Spatialization of Territorial Parcels for Efficient Land Management

Marco Aurélio BARBIERO and Carlos Antonio Oliveira VIEIRA, Brazil

Key words: Land Management, Multi-Purpose Territorial Cadastre, Spatialization of Territorial Parcels, LADM, CNEFE, CIATA.

SUMMARY

Georeferencing urban land parcels is essential for municipal management, supporting urban planning, land regularization, tax collection, integrated public services, and legal security in real estate transactions. It also promotes transparency and citizen participation while aligning administration with geographic information systems and federal regulations, including the *City Statute* and *National System for Territorial Information Management* (SINTER). Despite these benefits, georeferencing urban land parcel in Brazilian municipalities is limited. In 2019, only 21% of municipalities had a georeferenced cadastral database, while 59% collected only descriptive data, often without spatialization, and some lacked automation. This highlights the need for methods enabling large-scale spatialization of urban parcels and updated cadastral mapping. This work proposes a *Fit-For-Purpose Land Administration* methodology using *National Address Database for Statistical Purposes* (CNEFE) and municipal cadastre data combined with *OpenStreetMap*. Results show that georeferencing can achieve a maximum error limited to half the block dimensions when CNEFE points are adequately distributed, while discrepancies reveal potential cadastre integrity issues. Collected images and operation logs may support future *deep learning* applications, enhancing accuracy and automation.

RESUMO

O georreferenciamento de parcelas urbanas é essencial para a gestão municipal, apoiando o planejamento urbano, a regularização fundiária, a arrecadação tributária, a gestão integrada de serviços públicos e a segurança jurídica em transações imobiliárias. Também promove transparência e participação cidadã, alinhando a administração a sistemas de informação geográfica e às normas federais, incluindo o Estatuto das Cidades e o SINTER. Apesar desses benefícios, a cobertura de georreferenciamento nos municípios brasileiros é limitada. Em 2019, apenas 21% dos municípios possuíam uma base cadastral georreferenciada, enquanto 59% coletavam apenas dados descritivos, frequentemente sem espacialização, e alguns não utilizavam automação. Isso evidencia a necessidade de métodos que permitam a espacialização em larga escala das parcelas urbanas e a atualização do cadastro municipal. Este trabalho propõe uma metodologia de Administração Territorial Fit-For-Purpose utilizando dados do CNEFE e do cadastro municipal, combinados com *OpenStreetMap*. Resultados mostram que o georreferenciamento das parcelas pode alcançar erro máximo limitado à metade das dimensões da quadra, e que discrepâncias indicam possíveis problemas de integridade cadastral. Imagens e registros coletados podem apoiar futuras aplicações de *deep learning*, aumentando precisão e automação.

Spatialization of Territorial Parcels for Efficient Land Management

Marco Aurélio BARBIERO and Carlos Antonio Oliveira VIEIRA, Brazil

1. INTRODUCTION

There are many reasons for a municipality to carry out the georeferencing of urban land parcels. This practice directly contributes to urban planning by enabling the precise identification of land use and occupation, and it can improve the efficiency of tax collection. It also supports land regularization, provides a basis for the integrated management of public services, and strengthens legal security in real estate transactions. From an institutional perspective, georeferencing enables modern public administration, aligned with geographic information systems, and ensures compliance with federal regulations, such as the *City Statute* and the *National System for Territorial Information Management* (SINTER). Furthermore, it is a fundamental tool for reducing urban risks, formulating territorial social policies, and promoting transparency and social participation.

Despite the evident contributions offered by the presence of georeferenced data in urban cadastres, in 2019 only one-fifth of Brazilian municipalities (21% or 1,159 municipalities) had a georeferenced cadastral database, according to the Survey of Basic Municipal Information – 2019 (IBGE, 2019). The same survey showed that 59% of municipalities were limited to collecting descriptive information required for legal compliance and tax collection, without any concern for spatializing the parcel grid. There were also municipalities that did not use any type of automation or georeferencing.

Given such low rates of georeferencing, regardless of the reasons for this gap in municipal cadastres, it is urgent to seek alternatives that enable the large-scale spatialization of the urban land grid, as well as the production and updating of municipal cadastral mapping.

One of the most effective strategies to accelerate this spatialization process consists of leveraging already georeferenced databases and seeking methods of integration with municipal urban cadastres, even if the accuracy initially obtained is not ideal.

In Brazil, for instance, the *National Address Database for Statistical Purposes* (CNEFE), an initiative by IBGE (*Brazilian Institute of Geography and Statistics*) aimed at collecting and organizing address information for census purposes (IBGE, 2025), stands out as a relevant alternative to be used together with local cadastral mapping and other geotechnologies in the urban management of Brazilian municipalities.

This work proposes a *Fit-For-Purpose Land Administration* methodology (Enemark *et al*., 2021) for the large-scale spatialization of urban parcel grids, based on two public and free databases: CNEFE/IBGE and the municipal urban land cadastre.

2. THEORETICAL FRAMEWORK

The methodology presented here was named *SuperCIATA*, in reference to the CIATA Project (*Convênio de Incentivo ao Aperfeiçoamento Técnico-Administrativo de Municípios*), developed in Brazil in the 1970s (Cunha et al., 2019).

In addition to naming the methodology, *SuperCIATA* also designates the web application developed to demonstrate the different stages of transformation and integration of the original data until its consolidation into georeferenced cadastral information.

2.1 Urban Land Cadastre and the CIATA

The CIATA Project (*Convênio de Incentivo ao Aperfeiçoamento Técnico*) (SERPRO, 2025) was implemented in the 1970s by the *Secretariat of Economy and Finance* of the *Ministry of Finance*, with resources from the *Technical Assistance Program* (PRAT) and support from the *Federal Data Processing Service* (SERPRO). The initial objective was to assist small municipalities in the implementation of the *Municipal Territorial Cadastre* and thus increase the collection of their own revenues and reduce dependence on external resources (Silva, 2023).

Although its main focus was on improving municipal revenue collection, CIATA contributed to the implementation of cadastres in several Brazilian municipalities, serving as a basis for the development of broader territorial information systems.

2.1.1 CIATA Conceptual Model

CIATA was designed in a modular and flexible way to meet the specific needs of each municipality, covering modules for legal assistance, administrative assistance, urban real estate cadastre, fiscal cadastre, and revenue (Carneiro, 2003).

Considering the technical limitations of the 1970s and 1980s, especially the absence of commercial *Database Management Systems* (DBMSs), CIATA's data modeling followed the single complete record model prevalent at the time, in which all information about an object was stored in a single row of the file (Barbiero, 2025).

The single-record model of CIATA is now considered obsolete, having been replaced by the relational model (Date, 2004), but some important features for application were preserved in most modern cadastre systems: the "*property registration*" – a hierarchical primary key usually composed of *district*, *sector*, *block*, and *unit*; and the frontage and side dimensions of the parcels. With this data, the authors believe that it is possible to graphically represent the blocks and parcels that compose them proportionally.

2.1.2 CIATA and the LADM

LADM, short for *Land Administration Domain Model*, is a standardized international data model for representing land administration information. It was developed by the *International Federation of Surveyors* (FIG) and the *International Organization for Standardization* (ISO). LADM defines a set of objects (classes) and relationships that can be used to describe different aspects of land administration, such as: *property rights*, *land-use restrictions*, and *spatial information*. Adopting LADM can facilitate data integration between different systems and promote interoperability across countries (Enemark et al., 2021).

A deeper analysis of LADM easily reveals strong influences from the relational model proposed by Edgar Codd, in which each class or object is modeled in a separate data structure (Date, 2004), as verified in the ISO 19152 standard (ISO 19152, 2012). On the other hand, CIATA was predominantly implemented using a single file as its basis. Its data definitions – or at least some clues about them – can be found in the *Real Estate Cadastre Manual* (SERPRO, 2025).

Thus, a direct and detailed comparison between the two technologies is only possible at the conceptual level, as shown in Table 1.

Table 1. Conceptual comparison between CIATA and LADM

Characteristic	CIATA	LADM
Origin	Brazilian government initiative in the 1970s and 1980s.	International conceptual model developed by FIG and ISO.
Scope	Urban land cadastre in Brazil, with an initial focus on small municipalities.	Land administration in general, applicable to different countries and contexts.
Objective	Increase municipal revenue through an efficient technical cadastre system.	Provide a standard model for land administration, promoting interoperability.
Methodology	Flexible and adaptable, allowing each municipality to define the information to be collected.	Based on a conceptual model that defines standardized objects and relationships.
Technology	Limited to the technology available at the time, focused on data processing.	Technology-independent, can be implemented with different tools.
Georeferencing	Incipient and limited use of georeferencing. Few cases of effective spatial integration.	Georeferencing is a core part of the model, with strong emphasis on parcel representation.
Implementation	Discontinued in 1981 due to loss of funding.	Under continuous development and improvement, with international adoption.

The most visible contribution of LADM is the set of UML (*Unified Modeling Language*) diagrams that present the most relevant classes and attributes of a cadastral system. Classes are structures that define the data model of an object, while attributes are variables associated with classes, responsible for describing the specific characteristics of each object, according to LADM (ISO 19152, 2012).

In general terms, LADM divides the classes into three packages and one subpackage:

- **Party Packet:** This package represents the individuals, groups, and/or organizations related to the spatial units. The classes in this package include *LA_Party*, *LA_GroupParty*, and *LA_PartyMember*;
- **Administrative Packet:** This package deals with the rights, responsibilities, and restrictions to which each unit is subject. It includes classes such as *LA_RRR* (with its specializations *LA_Right*, *LA_Restriction*, and *LA_Responsibility*), *LA_BAUnit,* and *LA_AdministrativeSource*;
- **Spatial Unit Packet:** This package represents the spatial units, such as parcels, buildings, and infrastructure networks. It includes classes like *LA_SpatialUnit*,

LA_SpatialUnitGroup, *LA_Level*, among others;

– **Surveying and Representation SubPacket:** This subpackage, within the Spatial Unit Packet, is responsible for the geometric representations of spatial units and topological corrections through geographic information systems linked to databases. It includes classes such as *LA_Point*, *LA_SpatialSource*, *LA_BoundaryFaceString*, and *LA_BoundaryFace*.

The classes listed above are standardized by LADM (ISO 19152, 2012); however, each country may create additional classes to adapt the model to its specific needs.

It is also possible that some LADM classes are not represented in a cadastral application. Analyzing the fields defined by CIATA in light of LADM, for example, it is observed that the *Surveying and Representation SubPacket* was not included in the original CIATA project (Figure 1).

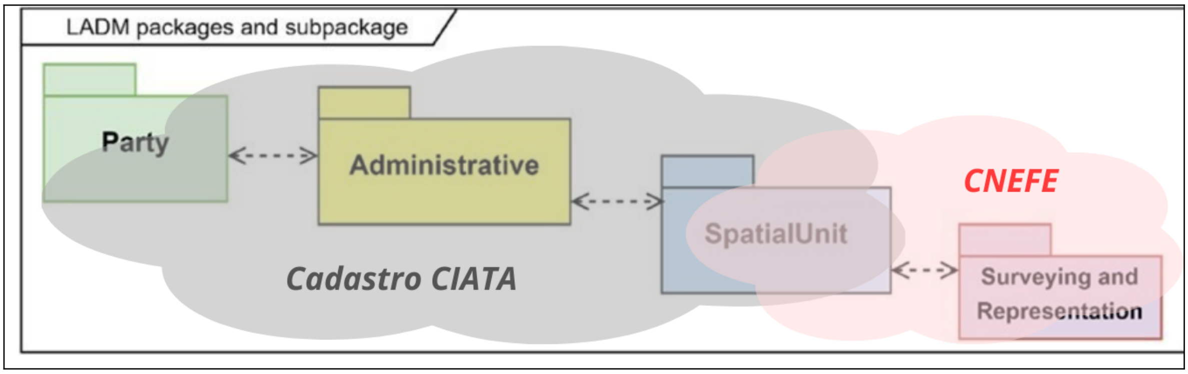

Figure 1. Basic LADM Classes (Packages) and Their Relationship with CIATA and CNEFE
Source: (Yılmaz; Alkan, 2024) and Authors' Adaptations.

2.2 CNEFE

According to IBGE (2025), *The National Address Database for Statistical Purposes* (CNEFE) is a nationwide database created in 2005. This database includes georeferenced addresses of households and establishments throughout the country.

CNEFE is administered by IBGE (*Brazilian Institute of Geography and Statistics*) and is fully updated during demographic censuses and periodically in other surveys. The data undergo a rigorous process of validation and standardization, ensuring its reliability (IBGE, 2024). Furthermore, the coverage of the data collection, which resulted in more than 103 million validated addresses in 2022, makes its georeferenced database one of the most comprehensive in the country.

In summary, starting with the 2022 Brazilians Census, for each address found in municipalities, census takers collected the coordinates of a point located on the street in front of constructed or under-construction units (IBGE, 2024).

It is important to highlight the concept of an address defined in the *Address Registration Standard Document* (IBGE, 2019): "*a text that allows the proper identification of a constructed unit within a municipality; that is, from this text, it must be possible to individualize and locate a point of interest within a given municipality*".

From the analysis of this definition of address, it can be inferred that not all objects of municipal cadastral interest are covered by CNEFE, since vacant lots and areas without specific use are not collected. On the other hand, the database includes addresses in rural areas, contains repeated georeferenced points, and includes attributes considered unnecessary, which need to be filtered for the effective processing of the *SuperCIATA* methodology.

Regarding georeferencing, according to the IBGE Methodological Note No. 01 (IBGE, 2024), the coordinates of the addresses are presented in the SIRGAS 2000 projection, in signed decimal degrees format with up to five decimal places. These coordinates were obtained using GNSS (*Global Navigation Satellite System*) devices embedded in the data collection equipment. Under normal data collection conditions (horizontal buildings, low-rise buildings, rural areas), the maximum error reaches 11.71 meters.

The CNEFE files for each Brazilian municipality are available on the official IBGE portal (IBGE, 2025) in *CSV* format (semicolon-separated) and *GeoJSON* format. Each record contains 35 attributes covering different dimensions of census interest.

Table 2 presents a comparison between CNEFE and LADM model, focusing on the structural and functional differences between them. In this context, it is highlighted that CNEFE does not include information related to the parties involved *(LA_Party* entities in LADM), nor does it represent the legal, technical, or administrative relationships between these parties and the registered addresses (*LA_RRR* — Rights, Restrictions, and Responsibilities). Moreover, the spatial representation is limited to indicating georeferenced points, without including information about the extent or boundaries of spatial units, such as areas or polygons.

In summary, the public data provided by IBGE through CNEFE can be interpreted, in light of LADM, as a partial representation focused on the basic administrative units (*LA_BAUnit*), lacking the legal and relational links that characterize a complete cadastral system according to the international standards of LADM (ISO 19152, 2012).

Table 2. Structural and functional comparison between CNEFE and LADM

Characteristic	CNEFE	LADM
Scope	Nationwide georeferenced address database managed by IBGE.	International conceptual model for land administration developed by FIG and ISO.
Data Collection	Periodically updated through demographic censuses and other surveys.	Not a database itself; the model can be implemented using different data sources depending on the context.
Attributes	35 attributes per record, focusing on census and statistical information.	Standardized objects and relationships covering rights, responsibilities, restrictions, and spatial units.
Georeferencing	Provides georeferenced points for addresses, but without defining parcel boundaries.	Includes comprehensive spatial representation (points, boundaries, parcels, and other spatial units).
Legal Relationships	Does not represent legal or administrative links between parties and addresses.	Explicitly models relationships through *LA_RRR* (Rights, Restrictions, and Responsibilities) and *LA_Party entities*.
Administrative Units	Represents basic administrative units (addresses) without legal or ownership information.	Provides a complete cadastral model including *LA_BAUnit* and its relationships with parcels and parties.
Interoperability	Limited to statistical and census purposes.	Designed to promote interoperability across countries and land administration systems following ISO 19152 standards.

2.3 OPEN STREET MAP – OSM

OpenStreetMap (OSM) is a collaborative project that provides open and free geographic data maintained by a global community of contributors. It contains detailed vector representations of urban elements such as streets (public roads), buildings, squares, among others (Haklay, 2010).

OSM was created in 2004 by Steve Coast in the United Kingdom as a response to the restricted access and high costs of proprietary geographic data (Haklay; Weber, 2008). Inspired by the collaborative model of Wikipedia, the project has established itself as the main free and open mapping initiative in the world, supported by a global community of volunteers (Goodchild, 2007).

The OSM database is organized into nodes, ways, and relations, which represent, respectively, discrete *points*, *lines* (streets and rivers), and *complex structures*. In this work, for the spatialization of municipal cadastral blocks, the coordinates of street segments (ways) are used, accessed through the Overpass API service. The Overpass API is a specialized query service for extracting OSM data using a language similar to SQL (Olbricht; Paulmann, 2015).

3. MATERIALS AND METHOD

3.1 Characterization of the Study Area

The municipality of Itabira, located in approximately 110 km from Belo Horizonte (The Minas Gerais State Capital), has a total area of 1,253.704 km², of which about 70 km² are urban. The estimated population in 2020 was 120,904 inhabitants.

The municipal real estate cadastre was outdated originally, consisting of alphanumeric information linked to the taxation system and geospatial data in digital format, but stored in disconnected databases. In this context, the need was identified to structure an updated conceptual model, in line with trends in the modernization of cadastral systems (LADM). This model should have the land parcel as the central object of the cadastre, in accordance with the definition of the International Federation of Surveyors (FIG), allowing its use as a basic reference for thematic cadastres and enabling integration with classes such as people, addresses, and taxes (Vieira et al., 2024).

3.2 Materials

This study uses three data sources: the CNEFE from IBGE, the municipal urban cadastre, and, when available, open data from *OpenStreetMap*.

To organize the different data sources, a naming convention was adopted by adding a domain identifier before table names: CN_ for CNEFE; CI_ for CIATA; and SC_ for SuperCIATA. The *SuperCIATA* application, built to test and demonstrate the methodology for integrating databases, also employs: the *MySQL database management system* (MySQL, 2025) for data storage; the *JavaScript programming language* (MDN, 2025) for performing calculations and developing pages; and the Leaflet library (Leaflet, 2025) for cartographic visualization. All these components are provided by the hosting service, and access to the application, documents, and code can be made via the URL: http://SuperCIATA.smuu.com.br.

3.3 Methods

3.3.1 Obtaining IBGE Points

The standardization, availability, and documentation of CNEFE provide simplicity to the process of importing its data into a relational database. In order to performer this task, it is

sufficient to download the file in CSV format and use the resources provided by *Database Management Systems* (DBMS) for import. At this initial stage, it is recommended to import the entire municipal file, proceeding later with the necessary filtering and structural transformations using SQL (*Structured Query Language*), a standardized programming language widely used for interacting with relational database management systems (Date, 2004).
Once the data are stored in a table named CN_PONTOS, it is possible to execute a single SQL query with multiple purposes: selecting the attributes of interest, excluding non-urban or low-quality addresses, concatenating attributes, and eliminating duplicate points. The code required for this task is shown in Table 3 and illustrates the efficiency and capability of SQL in handling and processing large volumes of data.

Table 3. SQL code for creating the CN_PONTOS_UNICOS

```
CREATE TABLE CN_PONTOS_UNICOS AS
SELECT
    COD_MUNICIPIO,
    COD_UNICO_ENDERECO,
    -- ID_QUADRA = concatenação COD_SETOR + NUM_QUADRA
    CONCAT(COD_SETOR, LPAD(NUM_QUADRA, 3, '0')) AS ID_QUADRA,
    -- ID_FACE = concatenação COD_SETOR + NUM_QUADRA + NUM_FACE
    CONCAT(COD_SETOR, LPAD(NUM_QUADRA, 3, '0'), LPAD(NUM_FACE, 2, '0')) AS ID_FACE,
    REPLACE(
        CONCAT(
            TRIM(COALESCE(NOM_TIPO_SEGLOGR,'')), ' ',
            TRIM(COALESCE(NOM_TITULO_SEGLOGR,'')), ' ',
            TRIM(COALESCE(NOM_SEGLOGR,''))
        ),
        '  ', ' '
    ) AS NOM_LOGRADOURO,     -- NOM_LOGRADOURO com limpeza de espaços
    NUM_ENDERECO,
    LATITUDE,
    LONGITUDE,
    ST_GeomFromText(
        CONCAT('POINT(', LONGITUDE, ' ', LATITUDE, ')'), 4326
    ) AS COORDS     -- Campo do tipo geometry (POINT)
FROM CN_PONTOS
WHERE NUM_QUADRA > 0
  AND NV_GEO_COORD < '4'
ORDER BY ID_QUADRA;
```

The **SC_PONTOS_UNICOS** table now presents new attributes:

- **ID_QUADRA = COD_SETOR + NUM_QUADRA**
- **ID_FACE = COD_SETOR + NUM_QUADRA + NUM_FACE**
- **NOM_LOGRADOURO = NOM_TIPO_SEGLOGR + NOM_TITULO + SEGLOGR + NOM_SEGLOGR**.

These modifications are necessary to align the concepts of CIATA and CNEFE. Other tasks in the methodology, such as calculating centroids and generating block identities, are performed through SQL queries available in Barbiero (2025).

3.3.2 Obtaining CIATA Data

Obtaining municipal cadastre data in the correct format can be quite challenging because, although most municipalities use a data model similar to CIATA, there is no standardization regarding the format or typology of attributes. In addition, it is very common for cadastre applications to be outsourced and store information in proprietary databases. In such cases, the administration must request that the software provider extract the records in CSV format

(*Comma Separated Values*) (*Open Knowledge*, [n.d.]) and, subsequently, make the necessary adjustments in the DBMS.
Doubts may arise regarding the obligation of third parties to provide data to the municipality. However, Brazilian legislation provides detailed guidance on this matter. According to Article 5 of the *General Data Protection Law* (LGPD, 2018), which regulates the processing of personal data in Brazil, two types of processing agents are distinguished: The ***Controller***, responsible for defining the purposes and means of processing personal data, and the ***Processor***, who carries out the processing according to the controller's instructions.
In the case of the urban cadastre, the municipality acts as the controller and may determine the operations to be carried out, including the export of information to open formats, with the possibility of charging compensation for the conversion work performed by the processor.

3.3.3 Obtaining Data from OpenStreetMap (OSM)

The street segment coordinates provided by *OpenStreetMap* are used to adjust the geographic slope of a road segment relative to the north–south axis and, by extension, of the polygon generated from municipal cadastre information.
With the centroid of the most populated face calculated and using the *Overpass API service*, a query is made in *OverpassQL* (OPENSTREETMAP, 2025) to return the coordinates of street segment points close to the centroid. Table 4 shows a query that searches for street way points within a radius of *around* meters from the latitude and longitude coordinates.

Table 4. Overpass API Query

```
`[out:json][timeout:25];
      (way["highway"]["name"~"${logradouro}",i](around:${raioMetros},${lat},${lng}););
      out body;>;out skel qt;`;
```

3.3.4 Integrating CNEFE and CIATA

The central idea of the *SuperCIATA* application is that the municipal territorial cadastre allows blocks to be represented as polygons (*Block Map*), and that CNEFE provides their approximate position. To achieve this, it is essential to establish correspondence between the blocks in the two cadastres.
This correspondence can be quite complex because municipalities freely define the values of the key components, without prior coordination with IBGE. Thus, there is no direct formula to relate the standard CIATA key to the CNEFE key. However, by combining the CIATA primary key with the street name of the addresses of lots/parcels, it is possible to identify which streets delimit the different blocks of the municipality.
By applying the same process to the CNEFE database, an attribute — **SC_ID_QUADRA** — is created, enabling linkage between the blocks in both databases.

3.3.5 Street Names

Since the strategy chosen to link the blocks in the two databases consists of using the names of the surrounding streets, it becomes essential to pay rigorous attention to these denominations. In any computerized system, the entry of incorrect data is a common occurrence. The proposed application has no authority over either database — that is, it cannot alter data in

the municipal cadastre or in CNEFE. Therefore, it is necessary to create a domain to standardize street names.
As IBGE data are already standardized, most of the preprocessing must be carried out on the municipality's database. The type of preprocessing required depends on individual analysis but will likely include the removal of accents and other special characters and the concatenation of fields.
The subsequent process involves identifying *toponymic equivalences* between the databases. The adopted strategy follows a hierarchical flow: (*i*) exact comparison; (*ii*) phonetic methods (Soundex/Double Metaphone) for unmatched records; and (*iii*) finally, similarity algorithms (Levenshtein) for residual cases — according to the methodology validated by Manning et al. (2008). Records irrecoverable by these methods require the use of artificial intelligence techniques, manual intervention, or controlled discarding (Macedo, 2023).
Once correspondence between street names is established, a table is created that includes **COD_MUNICIPIO**, the original street names from CIATA and CNEFE, as well as a unique identifier for the street.
The unique identifier of the street correspondence, called **SC_ID_LOGRADOURO**, is a hexadecimal digest value calculated by the *CRC32* function applied to the concatenation of the other attributes in the following order: COD_MUNICIPIO, NOM_LOGRADOURO_CIATA, and NOM_LOGRADOURO_CNEFE. This identifier will be incorporated into the address lists of both databases. After this attachment, the application starts identifying streets by SC_ID_LOGRADOURO in all database operations.

3.3.6 Creation of SC_ID_QUADRA

After generating the SC_ID_LOGRADOURO values, they are grouped in ascending alphabetical order into a vector, according to the blocks they delimit. This vector is inserted into the block tables of CIATA and CNEFE and becomes a unique identity of the block shared by both databases.
Once the SC_ID_QUADRA attribute has been appended to the tables, they can be correlated automatically.

3.3.7 Graphical Representation of Blocks

The procedure adopted to position the municipal cadastre block close to its real geographic location consists of aligning the midpoint of the CIATA polygon side with the centroid of the selected CNEFE block face and drawing the block according to the dimensions established in the CIATA cadastre.
The function that draws the block in *SuperCIATA* receives as parameters: the ordered list of street IDs; block face dimensions; block face centroids; and the number of unique points of the faces.
The block is drawn with the original dimensions of the CIATA cadastre, initially at an angle parallel to the equator, and then adjusted using OpenStreetMap data, when available. The accuracy verification of the *SuperCIATA* method will be carried out using the statistical metric *Root Mean Square Error* (RMSE). Its calculation consists of obtaining the mean of the squared differences between observed and estimated values and then extracting the square root

of this result (Bussab; Morettin, 2021). RMSE is widely used in the validation of spatial predictive models (Almeida et al., 2007).

3.3.8 Adjustment with OpenStreetMap

The OpenStreetMap data are used to adjust the angle of the CIATA drawing to the rotational angle of the street segment corresponding to the face used as an anchor. It is not always possible to obtain data with adequate precision because OSM is a cooperative database. In such cases, the block is drawn without adjustments.

4. RESULTS

In Figure 2 it is possible to visualize the centroid of the selected block (Figure 2(a)) and the graphical representation of the block adjusted by the angle obtained from OSM (Figure 2(b)).

(a) Centroid of the block face (b) Adjusted Block

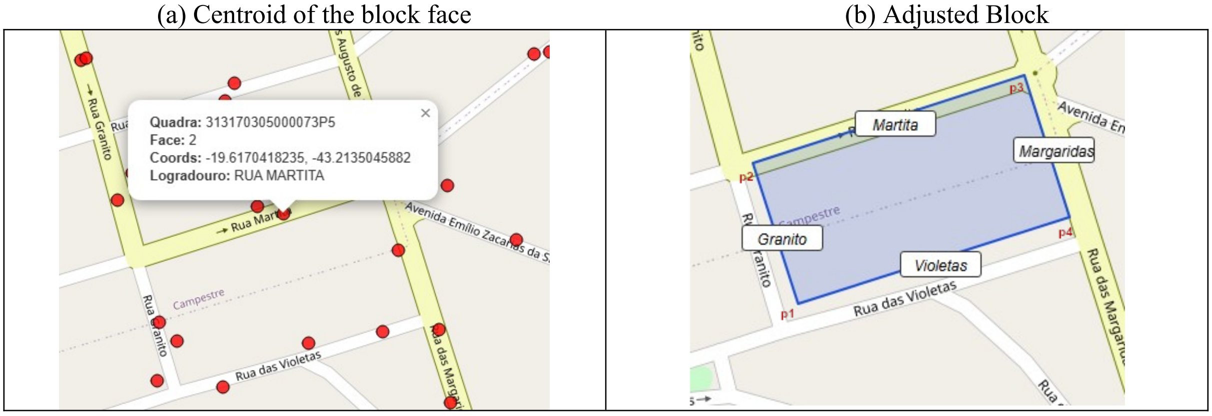

Figure 2. Centroid of the face and adjusted block[

Figure 3 shows the *shapefile* extracted from the Territorial Information System (SIT) of the municipality of Itabira - MG, which serves as a reference for validating the methodology. The data used for calculating the *mean error* and the *root mean square error,* presented in section 4.1, were extracted from this *shapefile*. The highlighted points represent the block vertices calculated by *SuperCIATA* (Figure 2(a)) and georeferenced by the municipality (Figure 3). It is observed that there is no exact coincidence among any of the three coordinate sources: municipality cadastre, IBGE, and *OpenStreetMap*.

Figure 3. Original shapefile from the Municipality of Itabira

Table 5 presents the errors, calculated using the statistical metric *Root Mean Square Error* (RMSE), that occur when the CIATA block is drawn using the block centroid as an anchor and after angle adjustment with OSM data. The data refers to a block in the municipality of Itabira-MG, and the reference values were obtained from the *shapefile* extracted from the municipality's *Territorial Information System* (SIT).

Table 5. Mean error with angle adjustment using OSM

Point	Latitude SIT-Itabira	Longitude SIT-Itabira	Latitude Adjusted	Longitude Adjusted	ΔLatitude (°)	ΔLongitude (°)	Distance (°)	Error (m)
1	-19.61782	-43.21399	-19.61779	-43.21398	0.000027	0.000004	0.0000273	3.03
2	-19.61725	-43.21417	-19.61723	-43.21417	0.000020	0	0.0000200	2.22
3	-19.61691	-43.21301	-19.61688	-43.21302	0.000034	0.000009	0.0000352	3.91
4	-19.61748	-43.21280	-19.61744	-43.21282	0.000032	0.000025	0.0000406	4.51

Metric	Value (m)
MAE (Mean Absolute Error)	3.42
RMSE (Root Mean Square Error)	3.53
Maximum Error	4.51

The observed errors are significant and may have several sources: limitations in the precision of the equipment used to collect coordinates; atmospheric interferences; displacement of the face centroid due to the irregular distribution of points; inconsistencies in parcel dimensions in the municipal cadastre; among others.
The reduction of these errors is possible through cadastre corrections or by conducting new coordinate collections at critical points, but each case must be analyzed individually.

5. CONCLUSIONS

The objective of this work was achieved, demonstrating the feasibility of using CNEFE data in association with a consistent municipal territorial cadastre and *OpenStreetMap* for the georeferencing of urban parcels at different levels of precision.
Considering that the quality of CNEFE data collection is homogeneous throughout the country, the georeferencing quality levels of *SuperCIATA* can be summarized as follows:

a) If an urban block has at least one constructed property on each face in the latest CNEFE survey, then a realistic municipal urban cadastre is sufficient to ensure the georeferencing of urban parcels with a maximum error limited to half of the block's dimensions. This condition establishes a minimum reliability benchmark for the proposed method;

b) If at least one of the block faces has a sufficient number of unique CNEFE points to generate a face centroid close to the real one, then a realistic municipal urban cadastre becomes sufficient to allow georeferencing of urban parcels with reasonable accuracy. This shows that the density of CNEFE points on each face is crucial to refine the spatial result;

c) Finally, if the representation of a CIATA block significantly diverges from the *SuperCIATA* map in shape and/or size, it is likely that the municipal cadastre has data integrity issues. This finding reinforces the potential of the methodology not only for georeferencing but also as a tool for auditing and validating the municipal cadastre.

Georeferencing with the precision obtained by the methodology can facilitate urban planning and compliance with various legal obligations. However, it is not sufficient, for example, to replace the mandatory georeferencing of properties for registration purposes.

As can be seen from the methodology description, the use of CNEFE and municipal cadastres in the CIATA standard restricts the application of *SuperCIATA* to municipalities in Brazil. Even so, the methodology can be applied in other countries, provided there is a database with georeferenced block information and another with block identification and lot dimensions. In the absence of georeferenced databases, municipal administrations may collect only the corner positions and associate them with the primary key of the municipal cadastre.

Regarding the use of artificial intelligence methods, *SuperCIATA* does not directly adopt them in the parcel geolocation process. However, the collection of images before and after adjustments, together with the detailed *operation log*, can serve as a solid and diverse training dataset for deep learning applications. Thus, future versions of the system may incorporate knowledge extracted from this database, enhancing both the accuracy and automation of the processes.

REFERENCES

ALMEIDA, C. et al. (Org.). Geoinformação em urbanismo: cidade real x cidade virtual. São Paulo: Oficina de Textos, 2007.

BARBIERO, M. A. SUPERCIATA - Georreferenciamento Municipal. Aplicação. Disponível em: http://superciata.smuu.com.br/. Acesso em: 28 ago. 2025.

BUSSAB, W.; MORETTIN, P. Estatística Básica. São Paulo: Editora Saraiva, 2021.

CARNEIRO, Andréa F. T. Cadastro imobiliário e registro de imóveis: a Lei n. 10,267/2001, Decreto n. 4,449/2002 e atos normativos do INCRA. São Paulo: S.A. Fabris Editor, 2003.

CUNHA, Eglaisa et al. O cadastro urbano no Brasil: histórico e evolução. Revista de Geografia e Ordenamento do Território, v. 0, n. 17, p. 55–74, 30 jun. 2019.

DATE, C. J. Introdução a Sistemas de Banco de Dados. 8. ed. Rio de Janeiro: Elsevier, 2004.
ENEMARK, S.; MCLAREN, R.; LEMMEN, C. Fit-for-Purpose Land Administration—Providing Secure Land Rights at Scale. Land, v. 10, n. 9, p. 972, 15 set. 2021.

GOODCHILD, M. F. Citizens as sensors: the world of volunteered geography. GeoJournal, v. 69, n. 4, p. 211–221, 30 nov. 2007.

HAKLAY, M. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. Environment and Planning B: Planning and Design, v. 37, n. 4, p. 682–703, ago. 2010.

HAKLAY, M.; WEBER, P. OpenStreetMap: User-Generated Street Maps. IEEE Pervasive Computing, v. 7, n. 4, p. 12–18, out. 2008.

IBGE. Cadastro Nacional de Endereços para Fins Estatísticos. Disponível em: https://www.ibge.gov.br/estatisticas/sociais/habitacao/38734-cadastro-nacional-de-enderecos-para-fins-estatisticos.html?=&t=o-que-e. Acesso em: 4 jun. 2025.

IBGE. Censo Demográfico 2022. Coordenadas Geográficas dos Endereços. Nota metodológica n. 01. Rio de Janeiro, 2024. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv102063.pdf. Acesso em: 4 jun. 2025.

IBGE. Padrão de Registro de Endereços - liv101639. Rio de Janeiro, 2019. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101639.pdf. Acesso em: 4 jun. 2025.

ISO 19152:2012(en), Geographic information — Land Administration Domain Model (LADM). 2012. Disponível em: https://www.iso.org/obp/ui/en/#iso:std:iso:19152:ed-1:v1:en. Acesso em: 6 dez. 2024.

LEAFLET - a JavaScript library for interactive maps. Disponível em: https://leafletjs.com/. Acesso: 10 ago. 2025.

MACEDO, D. Integração do Cadastro Único com Cadastro Nacional de Endereços para Fins Estatístico através da modelagem de um banco de dados espacial. [S. l.]: mds.gov.br, 2023. Disponível em: https://aplicacoes.mds.gov.br/sagi/pesquisas/documentos/relatorio/relatorio_270.pdf. Acesso em: 3 mar. 2025.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. Introduction to information retrieval. Cambridge: Cambridge University Press, 2008.

MDN. JavaScript. Disponível em: https://developer.mozilla.org/pt-BR/docs/Web/JavaScript . Acesso: 10 ago. 2025.

MySQL. Database Management System. Disponível em: https://www.mysql.com/. Acesso: 10 ago. 2025.

OLBRICHT, R.; PAULMANN, M. Overpass API. FOSS@HFT, 2015. 1 vídeo (44 min). Disponível em: https://av.tib.eu/media/17720 Acesso em: 17 jul. 2025.

OPENSTREETMAP WIKI. Overpass API/Overpass QL. Disponível em: https://wiki.openstreetmap.org/wiki/Overpass_API/Overpass_QL. Acesso em: 11 ago. 2025.

SERPRO. Superciata/PROJETO CIATA - MANUAL DO CADASTRO IMOBILIÁRIO.pdf. In http://superciata.smuu.com.br/docs/ProjetoCiata_ManualCadastroImobiliario_OCR.pdf . Acesso: 20 jul. 2025.

SILVA, Everton da (ORG.). Cadastro Territorial Multifinalitário aplicado à gestão municipal. Florianópolis, SC: Ufsc, 2023.

VIEIRA, C. A. O. et al. Proposição de modelos de governança e utilização de instrumentos de relações intergovernamentais para implementação do cadastro territorial multifinalitário. In: congresso brasileiro de cadastro multifinalitário e gestão territorial, COBRAC 2024, Florianópolis. Anais [...]. Florianópolis: UFSC, 2024

YILMAZ, O.; ALKAN, M. Applicability of spatial planning system package for the LADM Turkey country profile. Transactions in GIS, v. 28, n. 4, p. 858–883, jun. 2024.

BIOGRAPHICAL NOTES

Marco Aurélio BARBIERO - Federal Revenue Auditor of Brazil; Master's student in Territorial Management at PPGTG - UFSC; Postgraduate in Big Data from PUC-MG; Bachelor in Computer Science from the University of Passo Fundo – RS.

Carlos Antonio Oliveira VIEIRA – Professor Titular at the Federal University of Santa (UFSC) Catarina, PhD in Physical Geography, MSc. in Systems and Computing, BSc. Survey Engineering, B.Sc. In Public Administration.

CONTACTS

Marco Aurélio Barbiero
Universidade Federal de Santa Catarina
Rua Capitão Araújo, 551, 301
99010200, Passo Fundo - RS
BRAZIL
Phone: +55 54 991811915
E-mail: marco.barbiero@posgrad.ufsc.br
Website: http://superciata.smuu.com.br