A DEM Quality Dashboard

Ariza-López, Francisco Javier; García-Balboa, José Luis.
Dpto. Ingeniería Cartográfica, Geodésica y Fotogrametría
Universidad de Jaén
Jaén, Spain
Email: fjariza@ujaen.es, jlbalboa@ujaen.es

Reinoso-Gordo, Juan Franciso.
Dpto. Expresión Gráfica Arquitectónica y en la Ingeniería
Universidad de Granada
Granada, Spain
Email: jreinoso@ugr.es

***Abstract*—A more holistic and applied view of data quality is considered and to better communicate it, the use of dashboards is proposed. The use of dashboards is very common in business management and quality management, but not for communicating quality of geospatial data. The definition of the use case, the establishment of the key performance indicators and how to make their graphic representation in the dashboard are key processes to achieve this objective. This paper presents our considerations and progress focused on the development of this proposal.**

Keywords-Quality dashboard; data quality; fitness for use; DEM.

I. INTRODUCTION

There are several directly related digital models such as Digital Terrain Models (DTMs), Digital Surface Models (DSMs) and Digital Elevation Models (DEMs). Although the proposal made here may be valid for all of them, it will focus on the case of DEMs. A DEM is a bare-earth terrain void of vegetation and man-made features [1]. DEMs are a key data type for many applications domains because they provide the height component in Geographic Information System (GIS) analysis, the geomorphological description of the land, a reference surface for all hydrological applications (water cycle, erosion, floods, etc.), the basis for the development of forestry models, the base for agricultural parcel rating and are useful in every analysis task related to civil engineering.

The most usual way to evaluate the quality of a DEM is through Positional Accuracy Assessment Methods (PAAMs) (e.g., the National Standard for Spatial Data Accuracy). The use of PAAMs is the data quality approach preferred by official data producers. This is a datacentric perspective (internal quality). But this perspective does not adequately communicate the quality of a DEM data set when it is used in very common applications, such as the calculation of slopes, aspects, drainage networks, etc. This situation means that the data about the quality of a DEM supplied by producers is not understood by a large majority of users [2]. There are many references that indicate the existence of this handicap [3].

We believe that a more holistic view of the quality of a dataset is needed. But this holistic view must also be more applied, more focused on fitness for use. From the statistical point of view, a holistic perspective means a multivariate approach, as opposed to the univariate perspective (positional accuracy only). A more holistic view means considering all those aspects that define the whole, which is really complex. For this reason, the problem must be centered, and what the whole is must be well defined. In order to reach a more fitness for use orientation, we consider that the definition of use cases can be an adequate tool to define the framework of interest (the partial whole of interest). Basically, a use case is the description of an action or process with a certain level of formalization (e.g., using Unified Modelling Languages —UML— diagrams, or any other language). In addition, to better communicate quality, it may be appropriate to use graphic tools, which offer the user a friendlier view. In this sense, we are not referring to the representation of uncertainty (for example, in a terrestrial communication network using buffers of a certain width), but to the graphic expression of all the variables that are of interest to better communicate all the aspects the quality of a data set as a whole.

The above situation, where there is a need to communicate objective information about a complex situation as a whole, is not unique to the field of data quality. For example, in the field of business management it is very common to use numerous variables to define the operation or behavior of a company. In this case, it is usual to resort to the so-called “dashboards”. A dashboard is a type of graphical interface which often provides at-a-glance views of Key Performance Indicators (KPIs).

In the geomatics field, dashboards are used in business management, but there are few references related to data quality [4][5]. More common initiatives are related to the use of stars as a rating system (e.g., five stars for open data [6]). Recent DEM quality reviews [7][8] do not indicate the use of graphical elements (e.g., dashboards) to report DEM data quality. On the other hand, there are numerous works [2][3] indicating that users of DEM data do not have a good understanding of quality aspects. For all these reasons, there is a need to make communication regarding the quality of DEM data more understandable.

In line with all of the above, the objective of this work is to establish a dashboard proposal to offer a multivariate

 ISBN: 978-1-61208-983-6

quality assessment control panel for a specific use case of DEM data. Therefore, it is also our goal to establish a specific use case on which to develop the dashboard. First, we will define the use case, and then a dashboard proposal will be made. Therefore, the ultimate purpose of our contribution is to convey in a simple and understandable way the quality of the DEM data when considering a specific use case.

The structure of the paper is as follows: Section II presents major considerations that are needed to stablish a use case, Section III is a short discussion informs about our next steps. Finally, Section IV presents a brief conclusion.

II. DEFINING A USE CASE

Reference [3] recompiles several use cases of DEM data, and also the result of a worldwide survey on the main users and uses of DEM data. From this background, we consider that a use case that may be of great interest is the determination of a hydraulic network. A brief definition of the use case can be the following: given a DEM data product "DEMproduct", a well-defined sequence of known algorithms {A1, A2…} are applied to the DEMproduct dataset to obtain automatically a vector dataset corresponding to the drainage network. It should be noted that no operator intervention is considered. In this case, a multivariate quality assessment can be proposed if an adequate reference dataset exist or is created. The existence of the reference is a critical point for all accuracy assessments, but is out of the scope of this work.

Once the use case has been defined, basically through the set {data, algorithms, processes}, we have to define the content and configuration of the dashboard. In relation to its content, this tool must consider all those variables that are of interest from an applied perspective of the data generated and whose quality could be evaluated. For example, relevant aspects of the hydraulic network generated and that can be considered here are: the planimetric positional accuracy of the sections of the network, the positional accuracy of the nodes, the existence of surplus or missing edges, the accuracy of the altimetric profiles of the channels generated, the classification of the typology of the generated network, etc. We consider that user participation is necessary to decide which of these aspects are the most relevant KPIs. The ISO 19157 data quality model [9] will be applied to these KPIs to establish standardized definitions of quality elements and their measurements. Reference [10] presents an example that can adequately serve as a guide on the use of this geospatial data quality model. This example shows the combined use of "quality elements", "quality measures" and "data quality units" as defined by the ISO 19157 standard. In addition to the above, it must also be defined how the results are presented so that they communicate properly. Here the structure of the dashboard (quality dimensions) is as important as the tools (e.g., traffic lights, stars, histograms, box-plot graphics, etc.) that can be used depending on the type of variable being worked with and how you want to represent it. For example, there are variables that could be represented specialized or summarized in a value or in a graph. Another relevant aspect is to know if you want the dashboard to be simply informative (this is what results from the evaluation), or if it has a quality control perspective (the dataset passes/fails in relation to a specific quality element). In relation to graphical tools there is experience, but not with the perspective that is exposed here. We consider that user participation is also very relevant here.

Figure 1 shows an example of a section of two sections of a dashboard in which several KPIs related to the positional accuracy and the network-type classification have been included as traffic light signals.

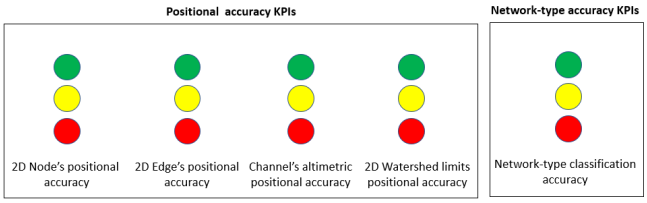

Figure 1. Example of two sections of a dashboard based on traffic light signals.

III. DISCUSSION

The approach of quality to users is a key aspect for the appropriate use of geospatial data. Therefore, the work that has been presented is aligned with a traditional desire in the field of geospatial data, which is to make quality more understandable for users. For this reason, it is very important that the aspects of quality to be considered are those that are really relevant KPIs from a "fitness for use" perspective, but can also be linked to the idea of functional quality (this idea is presented in another contribution of the authors in this congress). The work is just beginning, and the authors are preparing a survey to identify the KPIs for this case, and will subsequently survey various dashboard configurations as well.

The idea has been presented on a single use case, but it can be applied to any use case. Once the most defining use cases of DEM data are available, you can consider developing a complete dashboard for all of them, and such that it offers a holistic view of the quality of a DEM.

An especially critical aspect is the transition from elements of quality, as defined in the ISO 19157 model, to fitness for use, or usability in the jargon of ISO 19157. This is not a specific problem of our proposal, but it is a handicap that affects. The communication capacity focused on the use case would be much greater if this handicap were overcome.

IV. CONCLUSSIONS

The usefulness of dashboards is unquestionable in the fields of business management and quality management systems. This work presents an idea that already exists but has not received enough attention in the field of geospatial data quality. In order to advance in this line, it has been proposed to link it to use cases, such that they allow to limit the scope of work. The work is in progress, user participation is required to establish the KPIs and the dashboard configuration. We hope that these works will be the basis for offering a standardized communication tool for DEM's quality.

 ISBN: 978-1-61208-983-6

ACKNOWLEDGMENT

This work has been financed by the research project "Functional Quality of Digital Elevation Models in Engineering" of the State Agency Research of Spain. PID2019-106195RB-I00/AEI/10.13039/501100011033 (https://coello.ujaen.es/investigacion/web_giic/funquality4dem/).

REFERENCES

[1] D. F. Maune, Digital Elevation Model Tecnologies and Applications: The DEM Users Manual. ASPRS (2001).

[2] S. P. Wechler, "Perceptions of Digital Elevation Model Uncertainty by DEM Users," Urisa J. vol 15, pp. 57–64, 2003.

[3] F. J. Ariza-López, E.G. Chicaiza-Mora, J. L. Mesa-Mingorance, J. Cai, and J. F. Reinoso-Gordo, "DEMs: An Approach to Users and Uses from the Quality Perspective," Int. J. Spat. Data Infrastruct. vol 13, pp. 131–171, 2018.

[4] E. Folmer and W. Beek, "Data Quality for Use: A Linked Data Approach" The 3rd International Workshop on Spatial Data Quality (SDQ 2020) January 28th- 9th, pp 22-25, 2020.

[5] J. Nogueras-Iso, J. Lacasta, M. A. Ureña-Cámara, and F. J. Ariza-López, "Quality of Metadata in Open Data Portals," IEEE Access 2021, 10.1109/ACCESS.2021.3073455

[6] https://opendatahandbook.org/glossary/en/terms/five-stars-of-open-data/

[7] J. L. Mesa-Mingorance, and F. J. Ariza-López, "Accuracy Assessment of Digital Elevation Models (DEMs): A Critical Review of Practices of the Past Three Decades", Remote Sensing, 12(16):2630 2020, https://doi.org/10.3390/rs12162630

[8] Polidori L, and M. El Hage, "Digital Elevation Model Quality Assessment Methods: A Critical Review", Remote Sensing., 12(21):3522, 2020, https://doi.org/10.3390/rs12213522

[9] International Standardization Association, ISO 19157:2013 Geographic information — Data quality, ISO, 2013.

[10] F.J. Ariza-López, J. Rodríguez-Avi, J. F. Reinoso-Gordo, and I. A. Ariza-López, "Quality Control of As Built BIM Datasets Using the ISO 19157 Framework and a Multiple Hypothesis Testing Method Based on Proportions", ISPRS International Journal of Geo-Information, 8(12):569, 2019, https://doi.org/10.3390/ijgi8120569

.

 ISBN: 978-1-61208-983-6