# Analysing Spatio-Temporal Autocorrelation with LISTA-Viz

Frank Hardisty, Alexander Klippel
*GeoVISTA Center*
*Department of Geography*
*302 Walker Building*
*John A. Dutton e-Education Institute*
*2217 Earth & Engineering Sciences Building*
*The Pennsylvania State University, University Park, PA, 16801, USA*
hardisty@psu.edu
klippel@psu.edu

**Abstract** — Many interesting analysis problems (for example, disease surveillance) would become more tractable if their spatio-temporal structure was better understood. Specifically, it would be helpful to be able to identify autocorrelation in space and time simultaneously. Some of the most commonly used measures of spatial association are LISA statistics, such as the Local Moran's I or the Getis-Ord Gi*, however these have not been applied to the spatio-temporal case (including many time steps) due to computational limitations. We have implemented a spatio-temporal version of the Local Moran's I, and claim two advances: First, we exploit the fact that there are a limited number of topological relationships present in the data to make Monte Carlo estimation of probability densities computationally practical, and thereby bypass the "curse of dimensionality". We term this approach "spatial memoization". Second, we developed a tool (LISTA-Viz) for interacting with the spatio-temporal structure uncovered by the statistics which contains a novel coordination strategy. The potential usefulness of the method and associated tool are illustrated by an analysis of the 2009 H1N1 pandemic, with the finding that there was a critical spatio-temporal "inflection point" at which the pandemic changed its character in the United States.
Keywords—Spatio-temporal autocorrelation, Monte Carlo simulation, Moran's I

## Introduction

Effective methods for exploring space-time structure are needed to make sense of phenomena which contain both spatial and temporal referents (Andrienko et al. 2001). There is a particular need for methods which can not only represent spatio-temporal data visually and interactively, but can offer analytic judgements, by finding which spatio-temporal patterns in the data are significant in the statistical sense (Andrienko et al. 2007), also known as geovisual analytics (Kraak 2008). An example of a pressing spatio-temporal problem which could benefit from a geovisual analytics approach is that of understanding the spread of infectious disease.

We have therefore created a method for exploring spatio-temporal structure using an extension to one of the most popular statistical methods for spatial autocorrelation, the local Moran's I (Anselin 1995). This could be thought of as one of a class of LISTA statistics (for Local Indicators of Spatio-Temporal Association, after Anselin's LISA). We suggest that to make the computational problem of finding significance levels practical for larger data sets, the topology of the relations between the places be exploited.

We implemented this method in Java, and created an interactive tool, called LISTA-Viz, which exposes the statistical method in a manner that makes it easy to use, and which offers tight integration with many other tools in the GeoViz Toolkit (Hardisty 2005, Hardisty *et al.* Accepted) application. The user interface for LISTA-Viz is shown in Figure 1. This integration means that the spatio-temporal structures uncovered by LISTA-Viz can be analysed visually and computationally by the many other views and methods offered in the GeoViz Toolkit.

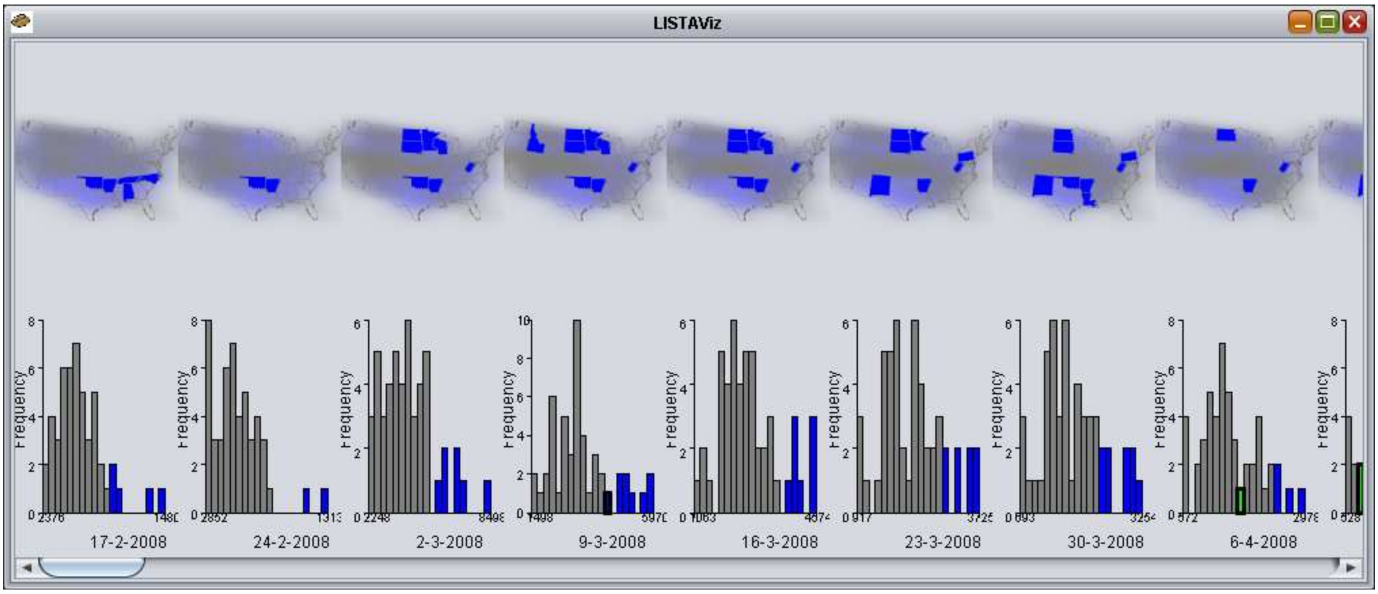


**Figure 1. LISTA-Viz user interface, showing space–time clusters of ILI for selected weeks in 2008 in the United States. In each map–histogram pair, those states that are part of a statistically significant space–time cluster of ILI rates are highlighted.**

The computational strategy described here is one of enumerating topological relationships and estimating the probability density function for each one, and using those with the observed patterns. This allows for much better performance, with a speed-up of two orders of magnitude or more for many common analysis cases. This increased efficiency allows the LISTA-Viz tool to offer visual analytic capability, whereby the analyst can interactively explore the data, and receive feedback on the statistical significance of observed spatio-temporal patterns. Fortunately, this increased performance does not carry any penalty in terms of accuracy; the results are functionally identical to results using the “naive” Monte Carlo approach. This is true because the method is essentially a caching one: it stores the results of intermediate computations and uses them where the same result is called for, thereby reducing the number of steps needed to calculate a result at the cost of increasing memory usage. We call this approach “spatial memoization”, after the memoization optimization strategy that is employed in computing (particularly functional computing).

In the remainder of this paper we describe some related work, the methods we used, our results, and then conclusions and future research.

## Related Work

There is a substantial body of work on both formal methods for delineating spatio-temporal structure, and on interactive visual methods for examining space-time. Some of the more important statistical methods are spatial statistics (Cliff *et al.* 2009), scan statistics (Kulldorff *et al.* 1995), geographically weighted regression(Brunsdon *et al.* 1998), Markov chain Monte Carlo (Brooks 1998), as well as Bayesian methods (Lawson 2006). Influential theory on visual analytics for spatio temporal problems includes analyses of how software can support these analytic tasks (Andrienko *et al.* 2008b).

### *Spatio-Temporal Autocorrelation*

The study of spatial autocorrelation has a substantial body of theory behind it now (Moran 1950, Cliff *et al.* 1973, Getis 2008, Cliff *et al.* 2009). Two of the most popular are LISA statistics such as Moran’s I and Scan statistics. LISA statistics are attractive in that they are able to describe areas of both positive and negative spatial autocorrelation. LISA statistics have the disadvantage of being difficult to compute probability values if large numbers of variables are attempted to be analysed, using either exact methods (Leung *et al.* 2003) or Monte Carlo simulation. Moran’s I in particular has been critiqued for exploratory use (Li *et al.* 2007). Saddle-point methods (Tiefelsdorf 2002) can be useful to reduce computation time for Local Moran's I if their preconditions are met. Normal approximations have been shown to lead to errors for local Moran's I (Bivand *et al.* 2009). While Moran's I does not detect spatial clusters *per se*, it can be usefully used in this manner by decomposing the results into clusters of high and low spatial autocorrelation (Zhang *et al.* 2007). Additionally, here are a number of methods that are able to delineate some forms of spatio-temporal structure.

Spatial scan statistics (Kulldorff *et al.* 1995) are popular because they can automatically identify clusters in space and time. However, they cannot detect moving clusters, and also suffer from computational complexity with large numbers of observations. GWR (Geographically Weighted Regression) is another promising avenue for spatio-temporal data analysis (Demsar *et al.* 2008), because of GWR's ability to distinguish how multivariate influences themselves vary across space and time.

***Interactive Exploration of Spatio-Temporal Structure***

Past work which the LISTA-Viz tool draws on includes two main branches: theory and software. A body of theory on how geovisualization and visual analytics for spatio-temporal phenomena should work has been developed (Andrienko *et al.* 2007). Concurrently, there are many spatio-temporal software environments that have blazed a trail into geovisual exploration and geovisual analytics (Guo *et al.* 2006, Demsar *et al.* 2008).

A typology of tasks and data types provides useful guidance (Andrienko *et al.* 2001). Dynamics, movement, and change have been identified as key issues in spatio-temporal analysis (Andrienko *et al.* 2008b). A recent special issue of CaGIS provides some of the most current thinking on the best means of geovisual space-time analytics (Andrienko *et al.* 2009).

Interactive dynamic maps and graphics have been seen as promising modes of inquiry for some time now (Andrienko *et al.* 2001). CommonGIS was one of the first systems to allow spatio-temporal data mining (Andrienko *et al.* 2004). The GeoDA package, which is free but not open source, provides easy-to-use software for exploring spatial dependency, and outstanding facilities for defining spatial weights

(Anselin *et al.* 2006). Evacuation scheduling is another interesting application of space-time modelling (Andrienko *et al.* 2008a).

Animated maps have been explored as vehicles for showing temporal data using time itself for decades now (Thrower 1961, Campbell *et al.* 1990, MacEachren *et al.* 1991, Koussoulakou *et al.* 1992). Space-time cubes provide one of the most popular means of visualizing spatio-temporal information (Huisman *et al.* 2009), which seems natural since they can depict both space and time directly. Time-waves are a new promising technique for depicting space-time(Leung *et al.* 2003), while good results have also been achieved using treemaps for the purpose of space-time representation as well (Slingsby *et al.* 2008, Slingsby *et al.* 2009). In sum, a variety of spatial statistical (and spatio-temporal statistical) methods have been developed on the one hand, and a body of theory has been developed on how analytical reasoning about spatially and temporally referenced phenomena can be better enabled by visual analytics on the other. However, the lack of computationally efficient statistical methods has hampered the use of spatio-temporal statistics in visual analytics applications. LISTA-Viz was developed as an approach and tool that allows spatio-temporal statistics to be used in an interactive manner, and therefore be incorporated in visual analytics.

## Methods

The contribution of LISTA-Viz consists of two parts, which we describe in turn. First we describe the LISTA-Viz computational methods, then how the interactive tool was realized in Java. The computational methods and the interactive tool could be realized separately in other environments, but they have advantages when used together, as described below. The computational method is described by first relating the steps in the method in detail, and then relating why this method will produce identical results at a computational savings, when compared to the naive Monte Carlo approach. The tool implementation is described by relating the programming architecture of both the LISTA-Viz tool and how it can fit into a larger computational ecosystem.

***Using Topology to Overcome the "Curse of Dimensionality"***

The core computational advance that we present here is a method for making Monte Carlo calculation of significance values practical for large numbers of observations, and large numbers of time slices. We describe first the "naive" Monte Carlo procedure, and then the alternative that we are proposing which exploits topology. Our approach enables saving up to orders of magnitude of computational time, with a cost of slightly increased memory usage. We contend that using topology will enable analysis not only of space-time Moran's I but also will be a viable strategy to render computationally practical other analysis methods that use Monte Carlo estimation of probability density functions for lattice data (commonly data for enumerated units such as states or provinces).

We call this approach "spatial memoization". In computing, the term "memoization" refers to an optimization strategy whereby the results of calling functions are stored, and then used when identical inputs are presented. It is related to caching, but is more specific in that it excludes techniques like page replacement and buffering. Spatial memoization refers to tailoring a memoization strategy for spatial inputs.

The naive approach to Monte Carlo calculation of local Moran's I values is as follows:

1. Create a list of each datum in the lattice (for example, census districts) and normalize it into standard deviation units, and a list for each time step.
2. Randomly re-order the lists and calculate the local Moran's I value for each location *n* times, where *n* is the number of Monte Carlo repetitions, recording the results each time. The space-time extension of the local Moran's I is straightforward, one simply adds the temporal neighbours in as part of the weights matrix.
3. Compare the local Moran's I values in each place with the Monte Carlo derived values, and report what proportion of the Monte Carlo values were larger than the given value. This is the equivalent of a p-value.

The difficulty with this method is the time required for calculation, as well as the amount of memory needed. The calculation time exponentially increases with the number of time steps. There is also the implicit difference that many time slices are likely to be examined (each with temporal neighbours) when looking for space-time structure, which implies that the calculations need to be repeated for each time step, whereas in the traditional spatial case, only a single variable at a time is examined. The need to evaluate many variables in turn makes it even more desirable to find performant methods of analysis.

The spatial memoization approach (which uses topology) is as follows:

1. Create a list of each datum in the lattice (for example, census districts) and normalize it to standard deviation units, and a list for each time step (this step is identical to that in the previously described procedure).
2. Enumerate the topological spatial relationships present in the data set. For example, it may be found that there are observations with no neighbours, with one neighbour, and so on.
3. For each observed spatial relationship found in step 2 (number of neighbours), perform a Monte Carlo simulation with the normalized data by randomly picking (without replacement) from the lists created in step one *n* times. Record the probability density for that relationship.
4. Compare each observed space-time local Moran's I value with the appropriate probability density and calculate a probability value (or p-value) from it.

This approach requires Monte Carlo repetitions to be performed for each topological relationship, rather than for each observation. Therefore, the advantage of this method should increase with the size of the data set, because the number of different spatial relationships (or topologies) is not likely to increase at the same rate that the number of observations does.

The numerical results of the LISTA-Viz approach are the same as those obtained following the traditional approach (within the limits of Monte Carlo simulation). This is true because the naive approach is essentially performing the same calculation repeatedly. The LISTA-Viz approach records, or caches, the results of calculations and stores them as an estimated probability density function that can be quickly called. We test the validity of this claim in the Results section below. The cost is that the estimated probability density functions must be stored, and therefore more space in memory is required. For the most common analysis cases, the increased performance would make the memory cost well worth paying.

***The LISTA-Viz Software Component***

The computational methods are realized using the standard Java libraries. The code for the methods is publicly available from the GeoViz Toolkit source code repository, and is released under the Library General Public License (LGLP). These methods are referenced from a component that leverages the map and histogram components of the GeoViz Toolkit. Also, the coordination strategies are used so that the results of the space-time Moran's I are able to be broadcast to all other components of the GeoViz Toolkit (See Figure 2). Note that a different coordination strategy was used internally within the LISTA-Viz component. This forms a part of the novel visual analytics strategy used by the LISTA-Viz component: the use of the introspective observer coordination (Hardisty *et al.* accepted) both internally to the visualization component and externally to automatically connect it to a larger visual analysis environment.

The introspective observer coordination strategy uses automatic interrogation of a component's capabilities, and matches that with the coordinator software design pattern. This allows components to be flexibly combined, which is essential to visual analytics, in which both humans and computers are creating relevant changes in the analysis system.

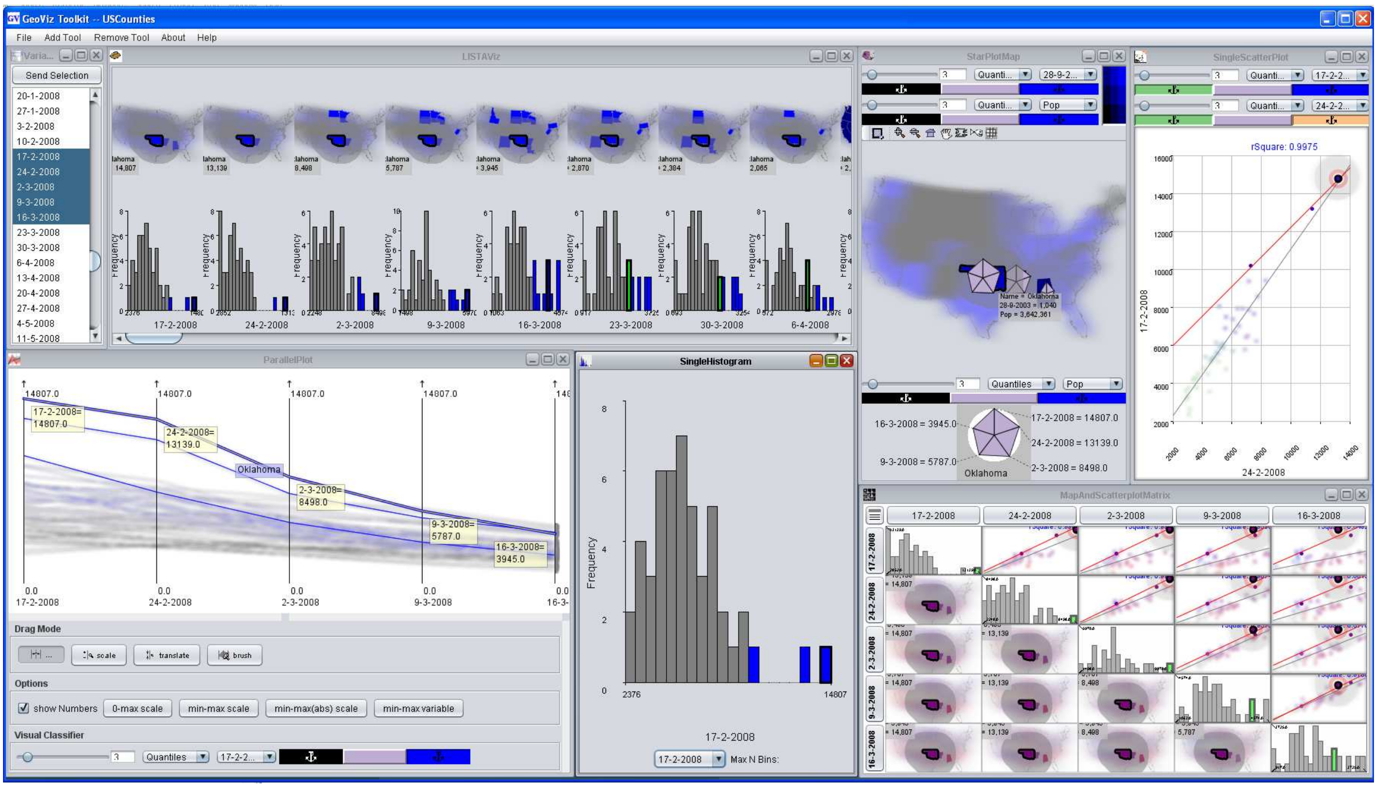


**Figure 2. LISTA-Viz component within the GeoViz Toolkit. The use of two coordinators facilitates visual analysis. An internal coordinator manages selections inside the LISTA-Viz component, while all of the components within the GeoViz Toolkit (including LISTA-Viz) are coordinated with the others, using a coordinator external to the LISTA-Viz component.**

In the case of the LISTA-Viz component, we created two coordinators: one internal, and one external. The internal coordinator allows the map-histogram pairs to receive the appropriate spatial and numerical data from external sources, and manages common visual and computational settings. The external coordinator is inside the GeoViz Toolkit, and it enables the space-time structures identified by the LISTA-Viz component to be broadcast to all other visual and computational components.

We describe here the process of how the realized LISTA-Viz component works in the hand of the user. The software can be used as either a stand-alone tool, or as part of the GeoViz Toolkit. In the case of LISTA-Viz being used alone, the spatial data must be supplied as a command-line argument, while if it is being used as a part of the GeoViz Toolkit, the data reading facilities provided by the GeoViz Toolkit can be used.

After the data is loaded, the system calculates the space-time Moran's I statistic, and selects those observations that meet the significance level that has been selected. At that point, each map-scatterplot pair will have a different set of observations selected, as seen in Figure 2. The user can then broadcast the set of selected observations associated with a particular map-scatterplot pair, which have been found to be part of a space-time cluster, to all listening components. In this manner these observations can be analyzed in other tools and views.

We combine the LISTA-Viz computational approach, using spatial memoization, and the LISTA-Viz software component, with its novel coordination strategy, because the two better enable geovisual analytics than either would alone. One of the goals of developing the computational approach was to allow for interactive detection of space-time structures in data, such as clusters. Interactivity demands a flexible coordination strategy that allows users to start computational processes, and interact with their results, without interfering with other simultaneous computational and visual processes. Multi-threading allows this, but it is difficult to create multi-threaded programs in a robust, extensible manner. The LISTA-Viz coordination strategy allows the computational processes to proceed independently, and extensibly, thus contributing to interactivity and to the geovisual analytic process.

## Results

Applying the newly developed LISTA-Viz method and tool allows us to show progress towards our goal of interactive exploration of spatio-temporal structure. . We were able to reduce computational time (compared with the naive Monte Carlo approach) so that it is feasible to use the Local Moran's I statistics with large data sets. To create an interactive environment fast computational methods are required. LISTA-Viz performance has been increased to allow for visual analytics, whereby the software user can interactively identify data of interest, and the tools can respond by identifying which parts of spatio-temporal structures in the data are statistically significant. Also, LISTA-Viz was able to identify interesting space-time clusters of influenza like illnesses, or flu outbreaks. We describe these two outcomes in more detail below.

***Computational results***

The computational advantages are considerable. Both the “naive” Monte Carlo approach and the topological, spatial memoization one take polynomial time to compute, that is, they take O(NSTM) calculations, where N is the number of observations in the data set, S is the number of time steps, T is the number of temporal neighbours to be examined, and M the number of Monte Carlo simulations. However, using the spatial memoization approach described above drastically reduces the size of N, with the advantage growing larger with the size of the data set. These results are shown in Table 1 below, which assumes 20 time steps of data, a temporal neighbourhood of 3, and 10,000 Monte Carlo simulations for both data sets, in addition to the factors presented in the table. In these tests, we used two data sets: U.S. states (n = 51), and U.S. counties (n = 3105).

| Observations | Topological Cases | Naïve approach | Spatial Memoization Approach | Expected Computational Savings |
|---|---|---|---|---|
| 51 | 9 | 3.E+07 | 5.E+06 | 5.6 |
| 3105 | 14 | 2.E+09 | 8.E+06 | 217.9 |

**Table 1: Expected Computational Savings of the Spatial Memoization approach .**

We then conducted tests using these two spatial data sets (U.S. states and U.S. counties). The results are shown below in Table 2. We found that the expected savings did not demonstrate themselves to the expected degree in small data sets, however a convincing speed-up was obtained for the larger data set. In the smaller data set (N = 51), calculating space-time Moran’s I values for a single slice using the naive approach took 314.5 milliseconds on average. Using the spatial memoization approach, it took 142.9 milliseconds, a speed-up of 2.2 times. While finding space-time Moran’s I autocorrelation in a data set that used the coterminous U.S. counties (N = 3105), it took an average of 16,820.25 milliseconds to find p-values for a single time slice, using a temporal neighbourhood of 3. Using the topology-based approach, the average time of calculation was reduced to 146.05, an improvement of two orders of magnitude. These experimental results confirms that the theoretical savings outlined above are realized in actual analysis scenarios, particularly in larger data sets. However, the speed-up is significant even for smaller data sets, especially those with a high temporal dimensionality.

| Observations | Topological Cases | Naïve approach, one slide, miliseconds | Spatial Memoization Approach, one slice, miliseconds | Computational Savings |
|---|---|---|---|---|
| 51 | 9 | 314.5 | 142.9 | 2.2 |
| 3000 | 14 | 16,820 | 146.05 | 115.1678 |

Table 2: Observed Computational Savings of the Spatial Memoization approach .

The spatial memoization approach will work will if and only if the number of topological cases or instances is smaller than the number of observations. This is the case for most data sets, for example, using a queen's case definition of contiguity, instances of U.S. states were found with 0 neighbours through 8 neighbours inclusive, thus 9 cases. Instances of U.S. counties were observed that had 0 to 11 neighbours inclusive, and additionally instances of counties with 13 and 14 neighbours existed, thus 14 topological cases. The ratio of the number of topological cases to the number of total cases $\frac{n(t)}{n}$ thus represents the computational savings.

***Case study***

We illustrate the advantages of this method and the LISTA-Viz tool with a case study of influenza incidence using the Google Flu Trends data. Influenza, or flu, is of interest for multiple reasons. First, there is the fact that influenza epidemics and pandemics have posed severe global health risks; the pandemic of 1918 killed an estimated 20 to 50 million people (Tumpey *et al.* 2005). Related to this is the fact that many public health organizations have recently become more interested in disease surveillance and have devoted increased resources to the task. There has been some success in the area (Huang *et al.* 2010), but the most popular tools (such as SaTScan) are not interactive. This raises the possibility that spatio-temporal analytics, in the form of interactive spatio-temporal cluster detection could improve public health by improving disease surveillance (Hardisty *et al.* 2008).

Reports by hospitals of the number of people admitted with influenza-like illnesses (ILI) is often the best indication available of where flu outbreaks are happening. ILI numbers are provided by the Centers for Disease Control (CDC) in the United States, but only to the level of nine areas of the United States, which limits their applicability to spatial analysis. The data in this study are those provided by the Google Flu Trends project, we thank Google for making this data publically available. The data is the estimated number of people per 100,000 in the relevant time period with influenza like illnesses (ILI). The estimated rates are provided for each state for each week from September 28$^{th}$, 2003 to the current time. (Previously, Google Flu had estimated the percentage of hospital admissions that were due to ILI.) We used the data set as provided by Google Flu Trends with no omissions. The H1N1 (Swine Flu) pandemic was a significant public health event in 2009.

Figure 3 illustrates the results of using the stand-alone LIST-Viz tool with this data. Examining the critical period in late September and early October when flu illnesses were peaking (including H1N1), and schools were closing, we can clearly see two separate significant patterns, one which is focussed on Texas and Louisiana, and a later one which includes states that surround Texas and Louisiana as well as Washington state.

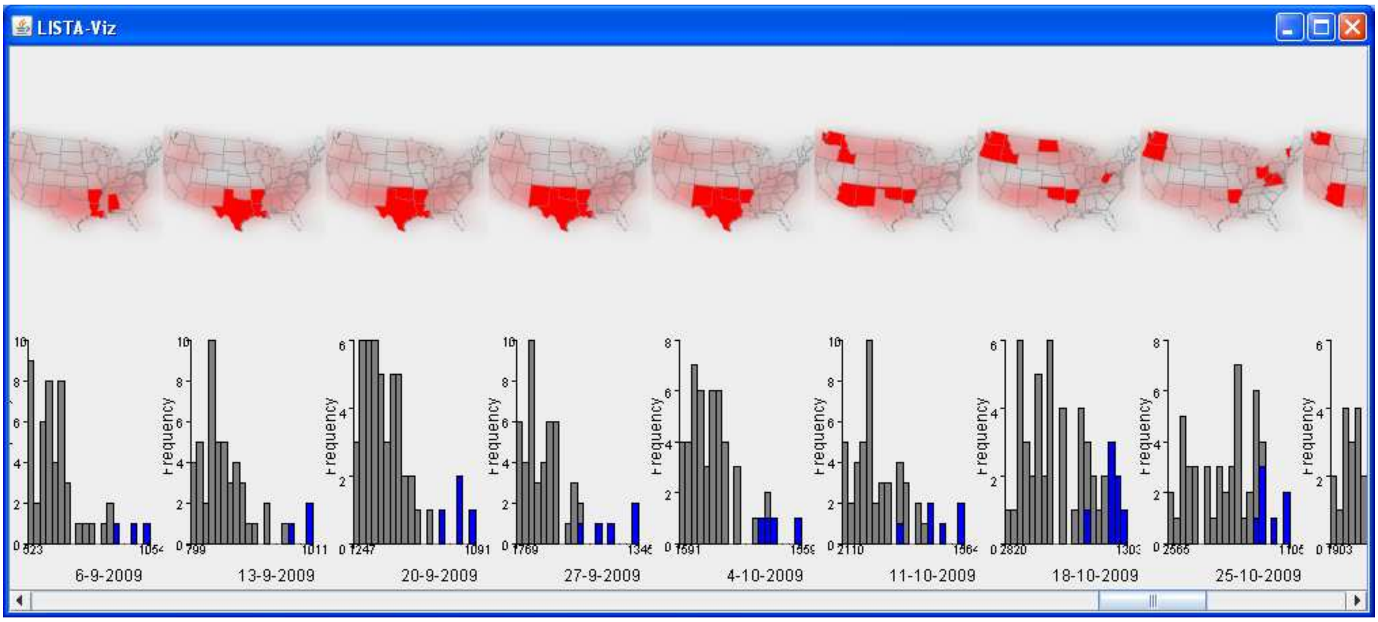


**Figure 3: LISTA-Viz showing spatio-temporal patterns in influenza incidence during the height of the 2009 H1N1 pandemic. Each map-histogram pair shows the ILI rate for a week during the period of highest infection rates. Those states within space-time clusters (as determined by the LISTA-Viz algorithm) are highlighted in each week.**

We then brought this data into the full GeoViz Toolkit for further analysis (see Figure 4). Examining the critical period when H1N1 infections were peaking, the interactive nature of these tools comes to the fore. By clicking on any of the periods in the LISTA-Viz tool, the significant observations are broadcast to the other visual-statistical tools. This allows us to see the trends from other points of view, and for example, to verify that the states that were part of a statistically significant cluster in late September were trending downwards at that point, as shown by the r-squared value of -0.539, as opposed to the rest of the states, which were still trending upwards. This shows that this week was truly an “inflection point” for the H1N1 pandemic of 2009, and that Texas and its surrounding states were the epicentre of this “inflection point”.

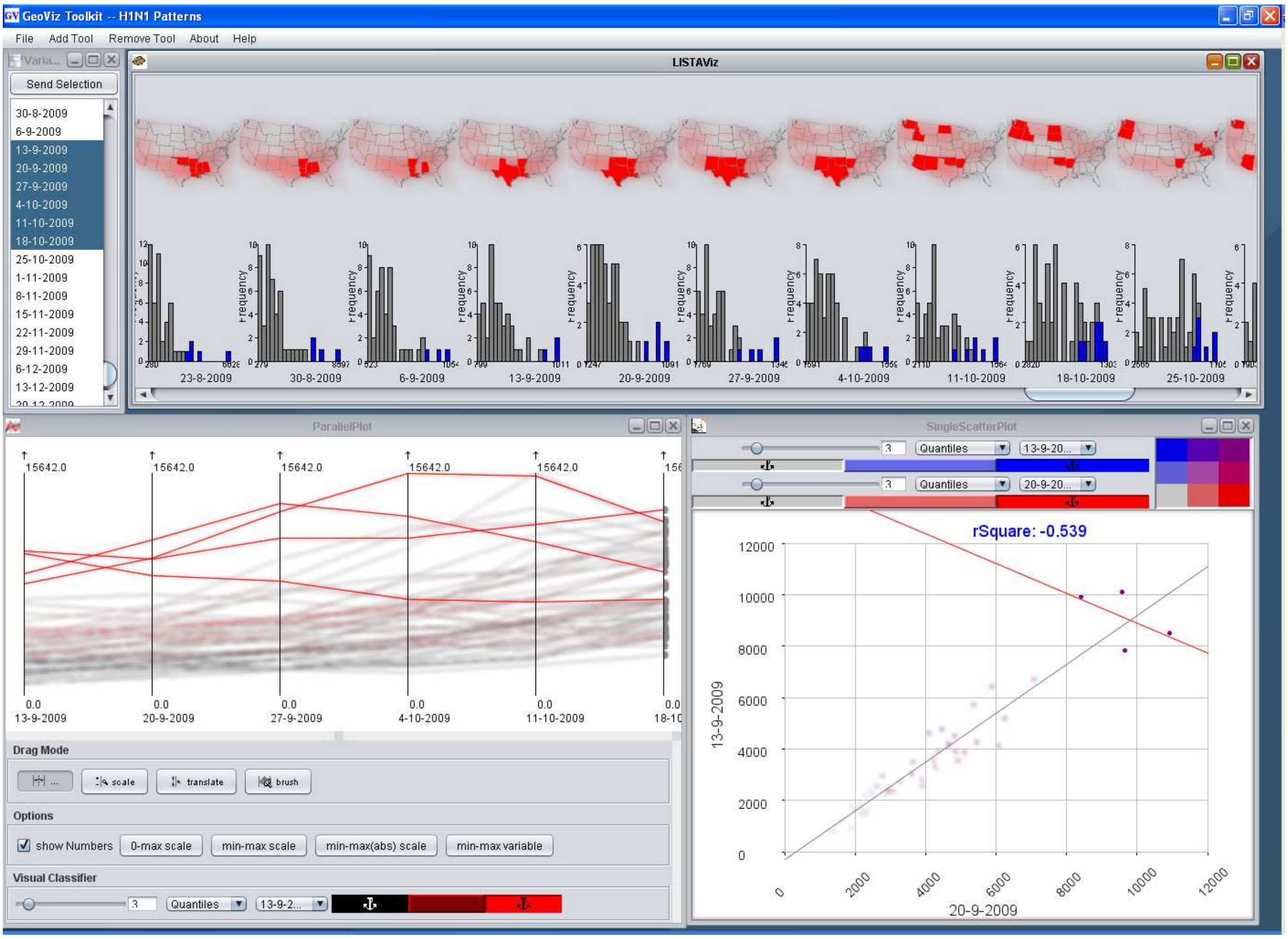

**Figure 4: LISTA-Viz in GeoViz showing spatio-temporal inflection point in the H1N1 pandemic of 2009. Three visual analysis components are shown: LISTA-Viz, a parallel coordinate plot, and a scatterplot. The LISTA-Viz component shows the space-time clusters for each time period, including the inflection point where the space-time pattern showed a dramatic change. The parallel coordinate plot shows the trajectory of ILI rates in the key cluster identified by LISTA-Viz, centred around Texas. The scatterplot shows the general upward trend for all states between the two key weeks in the inflection point, while it also highlights that the states in the key cluster have a decreasing trend.**

## Conclusions and Future Work

Considering the capabilities of this method and its associated tool, some logical extensions are evident. These include using the spatial memoization strategy with other spatial analysis methods, evaluating these methods for use in disease surveillance, trying alternative weight structures, and other topologies of coordination.

### *Using Spatial Memoization in Other Spatial Analysis Methods*

We have shown here that for the case of extending the local Moran's I statistic from spatial to spatio-temporal usage, the spatial memoization technique can provide useful computational savings. We also suggested that other related techniques should see similar benefits if this technique is applied. An interesting avenue for future research would be to apply these methods to other spatial analysis contexts, especially if an outside threshold for spatial dependence (a range) can be established. For example, we could apply this kind of method to geostatistics. The spatial memoization approach should yield computational gains, because what is needed for these methods is similar: estimated probability distributions. Geographically Weighted Regression (GWR) should show similar gains under similar conditions. Spatial statistics on point patterns would not be suitable unless the data is downgraded to gridded locations, and if the is a high enough density of events in the study area.

### *Evaluate Disease Surveillance Capability*

The goal of many public health agencies is to be able to monitor infectious disease outbreaks (Burkom 2003). Bringing interactive spatio-temporal exploration capabilities, including significance levels, has potential to aid the efforts of public health officials. Most large state health organizations have started developing infectious disease programs, however, most currently popular methods suffer from deficiencies(Shmueli *et al.*). Two deficiencies in particular can be addressed by visual analytics: current methods too run slowly, and they identify too many false positives. The techniques used in LISTA-Viz could help redress these problems, addressing the first by creating computational savings, and the second by allowing health analysts to quickly analyze clustering results.

### *Other Weight Structures*

In the currently reported research, all weights are of the binary, on or off variety. This has been done for reasons of computational efficiency. However, using our methods, it would be possible to use more elaborate weight structures, such as kernel approaches, where nearby neighbours have more influence than more distant observations, in both space and time. Doing so should not prove much more computationally costly as long as the method of enumerating the topological alternatives and estimating probability density functions is used. However, depending on the weight structures, as the number of unique spatio-temporal neighbourhoods approaches the full set of places and neighbours, the advantages of the topological approach will be lost.

### *Topologies of Coordination*

Finally, we turn to another kind of topology: the connections between the software components of a visual analytics system. The internal coordination model of LISTA-Viz differs from that of the GeoViz Toolkit. Further work on exploring how different topologies of component connection can aid visual analytic tasks may be fruitful.

## Acknowledgments

The support of the VACCINE Center (Visual Analytics for Command, Control, and Interoperability Environments, Grant 2009-ST-061-CI0001), the Contextual influences on the category construction of geographic-scale movement patterns (ConCat) grant from the National Science Foundation (Grant 0924534), the Vaccine Modeling Initiative grant from the Bill & Melinda Gates Foundation (Grant 49279), and a Wilson Travel Grant from the college of Earth and Mineral Sciences at the Pennsylvania State University, is gratefully acknowledged, however no endorsement is implied.

## References

Andrienko, G., Andrienko, N. and Bartling, U., 2008a, Visual analytics approach to user-controlled evacuation scheduling. *Information Visualization*, 7, pp. 89-103.

Andrienko, G., Andrienko, N., Dykes, J., Fabrikant, S.I. and Wachowicz, M., 2008b, Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research *Information Visualization*, 7, pp. 173-180.

Andrienko, G., Andrienko, N. and Gatalsky, P., 2004, Visual mining of spatial time series data. *Knowledge Discovery in Databases: Pkdd 2004, Proceedings*, 3202, pp. 524-527.

Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M.J., Maceachren, A. and Wrobel, S., 2007, Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science*, 21, pp. 839-857.

Andrienko, G., Andrienko, N., Jankowski, P. and Kraak, M.J., 2009, Special issue: GeoSpatial Visual Analytics. *Cartography and Geographic Information Science*, 36, pp. 223-224.
Andrienko, N., Andrienko, G., Savinov, A., Voss, H. and Wettschereck, D., 2001, Exploratory analysis of spatial data using interactive maps and data mining. *Cartography and Geographic Information Science*, 28, pp. 151-165.
Anselin, L., 1995, Local Indicators of Spatial Association - LISA. *Geographical Analysis*, 27, pp. 93-115.
Anselin, L., Syabri, I. and Kho, Y., 2006, GeoDa: An introduction to spatial data analysis. *Geographical Analysis*, 38, pp. 5-22.
Bivand, R., Muller, W.G. and Reder, M., 2009, Power calculations for global and local Moran's l. *Computational Statistics & Data Analysis*, 53, pp. 2859-2872.
Brooks, S.P., 1998, Markov chain Monte Carlo method and its application. *Journal of the Royal Statistical Society Series D-the Statistician*, 47, pp. 69-100.
Brunsdon, C., Fotheringham, S. and Charlton, M., 1998, Geographically weighted regression - modelling spatial non- stationarity. *Journal of the Royal Statistical Society Series D-the Statistician*, 47, pp. 431-443.
Burkom, H.S., 2003, Development, adaptation, and assessment of alerting algorithms for biosurveillance. *Johns Hopkins Apl Technical Digest*, 24, pp. 335-342.
Campbell, C. and Egbert, S., 1990, Animated cartography: Thirty years of scratching the surface. *Cartographica*, 27, pp. 24-46.
Cliff, A.D. and Ord, J.K., 1973, *Spatial Autocorrelation* (London: Pion).
Cliff, A.D. and Ord, J.K., 2009, What were we thinking? *Geographical Analysis*, 41, pp. 351-363.
Demsar, U., Fotheringham, A.S. and Charlton, M., 2008, Exploring the spatio-temporal dynamics of geographical processes with geographically weighted regression and geovisual analytics. *Information Visualization*, 7, pp. 181-197.
Getis, A., 2008, A history of the concept of spatial autocorrelation: A geographer's perspective. *Geographical Analysis*, 40, pp. 297-309.
Guo, D., Chen, J., MacEachren, A.M. and Liao, K., 2006, A Visualization System for Space-Time and Multivariate Patterns (VIS-STAMP). *Ieee Transactions on Visualization and Computer Graphics*, 12, pp. 1461-1474.
Hardisty, F., 2005, The GeoViz Toolkit: User-centered run-time coordination of geographic visualization components. In *Auto-Carto* (Las Vegas, NV: AAG).
Hardisty, F. and Conley, J., 2008, Interactive Detection of Spatial Clusters. *Advances in Disease Surveillance*, 5, p. 37.
Hardisty, F. and Robinson, A., Accepted, The GeoViz Toolkit: Using component-oriented coordination methods to aid geovisualization application construction. *International Journal of Geographic Information Science*.
Huang, S.S., Yokoe, D.S., Stelling, J., Placzek, H., Kulldorff, M., Kleinman, K., O'Brien, T.F., Calderwood, M.S., Vostok, J., Dunn, J. and Platt, R., 2010, Automated Detection of Infectious Disease Outbreaks in Hospitals: A Retrospective Cohort Study. *Plos Medicine*, 7.
Huisman, O., Santiago, I.F., Kraak, M.J. and Retsios, B., 2009, Developing a Geovisual Analytics Environment for Investigating Archaeological Events: Extending the Space-Time Cube. *Cartography and Geographic Information Science*, 36, pp. 225-236.
Koussoulakou, A. and Kraak, M.J., 1992, Spatio-temporal maps and cartographic communication. *Cartographic Journal*, 29, pp. 101-108.
Kraak, M.J., 2008, From geovisualisation toward geovisual analytics. *Cartographic Journal*, 45, pp. 163-164.
Kulldorff, M. and Nagarwalla, N., 1995, Spatial disease clusters - detection and inference. *Statistics in Medicine*, 14, pp. 799-810.

Lawson, A.B., 2006, Disease cluster detection: A critique and a Bayesian proposal. *Statistics in Medicine*, 25, pp. 897-916.
Leung, Y., Mei, C.L. and Zhang, W.X., 2003, Statistical test for local patterns of spatial association. *Environment and Planning A*, 35, pp. 725-744.
Li, H.F., Calder, C.A. and Cressie, N., 2007, Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*, 39, pp. 357-375.
MacEachren, A.M. and Dibiase, D., 1991, Animated maps of aggregate data - conceptual and practical problems. *Cartography and Geographic Information Systems*, 18, pp. 221-229.
Moran, P.A.P., 1950, A test for the serial independence of residuals. *Biometrika*, 37, pp. 178-181.
Shmueli, G. and Burkom, H., Statistical Challenges Facing Early Outbreak Detection in Biosurveillance. *Technometrics*, 52, pp. 39-51.
Slingsby, A., Dykes, J. and Wood, J., 2008, Using treemaps for variable selection in spatio-temporal visualisation. *Information Visualization*, 7, pp. 210-224.
Slingsby, A., Dykes, J. and Wood, J., 2009, Configuring hierarchical layouts to address research questions. *Ieee Transactions on Visualization and Computer Graphics*, 15, pp. 977-984.
Thrower, N., 1961, Animated cartography in the United States. *International Yearbook of Cartography*, 1, pp. 20-30.
Tiefelsdorf, M., 2002, The saddlepoint approximation of Moran's I's and local Moran's I-i's reference distributions and their numerical evaluation. *Geographical Analysis*, 34, pp. 187-206.
Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solorzano, A., Swayne, D.E., Cox, N.J., Katz, J.M., Taubenberger, J.K., Palese, P. and Garcia-Sastre, A., 2005, Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science*, 310, pp. 77-80.
Zhang, T.L. and Lin, G., 2007, A decomposition of Moran's I for clustering detection. *Computational Statistics & Data Analysis*, 51, pp. 6123-6137.