Data Quality Control in the Process Massive Land Survey According to the LADM-COL

Andrés GUARIN, Sergio RAMIREZ, Jonny SANCHEZ, Colombia
Daniel CASALPRIM, España

Key words: Data Quality Control, LADM, Interlis, QGis.

SUMMARY

In the new model of Cadastral Management in Colombia the decentralization of the cadaster is sought at several levels. The first one is to enable territorial entities to provide services as a Cadastral Manager. The manager will be responsible for attending all the procedures related to training, conservation or update of the cadastral information of the assigned municipalities. The second level is to enable cadastral operators, which are those entities that will be in charge of carrying out the information gathering in the field in a massive way and deliver it to the manager for validation, acceptance and formalization of the data.

The National Land Agency ANT, is the highest authority of the nation's land and is the entity responsible for carrying out activities for the allocation of rights and the regularization of different land tenure situations in the rural sector. In 2017, through Decree Law 902, the role of cadastral manager was assigned.

In this context, the ANT and USAID have developed a pilot project in the municipality of Ovejas - Sucre, where the intervention was carried out using the massive field data surveying methodology, in accordance with the technical product specifications defined by the cadastral authority.

In the process of consolidating the information, USAID delivered the structured information according to the LADM_COL model (version 2.2.1) in the INTERLIS * exchange format: xtf, the project assumed the challenge of being the first pilot to consolidate information in this format under the definitions made by the IGAC and the SNR. The delivery of the information in interlis format (*.xtf) and according to the LADM_COL (*.ili) model by the operator also implies a challenge for the institutions in the articulation of the information and the optimization of the processes of revision, verification and validation of the information delivered.

Data Quality Control in the Process Massive Land Survey According to the LADM-COL

Andrés GUARIN, Sergio RAMIREZ, Jonny SANCHEZ, Colombia
Daniel CASALPRIM, España

1. INTRODUCTION

A Land Administration System is a set of norms, standards, tools, processes and actors. According to this, to adopt a data model for land administration in Colombia will require institutional efforts to improve the value chain of the data managed by each entity. Additionally it should ensure the interoperability of the data with other institutions that are part of this model and consider the citizens as the end users of the information, since the vision of an efficient land administration system puts the citizen in a central role, and therefore, tries to improve the processes that impact on the citizen.

Older models of land management and land surveying are inadequate to meet the needs of a modern and sustainable society (Williamson, 2009). The paradigm of land administration focuses on providing tools to decision makers with the objective of the sustainable development (Williamson, 2009). This is one of the pillars of the transformation processes and vision of territorial approach that seeks, among other objectives, the integrated rural development.

The implementation of the new public policy of multipurpose cadaster, implies major challenges for entities that are part of the management of territorial objects. Among these challenges, we can find the redefinition and adjustment of the technological components such as information systems, that enable the information management, trough the adaption of the software components, hardware and communication networks to make them more efficient and to allow the exchange of information and the data publication based on the new strategic framework defined in the Land Administration Model.

The IGAC with the support of the Swiss Cooperation Project "Modernizing Land Administration in Colombia", financed through the State Secretary for Economic Affairs (SECO), has been working on the definition of a quality control scheme over the information, under the principles of efficiency and effectiveness. For this it is necessary to have technological tools that optimize the automated validation of data conforming to the LADM-COL model.

One of the main advantages of having LADM-COL modeled in format (*.ili) and that the data delivered by USAID is in format (*.xtf) is that it allowed to execute the validation of data in bulk at the semantic level and to apply business rules to be able to perform the quality control of the data delivered by the operator to ensure the compliance with the technical product specifications defined by the multipurpose cadaster.

The review of large data sets of information against the LADM model using INTERLIS and tools that allow the automation of the process is a great advance in the implementation of the multipurpose cadaster policy context.

Therefore, and in accordance with the agreement between the ANT and the IGAC, which aims to "combine institutional, technical, technological, human, logistic and financial efforts between the IGAC and the ANT, with the purpose of advancing the process of validation, approval and registration of the cadastral information, resulting from the land field surveying, in the cadastral database, within the framework of the implementation of the route of the OSP[1] plans in the municipalities or areas targeted by the ANT ".

2. METHODOLOGY FOR QUALITY CONTROL

According to the general purpose of the Multipurpose Cadaster V.2.1.1 which its main objective is the cadastral update, under the national land field survey and the contrast of the physical reality with the legal one, it is necessary and fundamental to define which These are the current legal situations at the national level, especially the informality that is currently in the municipality of Ovejas, which will be the pilot intervention territory.

The process of validation of compliance with the technical specifications of the product from the technological component, although it focuses specifically on certain annexes, is transversal in the process of quality control for the validation of the information corresponding to the other components, so that the data set validation process has several validation activities, as shown in Figure 1 below three main processes are performed. A first process of consistency in loading the information in relation to the delivery format and the location in the requested folder structure

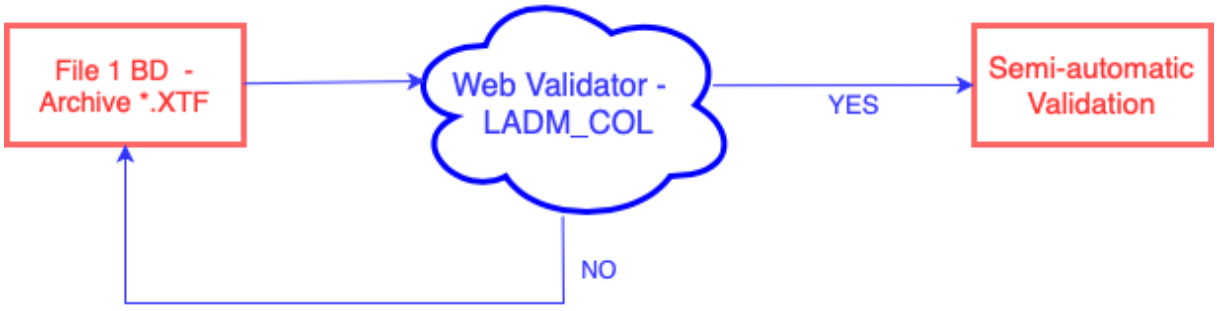


Figure 1. Validation Process

A second process corresponds to the application of a web validator on the data set, where the consistency of the information in relation to the semantics defined in the LADM_COL model is mainly validated and some logical validators are applied in relation to the codependency of variables and consistency of the information and finally a third validator that is carried out from

[1] Social Ordering of Property

the technological component is semi-automated because it is an assisted validator in which the validation of the topological rules on the data set of the intervention area is carried out and in the same way perform topological validations between deliveries.

As part of the validation process of the technological component in the information delivery scheme, a folder was created in the delivery structure called BD, which seeks that the information that is delivered by the operator is housed there. At this point, it is the operator who is responsible for ensuring the quality of the information and it is he who must guarantee the loading of the database in accordance with the LAMD_COL model and in the exchange format defined for the XTF multipurpose cadaster pilots.

Figure 2 review shows in detail the activities proposed to be carried out as part of the quality control process and that will be supported by the technological component, complying with the process of reviewing the information that is delivered by an operator and the return flow when the data does not comply or acceptance when the information delivered complies with the defined quality control.

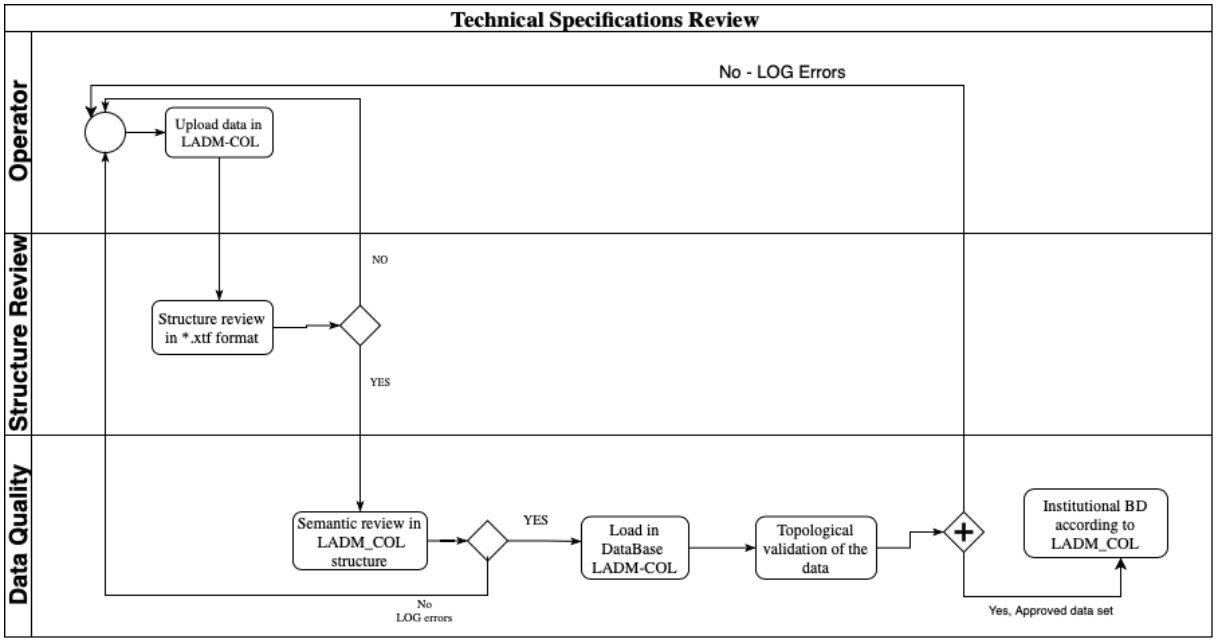


Figure 2. Workflow for quality control

2.1 Upload delivery to FTP

The activity of loading the information by the operator is carried out in accordance with the model structure LADM-COL in its version 2.2.1 and in *.xtf format, once the confirmation email of the load is received, the IGAC starts the download of the information that is stored in

the FTP and that will be the set of folders and files on which the quality control process will be carried out by the institutions.

2.2 Review structure and format (XTF) of delivery.

This activity includes two fundamental tasks in the framework of product quality control. A first review is that the information delivered by the operator and downloaded from the FTP are consistent and is loaded in the folder designated for loading the information from the database and the second control is that the requested delivery format is met. which as defined corresponds to the XTF format; In the methodology initially proposed if either of these two controls fails, the delivery is returned to the operator, to correct the findings found.

2.3 Systematized review Validator (ilivalidator).

As part of the process of the new vision of the multipurpose cadaster and the massive information gathering scheme, efforts are made to generate mechanisms that allow automation to be carried out in the validation of large volumes of information delivered by each of the operators, which is why that the IGAC has been working with the support of the Swiss cooperation project in the generation of a series of tools that will allow the institutions to carry out a first review process on the information delivered and once it complies with the technical specifications of the product, the thematic validation of the information.

At this point in the process chain of the review on compliance with the technical specifications, it relies on the review of the logical consistency of the data and the semantic validation of the information in accordance with the LADM_COL model.

3. RESULTS IN THE DATA QUALITY CONTROL IN OVEJAS – SUCRE MUNICIPALITY

Delivery by the information operator according to the LADM-COL model

For the delivery of the information, initially, a migration of the data from the operator to the model had to be made, because the same used its own model for the capture of the information. For this activity we used the plugin for QGIS version 3, LADM-COL Assistant, this tool allows you to migrate data massively through intermediate structures (alphanumeric information), and through FileGBD (geographic information).

In the case of the intermediate structure, an excel file was used with all the information consolidated by the operator that contained the property, rights, administrative and interested data, this structure was previously defined to automate the process and allowed to migrate the data Alphanumeric automatically, on the other hand the information was also migrated from postgreSQL of the information collected in property records and appraisals. In the case of the information of the FileGDB, the tool used to manage ETL (Extract, Transform and Load) was

used to pass the information of points of boundaries, boundaries, land, buildings and construction units.

Once these two sets of data were migrated, integration was made through the property codes generated by the operator.

The results of the migration were as follows:

Class	Quantity
Parcel	6343
Property chips	6343
Land	6190
Constructions	9762
Constructions Unit	9762
Boundaries	20850
Boundary points	52686
Lifting points	20868
Rights	6343
Interested	5116
Groups of interested parties	387
Construction appraisals	8086

Figure 3 shows the migrated geographic information:

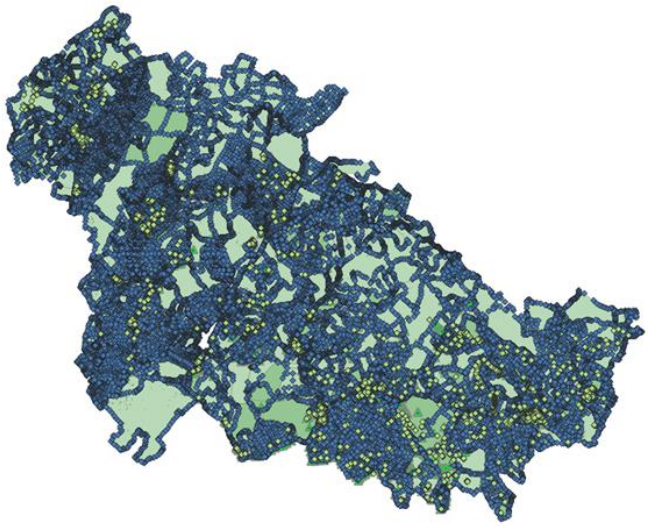


Figure 3. Information Municipality of Ovejas in the model LADM-COL

Looking closer, the following is observed (see figure 4):

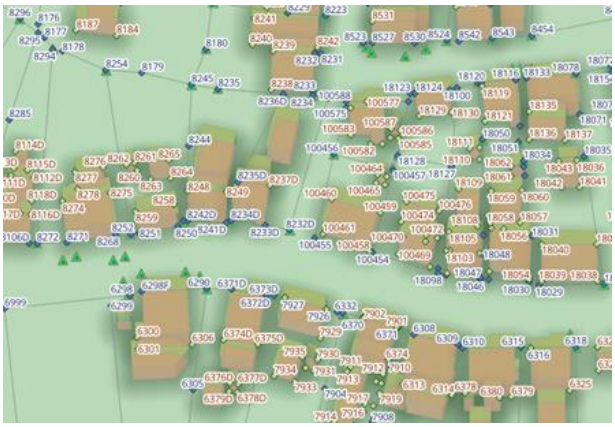


Figure 4. Information Municipality of Ovejas in the model LADM-COL

This was the information that the IGAC received to subsequently perform quality control, for this purpose a functionality of the LADM-COL Assistant was used, which allows validating rules defined by the institution to find both topological and alphanumeric information inconsistencies, in Figure 5 the different rules generated for this purpose are observed:

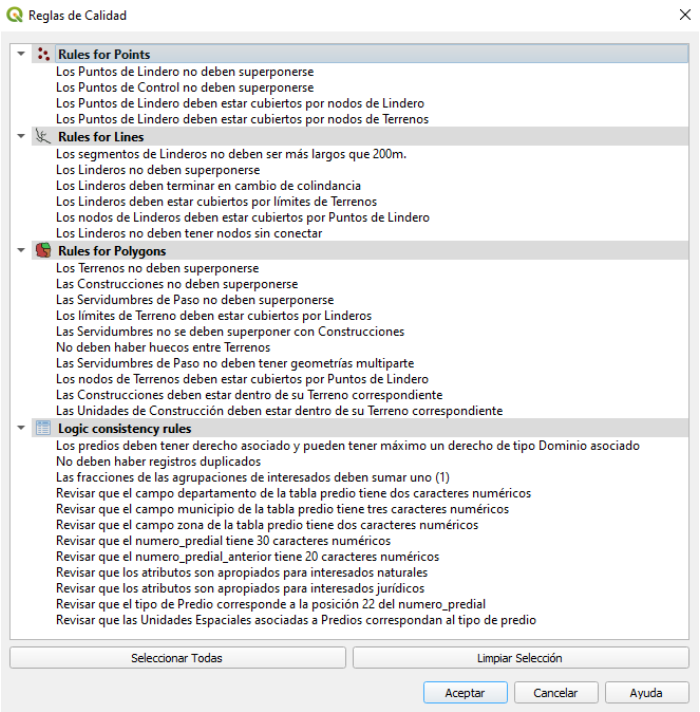
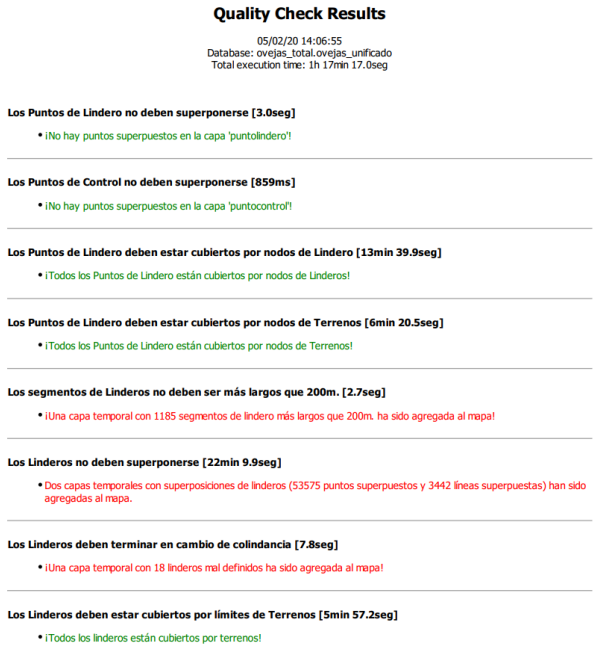


Figure 5. Topological validation and logical consistency rules

In total, 32 quality rules were validated, which subsequently yielded results in the geographical issue through some topological inconsistencies, as well as logic in the alphanumeric information, in figure 6 a result of validation execution is observed:



Quality Check Results

05/02/20 14:06:55
Database: ovejas_total.ovejas_unificado
Total execution time: 1h 17min 17.0seg

Los Puntos de Lindero no deben superponerse [3.0seg]

- ¡No hay puntos superpuestos en la capa 'puntolindero'!

Los Puntos de Control no deben superponerse [859ms]

- ¡No hay puntos superpuestos en la capa 'puntocontrol'!

Los Puntos de Lindero deben estar cubiertos por nodos de Lindero [13min 39.9seg]

- ¡Todos los Puntos de Lindero están cubiertos por nodos de Linderos!

Los Puntos de Lindero deben estar cubiertos por nodos de Terrenos [6min 20.5seg]

- ¡Todos los Puntos de Lindero están cubiertos por nodos de Terrenos!

Los segmentos de Linderos no deben ser más largos que 200m. [2.7seg]

- ¡Una capa temporal con 1185 segmentos de lindero más largos que 200m. ha sido agregada al mapa!

Los Linderos no deben superponerse [22min 9.9seg]

- Dos capas temporales con superposiciones de linderos (53575 puntos superpuestos y 3442 lineas superpuestas) han sido agregadas al mapa.

Los Linderos deben terminar en cambio de colindancia [7.8seg]

- ¡Una capa temporal con 18 linderos mal definidos ha sido agregada al mapa!

Los Linderos deben estar cubiertos por límites de Terrenos [5min 57.2seg]

- ¡Todos los linderos están cubiertos por terrenos!

Figure 6. Quality Rule Validation Report

This report generates the result of the defined rule as well as the execution time in the review of each one, this allows future scaling to determine or prioritize the verifications on the information delivered. In the case of geographical information, something like the following can be observed (see figure 7):

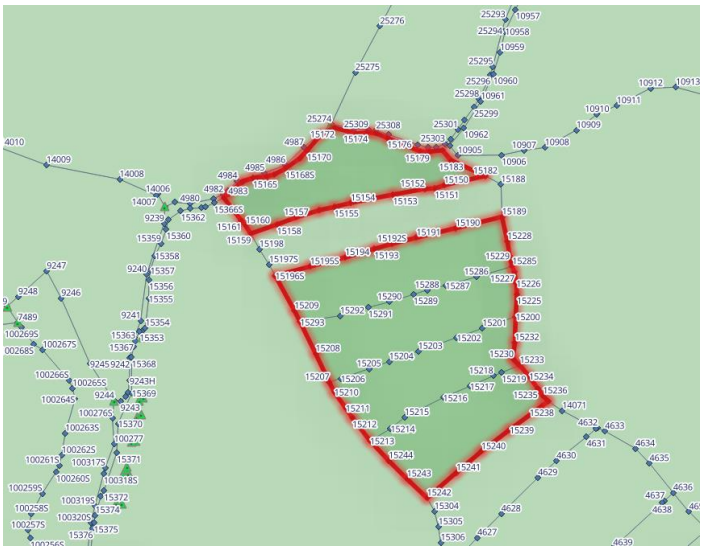


Figure 7. Result Border validation

In this case, there is an overlapping of boundaries, which are given because there is a matrix (previous) property a series of properties by which it was divided, this allows to partially identify the informality present in the area, likewise, if the land polygons are reviewed, this detail can be seen more clearly (see figure 8):

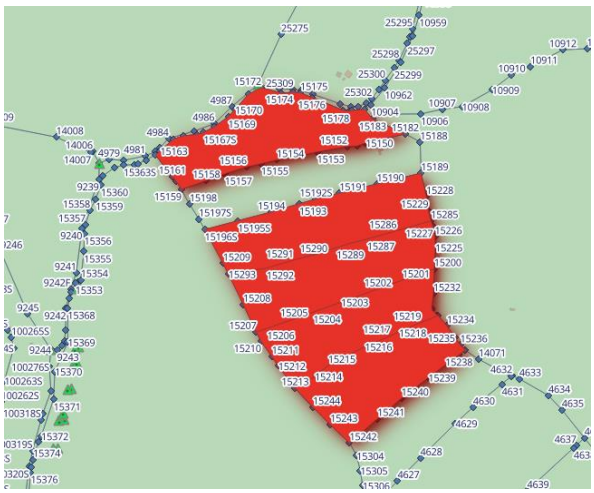


Figure 8. Result Land Validation

These details allow to identify not only possible topological errors but also administrative actions to which the IGAC must enter to update its cadastral information, in this case you can identify informal properties and how they should be created within the cadastral base in force within the national cadastral system . Details can also be found as constructions that belong to two properties at the same time, in many cases these correspond to shared bathrooms between the properties (see Figure 9):

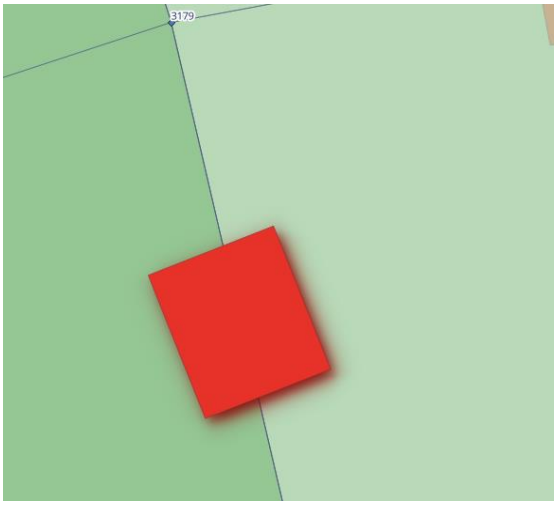


Figure 9. Result Construction and land validation

CONCLUSIONS

The tools developed for the data management and data validation against LADM, facilitated the revision of the information of the multi-purpose cadastral intervention in the municipality of Ovejas – Sucre. The time used in the review and validation of cadastral data were reduced.

In Colombia, cadastral information has been traditionally reviewed in formats such as Geodatabase, Shapefile and tables that are easy to manipulate, but with the implementation of the LADM and Interlis, that standaritation and validation was solved, given that the data is backed up in a database that maintains the semantic of the LADM.

Users in charge of the data review can access all cadastral data by accessing a single database, which contains the relationship of ownership between a plot (administrative unite) and a person (party).
The developed methodology for data review is useful for the implementation of the multi-purpose cadaster in Colombia because from the beginning the operator must structure their data according to the LADM and during the quality process, the information is reviewed in this model, and once it is approved, it is incorporated into the LADM institutional database.

During the validation process, the topological errors were almost entirely corrected, which allowed the identification ofspecific cases that were not necessarily errors themselves, but specific cases that despite being indicated as an error, should be corrected in a different way by the IGAC later, what reduced processes and time

REFERENCES

Departamento Nacional de Planeación. (2016). *Política para la Adopción e Implementación de un Catastro Multipropósito Rural-Urbano (CONPES 3859).* Bogotá, Colombia: DNP.

Departamento Nacional de Planeación. (2019). *Estrategia para la implementación de la política pública de Catastro Multipropósito (CONPES 3958).* Bogotá, Colombia: DNP.

FAO. (2005). *Access to rural land and land administration after violent conflicts* . Roma: FAO LAND TENURE STUDIES.

Germann, M., Steudler, D., Kaufmann, J., Lemmen, C., Oosteromm, P., & Zeeuw, K. (2015). "The LADM Based on INTERLIS". FIG Working Week 2015, Sofia, Bulgaria.

IGAC, SNR. (2016). *Conceptualización y especificaciones para la operación del Catastro Multiproposito.* Bogotá, Colombia.

ISO/TC 211 Geographic information. (2012). *Land Administration Domain Model (LADM). ISO/CD 19152.* Lysaker, Noruega: International Organization for Standarization.

Jenni, L., Guarín, A., Ziegler, S., & Bajo, V. (2017*). Development and Deployment of a LADM Implementing Toolkit in Colombia*. Washington, Annual Land and Poverty conference. World Bank.

Kaufmann, J., & Steudler, D. (1998). *Cadastre 2014 - A vision for a future cadastral system.* International Federation of Surveyors FIG, Commission 7.

Lemmen, C. (2012). *A Domain Model for Land Administration.* Delft: Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission 78 .

Lemmen, C., Oosterom, P., & Bennett, R. (2015). The Land Administration Domain Model. *Land Use Policy , 49*, 535–545.

Williamson, I., Enemark, S., Wallace, J., & Rajabifard, A. (2009). *Land Administration and Sustainable Development.* ESRI Press Academic.

BIOGRAPHICAL NOTES

Andrés P. Guarín López is a cadastral engineer and geodesist, with a master in information sciences and communications from the Universidad Distrital Francisco Jose de Caldas. During his professional occupation, he has been the technical leader for implementing several GIS projects, from analysis, design, and development to the implementation stage. He was engaged in drafting the conceptual design of the new Multipurpose Cadastre of Colombia and is currently working a consultant for the institution's Land Information Management.

Jonny C. Sánchez González is a cadastral engineer and geodesist, with a specialization in appraisals from the Universidad Distrital Francisco Jose de Caldas. During his professional occupation, he has worked on a project on land regularization. He currently works at the Agustin Codazzi Geographical Institute in the implementation of the Land Administration Domain Model and the new cadastre model in Colombia.

Sergio A. Ramírez Navarro is a technologist in surveying, surveying engineer and master's candidate in information and communications sciences - emphasis on geomatics from the Universidad Distrital Francisco José de Caldas, specialist in project management Universidad Piloto de Colombia, more than 11 years of experience in projects of Surveying, Geodesy, Cadastre, Cartography, LiDAR, among others. He is part of the technical team of the Implementation Agency BSF Swissphoto - Incige in the project "Modernization of Land Administration in Colombia" participating in the conceptualization of the LADM-COL model through the use of INTERLIS, also in the development of the plugin Assistant LADM-COL for QGIS, among others.

Daniel Casalprim. Specialist in Project Management of Map Production, Land Tenure and Property Rights projects, he has leaded Technical Teams and the implementation of innovative production structures in these areas during more than 15 years in Latin America, Africa and Europe, with different funding agencies. He has managed map production projects for

governmental agencies in South America, leaded a fiscal cadaster project in Cameroon and a land delimitation project in Ecuador (Programa Sigtierras). He has also been involved in the LIFT project in Ethiopia (DAI) through the evaluation of the Rural Land Administration System, funded by DFID. Currently he is leading the technical team involved in the Proyecto de Modernizacion de la Administración de Tierras in Colombia, funded by SECO (Swiss Cooperation).

CONTACTS

Andrés P. Guarín L.
Implementation Agency Bsf-Swissphoto - INCIGE
Av el Dorado No. 69 - 63, Edificio Torre 26, Oficina 406
Bogota D.C.
COLOMBIA
Phone: + 57 301 489 42 75
E-mail: andresguarinlo@gmail.com
Website: https://www.proadmintierra.info/

Jonny C. Sánchez G.
Agustin Codazzi Geographical Institute – IGAC
Carrera 30 48 – 51,
Bogotá D.C
COLOMBIA
Phone: + 57 320 467 66 78
E-mail: jonny.carlos.sanchez@gmail.com
Website: https://www.igac.gov.co/

Sergio A. Ramírez N.
Implementation Agency Bsf-Swissphoto - INCIGE
Av el Dorado No. 69 - 63, Edificio Torre 26, Oficina 406
Bogota D.C.
COLOMBIA
Phone: + 57 300 325 97 21
E-mail: seralra96@gmail.com
Website: https://www.proadmintierra.info/

Daniel Casalprim
Implementation Agency – Bfs-Swissphoto
Av el Dorado No. 69 - 63, Edificio Torre 26, Oficina 406,
Bogotá DC
COLOMBIA
Tel. +57 7040687
Email: daniel.casalprim@bsf-swissphoto.com
Website: https://www.proadmintierra.info