

# A Demo Manuscript as a ReviewerApp Study Case

Han Solo

Luke Skywalker

Obi-Wan Kenobi

## ABSTRACT

Geographic regression models of various descriptions are often applied to identify patterns and anomalies in the determinants of spatially distributed observations. These types of analyses focus on answering why questions about underlying spatial phenomena, e.g., why is crime higher in this locale, why do children in one school district outperform those in another, etc.? Answers to these questions require explanations of the model structure, the choice of parameters, and contextualization of the findings with respect to their geographic context. This is particularly true for local forms of regression models which are focused on the role of locational context in determining human behavior. In this paper, we present a visual analytics framework designed to support analysts in creating explanatory documentation that summarizes and contextualizes their spatial analyses. As analysts create their spatial models, our framework flags potential issues with model parameter selections, utilizes template-based text generation to summarize model outputs, and links with external knowledge repositories to provide annotations that help to explain the model results. As analysts explore the model results, all visualizations and annotations can be captured in an interactive report generation widget.

**Index Terms:** Spatial data analysis, geographically weighted regression, model explanation, visual analytics.

## 1 INTRODUCTION

Geographic models are often used to answer *why* questions regarding underlying spatial phenomena such as why do voter preferences generally exhibit strong spatial dependency. In order to answer such questions, researchers have developed sophisticated spatial modeling techniques and software, such as multiscale geographically weighted regression [21], to identify the determinants of the spatial patterns of data we observe in both the human and natural environments. However, these models generate output that needs local contextual information to be interpreted properly. Without such information, advanced spatial modeling techniques do not gain their full potential. While a variety of geovisual analytics systems [19] have been developed to explore spatial statistics with the help of interactive maps and narrative annotations, these applications tend to visualize only basic statistical results, and, to our knowledge, none of them explain spatial modeling results.

In this paper, we present a visual analytics framework designed to support spatial data modeling, analysis and reporting. Our approach is inspired by research in explainable machine learning [21] and a combination of narrative visualization [18] and storytelling techniques [7]. The framework provides multiple types of explanation support throughout the spatial analysis pipeline. In the model calibration stage, our framework interprets the functionality of the spatial model configuration and recommends parameter settings. Then, the framework summarizes and explains model outputs by adopting template-based text annotations, linking with external knowledge repositories to provide relevant contextual information. All visualizations, model results, and annotations can be captured

in an interactive report authoring widget, enabling analysts to generate documentation that explains their spatial analyses.

## 2 RELATED WORK

Our work focuses on facilitating spatial analysis by explaining the outputs of spatial modeling with contextual information through narrative visualization. In this section, we review related work on geographic analysis, narrative visualization, and model explainability.

### 2.1 Geographic Analysis

Various spatial prediction models [21] and spatial data analysis tools [24] have been developed to support geographic analysis. In this work, we focus on two local spatial models widely used in spatial analysis: Geographically Weighted Regression (GWR) [1], and its recent extension, Multiscale Geographically Weighted Regression (MGWR) [2]. GWR extends the classical linear regression model [20] by capturing spatial heterogeneity with influence spreading over the space in a constant scale. MGWR further improves GWR models where local influences are modeled in different spatial windows.

Support for a variety of spatial models is integrated into the most widely used geographic information systems (GIS). However, these systems and notebooks do not support integrated external knowledge sources to help contextualize models [6]. They also require analysts to judiciously choose the proper tools at every stage of the analytical pipeline to build a well-trained model. This process necessitates a solid foundation in geographical concepts and can be time-consuming, even for seasoned analysts. Our goal is to enhance the spatial modeling process through a deliberately designed user-friendly workflow and automatically generated explanatory narratives guided by domain experts to promote the understanding of local contextual information from models. Our work focuses on improving model explainability, enabling interactive report authoring, and supporting contextualization through narrative generation. This is important in local models such as GWR and MGWR because the main output from such models is a set of local parameter estimates from each process being modeled, and the spatial variation in these estimates needs to be explained in terms of the contextualized environment of each location by providing additional contextual information for users.

### 2.2 Narratives and Annotations

The design space of narrative visualization [18, 5] has been widely explored in the visualization community and a diverse set of storytelling and annotation methods have been developed to reveal observations in data and to convey key messages to an audience. Such narrative visualization techniques have been adopted by a variety of geovisual analytics applications [16] that have integrated story authoring tools with spatial data visualizations. For example, NewsViews [3] generates interactive maps with narrative annotations automatically from a given news article. Latif and Beck [10] introduce a bivariate map design that integrates template-based text annotations, and they later extend their work to investigate the interplay of text and visualization in geographic storytelling [11].

Many of these systems utilize annotations to enhance the visual narratives, and Kosara and Mackinlay [7] emphasize the importance of annotations for facilitating storytelling in visualization. Recent annotation work [15] has focused on the semi-automatic creation

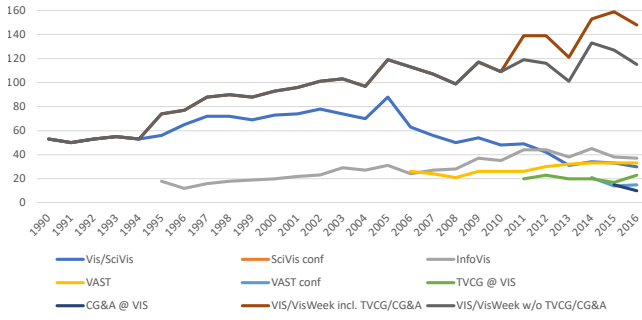


Figure 1: A visualization of the 1990–2016 publication data.

of presentation-like storytelling visualizations and explored mechanisms for automatically generating annotations by integrating deep-learning feature extraction techniques with a natural language generation process [9]. Annotations in such systems are used to describe salient patterns and facilitate storytelling [23].

### 2.3 Model Explainability

Narrative visualization is directly related to the concept of explainability, where the visualization authors seek to couple images and text to explain an underlying data analysis. With respect to model explainability, the visual analytics community has developed a variety of systems to support the interactive explanation of machine learning models (e.g. [22, 12]). Several model-independent approaches, e.g. EnsembleMatrix [25] and RuleMatrix [13], focus on the classifier’s input-output behaviors to provide insight into the model classification results. EnsembleMatrix provides a visual summary of the model outputs, RuleMatrix uses a matrix-based visualization to explain classification results, and Prospector [8] explores the relationship between feature values and predictions by using partial dependence diagnostics.

Most closely related to our work are the techniques that explain models from feature-level observation. Mühlbacher and Piringer [14] facilitated feature selection and optimization in regression models by partitioning the feature space into disjoint regions for visualization. Sedlmair et al. [17] proposed an abstract conceptual framework to discuss the visual parameter space analysis problems independent of the application domain. Goodwin et al. [4] extended visual parameter space analysis to the spatial domain, enabling the exploration of correlations between multiple variables that vary geographically at different spatial scales. While these techniques focus on explanations for domain experts, our work is designed to support explanations to experts and support their use of external information to contextualize these relationships and communicate their findings to a general audience. Our choice of text templates, as opposed to large language models, for narrative generation is to support control for reliability and reproducibility. All text generated must have a verifiable source and the generation of text for the analysis should always return the same results to ensure that the resulting analyses are not subject to misinformation.

### 3 STATISTICAL EVALUATION

A summary of the ratings on each evaluative dimension across conditions is shown in Table 1. The scores indicated that though drafts from the tool condition have passing quality, they still fall short from excellence.

In particular, a repeated measures ANOVA was conducted to investigate the influence of three conditions on the six measures. Our findings revealed a significant main effect on the score of Factor 1 ( $F(2, 46) = 4, p < .01$ ). Here is some text with statistical tests  $t(123) = .45, p = 0.65$ . Sometimes, the reported statistics

	df1	df2	Value	p-value
F	2	46	6.36	< 0.01
t		123	0.45	0.65
F	0.47	1.73	0.9	0.756

Table 1: Statistical testing results.

are so internally inconsistent that they can lead to a decision error  $F(0.47, 1.73) = 0.9, p = 0.756$ .

### REFERENCES

- [1] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4):281–298, 1996. doi: 10.1111/j.1538-4632.1996.tb00936.x 1
- [2] A. S. Fotheringham, W. Yang, and W. Kang. Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*, 107(6):1247–1265, 2017. doi: 10.1080/24694452.2017.1352480 1
- [3] T. Gao, J. Hullman, E. Adar, B. Hecht, and N. Diakopoulos. NewsViews: An automated pipeline for creating custom geovisualizations for news. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 3005–3014, 2014. doi: 10.1145/2556288.2557228 1
- [4] S. Goodwin, J. Dykes, A. Slingsby, and C. Turkay. Visualizing multiple variables across scale and geography. *IEEE Transactions on Visualization and Computer Graphics*, 22:599–608, 1 2016. doi: 10.1109/TVCG.2015.2467199 2
- [5] J. Hullman and N. Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2231–2240, 2011. doi: 10.1109/TVCG.2011.255 1
- [6] J. Hutt. Artificial reference paper 4. *Fake Journal* 2, 10(2):1–4, 2025. doi: 10.1080/13658816.2020.1720692 1
- [7] R. Kosara and J. Mackinlay. Storytelling: The next step for visualization. *Computer*, 46(5):44–50, 2013. doi: 10.1109/MC.2013.36 1
- [8] J. Krause, A. Perer, and K. Ng. Interacting with predictions: Visual inspection of black-box machine learning models. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 5686–5697, 2016. doi: 10.1145/2858036.2858529 2
- [9] C. Lai, Z. Lin, R. Jiang, Y. Han, C. Liu, and X. Yuan. Automatic annotation synchronizing with textual description for visualization. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1–13, 2020. doi: 10.1145/3313831.3376443 2
- [10] S. Latif and F. Beck. Interactive map reports summarizing bivariate geographic data. *Visual Informatics*, 3(1):27–37, 2019. doi: 10.1016/j.visinf.2019.03.004 1
- [11] S. Latif, S. Chen, and F. Beck. A deeper understanding of visualization-text interplay in geographic data-driven stories. *Computer Graphics Forum*, 40(3):311–322, 2021. doi: 10.1111/cgf.14309 1
- [12] Y. Lu, R. Garcia, B. Hansen, M. Gleicher, and R. Maciejewski. The state-of-the-art in predictive visual analytics. *Computer Graphics Forum*, 36(3):539–562, 2017. doi: 10.1111/cgf.13210 2
- [13] Y. Ming, H. Qu, and E. Bertini. RuleMatrix: Visualizing and understanding classifiers with rules. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):342–352, 2018. doi: 10.1109/TVCG.2018.2864812 2
- [14] T. Mühlbacher and H. Piringer. A partition-based framework for building and validating regression models. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):1962–1971, 2013. doi: 10.1109/TVCG.2013.125 2
- [15] D. Ren, M. Brehmer, B. Lee, T. Höllerer, and E. K. Choe. ChartAccent: Annotation for data-driven storytelling. In *IEEE Pacific Symposium on Visualization*, pp. 230–239, 2017. doi: 10.1109/PACIFICVIS.2017.8031599 1

- [16] A. Satyanarayan and J. Heer. Authoring narrative visualizations with ellipsis. *Computer Graphics Forum*, 33(3):361–370, 2014. doi: 10.1111/cgf.12392 1
- [17] M. Sedlmair, C. Heinzl, S. Bruckner, H. Piringer, and T. Möller. Visual parameter space analysis: A conceptual framework. *IEEE Transactions on Visualization and Computer Graphics*, 20:2161–2170, 12 2014. doi: 10.1109/TVCG.2014.2346321 2
- [18] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010. doi: 10.1109/TVCG.2010.179 1
- [19] L. Skywalker and O.-W. Kenobi. Artificial reference paper 2. *Fake Journal 2*, 10(2):1–4, 2025. doi: 10.1080/13658816.2020.1720692 1
- [20] D. Smeesters and J. Liu. Retracted: The effect of color (red versus blue) on assimilation versus contrast in prime-to-behavior effects. *Journal of Experimental Social Psychology*, 47(3):653–656, 2011. doi: 10.1016/j.jesp.2011.02.010 1
- [21] H. Solo and O.-W. Kenobi. Artificial reference paper 1. *Fake Journal 1*, 34(7):1378–1397, 2025. doi: 10.1080/13658816.2020.1720692 1
- [22] H. Solo and L. Organa. Artificial reference paper 5. *Fake Journal 2*, 10(2):1–4, 2025. doi: 10.1080/13658816.2020.1720692 2
- [23] A. Srinivasan, S. M. Drucker, A. Endert, and J. Stasko. Augmenting visualizations with interactive data facts to facilitate interpretation and communication. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):672–681, 2019. doi: 10.1109/TVCG.2018.2865145 2
- [24] M. Takatsuka and M. Gahegan. GeoVISTA Studio: A codeless visual programming environment for geoscientific data analysis and visualization. *Computers and Geosciences*, 28(10):1131–1144, 2002. doi: 10.1016/S0098-3004(02)00031-6 1
- [25] J. Talbot, B. Lee, A. Kapoor, and D. S. Tan. EnsembleMatrix: Interactive visualization to support machine learning with multiple classifiers. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1283–1292, 2009. doi: 10.1145/1518701.1518895 2