais da International Cartographic Association (ICA): linhagem, acurácia posicional, fidelidade do atributo ou acurácia temática, consistência lógica e completude. Estes indicadores são utilizados nas especificações da ISO 19113 e 19114 (ISO/TC211, 2003), estabelecendo princípios para identificação, avaliação e registro sobre a

qualidade da IG. A ISO 19113 fornece o conjunto de padrões indicadores de qualidade do geodado, enquanto a ISO 19114 orienta sobre a organização de procedimentos para a avaliação de métricas de qualidade do geodado.

O foco deste artigo é o indicador consistência lógica, conforme os subelementos definidos na ISO 19113 e previstos na Infraestrutura Nacional de Dados Espaciais do Brasil (INDE), avaliando a qualidade dos dados quanto aos aspectos conceituais, de domínio, de formato e de topologia.

Consistência Lógica

A consistência lógica informa sobre a manutenção de relações lógicas e topológicas dos dados espaciais. As análises incluem testes de valores válidos, testes gerais para dados gráficos (ex.: se os nós estão todos unidos, se os polígonos estão todos fechados) e testes topológicos específicos (ex.: se limites de polígonos vizinhos não estão se cruzando, se o sentido de fluxo não é contrário em elementos de uma rede) (Weber *et al.*, 1999).

Os testes de consistência lógica dos dados espaciais são necessários em várias etapas de manipulação dos dados, seja durante o processo de coleta ou após o seu processamento e análise. As incompatibilidades de consistência lógica mais verificadas nas informações geográficas são ocorrências de inconsistências geométricas, falhas geradas na digitalização e atualização, propagação de erros devido ao processamento ou não codificação das relações topológicas (Francisco, 2001).

Para Ceballos e Gatica (2012), a consistência lógica descreve o grau de confiança com o qual um determinado conjunto de dados cumpre especificações no que diz respeito à estrutura interna dos dados e da topologia. Considera ainda que a melhor forma de testar esse elemento de qualidade é através de seus subelementos: Consistência Conceitual, Consistência de Domínio, Consistência de Formato e Consistência Topológica. A proposta de Ceballos e Gatica (2012) para este teste é que:

- a) A análise da **Consistência Conceitual** visa verificar se os mesmos obedecem a um Modelo Conceitual que atenda às necessidades dos usuários. Por estar diretamente vinculada aos dados e seus atributos, propõe-se o uso da Norma “ISO2859-1 Procedimentos de amostragem para inspeção por atributos. Parte 1: Plano de amostragem para inspeção lote por lote, tabelados segundo o nível de qualidade aceitável (NCA)”. A NCA é um parâmetro do sistema de amostragem, nele o grau das amostras analisadas que não obedecem à conformidade não devem exceder 10% do total.
- b) Na **Consistência de Domínio** verifica-se a pertinência dos valores dos atributos ao domínio da base de dados;
- c) A estrutura de saída dos dados é testada na **Consistência de Formato**.
- d) E por fim, a **Consistência Topológica** revisa e corrige os erros de topologia.

Para Longley *et al.* (2013), testar a integridade topológica de um conjunto de dados é uma maneira útil de validar a qualidade geométrica dos dados e avaliar sua adequação para análise geográfica.

Sobre as inconsistências geométricas, Kainz (1997 apud Barros, 2011, p. 25) exemplifica como inconsistências geométricas as linhas duplicadas, linhas desconexas, linhas conectadas com ausência de pontos de interseção (nós), linhas estendidas excessivamente, polígonos sem centroides, polígonos abertos, pequenas áreas resultantes da sobreposição de mapas.

Na prática, os principais erros de inconsistência geométrica geralmente encontrados nos mapas são:

- Primitivas Geométricas duplicadas: Duplicidade de uma mesma feição representada (linha, ponto ou polígono).
- Ausência de sobreposição de nós (polígonos): Polígonos abertos ou com linha que não alcança ou ultrapassa o ponto de interseção necessitam a edição manual destes nós, aproximando-os ou juntando as linhas.
- Linhas descontínuas: Um objeto contínuo é representado descontinuamente ou com quebras.
- Presença de objetos curtos: Representação de feições invisíveis à escala do mapa, popularmente chamadas de “sujeiras de mapa”.
- Digitalização de número de pontos insuficiente: Uma representação curva depende do número de vértices utilizados. Consequentemente, o erro relativo à digitalização de linhas retas é muito menor que o resultante da digitalização de curvas complexas.
- Digitalização errônea de feição: A feição vetorial digitalizada não corresponde à realidade.
- Limites entre polígonos não definidos corretamente: As linhas que compõem um polígono se sobrepõem ou existe uma lacuna entre elas.

Ainda, para Kainz (1997 apud Barros, 2011, p. 26), um conjunto de dados separados em níveis de informações pode ser consistente dentro de cada nível de informação, mas inconsistente entre os outros níveis de informações e dados oriundos de diferentes escalas ou fontes cobrindo a mesma área.

Do ponto de vista de automatização do teste da consistência lógica, Gong and Mu (2000) desenvolveram uma ferramenta para a detecção de erros num banco de dados espacial através da análise da consistência usando relações lógicas entre vizinhanças espaciais e atributos de dados de diferentes fontes.

Na questão da consistência conceitual, Gática (2010) propõe uma tabela para a avaliação da consistência conceitual por meio da atribuição de cada entidades representada conforme suas características, devendo cada entidade e seu uso definir as

características aceitáveis do universo abstrato. Este tipo de validação é verificada em cada entidade e depende do universo abstrato e do conceito utilizado na criação de cada base de dados. Essa forma de avaliação da consistência conceitual foi utilizada no desenvolvimento deste artigo, e o seu nível de conformidade baseou-se na ISO 2859-1.

Teste da consistência lógica do CACI - Cadastro Comprometido com Intervenções

A análise de consistência lógica está diretamente relacionada à correta topologia dos dados. O objetivo é testar se os elementos geométricos representados (ponto, linha e polígono) não estão duplicados, se possuem homogeneidade, continuidade e conectividade e se os polígonos estão fechados (sem gaps). Além disso, verifica-se se o produto possui correto nível de generalização, no caso de mudança de escala.

As entidades geográficas codificadas utilizando modelos de dados vetoriais são designadas de feições, e esta será a convenção aqui adotada. Feições com mesmas características geométricas são armazenadas em uma mesma classe. Para o teste da consistência lógica dos produtos do CACI, foram selecionadas as feições apresentadas no Quadro 1.

Observou-se, nas feições selecionadas, que as informações a serem trabalhadas, além de desatualizadas, haviam passado por processos de generalização e/ou transformação do sistema geodésico de referência. Além disso, em sua grande maioria, não apresentavam metadados nem descrições quanto à qualidade da informação espacial.

Uma outra característica dos dados do CACI é o seu modelo de dados simples, derivados de trabalhos desenvolvidos em ambiente CAD, no qual entidades do mundo real são representadas simbolicamente como pontos, linhas e vetores de áreas. Segundo Longley *et al.* (2013), as estruturas de conjuntos de dados simples de polilinhas e polígonos é, por vezes, chamada de espaguete, quando linhas e polígonos podem se sobrepor, não havendo relacionamento entre nenhum dos objetos.

Embora o CACI tenha sido desenvolvido com o programa *Autodesk Map*, que permite fixar escala e incluir sistemas de referências, projeção e um banco de dados, os seus elementos individuais não possuem identificadores únicos, dificultando marcá-los com atributos, um requisito básico para aplicações SIG (Sistema de Informações Geográficas), desejável numa IDE. Os elementos do CACI elaborados em formato CAD só receberam identificadores quando exportados para o formato SDF (*Spatial Data File*), que permite estruturar um banco de dados espacial (*geodatabase*) para inserir no programa *MapGuide Open Source* que, por sua vez, serve de API (Interface de Programação de Aplicativos) e elo de ligação de um banco de dados. Essa foi a metodologia utilizada na estruturação do Sistema Integrado de

Uso do Solo (SIUS). Neste sistema, as feições são bloqueadas contra edições da sua geometria e topologia, impedindo a verificação qualitativa dos dados.

Quadro 1
Feições avaliadas quanto à consistência lógica

Feição	Origem	Ano de última atualização	SGR original	SGR Atual	Escala de Origem	Escala Generalizada
Hidrografia	Unibase	1998	SAD-69/69	SIRGAS 2000	1:1.000	1:10.000
Planimetria	Unibase	1998	SAD-69/69		1:1.000	1:10.000
Matas	Unibase	1988	SAD-69/69	SIRGAS 2000	1:1.000	1:10.000
Sistema Viário Principal	DER	2010	SAD-69/69	SIRGAS 2000	1:1.000	1:10.000
Educação	GERE	2008	SAD-69/69	SIRGAS 2000	1:2.000	1:10.000
Segurança Pública	SDS	1998	SAD-69/69	SIRGAS 2000	1:1.000	1:10.000
Lazer	Unibase	1997	SAD-69/69	SIRGAS 2000	1:1.000	1:10.000
Saúde	Unibase	1997	SAD-69/69	SIRGAS 2000	1:1.000	1:10.000
Conjuntos Habitacionais	FIDEM	2004	SAD-69/69	SIRGAS 2000	1:10.000	—
CSU	CEHAB	2004	SAD-69/69	SIRGAS 2000	1:10.000	—
Cemitério	Unibase	1997	SAD-69/69	SIRGAS 2000	1:1.000	1:10.000

Os modelos de dados de tipo topológicos são importantes pelo seu papel na validação dos dados, na modelagem de feições integradas, na edição e na otimização de consultas. Assim, tornam-se necessárias a criação e correção topológica do CACI, pelo fato de um layer em CAD não diferenciar estruturas topológicas entre pontos, linhas e polígonos.

Nas estruturas CADs, vários problemas devem ser preliminarmente corrigidos, antes da sua transformação em estrutura topológica. Como exemplo, em uma feição representada por um polígono simples na estrutura CAD, este é representado de forma semelhante àquela usada para elementos lineares, ou seja, por um par ordenado de coordenadas (x,y). Essa forma de representação apresenta como desvantagem o fato de que as linhas entre polígonos adjacentes devem ser digitalizadas duas

vezes, o que aumenta a possibilidade de ocorrência de erros. Além disso, as informações sobre os polígonos vizinhos são inexistentes, e as ilhas (polígonos contidos em polígonos) são apenas construções gráficas, não sendo possível verificar se a topologia está correta.

Quanto à estruturação do modelo conceitual utilizado neste trabalho, buscou-se atender às recomendações da ET-EDGV. Para a base de dados do CACI, foi proposta uma remodelagem de alguns temas que constituem esta base, de acordo com o modelo conceitual da ET-EDGV: territoriais básicos (relevo, vegetação e hidrografia), de infraestrutura de serviços (saneamento e transportes) e de infraestrutura social (educação, lazer, saúde e serviço social, segurança pública).

Observando as orientações da ISO 19113, e buscando a compatibilização conceitual do modelo ET-EDGV, procurou-se avaliar a qualidade dos dados quanto a Consistência Lógica sob aspectos conceituais, de domínio, de formato e de topologia (Figura 1).

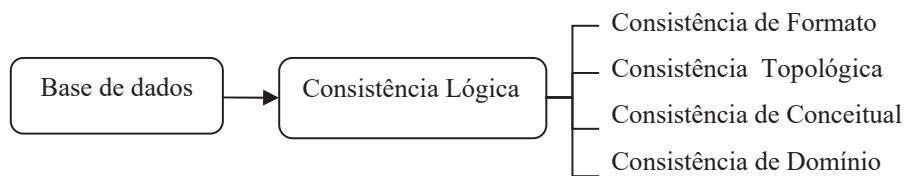


Figura 1. Subelementos testados para a validação da Consistência Lógica.

a) Consistência de Formato:

Nesse teste, foram avaliados os tipos de formatos válidos aceitos para armazenamento, de acordo com a estrutura física da base de dados desejada.

Os formatos, de maneira geral, variam conforme o software utilizado. Alguns formatos bastante utilizados nos modelos de dados vetoriais:

- a) *CAD*: DWG e DXF.
- b) *ESRI*: SHP e LYR

Entre os modelos de dados matriciais, os formatos mais aceitos são *TIFF*, *BMP*, *JPEG*, *MrSID*, *ECW*, *JPEG2000*, entre outros.

A importância da consistência de formato está em normalizar os dados e permitir a sua integração. Assim, os dados inicialmente apresentavam-se em formato *DWG* e foram convertidos para *SHP*, no caso das estruturas vetoriais, e *TIFF*, para as estruturas matriciais (ortofotocartas que deram origens aos dados vetoriais), já que tais formatos são reconhecidos pela maioria dos softwares de SIG.

b) Consistência Topológica:

Os dados vetoriais do CACI apresentaram erros grosseiros ocasionados pela digitalização de feições, identificados pela comparação entre vetores e ortofotos de origem.

Esses erros nos modelos de dados oriundos do processo de digitalização necessitam de correções preliminares, antes da geração e testes de validação da sua topologia. A Figura 2, apresenta um exemplo de erro identificado na feição de planimetria dos dados estudados.

O Quadro 2 apresenta os principais erros geométricos encontrados nas feições testadas.

Quadro 2
Erros de digitalização (inconsistência geométrica) encontrados
nos dados espaciais utilizados na área de estudo

-
- Primitivas geométricas duplicadas
 - Polígonos não fechados (nós sem sobreposição)
 - Linhos descontinuas (objetos com quebras)
 - Objetos curtos (não visíveis na escala da carta)
 - Digitalização de número de pontos insuficientes
 - Digitalização errônea
 - Generalização: excesso ou insuficiência de informações
 - Limites entre polígonos não definido corretamente
 - Nós não definidos
-

A detecção destas inconsistências geométricas foi realizada através da função *Drawing Cleanup* do programa *Autocad Map 3D 2013*. Essa ferramenta não é válida para correções de topologia, apenas de elementos geométricos em formato *CAD*.

A Figura 3 mostra a tela de opções do software *CAD* para essa verificação.

Na Figura 3 observa-se na primeira opção: *Objects to include in drawing cleanup*, que permite selecionar todos os elementos de um determinado nível de informação para verificação em conjunto ou por seleção manual, possibilitando a escolha de qualquer elemento de um determinado nível de informação. Existe ainda a opção de escolha do *layer* que se deseja verificar. Na segunda opção, objetos podem ser ancorados. Os objetos ancorados são pontos de referência e não são alterados ou movidos. Neste trabalho foi utilizada a opção automática, para a seleção dos elementos constantes de cada nível de verificação.

A tela seguinte apresenta os tipos de ações que se deseja executar e permitiu especificar a ordem de execução de cada tipo de tarefa de limpeza e correção geométrica, conforme pode ser visto na Figura 4.

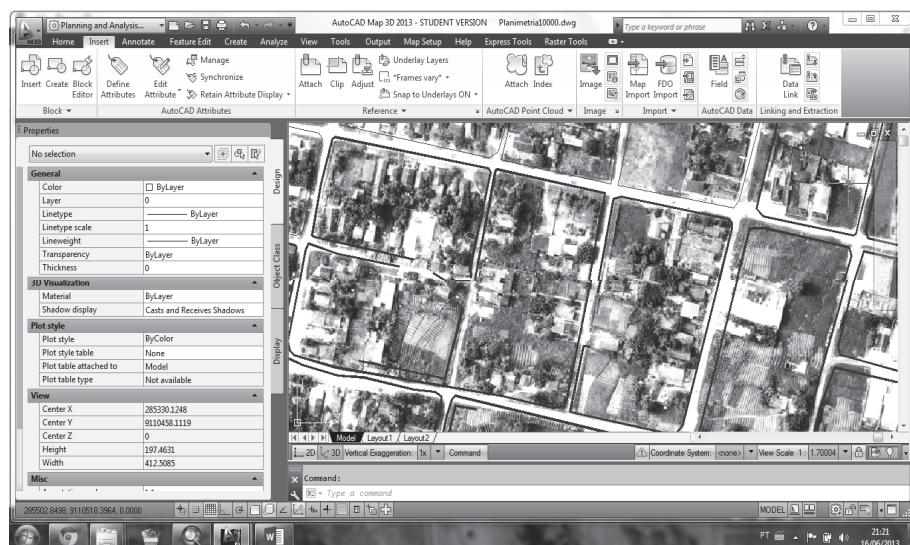


Figura 2. Exemplo de erro grosso identificado nos vetores que representam a feição de planimetria

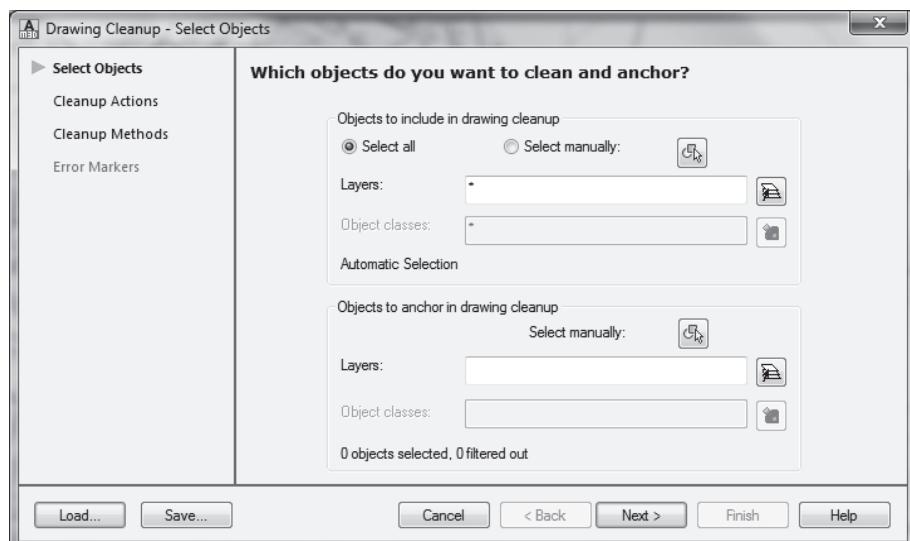


Figura 3. Tela principal da ferramenta *Drawing Cleanup* para verificação de consistência geométrica.

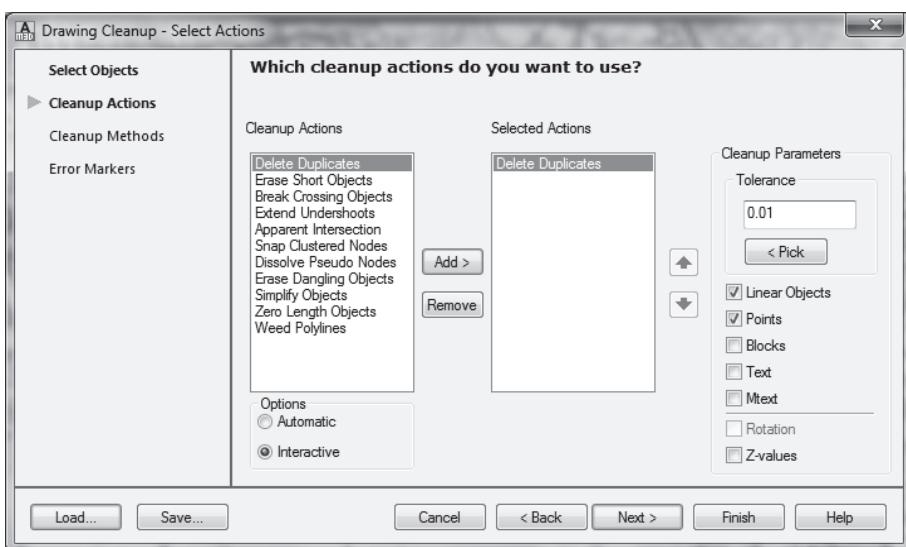


Figura 4. Tipos de opções de correções a serem efetuadas.

As opções de correções geométricas disponíveis pelo programa são:

- *Delete Duplicates*: apaga objetos duplicados;
- *Erase Short Objects*: localiza objetos menores que uma tolerância especificada e os elimina;
- *Break Crossing Objects*: quebra objetos e linhas que se cruzam;
- *Extend Undershoots*: estende linhas para que se encontrem com objetos;
- *Apparent Intersection*: estende linhas soltas para que se encontrem. A extensão das entidades é calculada com base na projeção dos caminhos naturais e limitada por um raio de tolerância;
- *Snap Clustered Nodes*: estende linhas em cruzamentos;
- *Dissolve Pseudo Nodes*: elimina nós desnecessários;
- *Erase Dangling Objects*: elimina linhas sem nexo que perturbam a formação de polígonos;
- *Simplify Objects*: permite a simplificação do desenho, eliminando nós, reduzindo o tamanho do arquivo;
- *Zero Length Objects*: Elimina objetos com tamanho igual a zero;
- *Weed polylines*: usada para adicionar ou remover vértices em polilinhas 3D (não usada neste trabalho).

A verificação das ações selecionadas pode ser realizada de forma automática ou manual. Optou-se pelo uso manual, pois nos casos da existência de objetos muito

próximos a outros, o programa pode eliminar ou estender estes objetos equivocadamente. Recomenda-se, ainda, que a execução de cada uma das ações desejadas seja realizadas individualmente, permitindo ao usuário avaliar os problemas encontrados, e intervir na solução ideal.

Outro fato que implicou na opção pela intervenção mal foi a definição das tolerâncias nas ações de correções geométricas. Nesse caso, o critério de tolerância considerado foi baseado na precisão gráfica, ou seja, a menor grandeza passível de ser representada num desenho em determinada escala. Considerando que a escala de trabalho era de 1:10.000, a precisão gráfica correspondente foi de 20cm. Assim, para elementos gráficos de dimensões inferiores a 20cm aplicou-se a ação *Erase Short Objects*.

De uma maneira geral, buscou-se avaliar em cada item um valor mínimo para uma tolerância lógica capaz de reconhecer a inconsistência, permitindo que a correção geométrica ocorra sem prejuízos à informação espacial.

A próxima tela do programa (Figura 5) ofereceu opções de modificar os objetos originais, corrigí-los ou apagá-los após as ações, usar ou não o layer original, e ainda realizar conversões geométricas dos objetos.

Por fim, foi apresentada a listagem dos erros de inconsistência geométrica localizados, conforme visualizado na Figura 6. É possível ir à ação e ver o quantitativo dos erros marcados, avaliá-los e assim definir qual o procedimento a realizar correção ou não.

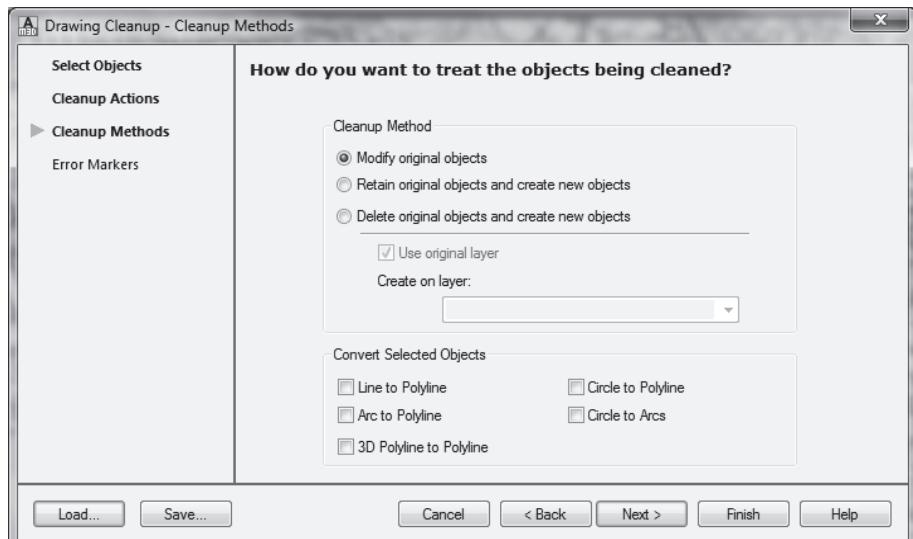


Figura 5. Tela de opções com os tipos de métodos de saída

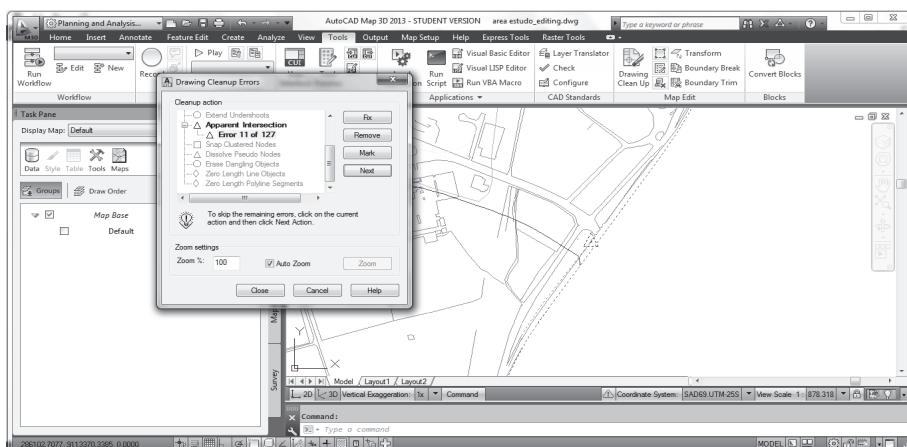


Figura 6. Relatório de erros.

No relatório dos erros de inconsistência geométrica foram encontrados 18.557 ocorrências, distribuídas conforme demonstrado no Quadro 3.

Quadro 3 Quantitativo de erros geométricos

<i>Inconsistência Geométrica</i>	<i>Quantitativo de erros</i>
Objetos duplicados	15169
Objetos curtos	185
Objetos que se cruzam	44
Linhas que necessitam estender até objetos	1887
Linhas soltas	127
Falsos nós	1080
Linhas sem nexo	62
Desenho necessitando ser simplificado	3

Concluídas as correções de inconsistência geométricas, procedeu-se à transformação de feições simples em topológicas, utilizando-se o programa *Autocad Map*. Este programa, no entanto apresenta limitações em relação aos tratamentos topológicos. Assim, o produto final vetorial foi exportado para *SHP*, de modo a se realizar os tratamentos e validações topológicas com o programa *ArcMap10*, mais eficiente para essa finalidade.

Para a criação de topologia no *ArcMap10*, foi gerado um banco de dados geoespaciais (*geodatabase*), que se trata de um diretório de sistema de arquivos onde todas as feições em *shapefile* (extensão SHP) foram importadas. Este procedimento foi realizado por meio da importação em um novo *feature dataset* com sistema de referência definido, criando um novo *feature class*, que corresponde a uma coleção de topologias.

A criação de topologias no *ArcMap10* foi realizada de forma simples (convencional), clicando com o botão direito do mouse sobre a *feature dataset*, e depois em *new – topology* (Figura 7).

Criada a topologia, escolheu-se um nome para o arquivo de relações topológicas e adicionou-se a este os *feature class* que foram trabalhados (Figura 8).

Ressalta-se que também foi possível adicionar as regras de validação de consistência topológica:

- *Must no overlap* (não deve haver interseções entre os polígonos)
- *Must not have gaps* (não deve haver fendas entre os polígonos)
- *Must not overlap with* (não deve haver interseções entre os polígonos de um *layer* com os do outro)
- *Must be covered by feature class of* (a área de um *layer* deve se sobrepor a área de outro *layer*)
- *Must cover each other* (deve haver a sobreposição mútua entre a área de um *layer* e a área de outro *layer*)

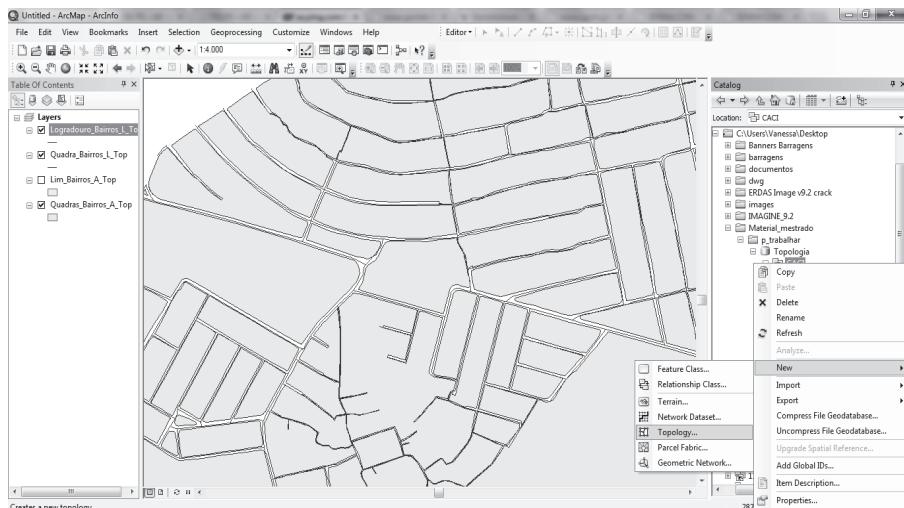


Figura 7. Criando uma topologia a partir do *feature dataset*.

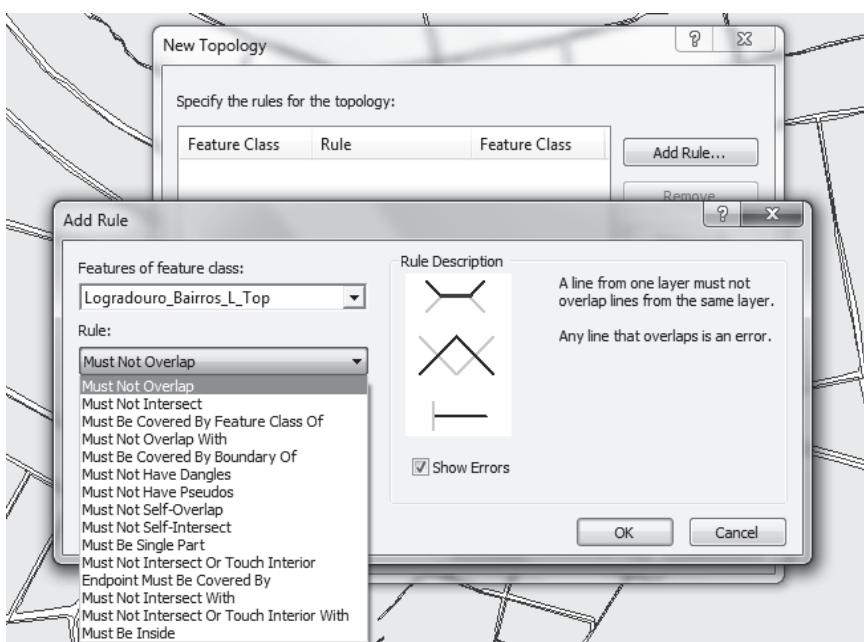


Figura 8. Regras de validações topológicas apresentadas no ArcMap10.

- *Must be covered by* (A área de um *layer* deve estar contida na área de outro *layer*)
- *Boundary must be covered by* (Os limites de uma área de um *layer* devem estar contidos nos limites de uma linha de outro *layer*)
- *Area boundary must be covered by bound* (Os limites de uma área de um *layer* devem estar contidos no limites de uma área de outro *layer*)
- *Contains points* (uma área de um *layer* de polígonos deve conter pelo menos um ponto de um *layer* de pontos)

Como a maior parte das correções topológicas já haviam sido realizadas no processo realizado em *CAD*, esperava-se um quantitativo inferior de erros para esta etapa, como confirmado no relatório do *ArcMap10*, o que totalizou 919 erros de inconsistência topológicas a serem corrigidos (Figura 9).

As checagens e correções foram feitas através da barra de ferramentas *topology* (Figura 10). Com a ferramenta *fix topology error tool*, os erros foram selecionados. A ferramenta *error inspector* registrou qual a regra topológica que não foi cumprida. Em seguida, o *ArcMap* disponibilizou algumas sugestões de correção, ou de marcação do erro como uma exceção à regra.

Finalmente, a base espacial foi considerada livre dos erros topológicos testados.

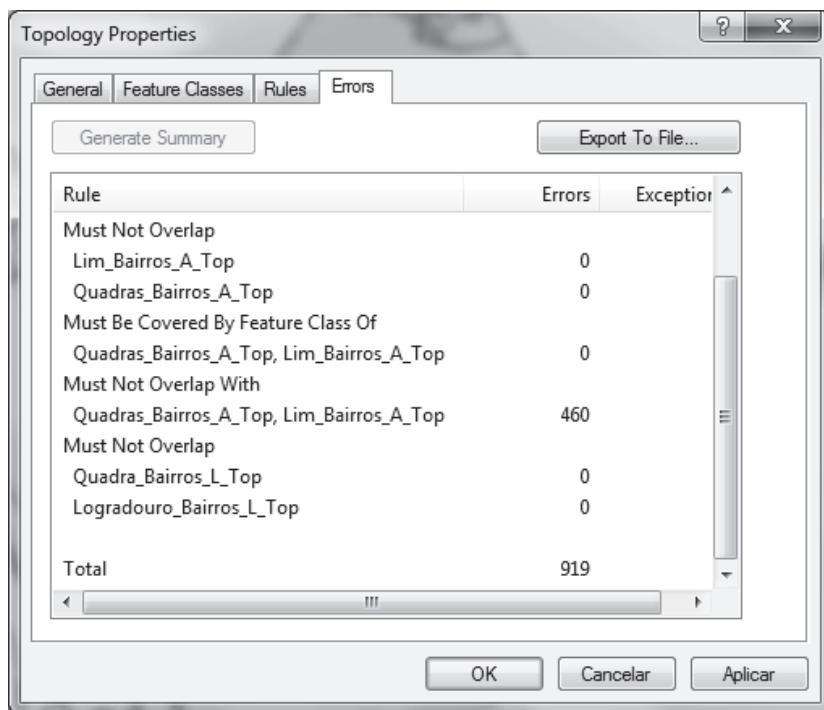


Figura 9. Tela de resultados dos erros de inconsistência topológica verificados.

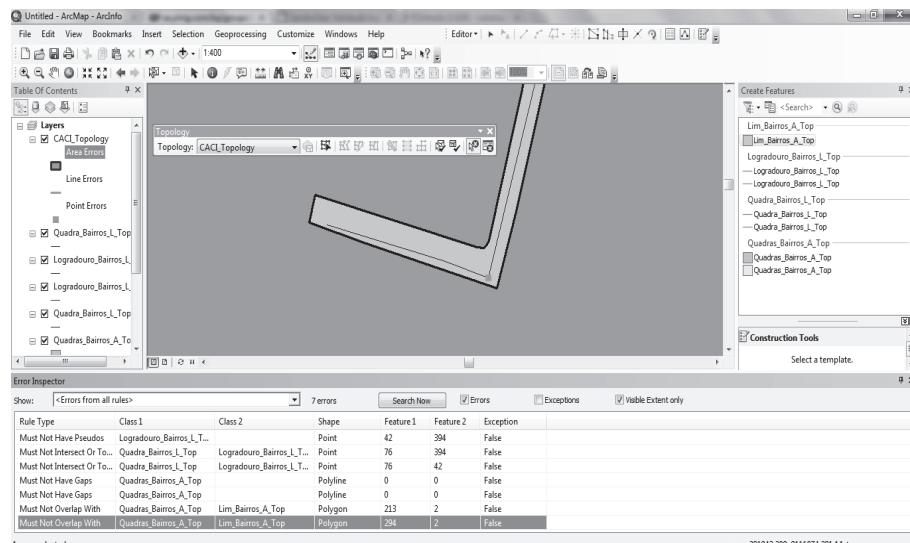


Figura 10. Processo de correção das inconsistências topológicas.

Código	Nome da feição	Categoria	Classe	Subclasse	Primitiva Geométrica	Nome_Tema	1:10.000	1:25.000	1:100.000	Valor
HID	Açude	Hidrografia	Massa_Dagua	Açude	Polygono (A)	HID_Açude A: alfamérico,80	X	X	X	0
HID	Rio	Hidrografia	Curso_Dagua	Corpo_d'água	Polygono (A)	HID_Curso_Dagua_A: alfamérico,80	X	X	-	1
HID	Rio	Hidrografia	Curso_Dagua	Corpo_d'água	Linha (L)	HID_Curso_Dagua_L: alfamérico,80	X	X	X	1
HID	Rio	Hidrografia	Trecho_Curso_Dagua	Corpo_d'água	Linha (L)	HID_T_Curso_Dagua_L: alfamérico,80	X	X	X	1
TRA	Sistema_Viário_Principal	Sistema de Transporte	Via_Rodovaria	-	Linha (L)	TRA_Via_Rodovaria_L: alfamérico,80	X	X	X	1
TRA	Trecho_Logradoiro	Sistema de Transporte	Arruamento	Vias_ou_Logrados	Linha (L)	TRA_Logradoiro_L: alfamérico,80	X	-	-	-1
TRA	Eixo_Logradoiro	Sistema de Transporte	-	-	Linha (L)	TRA_Eixo_Logradoiro_L: alfamérico,80	X	-	-	-1
TRA	Ponte	Sistema de Transporte	Ponte	-	Ponto (P) ou Linha (L)	TRA_Ponte_P: alfamérico,80	X	X	-	1
TRA	Ferrovia	Sistema de Transporte	Via_Ferreira	-	Linha (L)	TRA_Via_Ferreira_L: alfamérico,80	X	X	X	1
TRA	Hidrovia	Sistema de Transporte	Hidrovia	-	Linha (L)	TRA_Hidrovia_L: alfamérico,80	X	X	X	0
LIM	Limite_Municipal	Limite	Município	-	Polygono (A)	LIM_Município_A: alfamérico,80	X	X	X	1
LIM	Limite_Bairros	Limite	Bairro	-	Polygono (A)	LIM_Bairro_A: alfamérico,80	X	X	-	1
LIM	Limite_Qradas	Limite	-	-	Polygono (A)	LIM_Qrada_A: alfamérico,80	X	-	-	-1
REF	Rede_de_Reféncia	Pontos de Referencia	Pto_Ref_Geod_Topo	-	Ponto (P)	REF_Pto_Geod_P: alfamérico,10	X	X	X	-1
REL	Curve_de_Nível	Relevo	Curva_Nível	-	Linha (L)	REL_Curva_Nível_L:Real,5	X	X	X	1
VEG	Matas	Vegetação	Veg_Natural	Mata	Polygono (A)	VEG_Mata_A: alfamérico,80	X	-	-	1
VEG	Mangue	Vegetação	Veg_Natural	Mangue	Polygono (A)	VEG_Mangue_A: alfamérico,80	X	-	-	0
INF	Educação	Educação e Cultura	Org_Ensino	Area_Ensino	Polygono (A)	INF_Area_Ensino_A: alfamérico,100;	X	-	-	1
INF	Educação	Educação e Cultura	Org_Ensino	Edt_Ensino	Polygono (A) ou Ponto (P)	INF_Area_Ensino_P: alfamérico,100; INF_Edt_Ensino_A: alfamérico,100;	X	-	-	1
INF	Lazer	Educação e Cultura	Complexo_Lazer	Area_Lazer	Polygono (A) ou Ponto (P)	INF_Area_Lazer_A: alfamérico,100; INF_Area_Lazer_P: alfamérico,100;	X	-	-	1
INF	Lazer	Educação e Cultura	Complexo_Lazer	Campo_Nquadra	Polygono (A) ou Ponto (P)	INF_Campo_Nquadra_A: alfamérico,100; INF_Campo_Nquadra_P: alfamérico,100;	X	-	-	1
SAN	Rede_Esgoto	Abastecimento de Água e Saneamento	Complexo_Abastecimento_Agua	Rede_Esgoto	Linha (L)	INF_RedEsgoto_A: alfamérico,40;	-	-	-	0
SAN	Rede_de Água	Abastecimento de Água e Saneamento	Complexo_Abastecimento_Agua	Rede_Agua	Linha (L)	INF_Red_Agua_A: alfamérico,40;	-	-	-	0

Figura 11. Trecho da tabela análise da consistência da base de dados do CACI.

Attributes of facequadraCDUshp											
FID	Shape *	CSETCECO01	COUASEC001	CFACESHUM	CDSDADM	DSOF	CEP	CLOGRAC001	ILGPAVOFIC	ILGPAV	
*	Polyline	2070	14	2	4	420700014002	50000000	110507	Refúgio Campus da UFPE-4	Ref Campos da Utpe-4	
1	Polyline	2070	7	3	4	420700017003	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amiz	
2	Polyline	2070	12	1	4	420700012001	50000000	8516	Avenida Professor Artur de Sá	Av Prof Artur de Sa	
3	Polyline	2070	12	2	4	420700012002	50000000	20583	Avenida dos Visconde de São Leopoldo	Av dos Visconde de São Leopoldo	
4	Polyline	2070	12	4	4	420700012004	50000000	8214	Avenida da Arquitetura	Av da Arquitetura	
5	Polyline	2070	12	5	4	420700012005	50000000	97780	Avenida Jornalista Aníbal Fernandes	Av Jorn Aníbal Fernande	
6	Polyline	2070	12	6	4	420700012006	50000000	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	
7	Polyline	2070	12	8	4	420700012008	50000000	35956	Avenida Visconde de São Leopoldo	Av Visc de São Leopoldo	
8	Polyline	2070	12	9	4	420700012009	50000000	35957	Avenida Professor Luz Freire	Av Prof Luz Freire	
9	Polyline	2070	12	10	4	420700012010	50740530	35958	Rua Acadêmico Hélio Ribeiro	R Acadêmico Hélio Ribeir	
10	Polyline	2070	8	2	4	420700008002	50740590	20583	Avenida dos Economistas	Av dos Economistas	
11	Polyline	2070	18	3	4	420700018003	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amiz	
12	Polyline	2070	17	4	4	420700017004	50000000	110406	Refúgio Campus da UFPE	Ref Campos da Utpe	
13	Polyline	2070	17	5	4	420700017005	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	
14	Polyline	2070	18	2	4	420700018002	50000000	110469	Refúgio Campus da UFPE	Ref Campos da Utpe	
15	Polyline	2070	17	2	4	420700017002	50000000	110477	Refúgio Campus da UFPE-1	Ref Campos da Utpe-1	
16	Polyline	2070	16	2	4	420700016002	50000000	110465	Refúgio Campus da UFPE-2	Ref Campos da Utpe-2	
17	Polyline	2070	15	4	4	420700015004	50000000	110485	Refúgio Campus da UFPE-2	Ref Campos da Utpe-2	
18	Polyline	2070	14	4	4	420700014004	50000000	110485	Refúgio Campus da UFPE-3	Ref Campos da Utpe-3	
19	Polyline	2070	9	1	4	420700009001	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	
20	Polyline	2070	9	4	4	420700009004	50740550	8214	Avenida da Arquitetura	Av da Arquitetura	
21	Polyline	2070	9	2	4	420700009002	50740590	97780	Avenida Jornalista Aníbal Fernandes	Av Jorn Aníbal Fernande	
22	Polyline	2070	9	3	4	420700009003	50740560	97780	Avenida Jornalista Aníbal Fernandes	Av Jorn Aníbal Fernande	
23	Polyline	2080	4	1	4	420800004001	50740510	10568	Rua Dr. Barão de Monteiro	R Dr. Barão de Monteiro	
24	Polyline	2080	4	1	4	420800004001	50740510	10569	Rua Dr. Barão de Monteiro	R Dr. Barão de Monteiro	
25	Polyline	2080	4	3	4	420800004003	50741430	43036	Rua Mário Campelo	R. Mario Campelo	
26	Polyline	2070	8	4	4	420700008004	50740550	8214	Avenida da Arquitetura	Av da Arquitetura	
27	Polyline	2070	16	4	4	420700016004	50000000	110477	Refúgio Campus da UFPE-1	Ref Campos da Utpe-1	
28	Polyline	2070	16	3	4	420700016003	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	
29	Polyline	2070	16	1	4	420700016001	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	
30	Polyline	2070	15	2	4	420700015002	50000000	110493	Refúgio Campus da UFPE-3	Ref Campos da Utpe-3	
31	Polyline	2070	15	3	4	420700015003	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	
32	Polyline	2070	14	1	4	420700014001	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	
33	Polyline	2070	14	3	4	420700014003	50740570	34533	Avenida Reitor Joaquim Amazônas	Av Retor Joaquim Amaz	

Figura 12. Atributos da feição FaceQuadradas.

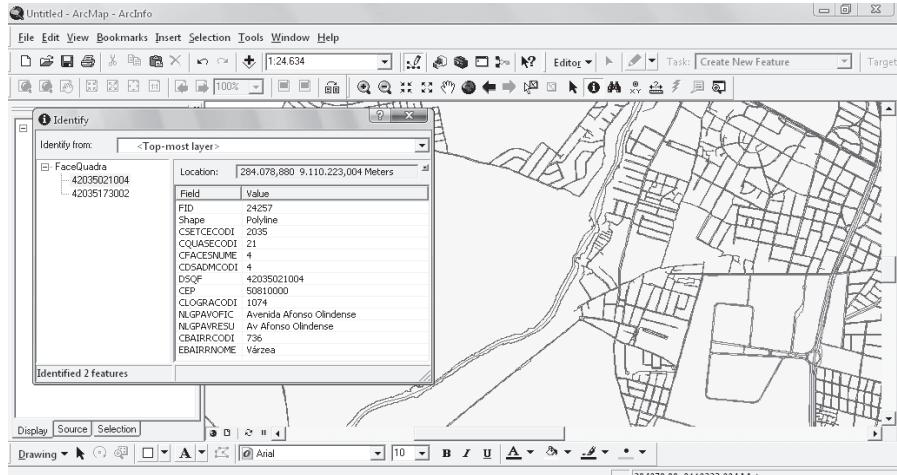


Figura 13. Informações de atributos da feição selecionada.

c) Consistência Conceitual:

A análise da consistência conceitual avaliou se as informações fornecidas na base de dados eram compatíveis ao universo abstrato desejável, definido previamente, e se ocorreu a relação lógica da informação entre os dados quanto ao seu uso e propósito. O universo abstrato considerado nessa pesquisa correspondeu às especificações técnicas previstas na ET-EDGV.

O Quadro 4 mostra os critérios utilizados para a avaliação da consistência conceitual, considerando a sua concordância com o planejamento pré-definido, nesse caso o atendimento às especificações da ET-EDGV.

Quadro 4
Valores para a avaliação da consistência conceitual

<i>Valor</i>	<i>Descrição</i>	<i>Definição</i>
1	Aceitável	A entidade concorda com o indicador estabelecido no universo abstrato
0	Não se aplica	Entidade não está presente no universo abstrato da carta controlada
-1	Omitida	A entidade foi omitida devendo ser corrigida através da integração de banco de dados da carta.

A Figura 11 apresenta os resultados de parte da análise da consistência conceitual da base de dados do CACI, considerando o modelo conceitual proposto na ET-EDGV.

Na análise dos conceitos (última coluna da Figura 11), o valor 0 (zero) corresponde às feições que não foram analisadas; o valor -1 (menos um), quando a feição não respeita o modelo conceitual aceitável, devendo ser corrigida; o valor 1 foi atribuído à feição que possui uma modelagem aceitável.

A partir dos conceitos obtidos na tabela de verificação do modelo conceitual, é possível calcular o desvio da consistência conceitual dos dados estudados. O percentual de não conformidade tolerável, previsto na ISO 2859-1, é de 10% e depende da quantidade de feições testadas:

$$\text{Desvio} = (100 * E) / T$$

sendo, T: Total de Conceitos
E: Soma Total de valores -1

Assim, a partir dos valores obtidos (a Figura 11 apresenta 20, mas o total de conceitos testados é 26), obtendo-se T= 26 e E =4, o que corresponde a um desvio de 15,38% entre os modelos conceituais do CACI e da norma ET-EDGV.

O resultado indica, portanto, a necessidade de adequação do modelo do CACI ao proposto pelas normas da INDE.

d) Consistência de Domínio:

A consistência lógica pode ser entendida como uma variável lógica (aprovação / reprovação) e serve para mostrar que não há casos imprevistos na base de dados. O teste deste indicador é feito com relação ao Controle de Códigos e de Atributos,

onde o primeiro garante que não existem casos com códigos que não estão no catálogo dos fenômenos admissíveis e o segundo garante que os atributos alfanuméricos que descrevem o objeto estão incluídos, e também que os seus valores pertencem ao domínio fornecido.

Assim, está diretamente relacionada à consistência conceitual e ao preenchimento da tabela de atributos. Por exemplo, as Figuras 12 e 13 mostram que a feição do tipo TrechoLogradouro será uma feição do tipo linha, que pode receber os seguintes atributos: CEP (tipo inteiro), Código do Logradouro (tipo inteiro), Nome do Logradouro (tipo caracteres), Diâmetro do logradouro (tipo real), entre outros.

As Figuras 12 e 13 apresentam ainda o exemplo da variável lógica aprovada para a feição FaceQuadra. Em relação ao controle de código ela é aceitável, pois é um dos fenômenos admissíveis no modelo conceitual do universo abstrato trabalhado, estando dentro de LimiteQuadra. Em relação ao controle de atributos, quando uma coluna é definida com o tipo inteiro, espera-se que todos os valores atribuídos a essa coluna sejam números inteiros, não podendo ser atribuída a essa coluna uma sequência de caracteres, e vice-versa.

A Consistência de Domínio é a responsável por detalhar as informações corretas quanto aos atributos de uma feição. Não se verificou inconsistência de domínio nas feições trabalhadas, porque de uma maneira geral o domínio é predefinido na elaboração dos dados, variando conforme a escala e tipo de feição. Além disso, suas tabelas de atributos já estavam todas definidas quanto ao tipo de declaração de variável e devidamente preenchidas.

Conclusões

Falhas de qualidade geram custos tangíveis e intangíveis. Quando são identificadas falhas na elaboração de um produto cartográfico, por exemplo, este poderá ser reprocessado ou até mesmo descartado da produção, precisando ser refeitas etapas por não atender as especificações.

Este trabalho apresentou a metodologia e resultados da análise de um destes indicadores, a consistência lógica, aplicados à base cartográfica do CACI - Cadastro de Interferências, utilizado para a concessão de autorizações de obras na Região Metropolitana de Recife-PE, Brasil. Outros testes são apresentados e discutidos em Maranhão (2013).

A verificação de consistência lógica, realizada conforme os subelementos observados na ISO 19113 e na INDE, avaliou a qualidade dos dados quanto aos aspectos de formato, topologia, conceituais e de domínio.

Quanto ao formato, foram realizadas conversões de dados vetoriais para SHP, e de dados raster para TIFF, formatos reconhecidos pela maioria dos programas de SIG. Do ponto de vista da topologia, foram utilizados procedimentos para verifica-

ção e eliminação de erros grosseiros, antes da realização das correções topológicas, até que a base analisada tenha sido considerada livre destes erros.

A partir da verificação do modelo conceitual, foi possível calcular o desvio da consistência conceitual dos dados estudados. O percentual de não conformidade tolerável, previsto na ISO 2859-1, é de 10% e depende da quantidade de feições testadas. O teste da base analisada apresentou um desvio de 15,38% entre os modelos conceituais do CACI e da norma ET-EDGV, indicando a necessidade de adequação do sistema ao proposto pelas normas da INDE.

A metodologia utilizada mostrou resultados eficientes permitindo, inclusive, compatibilizar os dados ao modelo conceitual previsto na INDE, através da ET-EDGV.

Bibliografía

- Ariza, F.J. y Alcázar, M.G. (2010). “Calidad e Información Geográfica Cadastral”, *Fórum Geográfico*, Ano III (3).
- Barros, E.R.O. (2011). “Uma Proposta para o Controle de Qualidade do Processo de Certificação de Imóveis Rurais”. Dissertação, Mestrado em Ciências Geodésicas e Tecnologias da Geoinformação, Universidade Federal de Pernambuco, Recife, 170 p.
- Ceballos, J.R. y Gatica G.R. (2012). “Implementacion de un Sistema de Gestión y Control de Calidad bajo las normas ISO 9001:2000, 19113, 19114, 19115 y 19138 em la producción cartográfica”, 2012. Disponível em :<http://www.powershow.com/view/28b565-YTVmZ/IMPLEMENTACION_DE_UN-SISTEMA_DE_GESTION_Y_CONTROL_DE_CALIDAD_BAJO_LAS_NORMAS_ISO_90012000_19113_191_powerpoint_ppt_presentation>, acessado em 12 de junho de 2013.
- Delavar, M.R. and Devillers, R. (2010). “Spatial Data Quality: From Process to Decisions”, *Transactions in GIS*, no. 14, pp. 379-386.
- Francisco, H.R. (2001). “Qualidade de dados espaço-temporal: estudo de caso no contexto da acurácia posicional e atualização”, Dissertação (Mestrado em ciências Cartográficas), Universidade Estadual Paulista, Presidente Prudente, 157 pp.
- Gatica, N.P. (2010). “SIG y Consistência Logica”, *Notas de Aula da Universidade de Santiago de Chile*, USACH, 2010. Disponível em: <<http://pt.scribd.com/doc/-105524431/SIG-y-Consistencia-Logica>> Acessado em 11 de julho de 2013.
- Gong P., Mu L. (2000). “Error detection through consistency checking”, *Geogr Inf Sci*, 6(2):188-193.

- ISO. ISO/TC211 (2010). *Guía de normas. Grupo Consultivo de Desarrollo Edición en español Comité ISO/TC 211: Información Geográfica/Geomática*, México, 116 pp.
- ISO. ISO 19113 (2002). *Geographic Information – Quality principles*, Geneva, Switzerland, 29 pp.
- Longley, P.A., Goodchild, M.F.; Maguire, D.J.; Rhind, D.W. (2013). “Sistemas e ciências da Informação geográfica”, 3a. ed., Porto Alegre, Bookman, 540 pp.
- Maranhão, V.C. "Modelo e controle de qualidade de uma Infraestrutura de Dados Espaciais para o estado de Pernambuco. *Dissertação (Mestrado)*, Centro de Tecnologia e Geociências, Universidade Federal de Pernambuco, Recife, 2013, 140 pp.
- Wan, Y., Shi, W., Gao, L. et al. (2015). “A general framework for spatial data inspection and assessment”, *Earth Sci Inform*, 8(4):919-935.
- Weber, E.; Anzolch, R.; Filho, J.L.; Costa, A.C.; Iochpe, C. (1999). “Qualidade de Dados Geoespaciais”, *Dissertação (Mestrado)*, Universidade Federal do Rio Grande do Sul, UFRS.