Proactive Spatiotemporal Resource Allocation and Predictive Visual Analytics for Community Policing and Law Enforcement

Abish Malik, Ross Maciejewski, *Member, IEEE*, Sherry Towers, Sean McCullough, and David S. Ebert, *Fellow, IEEE*

Abstract— In this paper, we present a visual analytics approach that provides decision makers with a proactive and predictive environment in order to assist them in making effective resource allocation and deployment decisions. The challenges involved with such predictive analytics processes include end-users' understanding, and the application of the underlying statistical algorithms at the right spatiotemporal granularity levels so that good prediction estimates can be established. In our approach, we provide analysts with a suite of natural scale templates and methods that enable them to focus and drill down to appropriate geospatial and temporal resolution levels. Our forecasting technique is based on the Seasonal Trend decomposition based on Loess (STL) method, which we apply in a spatiotemporal visual analytics context to provide analysts with predicted levels of future activity. We also present a novel kernel density estimation technique we have developed, in which the prediction process is influenced by the spatial correlation of recent incidents at nearby locations. We demonstrate our techniques by applying our methodology to Criminal, Traffic and Civil (CTC) incident datasets.

Index Terms—Visual Analytics, Natural Scales, Seasonal Trend decomposition based on Loess (STL), Law Enforcement

1 Introduction

The increasing availability of digital data provides both opportunities and challenges. The potential of utilizing these data for increasing effectiveness and efficiency of operations and decision making is vast. Harnessing this data with effective tools can transform decision making from reactive to proactive and predictive. However, the volume, variety, and velocity of these data can actually decrease the effectiveness of analysts and decision makers by creating cognitive overload and paralysis by analysis, especially in fast-paced decision making environments.

Many researchers in data visualization and visual analytics [37] have proposed interactive visual analytical techniques to aid analysts in these tasks. Unfortunately, most work in this area has required these casual experts (experts in domains, but not necessarily statistics experts) to carefully choose appropriate parameters from a vast parameter space, select the proper resolution over which to perform their analysis, apply appropriate statistical or machine learning analysis techniques, and/or understand advanced statistical significance testing, while accounting for the different uncertainties in the data and processes.

Moreover, the casual experts are required to adapt their decision making process to the statistical analysis space where they need to choose the appropriate time and space scales that give them meaningful analytical and predictive results. They need to understand the role that data sparsity, different distribution characteristics, data variable co-dependencies, and data variance play in the accuracy and reliability of the analytical and prediction results. In moving to this proactive and predictive environment, scale issues become even more important. Not only does the choice of appropriate scales help guide the users' perception and interpretation of the data attributes, it also facilitates gaining new insight into the dynamics of the analytical tasks [42] and the validity of the analytical product: a spatial resolution level that is too fine may lead to zero data input values with no predictive statistical value; whereas, a scale that is too coarse can overgeneralize the data and introduce variation and noise, reducing the value and specificity of the results. Therefore, it becomes critical for forecasting and analysis to choose statistically meaningful resolution and aggregation scales. Utilizing basic principles from scaling theory [42], and Norman's naturalness and appropriateness principles [26], we can both balance and harness these cognitively meaningful natural human-centered domain scales with meaningful statistical scales.

Therefore, in this paper, we present a visual analytics approach that provides casual experts with a proactive and predictive environment that enables them to utilize their domain expertise while exploring their problem and making decisions and predictions at natural problem scales to increase their effectiveness and efficiency in planning, resource allocation, and deployment. Our visual analytics framework [21, 22] provides interactive exploration of multisource, multivariate spatiotemporal datasets using linked views. The system enables the exploration of historic datasets and examination of trends, behaviors and interactions between the different spatiotemporal data elements. The focus of this paper, however, is to provide a proactive decision making environment where historic datasets are utilized at natural geospatial and temporal scales in order to guide future decisions and resource allocation strategies.

In our predictive visual analytics process, we allow users to interactively select and refine the data categories over which to perform their analyses, explore and apply meaningful geospatial (Sections 4.1-4.3) and temporal (Section 4.4) scales and aggregations, apply the forecasting process over geospace (Section 5), and visualize the forecasting results over their chosen geospatial domain. We utilize a Seasonal Trend decomposition based on Loess (STL) [9] approach (Section 3) that utilizes patterns of historical data and apply it in the geospatial domain to predict future geospatial incidence levels. Moreover, this approach provides domain-driven refinement of analysis and exploration to areas and time of significance (e.g., high crime areas or times).

The contributions of our work include these novel natural spatial and temporal analytical techniques, as well as a novel Dynamic Covariance Kernel Density Estimation method (DCKDE) (Section 4.2.2). These contributions can be applied to a variety of spatiotemporal datasets including distribution and logistics, public safety, public health, and law enforcement. We will utilize data from Criminal, Traffic, and Civil (CTC) incident law enforcement datasets in the examples throughout this paper. However, it should be noted that our technique is versatile and can be adapted for other relevant spatiotemporal datasets that exhibit seasonality.

- *Abish Malik, Sean McCullough and David S. Ebert are with Purdue University. E-mail: amalik|mccullo0|ebertd@purdue.edu.*
- *Ross Maciejewski and Sherry Towers are with the Arizona State University. E-mail: rmacieje|smtowers@asu.edu.*

Manuscript received 31 Mar. 2014; accepted 1 Aug. 2014. Date of publication 11 Aug. 2014; date of current version 9 Nov. 2014.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.
Digital Object Identifier 10.1109/TVCG.2014.2346926

2 Related Work

In recent years, there has been much work done in utilizing historic datasets for informing future actions and decisions of decision makers. Below, we discuss previous work in the field of visual analytics, and, since our chosen example domain and implementation is focused on crime data, we also explore previous work in criminology to provide a breadth of the related research areas.

2.1 Predictive Visual Analytics

There have been several visual analytics systems developed in recent years that support data analysis and exploration processes, and provide extensive data modeling and hypothesis generation tools (e.g., [28, 35]). More recently however, researchers have also started progressing toward creating visual analytics systems that incorporate predictive analytics in them. For example, Wong et al. [43] provide a visual interface and an environment that brings together research from several different domains to predict and assess the impact of climate change on U.S. power-grids. Muhlbacher and Piringer [25] provide a visual analytics framework for building regression models. Monroe et al. [23] utilize user-driven data visualizations that enable researchers to gain insights into large healthcare datasets.

Yue et al. [45] created an artificial intelligence based tool that leverages interactive visualization techniques to leverage data in a predictive analytics processes. Their time series modeling technique includes the use of the Box-Jenkins procedure [27]. Other time series modeling techniques extensively used include the ARMA (Auto Regressive Moving Average) [1] and ARIMA (Auto Regressive Integrated Moving Average) models. A summary of some other methods that involve geospatial modeling can be found in [11, 12]. Maciejewski et al. [20] utilize the seasonal trend decomposition by loess smoothing for generating temporal predictions for modeling spatiotemporal healthcare events. They also use the kernel density estimation technique for creating probability distributions of patient locations for use in healthcare data. Our work builds on these ideas where we utilize historic datasets to provide spatiotemporal forecasts into the future. The focus of our work is to explore the issues of geospatial and temporal scales so that casual experts can adapt their decision making process to the statistical analysis space. As such, we apply a user assisted data analysis approach to drive future decisions that helps prevent decision makers from getting over-burdened, while, at the same time, maximizes the utilization of their domain knowledge and perceptual capabilities.

2.2 Crime Hotspot Policing and Intervention

In recent years, there has been much research done that suggest the benefits of hot spot policing in preventing crime and disorder at these crime hotspots (e.g., [2, 3, 4]). Weisburd et al. [41] examine the effect and impact of crime hot spots policing and their findings suggest little negative effects and backlash among the residents of targeted areas of such policing efforts. Sherman [30] also explores the effects of police crackdowns (sudden increase in police presence in specific regions) among several case studies. He notes that while most of the crackdowns appeared to demonstrate initial deterrent effects, the effects decayed after short periods of time. Our work also enables law enforcement decision makers to identify and target crime hotspots by forecasting high probability crime regions based on historic spatiotemporal trends. Our work also factors in the temporal variations within the signals and, as such, provides dynamic hotspot locations for each predicted day.

Goldkamp and V$\hat{i}$lcic$\tilde{a}$ [15] provide insights into unanticipated negative effects of place-oriented enforcement intervention schemes on other societal aspects. They explored an intensive targeted enforcement strategy that was focused on drug crime and its related community effects and examined the overall side effects on the society. Sherman et al. [31] examine and provide an overview of the different aspects of predatory criminal activity at different spatial granularities and how these factors correlate with different aspects of the society. Bruin et al. [7] provide a toolkit that extracts the different factors from police datasets and creates digital profiles for all offenders. The tool then clusters the individuals against the created profiles by using a distance matrix that is built around different attributes (e.g., crime frequency, criminal history of the offenders).

2.3 Predictive Policing

There has been much work done in criminology to study criminal behaviors in order to develop models that predict various offense incidence levels at different spatial aggregation levels. Brown and Oxford [6] study methods that pertain to predicting the number of breaking and enterings in sub-cities and correlate breaking and enterings with different factors including unemployment rates, alcohol sales and previous incidents of crime. Yu et al. [44] also develop a crime forecasting model by employing different data mining classification techniques. They employ several classification techniques including Nearest Neighbor, Decision Tree and Support Vector Machines. Their experiments are run on two different data grid sizes, the 24-by-20 (approx. one-half mile square) and the 41-by-40 square grid cells (approx. one-quarter mile square). They note that the 24-by-20 grids consistently gave them better results than the 41-by-40 grids, which they attribute to the lack of sufficient information at the coarser resolution. Our technique also allows analysts to conduct their predictive forecasting at different spatial resolutions (e.g., over uniform spatial grids and natural underlying spatial boundaries) and temporal granularity levels (e.g., by day, week, month). Furthermore, our system also allows users to create spatial and temporal templates for use in the prediction process.

Monthly and seasonal cycles and periodic properties of crime are well known among criminologists [17]. Felson and Poulson [14] factor in the time of the day variation in the analysis of crime and provide summary indicators that summarize the hour-of-day variations. They provide guidelines for breaking the day into quartiles based on the median hour of crime. We use their guidelines in our work and provide default data driven time-of-day templates over which to forecast crime. We also utilize these techniques and incorporate the seasonality and periodicity properties of crime in order to provide spatiotemporal forecasts of future crime incidence levels.

3 Time Series Prediction using Seasonal-Trend Decomposition based on Loess (STL)

In order to model time series data, we employ the seasonal-trend decomposition technique based on a locally weighted regression (loess) methodology (STL), where a time series signal is considered to consist of the sum of multiple components of variation. To accomplish this, we first utilize the STL method [9, 16] to desynthesize the time series signal into its various components. An analysis of the underlying time series signal Y for CTC data reveals that a square root power transform stabilizes the variability and yields a more Normal distribution of time series residuals, which is a requirement to appropriately model the time series using STL. We consider the time series signal $\sqrt{Y}$ to consist of the sum of its individual components given by $\sqrt{Y_v} = T_v + S_v + D_v + R_v$, where, for the v-th time step, T_v is the inter-annual component, S_v is the yearly-seasonal component, D_v is the day-of-the-week effect, and R_v is the remainder variation component.

To predict using the STL method, we apply the methodology described in [20], where the fitted values $\hat{Y} = (\hat{y}_1, ..., \hat{y}_n)$ generated using the loess operator in the STL decomposition step are considered to be a linear transformation of the input time series $Y = (y_1, ..., y_n)$. This is given by $\hat{y}_i = \sum_{i=1}^{n} h_{ij} y_j \Rightarrow \hat{Y} = HY$, where H is the operator matrix whose (i, j)-th diagonal elements are given by $h_{i,j}$. In order to predict ahead by n days, we append the operator matrix H obtained from predicting ahead within each linear filter in the STL process with n new rows, and use this to obtain the predicted value. The predicted value for day $n+1$ is thereby given by $\hat{y}_{n+1} = \sum_{i=1}^{n} H_{n+1,i}\, Y_i$.

We use this concept of time series modeling and prediction and extend it into the spatiotemporal domain (see Section 5 for details). We further factor in for the sparcity of data in certain geographical regions, and devise strategies to alleviate problems resulting in prediction in these sparse regions (Section 4).

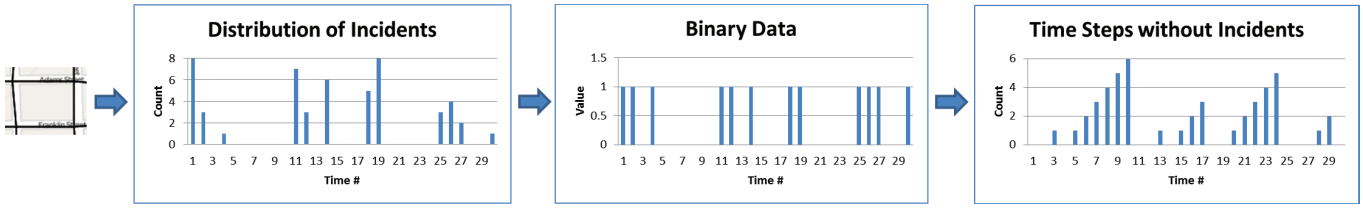

Fig. 1. Our geospatial natural scale template signal generation process. For each geospatial sub-division, the system generates a time series of the number of incidents, converts it into a binary signal, and processes the binary signal to generate the signal used to form the geospatial template.

4 Natural Scale Templates

In order to assist with the analysis process, we provide decision makers with natural scale templates that enable them to focus on appropriate geospatial and temporal resolution levels. These templates enable users to analyze their data at appropriate spatiotemporal granularity levels that help align the scale and frame of reference of the data analysis process with that of the decision making process. These templates also assist users in alleviating the impedance mismatch between data size/complexity and the decision makers' ability to understand and interact with data [29]. We support the creation of both geospatial and temporal templates in our system that facilitate the decision making process. A combination of the generated geospatial and temporal templates provide analysts with an appropriate starting point in the analysis process; thereby, eliminating the need to examine and analyze the entire spatiotemporal parameter space and reducing it to more manageable, appropriate scale levels. To be effective, the design of these scale templates must follow the appropriateness, naturalness, and matching cognitive principles [26]. As Wilkinson and Stevenson both point out [36, 40, 42], simple scaling theory techniques are not sufficient (e.g., axometric scaling theory), but provide useful guidance to primitive scales of reference. The combinations of these design principles and the guidance from these statistical scale papers, provide the motivation and basis for our natural scale templates described below.

4.1 Geospatial Templates

An underlying assumption with using STL to decompose time series is that the data are Normally distributed. The model predictions can get severely biased if this assumption is violated or if data are sparse. To remedy this, we provide methods that help guide users in creating geospatial scales that allow them to drill down to higher incidence regions that may provide better prediction estimates.

4.1.1 Geospatial Natural Scale Templates based on Spatiotemporal Incident Distribution

Our system allows users to narrow down the geographic space for the scope of analysis to regions with higher incidence counts and higher statistical significance for user-selected incident types. Our geospatial natural scale template methodology is shown in Figure 1. In order to generate geospatial templates, the system first fragments the geographic space into either uniform rectangular grids [6] or man-made spatial demarcations (e.g., census blocks). Then, for each subregion, the system generates a time series of the number of incidents that occurred within the subregion over time (e.g., by day, week, month). This signal is further cached for use later in the forecasting process. Next, we convert this time series signal into a binary signal across time, where a 1 represents that an incident occurred on a particular day and a 0 that no incident occurred. We then count the number of 0's between the 1's and progressively sum the number of 0's, outputting the result as another time series signal. As such, this signal is a representation of the number of time steps over which no incidents occurred for the given subregion.

This new time series signal is now utilized in the STL forecasting method (Section 3) and a predicted value is computed for the next day. It should be noted that the resulting time series for regions of lower incidence counts will not be sparse, and consequently, will generate higher predicted values. This process is repeated for all geospatial subregions and a unified picture is obtained for the next day. Finally, we filter out the regions with higher predicted values (low activity) by thresholding for the maximum value. The resulting filtered region forms the initial geospatial template. An example of a created geospatial template using this technique is shown in Figure 4 (Left).

4.1.2 User Refinement of Geospatial Template using Domain Knowledge

The geospatial template provides regions with relatively higher incident rates. The system further allows users to use their domain knowledge and interactively refine these template regions into sub-divisions. For example, users may choose to sub-divide the formed template regions by natural or man-made boundaries (e.g., state roads, rivers, police beats), or by underlying features (e.g., known drug hotspots). The system also allows users to explore the predicted future counts of the created sub-regions by generating an incidence count vs. time signal for each disjoint region and applying our forecasting methodology (Section 3) to find a predicted value for the next day. The results are then shown as a choropleth map to users (e.g., Figure 4 (Right)). These macro-level prediction estimates further assist decision makers in formulating high-level resource allocation strategies.

4.2 Kernel Density Estimation

One of the challenges with using the spatial distribution of incidents in a geospatial predictive analytics process is that it can exacerbate the problem of generating signals with low or no data values. To further refine our prediction model in geospace, we utilize a Kernel Density Estimation (KDE) technique to spread the probability of the occurrence of incidents to its neighboring regions. The rationale behind this is that criminology research has shown evidence that occurrence of certain types of crimes (e.g., residential burglary) at a particular region puts neighboring regions at an elevated risk [13, 18, 32].

Furthermore, crime also tends to be clustered in certain neighborhoods, and the probability of a crime occurring at a particular location can be highly correlated with the number of recent crimes at nearby locations. We incorporate this concept in a novel kernel density estimation method described in Section 4.2.2, where the kernel value at a given location depends on the locations of its k-nearest incidents. In addition, kernel density estimation methods take into account that crimes in low-crime or sparsely populated areas have low incidence, but non-zero probability. We utilize two interchangeable density estimation techniques in our implementation.

4.2.1 Kernel Scale based on Distance to the k-th Nearest Neighbor

To account for regions with variable data counts, we utilize a kernel density estimation technique and use a dynamic kernel bandwidth [33]. We scale the parameter of estimation by the distance from the point x to its kth nearest neighbor X_i. This is shown in Equation 1.

$$\hat{f}(x) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{max(h, d_{i,k})} K\left(\frac{x - X_i}{max(h, d_{i,k})}\right) \quad (1)$$

Here, N is the total number of samples, $d_{i,k}$ the distance from the i-th sample to the k-th nearest neighbor and h is the minimum allowed kernel width. We use the Epanechnikov kernel [33] to reduce calculation time, which is given by $K(\mathbf{u}) = \frac{3}{4}(1-\mathbf{u}^2)1_{(||u||\leq 1)}$. Here, the

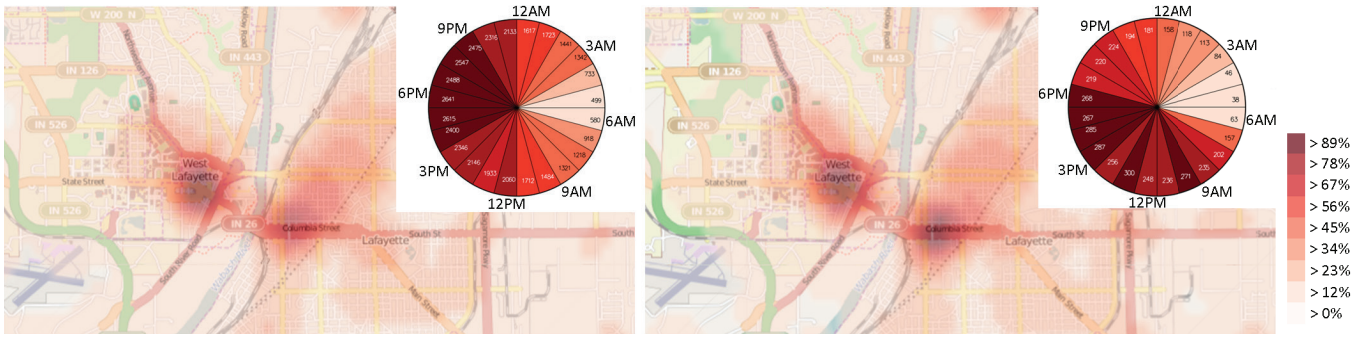

Fig. 2. Spatiotemporal distribution of historical CTC incidents for Tippecanoe County for (Left) 3/11/2012 through 3/10/2014, and (Right) for all Tuesdays in March in the past 10 years.

function $1_{(||u||\leq 1)}$ evaluates to 1 if the inequality is true and to 0 otherwise. In cases where the distance from the i-th sample to the k-th nearest neighbor is 0 (e.g., multiple calls from the same address), we force the variable kernel estimation to a minimum fixed bandwidth h. Making the kernel width placed at the point X_i proportional to $d_{i,k}$ gives regions with sparse data a flatter kernel, and vice-versa.

4.2.2 Dynamic Covariance Kernel Density Estimation Technique (DCKDE)

The kernel in the previous method is based on the distance from an incident location to its k-th nearest neighbor, which provides a flatter kernel for sparse regions. In a new kernel method, we use the information from all k-nearest neighbors to calculate the width of the kernel (rather than the most distant neighbor), thus reducing stochastic variation on the width of the kernel. As such, we fragment the geospatial region into rectangular grids and then utilize a Gaussian kernel at every grid node that is based on the covariance matrix of the location of the center of each node $\mathbf{X} = \{x, y\}$ and its k-nearest neighbor incidents [39]. Therefore, the kernel value is influenced by the k-nearest neighbors and provides a wider kernel in sparsely populated regions that enables the model prediction to be small but non-zero and also takes into account correlations between latitude and longitude; thus, improving the accuracy of the estimates. The value stored at each node location is given by $\delta(\mathbf{X}) = \frac{1}{2\pi|V|} e^{-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T V^{-1}(\mathbf{X}-\boldsymbol{\mu})}$, where $\boldsymbol{\mu} = \{\mu_x, \mu_y\}$ is the mean along the x and y directions of the k nearest neighbors and their covariance matrix V is defined as:

$$V = \begin{bmatrix} \sigma_x^2 & cov_{x,y} \\ cov_{x,y} & \sigma_y^2 \end{bmatrix} \tag{2}$$

Here, σ_x^2 and σ_y^2 is the variance along the x and y dimension respectively, and $cov_{x,y} = \sum_{i=1}^{k} \frac{(x_i-\mu_x)(y_i-\mu_y)}{k-1}$ is the sample covariance between x and y.

4.3 Neighbors with Similar Spatio-Demographics

For regions that generate a signal of lower statistical significance for the user selected categories, we provide the option to explore data in similar neighborhoods. For each census block, we utilize spatio-demographic census data to find those census blocks that exhibit similar spatial demographics. The rationale behind finding similar neighborhoods lies in the fact that regions with similar demographics tend to exhibit similar trends for certain types of crime [24, 34].

The process of finding similar census blocks for a given census block X includes computing the similarity distance from X to all neighboring census blocks that lie within a d mile radius from the centroid of X. The d mile radius constraint is imposed to factor in for Tobler's first law of geography [38] that suggests that near regions are more related to one another than distant regions. We use $d = 3.0$ miles in our implementation [8]. As such, the similarity distance between two census blocks A and B given k census data variables is given by $S_{A,B} = \sqrt{\sum_{i=1}^{k}(A(V_i) - B(V_i))^2}$, where $A(V_i)$ and $B(V_i)$ are the corresponding census data variable values (e.g., race, income, and age demographic data) for census blocks A and B respectively. Finally, the top N census blocks with the smallest similarity distance values are chosen as the similar census blocks for the given census block X. We use $N = 5$ as a default value in our implementation, but provide users with options to change this value on demand. We note that our future work includes extending this concept of finding similar neighborhoods to determining similar data categories for predictive purposes.

The system now provides users with the ability to generate *similar neighborhood* prediction maps where the prediction for a given census block X depends on the historic time series data of its N similar census blocks in addition to the past data of the census block X itself. Here, the input time series for the census block X used in the prediction algorithm is the per time step average of the N similar census block signals combined with the original signal from census block X. The resulting prediction maps incorporates the influence of incidence rates in neighborhoods that share similar spatio-demographic data.

4.4 Temporal Natural Scale Templates

As noted previously in Section 2.3, crime trends exhibit not only monthly and seasonal trends, but also shows day-of-the-week and hour-of-day variations. The prediction maps produced by the methods described so far provide prediction estimates over 24-hour periods. This information, albeit valuable to the law enforcement community in developing resource allocation strategies for their precincts, provides little detail of the 24-hour distribution of crime. In this section, we describe our approach to assist users in creating temporal scales.

4.4.1 Interactive Clock Display

Figure 2 (Top-Right) shows our interactive clock view that enables a radial display of temporal hourly data. The clock view provides a way for users to filter the data by the hour by interactively clicking on the desired hours, thereby filtering down the data for use in the prediction process. Users may use the clock view display to obtain a visual summary of the hourly distribution of the incidents and consequently make informed decisions on creating temporal templates over which good prediction estimates may be established.

4.4.2 Factoring in for Monthly and Day-of-the-Week Variations

In addition to utilizing the seasonal trend decomposition technique described in Section 3 to decompose the time series signals into its various components, we also utilize a direct approach where we allow users to create their own custom monthly and/or daily templates. Certain crimes tend to peak on certain days of the week (e.g., alcohol related violations tend to be higher over the weekend), whereas other crimes tend to be lower on other days (e.g., reported burglaries drop over the weekend). As such, we factor for these effects directly in the system and allow users to filter data specifically by month and/or by day-of-the-week. This further assists decision makers in developing and refining their resource allocation strategies.

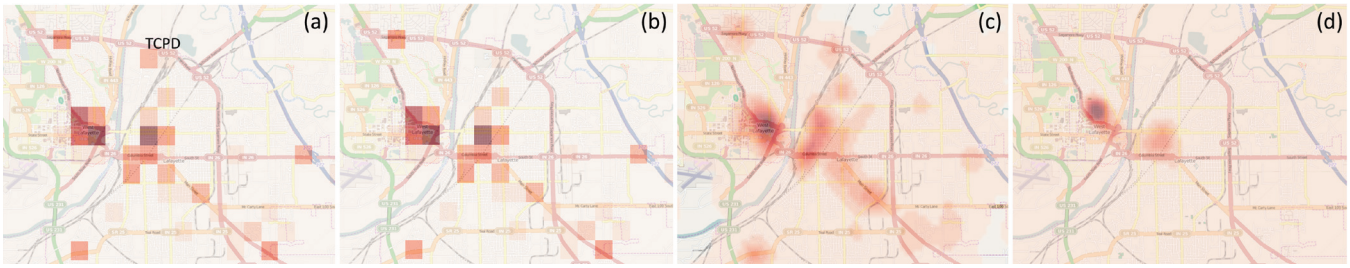

Fig. 3. Geospatial prediction results for 3/11/2014 for Tippecanoe County obtained using our STL forecasting methodology. (a) Predicted choropleth map for rectangular grids of dimension 64×64 using incidence count time series by day. (b) Refined predicted map after removing TCPD location from (a). (c) Predicted map using KDE based on the distance to the k-th nearest neighbor approach (Section 4.2.1). (d) Forecast map using DCKDE method (Section 4.2.2).

4.4.3 Refinement using Summary Indicators

We extend the method described in [14] to further assist users with refining and choosing appropriate hourly templates in the prediction process. In this method, the system computes the *median minute* of CTC incident for the selected 24-hour binning period that provides information about when exactly half of the incidents for the selected date range and offense types have occurred. Next, to get an indication of the dispersion of crime within the 24-hour period, the system computes the *first quartile minute* and *third quartile minute* for the selected data, which are the median times of the first and second halves of the 24-hour period from the median minute respectively. Finally, as temporal data can be inaccurate with many incidents that have missing time stamps, we provide users with an accuracy indicator to show the percentage of cases with valid time stamps. These summary indicators, along with the temporal templates described above, enable users to further refine their selected temporal templates for use in the prediction process. Example scenarios where these indicators are used are provided in Section 6.

5 Geospatial Prediction

The described visual analytics process involves a domain expert selecting appropriate data parameters, applying desired data filters and generating spatial and temporal natural scale templates using the methods described in Section 4. Next, the system incorporates the STL forecasting method (Section 3) and extends it to the geospatial domain to provide prediction estimates for the next N time steps (e.g., days, weeks, months). We now list the steps involved in our geospatial prediction methodology:

1. **Dividing geospace into sub-regions**: The first step in our methodology, just like in Section 4.1.1, involves subdividing geospace into either uniform rectangular grids of user specified resolutions or man-made geospatial boundaries.
2. **Generating the time series signal**: The system then extracts a time series signal for each sub-division. We allow two types of signals to be extracted for each sub-division: (a) incidence count vs. time step, and (b) kernel value vs. time step. Note that the signal generated in (a) is the same as that produced in Section 4.1.1 (Figure 1 (Distribution of Incidents)). The kernel values used in (b) are generated using any one of the methods described in Section 4.2.
3. **Forecasting**: The time series signal generated for each spatial unit is then fed through the STL process described in Section 3 where a forecast is generated for the next N time steps (e.g., days, weeks). This process is repeated for all region sub-divisions and prediction maps are finally obtained for the next N time steps.
4. **Visualizing results**: Finally, the results of our forecasting method are provided to the user either in the form of a choropleth map or a heatmap.

When users choose to fragment the geospace into uniform rectangular grids, we provide them with the ability to select the resolution level, or, in other words, the grid size of each grid. An incidence count vs. time step signal is then generated for each sub-region. It is important to note here that a grid resolution that is too fine may result in a zero count vs. time step signal that has no predictive statistical value. On the other hand, a grid resolution that is too coarse may introduce variance and noise in the input signal, thereby over-generalizing the data. An evaluation of our forecasting approach (Section 7) indicates that an average input size of 10 samples per time step provide enough samples for which our method behaves within the constraints and assumptions of our STL forecasting approach. We utilize this metric in our system in order to determine the applicability of our forecasting method for a particular sub-region.

Figure 3 shows a series of examples that demonstrate our geospatial prediction results using the methods described in this section. Here, the user has selected all CTC incidents for Tippecanoe County, IN, and is using 10 years' worth of historical data (3/11/2004 through 3/10/2014) to generate forecast maps for the next day (i.e., for 3/11/2014). Figure 3 (a) shows the prediction results when Tippecanoe County, IN is fragmented into rectangular grids of dimension 64×64. The input data for each sub-region consists of daily incidence count data over the last 10 years. This method, unlike the KDE methods, does not spread the probability to surrounding neighborhood regions when an incident occurs at a particular place. As a result, this method treats each region independently, and can be used when there are no correlations between geospatial regions (e.g., commercial vs. residential neighborhoods). This method can also be useful in detecting anomalous regions and regions of high predicted levels of activity. For example, the user notices something peculiar from the results in Figure 3 (a): a predicted hotspot occurs prominently over the Sheriff's office and county jail location (labeled as TCPD in Figure 3 (a)). This occurs because the default geospatial location of many incidents are logged in as the county jail, especially when arrests are associated with cases. To remedy for this, the user can refine the underlying geospatial template (Section 4.1.2) and dynamically remove this location from the geospatial template. The refined prediction map generated is shown in Figure 3 (b).

Figures 3 (c and d) show the predicted results of using the kernel density estimation based on the distance to the k-nearest neighbor approach (Section 4.2.1) and the DCKDE technique (Section 4.2.2), respectively. The KDE method applied to generate the prediction map in Figure 3 (c) provides a flatter kernel for relatively low-crime regions. As a result, the prediction map provides lower, but non-zero, predictions for these regions. The kernel width computed using this method is based on the distance from a point x to its kth nearest neighbor only. The DCKDE method, on the other hand, assumes that the probability of the occurrence of an incident at a particular location is correlated with the number of recent incidents at nearby locations. Accordingly, this method utilizes information from *all* k-nearest neighbors in calculating the kernel value. Thus, the regions with persistently higher incident concentrations generate focused hotspots when forecasting is performed using the DCKDE method. Finally, it should be noted that each method provides users with different insights into the dynamics of the underlying processes, and users can use their domain knowledge to further refine the results to make informed decisions.

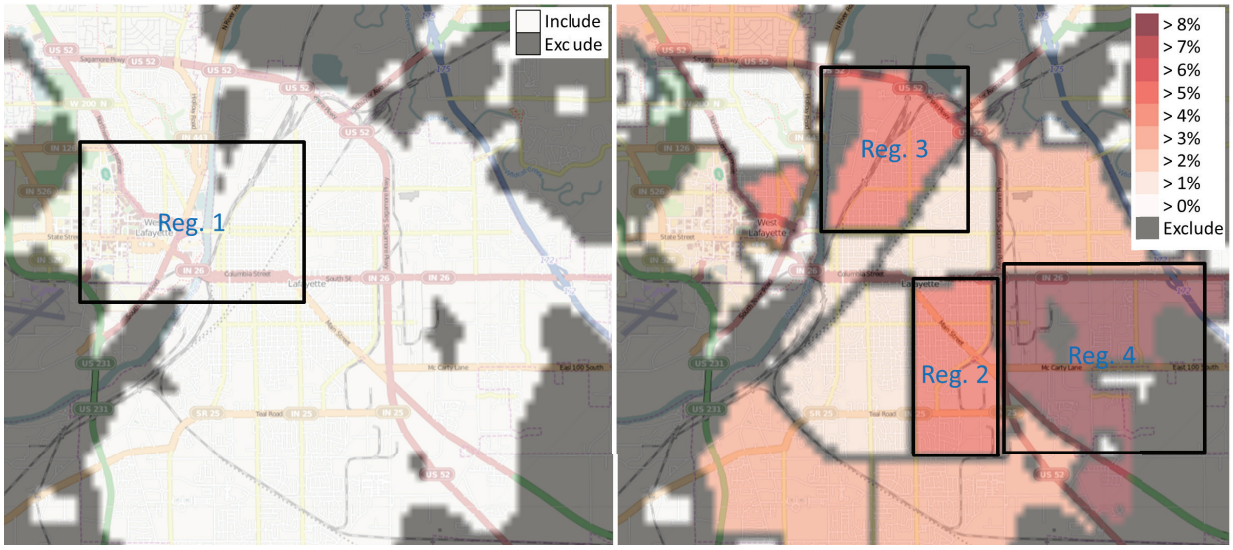

Fig. 4. (Left) Geospatial template generated for Tippecanoe County using 10 years' worth of historical data. (Right) Choropleth map showing the distribution of predicted incidents for 3/11/2014 by police beats for Tippecanoe County. Users may further select regions on the map (e.g., Reg. 1-4) to generate detailed predictions for the selected regions (Figure 5).

6 Case Study: Forecasting Future Criminal, Traffic and Civil (CTC) Incidence Levels

In this section, we demonstrate our work by applying our spatiotemporal natural scale template methodology to forecast for CTC incidence levels in Tippecanoe County, IN, U.S.A. This dataset consists of historical reports and provides several different attributes, including the geographic location, offense type, agency, date, and time of the incident. This dataset contains an average of 31,000 incidents per year for Tippecanoe County, and includes incidence reports for different categories of CTC incidents (e.g., crimes against person, crimes against property, traffic accidents). We use 10 years worth of historical data for this analysis. We provide a workflow when using our system in the analysis process.

Forecasting for all geospatial CTC incidents

Here, we describe a hypothetical scenario in which a law enforcement shift supervisor is using our system to develop resource allocation strategies for Tippecanoe County over the next 24 hour period for Tuesday, March 11, 2014. The supervisor is interested in developing a high-level resource allocation strategy, in particular, by police beats for the next 24 hour period. Law enforcement officers are generally assigned to a particular law beat and patrol their beat during their shift hours when not responding to a call for service. The supervisor is also interested in determining which hotspot locations to focus on for larger police beats. Finally, he also wants to refine the developed resource allocation strategy to factor in for the hourly variation of crime. To develop an appropriate resource allocation strategy, the shift supervisor performs several different analyses that are described in the following subsections. Although our example uses data for all CTC categories as inputs, users may filter their data using any combinations of CTC categories (e.g., crimes against property, person) to further refine their resource allocation strategy.

Overall daily resource allocation

The shift supervisor begins his process by visually exploring the spatiotemporal distribution of historical incidents using our system. When working through the system, the supervisor then visualizes the geospatial and hourly distribution of the incidents that occurred over the past 2 years, as shown in Figure 2 (Left). The supervisor notes several hotspots emerge for the selected period. The locations of these hotspots match with his domain knowledge of the area (e.g., city downtown regions, shopping center locations across town). The static image of the aggregate data, however, does not factor in the inherent spatiotemporal data variations, and basing a resource allocation decision on this image alone would be insufficient. The supervisor is also aware of the fact that police presence can act as a deterrent for certain types of crimes, and, therefore, wants to diversify and maximize police presence in these hotspot areas.

Next, the supervisor wants to factor for monthly and day-of-the-week patterns in his analysis. As such, he visualizes the geospatial and hourly distribution of all CTC incidents that occurred on any Tuesday in the month of March over the past 10 years (Section 4.4.2). The result is shown in Figure 2 (Right). The supervisor notes a slightly different geospatial distribution emerges as a result, with the intensity of hotspots shifting towards the east downtown Lafayette region. In this case, it also becomes apparent that for the 24-hour distribution, 10 AM, 1 PM and 3 PM-6 PM emerge as high activity hours.

Allocating resources by police beats

In order to narrow down the geospace and focus on relevant geographic locations, the supervisor decides to apply our geospatial template generation technique (Section 4.1) with all CTC incidents selected using 10 years' worth of historical data (i.e., from 3/11/2004 through 3/10/2014). The resulting geospace generated is shown in white in Figure 4 (Left). The supervisor notes that the resulting regions correspond to highly populated areas, and exclude areas of infrequent occurrences. Next, the system provides a total predicted number of incidents, N, for March 11, 2014 for the filtered geospatial region. This is done by generating a total incidence count vs. day time series signal using the past 10 years' worth of data and applying the STL forecasting method described in Section 3. Here, N is 59 incidents.

Next, the supervisor is interested in obtaining a high level overview of the distribution of the predicted incidents over geospace, and, in particular, by police patrol routes. As such, the supervisor uses our system and fragments the generated geospatial template using the city law beats shapefile. The resulting geospace is shown in Figure 4 (Right). In order to distribute the total predicted 59 incidents across police beats, the system computes an incidence count vs. day time series signal for each disjoint geospatial region and computes the predicted number of incidents n_i for each region (Section 3). Next, the probability of an incident within each disjoint region is calculated using the formula $p_i = n_i/N * 100$. The results of this operation are then shown to the user as a choropleth map, where each disjoint region is colored according to its value on a sequential color scale [5] (Figure 4 (Right)).

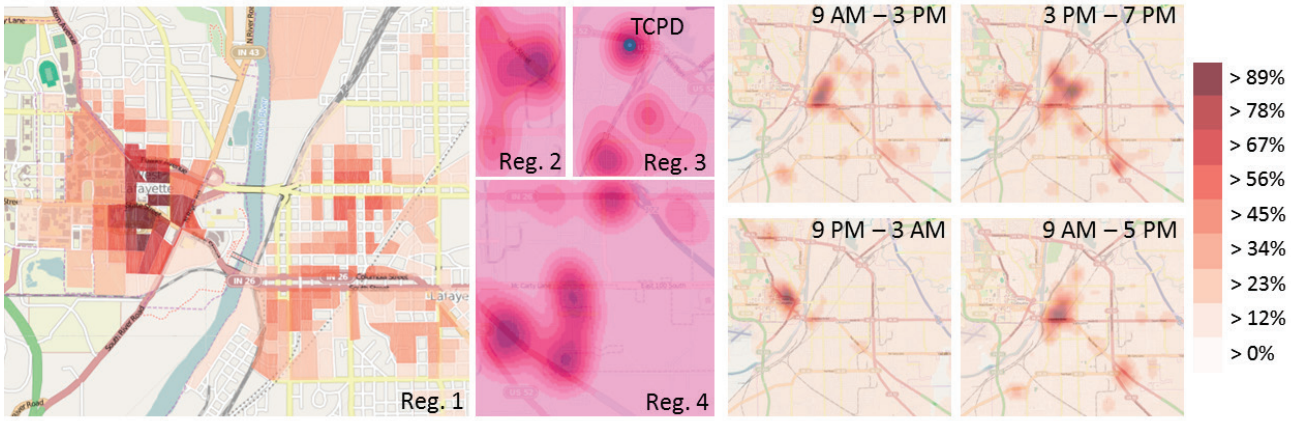

Fig. 5. User refinement of geospatial resource allocation strategy. The user has chosen to visualize predicted hotspots for regions labeled in Figure 4 (Regions 1 through 4), and for Tippecanoe County over hourly temporal templates.

Geospatial resource allocation strategy refinement using domain knowledge

While the high level police beat prediction map (Figure 4 (Right)) suggests putting a heavier emphasis on the eastern police beats of the city, the prediction results in Figure 3 indicate a more localized concentration of incidents at the city downtown locations. The shift supervisor may use these results and allocate higher resources to the eastern police beat of the city (Reg. 4 in Figure 4), and allocate a smaller number of resources, but at more concentrated locations in the downtown (Reg. 1 in Figure 4).

Now, the supervisor is interested in further refining her geospatial resource allocation strategy. First, she turns to the predicted hotspot regions in the city downtown regions (Reg. 1 in Figure 4). She decides to utilize the census blocks spatial boundary information and divides the geospace into census blocks. Next, she uses the method described in Section 5 to create a predicted choropleth map based on census blocks for the region. The result of this operation is shown in Figure 5 (Reg. 1). Here, the supervisor has chosen to use the kernel values obtained from the method described in Section 4.2.1 and spread them across the underlying census blocks for generating these results.

To obtain detailed predictions for the eastern city police beat region (Reg. 4 in Figure 4), the shift supervisor uses a different approach where she draws a region around the selected beat using the mouse and restricts the forecast to the selected region. The result of this operation is shown in Figure 5 (Reg. 4). From domain knowledge, she knows that this area has a high concentration of shopping centers. The hotspots obtained in Figure 5 (Reg. 4) align with these locations. Finally, the supervisor generates similar heatmaps for regions labeled as Reg. 2 and 3 in Figure 4, the results of which are shown in Figures 5 (Reg. 2 and 3), respectively. Note that the county jail location is once again a hotspot in Figure 5 (Reg. 3). With these detailed results in hand, the shift supervisor is able to devise an optimal resource allocation strategy for the next 24 hour period in Tippecanoe County.

Applying temporal templates

Finally, in order to refine her resource allocation strategy to different portions of the day, the shift supervisor chooses to apply the summary indicators method (Section 4.4.3). She finds that the first, median, and third quartile minutes for CTC incidents that occurred in the past 10 years were 9:25 AM, 3:11 PM and 7:28 PM respectively. She also notes that these indicators correspond with the hourly distribution of incidents using the clock view display in Figure 2. Therefore, the supervisor chooses two hourly templates using these summary indicators: (a) 9 AM through 3 PM, and (b) 3 PM through 7 PM. The supervisor also creates two other hourly templates: 9 PM through 3 AM to capture night time activity, and 9 AM through 5 PM to capture working hours of the day. She then uses the kernel density estimation method (Section 4.2.1) and re-generates prediction maps for March 11, 2014. These results are shown in Figure 5. As expected, the supervisor notes the shift in hotspot locations through the 24 hour period, which further enables the refinement of the resource allocation strategy for the different portions of the 24 hour period.

7 Model Evaluation and Validation

In order to evaluate our methodology, we conducted a series of statistical tests to understand the behavior and applicability of our approach in the spatiotemporal domain. Our validation strategy involved testing for the empirical rule of statistics, which describes a characteristic property of a Normal distribution: 95% of the data points are within the range $\pm 1.96\ \sigma$ of μ, where μ and σ are the mean and standard deviation of the distribution, respectively [10]. In order to help alleviate the challenges resulting due to the sparseness of the underlying data, we performed our analyses over a weekly data aggregation level. Our approach involved testing whether the 95% prediction confidence interval bound acquired for the geospatial predictions using our forecasting approach holds when compared against observed data [19]. This confidence bound would be violated if the variance of the observed data is higher (i.e., overdispersed data) or lower (i.e., underdispersed data) than that dictated by the prediction confidence bound. When the 95% prediction bounds are met as expected, and the data conforms to the Normal regime, the applicability of our spatiotemporal STL forecasting method is established.

Building on our STL based time series prediction discussion from Section 3, the variance of the fitted values $\hat{Y} = (\hat{y}_1, ..., \hat{y}_n)$ using the loess operator in the STL decomposition step is given by $Var(\hat{Y}_i) = \hat{\sigma}^2 \sum_{j=1}^{n} H_{ij}^2$ [20]. Here, $\hat{\sigma}^2$ is the variance of the input time series signal Y, and is estimated from the remainder term R_v. Subsequently, the variance for the predicted value $\hat{Y}_{n+1}$ for time step $n+1$ is given by $Var(\hat{Y}_{n+1}) = \hat{\sigma}^2(1 + \sum_{j=1}^{n} H_{n+1,j}^2)$. This provides the 95% prediction interval as $CI_{n+1} = \hat{Y}_{n+1} \pm 1.96\sqrt{Var(\hat{Y}_{n+1})}$.

Next, we performed a series of analyses at varied geospatial and temporal scales, and for different data categories. The geospace was first fragmented into sub-regions (either rectangular grids or using man-made boundaries), and time series signals were generated for each geospatial sub-region. In our analyses, we utilized a sliding time window of size 3 years (i.e., 3×52 weeks) that provided enough samples above the Nyquist frequency for the STL forecasting technique. Forecasting was performed using the methods described in Sections 5 and 7.1. We provide our evaluation methodology and results in the subsequent sub-sections.

7.1 Modified STL forecasting method to factor in for weekly data aggregation

As described earlier in Section 3, a time series signal $\sqrt{Y}$ can be considered to consist of the sum of its inter-annual (T_v), yearly-

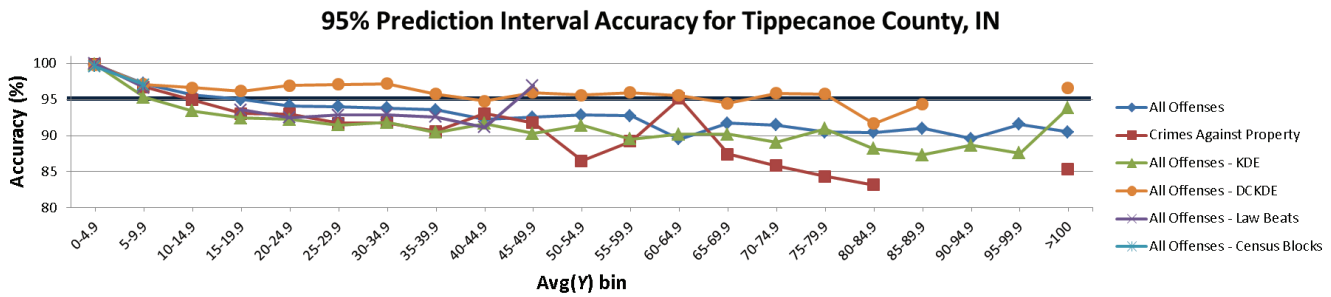

Fig. 6. 95% prediction interval accuracy vs. Avg(Y) for different CTC offenses for Tippecanoe County, IN. Here, geospace has been fragmented into rectangular grids of dimension $k \times k$ ($\forall\, k \in [1,128]$), and by law beats and census blocks.

seasonal (S_v), day-of-the-week (D_v), and remainder variation (R_v) components. However, since we used a weekly aggregation of data, the day-of-the-week component (D_v) must be excluded. Therefore, the time series signal gets modified to $\sqrt{Y_v} = T_v + S_v + R_v$. The prediction step, which involves predicting the value for week $n+1$, remains the same as given in Section 3.

7.2 95% prediction interval accuracy vs. input data average (Avg(Y))

In this method, the geospace was first fragmented into either: (a) rectangular grid regions of dimension $k \times k$ ($\forall\, k \in [1,128]$, with 128 chosen as upper threshold to provide a fine enough geospatial resolution), or (b) man-made geospatial regions (e.g., census blocks, census tracts, law beats, user-specified regions). For each geospatial region, we first generated the incidence count vs. week signal (denote this signal as Y) for a time window of n weeks beginning from the week of, e.g., 1/1/2009. We then used the modified STL forecasting method (Section 7.1) to calculate the 95% prediction interval CI for the predicted week $n+1$, and tested whether the observed data for week $n+1$ fell within the calculated 95% prediction interval for that geospatial region. The average of the input signal Y, $Avg(Y)$, was also calculated.

Next, the input time window was shifted by one week to generate the corresponding incidence count vs. week signal (so, this signal would begin from the week of 1/7/2009). We again computed $Avg(Y)$, and CI for the predicted week $n+1$. As before, we tested whether the observed data for the predicted week $n+1$ fell within the calculated 95% prediction interval. We repeated the process by sliding the time window till it reached the end of available data. For each $Avg(Y)$ value, we maintained two counters that kept track of the number of instances the observed data was within the 95% prediction interval ($C_{Correct}$), and the total instances encountered thus far (C_{Total}). Finally, $Avg(Y)$ values were binned, and $C_{Correct}$ and C_{Total} were summed for each bin. The 95% prediction interval accuracy for each $Avg(Y)$ bin is then given as $\frac{\sum_{bin} C_{Correct}}{\sum_{bin} C_{Total}} \times 100\%$.

7.3 Results and discussion

Figure 6 shows the 95% prediction interval accuracy results for different CTC offenses for Tippecanoe County, IN using the method described in Section 7.2. As can be observed from these results, when the average bin values are low (e.g., less than 10 input samples), the accuracy levels are higher than the expected 95% confidence bound. This indicates that the data are underdispersed for lower input values. In other words, the variance of the observed data is lower than that of the 95% prediction bound when the underlying data are sparse. This conforms to the expected behavior for predicting using our STL forecasting technique: the model predictions get biased if the underlying data are too sparse.

As the input signal average ($Avg(Y)$) values get larger (i.e., more than 10 samples per time step), the prediction accuracy starts to converge at around the expected 95% accuracy level. For example, the prediction interval accuracy for all offenses converges at around 93%. Also, note that the prediction accuracy using the DCKDE method (Section 4.2.2) converges close to the 95% accuracy level; thereby, indicating the efficacy of the technique. It should be noted that since the underlying processes being modeled here (e.g., CTC incidents) are inherently stochastic in nature, perfect 95% confidence bounds will not be achieved (as can be seen from the results in Figure 6). Furthermore, with an uncertain probability distribution of the underlying data, our application of the square root power transform may not guarantee homoscedasticity (i.e., stabilization of variability). This also contributes to our system not achieving perfect 95% confidence bounds. However, even though perfect confidence bounds are not achieved (as can be observed from Figure 6), the accuracy converges close to the 95% bounds. These results show that the underlying data are Normally distributed for higher values of $Avg(Y)$; thereby, satisfying the underlying assumptions of our method used to estimate the 95% confidence interval. This establishes the validity of the claims of our STL prediction methodology in the geospatial domain that the prediction modeling method works as expected as long as the underlying assumptions of the method are satisfied by the data.

Figure 6 shows the 95% prediction interval accuracy vs. input data average results (Section 7.2) for man-made geospatial regions (census blocks and law beats). These results show that the confidence bounds using census blocks are invariably higher than the expected 95% bound, which indicates that the underlying data are underdispersed. Census blocks are small geospatial units, typically bounded by streets or roads (e.g., city block in a city). The smaller $Avg(Y)$ values for census blocks in Tippecanoe County in Figure 6 (less than 10 input samples) further highlight the sparsity of input data. The combination of higher prediction interval accuracy levels and lower $Avg(Y)$ values are telltale for the data sparseness issues we have described, and suggest that the signals generated using census blocks have low predictive statistical power. This further underlines the need to intelligently combine geospatial regions of lower statistical values to obtain a signal of higher predictive power (e.g., as was done in Section 4.3). The 95% prediction interval accuracy results obtained using law beats in Figure 6, on the other hand, shows the accuracy converging at around the expected 95% confidence interval for higher $Avg(Y)$ values (more than 10 input samples). These results provide further evidence that as the underlying data values become larger and begin to conform to the Normal regime, our geospatial prediction methodology provides prediction estimates that are within the expected 95% prediction confidence interval. This further bolsters the applicability and validity of our STL prediction methodology in the geospatial domain.

We also applied the method described in Section 7.2 to all CTC incident category data and generated 95% prediction interval accuracy vs. the input signal average value ($Avg(Y)$) plots for different grid resolutions k. These results are shown in Figure 7. The results indicate that 95% prediction interval accuracy converges at or around the 95% confidence level for large enough $Avg(Y)$ values (i.e., for $Avg(Y)$ bigger than 10). The results indicate that our methodology behaves within the constraints of the Normal regime at higher $Avg(Y)$ values for the different grid dimensions. Also, note that smaller grid dimensions (k) correspond to larger geospatial sub-divisions; and accordingly, smaller k values generate signals of larger counts per bin (i.e., larger $Avg(Y)$

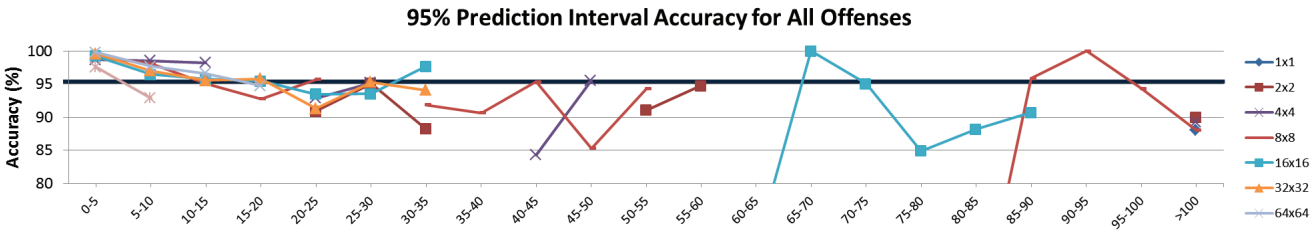

Fig. 7. 95% prediction interval accuracy vs. Avg(Y) for all CTC offenses for Tippecanoe County, IN. Here, geospace has been fragmented into rectangular grids of dimension $k \times k$ for various k values.

values), especially for regions with higher incidence rates. As can be seen from the results in Figures 6 and 7, the accuracy for higher $Avg(Y)$ values tend to be lower than the 95% prediction accuracy; thereby, indicating that the underlying data are slightly overdispersed. These results indicate that coarse scales can generate signals with too much variance, or combinations of multiple signals that overgeneralize the data. Furthermore, the signals generated at coarse scales can be affected by anomalies in underlying data (e.g., crime spikes during unusually high weathers, holidays). These can contribute to the non-Normality of the residuals, and produce an overdispersion of underlying data as compared to the assumptions of our model. It should be noted that although a slight data overdispersion is noticeable at coarse scales, they are deemed to be small enough to currently not warrant any correction. Finally, we note that further research is needed in order to determine the effects of these data overgeneralization issues at coarse scales and to devise strategies to mitigate for their effects.

7.4 Summary

Our model evaluation and validation strategy involved testing for the empirical rule of a Normal distribution where we tested whether the observed data conformed with the 95% prediction interval from our STL forecasting method at various geospatial scales. In order to cope with data sparseness issues, we performed our analysis at a weekly aggregation of data. Our results demonstrate the validity of our approach as long as the underlying assumptions of the underlying models are satisfied by the data. The results obtained using our DCKDE method are also promising. Our results also highlight the importance of performing analysis at appropriate scales, and demonstrate that the model predictions get severely biased when the underlying assumptions are violated by the data. We also explored the effects of data sparseness issues on our model predictions at fine geospatial scales. Our evaluation results show that the model predictions generated using input signals of 10 or more counts per time step on average tend to conform with the 95% prediction confidence intervals. We also highlight the effects of analysis performed at coarse scales, and show the data overgeneralization issues that occur at such scales. Although the results indicate a slight data overdispersion at coarse scales, the results show that the prediction accuracies from the model estimates still tend to converge at around the 95% confidence bounds. This further shows the effectiveness of our forecasting methodology in the geospatial domain. We also note that although our work enables hot spot policing and resource allocation strategy development, further evaluation is required to ascertain the efficacy of our predictive analytics framework when deployed in field. We leave this as future work.

8 Domain Expert Feedback

Our system was assessed by a police captain who oversees the operations and resource allocation of several precincts in a mid-sized police agency (of about 130 sworn officers) in the United States. In this section, we summarize the initial feedback received after conducting several informal interviews with him. The captain emphasized the need for a system that applies a data-driven approach to assist law enforcement decision makers in developing resource allocation strategies. He was impressed by the ability of the system to interactively generate various geospatial and temporal visualizations of historical datasets and forecast maps in real-time. Additionally, he also appreciated having the ability to dynamically apply any desired geospatial, temporal, and/or categorical filters on the data.

The captain stressed the need to carefully combine and aggregate different data categories for which reliable forecast maps could be generated. For example, he noted that a signal generated by combining two crime categories with different attributes (e.g., crimes against property and person) might introduce variability in the forecasting process and produce unreliable results. He further suggested that crimes of opportunity must be filtered out as these exhibit no discernable patterns. He asserted that different regions within the same city can exhibit different crime patterns due to the different underlying region dynamics. He expressed the importance for domain experts to create data category and spatiotemporal templates so viable prediction estimates can be computed using our methodology. Finally, the captain remarked that the predicted hotspot locations using aggregated CTC data occur at the known problem areas in the city.

9 Conclusions and Future Work

In this work, we have presented our visual analytics framework that provides a proactive decision making environment to decision makers and assists them in making informed future decisions using historical datasets. Our approach provides users with a suite of natural scale templates that support analysis at multiple spatiotemporal granularity levels. Our methods are built at the confluence of automated algorithms and interactive visual design spaces that support user guided analytical processes. We enable users to conduct their analyses over appropriate spatiotemporal granularity levels where the scale and frame of reference of the data analysis process and forecasting matches with that of the user's decision making frame of reference. It should be noted that while adjusting for the size of the geospatial and temporal scales is necessary, it is also important to adjust for the scale of the size of the dataset. A forecasting or analysis method that works well for one region with certain demographics and population densities may not have the same efficacy when applied to a different region. As such, our work explores the potential of visual analytics in providing a bridge so that different statistical and machine learning processes occur on the same scale and frame of reference as that of the decision making process.

Our future work includes developing new kernel density estimation techniques designed specifically for improving prediction forecasts. We further plan on improving our designed dynamic covariance kernel density estimation technique (DCKDE) to factor in for temporal distances to further enhance our STL based prediction algorithm. We also plan to incorporate data-driven methods that guide users in selecting between different choices provided by the system based on the underlying features of the data. We also plan on factoring in the influences and correlations among different variables to further refine our natural scale template generation methodology. Finally, we plan on conducting a formal user evaluation in order to understand the efficacy of our system in aiding domain experts to understand the properties of underlying data and their effects on the workings of the different underlying statistical processes.

Acknowledgments

This work was funded by the U.S. Department of Homeland Security VACCINE Center's under Award Number 2009-ST-061-CI0003.

REFERENCES

[1] G. Box and G. Jenkins. *Time series analysis: Forecasting and control*. Holden-Day, San Francisco, 1970.

[2] A. A. Braga. The effects of hot spots policing on crime. *Annals of the American Academy of Political and Social Science*, 578:pp. 104–125, 2001.

[3] A. A. Braga and B. J. Bond. Policing crime and disorder hot spots: A randomized controlled trial*. *Criminology*, 46(3):577–607, 2008.

[4] A. A. Braga, D. M. Hureau, and A. V. Papachristos. An ex post facto evaluation framework for place-based police interventions. *Police Quarterly*, 2012.

[5] C. A. Brewer. *Designing Better Maps: A Guide for GIS users*. ESRI Press, 2005.

[6] D. Brown and R. Oxford. Data mining time series with applications to crime analysis. In *IEEE International Conference on Systems, Man, and Cybernetics*, volume 3, pages 1453–1458 vol.3, 2001.

[7] J. S. d. Bruin, T. K. Cocx, W. A. Kosters, J. F. J. Laros, and J. N. Kok. Data mining approaches to criminal career analysis. In *Proceedings of the Sixth International Conference on Data Mining*, ICDM '06, pages 171–177, Washington, DC, USA, 2006. IEEE Computer Society.

[8] D. V. Canter. The environmental range of serial rapists. In D. V. Canter, editor, *Psychology in Action*, Dartmouth Benchmark Series, pages 217–230. Dartmouth Publishing Company, Hantshire, UK, January 1996.

[9] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning. Stl: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6:3–73, 1990.

[10] G. Cowan. *Statistical data analysis*. Oxford University Press, 1998.

[11] P. J. Diggle and P. J. Diggle. Statistical analysis of spatial point patterns. *London: Edward Arnold*, 1983.

[12] P. J. Diggle, J. Tawn, and R. Moyeed. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3):299–350, 1998.

[13] G. Farrell and K. Pease. *Repeat victimization*, volume 12. Criminal Justice Press, 2001.

[14] M. Felson and E. Poulsen. Simple indicators of crime by time of day. *International Journal of Forecasting*, 19(4):595–601, 00 2003.

[15] J. S. Goldkamp and E. R. Vilcica. Targeted enforcement and adverse system side effects: The generation of fugitives in philadelphia. *Criminology*, 46(2):371–409, 2008.

[16] R. Hafen, D. Anderson, W. Cleveland, R. Maciejewski, D. Ebert, A. Abusalah, M. Yakout, M. Ouzzani, and S. Grannis. Syndromic surveillance: Stl for modeling, visualizing, and monitoring disease counts. *BMC Medical Informatics and Decision Making*, 9(1):21, 2009.

[17] K. Harries. *Crime and the environment*. American Lecture Series; No. 1033. Charles C. Thomas Publisher, Limited, 1980.

[18] S. D. Johnson, W. Bernasco, K. J. Bowers, H. Elffers, J. Ratcliffe, G. Rengert, and M. Townsley. Space–time patterns of risk: a cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology*, 23(3):201–219, 2007.

[19] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li. *Applied linear statistical models*, volume 5. McGraw-Hill Irwin Chicago, 2004.

[20] R. Maciejewski, R. Hafen, S. Rudolph, S. Larew, M. Mitchell, W. Cleveland, and D. Ebert. Forecasting hotspots: A predictive analytics approach. *IEEE Transactions on Visualization and Computer Graphics*, 17(4):440–453, April 2011.

[21] A. Malik, R. Maciejewski, T. F. Collins, and D. S. Ebert. Visual analytics law enforcement toolkit. In *IEEE International Conference on Technologies for Homeland Security*, pages 222–228, 2010.

[22] A. Malik, R. Maciejewski, N. Elmqvist, Y. Jang, D. Ebert, and W. Huang. A correlative analysis process in a visual analytics environment. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 33–42, 2012.

[23] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman. Temporal event sequence simplification. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2227–2236, 2013.

[24] J. D. Morenoff, R. J. Sampson, and S. W. Raudenbush. Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence*. *Criminology*, 39(3):517–558, 2001.

[25] T. Muhlbacher and H. Piringer. A partition-based framework for building and validating regression models. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):1962–1971, Dec 2013.

[26] D. A. Norman. *Things that make us smart: Defending human attributes in the age of the machine*. Basic Books, 1993.

[27] R. Oppenheim. Forecasting via the box-jenkins method. *Journal of the Academy of Marketing Science*, 6(3):206–221, 1978.

[28] A. Rind, T. Lammarsch, W. Aigner, B. Alsallakh, and S. Miksch. Timebench: A data model and software library for visual analytics of time-oriented data. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2247–2256, 2013.

[29] G. Robertson, D. Ebert, S. Eick, D. Keim, and K. Joy. Scale and complexity in visual analytics. *Information Visualization*, 8(4):247–253, 2009.

[30] L. W. Sherman. Police crackdowns: Initial and residual deterrence. *Crime and Justice*, 12:pp. 1–48, 1990.

[31] L. W. Sherman, P. R. Gartin, and M. E. Buerger. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27–56, 1989.

[32] M. Short, M. Dorsogna, P. Brantingham, and G. Tita. Measuring and modeling repeat and near-repeat burglary effects. *Journal of Quantitative Criminology*, 25(3):325–339, 2009.

[33] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1986.

[34] S. J. South and S. F. Messner. Crime and demography: Multiple linkages, reciprocal relations. *Annual Review of Sociology*, 26(1):83–106, 2000.

[35] J. Stasko, C. Görg, and Z. Liu. Jigsaw: supporting investigative analysis through interactive visualization. *Information Visualization*, 7(2):118–132, 4 2008.

[36] S. S. Stevens. On the theory of scales of measurement, 1946.

[37] J. J. Thomas and K. A. Cook, editors. *Illuminating the Path: The R&D Agenda for Visual Analytics*. IEEE Press, 2005.

[38] W. R. Tobler. A computer movie simulating urban growth in the detroit region. *Economic geography*, pages 234–240, 1970.

[39] S. Towers. Kernel probability density estimation methods. *Proceedings of the Advanced Statistical Techniques in Particle Physics*, pages 107–111, 2002.

[40] P. F. Velleman and L. Wilkinson. Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, 47(1):65–72, 1993.

[41] D. Weisburd, J. Hinkle, C. Famega, and J. Ready. The possible backfire effects of hot spots policing: an experimental assessment of impacts on legitimacy, fear and collective efficacy. *Journal of Experimental Criminology*, 7(4):297–320, 2011.

[42] L. Wilkinson and G. Wills. *The Grammar of Graphics*. Statistics and Computing. Springer, 2005.

[43] P. C. Wong, L. R. Leung, N. Lu, M. Paget, J. C. Jr., W. Jiang, P. Mackey, Z. T. Taylor, Y. Xie, J. Xu, S. Unwin, and A. Sanfilippo. Predicting the impact of climate change on u.s. power grids and its wider implications on national security. In *AAAI Spring Symposium: Technosocial Predictive Analytics*, pages 148–153. AAAI, 2009.

[44] C.-H. Yu, M. W. Ward, M. Morabito, and W. Ding. Crime forecasting using data mining techniques. In *Proceedings of the IEEE 11th International Conference on Data Mining Workshops*, ICDMW '11, pages 779–786, Washington, DC, USA, 2011. IEEE Computer Society.

[45] J. Yue, A. Raja, D. Liu, X. Wang, and W. Ribarsky. A blackboard-based approach towards predictive analytics. In *AAAI Spring Symposium: Technosocial Predictive Analytics*, page 154. AAAI, 2009.