Consistency of Data in Cadastral Systems

Miodrag ROIĆ, Croatia

Key words: Data Consistency, Cadaster, Quality Control, Land Administration

SUMMARY

Cadastral systems should provide users with reliable data. However, cadastral systems established by a systematic approach in the 19th century are burdened with errors caused by long-term and uncoordinated maintenance in an analog form on paper. In addition to regular maintenance, systematic documentation reconstructions were carried out. Even during the reconstructions, mistakes were made, which were later converted into the electronic form. The social and political changes that significantly influenced the registrations in the cadasters in the past two centuries are briefly presented. Dynamic changes in land rights caused intensive registrations in the cadaster and thus, illegibility of documentation. Such illegibility documentation made it difficult to convert it into the electronic form, which led to new misinterpretations.

After the conversion into the electronic form, it was possible to detect and correct these errors by using the approach proposed in this paper. First, a typical analog documentation of the cadaster created by a systematic approach is presented and analyzed. The general approach to the digitization of descriptive data and the vectorization of the cadastral map are presented and discussed. Redundant data that were kept in analog documentation and those that were produced during the conversion to the electronic form were detected. A comparison of these redundant data (e.g., parcel area from the register and the area calculated from the vectorized map) is proposed to detect inconsistencies. It is also suggested to compare the constraints that the data should meet to detect inconsistencies. For example, the number of parcels on the cadastral map should be equal to the number of parcels in the register, having the same parcel IDs. A strategic approach to performing comparisons of redundant data was developed to implement error correction efficiently and quickly. A typical case of discovering the cause of the inconsistencies and a proposal for error correction are presented in detail. The presented quick detection and correction of errors increased the quality of data in the cadastral system and strengthened the user confidence in it.

Consistency of Data in Cadastral Systems

Miodrag ROIĆ, Croatia

1. INTRODUCTION

At the onset of the pandemic, many land administration organizations were already well equipped for the new normal, adapting to working at-distance, handling service spikes, responding swiftly to changing customer needs, and delivering novel data analytics services. Legal arrangements to support digital service delivery were often already in place or adaptable at short notice. The event even produced financial gains for many land administration organizations. Nevertheless, it also placed a spotlight on the issues of digital exclusion, data quality, standards, staff capacity, customer awareness, and partner collaboration. These are areas where investment is most urgently needed (Bennett et al., 2022).

Land information is the content of many registers and an important part of the (geo)information infrastructure. Most of the land information is registered into a cadaster, and all other registers are in some way related to it. Such information can hardly be established without a cadaster and is frequently established by obtaining cadastral data and adding a set of data that is significant to the resource for which the register is established. The land information registered in the cadaster is the most comprehensive and the highest quality set of existing geoinformation in many countries (Roić, 2012). Society and economy require the collection, storage, sharing, and analysis of large quantities of spatially and non-spatially referenced reliable data (Williamson, 2001).

However, in many countries, cadastral data were created approximately 200 years ago and were maintained on paper for a long time. Analogous data maintenance caused many errors when implementing changes, and there are often incorrect entries from the period when the cadaster was created. Research has shown that the data of the cadastral map are not consistent with the corresponding data in the register (Fetai et al., 2022; Karabin & Łuczyński, 2022; Roić et al., 2021; Taszakowski et al., 2018). The differences are significant, reduce trust in the cadastral system, and create difficulties for users.

Cadastral systems created by using a systematic approach (Henssen, 1995) have a seamless cadastral map and a corresponding register by cadastral municipalities. The cadastral map provides technical data and registers the corresponding descriptive data. The link is provided through a unique parcel ID.

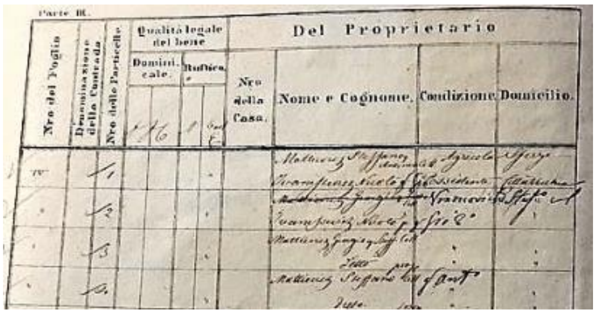
After the digitization of these data, redundant data are provided. They can be used to examine the consistency of data within the cadastral documentation. In this study, cadastral data created after the digitalization of analog documentation was analyzed. Redundant data indicating the inconsistency of data were detected. Guidelines for detecting inconsistencies were developed, and the possibilities of error detection with the help of inconsistencies were determined.

2. STRUCTURE AND CONTENT OF DOCUMENTATION

The basic technical aspects of any cadaster are an unambiguous definition of parcels and a related record providing the required data about these parcels. In many countries, the cadaster was established by using a systematic approach quite a few years ago. According to the technological possibilities of the time, the data were registered and maintained on paper. Parcels and fundamental data concerning them were mapped on the cadastral map, and the rest was registered in books in a tabular form. The descriptive data concerning the parcels and some social relations (e.g., taxpayers) were registered in the register.

The systematic establishment of the cadaster created various documents, including spatial representations, such as cadastral map sheets and lists showing the descriptive data of land features that could not be shown on the cadastral map. The position and relative mutual relationships of parcels, as well as other land features have since been shown on the cadastral map (Figure 1a). Descriptive data concerning parcels can be found in the lists. One of them is the list of cadastral parcels, which serves as the basis for creating further cadastral documents (Figure 1b).

a. Part of a cadastral map



b. Part of a list of parcels

Figure 1. Key cadastral documentation

With the use of the unique parcel ID, the link between the cadastral parcel on the cadastral map and the corresponding data in the list of cadastral parcels can be realized. The list of cadastral parcels is a tabular representation that contains all the cadastral parcels of one cadastral municipality arranged by ordinal numbers. Descriptive data concerning cadastral parcels are entered in the columns. These two documents are the key documents from which the further documents are created and referred to.

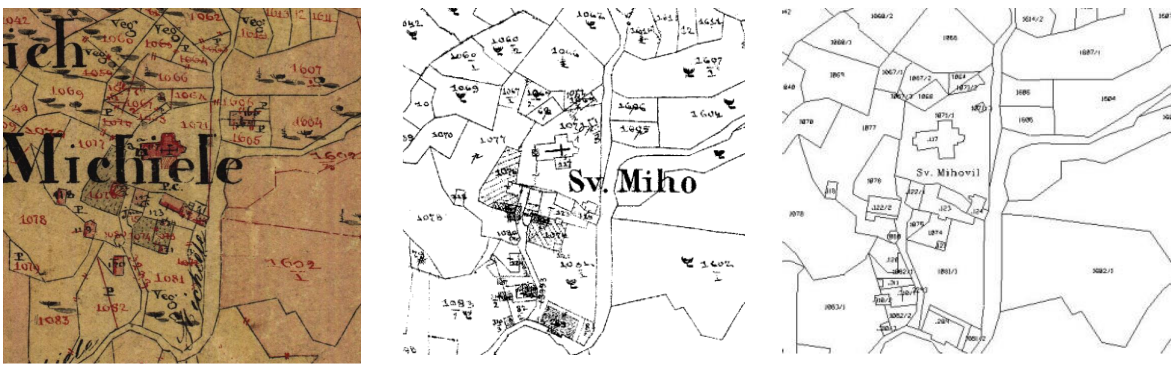
In addition to these two documents, data on taxpayers in possessory sheets were systematized. Furthermore, other documents were prepared that mainly served to fulfill the functions of the cadaster more efficiently. For example, a list of houses, an overview list of parcels by land use, records of area determination, and a register of all toponyms. These are the most common documents but depending on the individual country and the period when the cadaster was established, there may be others. Maintaining these documents was very demanding. It was often necessary to implement one change in several of them in a coordinated manner.

3. MAINTENANCE AND RECONSTRUCTION

Analog cadaster documentation is often damaged in the course of persistent maintenance and use, and it is necessary periodically to reconstruct it physically. Upon the implementation of the changes, documents become unreadable and illegible. This primarily applies to cadastral map sheets, but also to other documentation. The number of land changes is conditioned by social and political events. The turbulent history of Europe in the 20th century influenced the development of cadasters in certain countries. Significant changes in the cadastral systems of the countries of Central and Eastern Europe were caused by radical social and political changes at the end of the 20th century. These countries have had a cadaster since the 18th century, but in the 20th century, major land reforms were implemented in them, and many people were displaced due to war and other events. Two world wars and the introduction of a planned economy by the communist regimes changed the people to land relationships. At the end of the 20th century, with the fall of the communist regimes, these countries returned to a market economy. The return to the market economy in the countries was supported by the strengthening of land rights guarantees and thus, by the renewal of outdated and inappropriate cadasters.

The reconstruction is carried out in terms of data renewal, organization, and technology. The technological reconstruction of the cadaster nowadays refers mainly to computerization and the introduction of computerized data processing. All of the data are converted into the electronic form and stored in data repositories. Conversion into the electronic form is a demanding task in itself, and many countries perform transformation into a new projection coordinate reference system.

The first major systematic reconstruction of documentation, particularly cadastral map sheets, was carried out at the beginning of the 20th century by lithography. Worn-out sheets of the cadastral map (Figure 2a) were processed by a lithographic process in specialized lithographic departments. The content of the cadastral map was transferred to a new paper, while the canceled states registered during maintenance were omitted. All of the changes that were applied in red color, were mapped in the black color as the initial state. These sheets were returned to the cadaster offices for further maintenance (Figure 2b). Some of these sheets, which in the meantime had worn out again and had become illegible, were individually redrawn on new papers.



a. source map b. lithographed c. vectorized

Figure 2. Cadastral map reconstructions (Roić, 2012)

At the end of the 20th century, with the introduction of computers in the processing of cadastral data, intensive reconstruction of cadastral documents began again, this time by converting them into the electronic form. First, the registers, which were in the form of tables in books, were converted and further maintained by computer. Thereafter, the cadastral map sheets were scanned and vectorized and further maintained by computer. Furthermore, information technology methods protect data from physical damage and destruction. Thus, renovations such as lithography and digitization/vectorization are no longer required.

4. VECTORIZATION

The launch of computers into all forms of human activity did not bypass the cadaster. After the conversion of descriptive data, the cadastral map is also converted into the electronic form so that it can be further managed and maintained by computer. Descriptive data are suitable for organization and storage in a relational model. CAD and GIS tools are used to convert the cadastral map into the electronic form, and the data are stored in common data repositories

The project of analog cadastral map sheets' conversion into the electronic form is carried out by cadastral municipalities. This project consists of several interrelated units that may differ depending on the data and the chosen methodology. Conversion into the electronic form is performed exclusively for the entire cadastral municipality in compliance with the existing cadastral regulations. If it is outsourced, the project of conversion into the electronic form usually consists of taking the existing documents, evaluating the quality of documents, scanning, vectorization, controls and corrections, creating a technical report, and handing over the electronic cadastral map to the client.

In the process of conversion the existing cadastral map into the electronic form by vectorization (Figure 3), the analog cadastral map sheets are scanned. Raster files are corrected (stretch, scanning,...) by transformation into known theoretical dimensions. The sheets are georeferenced to the original coordinates, most often by affine transformation, which has been proven to be the most appropriate. This results in a seamless series of sheets of the entire cadastral municipality in the raster format. The coordinates of the boundary points of the land features from the cadastral map are determined from the georeferenced raster.

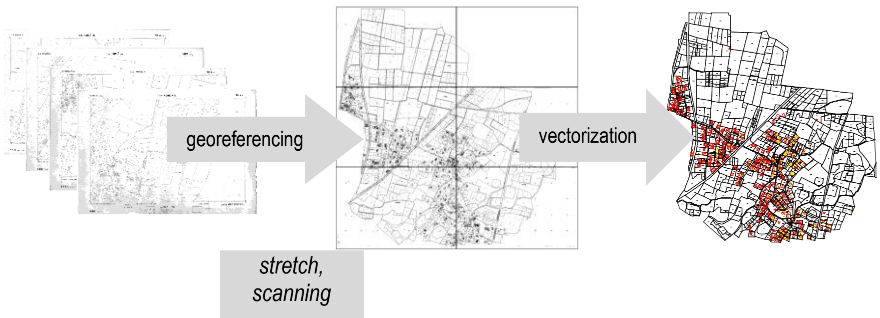


Figure 3. Conversion into electronic form by vectorization (Roić et al., 2002)

In the vectorization process, survey sketches, photo sketches, reports on changes, data on the borders of neighboring cadastral municipalities, identification materials (photographs, orthophotos, coordinates, positional descriptions of permanent geodetic points...), and other data (e.g., utility cadaster) can be used. The assessment of the quality of the existing documents refers to the assessment of the physical condition of map sheets and other documents. Each sheet should be inspected and evaluated for its content, physical condition, coordinate grid, geodetic network points, and actual dimensions of the sheet. The actual dimensions of the sheet are the dimensions of the mapped area of the sheet in the directions of the x (left and right edge) and y (bottom and top edge) axes and are measured precisely (by using a metal ruler).

Scanning is performed using a high-quality calibrated scanner with a positional scanning resolution of at least 300 dpi (dots per inch) unless otherwise requested because of special requirements. The map sheet is scanned in color using 8-bit indexed colors. The content of the map sheet "as it is" is vectorized. Changes that are reported during the execution of the vectorization project are implemented subsequently, after the electronic cadastral map becomes official. Therefore, it is important to carry out the vectorization project as quickly as possible.

Features and properties are not vectorized if they are no longer part of the cadaster, and for historical reasons, they are shown on the cadastral map that is being vectorized. Land features on the cadastral map are modeled with points, lines, text, and symbols. In the case of a point on several sheets of different scales, the coordinate is determined from the largest scale. Areas are described with lines and the associated text or symbols (IDs). The broken lines on the edges of the sheets are united. One line (vector) connects the points from one and the other adjacent sheet.

The lines complete the hierarchical network, and each line is stored once on the priority layer. Only the higher-order line is stored, and the priorities of the lines are boundary line, building line, land use, and other lines. All of the lines start and end at a discrete point or line intersection (node) uniquely determined by the number of the point. Thus, the correct topology is built. A correctly built topology allows us to calculate the areas of cadastral parcels, buildings, and land use from the coordinates of the characteristic points connected by lines.

The result of vectorization is a reconstructed cadastral map, now in the electronic form and seamless for the entire cadastral municipality.

5. REDUNDANT DATA

Electronic cadastral data enable us to carry out automatic controls. Automatic controls give us reliable information about the consistency of the cadastral map and register. To carry out automatic controls, we can use redundant information that was entered in the cadaster or that we produced in the digitization and vectorization process. Respecting the process and methodology of the systematic establishment of the cadastre, any data on the cadastral map must correspond to the data in the lists of the register. Redundant information refers to the number and IDs of polygonal features (cadastral parcels, buildings, and land use).

Primarily, logical consistency should be fulfilled. Each piece of land on the cadastral map must have a parcel ID. All of the parcel IDs on the cadastral map and in the register should be unique. The number of parcels on the cadastral map must be equal to the number of parcels in the register. The building mapped on the cadastral map must have an area registered in the register.

Because of the long-term separate maintenance of the register and the cadastral map, many data differences will be discovered. Queries on the data will enable their detection and help in understanding the cause.

Queries should be conducted for the following:

1. cadastral parcels that are in the list of cadastral parcels but are not mapped on the cadastral map,
2. cadastral parcels that are mapped on the cadastral map and are not in the list of cadastral parcels,
3. buildings that are registered in the list of cadastral parcels but are not mapped on the cadastral map,
4. buildings that are mapped on the cadastral map and are not registered in the list of cadastral parcels,
5. duplicate parcel ID on the map,
6. duplicate parcel ID in the register, and
7. land without parcel ID on the cadastral map.

The results of queries will be the inconsistencies in the cadaster that were caused by the maintenance of the cadastral map and register during the implementation of changes and/or the errors made by the cadaster establishment or reconstruction.

In addition to logical constraints, the compliance of areas should be investigated. For each parcel, when the cadaster is created or during maintenance, an area is determined and entered as official in the register (*Ar*). From the vectorization, we obtained the coordinates of all of the border points of the parcels and buildings and could calculate their areas (*Am*) with the GIS tools. They differed for various historical reasons, mainly because of graphic mapping and the determination of areas on analog maps. We considered these differences acceptable, but for them, we had to determine the criteria for determining the maximum acceptable difference. For graphically determined surfaces, the following general formula was applied:

$$\text{Tolerance} = 0.7 * \frac{M}{1000} * \sqrt{Ar}, \text{ where M is map scale} \quad (1)$$

Area differences greater than the acceptable ones indicated the presence of other errors that had to be detected and corrected. The causes could be the following: georeferencing of raster sheets, determination of coordinates during vectorization, or typing of areas from books into the register. To efficiently approach the detection of errors, it was best to present the results in a tabular form (Table 1). The data on parcel ID, area from the register, and area calculated from the vectorized cadastral map were entered in the first column. Then, in the next column, the difference between the area from the register and the calculated area from the cadastral map was calculated. Tolerances in the following column were calculated using formula 1, and then, the tolerance − absolute value of the difference was calculated. The presence of only negative values in this column indicated differences that were greater than the tolerance. For them, the cause was investigated, and an error in the data had to be found.

Most often, one error caused two or more unacceptable differences in areas. Therefore, it was better to strategically approach the detection of errors through relatively large differences. Column 7 presents the calculated percentage difference in relation to the total area according to the following:

$$R = \frac{(6)}{Ar} * 100 \quad (2)$$

Upon sorting the table by column 7, we observed that the parcels with the biggest differences appeared at the top.

Table 1. Area differences of cadastral parcels

Parcel ID	*Ar*	*Am*	Difference	Tolerance	Tolerance – abs(Difference)	R [%]	Note
1	2	3	4	5	6	7	8
*124	7	168	−161	5	−156	2229	
…							
*123	378	256	120	39	−81	21	
…							

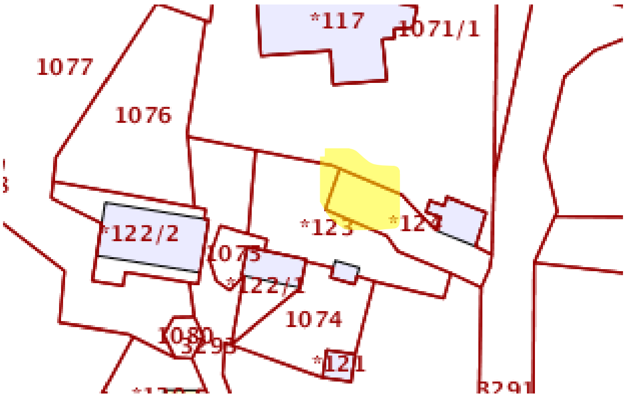
A similar methodology and table was applied to analyze the consistency of building areas and land use areas.

6. DISCUSSION

The implementation of the proposed comparisons uncovered the inconsistency of data in the cadastral system. For the biggest differences, it was usually the easiest to find the cause and correct the error. The huge unacceptable area difference for parcel *124, for example, was related to the difference in the neighboring parcel *123 (Table 1). The comparison of the registered parcel areas to the areas calculated from the electronic cadastral map revealed a huge unacceptable difference for parcel *123.

Upon comparing the electronic map (Figure 4a) and the map from which it was vectorized (Figure 4b), we concluded that the vectorization was done correctly. However, when we compared the archived cadastral map (Figure 4c), we found that on it, the connection sign (S), of the land marked in yellow in Figure 4a, was in a different location.

This implied that before the reconstruction of the cadastral map by lithography, this particular part of the land on the map belonged to parcel *124 and that with its area, the total area of the parcel *124 corresponded well with that in the register. This was also indicated by the fact that the area of the parcel *123 in the register was larger than that on the map. Upon correction of this error made during the reconstruction, both of the parcels had consistent areas within acceptable tolerances.



a) electronic map


b) vectorized map - paper


c. archived map

Figure 4. Example of cadastral system inconsistency

On the basis of the application of comparisons to authentic cadastral data, it was found that there were many such and similar errors. Most of them could be resolved within a short time of up to 10 min per case (Roić et al., 2021). Therefore, it is preferable that the work be performed by an expert who has experience in cadaster maintenance. This approach to improving the quality of data registered in the cadaster deals only with technical issues. The social relations of people to land (taxpayers, owners, ...) in relation to registered parcels are also an important factor of the cadaster. However, to the topic of this study, they were irrelevant and therefore were not considered.

Analyses can only be carried out in cadastral systems that were created by a systematic approach and that have seamless and complete documentation for a spatial unit (e.g., cadastral municipality). Depending on the cadastral data model in a particular country/administrative district, it may be necessary to slightly adjust certain comparisons or add new ones. Identical comparisons can be made for the technical data in the land book.

7. CONCLUSION

User trust in the cadastral system is very important. However, establishing a cadaster is an expensive and time-consuming task. Therefore, in many systems today, electronic data are obtained by using secondary methods, digitization, and vectorization of analog data.

The paper discussed how the cadastral system could be improved by using redundant data. The inconsistencies revealed by comparisons clearly pointed to the errors made in the past. This enabled a quick correction of errors, which enhanced the quality of data in the cadastral system and strengthened the user confidence in the system.

Today's electronic cadastral systems do not yet have built-in mechanisms for coordinated entry in all parts of the database. Consistent input is not ensured by appropriate built-in constraints. Therefore, it is recommended that the proposed comparisons also be carried out occasionally in an electronic cadastral environment.

REFERENCES

Bennett, R., Stöcker, C., & Asiama, K. (2022). *Digital transformation and land administration – Sustainable practices from the UNECE region and beyond; FIG PUBLICATION NO. 80*. FAO; UNECE (United Nations Economic Commission for Europe); https://doi.org/10.4060/CC1908EN

Fetai, B., Tekavec, J., Fras, M. K. K., Lisec, A., Fetai, B., Tekavec, J., Fras, M. K. K., & Lisec, A. (2022). Inconsistencies in cadastral boundary data: Digitisation and maintenance. *Land 2022, Vol. 11, Page: 2318*, *11*(12), 2318. https://doi.org/10.3390/LAND11122318

Henssen, J. (1995). Basic principles of the main cadastral systems in the world. *Modern Cadasters and Cadastral Innovations*, *May 1995*, 5–12. http://www.geo21.ch/cadastrallibrary/international/Henssen1995-BasicPrinciplesCadastralSystems.pdf

Karabin, M., & Łuczyński, R. (2022). The issues regarding the compliance of data contained in the cadastre and land register in Poland. *Volunteering for the Future - Geospatial Excellence for a Better Living*. https://www.fig.net/resources/proceedings/fig_proceedings/fig2022/papers/ts04b/TS04B_karabin_Luczynski_11532_abs.pdf

Roić, M. (2012). *Upravljanje zemljišnim informacijama - katastar (Land Information Administration: Cadastre)*. Geodetski fakultet. http://bib.irb.hr/prikazi-rad?&rad=593216

Roić, M., Cetl, V., Matijević, H., Kapović, Z., Mastelić-Ivić, S., & Ivšić, I. (2002). Prevođenje katastarskih planova izrađenih u Gauss-Kruegerovoj projekciji u digitalni vektorski oblik / postupci i procedure - tehničko izvješće o radovima na projektu. In *Republika Hrvatska - Državna geodetska uprava*. https://www.bib.irb.hr/129301

Roić, M., Križanović, J., & Pivac, D. (2021). An approach to resolve inconsistencies of data in the cadastre. *Land*, *10*(1), 1–20. https://doi.org/10.3390/land10010070

Taszakowski, J., Litwin, U., Doroz, A., & Misiak, K. (2018). Evaluation of the quality of cadastral data. *E3S Web of Conferences*, *63*, 00018. https://doi.org/10.1051/E3SCONF/20186300018

BIOGRAPHICAL NOTES

Miodrag Roić graduated with a degree in Geodesy from the University of Zagreb, Faculty of Geodesy. In 1994, he received a PhD from the Technical University Vienna. Since 1996, he has been a professor at the University of Zagreb, Faculty of Geodesy. He was Dean of the Faculty during the period spanning 2011–2015. The topics in which he specializes are Cadaster, Land Administration Systems, Engineering Geodesy, and Geoinformatics. He is a corresponding member of the German Geodetic Commission (DGK) and many other national and international scientific and professional institutions.

CONTACTS

Full Prof. PhD. Miodrag Roić
University of Zagreb, Faculty of Geodesy
Kačićeva 26
HR-10000 Zagreb
CROATIA
Tel. +385 (1) 4639 229
Email: miodrag.roic@geof.unizg.hr
Website: http://www2.geof.unizg.hr/~mroic/