Space, Time, and Hurricanes: Investigating the Spatiotemporal Relationship among Social Media Use, Donations, and Disasters

J. Sylvester, J. Healey, C. Wang, W. Rand
Center for Complexity in Business
Robert H. Smith School of Business
University of Maryland
{jsylvest, jhealey, cwang120, wrand}@umd.edu

Abstract

Social media is generated both in space and time, and in the context of disaster situations it can provide minute-by-minute and location-by-location awareness of the event and users' concerns and feelings about crises. In this paper, we investigate the relationship between social media data, web donations, and traditional media coverage. We start by examining the temporal relationships, then move on to spatial relationships and patterning. Finally, we explore spatiotemporal patterns, examining the confluence of these patterns. Our results indicate that social media can be used as an indicator of both spatial and temporal patterns of real world events such as donations to a large not-for-profit. Our models can be used to make recommendations to managers about how to target marketing messaging appropriately given the current state of social media.

1 Introduction and Background

As the use of social media has grown in recent years there has been an increasing effort turned toward studying the link between online and offline behavior. Online social networks affect both how people donate to charity campaigns [1] and how they respond to emergency events [2, 3]. In this paper we use data surrounding a natural disaster, Hurricane Irene, to study the link between social media activity and donations to disaster relief through two channels: SMS messaging and the web. We investigate the patterns of behavior of these activities in both time and space, as well as time and space together. Though both the temporal and spatial patterns of social media have been investigated in isolation, relatively little attention has been paid to how patterns emerge in both dimensions together.

For instance, Herrmann et al. [4] explore the temporal aspects of information spread online in relation to outside events. Yang and Leskovec [5] investigate the different forms that temporal patterns take in Twitter and blogging platforms. Grinberg et al. [6] look in to both temporal and spatial dynamics of online behavior using a combination of Twitter messages and Foursquare checkins, but do not combine these into an aggregate spatiotemporal measure of behavior, nor look at data external to social media. Li et al. [7] do consider simultaneous temporal and spatial patterns on two social media services, Twitter and Flikr, as well as demographic characteristics of the population; this paper differs by comparing spatiotemporal patterns of online with offline behavior.

This paper offers multiple perspectives on how one might approach the problem of spatiotemporal correlation of activities. We investigate approaches stemming from statistics, geographic information systems, and machine learning. We explore the hypothesis that patterns of activity on Twitter are more related to those of mobile phone usage than to a broader conception of internet usage.

In the end, we find that social media is definitely correlative and can potentially be used as a predictor of some offline events, such as donations. To that extent, in future work we will show how a manager can use these findings to make decisions about how to schedule messaging to perspective donors.

2 Data

The primary component of our data is a set of geotagged tweets we collected using TwEater, an open source program developed at the Center for Complexity in Business at the University of Maryland. We focussed on any tweets that contained hashtags (topic identifiers on Twitter) and raw keywords related to Hurricane Irene (i.e., #irene, irene, #hurricane, hurricane, #hurricaneirene). Though the original dataset contained many more tweets, we used only the geotagged tweets in this study. This produces 22,045 tweets with corresponding latitude and longitude, which was roughly 1% of all tweets containing the listed keywords. In addition, tweets falling in the Continental United States were reverse geocoded to extract their state and approximate ZIP Code.

In addition, we have information on donations received by a large, not-for-profit relief organization through both mobile phone messaging and web channels. The former consists of 10,073 mobile donations and the latter of 28,214 web donations. Mobile donations do not have any address data. To cope with this we estimate a ZIP code and latitude/longitude from the donating phone number based on the NANPA area code and exchange prefix. This was unnecessary for the web donation data since each record had an associated billing address, allowing us to geocode the latitude and longitude.

Finally, we make use of the number of traditional media stories published about Hurricane Irene, grouped at an hourly level. These data are collected from the websites of CNN, CBS News, and ABC News (which also includes a variety of AP reports). We use this data as an external in-

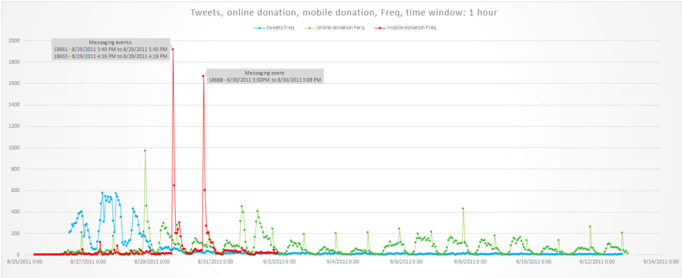

Fig. 1: Frequency of Tweets, web donations, and mobile donations

dicator of the overall amount of traditional media coverage.

All four data series (Twitter data, mobile donations, online donations, and mainstream news) are from 10:00 AM, 26 August, 2011 to 12:00 AM, 13 September, 2011, giving a coverage of approximately 21 hours before and 401 hours after Irene's US landfall.

3 Analysis

As mentioned above, we conduct three different types of analysis: (1) temporal analysis, (2) spatial analysis, and (3) spatiotemporal analysis. We will compare and contrast these three different analyses in the following sections.

3.1 Temporal

We use two different approaches for temporal analysis: a fixed effects and a variable autoregressive model.

3.1.1 Fixed Effects Model

We find that Tweets, mobile donations and web donations occurred regularly with a 24-hours time cycle. Hurricane Irene worked as an exogenous factor for the frequency of Tweets. Most of the Tweets occurred between 8/26–8/30 when Hurricane Irene battered the East Coast. Tweet frequency declined quickly after Hurricane Irene left, and finally became stationary. However, web donations were stationary and regular throughout this time period. For mobile donations, we notice that two spikes occurred in the series. Upon further investigation we discovered that the spikes in mobile donation activity appeared to be direct responses to the relief agency sending out solicitations for donations via text message. We believe these solicitations were the main cause of the peaks in the mobile donation series (Fig. 1).

We first apply a simple autoregressive, AR(p), model to the mobile donations and Tweets series respectively. For both of them, the model fits best when we include lag 1 and lag 24. Then we combine the two AR(p) models to create the following n-step-ahead forecasting model

$$y_t = \alpha_1 y_{t-1} + ... + \alpha_{t-q} y_{t-q} + \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + ... + \beta_p x_{t-p} + v_t + u + e_t \quad (1)$$

where y_t is the log of number of mobile donations at time t. x_t is the log of number of Tweets at time t, v_t represents the message event, e_i is the error term, i.e., when the relief agency specifically requested donations. We get the best fit with an adjusted $R^2 = 0.702$ when we include lag 1 and lag 24 for mobile donations and Tweets. The result shows that the message event worked as a strong exogenous factor for the mobile donations. This impulse intervention increased the number of mobile donations intensely. The regression results are shown in Fig. 2.

We apply the same procedure to the web donations and Tweets series. We get the best fit with adjusted $R^2 = 0.816$ when we include lag 1, lag 24, and lag 72 for web donations and lag 0 for Tweets. We find that more Tweets about the storm are associated with an increase in web donations.

3.1.2 Vector Autoregressive Model

We next test the temporal effects of social and traditional media mentions within the disaster relief context. We analyze the dynamic relationship between these forms of media and charitable donation behavior using a Vector Autoregressive Model with Exogenous Variables (VARX) [8]. This methodology accounts for the potentially endogenous relationship between traditional media interest, social media activity, and charitable donation behavior.

The first step in a VAR analysis is to test for unit roots for each of our variables. If a variable exhibits a significant unit root (Fig. 3), the process is not stationary, requiring the use of either differenced variables, in which we analyze the change in each variable from the prior time period, or a Vector Error Correction model in the case of cointegration, which is when two variables evolve with a common stochastic drift. We use an Augmented Dickey-Fuller test, a common method for testing for unit roots, to test whether our variables are stable or evolving. Over a 40 hour window, measured at an hourly level, we find that only Tweets exhibit a significant unit root. Therefore, we estimate a VARX model in differences due to the issues with stability for one variable and the lack of cointegration [8].

Our specific VARX model is given in Eq. 2, where ΔD_t is the change in total donations from hour $t-1$ through t, ΔT_t is the hourly change in tweets about Irene, and ΔM_t is the hourly change in traditional media stories, and [EX] represents a matrix of exogenous variables which are multiplied by an estimated matrix of coefficients β. We estimate a series of hourly indicators to control for differing media and donation behavior throughout the day. We also control for two events in which the relief agency asked for mobile donations through a text message appeal, which led to increased mobile donation behavior in the following hour.

$$\begin{bmatrix} \Delta D_t \\ \Delta T_t \\ \Delta M_t \end{bmatrix} = \sum_{j=1}^{40} \begin{bmatrix} \pi_{11}^j & \pi_{12}^j & \pi_{13}^j \\ \pi_{21}^j & \pi_{22}^j & \pi_{23}^j \\ \pi_{31}^j & \pi_{32}^j & \pi_{33}^j \end{bmatrix} \begin{bmatrix} \Delta D_{t-j} \\ \Delta T_{t-j} \\ \Delta M_{t-j} \end{bmatrix} \dots + [EX][\beta] + \begin{bmatrix} u_{D,t} \\ u_{T,t} \\ u_{M,t} \end{bmatrix} \quad (2)$$

We estimate each of the π_{ik}^j terms, which describe the nature of the relationship between the endogenous variables and their lagged terms. So, for example, π_{23}^j captures the effect of prior traditional media mentions on the number of tweets about Hurricane Irene in the current period. This model, therefore, allows us to estimate 9 distinct effects and evaluate their importance when examining the dynamic nature of the relationships between different forms of media and the donation behavior of consumers.

The next step in this analysis is to estimate this model

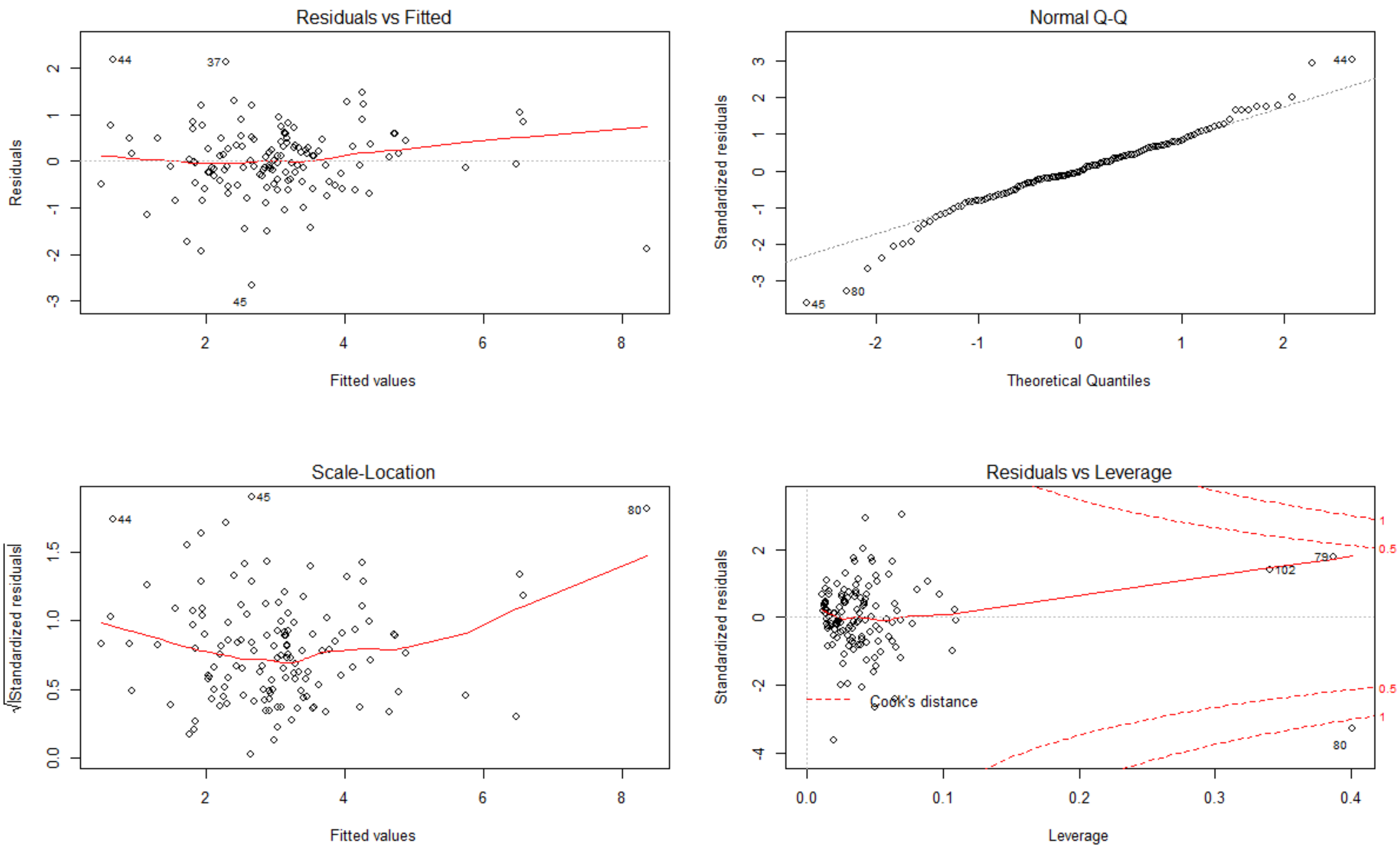

Fig. 2: Relationship of mobile donations to tweets.

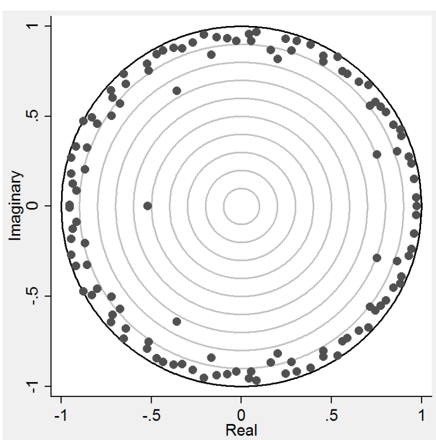

Fig. 3: Roots of the Companion Matrix

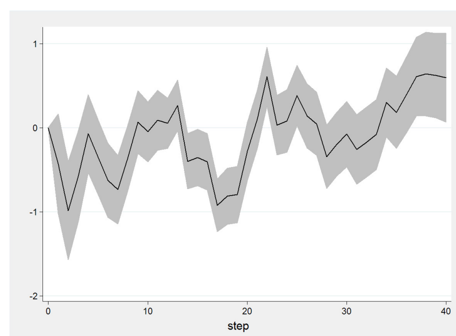

Fig. 4: The dynamic effect of tweets on donations.

and test that the results are stable. We find that this model explains a substantial amount of both donation behavior and media activity, with R^2 values of 0.97 for the change in the number of donations in the hour, 0.93 for the difference in the number of tweets about Hurricane Irene in an hour, and 0.89 for the number of traditional media mentions of the hurricane. We test the eigenvalue stability of the companion matrix for this model, which is a necessary condition for further analysis of the dynamic effects.

As this model displays stability, we now analyze the dynamic relationships using Impulse Response Functions. IRFs allow us to make inferences about the dynamic effects on a response variable of increasing an impulse variable by one unit. We provide graphs showing the Cumulative Impulse Response Functions for six of the nine tested endogenous relationships in Figures 4–9.

These results provide a number of important insights into the relationships between the two different forms of media and donation behavior. We find that additional tweets about the hurricane initially lead to lower donations for the charity (Fig. 4). However, eventually, prior tweets do result in a significant increase in donations that overcomes the initial drop in donation behavior. This result is interesting as it takes approximately 37 hours to see positive effects of tweets about the hurricane, which is somewhat surprising

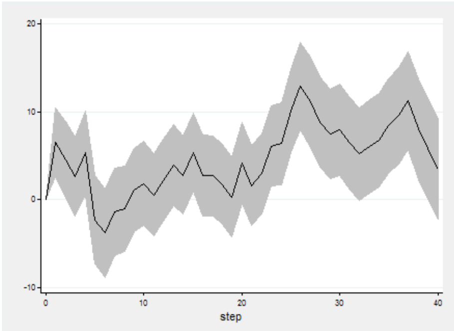

Fig. 5: The dynamic effect of traditional media on donations.

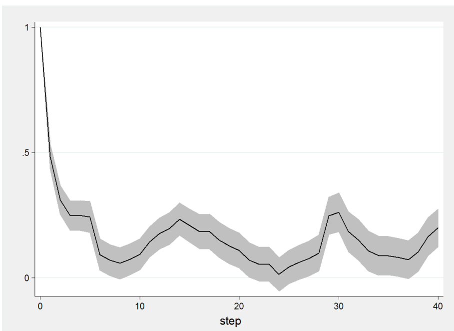

Fig. 6: The effect of adding a donation on future donations.

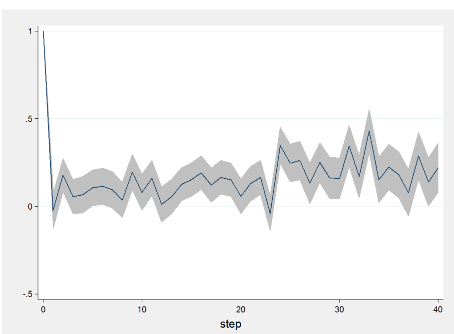

Fig. 7: The effect of adding a traditional media story on future traditional media coverage.

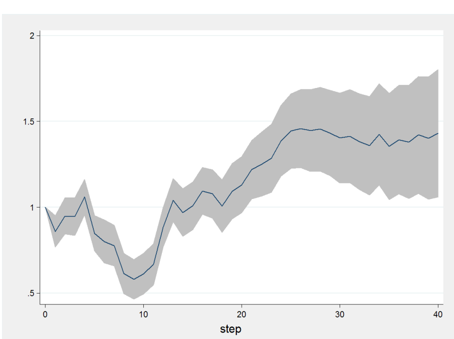

Fig. 8: The effect of adding a tweet on future tweets.

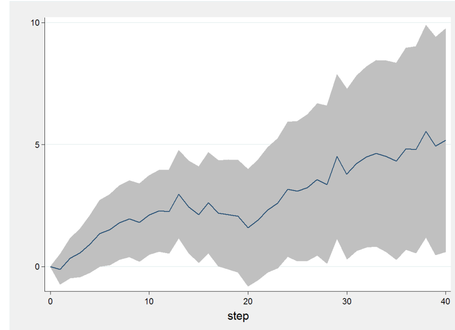

Fig. 9: The dynamic effect of traditional media mentions on tweets.

given the rapid nature of social media communications.

In comparison, traditional media stories have a relatively limited long-term effect on donation behavior (Fig. 5). The the relief agency sees an initial increase in donations due to traditional media stories. However, donation behavior resembles behavior in more traditional marketing contexts, such as sales and couponing, in which an initial burst in sales is followed by a trough, in which sales decrease due to consumers switching their purchasing to the sale period. This leads to the overall limited effect on total donations from increased traditional media exposure.

Increasing the difference in either the number of donations (Fig. 6) or of media stories (Fig. 7) results in relatively little permanent increase to the longer-term number of donations or media stories. Both of these variables start from the one unit added in order to calculate the IRF. However, the increase is soon lost to predicted lower future donations or media stories. Therefore increasing current donations or media stories has relatively little positive effect on the total number of donations or traditional media stories over the ensuing 40 hours. In contrast, and consistent with existing theory, the addition of a new Tweet about the hurricane is predicted to result in more Tweets about the hurricane in the future (Fig. 8). This captures the nature of information diffusion on Twitter, with users spreading information through social networks.

Finally, we find that further traditional media sources result in future Tweets (Fig. 9), demonstrating that while traditional media sources may not directly lead to increased donation behavior in the long-run, they can spawn conversations which lead to additional Tweets and thus additional conversations. This finding provides evidence for how firms should evaluate the impact of traditional media stories on consumer/donation behavior in the current environment of rapidly growing social media.

3.2 Spatial

For the spatial analysis, we apply both a manual segmentation and spatial autocorrelation statistics.

3.2.1 Manual Regional Characterization

Hurricane Irene had a different influence on different regions. Based on the track of Hurricane Irene, we divide the whole region into three areas, (1) Area A (Highly Impacted) — includes the states CT, DC, DE, MA, MD, ME, NC, NH, NJ, NY, PA, RI, VA, VT which were directly impacted by Irene, (2) Area B (Somewhat Impacted) — includes states that adjoined Area A (AL, FL, GA, KY, SC, TN, WV) and (3) Area C (Not Impacted) includes the rest of the states which were not directly impacted by Irene. Among them Area A occupied 28% of the overall mobile donations, 33% of the overall online danations, 79% of the overall Tweets and 26% of the overall population. Area B occupied 14% of the overall mobile donations, 9% of the overall online danations, 7% of the overall Tweets and 16% of the overall population. Area C occupied 58% of the overall mobile donations, 58% of the overall online danations, 13% of the overall Tweets and 58% of the overall population.

Our hypothesis is that Twitter is an indicator of interest in the hurricane. To test this we fit the following model

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + v_i + u + e_i \tag{3}$$

where y_i is the log of number of mobile donations (web donations) in state i. x_{1i} and x_{2i} is the log of number of Tweets and log of population in state i. v_i and e_i is the area level and error term. The results show that more Tweets and population result in more mobile donations and web donations. We also explore the difference of some characters between these three areas. By analysis of variance, we find that area B and C don't show a significant difference on Tweets/population, web donations/population and mobile donations/population. But they do have a significant difference between area A and C. Tweets/population, web donations/population and mobile donations/population in area A are higher. It indicates that people directly suffered from Irene tended to have more Tweets and tended to make more donations whether via mobile or web. We also find that mobile donations/Tweets and web donations/Tweets exist significant difference between area A, B and C, C is highest and A is lowest. We can infer this result since we know that most of Tweets were in area A.

3.2.2 Spatial Autocorrelation Statistics

We also calculate two measures of spatial auto-correlation common in the GIS literature: Moran's I [9] and Geary's C [10]. The results for each were similar, so we concentrate our discussion here on the latter. Geary's C is defined by

$$C = \frac{(N-1)\sum_i \sum_j w_{ij}(X_i - X_j)^2}{2(\sum_i \sum_j w_{ij})(\sum_i (X_i - \overline{X})^2)} \tag{4}$$

where X_i is the count of events in region i, N is the number of regions, $\overline{X}$ is the mean of all X_i and w_{ij} is the spatial weight between regions i and j. This weight is an inverse function of distance; closer regions are given higher weights. We adopt three methods of assigning weights: a threshold, a logistic sigmoid, and an exponential function. All three produced very similar outcomes, and so here we report the results using the threshold function, in which two regions were assigned a weight of 1.0 if they are less than 100 miles apart, and 0.0 otherwise. The range of C is $[0, 2]$, with 1.0 meaning there is no spatial autocorrelation present in the data. Values lower than 1.0 signify positive correlation (i.e. there tend to be large, contiguous areas of high and low density) while higher values represent negative correlation (areas of high and low density are intermixed).

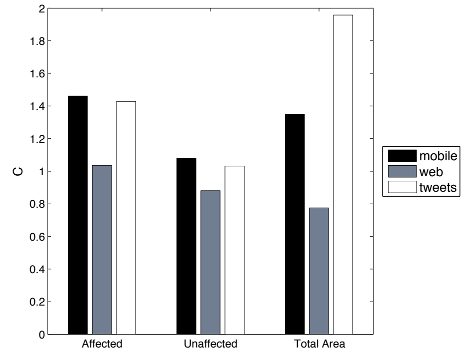

Fig. 10: Geary's C calculated for mobile and web-based donations and number of tweets per ZIP code.

We calculate the autocorrelation of the tweet, online and mobile donation series, aggregated at the ZIP code level. In addition, we partition the country into two regions, which we term the "affected" and "unaffected" zones of Hurricane Irene, with the former defined to be the ZIP codes in North Carolina, Virginia, Maryland, DC, Delaware, New Jersey, southern New York and New England. The results can be seen in Fig. 10. In the country as a whole and in both the affected and unaffected partitions, we find that the pattern of Twitter activity is more similar to that of mobile than web donations: both exhibit negative autocorrelation, while web donations are mildly positively autocorrelated. This is more pronounced in the area directly affected by Irene.

3.3 Spatiotemporal

In this section we conduct both a fixed effects model broken up by region, and a spatiotemporal analysis tool building off the work presented in the previous section.

3.3.1 Fixed Effects Model

To combine the benefit of spatial analysis and a temporal point of view we introduce the following spatiotemporal model to simultaneously estimate the spatial and temporal influence of Tweets on mobile donations,

$$\begin{aligned} y_{t,i} &= \alpha_1 y_{t-1,i} + ... + \alpha_{t-q} y_{t-q,i} \\ &\quad + \beta_0 x_{t,i} + \beta_1 x_{t-1,i} + ... + \beta_p x_{t-p,i} \\ &\quad + v_t + \gamma z_i + \delta_i + \mu_i + u + e_{t,i} \end{aligned} \tag{5}$$

where y_t is the log of number of mobile donations at time t, x_t is the log of number of Tweets at time t, z_i is the log of population at state i, δ_i is the area level, μ_i is the state level, v_t is the message event, e_i is the error term, and u is the intercept.

We get the best fit with adjusted $R^2 = 0.600$ when we include lag 1 and lag 24 for mobile donations and lag 0 for Tweets. We find that there exists significant differences between states. States from Area A which directly suffered from Irene tended to make more donations. Solicitations from the relief agency worked as the impulse intervention were the main cause of the donation peaks in the mobile donation series. We also introduce the area level into the regression slope to get a more sophisticated model as following. The result shows that there is no significant difference between regression slope obtained by the following model and our original model.

$$\begin{aligned}y_{t,i} = \alpha_{1,\delta_i} y_{t-1,i} + \ldots + \alpha_{t-q,\delta_i} y_{t-q,i} \\ + \beta_{0,\delta_i} x_{t,i} + \beta_{1,\delta_i} x_{t-1,i} + \ldots + \beta_{p,\delta_i} x_{t-p,i} \\ + v_t + \gamma z_i + \delta_i + \mu_i + u + e_{t,i}\end{aligned} \quad (6)$$

3.3.2 Spatiotemporal Cluster Analysis

To further examine the spatiotemporal patterns, we begin by running a k-means clustering algorithm [11] on all events, and recording the location of each cluster's centroid, c_i. We also calculate the standard deviation σ_i of the distance between each point in a cluster and its assigned centroid.

We then divide the data into equal frequency temporal bins of 1000 tweets. The start and end time of each window is used to bin the donation data. We re-run the clustering algorithm on only the events occurring in the current time window, using the overall cluster centroids as initial conditions. We then calculate the deviations of the centroids for the time window from the overall centroids according to $z_{it} = \|c_i - c_{it}\| (n_{it}\sigma_i)^{-1}$, where n_{it} is the number of events in cluster i and time window t. After repeating this procedure for all time windows, we take the mean over all z_{it} to be the final spatiotemporal variation of the data series.

Like all k-means based processes, this is subject to an arbitrary choice of the number of clusters. Following from choices in Section 4.2.1, we conduct our analysis here with 3 clusters. We find that the Twitter series has $z = 0.244$, while $z = 0.215$ for mobile donation and $z = 0.192$ for web donations. Once again, we see that the Twitter data is the most volatile, and that mobile behavior is more similar to Twitter than is online behavior in general.

4 Conclusions

In this research, we analyze the temporal, spatial, and spatiotemporal components of donation behavior and media activity within a hurricane relief context. We find a variety of evidence showing that social media activity is a leading indicator of donation behavior. Specifically, increases in Tweets lead to later increases in hurricane relief donations. However, possibly due to the nature of social media and donations, in which social media occurs during the storm followed by donations after the storm, we find a relatively substantial lag between Tweets and donations. The relationship between social media activity and donations are found for both online and mobile donations. From a spatial context, those users/individuals directly impacted by the storm display both greater Twitter activity and donate for disaster relief more often.

In addition to our results on the relationship between social media and donation behavior, we study the relationship between traditional media, social media, and donation behavior. While traditional media does not directly increase the number of donations to hurricane relief, it does increase social media activity, which has a positive relationship with donation behavior.

To refine our models, we plan to conduct a number of additional investigations. For instance, our regression method is dependent on the model settings and variable selection. In our case, we use forward selection. We will explore other variable selection methods such as the Akaike or Bayesian Information Criterion [12]. We will also explore a decomposition of time series such as a method based on rates of change or on predictability prior to the regression analysis.

Spatially, our cluster analysis method is sensitive to the initial choice of the number of clusters and size of temporal window. We are currently investigating how sensitive these parameters are. In addition, we are also exploring other clustering algorithms such as k-mediods [13] and DBSCAN, as well as alternative methods of aggregating the individual cluster-time window values into a single scalar measure.

Our results underline that both spatial and temporal components need to be considered when attempting to more accurately characterize both donation and social media behavior, and their interrelationships, in the disaster relief context. In future research on this topic, we plan to further develop our spatiotemporal results to provide more actionable strategic results, with the goal of better describing and predicting donation behavior based on the activity in both social and traditional media. In addition, we plan to build a decision support tool for managers interested in using this data and these models.

References

[1] Y. Mejova, V.R. Kiran, and M.C. Dougal. Giving is caring: Understanding donation behavior through email. In *17th ACM Conf. on Computer Supported Cooperative Work & Social Computing*, pages 1297–1307, 2013.

[2] A.L. Hughes and L. Palen. Twitter adoption and use in mass convergence and emergency events. *Int'l J. of Emergency Management*, 6:248–260, 2009.

[3] L. Palen, S. Vieweg, S.B. Liu, and A.L. Hughes. Crisis in a networked world: Features of computer-mediated communication. *Social Science Review*, 27:467–480, 2009.

[4] J. Herrmann, W. Rand, B. Schein, and N. Vodopivec. An agent-based model of urgent diffusion in social media. *SSRN*, 2297167, 2013.

[5] J. Yang and J. Leskovec. Patterns of temporal variation in online media. In *Proc. 4th ACM Int'l Conf. Web Search & Data Mining*, pages 177–186, 2011.

[6] N. Grinberg, M. Naaman, B. Shaw, and G. Lotan. Extracting diurnal patterns of real world activity from social media. In *Proc. AAAI Int'l Conf. on Weblogs & Social Media*, 2013.

[7] L. Li, M. Goodchild, and B. Xu. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography & Geographic Information Science*, 40(2):61–77, 2013.

[8] M.G. Dekimpe and D.M. Hanssens. Persistence modeling for assessing marketing strategy performance. In D. Lehmann and C. Moorman, editors, *Assessing Marketing Strategy Performance*. Marketing Science Institute, 2004.

[9] P.A.P. Moran. Notes on continuous stochastic phenomena. *Biometrica*, 37:17–23, 1950.

[10] R.C. Geary. The contiguity ratio and statistical mapping. *Incorporated Statistician*, 5(3):115–145, 1954.

[11] J.A. Hartigan and M.A. Wong. Algorithm AS 136: A k-means clustering algorithm. *J. Royal Statistical Soc. C*, 28 (1):100–108, 1979.

[12] K.P. Burnham and D.R. Anderson. *Model selection and multimodel inference: A practical information-theoretic approach.* Springer, 2002.

[13] H.-S. Park and J. Chi-Hyuck. A simple and fast algorithm for k-medoids clustering. *Expert Systems with Applications*, 36(2):3336–3341, 2009.