Automation of Indonesian Topographic Data Quality Control Using Data Reviewer

Danang Budi Susetyo*, Aji Putra Perdana
Center for Topographic Mapping and Toponym – Badan Informasi Geospasial, Indonesia

Abstract
Article 49 of Indonesian Law Number 4 Year 2011 states that Geospatial Information user has the right to know about the quality of data obtained. The statement shows that quality control (QC) must be done to every geospatial data produced by Geospatial Information Agency in order to ensure quality assurance (QA) to the data. Quality control is essential in spatial data production process to guarantee the data meet with specifications set. Implementation of QC achieved by fulfilling QC document of technical specification for each output. Spatial data QC can be done automatically and consistently using Data Reviewer in ArcGIS, by adapting to specification in QC document. This research mainly discusses the implementation of topographic mapping data quality control using ArcGIS Data Reviewer conducted by Geospatial Information Agency. These research results have been tested in QC implementation for Indonesian Topographic data production at medium scale in 2015 and proven by practices to support QC implementation in the steps of digitization and hypsography extraction.

Keywords: automation; quality control; Indonesian topographic data; ArcGIS data reviewer

INTRODUCTION
Article 49 of Indonesian Law Number 4 Year 2011 states that Geospatial Information user have the right to know about quality of data obtained. The statement shows that quality control (QC) must be done to every geospatial data produced by Geospatial Information Agency. Quality control aims to detect deviations so that Quality Assurance (QA) can be done to ensure these data meet the standards of quality (Wisnu, Riqqi, Harto, & Mulyana, 2014).

Categories of source errors create primary errors (positional errors and attribute errors) and secondary errors (logical consistency and completeness) (Devillers & Jeansoulin, 2013). It is aligned with The ICA Commission of Spatial Data Quality that defined seven main parameters of spatial data quality description and carried out other initiatives for standardization: positional accuracy, attribute accuracy, lineage, completeness, logical consistency, semantic accuracy, temporal information (Mauro Caprioli & Tarantino, 2003). At the ontological level, the internal consistency of the specifications is considered (Mostafavi, Edwards, & Jeansoulin, 2004), so that seven main parameters should be included in the spatial data specification that will be produced. The specification covers all the provisions contained in Terms of Reference. Moreover, the users of these data bases should be the ones who decide which of the available bases could satisfy their requirements, or in other words, what is the data quality level necessary for certain application (Joksic & Bajat, 2004), which is in line with the statement of M Caprioli, Scognamiglio, Strisciuglio, & Tarantino (2003) who stated that necessity of analysis and scalability issues must be considered in GIS data quality. The provisions are then listed in a document of quality control containing specification items,
whereas data produced can pass quality control if they are in accordance with all items in the document.

![Diagram](image)

**Figure 1.** The EPA Quality System Approach to Addressing Geospatial Data Applications (United States Environmental Protection Agency, 2003)

Although quality control parameters have been defined, the element of subjectivity between quality control officer may still happen, especially in visual review. Quality control automation can reduce that subjectivity. Tools that can support quality control automation for spatial data is ArcGIS Data Reviewer. ArcGIS Data Reviewer is also part of the Esri Production Mapping extension that can be purchased for ArcGIS Desktop (Esri Indonesia, 2014).

This research discusses about ArcGIS Data Reviewer application in quality control implementation of Indonesia Topographic Map. Quality discussed in this research is about data standards, the set of public data, and process standards, or in Gabriela, Stefan, & Aurelian (2010) defined as semantic level. Research result has been tested in topographical mapping scale 1:50,000 in Kalimantan, Indonesia, 2015, using 2D digitization method. In 2016, this research application re-tested in topographical mapping scale 1:25,000 using stereoplotting method. Quality control automation using ArcGIS Data Reviewer is proven to support quality control implementation for topographical mapping in Indonesia.

**METHODS**

This research is carried out by adjusting to QC parameter in Center for Topographic Mapping and Toponym, Geospatial Information Agency with tools provided in ArcGIS Data Reviewer. Customized steps are digitization/stereoplotting and hypsography extraction.

**RESULT AND DISCUSSION**

1. **Visual Review**

   Visual review is used to check the completeness of stereoplotting result against base data (such as radar or optical image data). Method used is to divide one sheet to some grids so that checking can be done more systematically. These grids are used for area visualization in one single checking, which then marked to avoid double QC in one area.

   Marking the wrong parts can be done in three ways:
   
   a. **Notepad**
      
      Notepad is used to mark errors based on three basic geometry in Geographic Information System: point, line, and polygon. QC officers mark the parts that they consider as deficient, whereas this perception of QC officer can be followed by operator. It minimizes the differences
in perception which usually occurred when QC result is only a circle or simply less detailed information.

Figure 2. Display of visual review using grid

Figure 3. Use of Notepad

b. Commit to Reviewer Table
Commit to Reviewer Table is used to mark a segment of the database which is the result of stereoplotting, whereas it then follows the segment that is marked as error. For example, if a stream segment incorrectly defined as seasonal stream when it should be defined as one line stream, Commit To Reviewer Table can directly mark the incorrect seasonal stream segment.
Figure 4. Commit To Reviewer Table

Below is an example of Commit To Reviewer Table use to mark a stream which should be plotted as two line stream, but is plotted as one line stream by operator. Mark given on the Commit To Reviewer Table follows the geometry of the stream that is considered incorrect.

Figure 5. Example Use of Commit to Reviewer Table

c. Flag Missing Feature
Flag Missing Feature is used to mark objects which has not been plotted. Its use is preferred to objects in forms of point or small polygon, so it can be marked by one point geometry. Flag Missing Feature is suitable to mark un-plotted building object, as shown in the figure below.

Figure 6. Flag Missing Feature

2. Automated Review
Automated review is used by adjusting tools in Data Check options to QC parameter. The adjustments are:
Stereoplotting steps
a. Feature Code (FCODE) in accordance with List of Topographic Elements Code
This QC parameter is checked using Domain Check. Domain Check is not used to check suitability of feature code in database attribute to the list of topographic elements code, but to check blank feature code (null). This deficiency can be detected because FCODE is attached with domain.

![Domain Check](image1)

**Figure 7. Domain Check**

b. Each object can only be captured once
This parameter requires no duplicate object. It can be checked using Duplicate Geometry Check, that can detect two or more objects with the same geometry and position.

![Duplicate Geometry Check](image2)

**Figure 8. Duplicate Geometry Check**

c. Stream should not dangle and must form a network
This parameter is relevant to topology rule stating that stream intersect should be snap (not undershoot or overshoot). This is suitable with Find Dangle Check tool.
d. All streams should align in one segment
   One segment means it is neither multipart or pseudonodes. This parameter is
   accommodated by Multipart Line Check and Unnecessary Nodes Check.

Figure 9. Find Dangle Check

Figure 10. (a) Multipart Line Check; (b) Unnecessary Nodes Check

e. All roads must be connected and form a network
   This parameter uses the same method as the feature used in stream (point d).

f. Runway and dock with planimetric detail > 12.5m x 12.5m should be plotted according
   to its form
   Evaluate Polyline Length Check can be used for this parameter, wherein if an area is
   less than 12.5m x 12.5m it will be detected as error because it is less than minimum
   polygon specification.
g. Conformity of bridge geometry to road and stream
   This parameter requires bridge to be placed in the intersect between road and stream.
   Check option that can be used for this parameter is Intersection on Geometry Check,
   where intersect must be done to road and stream feature.

h. Buildings sized > 12.5m x 12.5m are plotted using line
   This parameter concept is similar to runway (point f), so checking option to this
   parameter is the same, which is Evaluate Polyline Length Check.

i. Land cover line must snap to road or stream if they meet each other
   This parameter also uses Find Dangle Check, but involves three forming elements
   of land cover: open land vegetation line, hydrography, and transportation. These features
   are first merged to one feature, and then processed using Find Dangle Check, with more
   focus on checking to open land vegetation line.

j. Each land cover has annotation (point) in accordance with the type of land cover
   This parameter is checked using Topology Rules Check, more accurately with Contains
   One Point rule. The concept is that every land cover polygon must have one label point.
   Error will be detected if one polygon has more than one label point or no label point at
   all.
Technically, the checking is done by creating temporary land cover polygon from the combined file of constituent elements of land cover processed at point i. It is then checked against points of label made by operator using Topology Rules Check.

![Figure 13. Topology Rules Check](image)

k. Intersect line features have the same elevations in the vertex
   This parameter can be checked using Different Z at Intersection Check, where vertex intersection with different elevation can be detected.

![Figure 14. Different Z at Intersection Check](image)

1. The outline of stagnant waters (lakes, ponds, marshes, etc) has the same elevations each vertexes
   This parameter is referred to plotting concept of hydrography feature wherein plotting is done on the borderline between water and land, assuming that the outline of stagnant water area is its surface, which means each vertex has the same elevation. This parameter can be checked using Adjacent Vertex Elevation Change Check.
Figure 15. Adjacent Vertex Elevation Change Check

m. Vertical elevation values of the stream and irrigation area consistently decreases from upstream to downstream
   This parameter can be checked using Slope Direction Change (Monotonicity) Check, which can also identify errors of stream flow direction.

Figure 16. Slope Direction Change (Monotonicity) Check

Hypsography extraction steps
a. Contour is in accordance with hydrography and building features
   This parameter states that contour is not allowed to intersect building (if it’s plotted as line), and if contour intersect with stream, it should be indented upstream. The conformity of contour to building can be checked using Geometry on Geometry Check.
b. No dangle contour, with exception of auxiliary contour
Checking to this parameter is similar to those features that are not allowed to dangle, so it can use Find Dangle Check.

c. Contour interval 10m
This parameter depends on the scale; however, since this research uses scale 1:25,000 databases, so contour interval used as sample is 10m. This can be checked using Execute SQL Check, by performing query to contour value which is not a multiple of 10.

d. Index contour interval 50m
This parameter can also be checked using Execute SQL Check, by adjusting the value to 50 and first performing definition query to index contour, so that only index contour will be displayed.

e. Auxiliary contour interval 5m
This parameter is similar to point d, but only auxiliary contour will be displayed, and checked value was adjusted to 5.

f. Contours do not intersect each other
This parameter is also checked using Geometry on Geometry Check, to identify intersect contour.
g. Contour does not intersect the same stream segment more than once
   This parameter can be checked using Evaluate Intersection Count Check.

![Image of Evaluate Intersection Count Check]

**Figure 19. Evaluate Intersection Count Check**

h. Contour does not intersect waterline
   This parameter concept is similar to point a and f, with input features are contour and hydrography, so checking option used is Geometry on Geometry Check.

i. Spot height value is not same with contour value
   This parameter can be checked using Table to Table Attribute Check, where elevation value from these features can be identified as error.

![Image of Table to Table Attribute Check]

**Figure 20. Table to Table Attribute Check**

3. Batch Review
   Batch review can simplify and accelerate automated review. The concept is to create batch jobs, which are groups of checks that can be saved and run against the data to validate many specifications simultaneously (Esri Indonesia, 2014).
CONCLUSIONS
Article 49 of Indonesian Law Number 4 Year 2011 states that Geospatial Information users have the right to know the quality of data obtained. It becomes the main reason for the need of quality control in spatial data production. Quality control automation using tools can minimize subjectivity. Tools that can support quality control automation for spatial data is ArcGIS Data Reviewer.

ArcGIS Data Reviewer has three functions: visual review, automated review, and batch review. Visual review is used to check the completeness of stereoplotting result against base data (such as radar or optical image data). Automated review is used by adjusting tools in Data Check options to QC parameter. Batch review can simplify and accelerate automated review, by creating batch jobs, which are groups of checks that can be saved and run against the data to validate many specifications simultaneously. By using these three methods, spatial data quality control can be done more quickly and systematically.

ACKNOWLEDGMENTS
We would like to show our gratitude to Center for Topographic Mapping and Toponym, Geospatial Information Agency for data used in this research, and opportunity to apply this research result to medium scale topographic mapping job in 2015 and 2016.

REFERENCES


