

# A COMPARATIVE STUDY OF TECHNIQUES OF VISUALIZATION OF DISTRIBUTIONS FOR GEOGRAPHICAL DATA

Edcley Silva

Master in Computer Science  
Centro de Informática

Universidade Federal de Pernambuco  
Recife, Pernambuco 2126-8430  
Email: ejs2@cin.ufpe.br

Nivan Ferreira

Assistant Professor  
Centro de Informática

Universidade Federal de Pernambuco  
Recife, Pernambuco 2126-8430  
Email: nivan@cin.ufpe.br

Fábio Miranda

Post Doctoral Research Associate  
Tandon School of Engineering

New York University  
New York, New York +1 646-997-3600  
Email: fmm283@nyu.edu

**Abstract**—Currently, technological advances have revolutionized the way natural phenomena are studied. Natural phenomena can be represented through distributions of geographic data that are a rich source of information and can be explored in different ways. One of them is the representation of uncertainty through the distribution of probability. Modeling the uncertainty of this type of distribution and representing it in geographic visualization is complicated because maps (common types of geographic visualization) need the visual environment to represent geographic space and there are not many ways to represent any other information. One of the ways often used as a solution is statistical summarization such as mean, but summarizing the data alone may can hide the data's behavior and generates ambiguity. The concealment of the uncertainty of the data in visualization can be justified by the way the uncertainty is represented that may not be understood by the user. Technical proposals have been proposed to represent distributions, but generally they only represent the presence and spread of uncertainty recently others approaches based on probability of proportion of data, animation and interaction have proposed to make quantification of probability, but have not been used or compared formally for geographic data. The objective was qualitatively compare main approaches to visualize probability distributions on a geographical scenario (includes factors such as distance, size and variation), using the recent proposed approaches in the context of abstract data, analytical tasks and user study. The results show which approach has the better performance in the presented cases.

## I. INTRODUCTION

An ubiquitous aspect in data analysis (particularly in geographic ones), which is commonly ignored [1], is the presence of randomness or uncertainty. It is uncertainty comes from the intrinsic randomness present in natural phenomena and variability present in the data due to completeness, consistency, errors, modeling, sampling and etc. [2]. An existing challenge is to model this uncertainty using probability distributions. Thus, a central problem in data visualization geography is like supporting the visual exploration of probability distributions on maps since the visual channels of the maps that are generally used to represent geographic data is often busy. The most common solutions for representing probability distributions in

data geographic regions make use of summary measures, or here called statistical summaries, as the mean and the variance [3]. These summaries are efficient in reducing the amount of data to be analyzed, however, it is possible to have completely different data collections with the same statistics as mean and variance [4]. This not only makes representation ambiguous, but it can also hide important patterns present in the data [5]. Furthermore, and most importantly, these representations fail to be well interpreted, even by individuals with statistical training [6] [7] [8]. This difficulty is in the fact that users of such visualizations must make mental inferences, in order to perform analytical tasks (comparison, sorting, etc.), which is a non trivial and, in many cases, impossible. For this reason, it has been defended by many researchers a more reliable representation of the distribution of the data represented and not only statistical summaries [9] [10]. Some recent work has proposed solutions to this problem of quantifying uncertainties but in the context of abstract data. These are based on the use of sampling and interactivity tools [11] or in the use the frequentist view of probability to build visualizations [12] [13]. The great innovation of the proposals made was to communicate uncertainty and quantify it, making visualizations dynamic or more intuitive. However, as far as is known, they have not been applied to geographic data and, also, have not been formally compared.

## II. RELATED WORK

Data distributions are essential objects in data analysis. These mathematical objects are commonly used in addition to representing data sets modeling the variability of the set and also to model the randomness / uncertainty of phenomena quantification tasks. In this section we present the way how the data distributions has been used recently in others works.

### A. Viewing Data Distributions

Interactive data visualization systems in general make great use of aggregations to summarize data distributions and, thus, build visual representations in different areas. We highlight the

works of [14] [15] [16] [17] [18], which are clear application of the search for better exploration of results of modeling and predictions, in addition to the attempt to enable the study of the variability of model predictions for different approaches and parametric configurations. In addition, there are also strategies to support interactive visualization large amounts of data, either by sampling or by building data structures as in the work of [19] on Gaussian Cubes. implementations efficient data cubes to store statistics considered important. Another strategy for this problem is the proposal by [20], who use progressive sampling, which is a technique of approximate the value of the query by aggregations and pre-calculations of a data set, so that in large volumes of data the user does not have to wait until the end of the processing the entire base to have an answer. In general, these techniques use statistical summaries to represent distributions, mainly statistics at the moment which despite being a flexible representation can be problematic because it hides the variability (the uncertainty) giving people a false sense of precision about the data [21], represent few characteristics of a distribution, for example, the mean, the variance [8], quantiles [22] and entropy [23] and is difficult to interpret in tasks [7], even for individuals with statistical training [6]. In fact, analytical tasks view centers how to compare distributions are difficult to do because users need to interpret all questions and conclusions in a probabilistic, even with the aid of visualizations, can still be very difficult given the great complexity of the analysis process. For abstract data [11] they present a tool, based on interactive annotations, to represent uncertainty. This tool is based on the idea of quantification automatic probability of some analytical tasks through interactions. [12] present the Hypothetical Outcome Plots that makes use of animation for depict possible samples of the distribution and that can be considered more intuitive. The counterpart is that the use of this technique is limited to a certain amount data and requires a lot of attention to follow the animations. Already [24] we bring the quantile dot plot that tries to improve the probability estimate accumulated, using an adaptation of dot plots (dot plot). In this case, there is a need to perform a count which requires greater effort. All the solutions presented brought some benefit to quantification but were not tested with analytical tasks or formally compared in a geographic data environment. Subsection text here.

### 1) Distribution Display Techniques For Geographical Data:

In the geographical context, [25] classify the various methods for visualizing distributions / uncertainty between static, dynamic and interactive, and the types of uncertainty that they can represent in three: positional, attribute-value and temporal. Attribute value uncertainty is the property of spatial data that reflects the correctness/variability of attributes associated with spatial positions. Temporal uncertainty is related to the change in data over time and expresses the occurrence of the data in comparison with the required frequency of updates. In this paper, the focus will be on attribute value uncertainty, as this is the most common scenario in data analysis.

A classic example of a static approach is the choropleth

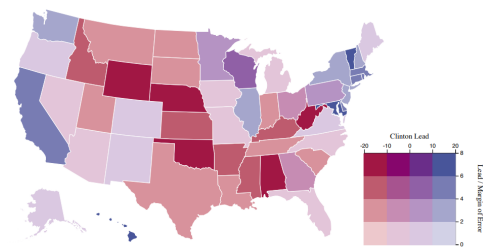


Fig. 1. Use of multidimensional scale (x and y) to display uncertainty by [8].

technique, which is basically a visual representation of a statistical summary of a distribution (corresponding to a region of the map) with colors. In a choropleth, generally, the distribution average of each region is mapped to a color that is used to fill the pixels that make up the corresponding polygon and the limits of polygons give rise to the concentration of occurrence of an event. Many variations of this technique have been proposed to represent more information about distributions. As the multiple view [2], where more than one visualization is used to represent different values and possible correlations, therefore requiring a great cognitive effort. Another still appears in the work of [26]. To quantify uncertainties in a very simple way, this work brings the use of bi-dimensional colors/legends to not only represent the average but also some measure of spreading (see Fig.1). Another static method is that of glyph that uses objects or symbols that represent information in the context in which they are inserted (see Fig.2). Glyphs have been adapted to try to represent more than one variable and thus portray possible uncertainties. An example is the blur technique (Blur). However, a common

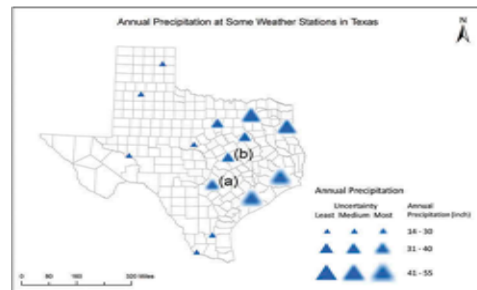


Fig. 2. Glyph application on map by [27].

problem in glyphs is overplotting, which is the insertion of one object on top of the other, generating inhibition of elements, due to the need to represent a lot of information and have little space. In addition, due to the scale of the map and lack of alignment between glyphs, it is often difficult to make comparisons between these graphic elements and, therefore, to estimate the probability in user tasks with precision. Another static technique is the point dispersion technique. It is based on the scatterplot technique, which is a traditional scatterplot that represents each object in a data set with a point, positioned in two continuous orthogonal dimensions x and y [28]. The distribution values are represented by the color and number

of points in a given location, but this technique, although widely used in geographic visualization systems, suffers from the same problems as glyphs. Subsubsection text here.

The interactive techniques can be considered an evolution of the previous ones because they make use of techniques like choropleth but with the aid of a cursor for example that allows the user to carry out some interactions. [5] use a cursor to allow the quantification of uncertainty using quantiles of the distribution or CDF of each distribution. However, the authors do not make an assessment of how This approach is effective. Another proposal by [29], uses interaction to communicate the type of attribute uncertainty and how they relate and impact the behavior of geographic regions to make predictions. I understand that a high learning curve is necessary because exploration involves many actions and understanding the representations may not be so simple. The same happens in the work of [30] but this time the map does not reflect a choropleth, but a heat map.

Finally [31] use an dynamic technique. They use animation flows, which change frequently to show the actual displacement of the taxis and try to make predictions with quantification. When comparing this technique with a static one, users reported that perceiving uncertainty with is easier with uncertainty but quantifying it is more difficult.

### III. PRELIMINARIES AND EXPERIMENT DESIGN

We use a research methodology for experiment design based on [11] and [12]. And in this section, all the details of the methodology used are presented.

#### A. Analytic Tasks

The tasks that were used in the study were defined based on the study by [32], of the types of tasks of greatest interest to the user. In their work, [11] they discuss the need to translate these tasks into probabilistic versions when dealing with probability distributions. In this way, the goal is to support the quantification of the probability of some fact (object of the analytical task) to happen. In the context of geographic data, [33] in their work they classify these tasks by performing a refinement, highlighting for example comparison tasks as being key factors in identifying patterns and behaviors regarding techniques and the data itself. Based on these and other works, analytical tasks were selected to assess visualizations in the geographical context, the tasks are presented in the table. Be  $A..Z$  individual areas with a data distribution and  $c$  a constant the tasks are: **Extract Value** ( $P(A \leq c), P(A \geq c), P(c_1 \leq A \leq c_2)$ ): Given an area find a probability that the area have data between the specific range. **Compare Distributions** ( $P(A \leq B)$ ): Given a pair of geographical region A and B, estimate the probability of a sample in the first region being smaller than one of the second region. **Find extremes** (maximum and minimum): Among all areas of geographic data ( $A_1, \dots, A_n$ ), the maximum and minimum under a given condition must be informed ( $P(A_i \leq c)$ ). **Estimate the arithmetic mean** of a given distribution (calculate derived value);

#### B. The Datasets

In this paper, two datasets of data distributions of selected geographic events are used to simulate real scenarios of geographic data analysis. The first dataset refers to distributions of rainfall index observed in the cities of the state of Pernambuco - BR, composed of 184 regions, with rainfall data for all months of 2016-2018. Totaling a set of 36 values for each region. Data are collected through meteorological stations and made available in an open manner by the regulatory agency - Agência Pernambucana de Águas e Clima through its website<sup>1</sup>. The second dataset refers to taxi rides in the city of NYC (USA) using only the rides that start in some neighborhoods on the island of Manhattan. 29 neighborhoods with data for the months of 2018 totaling 12 values for each neighborhood. Data is collected through the monitoring of taxis and is made available in an open manner by NYC City Hall through the New York City Taxi and Limo Commission<sup>2</sup>.

#### C. Visual Metaphors Applied

In this paper, the common case of visual analysis comes from geographic data, where there are limited geographic regions (e.g., neighborhoods, cities, states, etc.) and a probability distribution associated with each of these regions. The most commonly used visual metaphor in this case are choropleths and their variants (choropleth maps) so we will approach choropleths with adaptations of 3 abstract data approaches, complementing the most frequent approaches to those of interaction: animation (here called hypothetical outcomes maps HOM), interaction (interactive map IM) and density of points here called dot map (DM) (besides being widely used in maps, this technique has principles of glyphs but better treating some problems such as overplotting).

The DM is based on the idea of dot plots [21] for the spatial context. To visually represent the distribution corresponding to a particular region, this technique draws points positioned randomly within the region. Each point corresponds to a sample of the distribution. The dots are colored according to the value of the sample associated with it. The greater the incidence of a value in the distribution of that region, the more points received a certain color. The number of points is proportional to the area in m<sup>2</sup> of the polygon.

HOM, on the other hand, uses a frequentist approach based on [12] taking samples from different distributions as input to choropleth as a possible hypothetical scenario. HOM uses animations to display several hypothetical maps. It is possible to sample data in several ways in this paper we use data correlation sampling. Thus, the resulting animation presents 12 frames that are presented at random (at the end, the animation resumes). The exposure time for each frame was defined as 660ms. This time was established through a pilot test with users.

The IM is based on the proposals of [11] [5]. In this technique, a number of tasks are supported through implemented

<sup>1</sup><http://www.apac.pe.gov.br/meteorologia/monitoramento-pluvio.php>

<sup>2</sup><https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>

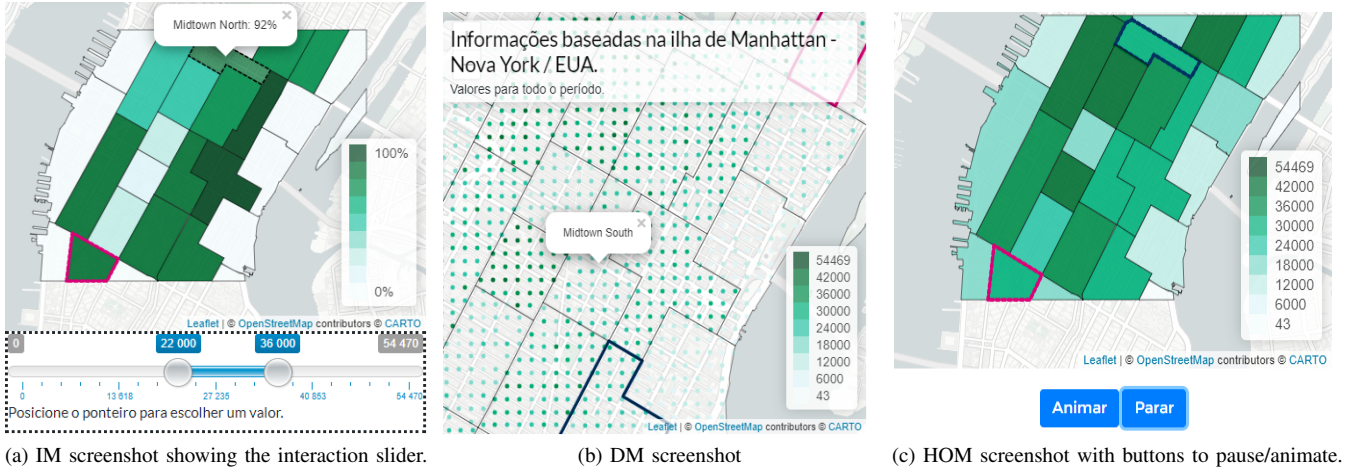


Fig. 3. The Interface created for the visual metaphors applied

interactions making it possible to quantify the probability that the fact described by the task actually occurs. In the present study, the tasks of extracting value and comparison were fixed. To support the first task, the interface features a slider, in which the user can define a range  $[c_1, c_2]$  values. The system then uses this interval to implicitly quantify the value of  $Pr(c_1 \leq A \leq c_2)$  each region. The user can change the interval value and interactively the associated choropleth map is updated. To support the comparison task, the user can double-click on a region  $A$  and all other regions on the map are compared with  $A$  in terms of distribution, the process result is  $Pr(A \leq B)$  where  $B$  represents each one. regions and their respective data sets. The three techniques were implemented on the web platform, in the JavaScript language, using the libraries D3.js, Leaflet.js and Turf.js. and the Fig.3 shows the system created for execute this study.

For the comparative study between the techniques, there are some hypotheses about their performance to be validated from the data collection in the study with the user. Are they:  $H1$  - Among the 3 techniques, visualization with interaction is the most accurate for the tasks of extracting value, comparing and finding extremes.  $H2$  - Point mapping techniques and hypothetical maps should have lower confidence measures than the interaction technique except for the task of estimating average.  $H3$  - The user's response time in the interaction technique tends to be less than or equal to the time recorded in the other techniques, except for the task of estimating average.  $H4$  - The user's response time tends to increase the greater the geographical distance between the regions to be compared.  $H5$  - Factors such as distance between polygons, size of regions and the magnitude of the variance of the distribution of each region negatively impact the accuracy of the user for the answers of the tasks in the techniques of point map and hypothetical maps for the comparison task.

#### D. Execution of case study with user

For the data collection of this paper, a questionnaire inserted in the visualizations divided into three parts was used. The first

part (with 6 questions) of the questionnaire was qualitative and deals with personal issues that aim to map the participant's profile. The second part (26 open questions about the analytical tasks) of the questionnaire was quantitative and concerns questions that involve the user study part itself, the tasks, and aims to measure the efficiency of the technique under study to represent risk, uncertainty and make comparisons as seen in studies literature and cited in this paper. In this part, time and precision were still collected to measure the accuracy of the bets. The third part of the questionnaire was qualitative and used a textual field for the participant to write and give their opinion under the visualization. Between part 1 and part 2, training was allocated regarding aspects of visualization, that is, how the technique would work so that the user would be able to learn before using it.

TABLE I  
SAMPLE QUESTIONS FOR EACH PROPOSED ANALYTICAL TASK.

Analytical task type	Question example
Extract Value	What is the probability to rain on A between 50 and 150mm?
Compare Distributions	What is the probability of raining less in A than in B?
Find extremes	Select the name of the region most likely from the list raining below 100mm?
Estimate the arithmetic mean	Estimate the average rainfall in A.

Each metaphor was hosted individually to ensure that each user used only one metaphor. The tests were carried out under supervision, being randomized where neither the researcher nor the user had control over the direction of the metaphors to users and the Think Aloud protocol was also used to capture the users' natural comments and reactions. The Tab.I shows some examples of questions introduced to the users. For the task of comparing, we also included some factors: distance between areas A and B (1 long case 2 short case), size of

polygons A and B (1 large case 1 small case) and data variance (1 low case 2 case high).

#### IV. RESULTS AND DISCUSSIONS

The results are formed by a population that had to obey some criteria to avoid bias among which we mention 2: Not having previous contact with the solution and having a minimum level of education at graduation. The population consists of 66 (sixty-six) participants, alternately divided (as mentioned in the previous section) into 3 groups of 22 people, where each group represents one of the visualization metaphors studied. The population consisted of 41 male and 25 female individuals, of the 66 individuals 39 said they had experience with statistics against 27, 43 said they had experience with visualization against 23, and of the 66, 36 had undergraduate education while 30 postgraduate.

##### A. Qualitative Results

Among all the comments collected, either by the field in step 3 of the study, or by the think aloud protocol, we highlight some to summarize in general the opinions of users, these opinions help to identify indications that some hypotheses will or will not be proven. In IM in general, users were afraid to trust the tool's interaction functionality: *"I was a little afraid to trust the tool, but when you learn and understand its use, it is much easier, even so I tried to check it out a test to be sure and increase my confidence."* For DM, users in general found it a more difficult process to find a logic to answer the questions: *"At first I found the points a little more complicated, I found it difficult to estimate, so I counted the colors of the balls, looked up from above, when I zoomed in I could see better, as time went by I was more confident, some questions can be confusing if I don't pay attention, I think it would be interesting to see the values of each ball"*. Although users do not know about the factors implicated in the task of comparing areas, the DM highlighted several times, as in the following comment: *"You have to analyze well because there are regions that are closer and others more distant, in some cases it is necessary to zoom in, by zooming in, you have a better idea"*. At HOM, users also noticed the factors and felt more difficult to follow the transitions and make visual inferences as we will see below: *"It is a lot to analyze and observe, so many variables passing at the same time, I missed using paper or calculator, in a way is tiring and the question of answering as quickly as possible also creates insecurity"*. And yet another complements about HOM: *"you are insecure [with the metaphor of HOM] especially those who are used to calculating that it is my case"*. These comments indicate that the quantitative results should mainly reflect H2 and H5.

##### B. Quantitative Results

To analyze the data, all the quantitative data collected (accuracy in relative error, measured time and confidence) were subjected to statistical analysis according to the groups of each task composed of the results of the questions of each group. The Kolmogorov-Smirnov (KS) normality tests

with 95% accuracy indicated 0.000 significance for all cases, strongly rejecting the hypothesis that the data would follow a normal distribution. In the case of samples that are independent and do not follow a normal distribution, the Kruskal Wallis (KW) tests were applied with 95% accuracy, since each test would consist of 3 samples that represented each of the metaphors in order to find significant differences or not. In case the difference hypothesis between samples was positive, the pairs were submitted to the Wilcoxon test (W) with 95% accuracy.

1) *Extract Value Results:* In this task in terms of accuracy for the rainfall dataset, the metaphors had mean and variance, respectively: IM 0.03 & 0.02, DM 0.54 & 0.04, HOM 0.40 & 0.39. The p-value of KW was 0.000 (very high), whereas the maximum p-value of W for pairs was 0.006.

In the taxi dataset, the metaphors had mean and variance, respectively: IM 0.03 & 0.17, DM 1.45 & 2.34, HOM 0.70 & 2.33. The p-value of KW was 0.000 (very high) and the maximum p-value of W for the pairs was 0.007. These results allow us to say that IM performed better confirming H1. A similar result was observed for time and the confidence IM has a time less than or equal to the other metaphors also allowing H2 and H3. But one thing stands out is the results of time and confidence show that the behavior is very similar between DM and HOM, raising a possible difficulty comparable between them.

2) *Compare Distributions Results:* For comparison, the accuracy for the rainfall dataset had mean and variance, respectively: IM 0.03 & 0.19, DM 0.53 & 0.47, HOM 0.59 & 0.89. The p-value of KW was 0.000 (very high), since the p-value of W for the pairs brought a novelty IM had a significant difference with the other pairs (p-value of 0.000) but the DM-HOM pair did not have meaning that is very similar behavior.

In the taxi dataset, metaphors had mean and variance, respectively: IM 0.05 & 0.26, DM 0.59 & 0.42, HOM 0.53 & 0.45. The p-value of KW was 0.000 (very high), the W test had the same behavior in both datasets, only the DM-HOM pair did not show significance with a p-value of 0.093. These results give greater certainty that comparing IM distributions has a better performance confirming H1. For the time and the confidence in both IM datasets it has better performance allowing also to confirm H2 and H3.

3) *Find Extremes Results:* For the find extremes task, the accuracy for the rainfall dataset had mean and variance, respectively: IM 0.05 & 0.03, DM 0.25 & 0.43, HOM 0.26 & 0.45. The p-value of KW was 0.001 (quite high), whereas the p-value of W for the pairs showed that only the DM-HOM pair had no significance with a p-value of 0.632.

In the taxi dataset, metaphors had mean and variance, respectively: IM 0.21 & 0.36, DM 0.47 & 0.45, HOM 0.35 & 0.43. The p-value of KW was 0.000 (very high), the W test had the same behavior in both datasets, only the pair DM-HOM did not show significance with a p-value of 0.064. These results point to IM as the best performing metaphor confirming H1. For the time in the two datasets the times were very similar, which is justified by the need for users to think more before

answering to get an idea the KW of the taxi dataset registered a p-value of 0.641. The confidence in both datasets followed the same pattern already observed in other tasks where IM has a better performance against H2 and H3.

### C. Estimate the arithmetic mean

For the estimate the arithmetic mean task, the accuracy for the rainfall dataset had mean and variance, respectively: IM 1.17 & 1.55, DM 0.97 & 0.84, HOM 0.34 & 0.44. The p-value of KW was 0.001 (quite high), whereas the p-value of W for the pairs showed that only the IM-DM pair had significance with a p-value of 0.777. With the help of averages and variances it is possible to see that HOM had the best performance.

In the taxi dataset, metaphors had mean and variance, respectively: IM 0.16 & 0.24, DM 0.21 & 0.15, HOM 0.07 & 0.01. The p-value of KW was 0.003 (quite high), the W test showed that only the pair DM and HOM has a significant difference with a p-value of 0.000. These results show that once again IM cannot overcome the performance of other metaphors against H1. For the time we observed a greater expenditure of time to answer the questions and a much lower confidence in IM than we saw in the other tasks, which confirms the numerical results and contradict H2 and H3.

### D. Factors of The Task To Compare

In order to analyze the factor variables, the task of comparing distributions was chosen because it involves two regions simultaneously. To choose the regions to be tested, an analysis and manual selection of the two data sets was necessary to isolate the variables and select cases of regions that, when compared by the user, satisfy the analysis of the factors. At the end for the 3 factors and their variations (totaling 6 different scenarios) there is a set of 12 selected regions (2 to be tested in each scenario), for each data set. We initially studied only metaphors for DM and HOM.

1) *Results DM factors:* We tested the pairs of factors: long and short distance, high and low variance and large and small size. The results of W show that only the distance factor (long and short distance pair) in the points map has a significant difference (p-value: 0.007), confirming H5 only for distance. For the time H4, the lowest p-value of W found for the pairs in both datasets was not significant: 0.167. So H4 is false in this scenario.

2) *Results HOM factors:* In the case of HOM, we observed a behavior of factors similar to what we saw in the DM. The results of W show that only the distance factor (long and short pair) presents a significant difference in both datasets with (maximum p-value: 0.019), confirming H5 only for distance. For time H4, the highest p-value of W found for pairs in the two data sets was not significant: 0.173. So H4 as in DM cannot be accepted.

Thus, the results confirm that of the three techniques, the visualization with interaction was the one that generated the greatest accuracy in the user's responses to the tasks, with the exception of estimating the average confirming the H1

hypothesis. The mean confidence values allow us to affirm that the point map and the hypothetical maps had lower confidence measure indexes than the interaction technique as predicted in H2. With the exception of the task of estimating the average, H3 is proven because the user's response time in the interaction technique was less than or equal to the time recorded by the other techniques. H4 has not been proven, as the user's response time has not increased despite the geographic distance between the regions in the techniques. Regarding the factors of distance, size and variation of the distribution of each polygon, only distance proved to be a factor that negatively impacts the accuracy of the user for the answers of the tasks in the techniques of point map and hypothetical maps. Partially confirming H5, since this hypothesis predicted that all factors would have a significant difference in terms of error.

## V. CONCLUSION

The collection and analysis of the data indicated that the interaction technique has better performance in accuracy (relative error), time and confidence measure for the tasks of extracting value, comparing distributions, and finding extremes. Statistically, the results of the interaction technique showed significant differences in relation to the other techniques for accuracy, time and confidence measure. The only task that the interaction technique did not obtain good results was the task of estimating the average, in this case the hypothetical map technique has the best results. It was also observed that a common behavior of point map techniques and hypothetical maps in some tasks. Another fact also highlighted is the fact that the interaction technique automates the process of calculating and making the answer available in a simple way, which generates a certain distrust in users about the ease of obtaining answers. As limitations we highlight that because it is a voluntary study that needs some kind of follow-up and it takes an average of 30 minutes to answer the number of volunteers was limited to 66. Also because it was a voluntary recruitment it was not possible to balance the number of participants between profiles. The database used also needs to be further studied to be further explored. In addition, there is a need to investigate the influence of the number of regions on responses. Another limitation is to test whether the ability to quantify uncertainty actually leads to better decisions in practice. Personal factors of users can also represent limitations, users with natural personalities of insecurity. In addition, it is intended to study additional variations of metaphors, in particular to assess whether the point mapping technique with the organization of the distribution of points within polygons, following a pattern, positively influences the performance of tasks. Finally, new tests should also be conducted comparing the techniques with synthetic data.

## REFERENCES

- [1] J. Hullman, X. Qiao, M. Correll, A. Kale, and M. Kay, "In pursuit of error: A survey of uncertainty visualization evaluation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1,

- pp. 903–913, 01 2019. [Online]. Available: <https://doi.org/10.1109/tvcg.2018.2864889>
- [2] A. M. MacEachren, A. Robinson, S. Hopper, S. Gardner, R. Murray, M. Gahegan, and E. Hertzler, “Visualizing geospatial information uncertainty: What we know and what we need to know,” *Cartography and Geographic Information Science*, vol. 32, no. 3, pp. 139–160, 01 2005. [Online]. Available: <https://doi.org/10.1559/1523040054738936>
  - [3] C. Kinkeldy, A. M. MacEachren, and J. Schiewe, “How to assess visual communication of uncertainty? a systematic review of geospatial uncertainty visualisation user studies,” *The Cartographic Journal*, vol. 51, no. 4, pp. 372–386, 09 2014. [Online]. Available: <https://doi.org/10.1179/1743277414y.0000000099>
  - [4] J. Matejka and G. Fitzmaurice, “Same stats, different graphs,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI 17*. ACM Press, 2017. [Online]. Available: <https://doi.org/10.1145/3025453.3025912>
  - [5] C. A. de Lara Pahins, N. Ferreira, and J. Comba, “Real-time exploration of large spatiotemporal datasets based on order statistics,” *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2019. [Online]. Available: <https://doi.org/10.1109/tvcg.2019.2914446>
  - [6] S. Belia, F. Fidler, J. Williams, and G. Cumming, “Researchers misunderstand confidence intervals and standard error bars,” *Psychological methods*, vol. 10, no. 4, p. 389, 2005. [Online]. Available: <https://doi.org/10.1037/1082-989X.10.4.389>
  - [7] G. Cumming and S. Finch, “Inference by eye: confidence intervals and how to read pictures of data,” *American psychologist*, vol. 60, no. 2, p. 170, 2005. [Online]. Available: <https://doi.org/10.1037/0003-066X.60.2.170>
  - [8] M. Correll and M. Gleicher, “Error bars considered harmful: Exploring alternate encodings for mean and error,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2142–2151, 12 2014. [Online]. Available: <https://doi.org/10.1109/tvcg.2014.2346298>
  - [9] G. A. Rousselet, J. J. Foxe, and J. P. Bolam, “A Few Simple Steps to Improve the Description of Group Results in Neuroscience,” *European Journal of Neuroscience*, vol. 44, no. 9, pp. 2647–2651, 2016. [Online]. Available: <https://doi.org/10.1111/ejn.13400>
  - [10] T. L. Weissgerber, N. M. Milic, S. J. Winham, and V. D. Garovic, “Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm,” *PLoS Biology*, vol. 13, no. 4, p. e1002128, 2015. [Online]. Available: <https://doi.org/10.1371/journal.pbio.1002128>
  - [11] N. Ferreira, D. Fisher, and A. C. König, “Sample-oriented task-driven visualizations,” in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI 14*. ACM. ACM Press, 04 2014, pp. 571–580. [Online]. Available: <https://doi.org/10.1145/2556288.2557131>
  - [12] J. Hullman, P. Resnick, and E. Adar, “Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering,” *PLOS ONE*, vol. 10, no. 11, pp. 1–23, 11 2015. [Online]. Available: <https://doi.org/10.1371/journal.pone.0142444>
  - [13] J. Hullman, Y.-S. Kim, F. Nguyen, L. Speers, and M. Agrawala, “Improving comprehension of measurements using concrete re-expression strategies,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI 18*. ACM. ACM Press, 04 2018, pp. 1–12. [Online]. Available: <https://doi.org/10.1145/3173574.3173608>
  - [14] C. C. Freifeld, K. D. Mandl, B. Y. Reis, and J. S. Brownstein, “HealthMap: Global infectious disease monitoring through automated classification and visualization of internet media reports,” *Journal of the American Medical Informatics Association*, vol. 15, no. 2, pp. 150–157, Mar. 2008. [Online]. Available: <https://doi.org/10.1197/jamia.m2544>
  - [15] N. Ferreira, L. Lins, D. Fink, S. Kelling, C. Wood, J. Freire, and C. Silva, “BirdVis: Visualizing and understanding bird populations,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2374–2383, 12 2011. [Online]. Available: <https://doi.org/10.1109/tvcg.2011.176>
  - [16] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva, “Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2149–2158, 12 2013. [Online]. Available: <https://doi.org/10.1109/tvcg.2013.226>
  - [17] J. Poco, A. Dasgupta, Y. Wei, W. Hargrove, C. Schwalm, R. Cook, E. Bertini, and C. Silva, “Similarity explorer: A visual inter-comparison tool for multifaceted climate data,” in *Computer Graphics Forum*, vol. 33, no. 3. Wiley Online Library, 2014, pp. 341–350. [Online]. Available: <https://doi.org/10.1111/cgf.12390>
  - [18] D. N. Williams, T. Bremer, C. Doutriaux, J. Patchett, S. Williams, G. Shipman, R. Miller, D. R. Pugmire, B. Smith, C. Steed *et al.*, “Ultrascale visualization of climate data,” *Computer*, vol. 46, no. 9, pp. 68–76, 2013. [Online]. Available: <https://doi.org/10.1109/MC.2013.119>
  - [19] Z. Wang, N. Ferreira, Y. Wei, A. S. Bhaskar, and C. Scheidegger, “Gaussian cubes: Real-time modeling for visual exploration of large multidimensional datasets,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 681–690, 2016. [Online]. Available: <https://doi.org/10.1109/TVCG.2016.2598694>
  - [20] D. Moritz, D. Fisher, B. Ding, and C. Wang, “Trust, but verify: Optimistic visualizations of approximate queries for exploring big data,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI 17*. ACM. ACM Press, 05 2017, pp. 2904–2915. [Online]. Available: <https://doi.org/10.1145/3025453.3025456>
  - [21] M. Kay, T. Kola, J. R. Hullman, and S. A. Munson, “When (ish) is my bus? user-centered visualizations of uncertainty in everyday, mobile predictive systems,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI 16*. ACM Press, 05 2016. [Online]. Available: <https://doi.org/10.1145/2858036.2858558>
  - [22] H. Wickham and L. Stryjewski, “40 years of boxplots,” *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–17, 01 2011. [Online]. Available: <http://vita.had.co.nz/papers/boxplots.pdf>
  - [23] K. Potter, S. Gerber, and E. W. Anderson, “Visualization of uncertainty without a mean,” *IEEE computer graphics and applications*, vol. 33, no. 1, pp. 75–79, 01 2013. [Online]. Available: <https://doi.org/10.1109/MCG.2013.14>
  - [24] M. Fernandes, L. Walls, S. Munson, J. Hullman, and M. Kay, “Uncertainty displays using quantile dotplots or CDFs improve transit decision-making,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI 18*. ACM. ACM Press, 04 2018, pp. 1–12. [Online]. Available: <https://doi.org/10.1145/3173574.3173718>
  - [25] M. Ślusarski and M. Jurkiewicz, “Visualisation of spatial data uncertainty. a case study of a database of topographic objects,” *ISPRS International Journal of Geo-Information*, vol. 9, no. 1, p. 16, dec 2019. [Online]. Available: <https://doi.org/10.3390/ijgi9010016>
  - [26] Y. Zheng, W. Wu, Y. Chen, H. Qu, and L. M. Ni, “Visual analytics in urban computing: An overview,” *IEEE Transactions on Big Data*, vol. 2, no. 3, pp. 276–296, 09 2016. [Online]. Available: <https://doi.org/10.1109/tbdata.2016.2586447>
  - [27] R. Scholz and Y. Lu, “Uncertainty in geographic data on bivariate maps: An examination of visualization preference and decision making,” *ISPRS International Journal of Geo-Information*, vol. 3, no. 4, pp. 1180–1197, oct 2014. [Online]. Available: <https://doi.org/10.3390/ijgi3041180>
  - [28] A. Sarikaya and M. Gleicher, “Scatterplots: Tasks, data, and designs,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 402–412, 01 2018. [Online]. Available: <https://doi.org/10.1109/tvcg.2017.2744184>
  - [29] Z. Huang, Y. Lu, E. Mack, W. Chen, and R. Maciejewski, “Exploring the sensitivity of choropleths under attribute uncertainty,” *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2019. [Online]. Available: <https://doi.org/10.1109/tvcg.2019.2892483>
  - [30] W. Chen, Z. Huang, F. Wu, M. Zhu, H. Guan, and R. Maciejewski, “VAUD: A visual analysis approach for exploring spatio-temporal urban data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 9, pp. 2636–2648, 09 2018. [Online]. Available: <https://doi.org/10.1109/tvcg.2017.2758362>
  - [31] C. Keßler and E. Lotstein, “Animation as a visual indicator of positional uncertainty in geographic information,” in *Lecture Notes in Geoinformation and Cartography*. Springer International Publishing, 03 2018, pp. 365–382. [Online]. Available: [https://doi.org/10.1007/978-3-319-78208-9\\_19](https://doi.org/10.1007/978-3-319-78208-9_19)
  - [32] R. Amar, J. Eagan, and J. Stasko, “Low-level components of analytic activity in information visualization,” in *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*. IEEE, 2005. [Online]. Available: <https://doi.org/10.1109/infvis.2005.1532136>
  - [33] N. Andrienko and G. Andrienko, *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer-Verlag, 01 2006. [Online]. Available: <https://doi.org/10.1007/3-540-31190-4>