

Using coarse GPS data to quantify city-scale transportation system resilience to extreme events

Brian Donovan

Department of Civil and Environmental Engineering
University of Illinois at Urbana-Champaign
bpdonov2@illinois.edu

Dan Work (corresponding author)

Department of Civil and Environmental Engineering and Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
dbwork@illinois.edu

4731 words + 6 figures + 2 tables = 6731 words

ABSTRACT

This article proposes a method to quantitatively measure the resilience of transportation systems using GPS data from taxis. The granularity of the GPS data necessary for this analysis is relatively coarse; it only requires coordinates for the beginning and end of trips, the metered distance, and the total travel time. The method works by computing the historical distribution of pace (normalized travel times) between various regions of a city and measuring the pace deviations during an unusual event. This method is applied to a dataset of nearly 700 million taxi trips in New York City, which is used to analyze the transportation infrastructure resilience to Hurricane Sandy. The analysis indicates that Hurricane Sandy impacted traffic conditions for more than five days, and caused a peak delay of two minutes per mile. Practically, it identifies that the evacuation caused only minor disruptions, but significant delays were encountered during the post-disaster reentry process. Since the implementation of this method is very efficient, it could potentially be used as an online monitoring tool, representing a first step toward quantifying city scale resilience with coarse GPS data.

INTRODUCTION

Motivation

In recent years, resilience of city infrastructure has gained a great deal of attention [1]. When disasters and other extreme events occur, infrastructure may fail, incurring large human, economic, and environmental costs. This is especially relevant for transportation infrastructure, since it is crucial for city evacuations and emergency services in post-disaster environments. Methods are needed to quantitatively monitor the transportation infrastructure in terms of its ability to withstand and recover from such events. Measuring the performance of city-scale infrastructure with traditional traffic sensors is cost-prohibitive due to relatively high installation costs, but many cities already collect data from taxis equipped with GPS sensors. Though this analysis could be performed with any GPS data, taxi data is publicly available in many cases. The New York City dataset used in this analysis gives interesting insights about the performance of infrastructure during Hurricane Sandy and other major events.

The goal of this article is to develop and implement a method for measuring resilience of city-scale transportation networks using only publicly available taxi datasets. The technique is designed with the following characteristics:

1. **The method can be applied at the city-scale, or larger.** Extreme events such as hurricanes have the ability to affect an entire city. For this reason, it is important to examine impacts at a high-level city view, rather than the level of individual vehicles or streets.
2. **The method measures network performance quantitatively, in terms of recovery time and peak pace deviations.** Recovery time and peak performance degradation are fairly standard quantities of interest in the resilience literature [2, 3]. While travel times are a natural performance measure for transportation networks, we instead use *pace*, or travel time per mile. This normalization accommodates the varied length of taxi trips within a city.
3. **The method accommodates inherent variability in traffic conditions and data.** The available data is full of noise and depends on many unmodeled human factors. As a result, the method evaluates events that cause statistically significant disruptions, in order to separate the signal from the noise.
4. **The method is computationally tractable.** Since taxi trips occur very frequently in large cities, the amount of data available for analysis is large. In order to be tractable, the computation should be $O(N)$, where N is the number of taxi trips, and ideally require only one pass through the raw data. Of practical significance, these single-pass algorithms could also be used to process the data in a realtime stream.

Related Work

In recent years, the study of resilience has gained popularity in the systems engineering community. Haimes [3, 4, 5] gives a framework for assessing resilience, which focuses on modeling a system and the possible outcomes of various events. He asserts that a resilient system should suffer only slight degradation during an event, then rapidly recover. Reed et al. [6] note that the quality of

service abruptly drops during an event, then exponentially decays back to typical values. They suggest that an appropriate resilience measure is the integral of this exponential curve. Authors in the related field of risk analysis emphasize the importance of unknown factors while assessing resilience [2, 7].

Though there is no precise consensus on the definition of resilience, *peak disruption* and *recovery time* are consistently discussed quantities. In other words, peak disruption measures how far the quantity of interest deviates from typical values, and recovery time measures how long it takes to return to typical values. Most of these works also emphasize that resilience must be measured with respect to a given event and quantity of interest. For example, one case study used the number of functioning nodes in a power grid as the quantity of interest, assessing resilience against hurricanes and minor events [8]. Several authors have proposed quantities of interest for transportation systems. Omer et al. [9] proposed a method which measures the resilience of a road-based transportation network in terms of travel times between cities. Chang et al. [10] evaluated a post-earthquake transportation network in terms of accessibility and coverage. This is partly based on an accessibility metric devised by Allen et al. [11], which considers travel times between various regions of a city. Thus, travel time is a standard quantity on which to measure resilience. This article will use the related quantity of pace, or travel time per mile.

A distinct set of studies use large amounts of data to extract useful information about urban systems. The work most closely related to resilience is a study by He and Liu [12], which uses loop detector data to measure the effect of the I-35W bridge collapse in Minneapolis in 2007. Geroliminis et al. [13] use loop detector data, combined with 500 GPS vehicles to extract macroscopic traffic properties from an urban-scale transportation network. Other works use GPS traces of mobile devices to analyze movement patterns of crowds during typical days and atypical events [14, 15]. Castro et al. [16] present a method for inferring current and future traffic states from taxi GPS data. Zheng et al. [17] propose a method that tracks taxi trips between various regions of a city and identifies flawed urban planning. Another study measures temporal patterns in the density of taxi pickups and dropoffs to identify the social function of various city regions [18]. They point out that unusual output can be used to detect events like holidays. Chen [19] specifically focuses on identifying anomalous taxi trajectories, in order to detect fraud or special events. Ferreira et al. [20] created a graphical querying tool which can be used to count taxi trips between arbitrary geometrical regions as a function of time. They noted the drop in the frequency of taxi trips during Hurricane Sandy and Hurricane Irene, pointing out that the Irene-related drop was more significant, but the Sandy-related drop was longer lasting. By examining pace, we confirm that Hurricane Sandy had a longer recovery time, but find the contrasting result that Hurricane Sandy also has a more significant peak disruption.

Outline and Contributions

The contributions of this work are as follows. In Section 2, a method is proposed to use taxis as pervasive city-scale resilience sensors. This method detects unusual events and measures them in terms of peak disruption and recovery time. It introduces paces between regions of the city as the key performance measure, and it uses the historical pace distribution to detect and measure extreme events. In Section 3, the method is applied to a four-year dataset from New York City to identify and compare properties of events such as Hurricane Sandy. Of practical significance, the analysis identifies the relative efficiency of the pre-Sandy evacuation, contrasted with the gridlock of post-Sandy reentry. Conclusions and future work is summarized in Section 4. As a technical

contribution, all code [21] and data [22] used in this analysis are made publicly available.

METHODOLOGY

Overview

The proposed technique to measure city-scale resilience of the transportation network in response to various events by examining taxi trip data is done in three steps. In section 2.2, individual taxi trips are aggregated by origin-destination pairs in order to measure typical paces between various regions of the city. This aggregation technique makes it possible to extract city-scale features at various points in time, since it is difficult to measure resilience from individual trips. Section 2.3 imposes a one-week periodic pattern on the paces. Defining the variance across weeks as a probability distribution allows for a description of typical behavior that accounts for noisy day-to-day variations. Finally, Section 2.4 uses these distributions to quantify how typical or atypical the pace is at a particular point in time. Atypical times (e.g. the 5% most unlikely points in time) are flagged as events, and they are examined in more detail.

Extraction of Time-Series Features from Aggregated Trips

In the first stage of analysis, trips are grouped by their geographic locations and times of occurrence. More specifically, the city is divided into a small number, k , of large regions. This allows each taxi trip to be labeled as one of k^2 unique origin-destination pairs. Time is discretized into hours, so a large sample of trips can be gathered at any point in time. The start zone, end zone, and departure time are used to partition all of trips into subsets. The variable $T_{i,j,t}$ denotes the set of all trips from zone i to zone j at time t :

$$T_{i,j,t} = \{r | o(r) \in z(i), d(r) \in z(j), \lfloor s(r) \rfloor = t\}, \quad (1)$$

where $o(r)$ is the origin of trip r , $d(r)$ is the destination of trip r , $z(i)$ is the geographic region of zone i , and $\lfloor s(r) \rfloor$ is the start time of trip r , rounded down to the hour. It is assumed that i and j are both in $\{0, 1, \dots, k-1\}$. Once these subsets of trips are defined, macroscopic traffic features can be extracted from them. Of particular interest is the expected travel time between two regions. However, travel times of individual vehicles between two regions are not uniform, due to the varying lengths of trips that connect the same regions. Much of this variation can be accounted for by normalizing against distance. In this way, the *average pace* is computed for each trip subset $T_{i,j,t}$. Trips are weighted by their distance, since longer trips give more information about the state of traffic. In this way, the distance-weighted average pace, $P(i, j, t)$, of taxis from zone i to zone j at time t is computed:

$$P(i, j, t) = \frac{\sum_{r \in T_{i,j,t}} l(r)p(r)}{\sum_{r \in T_{i,j,t}} l(r)} = \frac{\sum_{r \in T_{i,j,t}} l(r) \frac{u(r)}{l(r)}}{\sum_{r \in T_{i,j,t}} l(r)} = \frac{\sum_{r \in T_{i,j,t}} u(r)}{\sum_{r \in T_{i,j,t}} l(r)}, \quad (2)$$

where $u(r)$ is the travel time of trip r , $l(r)$ is the metered length of trip r , and $p(r) = \frac{u(r)}{l(r)}$ is the pace of trip r . For a fixed value of t , all k^2 distance-weighted average paces collectively form the *mean pace vector*. This vector is a function of time, and contains the k^2 pace values at a particular point in time. Specifically, the n th element is given by

$$\mathbf{A}(t)_n = P\left(\left\lfloor \frac{n}{k} \right\rfloor, n \bmod k, t\right), \quad (3)$$

where $n \in \{0, 1, 2, \dots, k^2 - 1\}$.

Identification of City-Scale Typical Behavior

The mean pace vector, $\mathbf{A}(t)$, has a strongly periodic weekly pattern. During rush hour, the pace is high, especially in dense downtown regions, and at night the pace is low. On weekends, the rush hour is less extreme. However, the mean pace vector has some variance around this periodic pattern, so it is viewed as a distribution conditioned on time. For example, the mean pace vector for all Tuesdays at 3pm will be slightly different, especially during an unusual event. To facilitate this grouping, the *reference set* Q_t is defined for all times t . This set contains all of the mean pace vectors which occur at the same point in the periodic pattern as $\mathbf{A}(t)$, except for $\mathbf{A}(t)$ itself. Intuitively, when deciding how typical the traffic data is at time t , that data should not be used as part of the definition of typical. Since there are 168 hours in a week, the reference set can be defined as

$$Q_t = \{\mathbf{A}(h) | h \equiv t \pmod{168}, h \neq t\}. \quad (4)$$

The reference set Q_t makes it possible to compute the expected value of the mean pace vector $\mu(t)$ as well as the covariance matrix $\Sigma(t)$. This covariance matrix is important because it quantifies the noisy day-to-day fluctuations in the mean pace vector, outside of the event at hand, and how the dimensions correlate. The time-dependent sample mean and covariance matrices can be defined as:

$$\begin{aligned} \mu(t) &= \frac{1}{|Q_t|} \sum_{\mathbf{A} \in Q_t} \mathbf{A} \\ \Sigma(t) &= \frac{|Q_t|}{|Q_t|-1} \left(\sum_{\mathbf{A} \in Q_t} \frac{\mathbf{A}\mathbf{A}^\top}{|Q_t|} - \mu(t)\mu(t)^\top \right). \end{aligned} \quad (5)$$

If an independence assumption is desired, the diagonal components of these matrices can be extracted. However, it is likely that many of the k^2 dimensions of $\mathbf{A}(t)$ are highly correlated, so the full covariance matrix is used for the remainder of the analysis. For example, trips that start or end in the same region often have highly correlated paces. Together, $\mu(t)$ and $\Sigma(t)$ make it possible to define a probability distribution, conditioned on time. A normally distributed random variable is defined for each hour:

$$\mathbf{X}_t \sim \mathcal{N}(\mu(t), \Sigma(t)). \quad (6)$$

Detection of Deviations from Typical Behavior

The random variables \mathbf{X}_t capture the historical distribution of the mean pace vector given the time of the week, so they can be used to determine if a mean pace vector is typical or atypical. If a mean pace vector $\mathbf{A}(t)$ is sufficiently far from its corresponding expected value μ_t , it will have a low probability density, given by

$$P(\mathbf{X}_t = \mathbf{A}(t)). \quad (7)$$

An unfortunate property of the probability density given in (7) is that the random variables \mathbf{X}_t can have different variance at different times of the day (i.e. different values of t). Specifically, the covariance tends to be high during the day and low at night. The probability mass is more

spread out during the day, so the probability density will be lower on an average day than an average night. Rather than using different thresholds to define unusual nighttime and daytime behavior, the two quantities are normalized. Specifically, the *relative conditional probability density*, $R(t)$, divides the probability density by the maximum possible probability density. This is equivalent to a *relative likelihood function* [23], except it is viewed as a function of the observation $\mathbf{A}(t)$ instead of the parameters $\mu(t)$ and $\Sigma(t)$. Formally,

$$R(t) = \frac{P(\mathbf{X}_t = \mathbf{A}(t))}{\max_{\mathbf{x}} P(\mathbf{X}_t = \mathbf{x})}. \quad (8)$$

Since the multivariate Gaussian distribution achieves maximum probability at its mean, this division simply has the effect of removing the normalization constant $((2\pi)^k |\Sigma|)^{-\frac{1}{2}}$, where $|\Sigma|$ is the determinant of Σ and k is the number of dimensions. The relative conditional probability density, $R(t)$, is the key measure of typical versus atypical behavior, and is used to measure recovery time. When $R(t)$ falls below a given threshold, an event is detected. The event is declared complete when $R(t)$ returns to a value higher than the threshold. In this work, the choice of the threshold is the 5% quantile of $R(t)$, but this value can easily be lowered to detect smaller events or raised to detect only the most severe events. The function $R(t)$ is a fairly noisy, which means that it can occasionally *thrash* over the threshold. In other words, $R(t)$ may rise above the threshold, then immediately drop back below it, effectively breaking the event into two pieces. To prevent this, consecutive events separated by fewer than six hours are merged. Figure 1 illustrates this process.

Once the recovery time of an event is computed, other properties can be computed. For example, it is possible to compute the maximum pace deviation, or the slowest type of trip during the event. Thus, each event can be described with a set of meaningful statistics. Comparisons between various events make it possible to describe which types of events the city is resilient to, and where there is room for improvement. For longer-lasting events like Hurricane Sandy, it is possible to examine different stages of the event in greater detail.

APPLICATION TO HURRICANE SANDY WITH NEW YORK CITY TAXI DATA

In this section, the previously described methodology is applied to a dataset of New York City taxi trips. This dataset, which was obtained through a *Freedom of Information Law* (FOIL) request, covers four years of operation and details over 700 million trips. Many events are detected within this four year span and compared quantitatively. Special attention is given to the Hurricane Sandy and some interesting properties are discovered.

The Dataset

The data used in this analysis takes the form of a massive table where each row represents a single taxi trip. Table 1 gives a small sample of this data. Note that this data format is the minimum amount of information required to perform the analysis. Other datasets may contain, for example, periodic GPS updates, but this is at least as much information as the New York City data. As there are several entries per second for four years, the raw data takes up about 116GB in text CSV format. We have made this large dataset publicly available [22].

The dataset also contains a large number of errors. For example, there are several trips where the reported meter distances are significantly shorter than the straight-line distance, violating Euclidean geometry. Additionally, many trips report GPS coordinates of (0,0), or cover impossible

Event Detection – Thrashing

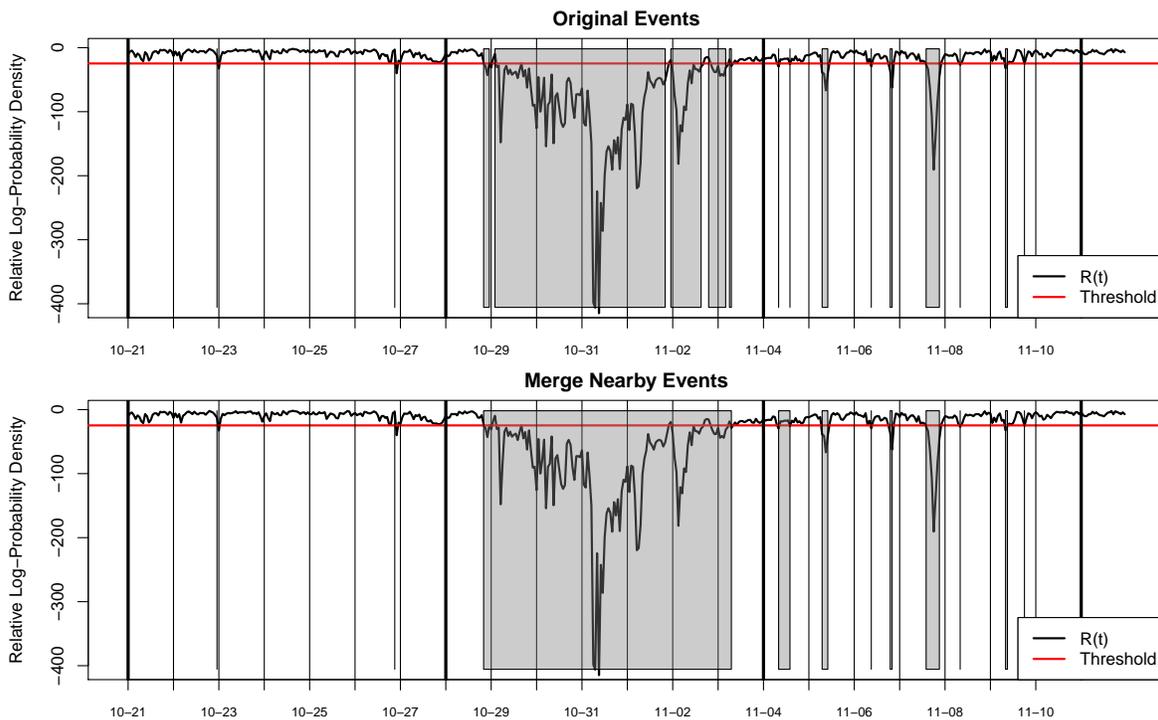


FIGURE 1 : Demonstration of event detection. Events are detected when $R(t)$ goes below the threshold, but thrashing often occurs. The top graph shows that this thrashing causes events to be divided into several pieces. For this reason, events with fewer than six hours between them are merged, as shown in the bottom graph. Note that many small events are detected, but it is generally more interesting to look at the longer, more significant events.

pickup datetime	dropoff datetime	duration (sec)	distance (mi)	pickup lon	pickup lat	dropoff lon	dropoff lat
2013-05-01 00:02:11	2013-05-01 00:14:28	737	2.9	-74.00	40.74	-74.01	40.71
2013-05-01 00:02:12	2013-05-01 00:12:31	618	1.8	-74.00	40.73	-73.98	40.72
2013-05-01 00:02:12	2013-05-01 00:07:39	326	1.3	-73.97	40.76	-73.96	40.77
2013-05-01 00:02:13	2013-05-01 00:04:35	141	0.6	-73.99	40.75	-74.00	40.75
2013-05-01 00:02:14	2013-05-01 00:04:09	115	0.5	-73.98	40.75	-73.99	40.74

TABLE 1 : A small subset of the data used in this analysis. Each row corresponds to an occupied taxi trip.

distances, times, or velocities. All of these types of obvious errors are discarded and account for roughly 7.5% of all trips.

After removing errors, the dataset is then filtered to remove data outside of the scope of the analysis. For example, there are many trips which start in Midtown, travel over 50 miles, then end less than a block from their starting points. These trips are entirely possible, but unlikely to be representative of Midtown-to-Midtown trips because they likely drove many miles in other areas. This filter is implemented by thresholding the *winding factor*, or metered distance over straight-line distance. Other trips which last less than 60 seconds are unlikely to give accurate pace estimates because the initial non-driving time becomes more important. These types of trips are also removed, accounting for roughly 4% of the original data. Figure 2 shows histograms of all trip features considered for filtering, as well as the thresholds used for invalid data. Additionally, the entire months of August and September 2010 were discarded due to a high number of errors.

Computational Issues

Due to the size of the dataset, an efficient software implementation of the analysis is crucial. This section discusses the algorithmic and practical aspects of the analysis, using the New York City taxi dataset as an example. In this way, concrete figures can be used for quantities like runtime or data size. More general concepts like time complexity do not depend on the dataset.

The first step described in Section 2.2 is the most computationally expensive. All of the 697,622,444 individual trips are aggregated into 35,064 mean pace vectors - one for each hour in the four-year dataset. Since the trip data is sorted chronologically, it is possible to compute these mean pace vectors in a single pass. Recall from (2) that the mean pace computation involves the sum of trip durations and the sum of trip distances. Thus, these two sums are initialized to zero for each of the 16 types of trips. Each time a trip is read from the file, the relevant sums are incremented. The error filtering from Section 3.1 can also be performed at this stage, so an

Data Filtering

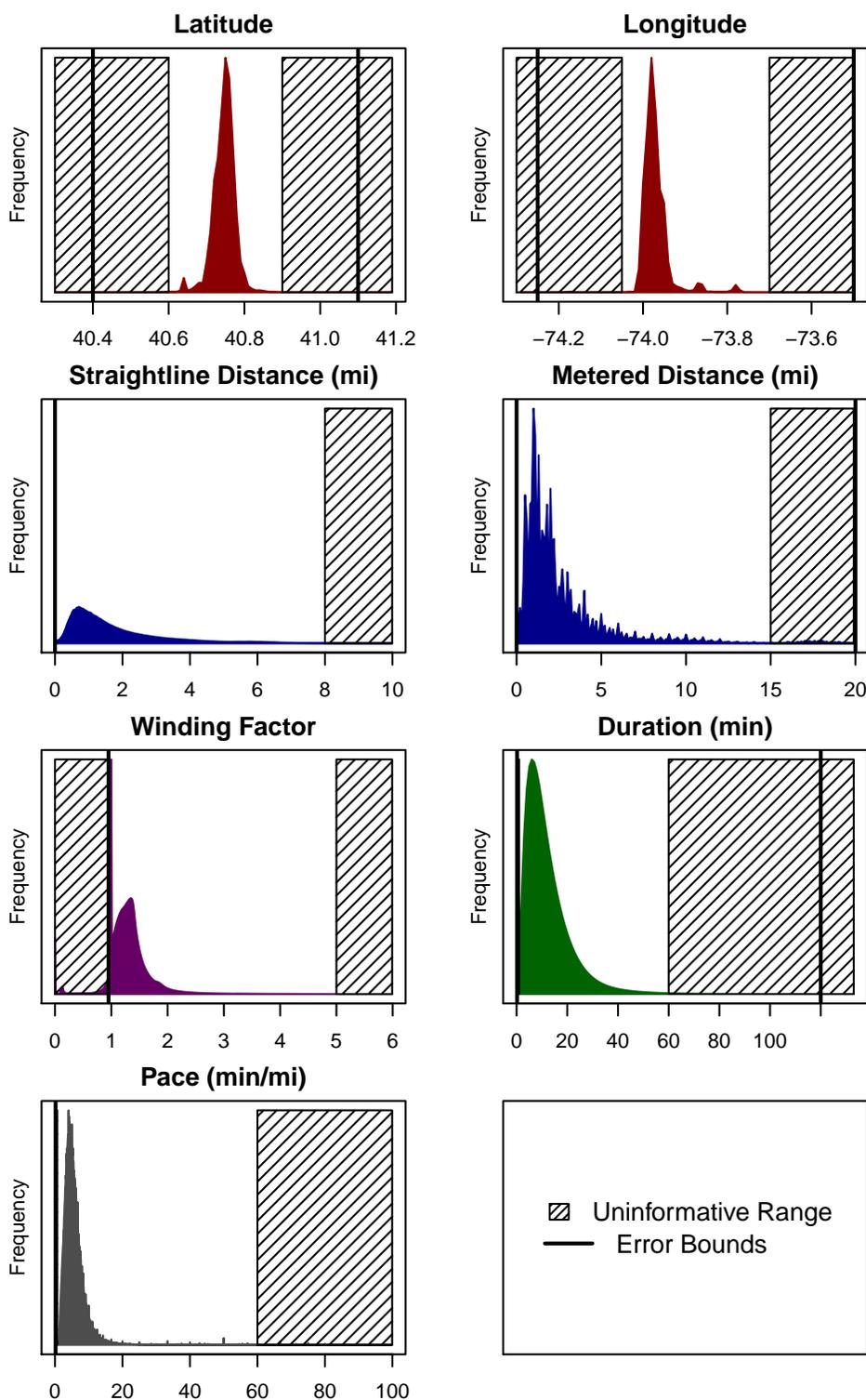


FIGURE 2 : Distributions of individual features of taxi trips. Simple thresholds are used to filter trips that contain errors, or are otherwise uninformative. Note that the *straightline distance* is the Euclidean distance between start and end coordinates, while the *metered distance* is the value returned by the odometer. The *winding factor* is the metered distance divided by the straightline distance. A winding factor less than 1 is geometrically impossible, and a large value indicates that the taxi did not proceed directly to its destination.

Algorithm 1 Online Mean Pace Vector Extraction

<pre> prev_hour := -1 sum_duration := zeros(NUM_TYPES) sum_distance := zeros(NUM_TYPES) for all trip ∈ chronological_trips do while trip.hour > prev_hour do output($\frac{\text{sum_duration}}{\text{sum_distance}}$) sum_duration := zeros(NUM_TYPES) sum_distance := zeros(NUM_TYPES) prev_hour += 1 end while if trip.isValid() then $i \leftarrow \text{category}(\text{trip.pickup}, \text{trip.dropoff})$ sum_duration[i] += trip.duration sum_distance[i] += trip.distance end if end for </pre>	<ul style="list-style-type: none"> ▸ Start at beginning of time ▸ Initialize sums to 0 ▸ Initialize sums to 0 ▸ Loop over all trips ▸ If previous hour is complete: ▸ Output mean pace vector ▸ Reset sums to 0 ▸ Reset sums to 0 ▸ Advance to next hour ▸ Data filtering ▸ Determine trip type ▸ Update distance sum ▸ Update duration sum
--	--

additional pass of the dataset is not required. When the start hour of the current trip (rounded) is greater than the start hour of the previous trip, the sums are complete for the previous hour. The mean pace vector is computed by division and output, then the sums are reset to zero. Thus, the computation is one large loop over the entire dataset. A short pseudocode is given in Algorithm 1. Note that NUM_TYPES is 16, since there are four regions.

Since each trip is accessed only once, the computation is $O(N)$, where N is the total number of trips. The computation of each hour timeslice is independent, making it possible to employ parallel processing if the data is partitioned ahead of time. The analysis was implemented in python (source code available at [21]) and run on an 8-core 2.5GHz machine with 24GB of ram. The extraction of all 35,064 mean pace vectors took about 75 minutes, using roughly 40MB of RAM for each of the eight processes. The fact that the runtime is much shorter than the real timespan of the dataset combined with the single-pass property means that this preprocessing could be performed in realtime. In other words, this system could realistically collect trips as they occur, update the relevant sums, then output the mean pace vector at the end of the hour.

The remaining computations involve mean pace vectors instead of raw trip data. They also have linear time complexity and are much faster than the preprocessing. Recall from (4) and (5) that, at a particular hour, the mean and covariance need to be computed for *all hours in the periodic pattern except that hour*. The naive implementation of this calculation has a quadratic time complexity, since each mean pace vector must be compared against every other mean pace vector. However, it is possible to compute all of these quantities in linear time. Instead of directly computing the mean of all values except $\mathbf{A}(t)$, the sum of all values including $\mathbf{A}(t)$ is computed up front. Then, in the loop, $\mathbf{A}(t)$ is subtracted from this sum. Formally, the *inclusive reference set*, Q_{t+} , is defined in a similar way to (4), except that it includes the mean pace vector $\mathbf{A}(t)$. In other words,

$$Q_{t+} = \{\mathbf{A}(h) | h \equiv t \pmod{168}\} = Q_t \cup \{\mathbf{A}(t)\}. \quad (9)$$

Unlike the reference set from (4), the inclusive reference set is identical for values of t that occur at the same point in the periodic pattern. Thus, Q_{t+} and the sum of all vectors in Q_{t+} only need to be computed once. To compute the sum of all vectors in Q_{t+} *except* $\mathbf{A}(t)$, it is sufficient to subtract $\mathbf{A}(t)$ from this sum. Thus, the mean computation can be written as

$$\mu(t) = \frac{1}{|Q_t|} \sum_{\mathbf{A} \in Q_t} \mathbf{A} = \frac{1}{|Q_{t+}| - 1} \left(\left(\sum_{\mathbf{B} \in Q_{t+}} \mathbf{B} \right) - \mathbf{A}(t) \right). \quad (10)$$

A similar technique is used for the sum of outer products in the covariance computation. This method avoids redoing most of the addition in each iteration, allowing for a significant improvement on large datasets. Once $\mu(t)$ and $\Sigma(t)$ are computed, $R(t)$ can be computed in constant time. Thus, the entire operation runs in linear time. On the same machine, this computation ran in less than 30 seconds, producing the timeseries of $R(t)$. Again, this operation would be feasible in a real-time system. However, it is worth noting that it may be desirable to re-generate old values of $R(t)$ in light of new information.

Once $R(t)$ is generated, the event detection described in Section 2.4 can also be performed in linear time. Events and spaces between events are stored as a linked list, where each node contains the start time and end time. Scanning through $R(t)$ chronologically, a new node in the linked list is generated each time $R(t)$ crosses above or below the threshold. Then, to remove short spaces between events, this linked list is iterated upon. Each time a non-event node of less than the desired duration is discovered, that node and its two neighbors are replaced with one larger node. On the same machine as the previous computations, it took less than one second to perform the event detection.

Extraction of Pace Features

The map of New York City is first split into four large regions, shown in Figure 3. For the remainder of the analysis, the zones will be referred to in the following way: *Upper Manhattan* (U), *Midtown* (M), *Lower Manhattan* (L), and *East of the Hudson River* (E). Note that the Eastern region is connected only by bridges and tunnels and thus problems with this infrastructure will tend to increase travel times between this region and others. Specifically relevant to Hurricane Sandy is the Lower Manhattan region, since it experienced severe flooding and power outages. Choosing four large regions in this way satisfies the first goal outlined in Section 1.1 because it defines meaningful city-scale properties. Instead of looking at every street in New York under a microscope, it defines large areas with key geographic and infrastructural properties. The travel times between these regions reflect the overall performance of city-scale transportation infrastructure. It is worth noting that the methodology allows for an arbitrary choice of regions. This implementation simply chooses zones which are useful for detecting the types of events that occur in New York City.

Recall that a taxi can take one of 16 possible trips between these regions. Aggregating these trips by type and hour as in Section 2.2 produces the 16-dimensional mean pace vector, $\mathbf{A}(t)$, at all points in time. Figure 4 shows three typical weeks of mean pace vectors, revealing the expected periodic pattern.

Analysis of Events

As detailed in Section 2.3, the expected behavior is generated for all times t according to $\mu(t)$ and $\Sigma(t)$. An interesting way to view the mean pace vector $\mathbf{A}(t)$ is by standardizing it, element by

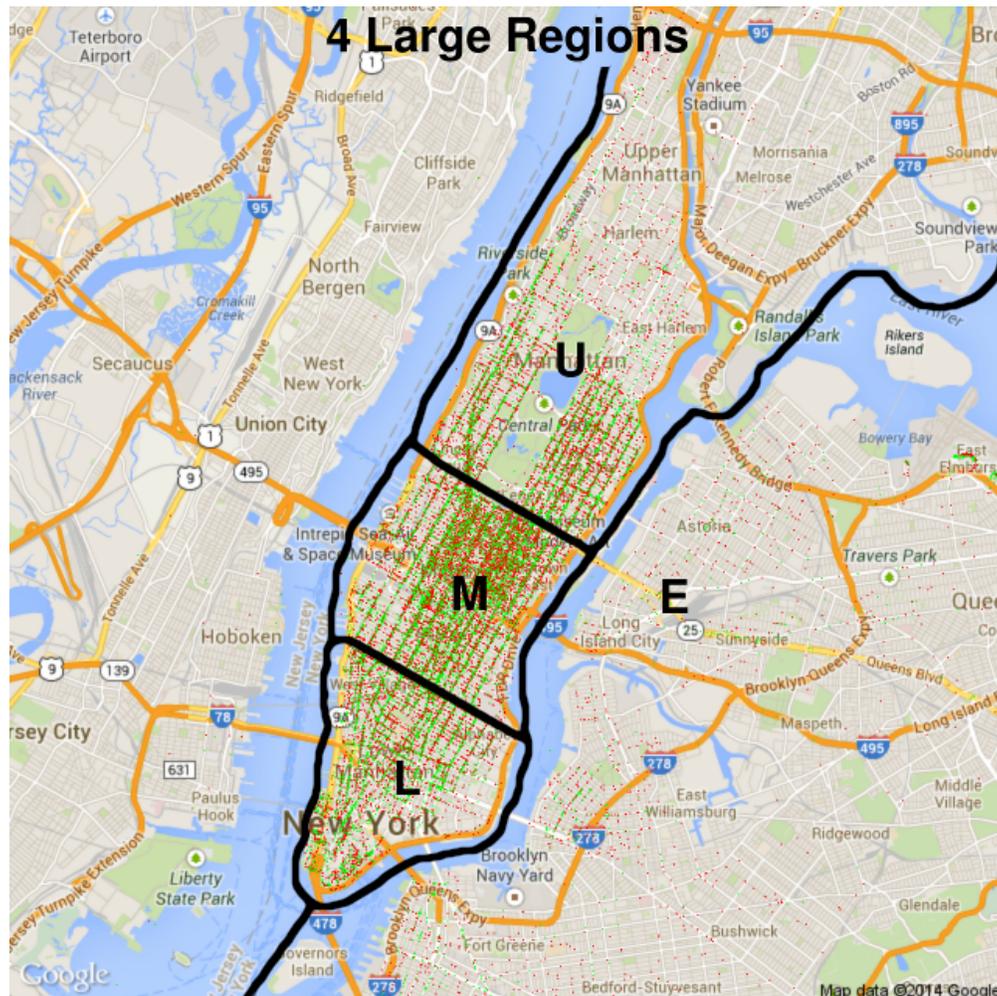


FIGURE 3 : Division of New York City into four large regions denoted U, M, E , and L . A random sample of 0.01% of the taxi trips in 2012 are shown. Pickups locations are marked in green, and the corresponding dropoffs are marked in red. The majority of trips occur in Manhattan, with especially high concentration in the Midtown region.

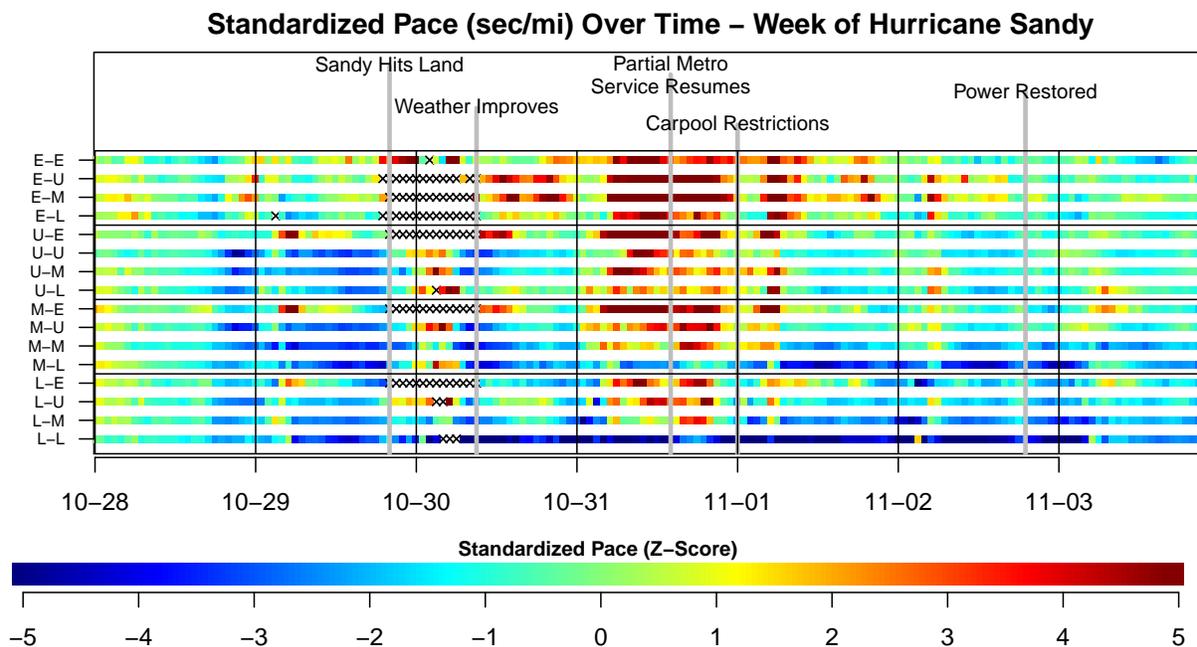


FIGURE 5 : The standardized pace vector during the week of Hurricane Sandy, 2012. Labels are included to show the times of specific phases of the event [24]. An average week would have values of zero everywhere, but significant deviations are shown during the week of Hurricane Sandy. Missing data (hours where there are less than five occurrences of a given trip) are marked with black Xs.

element, producing the *standardized pace vector*. The i th element of this vector is given by

$$\mathbf{S}(t)_i = \frac{\mathbf{A}(t)_i - \mu(t)_i}{\sqrt{\Sigma(t)_{i,i}}}. \quad (11)$$

Intuitively, the standardized pace vector tells how many standard deviations away from the mean the pace of each category of trips is at time t . In other words, it is possible to identify the trips that are going slower or faster than expected, and how significant this difference is. Figure 5 shows the standardized pace vector during the week of Hurricane Sandy. This figure gives some intuition on the behavior of various regions of the city during and after the hurricane. It also includes labels indicating the occurrences of various phases of the event, obtained from a post-Sandy study from NYU [24]. The most notable finding is that the slowest traffic occurred on Wednesday October 31st, almost two days after the hurricane struck land. On this day, some airports, buses, and commuter rails attempted to resume normal service, but much of the infrastructure was still damaged [24].

However, it is even more surprising that Midtown-to-Lower Manhattan and Lower Manhattan-to-Lower Manhattan travel times are significantly *lower* than expected during this time. The pace of these trips remains almost five standard deviations below the mean until Saturday the 3rd, despite the severe flooding and power outages in Lower Manhattan. The fact that a hurricane can actually make traffic move faster in some areas of the city indicates that the usage of the infrastructure changed. It is likely that the hurricane decreased demand on the transportation network in Lower Manhattan until the infrastructure recovered.

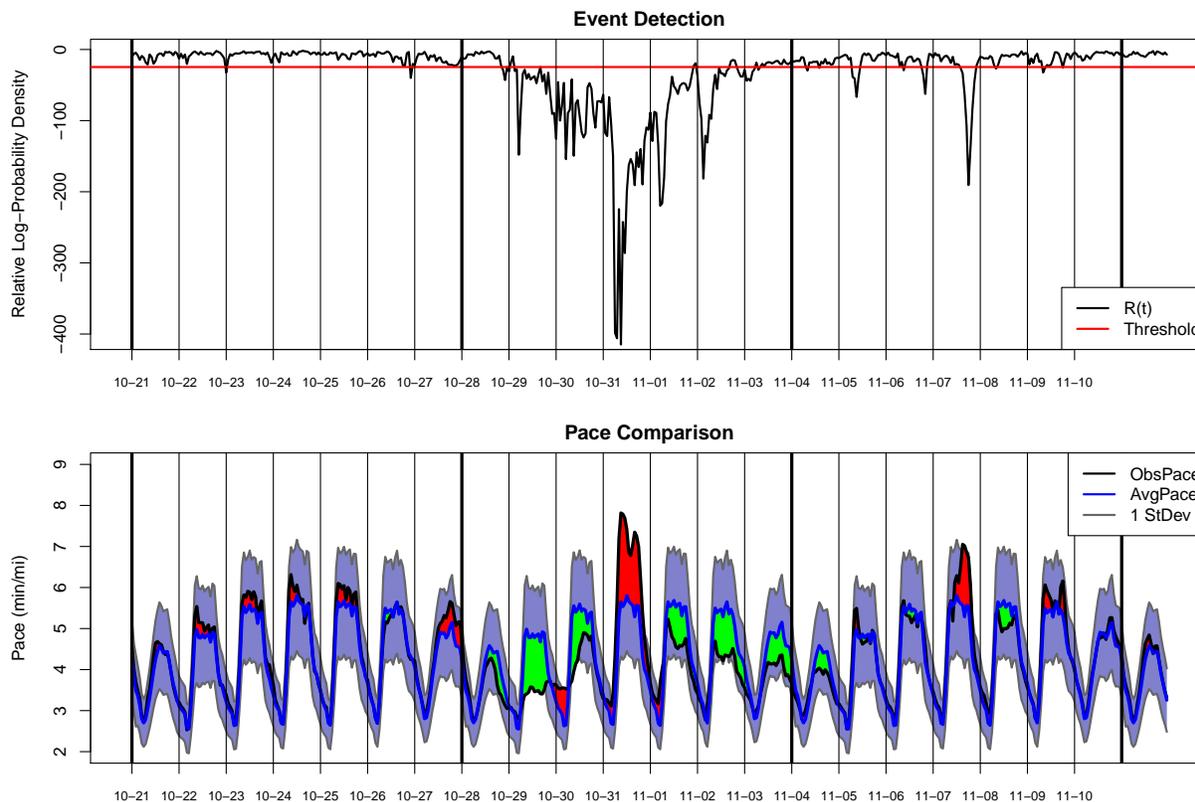


FIGURE 6 : Probabilistic detection and measurement of the event Hurricane Sandy. The relative log-probability density, $R(t)$, is plotted in the top figure and events are detected when it goes below the threshold. For comparison, the average pace of all taxis in the city is plotted below and compared to the expected value. Green areas indicate that travel times are low, but red indicates that they are unusually high.

This standardized pace vector gives a meaningful interpretation of unusual travel times between various regions of the city. However, it fails to account for correlations between these typical travel times i.e., the off-diagonal elements of $\Sigma(t)$. However, the full covariance matrix is considered by the relative likelihood function $R(t)$. As described in Section 2.4, events are detected when $R(t)$ goes below a threshold for a significant period of time. Figure 6 shows this process, along with the average pace of all taxis. Table 2 shows the top ten events, sorted by duration. At the top of the list is Hurricane Sandy, taking over five and a half days for travel times to return to normal. This is over three times the recovery time of Hurricane Irene. This agrees with the results of [20], which showed that the total number of Manhattan taxi trips returned to normal more quickly during Hurricane Irene than Hurricane Sandy. At its worst, Sandy added over two minutes to each mile driven by taxis in the city, while Irene added less forty seconds. This is in contrast to the results of [20], which showed that the peak drop in the number of taxi trips was greater during Hurricane Irene. The blizzard of December 2010, while shorter, added four minutes of travel time to each mile at its peak.

Event	Start Time	Duration (hours)	Max (min/mi)	Min (min/mi)	Worst Trip
Sandy	2012-10-28 21:00:00	132	2.25	-1.6	E → M
Blizzard	2010-12-26 13:00:00	112	4.41	0	M → M
Blizzard	2011-01-31 08:00:00	49	2.04	0	E → E
Irene	2011-08-27 13:00:00	43	0.64	-1.66	E → E
Unknown	2013-10-12 03:00:00	33	1.09	0	E → L
Blizzard	2013-02-08 06:00:00	26	1.54	-0.58	E → E
Blizzard	2010-02-10 06:00:00	24	0.67	-1.01	E → E
New Years	2012-12-31 15:00:00	20	1.42	-2.66	E → M
Memorial	2011-09-09 08:00:00	19	1.66	0	U → U
Blizzard	2011-01-28 02:00:00	18	2.57	0	L → L

TABLE 2 : Comparison of New York City transportation infrastructure resilience to the 10 longest events. The duration in hours, and the maximum/minimum pace deviation in minutes/mile is given for each event. Note that a positive number indicates a delay while a negative indicates a decreased pace. The final column indicates which of the 16 trips most frequently had the highest standardized pace during the event. Labels for events (the first column) are determined manually (cf. [25]).

CONCLUSION

This analysis has shown that it is possible to detect and measure the effects of unusual events on transportation infrastructure using only taxi GPS data. This is a first step toward assessing and improving city-scale resilience. Of key importance, the method is extremely low cost, because it does not require the installation of any additional sensors or purchase of private GPS datasets. This method proposes computing *origin-destination paces*, or average travel time per mile between various regions of the city. The effects of various events are quantified by the sizes and durations of pace deviations from typical values. Importantly, this measurement is done in a probabilistic way, so significant events can be distinguished from random day-to-day fluctuations.

The given method is applied to a dataset from New York City, and Hurricane Sandy is analyzed in detail. The analysis shows this was the longest event in the four year dataset, and one of the most severe in terms of peak pace deviation. At its worst, Hurricane Sandy caused over two minutes of delay per mile, but actually resulted in *faster* traffic for most of its duration. Most interestingly, the spike in delay occurred two days after the hurricane struck, as many residents migrated back into the city. This re-entry process was extremely slow when compared to the evacuation process before the hurricane, suggesting that more traffic management might be necessary following an event. The analysis of an extreme event like Hurricane Sandy demonstrates the ability of the proposed method to capture and describe atypical city-scale properties of the transportation network.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant No. CNS-1308842.

References

- [1] D. Matherly and N. Langdon, "A guide to regional transportation planning for disasters, emergencies, and significant events," Tech. Rep. 777, National Cooperative Highway Research Program, 2014.
- [2] T. Aven, "On some recent definitions and analysis frameworks for risk, vulnerability, and resilience," *Risk Analysis*, vol. 31, no. 4, pp. 515–522, 2011.
- [3] Y. Y. Haimes, "On the definition of resilience in systems," *Risk Analysis*, vol. 29, no. 4, pp. 498–501, 2009.
- [4] Y. Y. Haimes, "On the complex definition of risk: A systems-based approach," *Risk Analysis*, vol. 29, no. 12, pp. 1647–1654, 2009.
- [5] Y. Y. Haimes, "Responses to Terje Aven's paper: On some recent definitions and analysis frameworks for risk, vulnerability, and resilience," *Risk Analysis*, vol. 31, no. 5, pp. 689–692, 2011.
- [6] D. A. Reed, K. C. Kapur, and R. D. Christie, "Methodology for assessing the resilience of networked infrastructure," *IEEE Systems Journal*, vol. 3, no. 2, pp. 174–180, 2009.
- [7] S. Kaplan and B. J. Garrick, "On the quantitative definition of risk," *Risk Analysis*, vol. 1, no. 1, pp. 11–27, 1981.
- [8] M. Ouyang, L. Dueñas-Osorio, and X. Min, "A three-stage resilience analysis framework for urban infrastructure systems," *Structural Safety*, vol. 36–37, pp. 23–31, 2012.
- [9] M. Omer, A. Mostashari, and R. Nilchiani, "Assessing resilience in a regional road-based transportation network," *International Journal of Industrial and Systems Engineering*, vol. 13, no. 4, pp. 389–408, 2013.
- [10] S. E. Chang and N. Nojima, "Measuring post-disaster transportation system performance: the 1995 Kobe earthquake in comparative perspective," *Transportation Research Part A: Policy and Practice*, vol. 35, no. 6, pp. 475–494, 2001.
- [11] W. B. Allen, D. Liu, and S. Singer, "Accessibility measures of US metropolitan areas," *Transportation Research Part B: Methodological*, vol. 27, no. 6, pp. 439–449, 1993.
- [12] X. He and H. X. Liu, "Modeling the day-to-day traffic evolution process after an unexpected network disruption," *Transportation Research Part B: Methodological*, vol. 46, no. 1, pp. 50–71, 2012.
- [13] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 759–770, 2008.
- [14] F. Calabrese, F. C. Pereira, G. D. Lorenzo, L. Liu, and C. Ratti, "The geography of taste: analyzing cell-phone mobility and social events," in *Pervasive Computing*, pp. 22–37, Springer, 2010.

- [15] F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, and C. Ratti, “Real-time urban monitoring using cell phones: A case study in Rome,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 141–151, 2011.
- [16] P. S. Castro, D. Zhang, and S. Li, “Urban traffic modelling and prediction using large scale taxi GPS traces,” in *Pervasive Computing*, pp. 57–72, Springer, 2012.
- [17] Y. Zheng, Y. Liu, J. Yuan, and X. Xie, “Urban computing with taxicabs,” in *Proceedings of the 13th International Conference on Ubiquitous Computing*, pp. 89–98, ACM, 2011.
- [18] G. Qi, X. Li, S. Li, G. Pan, Z. Wang, and D. Zhang, “Measuring social functions of city regions from large-scale taxi behaviors,” in *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011*, pp. 384–388, 2011.
- [19] C. Chen, D. Zhang, P. Samuel Castro, N. Li, L. Sun, and S. Li, “Real-time detection of anomalous taxi trajectories from gps traces,” in *Mobile and Ubiquitous Systems: Computing, Networking, and Services* (A. Puiatti and T. Gu, eds.), vol. 104 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 63–74, Springer Berlin Heidelberg, 2012.
- [20] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva, “Visual exploration of big spatio-temporal urban data: A study of New York City taxi trips,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2149–2158, 2013.
- [21] B. Donovan, “Published code.” <https://github.com/UIUC-Transportation-Data/gpsresilience>.
- [22] B. Donovan and D. Work, “New York City taxi data 2010–2013.” <https://uofi.box.com/s/zmggziub40wx1bq2h9bq>.
- [23] D. A. Sprott, *Statistical Inference in Science*. Springer, 2000.
- [24] S. Kaufman, C. Qing, N. Levenson, and M. Hanson, “Transportation during and after Hurricane Sandy,” tech. rep., Rudin Center for Transportation, NYU Wagner Graduate School of Public Service, 2012.
- [25] “Lmgtfy.” <http://lmgtfy.com/?q=new+york+city+october+28+2012>.