ClimaInfo: Information Visualization Dashboard for Climatological and Environmental Data

Camile Maria Cunha Aguiar[1(✉)], Roberto Yuri da Silva Franco[1], Divino Vicente Silvério[1], Tiago Davi Oliveira de Araújo[3], Eduardo Queiróz Marques[1], and Bianchi Serique Meiguins[2]

[1] Federal Rural University of the Amazon, Pará, Brazil
camilemaria55@gmail.com, eduardobio@gmail.com {roberto.franco, divino.silverio}@ufra.edu.br
[2] Federal University of Pará, Pará, Brazil
bianchi@ufpa.br
[3] University of Aveiro, Aveiro, Portugal
tiagodavi70@ua.pt

Abstract. Brazil is a continental country with great potential for agribu-siness according to the specificities of each region. Many of these agribusiness scenarios require climatological data to make more assertive decisions regarding business and public policies. However, the country's northern region, particularly the state of Pará, faces several challenges with the unavailability of climatological data, mainly due to insufficient weather stations. Furthermore, the download from the weather station is limited to a short period, and information is not ready to be used, given several gaps and other issues that require considerable technical knowledge to solve. On the other hand, remote sensing data, while providing a comprehensive dataset covering the entire globe, lacks the precision of station data and isn't easily accessible to the general public. Thus, this article aims to present a web information visualization application named ClimaInfo to present read to use historical climatological and environmental datasets derived from remote sensing products with global coverage, mainly for micro-regions or cities in Brazil that do not have the necessary infrastructure to collect these data. . The remote sensing data are gathered, processed, and presented in different scenarios of interactive dashboards to communicate and support the decision-making processes. Finally, a ClimaInfo communicability assessment was carried out to indicate whether the participant could understand all the elements in the interface. In conclusion, ClimaInfo bridges the gap in data accessibility and enhances the capabilities of stakeholders to make informed decisions through user-friendly web-based applications.

Keywords: Information visualization · Climatology · Environmental · Dashboard · Assessment

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
O. Gervasi et al. (Eds.): ICCSA 2024 Workshops, LNCS 14819, pp. 381–393, 2024.
https://doi.org/10.1007/978-3-031-65282-0_25

1 Introduction

The state of Pará faces a significant challenge in obtaining climate and environmental data; this lack is evident in several available applications, making it challenging to conduct research more quickly [2] and make important decisions in agribusiness and public policies. This data is valuable, as it can help agricultural producers in a given micro-region or city make more assertive decisions based on the climate forecast and the farm calendar [13]. Consequently, the availability of these data and its dissemination practically and intuitively becomes an essential service for society.

For example, Capitão Poço, a city of Pará, has two meteorological stations, one from the National Institute of Meteorology (INMET) and the other from Purple Air, resulting from a research project. However, the data collected by these stations present problems, such as missing data, storage failure, sensor malfunction, equipment calibration problems, and other problems [1]. Therefore, the data is not entirely reliable for research and other applications.

An information visualization application was developed to assist in accessing, representing, and interpreting climatological and environmental data to mitigate this challenge. To overcome the issue of missing data, remote sensing data from ERA5, CHIRPS, and SRTM obtained through Google Earth Engine [6] were utilized, enabling the collection of a series of data at different time intervals. Among the variables that compose the proposed set of data, the following stand out: precipitation, temperature, slope, and elevation, focusing on historical data of minima, averages, and maxima for the cities of Pará.

The ClimaInfo web application's main interface is an interactive dashboard for data visualization. It uses intuitive visualization techniques that facilitate data understanding, even for users without a technical background. Furthermore, it allows data to be downloaded in CSV format and used freely without access restrictions.

Finally, a communicability assessment of the ClimaInfo application was carried out to determine the ease of understanding the proposed data visualizations, the application's functionality, and the available interactions.

Thus, this project offers three main contributions: the availability of a dataset of climatological and environmental data, the development of a web application, ClimaInfo, for disseminating data through interactive dashboards, and the dissemination of climatological and environmental data from cities in the state of Pará.

2 Background and Related Works

Information visualization involves representing complex data through visual elements and creating abstract representations that facilitate understanding of the data and associated information [15]. Such visual representations are remarkable in communicating denser information, as they allow large volumes of data to be condensed and interpreted more clearly and accessible in a simple visualization [9].

For instance, when analyzing the impact of climate change, the most common representation is a choropleth map, as in Fontana's work [3], because it aims to enable the user to understand the climatic conditions of a particular location quickly.

Commonly, information visualization tools use a dashboard because its use combines various visualization techniques, in addition to seeking to mitigate potential problems in data reading, and offers a simpler understanding of information to the user [10], given that each applied technique highlights the representation of a specific group of data.

Some fundamental steps compose the views of visualization techniques to make data into perceptible visual representations [7]. Data Transformation and Analysis, where data is structured and standardized, defining attributes such as name, type, range, and semantics. Selection and Filtering, where focus data is selected based on data modeling results. Mapping, where focus data is linked to visual representations, allows the perception of patterns and characteristics. Rendering, where data and its visual attributes are transformed into images, and additional information is added to assist in understanding the visualization. User Interface Control, where user interaction with the visualization occurs, exploring data from different perspectives.

Some ways to evaluate visualization tools are assessments that use the System Usability Scale (SUS) method or the Communicability Evaluation Method (CEM). SUS is a 10-question questionnaire that evaluates user effectiveness, efficiency, and satisfaction [12]. The test involves participant training, specific task performance, and questionnaire response. The test has these stages, which are qualitative and quantitative evaluations. Scores with different usability levels range from 0 to 100 [12].

CEM evaluates whether the system's functionalities and interface are understandable to the user [8]. The evaluator defines tasks for the participant and observes communication breakdowns. The SUS questionnaire was also applied, and the results analysis was similar to that of traditional SUS. Participant support is limited in CEM, with only initial training, but training also occurs before the test begins, based on a video recorded by the evaluator.

3 Methodology

The first step involved data collection. Due to the unavailability of weather station data for many municipalities and the gaps in the available datasets, we chose to use data obtained through remote sensing products. Thus, climatic and environmental data were extracted using the Google Earth Engine (GEE) [6], including temperature, precipitation, slope, and elevation. Precipitation and air temperature datasets were acquired from CHIRPS [5] and ERA5-Land [11]. The slope and elevation data were obtained from the Shuttle Radar Topography Mission (SRTM), which has a spatial resolution of 900 m per pixel and global coverage. Using the GEE platform, we calculated for each of the 144 municipalities in Pará state, Brazil, the monthly average, minimum, and maximum values

for temperature and precipitation from 1965 to 2022. We computed average, minimum, and maximum values for each municipality for slope and elevation that do not vary over time. Some information is presented in Table 1.

Table 1. Table with information on data extracted from GEE.

Variables	Source	Temporal Resolution	Spatial Resolution
Temperature	ERA5	1965–2022	0.5°
Precipitation	CHIRPS	1981–2022	0.1°
Slope	SRTM	N/A	30 m
Elevation	SRTM	N/A	30 m

Thus, the application's database consists of sensor-derived data, which may have less precision but provide a complete series with no gaps. Data obtained from weather station sensors will also be processed and organized to eventually complement the application's database, thus forming an accurate and complete dataset from multiple sources (Fig. 1).

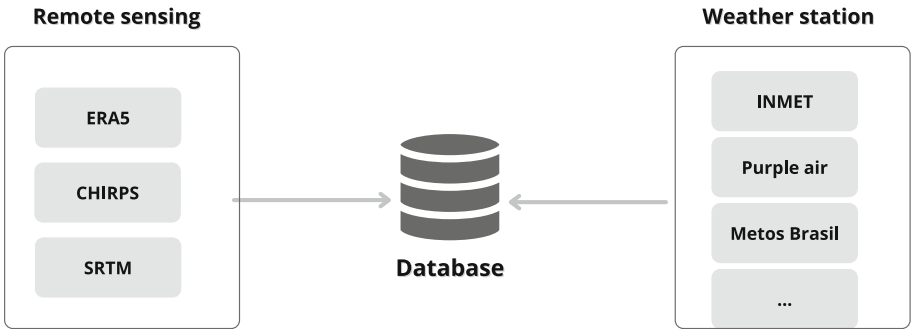

Fig. 1. Scheme of the application's database.

The requirements are divided into functional and non-functional requirements. Functional requirements (FR) outline the system's functionalities and determine user demands [14]. In contrast, non-functional requirements (NFR) specify the product's needs, such as security, which is crucial for software but addresses a software need rather than a user need [14].

Functional requirements include the monthly and annual availability of Pará state data, including precipitation, air temperature, elevation, terrain slope averages, minima, and maxima, represented by simple bars, stacked bars, and choropleth maps. They also include filters by municipality and time interval.

Non-functional requirements are essential elements for the success of the system, extending beyond specific functionalities. They address crucial aspects

such as security, ensuring data protection, and system integrity against potential threats. Additionally, they emphasize system responsiveness, ensuring it efficiently and promptly responds to user requests. A high level of usability is also a fundamental non-functional requirement, ensuring the system is intuitive and easy to use for a wide range of users. Effective communication with the user is another important aspect, ensuring the system provides clear and understandable feedback, and facilitating user interaction with the application.

The project's architecture aims to facilitate the creation and design of the web application. The architecture consists of three main modules. Preparation module: responsible for organizing and preprocessing all data to be used in the tool, whether from remote sensing or automatic sensors from weather stations, adapting them to a specific standard. Representation module: creates information visualizations with established visualization techniques. Interaction module: renders visualizations with data and allows different views and interactions, such as filtering by time and municipality (Fig. 2).

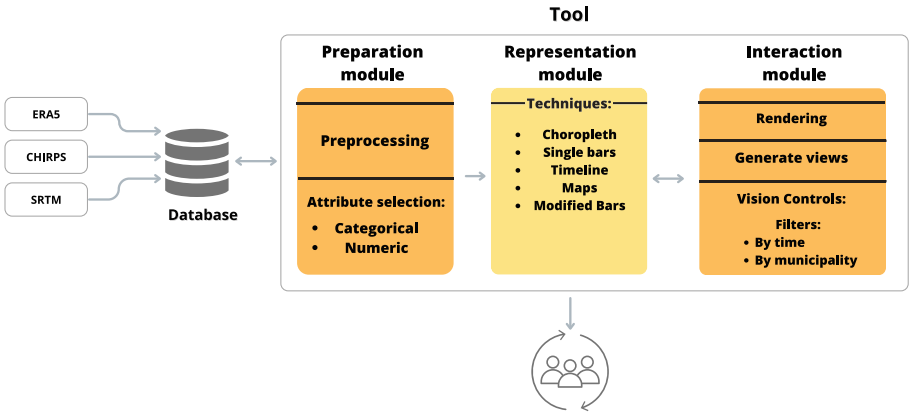

Fig. 2. The figure illustrates how the project's architecture works.

4 Prototype - ClimaInfo

Initially, the Django framework in Python was used to execute and develop the backend. Subsequently, the interface, the frontend, was developed using HTML, CSS, and JavaScript (JS) to create the menu with functionalities such as data entry and listing related to cities, temperature, precipitation, slope, and elevation. Additionally, each data entry and listing screen was created to allow users to perform CRUD (create, read, update, and delete) operations on the mentioned data. These functionalities were implemented so administrators can register and manipulate the data, but the general public will not have access; identification and validation are required to perform such actions. It's worth noting that users

can download all data on all listing screens by clicking a button. This functionality makes the data public and easily accessible to the entire community, enabling everyone to collaborate in filling the gaps in third-party databases.

Subsequently, visualization techniques were developed and manipulated using the D3.js library, as determined in the functional requirements survey. The mapping technique was developed to view the map of Pará and the choropleth map, as well as simple bars and stacked bars. Therefore, with the map, the user can filter which municipality they want to view and select it by clicking on the region, thus generating simple bars and stacked bar charts. There is also filtering by time, which is determined by a range that the user can control.

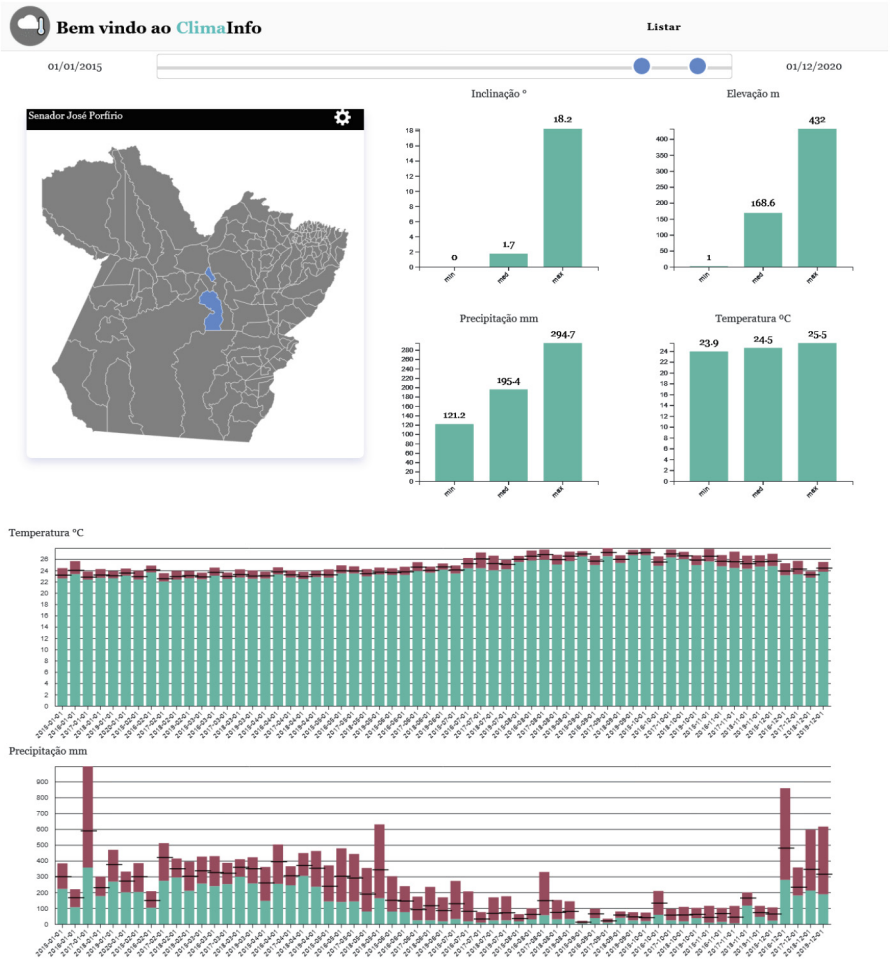

Fig. 3. The application's homepage features a map of Pará and simple and stacked bar visualizations. These elements appear upon interaction with the map, along with a time window for adjusting the months the data will be displayed.

Bar plots show the variables' average, minimum, and maximum data (Fig. 3), where environmental variables like slope and elevation do not vary over time, while precipitation and temperature do. From this, the historical average, minimum, and maximum for the period chosen by the user are presented for climatic variables. Stacked bars are used to visualize the entire chosen time series of environmental variables, with the maximum represented by the red color, the minimum by the blue color, and the average by the black color (Fig. 3).

The choropleth map technique was developed to assign colors to the map according to the data and user preferences. The user can select the desired variable, such as minimum, maximum, or average temperature, and define the initial and final colors, where the initial color represents the lowest recorded amount. In contrast, the final color represents the highest amount. In this sense, the colors vary according to the data of each municipality and the selected period (Fig. 4). All these functionalities can be tested by accessing the ClimaInfo application (http://apps.roberto.eti.br/climainfo/).

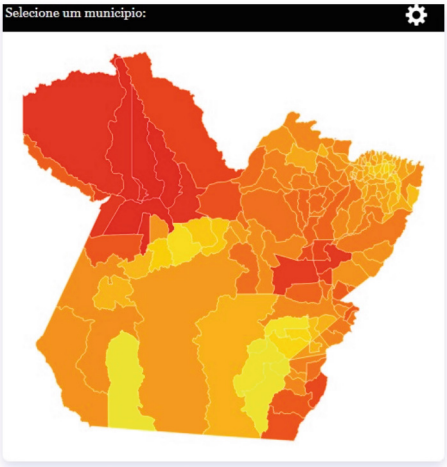

Fig. 4. The Choropleth technique was applied to the map for visualization of the maximum temperature over the selected period to display the municipalities with the highest temperature and make a comparison in the Pará region.

5 Evaluation Methodology

The ClimaInfo communicability evaluation aimed to verify if participants could comprehend the information presented in the dashboard. The process involved several steps, beginning with the application and Informed Consent Form (ICF) presentations. The ICF ensured participants that their data and identity would be protected, and they could withdraw from the evaluation if uncomfortable,

with their data not being stored. Subsequently, participants underwent training to understand the application, facilitated by an instructional video provided by the evaluator, given that the evaluation was conducted via an online questionnaire.

Furthermore, a participant identification stage was conducted to collect information about name, age, gender, education level, location, and whether they were experts in information visualization or climatology. This step was crucial for understanding user-profiles and determining whether they were part of the target audience, which includes students conducting climate-related research and experts in the two areas addressed in this project.

Participants were then asked to perform specific tasks akin to a qualitative test. Each task aimed to determine whether the participants easily understood the data presented in the visualizations. Static images accompanied the questions to assist the participants. Examples of questions include identifying periods with the lowest precipitation amount and selecting regions with the highest amount based on color shading.

1. What is the period with the lowest precipitation amount? (Considering that the options are cutouts from the image below).
2. Considering that lighter shades represent lower amounts and darker shades represent higher amounts, select the option that indicates the region with the highest precipitation amount.
3. What is the maximum temperature amount indicated in the visualization?
4. Which municipality do the presented visualizations refer to?
5. What period will be represented in the visualizations based on the image below?
6. Considering that precipitation and temperature vary depending on the selected time window, answer:
   (a) The minimum, average, and maximum values indicated in the visualization refer to the entire time series of data.
   (b) The minimum, average, and maximum values indicated in the visualization refer to the user-selected time window.
7. The image below represents the visualizations in the chosen time series from 01/01/2000 to 01/12/2015. If the user modifies the time window, will the slope and elevation amounts also be altered? Why?
8. Choose the option that shows the maximum temperature amount recorded in November 2012.

The first three questions aim to verify if the chosen visualization technique could convey the information to the participant, and the last question. All of them were built using a different data visualization technique to test all available ones in the application. The fourth and fifth questions are to analyze if the participant understood the filtering functionality existing in the application; both filters, by time and by municipality, were tested. The sixth and seventh questions are more related to the data; it was an attempt to verify if the participants could understand the data in the application just by visualizing its representation.

An adapted version of the System Usability Scale (SUS) questionnaire was also administered to assess application usability and obtain quantitative results. The SUS questions included inquiries about application usage frequency, perceived complexity, ease of use, need for technical assistance, integration of map functions, consistency in data representation, learning curve, confidence in application usage, and the need to acquire new skills.

1. I would like to use the ClimaInfo tool frequently.
2. I find the ClimaInfo tool unnecessarily complex.
3. I found the ClimaInfo tool easy to use.
4. I would need assistance from someone with technical knowledge to use the ClimaInfo.
5. The various map functions are well integrated into the ClimaInfo tool.
6. The ClimaInfo tool represents the data inconsistently.
7. People will learn how to use the visualizations quickly.
8. I found the ClimaInfo application complex to use.
9. I felt confident using the ClimaInfo.
10. I needed to learn several new things before being able to use the ClimaInfo.

Finally, participants were invited to highlight the positive and negative aspects of ClimaInfo in the feedback stage, providing insights for future improvements. Participants were also thanked for their contribution and participation in the evaluation.

6 Results and Discussion

In communicability evaluation, the target audience was established as individuals who already possess a minimum level of prior knowledge. Thus, participants include those understanding areas such as biology, forestry engineering, and agronomy, as well as experts in information visualization and climatology, to understand how these areas can be better addressed in the application. The goal was to ensure that the application was initially accessible to undergraduate students to make all necessary adjustments for the application to become practical for the entire community regardless of their technical knowledge level in areas such as climatology or information visualization.

As suggested by Forsell [4], it is ideal to have a group with a minimum of 12 people for information visualization evaluations. Based on this, there were a total of 35 participants, who were separated and reorganized into three groups: group 1 consisted of all participants, totaling 31, considering that the selected participants were only those with the desired profile, as four were from the healthcare field, thus not fitting the target audience; group 2 consisted of undergraduate students, with 14 participants; and group 3 consisted of experts, with 13 participants, who consider themselves experts in one of the two areas addressed in this project, information visualization and climatology. The other four participants who complete Group 1 do not belong to Group 2 or 3.

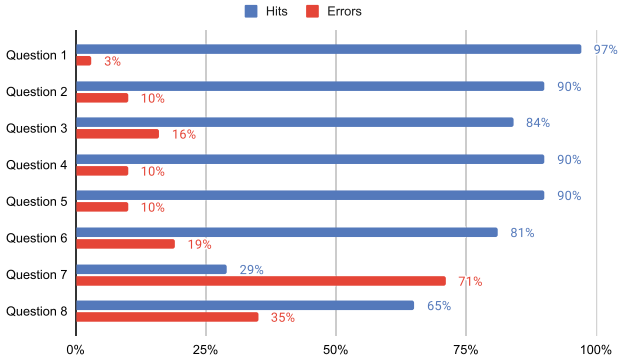

Fig. 5. The figure presents the percentage of errors and correct answers of the group with all participants, where the red bar represents the percentage of errors made by the participants in the applied tasks, and the blue bar represents the correct answers. (Color figure online)

In group 1, the age ranged from 21 to 35 years old, and regarding the educational level of the participants who took part in the evaluation, about 35% had incomplete undergraduate degrees, 29% had a doctorate, and the other levels had a low percentage. In carrying out the proposed tasks, the participants obtained a high correctness rate in seven questions; however, there was an error rate in the seventh question, and after a thorough analysis, it was found that interaction with the application was necessary. For this reason, this task was discarded (Fig. 5). The SUS result calculated for this group was 80.2, classified as "A−" on the scale, which means a good level of usability, but it indicates that there are measures to be improved.

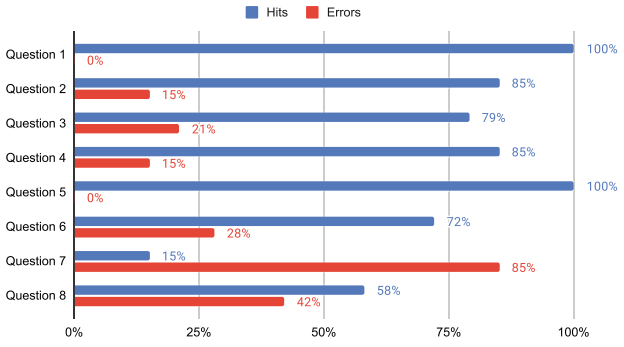

Fig. 6. The figure presents the percentage of errors and correct answers of group 2, where the red bar represents the percentage of errors made by the participants in the applied tasks, and the blue bar represents the correct answers. (Color figure online)

Group 2 gathered over 70% of participants between 21 to 30 years old, with the educational level divided between completed high school and incomplete undergraduate degrees; in other words, no participant in this group has a completed undergraduate degree. This group had the lowest correctness rate in the tasks compared to the other groups (Fig. 6), possibly because the training phase was ignored, as there was feedback that a video teaching how to use the application was necessary. Furthermore, the group with the lowest SUS result, with a value of 79.8, was classified as "A−".

Group 3, composed of people who qualified as experts in the areas addressed in the application, had an age range of 26 to 35 years old, with education corresponding to doctorates, masters, and completed undergraduate degrees. This group achieved the highest percentage of correct answers compared to the previous groups, even in the discarded seventh question(practically 60% got it right), demonstrating high performance in solving the proposed tasks (Fig. 7). Moreover, the method calculation pointed to a result of 87.5, classifying the application as "A+" on the SUS scale.

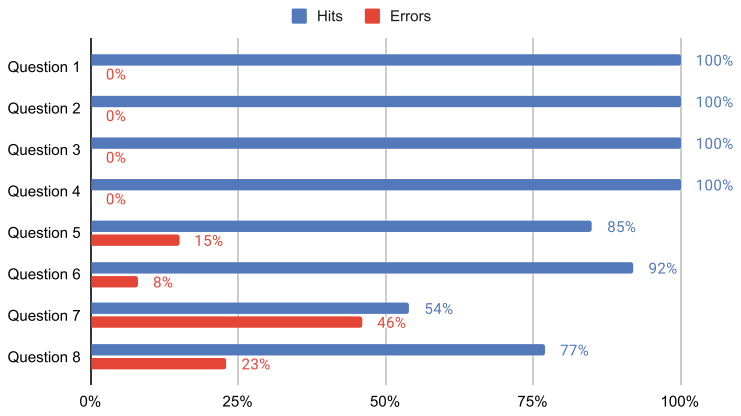

Fig. 7. The figure presents the percentage of errors and correct answers of group 3, where the red bar represents the percentage of errors made by the participants in the applied tasks, and the blue bar represents the correct answers. (Color figure online)

As mentioned earlier, the SUS calculation for the expert group yielded a score of 87.5, placing the application in the 'A+' category on the scale. In contrast, the student group obtained the lowest SUS score of 79.8, resulting in an 'A−' classification. The overall SUS score for all groups was 80.2, also falling within the 'A−' range. While this suggests a commendable level of usability, it also hints at areas where enhancements could be made. It's believed that the higher score attained by the expert group might be attributed to their familiarity with the dataset, thereby streamlining their use of the tool. This correlation is illustrated in Fig. 8.

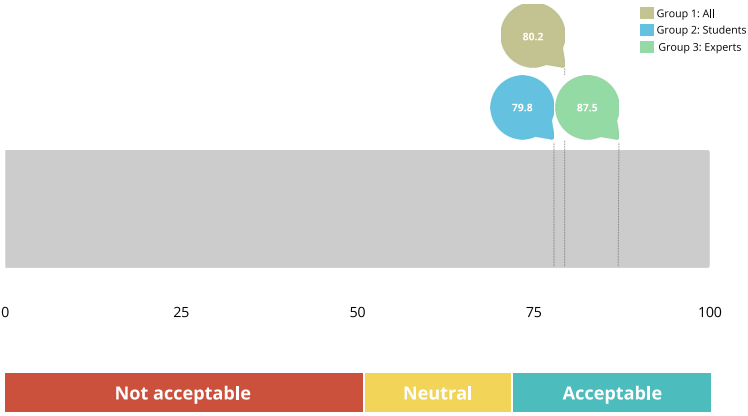

Fig. 8. The figure represents the system usability scale score for the three groups, the expert group achieved the highest score.

7 Final Remarks and Future Works

The ClimaInfo application was developed to address the problem of unavailable climatic data in municipalities in the state of Pará. Through visualization techniques, the application provides representations of minimum, average, and maximum historical temperature and precipitation series and slope and elevation, thus filling gaps in the database. Thus, providing the download of the data used in the application helps fill the existing gaps in the database.

The ClimaInfo communicability evaluation, conducted with different groups, demonstrated that the application is a communicative and user-friendly application. The SUS results indicated a usability level between A− and A+, confirming its usefulness for the target audience.

As future improvements, it would be interesting to display data series from across Brazil and implement new visualization techniques to present the data in different ways and increase interactivity. Additionally, introducing the functionality to compare data from different municipalities would provide users with a more simplified and comprehensive analysis.

References

1. Baba, R.K., Vaz, M.S.M. G., Costa, J.: Correção de dados agrometeorológicos utilizando métodos estatísticos. Revista Brasileira de Meteorologia, vol. 29, pp. 515–526. SciELO Brasil (2014). https://doi.org/10.1590/0102-778620130611
2. Costa, J.C., Pereira, G., Siqueira, M.E., Cardozo, F.D.S., Silva, V.V.D.: Validação dos dados de precipitação estimados pelo chirps para o Brasil. Revista Brasileira de Climatologia, vol. 24 (2019). https://doi.org/10.5380/abclima.v24i0.60237
3. Fontana, E.: Projeta toolbox: módulo de visualização de dados climáticos. Universidade de Passo Fundo (2021)

4. Forsell, C.: A guide to scientific evaluation in information visualization. In: IEEE 14th International Conference Information Visualisation, pp. 162–169. (2010). https://doi.org/10.1109/IV.2010.33
5. Funk, C.C., et al.: A quasi-global precipitation time series for drought monitoring (2014). https://doi.org/10.3133/ds832
6. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R.: Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. **202**, 18–27 (2017). https://doi.org/10.1016/j.rse.2017.06.031
7. Liu, Shixia, Cui, Weiwei, Wu, Yingcai, Liu, Mengchen: A survey on information visualization: recent advances and challenges. Visual Comput. **30**(12), 1373–1393 (2014). https://doi.org/10.1007/s00371-013-0892-3
8. Michelon, G.K., Barbosa, M.M., Hoffman, A.B.G.: Avaliação da comunicabilidade do website moodle. Revista Eletrônica Científica Inovação e Tecnologia **7**(16), 150–160 (2016)
9. Nascimento, H.A.D., Ferreira, C.B.R.: Uma introdução à visualização de informações. Rosana Horio Monteiro (2011)
10. Rolim, D.A.A.: Dashboards para desenvolvimento de aplicações e visualização de dados para plataformas de cidades inteligentes. Dissertação (Mestrado) - Universidade Federal do Rio Grande do Norte (2020)
11. Sabater, J.M.: Era5-land hourly data from 1950 to present. In: Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (2019). https://doi.org/10.24381/cds.e2161bac
12. Sauro, J., Lewis, J.R.: Quantifying the User Experience: Practical Statistics for user Research. Morgan Kaufmann (2016)
13. Schmalfuss, L.S., Santos, R.L.: Estação climatológica da uefs: uma proposta de ligação da ciência com a comunidade feirense. Anais da Jornada de Extensão da UEFS, vol. 1, no. 1 (2019)
14. Sommerville, I.: Engenharia de Software, 9th edn. Pearson Prentice Hall, São Paulo (SP) Brasil (2011)
15. Spence, Robert: Information Visualization. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07341-5