Gentrification. 2025; 3:93
doi: 10.62486/gen202593
ISSN: 3046-4137

ORIGINAL

Analysis of dynamic scenarios generated by urban growth in the city of Pasto using a predictive model supported by a software product

Análisis de escenarios dinámicos generados por el crecimiento urbano en la ciudad de Pasto mediante un modelo predictivo respaldado por un producto software

Jheison Edilson Arteaga Quistial[1] ✉, Miguel Ángel Velásquez Bravo[1] ✉, Omar Alexander Revelo Zambrano[1]

[1]Universidad CESMAG, Facultad de Ingeniería, Ingeniería de sistemas. Pasto – Nariño, Colombia.

Cite as: Arteaga Quistial JE, Velásquez Bravo M Ángel, Revelo Zambrano OA. Analysis of dynamic scenarios generated by urban growth in the city of Pasto using a predictive model supported by a software product. Gentrification. 2025; 3:93. https://doi.org/10.62486/gen202593

Submitted: 18-03-2024 **Revised:** 24-07-2024 **Accepted:** 05-07-2025 **Published:** 06-07-2025

Editor: Prof. Estela Hernández-Runque

ABSTRACT

The objective of this research is to analyse urban growth in Pasto using a predictive model in order to identify expansion patterns and anticipate future scenarios. The study responds to the need to manage rapid urban change in the city, which has gone from being mainly agricultural to urban in the last two decades. Through the analysis of urban data, it seeks to facilitate planning and informed decision-making, promoting sustainability and organised development.

Keywords: Urban Growth; Predictive Model; Urban Repository; Data Analytics; Web Application; Residential Projects.

RESUMEN

El objetivo de esta investigación es analizar el crecimiento urbano en Pasto mediante un modelo predictivo, con el fin de identificar patrones de expansión y anticipar escenarios futuros. El estudio responde a la necesidad de gestionar el rápido cambio urbano en la ciudad, que ha pasado de ser mayormente agrícola a urbano en las últimas dos décadas. A través del análisis de datos urbanísticos, se busca facilitar la planificación y toma de decisiones informadas, promoviendo la sostenibilidad y el desarrollo organizado.

Palabras clave: Crecimiento Urbano; Modelo Predictivo; Repositorio Urbanístico; Analítica de Datos; Aplicación Web; Proyectos Residenciales.

INTRODUCTION

Currently, the disorganized expansion of both formal and informal housing growth in the city of Pasto has generated an urban landscape marked by significant densification in the peripheries and rural areas.[1,2,3,4,5] This reality poses substantial challenges for the city's territorial planning and management, manifesting itself as an evident transition from an agricultural to an urban environment, especially noticeable between the years 2000 and 2022. The implementation of occupation models in this period has caused a profound structural transformation in the urban-territorial configuration of Pasto.[6,7,8,9,10] This change is clearly manifested in the complexity of the urban morphology, where peripheral points converge towards a compact footprint.[11,12,13,14,15] The predominant focus on peripheral areas as a solution for urban development highlights the urgent need to address the challenges arising from territorial expansion and cohesion in the city.[16,17,18,19,20]

Over more than two decades, numerous urban simulation and modeling studies, from the contributions of

Tobler, Couclelis, Takeyama, and White and Engelen to more contemporary research by Feng, Liu, and Batty and Feng and Tong, have highlighted the functionality of these approaches. However, the current dispersion of the information systems of the city of Pasto in various sources, some outdated or not purified, has generated information inconsistent with the reality of the city. Likewise, the available data exhibit an intrinsic dynamic in constant evolution, undergoing transformations closely linked to the urban growth processes in the expansion areas.[21,22,23,24,24,25] The lack of reliable and easily accessible sources complicates the task of establishing direct and precise relationships with updated and timely data. The complications inherent in the use of and access to information act as barriers that restrict the smooth and functional transmission of the corresponding repository.[26,27,28,29,30] Therefore, the research proposal entitled "Analysis of urban growth in the expansion zones of Pasto 2000-2022: implications of the compact territorial model proposed by the land use plans in urban planning, by means of a predictive model".[31,32,33,34,35] The research is in process of execution, in which research processes of great relevance have been obtained, highlighting the development of quality software products.[36,37,38,39,40] Thus highlighting the development of a web portal as a consultation tool, which was presented at the III International Meeting of Semilleros of CESMAG University, obtaining recognition by the evaluators as an outstanding work for its innovation in the field of architecture and systems engineering.[41,42,43]

With the above, several factors to improve have been identified; therefore, the initiative of developing an interactive web page arises applying innovative technologies such as predictive models,[44,45,46] which are supported by artificial intelligence, allowing to highlight the urban growth in the city of Pasto.[47,48,49] Likewise, its free use and easy access and understanding of the information will be its main attraction.[50,51,52] This approach will not only include the consolidation of validated data, but also its visualization through interactive software that aims to collect reliable data and present the information in an accessible way, providing users with a detailed understanding of the evolution of urban growth in Pasto.

The ongoing research seeks to identify and analyze in detail the transition of Pasto from an agricultural to an urban environment, recognizing the complexity and implications that this change has had on the configuration of the city over time. It seeks to comprehensively examine how this phenomenon has left a significant mark on the urban-territorial structure, influencing the planning and development of the city. The research process delves into the emerging morphology, highlighting the key elements that have contributed to the transformation, from the occupation models implemented to the dynamics of formal and informal growth.

Problem statement

With the enactment of Law No. 388 of 1997, which is established as the main regulation in territorial planning, Colombia began a new stage in the planning and management of cities and urban centers in the country. However, these regulations are effective only if the hierarchy and preponderance of urban land compared to other types of land is understood. In other words, planning focuses on the materialization, densification and development of the urban space we know as cities. From another perspective, a lack of optimal management tools for the development of rural and suburban areas, whose vocation is specifically focused on rural, agricultural and tourism activities, is identified.[53,54,55]

Currently, information systems are available to preserve data and variables related to urban development patterns in Pasto. However, this information comes from various sources and, in some cases, has not been subject to updating, cleaning or modification, resulting in the presence of data inconsistent with the real situation in the city. In addition, it is crucial to highlight that this information does not remain static, but experiences a constant dynamism and evolution in parallel to the urban growth processes in the expansion areas of Pasto. The lack of reliable sources available for direct consultation by users makes it difficult to access updated and timely data. In the context of a city like Pasto, characterized by the management of multiple repositories stored in various sources, the use of information becomes more complex hindering access to it. Considering data analytics and the collection of information in an accurate and concise manner becomes a fundamental reference to achieve an effective development and understanding of the situation during the periods 2000 and 2022. "Pasto is one of the cities in Colombia that has undergone a process of transit from an agricultural to an urban landscape. This fact became more evident with the occupation models implemented between 2000 and 2022, a topic that was addressed by experts in the program OGU, of Radio UNAL". Considering the above quote as a reference, which evidences the significant transition that Pasto has experienced over the years, it is imperative to conduct an analysis and information gathering to determine the factors that have contributed to this process of change and growth.

How has the city of Pasto experienced its urban growth between 2000 and 2022, evaluated through dynamic scenarios and supported by a predictive model?

General Objectives

To determine interactive scenarios of urban growth in the city of Pasto through a predictive model, supported by a software product.

METHOD

Paradigm

The research focuses on the positivist paradigm based on replicability, which involves the predictive model fed by a reliable data source collected during the research. A fundamental characteristic of this paradigm is causality, due to the fact that the research process emphasizes on identifying the cause, which is urban growth, and its effects, which correspond to changes in the dynamic scenarios of the city of Pasto.

Approach

The research approach aims at creating a software product that allows the identification of urban growth in the city of Pasto. For this purpose, a quantitative approach is taken into account, which is characterized by the study of a matrix resulting in an information repository containing more than 30 000 urban data. This data will be the driving force for the predictive model to be fed. From the analysis of this data, the quantitative approach allows us to identify which sectors of Pasto have experienced greater urban growth, as well as to determine which types of constructions have been more constant between 2000 and 2022. The importance of this analysis lies in the fact that, by identifying patterns and trends, the model will be able to make informed projections about the future of urban development in the city, allowing the information obtained to be effectively visualized in the software product developed.

Method

Since this research encompasses both quantitative and qualitative aspects, it can be classified as a mixed research. In this sense, various methods are employed, including documentary research in combination with field research, which in turn is broken down into surveys, observation and data analysis.

For the implementation of the predictive model, the validation, evaluation and analysis of the results, aspects that were contemplated during the data collection stage, are considered essential.

Type of research

The research is exploratory, given the complexity of analyzing dynamic scenarios in the framework of urban growth in the city of Pasto and the possibility of having a limited history of previous research. It is proposed to start with an exploratory research that allows an in-depth analysis of the dynamics of urban growth in Pasto, identifying key and relevant factors that allow the correct study of such dynamics.

Once an adequate level of understanding of the dynamics of urban growth in Pasto has been reached, we proceed to an experimental type of research. In this phase, the predictive model is developed and validated using previously processed data. In addition, we evaluate how various key factors influence urban growth.

Research design

The research adopts the quasi-experimental design, which allows the manipulation of specific independent variables, such as urban policies or urban planning, important to analyze their impact on urban growth in Pasto. This design facilitates the application of controlled interventions, followed by the evaluation of their effects on the dynamics of urban growth in the aforementioned city.

In this context, study groups representing different urban scenarios or conditions are defined. Key variables related to urban growth are manipulated to observe how they influence both the predictive model and the results obtained.

This design enables a more controlled and structured approach capable of analyzing the dynamics of urban growth in the city of Pasto, facilitating the evaluation and validation of the predictive model supported by the developed software.

Population

The target population of this research focuses on urban growth in the city of Pasto, addressing aspects related to buildings, condominiums, residential complexes and housing in general.

Sample

The sample covers the period between the years 2000 and 2022. During this interval, an exhaustive analysis was carried out focused on the study and evaluation of urban growth with respect to the multiple urban variables obtained in the research process.

Data collection techniques

The first phase focused on conducting a survey of people who have knowledge in the area of housing construction in the city of Pasto in the periods 2000-2022, in order to identify the residential projects that have been built and developed over the years, as well as each characteristic that identifies each residential project.

Likewise, a thorough research in virtual sources was carried out to identify several key factors needed for the study of the urban growth of the city.

Validity of the collection techniques

The techniques to be employed are based on the synergy between urban growth and technology. The creation and implementation of a survey represents the primary phase to determine the areas of the city of Pasto that should be analyzed in order to subsequently work on them and collect the necessary information for the implementation of a predictive model defined by urban variables, which include relevant information for training and development. The analysis of the paper supports the feasibility of the study and the integration of urban growth and technology, with the purpose of offering a perspective on how to identify urban growth in the city of Pasto.

Reliability of the collection techniques

The reliability of the information is based on the analysis of the information and its coherence with the context and the questions posed in the survey. To achieve this, a detailed explanation of the purpose and aim of the survey or the topics to be covered will be provided, either in visual or face-to-face format. This will ensure that the respondent can provide valid and accurate information. Similarly, in data collection through virtual tools, comparative filters will be applied using various sources of information to identify patterns of consistency that validate the reliability and validity of the information provided.

Data collection instruments

In the data collection process, advanced technological tools and various online sources will be used. These tools will include platforms such as Google, Google Scholar, Google Maps and Google Earth, which will allow obtaining geospatial data (study variables: longitude and latitude of residences located in Pasto), maps and satellite views to visualize the urban environment of Pasto and track its evolution over time. Likewise, home price information pages and websites that offer visual resources to identify what the city of Pasto was like before the year 2000 will be explored, together with the Instituto Geográfico Agustín Codazzi (IGAC) as a primary source to guarantee the reliability of the information. These sources will be crucial for the development and population of an information matrix with valid and accurate data to support the analysis of urban growth in the city.

Additionally, the surveys that will be carried out in the city of Pasto, aim to collect essential information for the research process, allowing to analyze and study the urban growth of the city, focusing also on the variables of study of the research:

- Name of the Project (Residence, Building or Complex): where it is sought to obtain specific information on the denomination of the residence, building or housing complex located in a certain area. Seeking to understand how these spaces are identified and differentiated within the city.
- Type of Housing (Morphology): allowing to know in detail the morphology of the dwellings, that is, if they are multi-family or single-family. This classification will help to understand the predominant residential structure in the city of Pasto, providing valuable information on architectural diversity.
- Housing Stratum: allowing to identify the stratum to which each residence, building or residential complex belongs. This variable is crucial to analyze the socioeconomic distribution in the city, making it possible to understand the diversity of strata present.
- Housing Size: in this section, we intend to obtain information about the dimensions of the dwellings, expressed in square meters. This will make it possible to understand the variability in the size of residences or buildings in the area, providing valuable insights into the distribution of residential space.

Only a few items obtained through the previous surveys were recorded in the repository, while the majority were collected through an exhaustive documentary review. Reliable sources on websites, analysis through Google Maps and the support of confidential and valuable information provided by EMPOPASTO were used, which allowed complementing and validating the information in a comprehensive manner.

RESULTS

The creation of an urban data repository in the city of Pasto has been fundamental for the analysis of urban growth between 2000 and 2022. This repository compiles detailed information on residential projects, including location, housing characteristics and risk contexts, serving as an essential tool for urban planners and local authorities in the identification of expansion patterns and evaluation of urban policies.

To efficiently manage this information, an intuitive web application was developed, allowing users to access relevant data and visualize the results of a predictive model. This model, designed to analyze urban growth, uses machine learning techniques and has been validated with metrics such as the coefficient of determination (R^2) and the mean square error (MSE), demonstrating its effectiveness in forecasting future trends.

The integration between the web application and the predictive model is facilitated by an API developed with Flask, which allows data to be sent and predictions to be obtained in real time. Each of these components has been developed following the following specific methodologies:

- Data Warehousing for the creation and structuring of the data repository.
- Scrum for the development of the web application, allowing an agile and adaptive approach.
- CRISP-DM for the development of the predictive model, ensuring a rigorous analysis based on historical data.

This combination of methodologies and tools established a robust framework for the planning and sustainable development of Pasto, fulfilling each of the specific objectives of the research.

Creation of an urban data repository

After arduous work and meticulous information gathering, a data repository was compiled that groups all the residential projects built in the city of Pasto between 2000 and 2022. This repository, now complete and well documented, represents an invaluable source for any entity interested in the study and analysis of urban growth in the region.

The repository contains an extensive collection of geospatial data representing residential projects built between the years 2000 and 2022, which also includes information or urban characteristics such as: The risk or threat zone, main ecological structure, number of housing units, housing type (morphology), housing type (subsidy), housing stratum, housing size, housing price, year of construction, population density (hab/Hec), facilities (m), location of the residential project (Longitude), location of the residential project (Latitude). This resource provides a detailed view of Pasto's urban growth between 2000 and 2022, serving as a key tool for urban planners, academics and authorities to identify expansion patterns and evaluate policies. In addition, it supports strategic decisions on sustainable development and infrastructure improvement, and is also valuable for real estate companies and builders in understanding the residential market.

In the context of the creation of the urban repository, the Data Warehousing Methodology was chosen as the most appropriate to carry out an effective process. This methodology is structured in the following key phases:

Data Collection

- Source Identification: in this phase, it was determined which data and sources were needed to enrich the urban repository. Data collection was carried out from various sources, including documentary review and obtaining geospatial information using Google Maps and Google Earth. In addition, Empopasto contributed by providing urbanistic information relevant to the research, which was fundamental for the construction of the data repository.
- Data Extraction: where the data collection process was carried out from the identified sources.

Data Transformation

- Data Cleaning: phase focused on eliminating duplicate data, correcting errors and standardizing formats.
- Standardization: key stage to ensure that the data follows a consistent and structured format necessary for its understanding.

Data warehousing

- Database Design: where the database structure was defined, considering the tables that could be used as well as the relationships between them.
- Data Loading: process of inserting the transformed and debugged data into the repository as such, to be manipulated by the web application for its management and by the predictive model to make predictions.

Documentation and metadata

- Metadata creation: phase where the process of documenting the source, format and context of the data was carried out to facilitate its use and understanding.
- Use Guides: a useful guide was developed to understand the structure and characteristics of the data and the urban planning repository.

Maintenance and updating

- Regular updating: stage where a schedule and the necessary structure for the incorporation of new data in the repository were established.
- Quality Monitoring: phase focused on the periodic review of data quality.

Accessibility and updating

- User Interface: an interface was created in the web application to facilitate data access and consultation, also guaranteeing the correct management and security of the data.

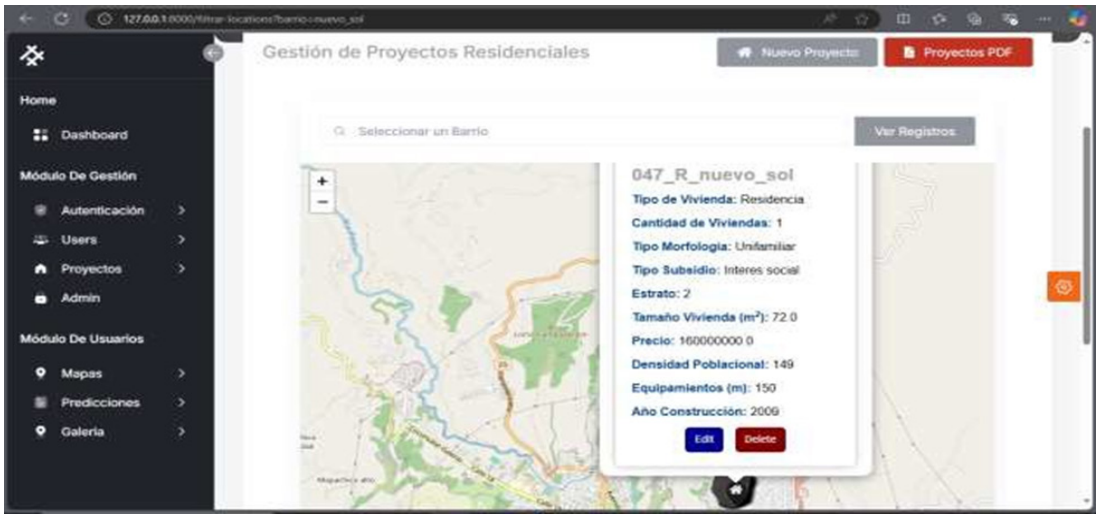


Figure 1. Interface of the web application that allows the management of the urban planning repository

- Visualization tools: stage where statistical graphs and a Dashboard were implemented to analyze and visualize the data interactively.

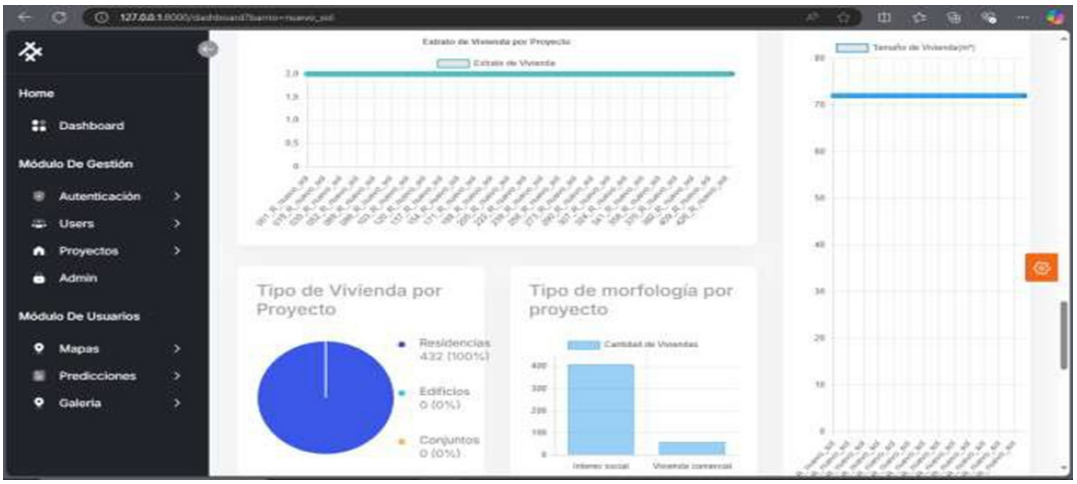


Figure 2. Interface of the web application where it allows the visualization of statistical graphs by neighborhood (1 neighborhood only)

Figure 3. Interface of the web application where it allows the visualization of statistical graphs comparing all the neighborhoods of the city

Validation in the collection of information for the urban planning repository: support from Empopasto

Data collection for the urban planning repository is a fundamental process that relies on the truthful and accurate information provided by Empopasto. Empopasto's experience and commitment to data quality ensures that the information collected is not only reliable, but also accurately reflects the characteristics and needs of the urban environment. This contribution becomes an essential pillar for the development of the research process, ensuring that future decisions are based on solid and relevant evidence.

Information validated and provided by Empopasto:

Source: EMPOPASTO

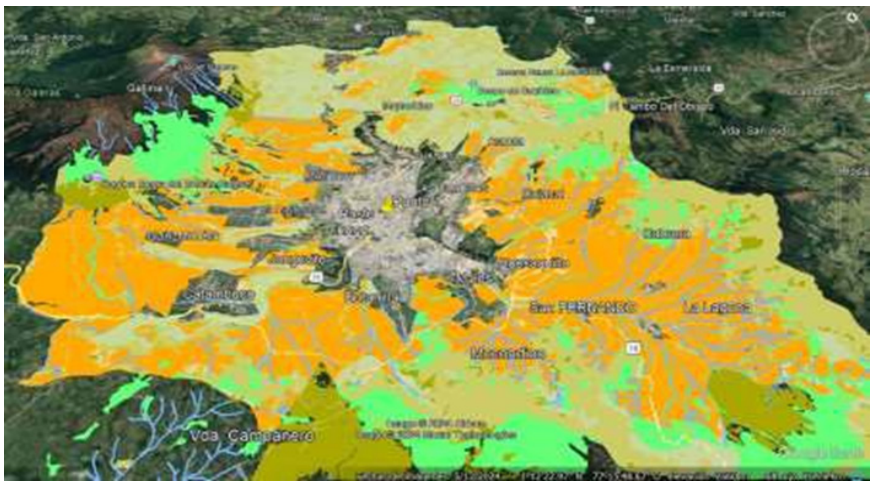
Figure 4. Projection of a point in google earth working risk or threat zone

COLOR IDENTIFICADOR	Tipo de zona
	Zona con amenaza por inundacion
	Zona con amenaza por deslizamiento
	Zona por colapso por mineria
	Zona por colapso por mineria
	Zona Sin Riesgo

Source: EMPOPASTO

Figure 5. Classification of risk zones according to an identifying color

By placing a point on the map and assigning it a specific color, a risk zone corresponding to that residential project is identified. This action allows a clear and precise selection of the affected area, facilitating the diligence of the value in the "risk_zone" column of the repository. Thus, an efficient management of the information is guaranteed, allowing a deeper analysis of the associated risks in the city of Pasto.



Source: EMPOPASTO

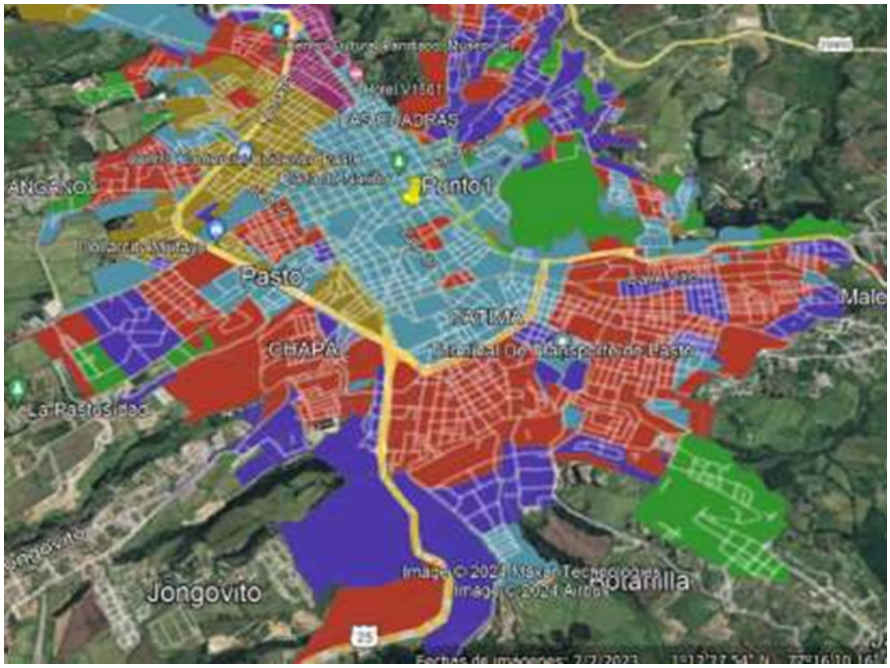
Figure 6. Projection of a point in google earth working main ecological structure

COLOR IDENTIFICADOR	Estructura Ecológica
	Areas agricolas
	Drenaje doble
	Areas de Amenazas Naturales
	Drenaje sencillo
	Areas forestales protectoras
	Areas Agrosilvopastoriles
	Sin zona ecológica

Source: EMPOPASTO

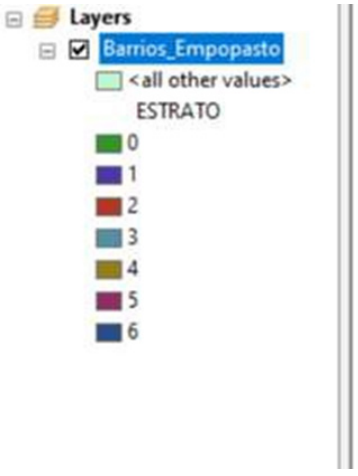
Figure 7. Classification of ecological structures according to an identifying color

By placing a point on the map and assigning it a specific color, a zone is established with respect to the ecological structure for that residential project. This action allows a precise selection of the area in question, facilitating the diligence of the value in the "ecological_structure" column of the repository. In this way, a clear organization of the information is ensured, which contributes to a better understanding and analysis of the ecological elements present in the territory.



Source: EMPOPASTO

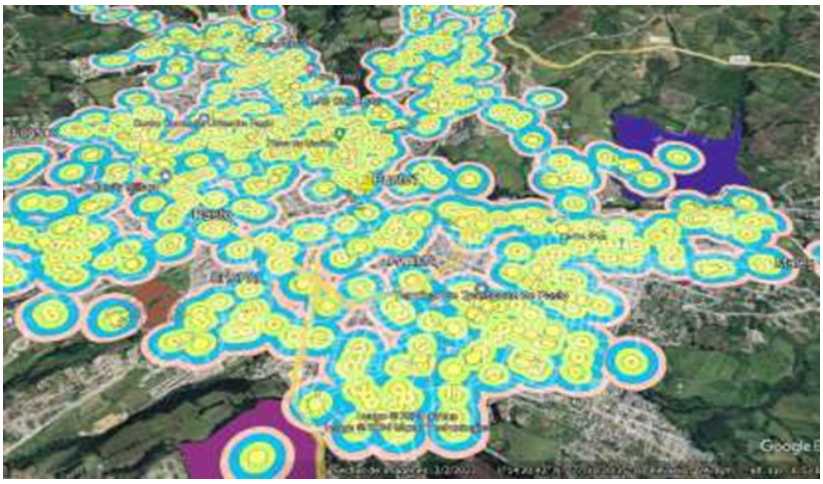
Figure 8. Projection of a point in google earth working stratum



Source: EMPOPASTO

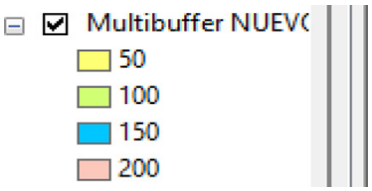
Figure 9. Classification of Pasto strata according to an identifying color

By placing a point on the map and assigning it a specific color, the stratum for that point or residential project is defined. This action allows for the precise selection of the corresponding area, facilitating the diligence of the value in the "housing_stratum" column of the repository. Thus, a clear categorization of the information is guaranteed, which contributes to a deeper analysis of the socioeconomic characteristics of the population in the territory.



Source: EMPOPASTO

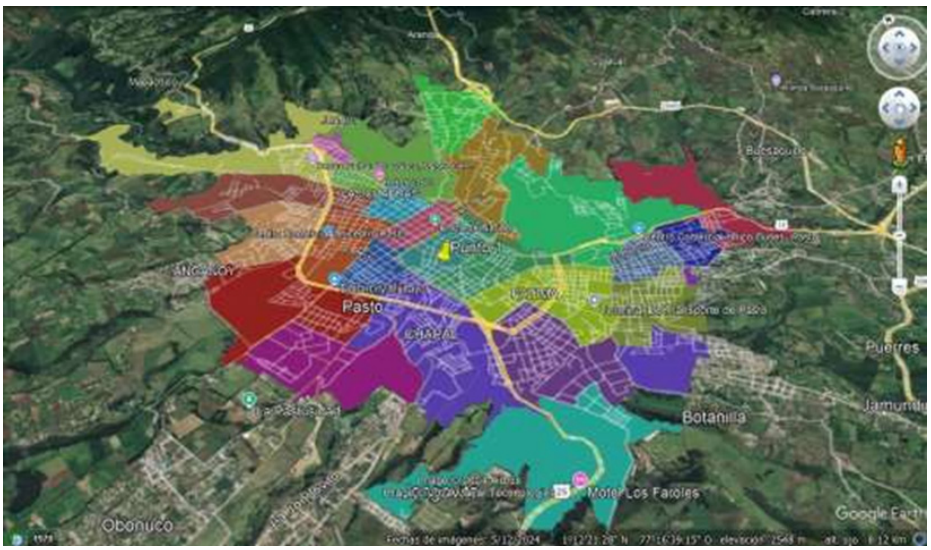
Figure 10. Projection of a point in google earth working equipments



Source: EMPOPASTO

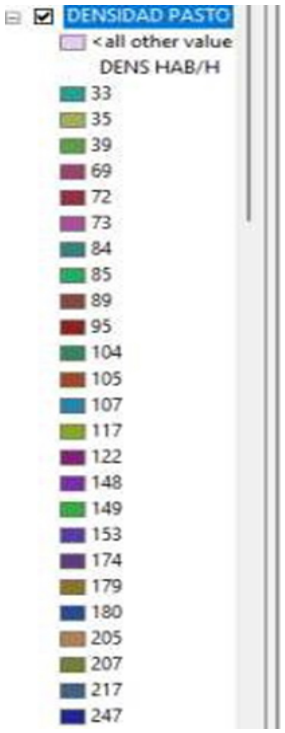
Figure 11. Classification of facilities (in meters) according to an identifying color

By placing a point on the map and assigning it a specific color corresponding to the equipment zone, the equipment value (in meters) for that point is determined. This process ensures that the information collected is accurate and adequately reflects the availability of services in the selected area. This action allows for the precise selection of the corresponding space, facilitating the diligence of the value in the "equipment" column of the repository. In this way, a clear organization of the information is ensured, allowing a more complete analysis of the availability and distribution of services and resources in the territory.



Source: EMPOPASTO

Figure 12. Projection of a point in google earth working density



Source: EMPOPASTO

Figure 13. Classification of densities (hab/Hec) according to an identifying color

By locating a point on the map and assigning it a specific color related to population density, the concentration of inhabitants in that area and the density value for that point are clearly defined. This action allows the corresponding value to be accurately recorded in the "population_density" column of the repository. This ensures that the information accurately reflects the demographic distribution, facilitating a deeper analysis of the social and urban dynamics of the territory.

Validation of the collection by means of a survey of an expert in construction

Validation of the data collection through the survey of a construction expert is a crucial step to ensure the quality and relevance of the information obtained. By consulting a professional with experience in the sector, the aim is to corroborate the accuracy of the data collected and to obtain valuable perspectives that enrich the analysis. This validation not only reinforces the credibility of the research, but also ensures that the decisions and recommendations derived from the study are based on technical and practical knowledge. To view the interview uploaded to YouTube click on the link (Link Video), (Access to the interview script).

Answers obtained during the interview

Block 1: Professional Experience and Urban Growth

1. Could you start by telling me a little about your professional experience in the area of construction and installation in Pasto?

The expert is an electrical technician with extensive experience in construction, and has had the opportunity to closely observe the progress of Pasto's construction and society throughout his career.

2. From your experience, how would you describe the urban growth of the city of Pasto since 2000?

Describe the urban growth in Pasto as accelerated, driven largely by population growth and migratory factors. Urban expansion has been remarkable compared to previous years, and the city has had to adapt quickly to this pressure.

3. What do you think are the main factors that have driven this growth?

Among the factors mentioned are migratory displacement (including the impact of Venezuelan immigration), regional internal displacement, and the search for better economic opportunities in Pasto, especially by inhabitants of rural areas.

4. What would you say are the sectors or neighborhoods of the city that have grown the most in the last 20 years?

The expert points out several key areas:

- Surroundings of the San Pedro Hospital and the areas near the stadium.
- Southeastern areas, including sectors such as Santa Monica.
- Commune 10, including sectors such as Nuevo Sol, Nueva Aranda and San Diego, which have experienced significant growth in construction.

Block 2: Neighborhoods, Residential Complexes and Development in Pasto

1. Can you tell me about the development of neighborhoods, residential complexes or condominiums in Pasto since 2000?

The expert validates the classification work that has been done in terms of residential complexes and condominiums. He emphasizes that the use of advanced technology and software has allowed for a detailed and accurate analysis of housing development.

2. Could you confirm if the classification of housing, complexes and buildings that we have worked on corresponds correctly to constructions after the year 2000?

Yes, the expert confirms that the classification of dwellings, complexes and buildings is appropriate and accurately reflects the characteristics of constructions built after 2000.

He emphasizes that a thorough analysis has been carried out, supported by advanced technological tools and specialized software, which allow validating the accuracy of the classification and the data collected. In addition, he mentions that the process has considered the evolution in the typology of buildings, which allows the analysis to be adjusted to the specific characteristics of new urban developments in Pasto and helps to better visualize the differences in housing growth in the city. This combination of technology and a detailed approach ensures that the data accurately represent recent urban development in Pasto.

3. From your experience, have you noticed any significant changes in the type of housing or residential complexes being built in Pasto since 2000 compared to before?

There has been a notable change in housing. There is a diversification of socioeconomic strata in new constructions, adapting to market demands, where there are differences in middle, low and high strata, and the proliferation of commercial and residential buildings.

4. Do you consider that the data we have used accurately reflect the urban growth in Pasto after 2000?

Yes, I consider that the data reflect the growth well, observable in the amount of new constructions for both housing and commerce in the city.

5. Do you think that the city of Pasto will tend to grow more in buildings, complexes or condominiums?

The expert anticipates a growth in horizontal property due to the limitation of available land. He mentions that previously larger houses predominated, but nowadays new constructions are oriented towards more compact horizontal buildings and developments.

Development of an optimal web application with quality standards

After an intensive period of development and testing, a web application was successfully created that meets high quality standards and is capable of managing information related to users, roles, user permissions, reports and residential projects. The application, in addition to its management functions, allows displaying the results of the predictions made by the predictive model, created to analyze the urban growth of the city of Pasto.

The web application was designed to be intuitive and easy to use, allowing registered users to access key information in a secure and efficient manner. Among its functionalities, in order to fulfill objective 2 of the research are:

1. Visualize the information of the urban planning repository and topics related to urban growth (home page).
2. Access to the system through user authentication.
3. Assigning roles within the system.
4. User management within the system.
5. Management of each user's personal information.
6. Manage the urban repository.
7. Visualize the information of the urban planning repository.
8. Develop and interact with a Chat Bot in the web application.
9. Implement an API.
10. Send relevant information to make urban growth predictions.
11. Saving the predictions made by the predictive model.
12. Search for the predictions made by the predictive model.
13. Visualize the predictions made by the predictive model.

For the development of the web application, the Scrum methodology was used, implementing a total of four sprints. During this process, a product backlog was created to organize the user stories, which allowed structuring and prioritizing the tasks effectively. This organization facilitated the agile development of the application, ensuring that objectives were met and functionalities were delivered continuously.

Planning Sprints 1,2,3,4

	Sprint 1	Sprint 2	Sprint 3	Sprint 4
FECHA INICIO	12/02/2024	26/02/2024	11/03/2024	1/04/2024
FECHA FIN	26/02/2024	11/03/2024	1/04/2024	10/05/2024

Figura 14. Start and End Dates Sprints 1,2,3,4

Product Backlog developed

ID	*Como…*	*Necesito…*	*Para…*	*Prioridad*	*Sprint*	*Estado*
HU-1	Super Administrador, Adminstrador,Usuario	Acceder al sistema mediante la autenticación de usuarios	Poder registrarse	Should	1	**Terminado**
HU-2	Super Administrador, Adminstrador,Usuario	Acceder al sistema mediante la autenticación de usuarios	Poder Iniciar sesión	Should	1	**Terminado**
HU-3	Super Administrador, Adminstrador,Usuario	Recuperar la contraseña en caso de olvidarla o perderla	Ingresar al sistema	Should	1	**Terminado**
HU-4	Super Administrador, Adminstrador,Usuario	Visualizar y gestionar la información personal	Controlar y organizar los datos personales de cada usuario del sistema	Could	1	**Terminado**
HU-5	Super Administrador, Administrador, Usuario	Descargar gráficos de barras de la información urbanística	Generar informes y proporcionar una representación visual de los datos urbanísticos para su análisis y presentación	Could	1	**Terminado**

HU-6	Super Administrador, Adminstrador,Usuario	Realizar predicciones de crecimiento urbano de la ciudad de Pasto en un determinado año mediante un Chat Bot	Obtener gráficos de crecimiento urbano de la ciudad en un determinado año en especifico	Must	2	**Terminado**
HU-7	Super Administrador, Adminstrador,Usuario	Realizar predicciones de crecimiento urbano por zonas de la ciudad mediante un Chat Bot	- Ver un panorama de crecimiento urbano en diferentes zonas de Pasto en un mapa interactivo	Must	2	**Terminado**
HU-8	Super Administrador, Adminstrador,Usuario	Contar con una función en el chat bot que permita ver la información de los mapas del sistema	El Chat Bot debe contar con una función para ver los mapas del sistema -Mapa de comunas -Mapa de densidades -Mapa de estratos -Mapa de equipamientos -Mapa de puntos **(Opción 3 del Chat Bot)**	Could	2	
HU-9	Super Administrador, Adminstrador,Usuario	Contar con una función en el chat bot que permita salir de las opciones de menú	El Chat Bot debe contar con una función para salir de las opciones de menú **(Opción 4 del Chat Bot)**	Could	2	**Terminado**

HU-10	Super Administrador, Adminstrador,Usuario	Guardar las predicciones de un determinado año en la base de datos del sistema	El sistema debe perimir almacenar las predicciones realizadas en la base de datos para brindar la seguridad de la información	Must	2	**Terminado**
HU-11	Super Administrador, Adminstrador,Usuario	Guardar las predicciones por zonas en la base de datos del sistema	El sistema debe perimir almacenar las predicciones realizadas en la base de datos para brindar la seguridad de la información	Must	2	**Terminado**
HU-12	Super Administrador, Adminstrador,Usuario	Buscar las predicciones de crecimiento urbano de la ciudad almacenadas en el sistema	Encontrar los gráficos de las diferentes predicciones de crecimiento urbano de la ciudad realizadas en el sistema	Should	2	**Terminado**
HU-13	Super Administrador, Administrador, Usuario	Buscar las predicciones de crecimiento urbano por zonas almacenadas en el sistema	Encontrar los gráficos de las diferentes predicciones de crecimiento urbano por zonas realizadas en el sistema	Should	2	**Terminado**

https://doi.org/10.62486/gen202593 ISSN: 3046-4137

HU-14	Super Administrador, Administrador, Usuario	Visualizar las predicciones de crecimiento urbano de la ciudad almacenadas en el sistema	Ver diferentes gráficos de crecimiento tales como: - Gráfico de crecimiento urbano real de Pasto - Gráfico de crecimiento urbano predicho por el modelo predictivo a desarrollar - Gráfico comparativo entre crecimiento urbano real y crecimiento urbano predicho de Pasto - Gráfico que contiene la tabla de predicciones con los porcentajes de crecimiento urbano de Pasto	Should	3	**Terminado**
HU-15	Super Administrador, Administrador, Usuario	Visualizar las predicciones de crecimiento urbano por zonas almacenadas en el sistema	- Ver un gráfico de barras demostrando el crecimiento residencial respecto a ciertas variables urbanísticas - Ver en un mapa interactivo el crecimiento urbano por zonas, mostrando las predicciones almacenadas en la base de datos	Should	3	**Terminado**

HU-16	Super Administrador	Asignar y gestionar los roles dentro del sistema	Poder tener control de la visualización o gestión de la información	Must	3	**Terminado**
HU-17	Administrador	Gestionar usuarios dentro del sistema	- Controlar el acceso a la información	Must	3	**Terminado**
HU-18	Super Administrador	Gestionar la información del repositorio urbanístico mediante los procesos: - Añadir - Editar - Eliminar - Buscar	Controlar y organizar toda la información de las variables urbanísticas del repositorio de datos	Must	4	**Terminado**
HU-19	Super Administrador, Administrador, Usuario	Visualizar la información del repositorio urbanístico	Poder obtener un panorama de la información urbanística de la ciudad de Pasto	Should	4	**Terminado**
HU-20	Super Administrador, Administrador	Realizar informes de la información del sistema, relacionada con cada uno de los módulos (usuarios, proyectos residenciales)	Tener un control de la información urbanística en documentos organizados y estructurados	Should	4	**Terminado**

Figura 15. Product Backlog of Sprints 1, 2, 3, 4

In addition, to facilitate the sending of information from the web application to the predictive model, it was necessary to create a connection bridge. For this reason, an API was developed to enable this communication effectively.

Development and implementation of a functional API

An API (Application Programming Interface) has been successfully implemented using Flask that acts as a connection between the predictive model and the web application. This API allows sending data from the web application to the predictive model, providing a mechanism for users to send urban information through forms in the web application to subsequently receive the predictions of the predictive model and thus visualize them in the web application.

API Features

- Interaction with the Predictive Model: the API can load the predictive model from the web application storage or operating system storage and thus make predictions based on the data it receives from the web application. This allows users to get real-time results.
- Inter-Application Communication: the API uses Flask to create an endpoint that the web application can use to send data. This endpoint is designed to be secure and efficient, allowing seamless communication between the web application and the model.
- Forms Integration: the web application has two forms that allow users to send information from the web application to the model. First, there is a form where a specific year can be entered to make the respective urban growth prediction.

The information provided by users is sent to the API, where the prediction is generated using a predictive model. This prediction is displayed in a graph representing the projected urban growth and the annual variation from 2022 to a specific year. In addition, a statistical graph is stored in the "images" folder of the web application, which illustrates the actual growth, the model's prediction in relation to the actual growth, and a comparison between the two, as well as the growth prediction from 2022 to the year entered by the users.

Secondly, there is a form where you can enter certain information related to different urban variables and thus make predictions to identify areas with higher or lower urban growth evaluated from a growth index.

The information provided is sent to the API, where the model makes the prediction based on the data entered, including type of housing, number of dwellings, morphology, price, population density, type of subsidy, longitude and latitude. These data, sent from the web application, are analyzed by the model to generate a prediction. As a result, a geospatial point is created on an interactive map that represents a growth index based on the urbanistic variables entered. This provides an overview of the areas of Pasto with the highest and lowest urban growth.

- Response and Error Handling: the API is configured to handle responses properly and report errors if something goes wrong. This ensures that users get clear and useful responses when interacting with the web application.

Development and implementation of the UrbanPredictor (ChatBot) assistant tool

UrbanPredictor is a chatbot designed to improve the user experience in the JMM URBAN VISION web application, facilitating the sending of data to the API in an efficient way. This assistant has several functions that allow to optimize the interaction and the collection of relevant information for urban prediction.

1. Sending Prediction Data by years:
    - Select Option 1: sending data to make a prediction of urban growth in a given year.
    - Fill in the prediction ID.
    - Fill in the prediction year.

2. Sending of forecast data by zones:
    - Select Option 2: sending data to make a prediction of urban growth by zones in the city of Pasto.
    - Fill in the project code.
    - Fill in the type of housing.
    - Completion of the morphology.
    - Completion of the housing subsidy.
    - Completion of the size of the dwelling.
    - Completion of the price of the house.
    - Fill out the population density.
    - Fill in the longitude and latitude of the residential project.

Test Cases for the Web Application, API and ChatBot

Extensive test cases were conducted for the web application, the API and the chatbot, ensuring that all functions operate as expected. These test cases allowed verifying that each of the functionalities described in the user stories and in the Product Backlog are correctly implemented, thus guaranteeing the quality and efficiency of the system. This process not only validates the correct functioning of the tools, but also ensures an optimal experience for the end user, (Access to test cases).

Validation in the use of the web application by experts.

For the validation of the web application, a survey process was carried out in order to evaluate its functionality. The surveys highlighted aspects such as comprehension and ease of use, as well as the appreciation

of the design, colors, interfaces and typography. Also, the performance of the application was inquired about, evaluating if it was agile and fast for the respondents, who are experts in the area of computer science and have studied or are studying careers related to this discipline.

Results obtained from the surveys

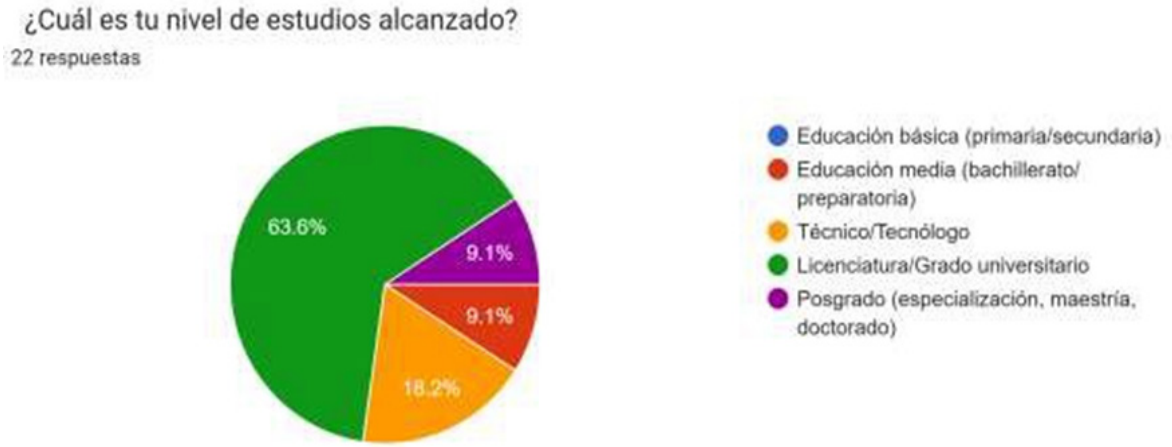


Figure 16. Pie chart of the respondents' level of studies

Figure 17. Pie chart of respondents' years of experience in the use of applications in general

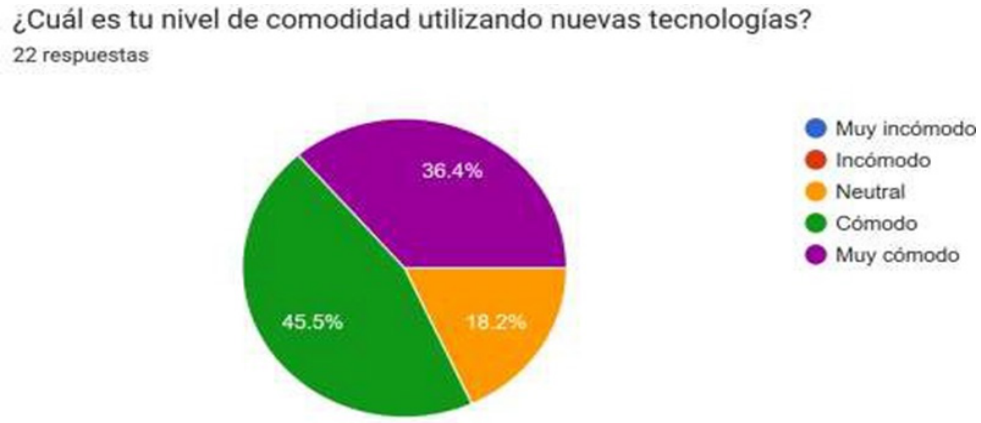


Figure 18. Pie chart of the level of comfort in using new technologies by respondents

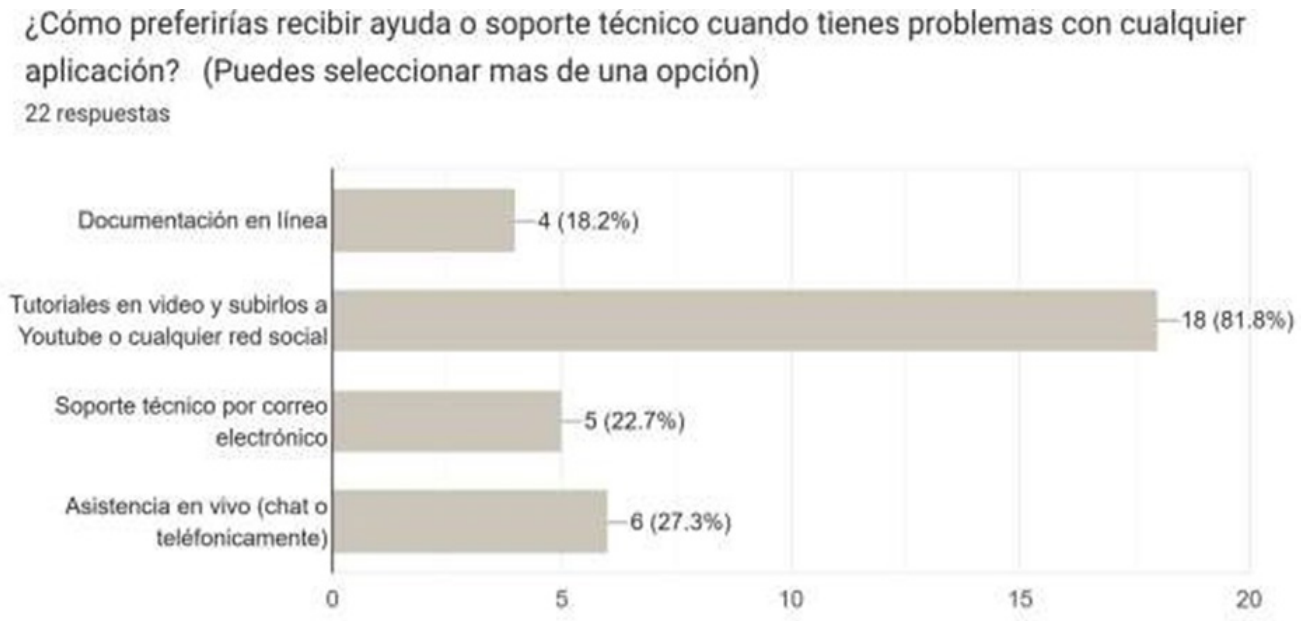


Figure 19. Pie chart of respondents' preference of communication channel for receiving support

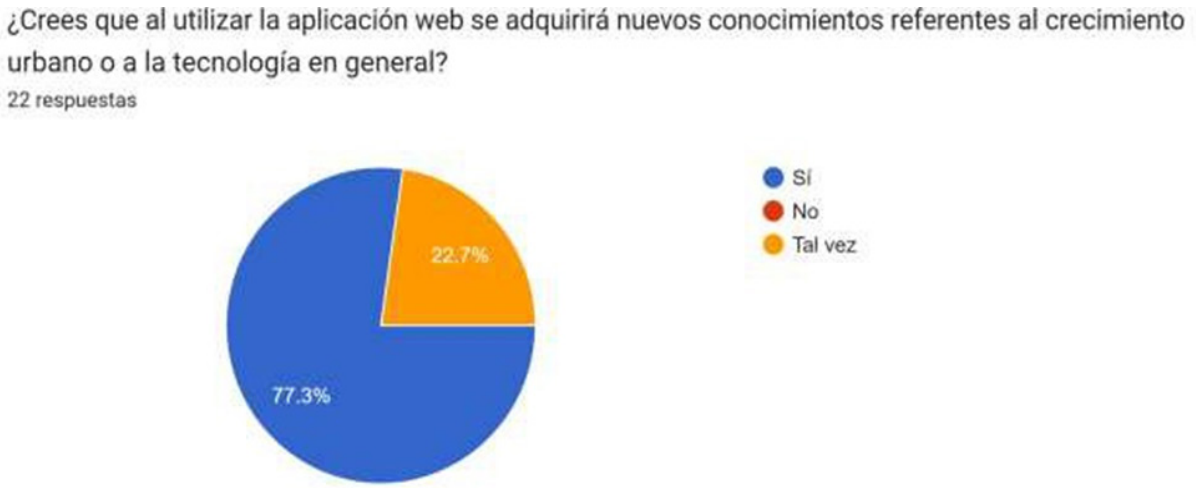


Figure 20. Pie chart to identify the main motivation in the use of the web application by the respondents

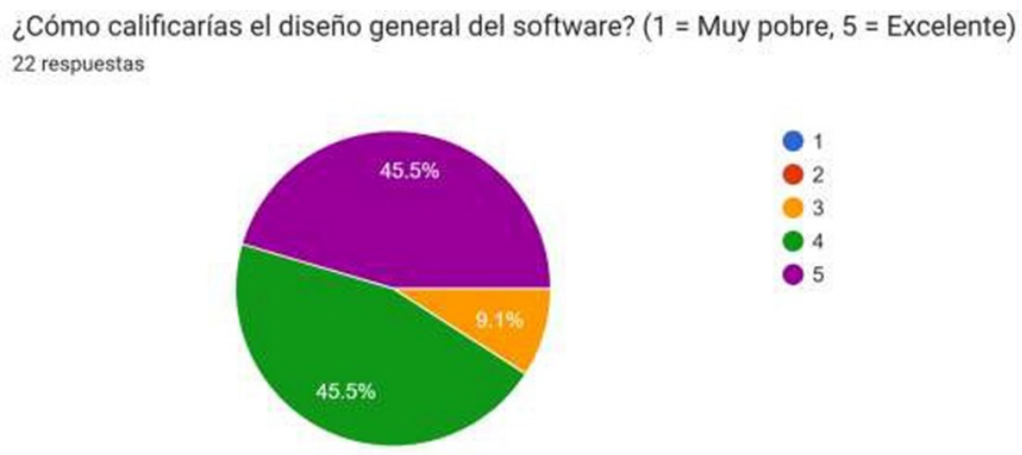


Figure 21. Pie chart to identify the likelihood that respondents will acquire new knowledge regarding urban growth

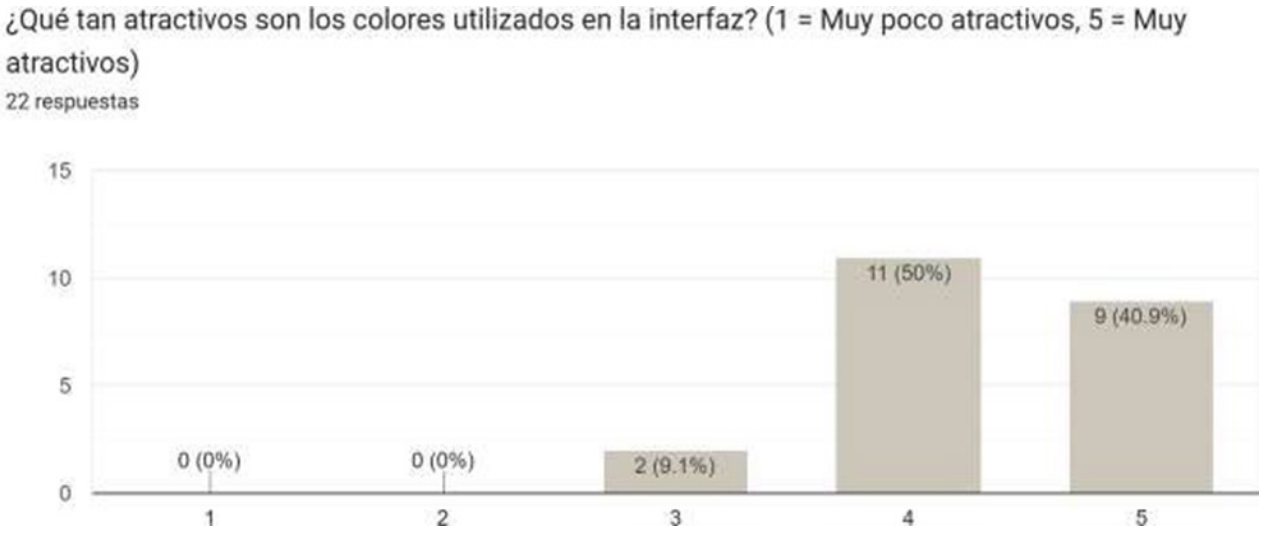


Figure 22. Respondents' rating of the overall design of the web application

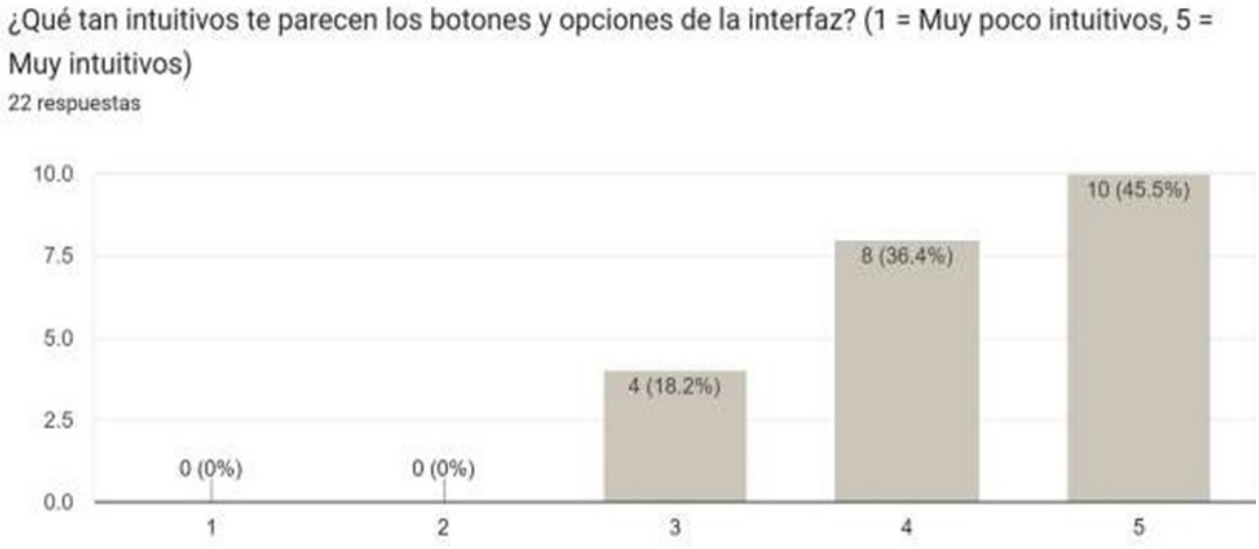


Figure 23. Bar chart rating the level of attractiveness of the colors of the web application by respondents

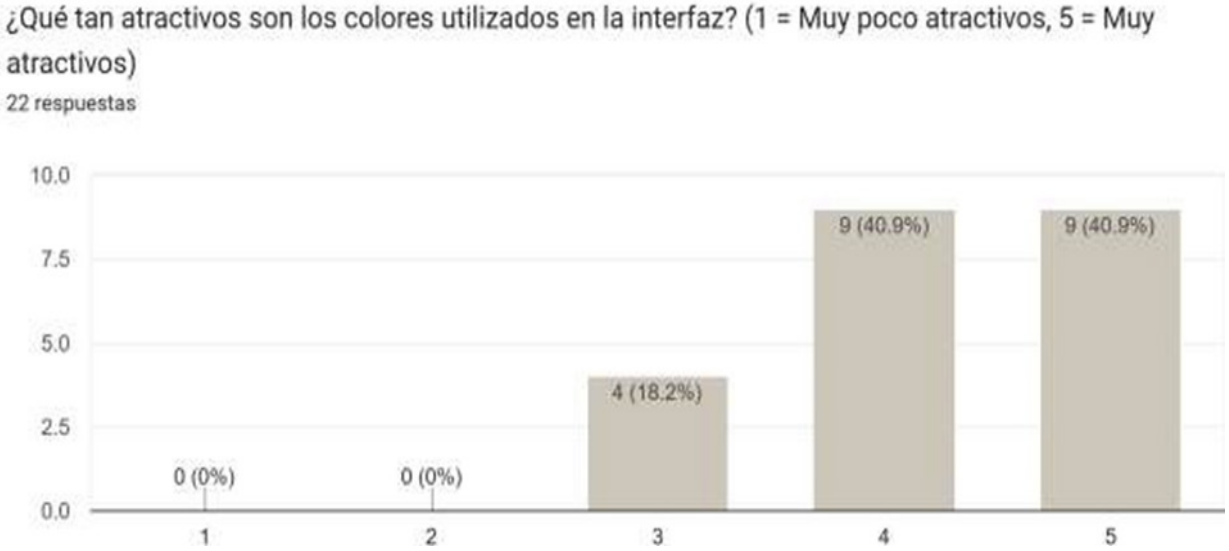


Figure 24. Rating of the level of attractiveness of the buttons and options menu by the respondents

https://doi.org/10.62486/gen202593 ISSN: 3046-4137

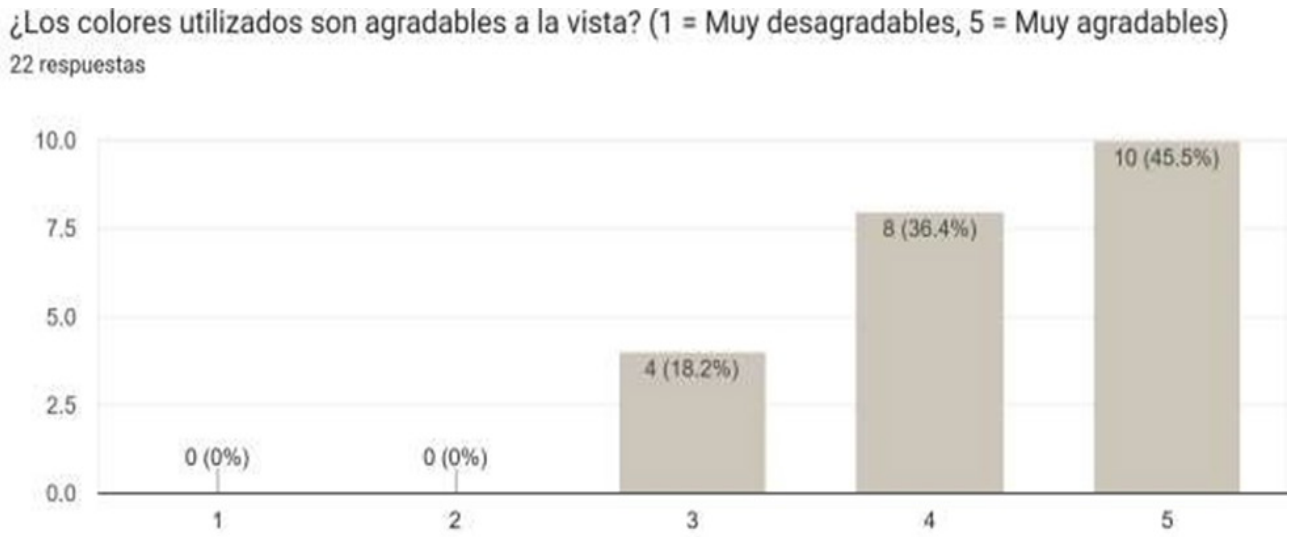


Figure 25. Rating of the level of attractiveness of the colors of the web application interfaces

Figure 26. Rating of the colors of the web application with respect to whether they are pleasing or not to the eye

Figure 27. Rating of the typography of the web application by the respondents

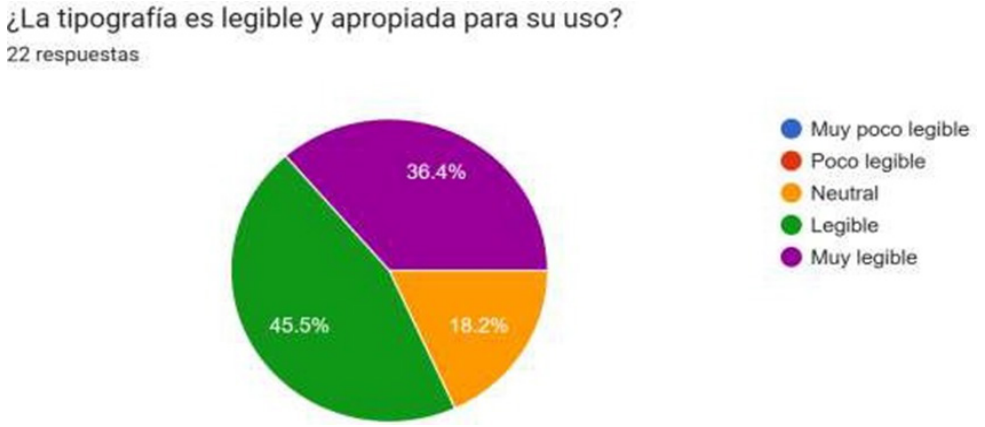


Figure 28. Determination of the readability of the typography of the web application

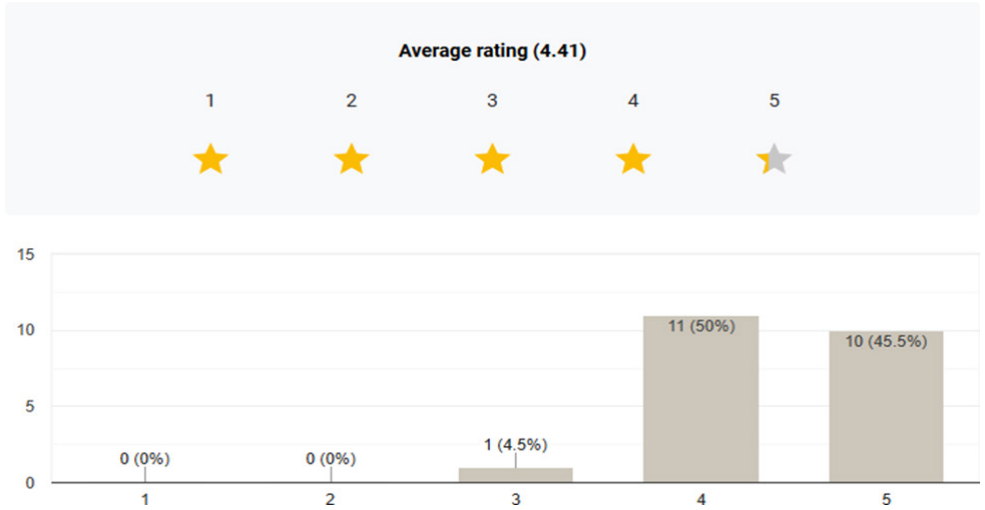


Figure 29. Rating of the user-friendliness of the web application

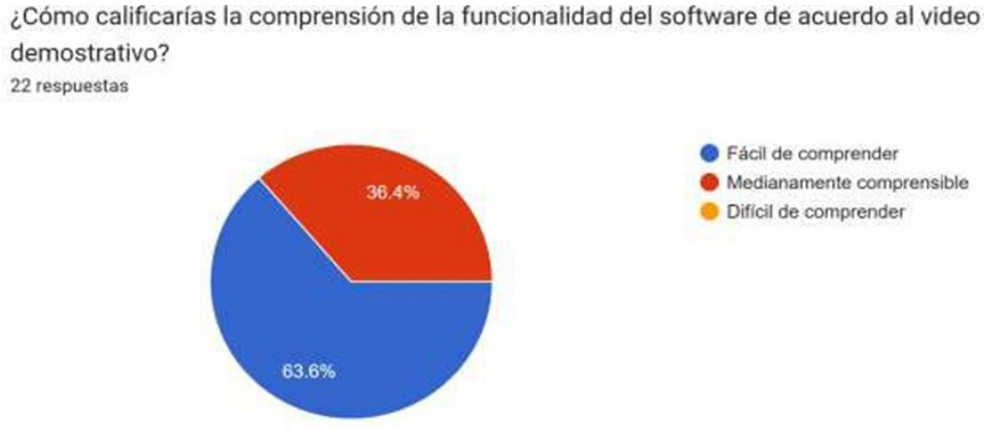


Figure 30. Rating of the level of understanding of the functionality of the web application

Development of an efficient predictive model

After a rigorous and successful development process, the creation of a predictive model capable of making predictions on the urban growth of the city of Pasto for the period 2000-2022, as well as for future scenarios, was completed. This model was built with the additional purpose of finding patterns and trends in the data collected, thus making it possible to foresee the possible urban development of the city in the coming years.

The model is based on a robust set of historical data on residential projects, land use, infrastructure and other key factors related to urban growth in Pasto. Using advanced machine learning techniques and statistical analysis, the model was trained to identify significant relationships between these variables. This provides a basis for predicting the urban growth of the city in the coming years, as well as the areas with more or less growth.

The effectiveness of the model was tested using metrics such as the coefficient of determination (R^2) and the mean square error (MSE), which indicate the model's performance with respect to the data and the accuracy of its predictions, respectively. During the evaluation process, the model achieved an R^2 of 0,96 and MSE of 0,003, demonstrating a strong correlation between the model's predictions and the actual data. The MSE remained within an acceptable range, indicating that the prediction errors were low and that the model could be effectively generalized to future data, furthermore, in the cross validation process good results were obtained, with an R^2 coefficient of 0,94 and an MSE of 0,004 demonstrating an efficient performance of the model with new urban data of the city of Pasto.

For the development of the predictive model, the CRISP-DM (Cross- Industry Standard Process for Data Mining) methodology was used as a reference framework. This methodology provided a structured and systematic approach covering all phases of the process, from business understanding and data preparation to modeling and evaluation. By following the CRISP-DM steps, it ensured that each stage of development was rigorous and aligned with the project objectives, thus facilitating the generation of a robust and effective predictive model.

Breakdown of the phases or stages considered for the development of the predictive model:

1. Business Understanding
    - Research Objectives.
        a) General objective: to determine interactive scenarios of urban growth in the city of Pasto by means of a software product supported by a predictive model.
        b) Specific objectives:
            - Generate an information repository of interactive urban growth scenarios in the city of Pasto.
            - To develop a software that allows managing the information of the different dynamic scenarios of urban growth in the city of Pasto, as well as the information obtained during the research process.
            - Apply the predictive model using data analytics as a simulation module in the software product.
            simulation module in the software product.

    - Definition of the objectives of the model and the problem to be solved.
        a) Analyze and predict the urban growth of the city of Pasto in general.
        b) Analyze and predict urban growth by zones of the city of Pasto.
        c) Definition of the problem question: How has the urban growth of the city of Pasto evolved between 2000 and 2022, evaluated by means of dynamic scenarios and a predictive model supported by a software product?

Note: the creation of the predictive model seeks to solve specific objective 3 of the research.

2. Data Understanding
    - Selection of data collection techniques.
    - Collection of necessary and relevant urban data for the predictive model.
    - Exploration and analysis of the data to obtain information on its quality, structure and content.
    - Identification of problems or limitations in the data, such as missing values or inconsistencies.

Cantidad de viviendas en total	Nombre del proyecto	Tipo de vivienda (BIFAMILIAR)	Tipo de vivienda (MULTIFAMILIAR)	Tipo de vivienda (UNIFAMILIAR)	Tipo de vivienda (VIVIENDA COMERCIAL)	Tipo de vivienda (INTERES SOCIAL)

Estrato vivienda(0)	Estrato vivienda(1)	Estrato vivienda(2)	Estrato vivienda(3)	Estrato vivienda(4)	Estrato vivienda(5)	Estrato vivienda(6)	Estrato vivienda(N/A)	Precio promedio viviendas

Año de inicio 2003	Año de inicio 2005	Año de inicio 2006	Año de inicio 2008	Año de inicio 2009	Año de inicio 2010	Año de inicio 2011	Año de inicio 2012	Año de inicio 2013

Año de inicio 2014	Año de inicio 2015	Año de inicio 2016	Año de inicio 2017	Año de inicio 2018	Año de inicio 2019	Año de inicio 2020	Año de inicio 2021	Año de inicio 2022

Figure 31. Headings of the urban repository grouped by neighborhoods

3. Data Preparation
    - Cleaning and preprocessing of the data to make them suitable for analysis.
    - Transformation of the data as needed, applying MinMaxScaler() as the scaling process, in addition to Flatten for the data flattening process.
    - Creation of training and test data sets.

```python
# Lectura del archivo excel
df1 = pd.read_excel(file_path)

# Eliminar filas con valores nulos del repositorio que esta en el archivo
df = df1.dropna()
```

Python

Figure 32. Cleaning of null values from the urban repository

```python
scaler = MinMaxScaler()  # Escalar en el rango [0, 1]
datos_num = scaler.fit_transform(datos)
```

Python

Figure 33. Scaling of data from the urban planning repository.

```python
# Convertir `y` a 1D usando `flatten`
y_1d = y.flatten()
```

Python

Figure 34. Flattening of the target variable

```python
# Separacion de los datos en datos de entrenamiento y de prueba
X_train1, X_test1, y_train1, y_test1 = train_test_split(X_scaled1, y_1d, test_size=0.2, random_state=42)
```

Python

Figure 35. Separation of the urban planning repository data into training and test data

4. Modeling
    - Selection of the appropriate machine learning techniques and algorithm for modeling.
    - Training of the model and adjustment of the hyperparameters to improve its performance.
    - Validation of the model using the test data set.

Modelo de aprendizaje automático	r2	MSE
DecisionTreeRegressor()	0,965	0,003
RandomForestRegressor()	0,908	0,990
LinearRegression()	0,198	8,673
Ridge()	0,200	8,656
SVR()	0,424	6,227
Lasso()	-0,005	10,877
KNeighborsRegressor()	0,916	0,899

Figure 36. Comparison of regression machine learning models

https://doi.org/10.62486/gen202593 ISSN: 3046-4137

Note: before the creation of the model, a comparison of different machine learning models was performed to select the most suitable one for the current research. This process was carried out by comparing their r2 and MSE coefficients as shown in figure 36, so that the closer r2 (coefficient of determination) is to 1 and MSE (average of the squared differences between the predicted and actual values) is the lowest possible value, the better model will be identified. Therefore; DesicionTreeRegressor () is the best fit to the above conditions.

```python
modelo_arbol = DecisionTreeRegressor(
    random_state=42,
    min_samples_leaf=1,  # Mínimo 1 muestra por hoja
    min_samples_split=2  # Mínimo 2 muestras para dividir un nodo
)
```

Python

Figure 37. Creation of the model (regression decision tree)

```python
modelo_arbol.fit(X_train1, y_train1)
```

Python

```
DecisionTreeRegressor
DecisionTreeRegressor(random_state=42)
```

Figure 38. Training of the model (regression decision tree)

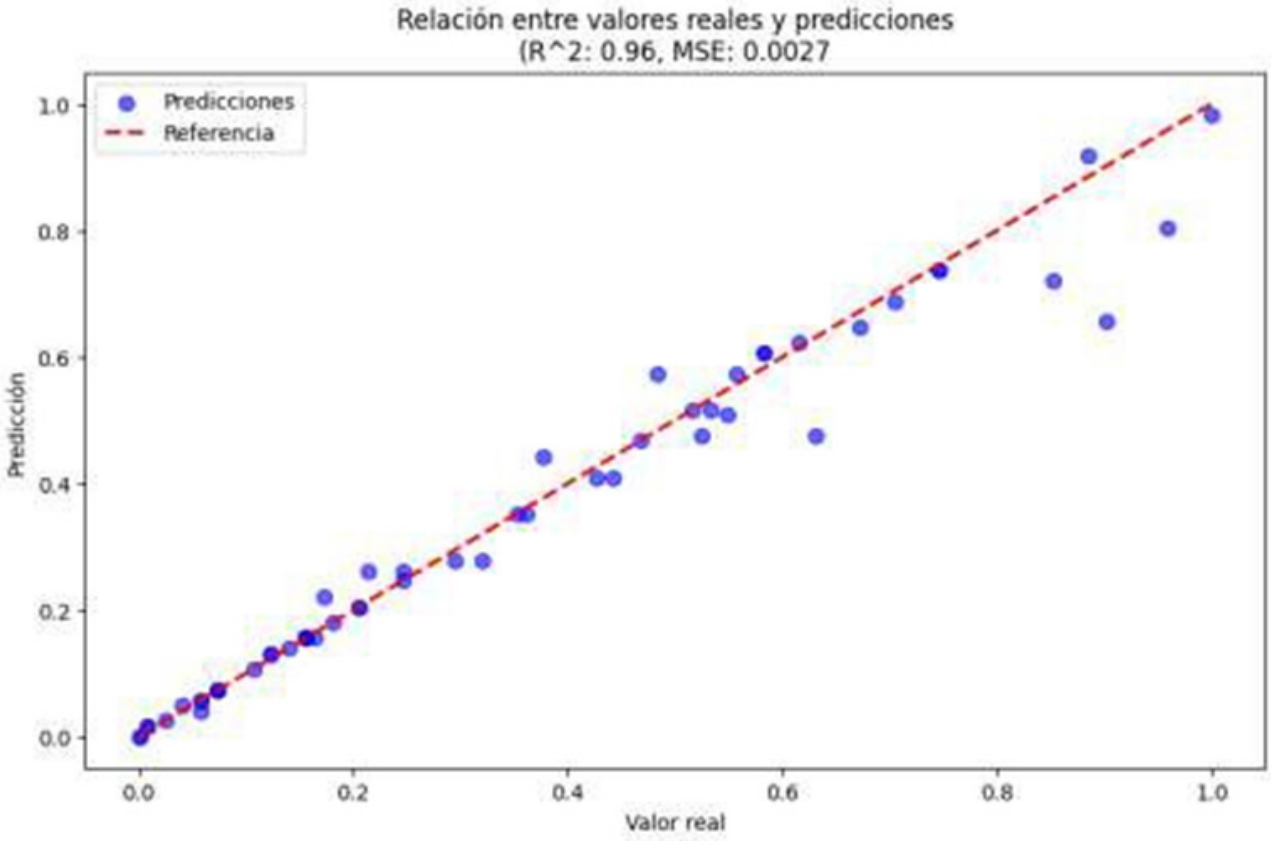


Figure 39. Model validation (regression decision tree)

5. Evaluation
    - Evaluation of the model's performance by verifying whether it meets the business objectives.
    - Performing additional tests to ensure that the model is robust and generalizable.

```
# Realizar validación cruzada con 9 pliegues para calcular el MSE
cv_scores = cross_val_score(modelo_arbol, X_scaled1, y_1d, cv=6, scoring='neg_mean_squared_error')

# Convertir a valores positivos para obtener el MSE real
cv_scores = -cv_scores  # Convierte el MSE a positivo

# Imprimir resultados del MSE
print("MSE Scores:", cv_scores)  # MSE para cada pliegue
print("Mean MSE:", np.mean(cv_scores))  # Promedio de los MSE
```

✓ 0.0s Python

```
MSE Scores: [0.00424533 0.00126081 0.00027446 0.00047316 0.00800946 0.005312  ]
Mean MSE: 0.0032625382151522607
```

Figure 40. Cross-validation of the model's MSE (regression decision tree)

```
# Realizar validación cruzada con 9 pliegues para calcular el MSE
cv_scores = cross_val_score(modelo_arbol, X_scaled1, y_1d, cv=6, scoring='r2')

# Convertir a valores positivos para obtener el MSE real
cv_scores = cv_scores  # Convierte el MSE a positivo

# Imprimir resultados del MSE
print("MSE R2:", cv_scores)  # MSE para cada pliegue
print("Mean R2:", np.mean(cv_scores))  # Promedio de los MSE
```

✓ 0.0s Python

```
MSE R2: [0.92234224 0.97723786 0.99597407 0.98494665 0.89872006 0.90938974]
Mean R2: 0.9481017704147909
```

Figure 41. Cross-validation of the model's r2 (regression decision tree)

6. Deployment
   - Deployment of the model in a runtime environment.
   - Review of the entire prediction process to ensure that the model is suitable for deployment.
   - Interpretation of the model results.

```
 * Serving Flask app '__main__'
 * Debug mode: on
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
Press CTRL+C to quit
```

Figure 42. Server where the model is executed (regression decision tree)

7. Support of a teacher expert in artificial intelligence for the development of the model
   - The expert teacher Héctor Andrés Mora Paz has extensive experience in the field of artificial intelligence and machine learning, supported by his experience in data analysis projects and development of predictive models. His knowledge in advanced AI methodologies has been fundamental to guide the creation of the model, providing accurate strategies and key optimizations to improve the accuracy and efficiency of the predictions.
   - It is important to consider the expert faculty member's resume, where he details his professional experience and achievements in the area of artificial intelligence supporting his research support (Link resume).

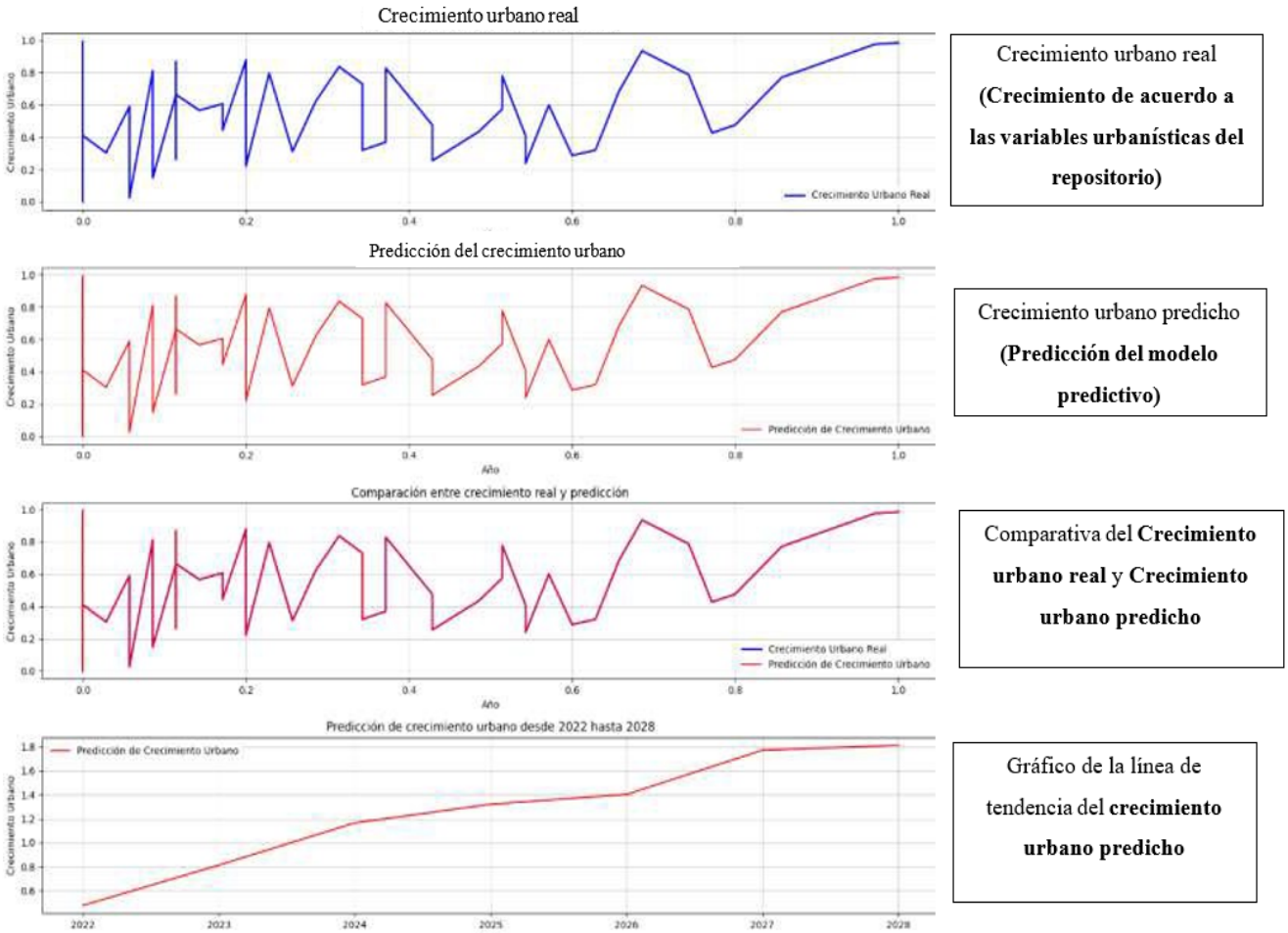


Figure 43. Graph of actual urban growth and predicted urban growth generated by the predictive model

Additional research and achievements

Throughout the development process of the project, it has obtained important recognitions in regional expositions, highlighting the participation in the Cesmag University and in research seedbeds. The project has also participated in national and international congresses, such as the Safaris Tech International Congress. Among the outstanding achievements is a first place in the call for research projects in 2022. In addition, second place was achieved in the Regional Seedbed Conference, thus qualifying for participation in Ibarra, Ecuador, as well as third place in the National Seedbed Congress held at the Cesmag University. Likewise, two research articles have been developed documenting the development of the project, (Access to certifications and achievements), (Access to developed articles).

Installation, requirements and configuration manuals for the web application, api, chat bot and predictive model:

- User manual.
- Installation manual.
- Cataloging sheet.
- Technical document of requirements.

Analysis of results

The analysis of dynamic scenarios related to the urban growth of the city of Pasto has been carried out through the study of several urban variables. These variables, which provide a comprehensive framework for understanding the dynamics of development in different areas, will be represented through statistical graphs, which will allow to effectively visualize the trends and relationships between the factors that influence urban growth.

Breakdown of urbanistic variables and tabular and/or statistical results

Project Name: identification of each residential project, fundamental to reference and analyze the number of records referring to projects built between the years 2000 and 2022.

Table 1. Total number of residential projects in the city of Pasto	
Total number of residential projects built between 2000 and 2022	17 047

The above table presents the total number of projects to be analyzed, also considering the urban variables obtained during the research process. This approach allows for a comprehensive study of the urban growth of the city of Pasto between 2000 and 2022. It is worth mentioning that this value is obtained through the query "SELECT count(`project_name`) FROM `locations`;" in the urbanistic repository managed in phpmyadmin and executed in port 80 of the apache server.

Housing Type: characterization referring to the architecture of housing, such as residences, buildings, complexes, which helps to understand the diversity in the housing supply and its effect on the development of different areas.

Figure 44. Comparison of Residences, Buildings and Complexes in Pasto

The graph shows a comparison in the total number of residential projects, buildings and complexes. Noting that residences have a significantly higher number of projects, followed by complexes and then buildings, the analysis of this distribution could focus on various aspects of urban growth, considering how different types of constructions impact urban development.

In conjunction, considering urban growth, the most appropriate architecture could be a mixed approach between residential complexes and buildings. This is due to the following factors:

- Space Optimization: buildings (with 432 projects) occupy less floor space for the same amount of people compared to single-family residences. This allows for greater population density and optimizes land use, reducing the horizontal expansion of the city. Residential complexes (3367 projects) allow maintaining a low-rise environment, but with the advantage of grouping several dwellings, creating mini-communities and improving land use efficiency compared to single-family residences.
- Sustainability and Environmental Preservation: a mixed approach, favoring clusters and buildings in urban areas, can help reduce uncontrolled urban sprawl into rural or ecological reserve areas. This is key to sustainability, as it reduces the need for new transportation, water and energy infrastructure in areas far from the urban center.
- Access to Infrastructure and Services: by having higher population density in less space, as in the case of buildings and complexes, the city can concentrate and improve public services (transportation, education, health) in a smaller area, which reduces costs and improves access to infrastructure.
- Flexibility and Adaptability to Growth: residential clusters offer a middle ground: they allow a certain level of density without giving up a low-rise design, more suitable in neighborhoods or areas where infrastructure does not yet support tall buildings. In areas with high demand and good infrastructure, high-rise buildings would be the optimal choice, maximizing the capacity of the area without the need for horizontal expansion. Thus, using phpmyadmin and the query, "SELECT DISTINCT(`id_neighborhood`) FROM `locations` where `housing_type`='Building' ORDER BY `locations`.`id_neighborhood` ASC;" we get the list of neighborhoods that meet the first condition where the residential project is a building. Now; to know which are the neighborhoods or projects that belong to a set we perform the following query "SELECT DISTINCT(`id_neighborhood`) FROM `locations` where `house_type`='Set' ORDER BY `locations`.`id_neighborhood` ASC;", thus obtaining the neighborhoods that meet the second condition.

Now to be clear about the neighborhoods that have had an impact on their urban growth, considering the type of housing of each project we have:

- First, we have the list of all the neighborhoods with their respective id that have been obtained during the research.
- Secondly, we have the lists of the neighborhoods obtained in the previous consultations considering the type of housing variable.

Risk or Hazard Zone: evaluation of areas susceptible to natural disasters, essential to identify limitations in urban growth and plan safely.

Figure 45. Number of projects considering their risk zone

The graph shows the number of projects classified by their exposure to different risk zones: projects without risk zone, landslide hazard, mining collapse, and flood hazard. The final analysis for urban growth considering these data could focus on the following aspects:

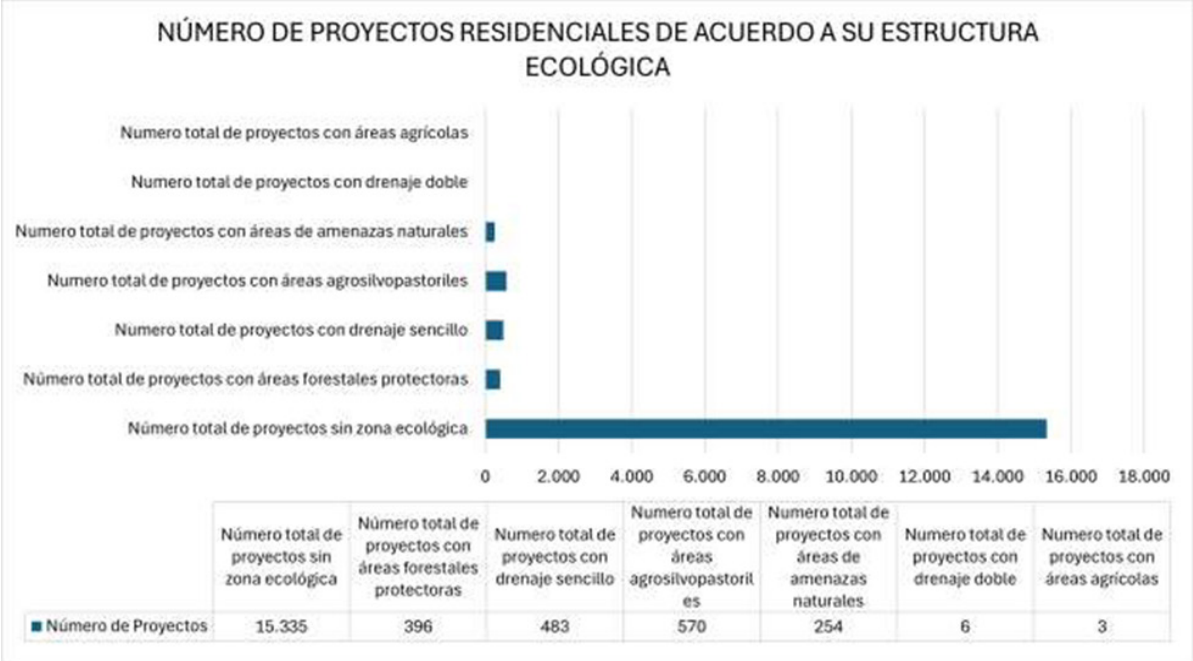
- Prioritization of Safe Zones: with an overwhelming number of projects (14 688) that are not at risk, it is clear that most of the constructions are in safe zones. This is positive for urban development, as projects in low-risk areas tend to be more sustainable and require fewer interventions to mitigate environmental risks.
- Zoning and Land Use Policy Consideration: areas at risk of landslide, mining collapse, or flooding represent a small fraction of projects, but it is important that these areas have strict regulations. Zoning policies should prohibit or restrict construction in high-risk areas to avoid future damage and high mitigation costs.
- Mitigation Infrastructure Planning: for projects in risk zones (especially the 1142 threatened by mining and 1057 by flooding), urban growth should integrate mitigation infrastructure (retaining walls, adequate drainage, among others). Construction in these areas should include detailed assessments to reduce the impact of potential disasters.
- Environmental Protection and Sustainable Development: avoiding urban expansion in at-risk areas promotes more sustainable and less invasive urban development, conserving natural areas and minimizing the modification of sensitive ecosystems.
- Incentives for Relocation or Redevelopment of Projects: areas with significant risks, such as mining or landslides, could benefit from incentives to relocate existing projects or retrofit them with safety improvements. This reduces risk exposure and improves the city's resilience to potential disasters.

Now to be clear about the neighborhoods that have had an impact on their urban growth, considering the risk zone of each project we have: using phpmyadmin and the query, "select DISTINCT(`id_neighborhood`) from locations where `risk_zone`='No Risk' ORDER BY `locations`.`id_neighborhood` ASC;" we obtain the list of neighborhoods that meet the condition where residential projects have not been built in an area with any type of natural risk, allowing us to identify the neighborhoods that have had a positive urban growth.

- In the first place, we have the list of all the neighborhoods with their respective id that have been obtained during the research.

- Secondly, there is a list of the neighborhoods obtained through the consultations, considering also the risk zone variable.

Main Ecological Structure: identification of natural elements that sustain the ecosystem, vital to ensure that urban development does not compromise biodiversity.



	Número total de proyectos sin zona ecológica	Número total de proyectos con áreas forestales protectoras	Numero total de proyectos con drenaje sencillo	Numero total de proyectos con áreas agrosilvopastoriles	Numero total de proyectos con áreas de amenazas naturales	Numero total de proyectos con drenaje doble	Numero total de proyectos con áreas agrícolas
Número de Proyectos	15.335	396	483	570	254	6	3

Figure 46. Number of projects considering their ecological structure

The graph shows the number of residential projects classified according to their ecological structure. Most of these projects (15 335) are in areas without ecological structure, while a much smaller number are in areas with ecological characteristics such as protective forest areas, agricultural areas, drainage (single or double), and natural hazards.

Here are the key points for urban growth analysis:

- Predominance of Areas without Ecological Structure: the high number of projects in zones without ecological structure indicates that urban expansion has occurred mostly in areas that do not have significant ecological value. This may facilitate construction without environmental constraints, but could also indicate a loss of balance in the urban-ecological relationship of the city.
- Opportunity Green Infrastructure and Sustainable Development: since very few projects are developed in areas with forest or agricultural protection, urban growth could benefit from integrating green infrastructure in these areas. This could include the creation of ecological corridors, green spaces and buffer areas that promote a healthier urban environment with less environmental impact.
- Mitigation of Environmental Risks in Natural Areas: only 254 projects are in natural hazard areas and 570 in agro-silvopastoral zones. This suggests that, although there is a presence of projects in sensitive areas, the number is low. Nevertheless, construction in these areas should be carefully planned so as not to alter ecosystems, reduce risks and conserve natural resources.
- Consideration of Drainage Structures: the small number of projects with drainage (single or double) suggests a possible lack of drainage infrastructure in new construction. Urban growth should ensure that future projects in areas susceptible to flooding include adequate drainage systems to avoid erosion and water accumulation problems.
- Protection of Forest and Agricultural Areas: with only 396 projects in protective forest areas and 6 in agricultural areas, it appears that urban growth has not severely impacted these areas. However, this also reflects the need to implement stricter conservation policies to protect the remaining ecosystems and avoid urbanization of these areas in the future.

Thus, considering an urban growth with a sustainable approach and focusing on maintaining a balance between expansion and ecological conservation, the neighborhoods that present protective forest areas or agricultural areas are obtained respectively: in the first instance, by performing the following query in the urban repository; “SELECT DISTINCT(`id_barrio`) FROM `locations` WHERE `ecological_structure`=`Protective forest

areas`;" , similarly you have the following query; "SELECT DISTINCT(`id_neighborhood`) FROM `locations` WHERE`ecological_structure`=`agricultural areas`;"

- First, we have the list of all the neighborhoods with their respective ids that have been obtained during the research.
- Secondly, we have the lists of the neighborhoods obtained during the previous queries taking into account the ecological structure variable.

Number of dwellings: total number of housing units, which allows us to analyze the capacity of each area to accommodate population and its growth.

Table 2. Number of housing units in Pasto's residential projects

Total number of dwellings considering each residential project	27 653

The above table showing the total number of dwellings is a key component in the study of urban growth, as it provides essential data for analysis, planning and informed decision making, the value shown is obtained through the query; "SELECT sum(`number_dwellings`) FROM `locations`;", in the urban repository.

Housing Type (Morphology): characteristic focused on the design and shape of the buildings, where they can be categorized as single-family, two-family, and multi-family.

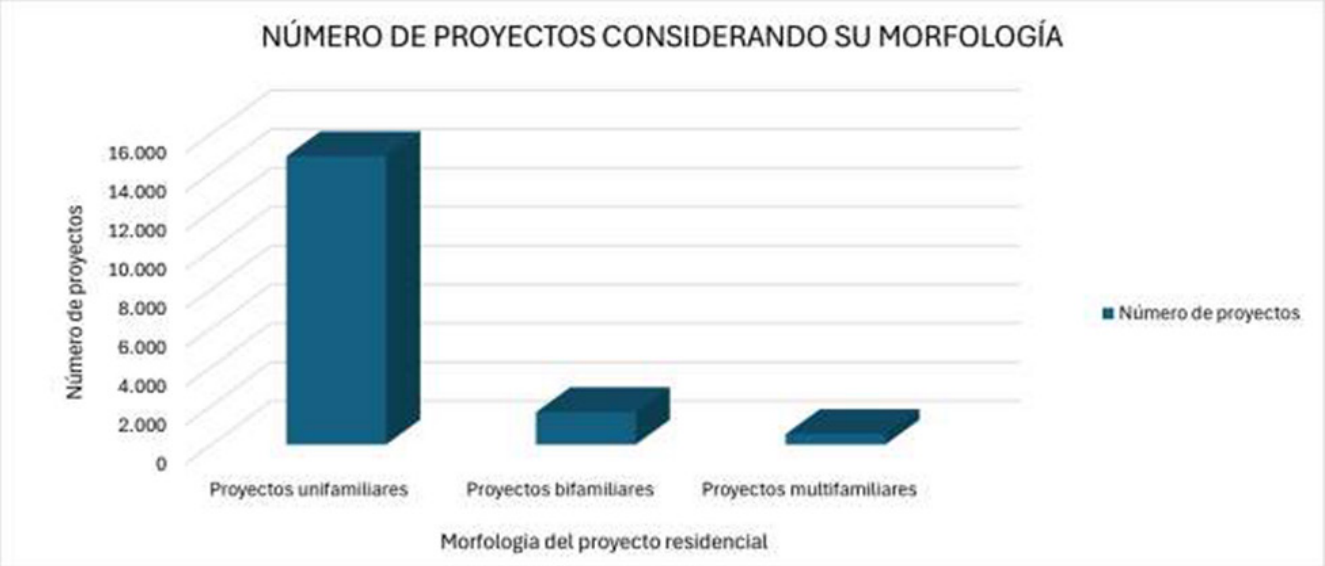


Figure 47. Number of projects considering their morphology

For sustainable and efficient urban growth, the most suitable residential morphology types would be multi-family and two-family projects. Here we explain why these types are beneficial in the context of urban growth:

Multifamily Projects (Apartment Buildings):

- Land Use Efficiency: multifamily projects allow a large number of units to be concentrated in a smaller space, which reduces horizontal sprawl and land consumption.
- Shared infrastructure and services: by being in a single building or complex, infrastructure (water, electricity, transportation, etc.) can be distributed more economically and efficiently.
- Reduced urban costs: by concentrating more people in smaller areas, governments can reduce costs in the construction and maintenance of infrastructure such as roads, lighting, and sanitation systems. Environmental sustainability: multifamily housing buildings reduce the ecological footprint by limiting land and energy use, promoting a compact city model. Two-Family Projects (Duplex or Townhouses).
- Medium Density: two-family projects allow for intermediate density, providing more private space than a multi-family building, but being more efficient in land use than single-family.
- Design flexibility: townhouses or duplexes can accommodate different family needs, allowing for more affordable housing without the need for large tracts of land.
- Cost-sharing: in many areas, two-family projects share walls and other elements, which reduces construction costs and makes more efficient use of resources. Thus, for sustainable urban growth, multifamily projects should be prioritized in central areas or areas with good connectivity, and two-family

projects in peripheral areas or in neighborhoods requiring medium density. This approach optimizes land use, reduces expansion costs, and supports greener and more efficient urbanization.

To obtain the residential projects that meet the above conditions, the following queries have been performed; to obtain the neighborhoods with multifamily projects: "SELECT DISTINCT(`id_neighborhood`) FROM `locations` where`morphology_type`='Multifamily';"; now to obtain the neighborhoods with two-family projects we have: "SELECT DISTINCT(`id_neighborhood`) FROM `locations` where`morphology_type`='Two-family';".

- En first, you have the list of all the neighborhoods with their respective ids that have been obtained during the research.
- En secondly, we have the list of the neighborhoods obtained in the previous queries considering the morphology variable.

Housing Type (Subsidy): classification of housing according to government subsidies and monetary value (COP) of each residential project, which impact housing affordability and may alter market dynamics in certain areas.

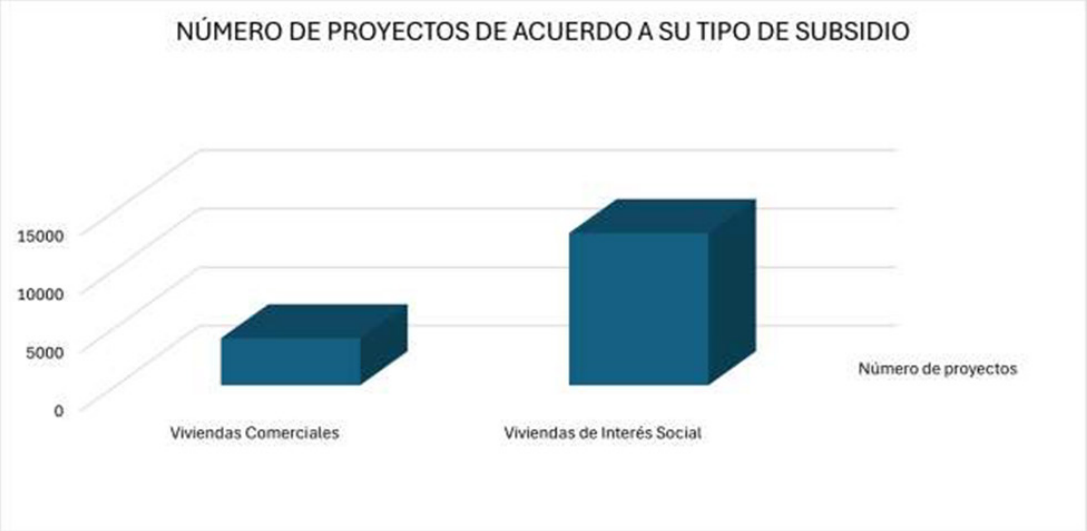


Figure 48. Number of projects considering their type of subsidy

The graph shows that most of the projects are for social interest housing (VIS), while a smaller proportion corresponds to commercial or non-SIV housing. Based on this information, and considering the definitions of VIS and non VIS:

- Boosting access to housing: the focus on SIV housing suggests that urban growth is oriented towards accessibility and support for sectors of the population with fewer resources. This is positive for inclusive urban growth, as it allows more people to access cost-regulated housing.
- Sustainability and urban density: SIVs tend to promote efficient urban land occupation rather than encouraging the horizontal sprawl that characterizes higher-cost single-family or commercial housing. This is key to sustainable urbanization, as it helps to concentrate the population in consolidated urban areas.
- Infrastructure and service challenges: given the high number of SIV projects, it is necessary for cities to invest in infrastructure and services (transportation, education, health, etc.) in the areas where these projects are developed, in order to guarantee the quality of life of their inhabitants and avoid saturation of resources.
- Balance in urban development: although SIVs are important for social inclusion, it is also necessary to balance them with non-SIV housing to diversify urban areas, which can attract investment and promote a socioeconomic mix that benefits the city's development. In this way, for urban growth driven by an inclusive strategy, social interest projects should be prioritized as an appropriate approach to improve access to housing and can be sustainable if accompanied by investments in infrastructure and public services.

To obtain the residential projects that meet the above conditions, the following queries have been performed; to obtain the neighborhoods with social interest projects: "SELECT DISTINCT(`id_neighborhood`) FROM `locations` where `subsidy_type`='Social interest',".

- First, we have the list of all the neighborhoods with their respective id that have been obtained during the investigation.
- Secondly, we have the lists of the neighborhoods obtained from the previous query considering the variable Type of housing according to the subsidy.

Housing stratum: socioeconomic level associated to each housing unit, which allows understanding the demographic profile and the needs of the residents.

Table 3. Number of projects considering their socioeconomic stratum

Stratum	Number of projects
Stratum 0	2465
Stratum 1	5513
Stratum 2	5378
Stratum 3	3033
Stratum 4	531
Stratum 5	102
Stratum 6	25

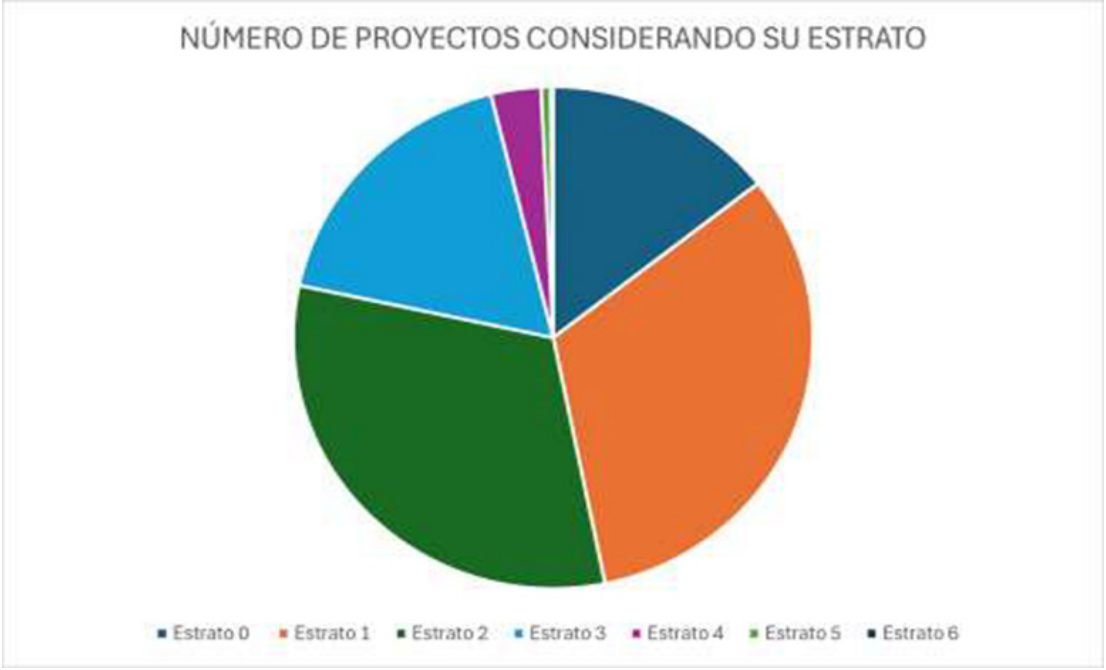


Figure 49. Pie chart of the number of projects considering socioeconomic stratum

The previous graph and tabular data shows the distribution of residential projects according to socioeconomic stratum, indicating how urban growth is being configured according to this factor. In terms of urban growth, the analysis would be as follows:

- Focus on low and middle strata: the largest number of projects is concentrated in strata 1, 2 and 3. This indicates a focus on the development of affordable housing for low and middle income sectors of the population. This type of distribution promotes inclusive urban growth, allowing people of different socioeconomic levels to access housing in urban areas.
- Inequality in distribution by high strata: the high strata (strata 5 and 6) have a lower representation compared to the lower strata. This suggests a lower investment in areas for high-income sectors, which could indicate an urban orientation towards covering basic needs and promoting housing equity. However, this lack of projects in the upper strata could also limit social and economic diversity in certain areas.
- Boosting densification: projects in the lower and middle sectors are generally concentrated in more densely populated or expanding areas, which can promote greater densification in already developed urban areas. This is beneficial for more efficient land use and facilitates the provision of basic services and infrastructure.

- Challenges for infrastructure and services: the concentration of projects in the lower and middle strata implies the need for adequate and sufficient infrastructure in the areas where these projects are developed, such as transportation, education and health, to avoid overloading urban resources and ensure an adequate quality of life.
- Promoting social cohesion: by facilitating access to housing for the lower and middle sectors, urban growth can promote greater social cohesion, integrating different sectors of society into the urban fabric and avoiding the exclusion of certain groups from urban areas.

In this way, the focus on low- and middle-sector projects contributes to inclusive and accessible urban growth. However, it is important to maintain the balance between strata and improve infrastructure in these areas to support sustainable growth and improve the quality of life of their inhabitants.

To obtain the residential projects that meet the above conditions, the following queries have been performed; to obtain the neighborhoods with projects that present stratum 1: "SELECT DISTINCT(`id_neighborhood`) FROM `locations` WHEREhousing_stratum`='1';", for neighborhoods with projects that present stratum 2: "SELECT DISTINCT(`id_barrio`) FROM `locations` WHERE `estrato_vivienda`='2';", for neighborhoods with projects that have stratum 3: "SELECT DISTINCT(`id_barrio`) FROM`locations` WHERE `estrato_vivienda`='3';".

- First, you have the list of all the neighborhoods with their respective id that have been obtained during the investigation.
- Secondly, you have the list of the neighborhoods that were obtained from the previous queries where the variable stratum has been considered.

Housing Size (m^2): living area, important to evaluate the quality of life and the efficient use of space in each area.

Figure 50. Comparison of minimum value of housing size in (m2) vs. maximum value of housing size in (m2)

Housing Price: acquisition cost, which influences demand and may reflect real estate market trends in different areas.

Table 4. Minimum, maximum and average value of Pasto's housing prices (COP)

Minimum value of housing prices in Pasto (COP)	$ 100 000 000
Maximum value of housing prices in Pasto (COP)	$ 950 000 000
Average home price in Pasto (COP)	$ 278 573 062,71

Minimum Value

- SQL query: "SELECT MIN(precio_vivienda) FROM locations;"
- Definition: the lowest price of housing in Pasto, determined at $100 000 000.
- Research Utility: identification of Market Segments which allows detecting the most affordable housing segment, essential to evaluate the supply of housing aimed at the low-income population.

- Affordability Analysis: informs on the availability of affordable housing options and supports the development of policies to promote access to low-cost housing.

Maximum Value

- SQL query: "SELECT MAX(housing_price) FROM locations;"
- Definition: the highest price of housing in Pasto, set at $950 000 000.
- Research Utility: luxury Niche Detection allowing to identify the luxury housing segment, allowing to analyze how this sector impacts urban growth.
- Investment Guidance: provides key data for investors interested in luxury projects, helping to define investment opportunities in the real estate market.

Average Value

- SQL query: "SELECT AVG(housing_price) FROM locations;"
- Definition: the price average of of the homes at Pasto, calculated at $4 748 835 000 000.
- Research Utility: general evaluation of the market allowing to give a broad view of the cost of housing, fundamental to understand the price dynamics in Pasto.
- Project Planning: facilitates the development of projects aligned with the average price and the purchasing capacity of the population, optimizing the housing supply.
- Public Policy: helps policy makers to understand the market, allowing the creation of strategies to improve equity in access to housing.

Start Year: date of construction of the projects, relevant to identify growth patterns over time and their relationship with urban development.

Table 5. Oldest, current and average year of housing construction

Oldest year of housing construction housing (2000-2022)	2003
Most current year of housing construction (2000-2022) 2003 housing construction (2000-2022)	2022

From the previous graph, the maximum year of housing construction can be determined, which is obtained by consulting the database with the following query: "SELECT max(start_year) FROM locations;". This year is relevant for the research project, since it marks the limit of the research process. The urban data collection covered from 2002 to 2022, thus respecting this range. In addition, the minimum year recorded is 2003, obtained through the query: "SELECT min(start_year) FROM locations;". This value also fits the minimum limit of the research range, since 2003 is between 2000 and 2022.

In addition, by focusing on urban planning, a significant impact on the urban growth of a city is generated. For urban planners and decision-makers, this analysis provides information on the periods in which concentrated construction efforts have been concentrated. This information is valuable, as it can guide future investments in infrastructure and public services.

Population Density (inhab/Hec): number of inhabitants per hectare, which helps to analyze the demographic pressure on the development of different areas.

Table 6. Minimum and maximum value of population density

Minimum value of population density in a given area or zone of Pasto (inhab/Hec)	33 inhab/Hec
Maximum value of population density in a given area or zone of Pasto (hab/Hec)	247 inhab/Hec

From the previous statistical table, the minimum and maximum values of population density for the city of Pasto are obtained through the queries respectively; "SELECT min(`population_density`) FROM `locations` where `population_density`!=0;" and "SELECT max(`population_density`) FROM `locations`; highlighting a variation between them indicating a diversity in the occupation of the territory and in the distribution of the population. This variation in population density has several implications for urban growth:

- Planning and Development.

- High density areas may require more intensive urban planning, with investments in infrastructure, transportation and public services to serve a larger population.
- Expansion Zones
- Areas with low density could be seen as potential areas for future growth, which could guide decisions on where to focus development and investment.
- Sustainability
- Variability in density can also influence the sustainability of urban growth. High-density areas should be carefully managed to avoid problems such as congestion, pollution, and lack of adequate services.

Facilities (m): services and facilities available in each area, which are crucial in determining the quality of life and attraction of new residents.

Table 7. Minimum and maximum value of facilities

Minimum value of facilities present in the city of Pasto	0 m
Maximum value of facilities present in the city of Pasto	200 m

From the table above, it can be inferred that if a residential project has a low value in facilities, it will be closer to an urban facility, such as a park, hospital, school, among others. On the other hand, if a residential project has a high equipment value, this indicates that it is farther away from these facilities, which could hinder the quality of life of its residents.

Location of the Residence (Longitude and Latitude): geographic coordinates that allow mapping and analyzing how the projects are distributed in the territory of Pasto.

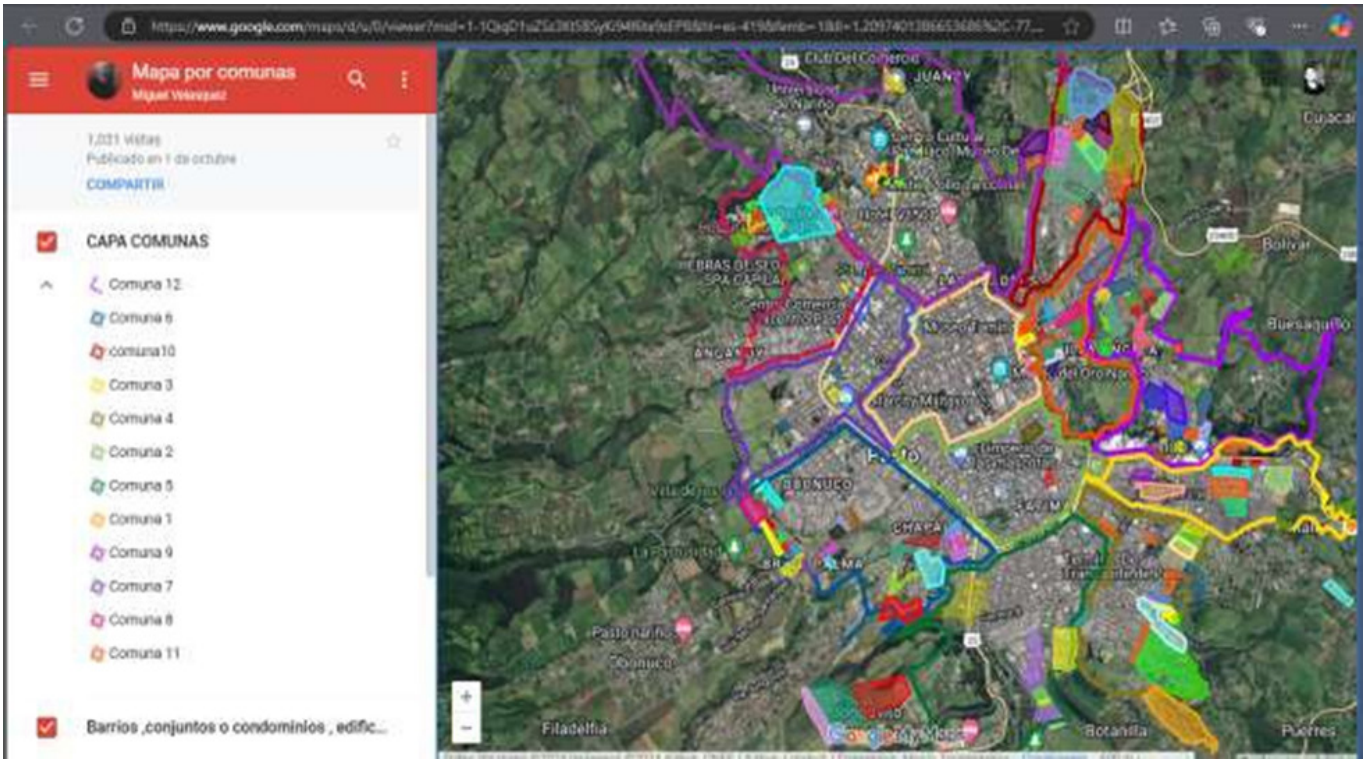


Figure 51. Interface of the web application where the residential projects of the city of Pasto are visualized grouped by neighborhoods and communes

Predictive analysis over a range of years:

Figure 52 shows the predicted urban growth trend from 2022 (Bar 1) as a starting point, until 2027, which would be the year entered by the user, where the predictive model makes the respective prediction by obtaining this information from the web application, allowing it to be visualized in the system in the same way. Thus, it can be said that the predicted urban growth is obtained by the model from the urban variables with which it was trained.

Figure 52. Graph of predicted urban growth

Figure 53. Graph of the annual difference of predicted urban growth

The above graph shows the annual difference of the predicted urban growth of each year with respect to the previous one, thus showing the difference value for each year, thus obtaining an overview of the years that grow more or less according to the previous year.

It should be noted that the prediction values are generated by the predictive model and are sent back to the web application through the api, as shown in the following figure.

Figure 54. Workflow of the web application, Chat Bot and the predictive model

Año	Crecimiento Urbano Predicho	Diferencia Anual	Categoría de Crecimiento
2022	0.48	0.00	
2023	0.81	0.34	Crecimiento Bajo
2024	1.16	0.35	Crecimiento Bajo
2025	1.32	0.16	Crecimiento Bajo
2026	1.40	0.08	Crecimiento Bajo
2027	1.77	0.37	Crecimiento Bajo

Figure 55. Tabular data of the annual difference for each year

The predicted urban growth and annual difference values generated by the predictive model are tabulated by the model itself and then categorized. For the determination of the categorization levels, the following is considered:

Classification of levels according to percentiles:

Formula to consider: P=(k*n)/100

- Low predicted urban growth: categorized as “Low Growth” where it is assigned the color red, considering it also with 25 % according to the percentile.
- Stable predicted urban growth: categorized as “Medium Growth” where it is assigned the color orange, considering it also with 50 % according to the percentile.
- High predicted urban growth: categorized as “High Growth” where it is assigned the color red, considering it also with 75 % according to the percentile.
- It should be noted that the white color represents the starting point of the prediction value.

Predictive analysis by city zones

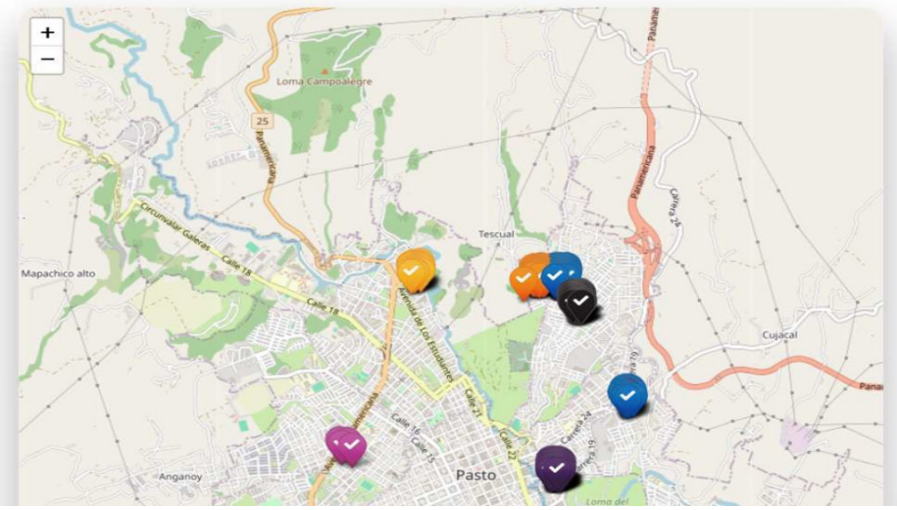


Figure 56. Map of predicted urban growth considering the Type of housing of residential projects

Figure 57. Levels of horizontal or vertical growth of residential projects

The above graphs highlight the concentration of points in seven zones of the city, where each zone corresponds to a neighborhood. This concentration is represented by an identifying color that highlights the level of growth, either vertical or horizontal, of each area. The methods used to obtain these levels are explained below:

To define the levels, a query was performed in the repository using the following instruction: "SELECT `home_type`, `prediction`, `neighborhood_id` FROM `predictions` GROUP BY `neighborhood_id` ORDER BY `predictions`.`prediction` ASC;", This query sets the levels based on the value of the prediction, which in this case represents the predicted urban growth. In addition, it is related to the most predominant projects in each neighborhood, considering the type of housing. Thus, it is evident that the type of housing directly influences the predicted urban growth, as illustrated in the figure below.

	tipo_vivienda	prediccion 1	barrio_id
Editar Copiar Borrar	Residencia	0.010071375462131	1
Editar Copiar Borrar	Residencia	0.010197141327536	2
Editar Copiar Borrar	Residencia	0.010225414060676	3
Editar Copiar Borrar	Residencia	0.010270778947074	142
Editar Copiar Borrar	Conjunto	0.010444525959599	110
Editar Copiar Borrar	Edificio	0.025777366951184	263
Editar Copiar Borrar	Edificio	0.16296004495575	228

Figure 58. sql query comparing housing type with predicted urban growth value

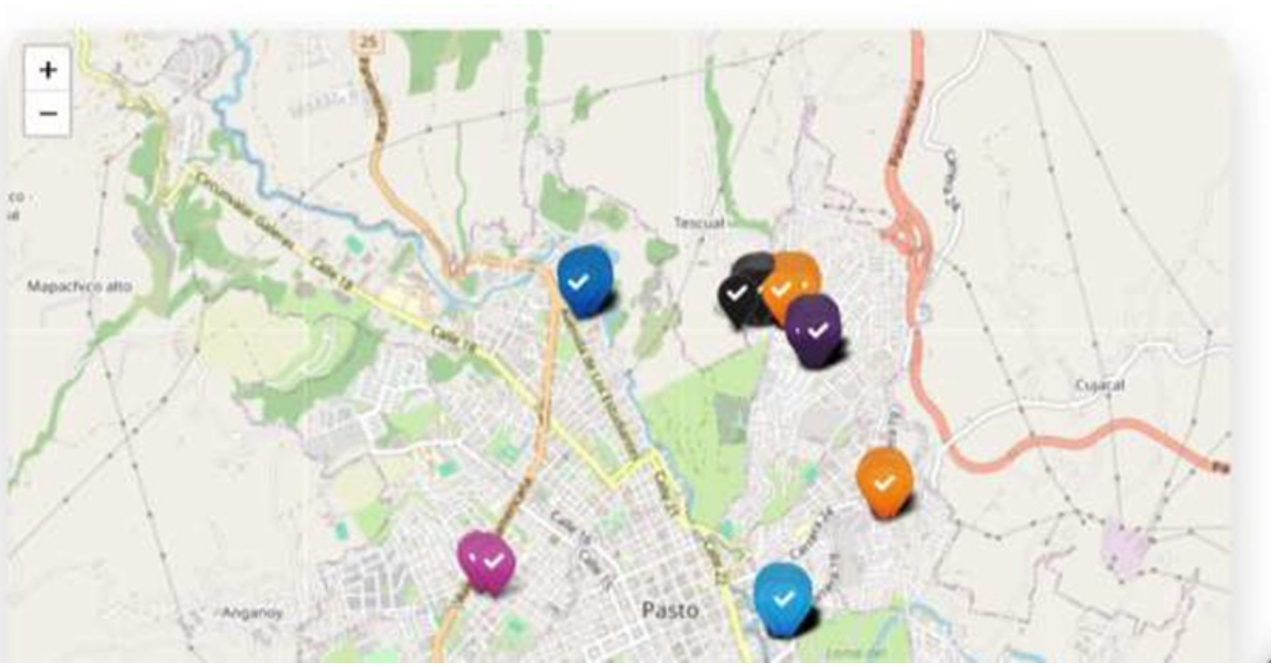


Figure 59. Map of predicted urban growth considering the morphology of residential projects

In the previous figure, it can be observed that neighborhoods with predominance of Residence type dwellings present the lowest predicted values. In contrast, the complexes show slightly higher prediction values compared to the residences. On the other hand, buildings present the highest prediction values. It is important to note that the prediction values for residences, buildings and clusters may vary, since the model was trained with multiple variables. For example, in the neighborhoods identified with IDs 263 and 228, the prediction value changes due to another variable: the stratum. However, in this case, we are specifically comparing the predicted urban growth (prediction column) with the type of housing. Finally, predicted urban growth can be analyzed according to horizontal or vertical growth in the following way; if, for example, buildings have a higher predicted growth, this could indicate a tendency towards verticalization in certain areas. On the other hand, if residences have a high prediction, it may reflect a more horizontal development.

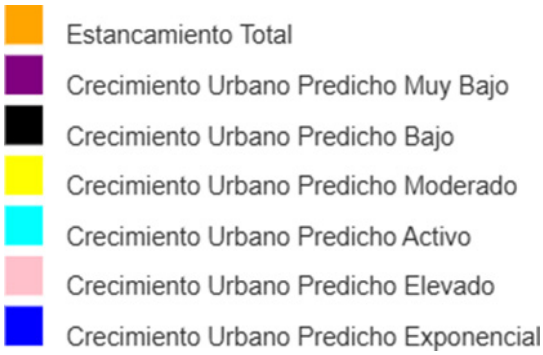


Figure 60. Levels of predicted urban growth considering the morphology of residential projects

The above graphs highlight the concentration of points in seven zones of the city, each representing a neighborhood. This grouping is visualized through a range of colors that identifies and highlights the level of predicted urban growth, either total stagnation or exponential, in each area.

To define the levels, a query was performed in the repository using the following instruction: “SELECT `morphology`, `prediction`, `neighborhood_id` FROM `predictions` GROUP BY `neighborhood_id` ORDER BY `predictions`.`prediction` ASC;”, This query sets the levels based on the value of the prediction, which in this case represents the predicted urban growth. In addition, it is related to the most predominant projects in each neighborhood, considering the morphology of the dwellings. Thus, it is evident that morphology directly influences the predicted urban growth, as illustrated in the figure below.

←T→				morfologia	prediccion ▲ 1	barrio_id
☐	Editar	Copiar	Borrar	Unifamiliar	0.010071375462131	1
☐	Editar	Copiar	Borrar	Unifamiliar	0.010197141327536	2
☐	Editar	Copiar	Borrar	Unifamiliar	0.010225414060676	3
☐	Editar	Copiar	Borrar	Unifamiliar	0.010270778947074	142
☐	Editar	Copiar	Borrar	Unifamiliar	0.010444525959599	110
☐	Editar	Copiar	Borrar	Multifamiliar	0.025777366951184	263
☐	Editar	Copiar	Borrar	Multifamiliar	0.16296004495575	228

Figure 61. sql query comparing the morphology with the predicted urban growth value

In the figure above, it can be seen that neighborhoods with predominantly Unifamilair type housing have the lowest predicted values. In contrast, projects with Multifamily morphology show higher prediction values compared to single-family. It should be noted that the prediction values for single-family, two-family and multifamily projects may vary, since the model was trained with multiple variables. For example, in the neighborhoods identified with IDs 263 and 228, the prediction value changes due to another variable: strata. However, in this case, we are specifically comparing predicted urban growth (prediction column) with morphology. Finally, predicted urban growth can be analyzed according to horizontal or vertical growth in the following way; if multi-family morphology shows higher predicted growth, it could imply that areas are looking to increase population density by residential buildings instead of single-family dwellings.

https://doi.org/10.62486/gen202593 ISSN: 3046-4137

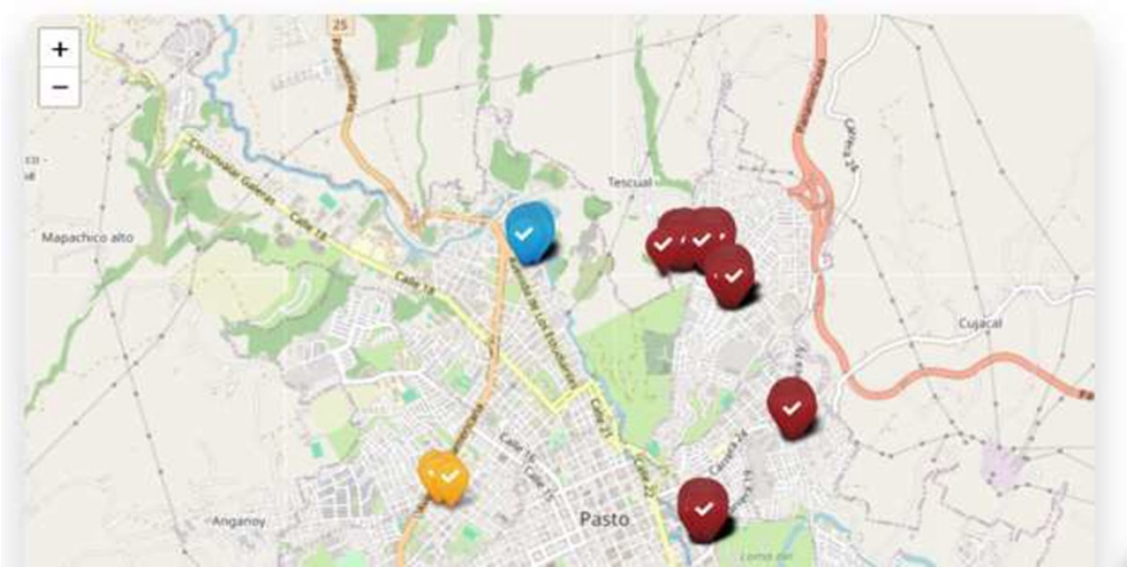


Figure 62. Map of predicted urban growth considering the stratum of residential projects

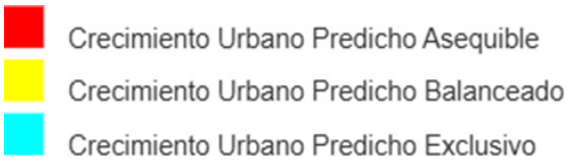


Figure 63. Levels of predicted urban growth considering the stratum of residential projects

The above graphs highlight the concentration of points in seven zones of the city, each representing a neighborhood. This grouping is visualized through a range of colors that identifies and highlights the level of predicted urban growth with respect to the housing stratum.

- Predicted Affordable Urban Growth: high prediction of predicted urban growth in low strata areas (0,1,2,3).
- Balanced Predicted Urban Growth: moderate predicted urban growth in middle strata areas (4).
- Exclusive Predicted Urban Growth: characterized by high predicted growth in high strata areas (5,6).

To define the levels, a query was performed in the repository using the following instruction: “SELECT `strata`, `prediction`, `neighborhood_id` FROM `predictions` GROUP BY `neighborhood_id` ORDER BY `predictions`.`prediction` ASC”, This query sets the levels based on the value of the prediction, which in this case represents the predicted urban growth. In addition, it relates to the stratum of residential projects. Thus, it is evident that the stratum directly influences the predicted urban growth, as illustrated in the figure below.

←T→		estrato	prediccion ▲ 1	barrio_id
☐	Editar Copiar Borrar	2	0.010071375462131	1
☐	Editar Copiar Borrar	1	0.010197141327536	2
☐	Editar Copiar Borrar	2	0.010225414060676	3
☐	Editar Copiar Borrar	1	0.010270778947074	142
☐	Editar Copiar Borrar	3	0.010444525959599	110
☐	Editar Copiar Borrar	4	0.025777366951184	263
☐	Editar Copiar Borrar	5	0.16296004495575	228

Figure 64. sql query comparing the stratum with the predicted urban growth value

In the figure above, it can be seen that neighborhoods with predominantly housing in strata 0, 1, 2 and 3 present a rather low predicted value. In contrast, the variation in prediction starts from stratum 4 onwards. In addition, stratum 5 is the most influential in the prediction value of the model. Finally, the predicted urban

growth can be analyzed according to strata in the following way; if the lowest strata show a high urban growth prediction, it could reflect efforts in affordable housing policies. On the other hand, if the higher strata have higher predictions, this suggests urban growth in more exclusive areas.

The present research has satisfactorily fulfilled the hypothesis: "It is possible to determine the interactive scenarios of urban growth in the city of Pasto through a predictive model, supported by a software product". Throughout the development of the study, different dynamic urban growth scenarios were generated and evaluated, which were validated by experts in the area, ensuring the robustness and reliability of the results obtained.

The results presented have been thoroughly analyzed, demonstrating that the predictive model implemented is not only capable of reflecting the current realities of urban growth in Pasto, but also of anticipating future development scenarios. This analysis has been carried out by adequately fulfilling the specific objectives of the research, such as the creation of a repository of relevant information and the development of an intuitive software that manages this information effectively.

The approach adopted in this study provides a valuable tool for urban planners and decision makers, facilitating urban planning based on accurate data and interactive simulations. In this way, it contributes to the sustainable development of the city, ensuring that the decisions made are aligned with the needs and expectations of the community. Finally, the work carried out not only meets the proposed objectives, but also establishes a solid foundation for future research in the field of urban growth, promoting a proactive and evidence-based approach to the management of urban development in Pasto.

CONCLUSIONS

The determination of interactive scenarios has provided a better understanding of the dynamics of urban growth in Pasto. By integrating a predictive model supported by specialized software, the visualization of possible future scenarios has been facilitated, providing urban planners and decision makers with a valuable tool to anticipate and manage growth effectively. This proactive approach helps to plan urban development in a sustainable manner, considering various variables that affect the city.

The creation of an information repository has been instrumental in consolidating data related to urban growth scenarios. This repository not only compiles historical information, but also includes projections based on the predictive model, which facilitates a comprehensive analysis of development trends. In addition, this resource becomes an essential tool for researchers and planners, allowing them to access relevant and up-to-date data to support informed decision making.

The development of specialized software for managing information on dynamic urban growth scenarios has been a significant advance in research. This software allows organizing, analyzing and visualizing data in an intuitive way, facilitating interaction with the information. By integrating the information collected during the research process, it ensures that users can easily access relevant data and perform comparative analysis, which improves the ability to respond to the challenges of urban growth in Pasto.

Predictive modeling using data analytics has proven to be effective as a simulation tool in the software, allowing the impact of various variables on urban growth to be evaluated. This approach facilitates a detailed analysis that supports decision making, allowing urban planners to consider multiple scenarios and outcomes in their sustainable development plans.

RECOMMENDATIONS

Expand the Repository Database

It is recommended that the information repository be expanded to include additional data, such as historical records of urban policies, addition of new urban planning variables, case studies of past interventions, and updated demographic data. This expansion will provide a richer context for future analyses and allow for a better understanding of the factors that have influenced urban growth over time.

Develop Advanced Functionalities in the Software

To improve the management of information on dynamic urban growth scenarios, it is suggested to implement advanced functionalities in the developed software. This may include more sophisticated simulation tools, additional predictive analysis capabilities, and options for users to generate customized reports that align with their specific needs in urban planning.

Perform Long-Term Simulations

It is recommended that long-term simulations be conducted using the predictive model to assess the impact of different urban policies on growth in Pasto. These simulations will help decision makers visualize the potential consequences of their strategies over time, allowing for proactive adjustments in urban planning.

https://doi.org/10.62486/gen202593 ISSN: 3046-4137

Encourage the Use of Sustainable Growth Models

It is suggested that future studies incorporate sustainable development principles into urban growth models. When evaluating dynamic scenarios, sustainability indicators such as efficient use of resources, protection of green areas and social equity can be included, which will contribute to a more responsible and balanced urban development.

BIBLIOGRAPHIC REFERENCES

1. «Entender la relación ciudad - territorio es comprender la transformación de Pasto». Radio UNAL. https://ieu.unal.edu.co/en/rssieu/tag/Entender

2. Apablaza M, Henriquez C. URBAN PLANNING AND GROWTH: dislocations and URBAN SUSTAINABILITY CHALLENGES OF REGIONAL SANTIAGO, METROPOLITAN. Rev electron geogr cienc soc. 2010;14.

3. Ulloa-Espíndola R, Pérez-Albert Y. Validación de un modelo de predicción del crecimiento urbano en Quito (Ecuador). EURE (Santiago). 2022;48(144):1–27. https://www.scielo.cl/scielo.php?pid=S0250-71612022000200006&script=sci_arttext

4. González et al. Análisis del crecimiento urbano en las zonas de expansión de Pasto 2000-2022. Convocatoria Universidad CESMAG; 2022.

5. Antonio L. Un modelo de crecimiento urbano vertical con factores característicos basado en inteligencia artificial. Uaemex.mx; 2021. http://ri.uaemex.mx/handle/20.500.11799/112219

6. Santos D. 20 ejemplos de páginas interactivas excepcionales. Hubspot.es; 2022. https://blog.hubspot.es/website/ejemplos-paginas-interactivas

7. Ley 388 de 1997. https://www.minambiente.gov.co/wp-content/uploads/2021/08/ley-388-1997.pdf

8. Administradorieu. Entender la relación ciudad - territorio es comprender la transformación de Pasto. http://ieu.unal.edu.co/medios/noticias-del-ieu/item/entender-la-relacion-ciudad-territorio-es-comprender-la-transformacion-de-pasto

9. Cañas EA, Díaz IC. Estrategias para la gestión de grandes volúmenes de datos por medio de Big Data. MVM Ingeniería de Software.

10. El Tiempo. Ahora se puede predecir el crecimiento urbano de las ciudades. 2021. https://www.eltiempo.com/colombia/medellin/innovacion-paisa-permite-predecir-el-crecimiento-urbano-de-las-ciudades-593757

11. Universidad EAFIT. Investigadores de EAFIT pronostican el crecimiento urbano y la demanda de agua en las ciudades. https://www.eafit.edu.co

12. González et al. Análisis del crecimiento urbano en las zonas de expansión de Pasto 2000-2022. Convocatoria Universidad CESMAG; 2022.

13. Caprotti F, et al. The New Urban Agenda: key opportunities and challenges for policy and practice. Urban Res Pract. 2017;10(3):367–78.

14. Blei AM, Angel S, Civco DL. Urban Expansion in a Global Sample of Cities, 1990–2014. Working Paper WP18AB2. Lincoln Institute; 2018.

15. Haller A. Los impactos del crecimiento urbano en los campesinos andinos. Espacio y Desarrollo. 2023.

16. Principi N. Evaluación multicriterio aplicada a la expansión urbana en Luján. UNLu. 2021.

17. Benjamín F, Antonio J. Modelo de evaluación de la capacidad de acogida del territorio con SIG y técnicas de decisión multicriterio. Investig Geogr. 2013;(60):69.

18. Bravo Cobeña CM, Valdivieso Guerra PEA, Arregui Pozo R. Propuesta de modelos predictivos en la planificación territorial. Rev Electr Geogr Cienc Soc. 2007;XI(245).

19. Gómez Delgado CM. Diseño de un modelo basado en agentes para simular el crecimiento urbano. Boletín A G E. 2016;(70).

20. Abrego Almazán D, Medina Quintero JM, Sánchez Limón ML. Los sistemas de información en el desempeño organizacional. Investig Adm. 2015;(115):7–23.

21. Castillo G. Los sistemas de información gestionando grandes volúmenes de datos. Scielo.org.co. 2023.

22. Bravo Cobeña CM, Valdivieso Guerra PEA, Arregui Pozo R. Los sistemas de información en la toma de decisiones gerenciales. ECA Sinergia. 2018;9(2):45.

23. Moreno-Cevallos JR, Dueñas-Holguín BL. Sistemas de información empresarial: la información como recurso estratégico. Dominio Cienc. 2018.

24. Orellana B, Ormeño S, Rodón Módol J. Impacto de la automatización sobre el desempeño. Esan.edu.pe. 2014.

25. Prieto A, Martínez M. Sistemas de información en las organizaciones. Rev Cienc Soc (Ve). 2004;X(2):322–37.

26. Jaramillo Valbuena S, Cardona SA, Fernández A. Minería de datos sobre streams de redes sociales. Inst Investig Bibliotecológicas. 2015;(33):63–74.

27. Jaramillo A. Aplicación de técnicas de minería de datos para determinar interacciones estudiantiles. Espol.edu.ec. 2023.

28. Bazant J. Procesos de expansión y consolidación urbana de bajos ingresos. Bitácora Urbano Territorial. 2008;13(2):117–32.

29. Alexander. Modelo para la evaluación multicriterio de tecnologías en salud. Unal.edu.co. 2021.

30. Lucia O. La evaluación multicriterio social y su aporte a la conservación de bosques. Rev Fac Nac Agron Medellín. 2023;58(1):2665–83.

31. Marcela L, Stevens G. Aplicación de un modelo predictivo para el cambio de cobertura urbana. Udistrital.edu.co. 2022.

32. Sánchez et al. Dinámica del uso del suelo y planificación sustentable. Scielo.org.co. 2023.

33. Enrique J. Acerca de los sistemas de información geográfica. Univ del Valle. 2013.

34. Manuel J, Gitierrez P. Desarrollo de un sistema de información para gestión de proyectos. UCatólica Colombia.

35. Lucía M, Elvirita L. Calidad de la gestión del sistema de información de una IPS. Cuader Lat Admin. 2016;9(17):58–71.

36. Oviedo Carrascal AI, Jiménez Giraldo J. Minería de datos educativos: análisis del desempeño estudiantil. Rev Politécnica. 2019;15(29):128–40.

37. Michele K, Yurani T. Evolución socio-espacial del área urbana de Sotomayor. Udenar.edu.co. 2021.

38. Benitez K, Gomez D. Zonificación por susceptibilidad a fenómenos de inundación. Udenar.edu.co. 2015.

39. Mauricio D, Jaramillo Molina C, Alberto C. Evaluación espacial de centros de distribución de café. Rev EIA. 2022;19(38):10.

40. Muñoz H, Revelo L. Implementación del sistema de información 'Open bravo'. Udenar.edu.co. 2015.

41. Lucía M, Elvirita L. Calidad de la gestión del sistema de información de una IPS. Cuader Lat Admin. 2016;9(17):58–71.

42. Carolina L, Sánchez JJ. Mapa de fallas de los volcanes Chiles-Cerro Negro. Bol Geol. 2017;39(3):71–86.

43. Pereira T, Clara. Minería de datos en supervivencia de cáncer cervical. Univ Salud. 2023;14(2):117–29.

44. López GA. ¿Las infraestructuras como proyecto de ciudad? Ciudades (Valladolid). 2023;(11):105–32.

45. Pérez Bustamante L, Salinas Varela E. Crecimiento urbano y globalización en Concepción, Chile. 2007.

46. Herrmann MG, Van Klaveren A. Disminución de participación ciudadana y planificación urbana. EURE. 2016.

47. Vista de Hacia una política pública de desarrollo urbano sostenible en el Perú. Paideia. http://45.231.72.143

48. De Población En et al. Repositorio UC. https://repositorio.uc.cl

49. Rodríguez P, Palomino N, Mondaca J. El uso de datos masivos para políticas públicas. BID. https://publications.iadb.org

50. Modelando el crecimiento de ciudades medias. Google Books. 2014.

51. Carlos J. Análisis de modelos predictivos basados en visión computacional. Unam.edu.pe. 2019.

52. Fernando L, Carlos J. Desarrollo de software educativo sobre Ubuntu 13.10. Unach.edu.ec. 2016.

53. Cataldi Z, Lage F, Pessacq R, García Martínez R. Ingeniería de software educativo. UBA. https://d1wqtxts1xzle7.cloudfront.net

54. Varella G, Dos Santos M, Francisco C, Cassettari E. Propuesta de priorización de requisitos de software. ResearchGate. 2020.

55. Aguilera Benavente F. Predicción del crecimiento urbano con SIG y autómatas celulares. Geofocus. 2006;(6):4.

FINANCING

None.

CONFLICT OF INTEREST

None.

AUTHORSHIP CONTRIBUTION

Conceptualization: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Data curation: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Formal analysis: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Research: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Methodology: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Project Management: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Resources: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Software: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Supervision: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Validation: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Visualization: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Writing - original draft: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.

Writing - proofreading and editing: Jheison Edilson Arteaga Quistial, Miguel Ángel Velásquez Bravo, Omar Alexander Revelo Zambrano.