Weather's Impact on Travel Time and Travel Time Variability in New York City

3 authors:

M. Anil Yazici
Stony Brook University

Camille N. Kamga
City College of New York

Abhishek Singhal
City College of New York

Weather's Impact on Travel Time and Travel Time Variability in New York City

M. Anil Yazici, Ph.D. (Corresponding Author)
Research Associate
Region-2 University Transportation Research Center
City College of New York
160 Convent Avenue, Marshak Building, Suite MR-910
New York, NY, 10031
Tel: (212) 650-80-71
Fax: (212) 650-8374
E-mail: yazici@utrc2.org

Camille Kamga, Ph.D.
Assistant Professor
Department of Civil Engineering, City College of New York
160 Convent Avenue, Marshak Building, Suite MR-917
New York, NY, 10031
Tel: (212) 650-80-87
Fax: (212) 650-8374
E-mail: ckamga@utrc2.org

Abhishek Singhal
Graduate Assistant
Region-2 University Transportation Research Center
City College of New York
160 Convent Avenue, Marshak Building, Suite MR-910
New York, NY, 10031
Tel: (212) 650-80-71
Fax: (212) 650-8374
E-mail: singhal@utrc2.org

Word count: 5213 Words + 3 Figures + 4 Tables = 6963
Abstract: 176

Transportation Research Board's 92nd Annual Meeting, Washington, D.C., 2013

ABSTRACT

In this study, the impact of weather conditions on travel time and travel time variability in New York City is investigated using Classification and Regression Trees (C&RT). For this purpose, taxi GPS data provided by the New York City Taxi and Limousine Commission (TLC) with more than 370 million records is merged with historical weather data. For all day-of-week (DOW), time-of-day (TOD) and weather condition categories, the impact of weather on the mean and mode of travel time distributions and on the coefficient of variation as a measure for variability are analyzed. It is found that the level of travel time variability changes across DOW-TOD-weather categories and that weather has a higher impact on travel time and variability during less congested periods. The literature has shown that inclement weather slows traffic: a major finding of this study is that it also reduces traffic variability, a finding that would seem counter-intuitive. Using a rich dataset and the appropriate analytical methods, the present study contributes valuable insights to the understanding of travel time variability in an urban context.

INTRODUCTION

Travel time between two arbitrary points in a transportation network can take different values based on the time of day (e.g., peak *vs.* off-peak), day of the week (e.g., weekday *vs.* weekend), as well as for two vehicles traveling the same route at the same time but at different speeds. The literature refers to these variations as *inter-period*, *inter-day* and *inter-vehicle* travel time variability, respectively (1,2). Although *travel time variability* and *travel time reliability* are generally used interchangeably, the concept of reliability refers to a failure of the system under consideration. For road networks, such failure is difficult to define. However, travel time variability can be observed for a given period of travel time data and analyzed to assess a particular transportation system's reliability. In other words, travel time variability can serve as input for the analysis of travel time reliability, without reference to the "system failure" aspect of many engineering notions of reliability. A technical report for the Federal Highway Association (3) identifies travel time variability as a crucial component of travel time reliability and approaches reliability from the perspective of traffic congestion and delay. The report defines travel time reliability as a formal way of describing non-recurrent delays and identifies seven main factors contributing to congestion and travel time variability: 1) Capacity, 2) Traffic Incidents, 3) Work Zones 4) Weather, 5) Traffic Control Devices, 6) Special Events, and 7) Fluctuations in Normal Traffic.

This paper selects *weather* from the above list and analyzes its impact on travel time variability by day-of-week (DOW) and time-of-day (TOD) considerations. The paper uses New York City (NYC) taxi GPS data provided by the NYC Taxi and Limousine Commission (TLC) and weather data extracted from www.wunderground.com to calculate travel times for different time periods and weather conditions. *Mean travel time* is typically the standard measure used to investigate overall change in travel times across different time and weather conditions. The *coefficient of variation* measures the dispersion level and variability of the data, which is also treated here as a reliability measure. *The statistical mode of travel time*, in contrast, is not commonly used to measure mean or median travel times. Nevertheless, as discussed in (4), travel decisions are typically made based on momentary recollections of past experience, and the mode of the travel time distribution can provide a more important inference than the mean travel time. In this paper, the three aforementioned distributional characteristics of travel times--*inter-period*, *inter-day* and *inter-vehicle* variability--are analyzed within different DOW, TOD and weather conditions by using classification and regression trees (C&RT).

The outline of the paper is as follows: First, the paper presents a review of the literature on travel time variability. Then it provides information about the NYC taxi GPS and weather data and gives a brief summary of the classification and regression tree (C&RT) methodology. Finally, it discusses the findings in detail and presents conclusions.

LITERATURE REVIEW

Travel time variability is a concept that is used almost interchangeably with travel time reliability (5,6). Travel time variability can be defined as the random variation in travel time (7), or as the uncertainty in journey trip times (8), without requiring a system failure. Researchers employ probe vehicles, Bluetooth, GPS and similar technology to analyze the aforementioned randomness in travel time (2, 7,9,10, 11,12, 13). Studies agree that both average travel times and travel time variability vary according to the time-of-day, day-of-week, and weather conditions. The impact of weather on traffic has been accepted as a fact in transportation engineering and several studies focus specifically on weather impacts on travel time variability (11,13). Chien and Kolluri (11) study travel time variability on New Jersey freeways, while Tu et. al (13) have done the same type of analysis for freeways in the Netherlands. Both studies measure the variability with standard deviation and conclude that adverse weather increases travel time variability. Chien and Kolluri (11) point out that the impact of adverse weather is more significant during peak periods than that during off-peak periods. In contrast, Tu et. al (13) conclude that adverse weather leads to higher travel time variability at lower demand levels. If lower demand levels can be considered as off-peak conditions, the findings of these two studies sound contradictory. Kwon et. al. (14) extend The Federal Highway Administration's (FHWA) "congestion pie" idea to travel time reliability and introduce the "unreliability pie." "Congestion pie" includes percentage contribution of selected factors to congestion, and weather is one of the factors selected for both the congestion and unreliability pies. Kwon et. al. (14) find that weather has relatively little impact on travel time unreliability, and weather is not significant for the "noon" period. In other words, the results of Kwon et. al. agree with those of Chien and Kolluri (11). Peer et. al (12) investigates the relationship between mean duration and variability (again, represented by standard deviation). They find that unit increase in mean travel time is associated with a larger increase in variability during free-flow conditions than when there is hyper-congested traffic. Thus, Peer et. al (12) also recognize the different characteristics of variability during peak and off-peak periods. Peer et. al (12) question some of their results' divergence from previous studies and conclude that travel time variability on freeways and urban roads have different characteristics. They claim that in urban networks travel time variability is already high due to factors such as traffic lights, and an increase in mean delay may not be fully reflected on the travel time variability. Franklin and Kalstrom (15) also confirm the difference between freeway and urban travel time characteristics based on data from Stockholm. A paper by the authors of the present study (9) looks at the impact of DOW-TOD on travel time reliability in NYC and also reaches a similar conclusion: travel time reliability in urban networks is subject to different parameters compared to freeways. These findings align with previous studies that found that variability is lower during peak periods and higher for off-peak periods.

As presented above, the literature includes studies that investigate the impact of weather on travel time variability for freeways, but the authors are not aware of a study that examines an urban setting. For this purpose, the current study merges an extensive NYC taxi trip data set with historical weather information to study the impact of weather on travel time and on travel time variability in NYC's urban transportation network.

DATA

This paper uses the taxi travel times available through the New York City taxi GPS data obtained from the TLC and assumes that these reflect travel time patterns in the city. The dataset has more than 370 million taxi trips covering the period from January 1, 2009 to November 28, 2010.The vast majority of trips originate from and end in NYC (the main origin and destination points being in Manhattan). An additional small portion of the dataset includes other New York destinations (e.g., Long Island) and New Jersey. Overall, the data include trips for all possible time periods and traffic conditions, as well as almost all regions in and around the city. For each taxi trip, travel time is calculated using pick-up/drop-off date and time. As discussed in (6) while studying travel time variability/reliability, one has to use either a representative trip length, or to conduct analysis in length-neutral terms. Lomax et. al. (6) suggest using "travel rate" (in minutes per mile) as a length-neutral surrogate for trip time variation, and the current paper adopts this length-neutral measure by dividing each trip travel time by the trip length. This conversion makes it possible to perform a system-wide analysis of travel times rather than using selected origin-destination pairs or a representative trip length.

Data on weather conditions is gathered from the www.wunderground.com website, and includes various weather-related information such as temperature, wind speed and its direction, precipitation, etc. In the current paper only the "weather condition" field is used. The Wunderground website classifies weather conditions based on precipitation, cloud structure, visibility, and so on. It also categorizes rain and snow conditions based on precipitation levels. For the purposes of this study, main considerations are with the existence, type, and level of precipitation (e.g. clear, light rain, rain, heavy rain, light snow, snow, and heavy snow). Conditions such as "overcast" or "partly cloudy" do not include precipitation and they are assumed to be no different from "clear" conditions as far as traffic conditions are concerned. "Fog" conditions, on the other hand, may have an impact on traffic flow for freeways, but it was assumed that fog would have a negligible effect for travel within New York City. Hence, the following re-categorization for weather conditions is performed:

Existing Categories→Assigned Categories

Existing Categories	Assigned Categories
Clear	Clear
Partly Cloudy	Clear
Scattered Clouds	Clear
Mostly Cloudy	Clear
Overcast	Clear
Fog	Clear
Haze	Clear
Light Rain	Light Rain
Light Freezing Rain	Light Rain

Rain→Rain
Heavy Rain→Heavy Rain
Light Snow→Light Snow
Snow→Snow
Heavy Snow→Heavy Snow
Unknown→N/A

Wunderground updates weather conditions at least every hour and more frequently if there is any change to report. When combining taxi trip and weather dataset, it was assumed that the weather conditions at the beginning time of each trip would be valid for the whole trip. Considering that around 75% of taxi trips are completed within 15 minutes, the probability of weather condition changing over the course of a single trip is very low.

Erroneous and Missing Records

In order to use the taxi trip data, records with possible recording errors (e.g. trip records with zero travel time or zero distance) were deleted. There were also records with physically impossible travel times (e.g., average speed in multiples of the free flow speed or of the speed limit) or unreasonably slow travel speeds (e.g., vehicle speeds that were fractions of the average walking speed). As a general rule, records with an average speed higher than 1 minute per mile or 60 mph (which is a very high speed even for emergency vehicles in NYC traffic) and those with a speed lower than 40 minutes per mile (or 1.5 mph, i.e., half of the average human walking speed of 3 mph) are deleted from the dataset. Similarly, trip records corresponding to “unknown” or “N/A” weather conditions were also excluded.

CLASSIFICATION AND REGRESSION TREES

Classification and Regression Trees (C&RT) is the name of a non-parametric model proposed by Breiman, Friedman, Olshen, & Stone (16). C&RT is used in variety of fields spanning statistics, data mining, artificial intelligence, and machine. C&RT takesX_i’s as categorical or continuous input to predict the target variable, Y, which can also be either continuous or categorical. The calculated tree is a *classification* tree when Y is a categorical variable, and a *regression* tree when Y is continuous. Based on an error tolerance level, the algorithm partitions the target variable into more homogenous clusters by constructing a binary tree that narrows the decision space at each node. The calculated tree offers an easy to interpret tree in which one can start from the top node, follow the branches based on the split criteria and reach down to the terminal node that gives the regression (or classification) value. The C&RT algorithm expands the tree to the level where it fully represents the target variable. As a result the constructed tree may have a large number of nodes (=high tree complexity). The tree may also include all the decision variables in the final tree but a fully developed complex tree with all input variables does not necessarily yield good predictions for another dataset, and it also complicates the interpretation. The C&RT output does not provide significance levels on decision variables due to its non-parametric nature. Model validation and decision variable selection are performed via cross validations for different levels of tree complexities. To determine the “optimal” tree size, based on the trade-off between the tree costs after *cross validation* and *re-substitution*. Re-substitution cost refers to the tree “pruning”. Pruning starts from the terminal nodes in the C&RT tree and child nodes are pruned one by one. More pruning results in simpler tree, however, the pruned tree introduces a cost associated with misclassifying the target variable. On the other hand, the cross validation error initially decreases as the tree complexity increases, however starts increasing after certain degrees or levels of tree complexity. The optimal tree can be selected on the minimum cost tree or the tree with the smallest number of nodes which is around one standard deviation away from the minimum cost tree. In this study, MATLAB statistics toolbox C&RT functions are used. Following MATLAB’s default settings (17), ten-fold cross validation is performed to calculate the minimum tree.

THE ANALYSIS AND FINDINGS

The first step of the analysis involves extracting the travel time distributions of the trips that fall into certain DOW-TOD-weather categories, as well as calculating the mean, mode and coefficient of variation (CoV) for each distribution. Then the C&RT are formed using each DOW-TOD-Weather category as the categorical input X_i with the target variable Y representing mean, mode, or CoV of the corresponding travel-time distribution. The resulting trees are shown in FIGURE 1, FIGURE 2 and FIGURE 3 for the mean, mode and CoV respectively. A general discussion of the tree results follows the figures. Then the impact of weather on travel time and travel time variability is discussed by comparing the C&RT results for selected TOD periods across all days and all weather conditions. Please note that rare weather conditions, such as heavy snow, tend to occur only a few times in a season, and these rare conditions occur for only a few different TODs. For this reason, the input for C&RT includes some DOW-TOD-weather categories that do not have any trip records. C&RT produces estimates for these rare categories even where there are no recorded observations.

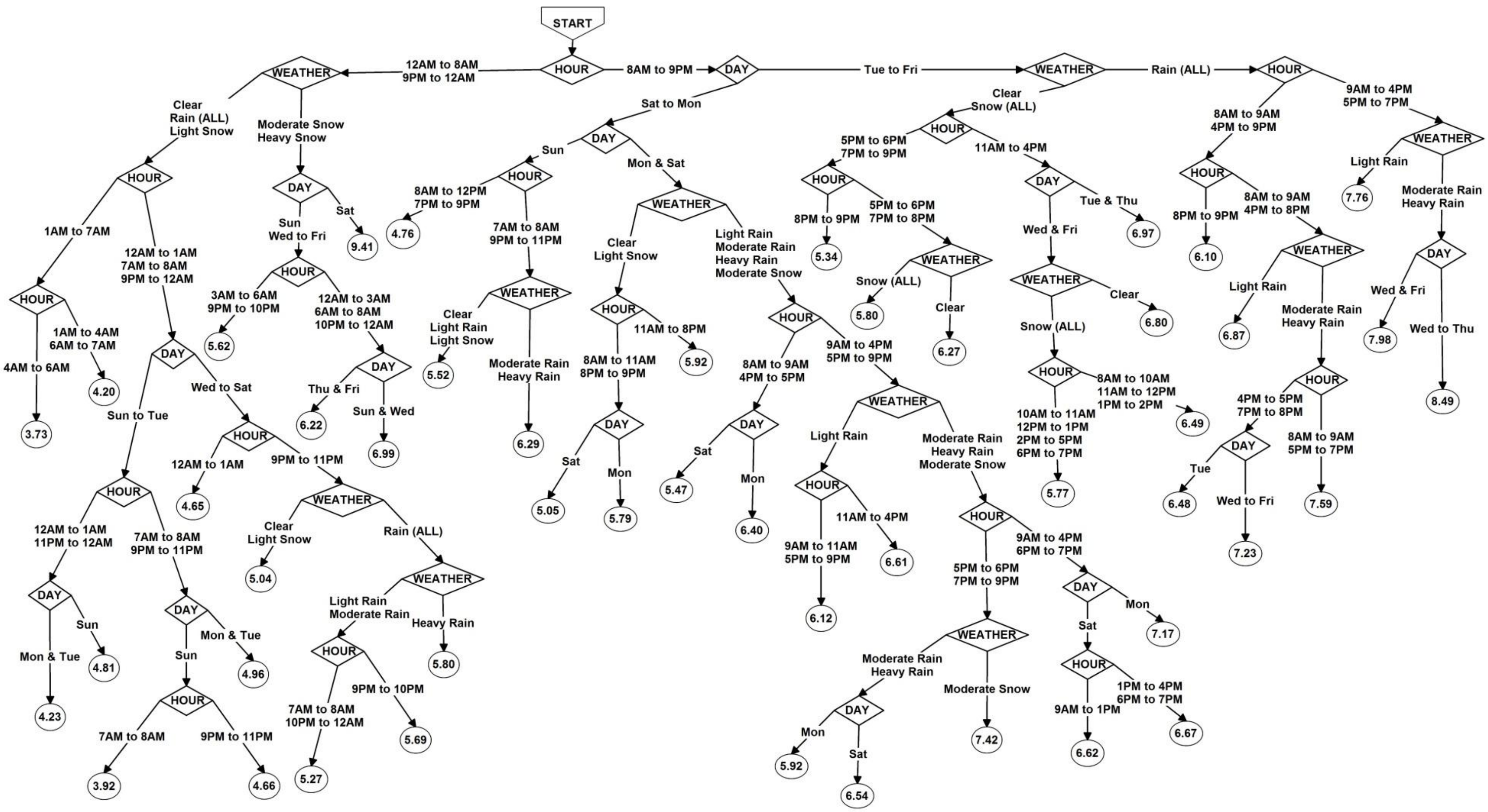

FIGURE 1C&RT for the Mean Travel Time

FIGURE 1 shows that the fastest travel time period (3.73 minutes per mile) is 4AM to 6AM for all days and in all weather except for conditions of moderate and heavy snow. The slowest period (9.41 minutes per mile) is the Friday midnight to Saturday 8AM and 9PM to midnight on Saturday under moderate or heavy snow. The first split in C&RT is by the hour. The 8AM-9PM period can be generally considered to be high activity hours (travel to and from work and after-work hours) and the remaining are low activity hours. This contrast between day and night hours constitutes the first split in the mean travel time tree. There is no clear cut between weekday and weekend days. Meanwhile, there is also no consistent pattern on weekend days. The mean travel times for TOD periods on Sunday and Saturday are distinct from each other. There are also TOD-weather periods that have the same mean travel time regardless of DOW. For instance, during clear, rain and light snow conditions alike, mean travel time between 1AM to 7AM does not depend on DOW.

The mode travel time tree shown in FIGURE 2 is relatively less complex compared to the mean travel time tree. The first split is exactly the same, and the subsequent split nodes are similar as well. The lowest mode travel time (3.09 minutes per mile) occurs during the 3AM-6AM period for all days under clear and all types of rain and light snow conditions. The highest mode travel time is 8.23 minutes per mile and occurs at the same period of maximum mean travel time. Since the travel-time distributions are positively skewed, the travel time mode is smaller than the mean travel time for all periods. Both the mode and mean travel time trees provide output that is realistic in terms of the expected DOW-TOD congestion patterns. In general (with few exceptions), both regression trees yields higher travel times as the weather severity increases, e.g., clear → light rain → moderate rain → … →heavy snow.

While calculating the trees, the target variable is partitioned based on an error tolerance. Since more than one significant digit for a travel time value refers to a precision of seconds, error tolerance for mode and mean travel times need not be very small. Optimal CoV tree has less complexity relative to both the mean and mode travel time trees, although the error tolerance assigned for CoV tree is smaller (10^{-6}, MATLAB's default value) than the error tolerance assigned for the mean and mode travel time trees (10^{-1}). The first split in the CoV tree is similar to the mode and mean travel time trees, dividing the day into midnight to 7AM and 7AM to the end of the day. The lowest variation is estimated for Wednesday and Saturday between 12AM and 7AM under moderate and heavy snow. The highest variation is between 4AM and 6AM on Monday, Tuesday and Thursdays under clear and all types of rain and light snow conditions.

Overall, the split nodes of the trees give some idea about the DOW-TOD-weather patterns for the travel time measures at hand, and they can illustrate the modeling capabilities of the C&RT approach. In general, the trees do not diverge significantly from the expected travel times in New York City, that is, it would be anticipated that there would be low travel times during the early morning and late evening hours, and higher travel times at business hours. However, the analysis of the values at the terminal nodes provides the desired results to provide insights on the impact of weather conditions on travel time variability. This analysis is presented in the following section.

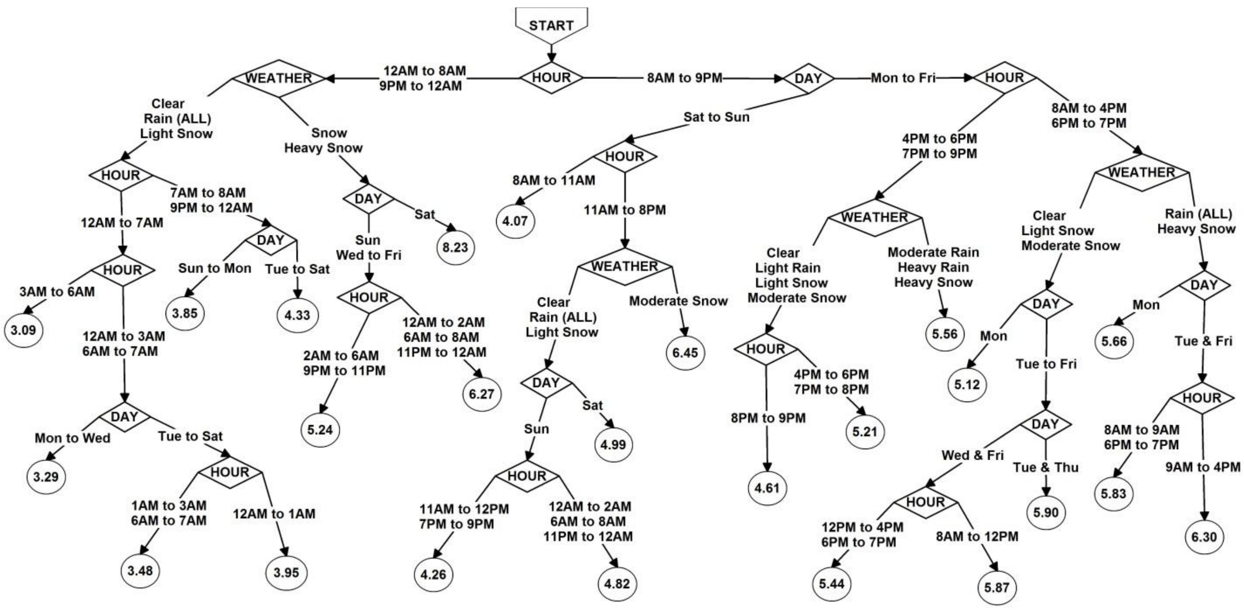

FIGURE 2C&RT for the Mode Travel Time

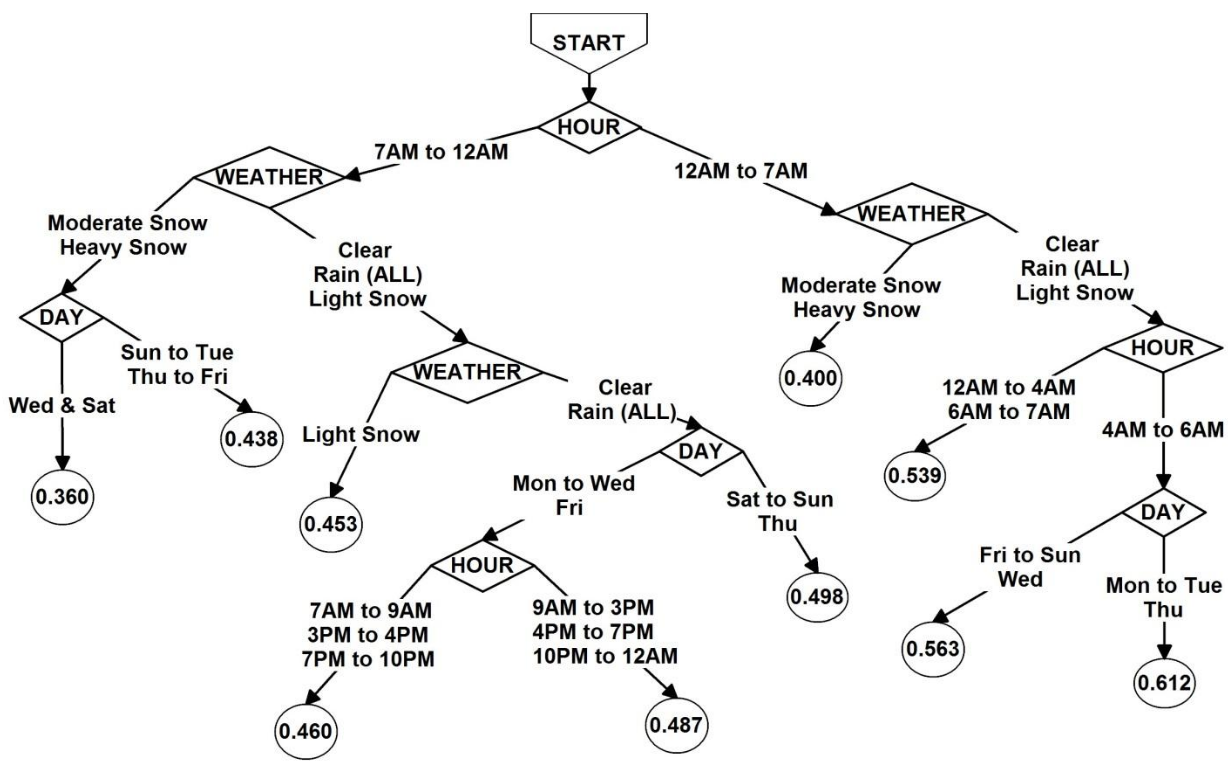

FIGURE 3C&RT for the Coefficient of Variation (CoV)

Weather Impacts on Travel Time and Variability

To investigate and quantify the impact of weather on travel time, calculated C&RT results are presented for selected TOD periods across all days and weather conditions. A total of six one-hour time periods are examined. These six periods represent different TOD cases:

12AM – 1AM: Late night
4AM – 5AM: Early morning
8AM – 9AM: Morning peak
12PM – 1PM: Midday
4PM – 5PM: Evening peak
8PM – 9PM: Evening/night

For each DOW-TOD-weather category, the calculated measure is compared with a clear-weather value to quantify the impact of weather. The weekday results are presented in TABLE 1, TABLE 2, and TABLE 3 for the travel-time mean, mode and CoV, respectively. TABLE 4 shows the weekend results for all three measures. In all tables, the weather categories are represented by:

1 – Clear
2 – Light Rain
3 – Moderate Rain
4 – Heavy Rain
5 – Light Snow
6 – Moderate Snow
7 – Heavy Snow

Due to space considerations, weather is abbreviated as "W" and the percentage of change with respect to clear weather conditions are represented by "%" where necessary.

As shown in TABLE 1, the impact of weather on the mean travel time is generally higher for late night and early morning periods and for off-peak periods during weekdays. This observation is also valid for the mode for travel time on weekdays (TABLE 2). For example, the mean travel time changes are at most 12% and 8% for the morning and evening peaks, respectively. In contrast, the highest increase for both the mean and mode for travel times are calculated for the 12AM-1AM and 4AM-5AM periods, which goes up to 59% and 80%, respectively. A similar observation is also valid for the travel-time mode on weekdays (TABLE 2). In addition, the 12AM-1AM and 4AM-5AM periods do not exhibit any change until conditions become extreme. For these periods there is no change in mean travel time during clear, all levels of rain and light snow conditions, whereas more severe conditions such as moderate and heavy snow result in significant travel-time effects. For other periods, travel time differences can be observed even between increased severity levels of the same weather condition, e.g., light and moderate rain. There are also a few cases where the mean travel time decreases, especially for some snow conditions. For example, all snow conditions cause a 15%

decrease in mean travel times on Wednesdays and Fridays for the 12PM-1PM and 4PM-5PM periods. An interesting finding is that the travel-time mode does not show any decrease for the same periods and weather conditions. Instead, the calculations show increases in mode travel time. (Please see TABLE 2). Decreases in the mode are observed as well but only at a magnitude of 1% and for a few cases. For weekends (TABLE 4), weather's impact on mean travel times show distinct patterns. For example, weather is found to have no impact on Sunday for the 8AM-9AM and 8PM-9PM periods, but the same is not true for Saturday. Weather severity has more of an impact on Saturday. In contrast, the impact of weather on mode travel time for Saturday and Sunday are more consistent. In general, it can be concluded that only moderate and heavy snow affect the travel time mode.

Overall, the presented findings show that the effects of weather on travel time vary depending on TOD and DOW. During periods with higher congestion, i.e. morning and evening traffic peaks, the impact of weather is less than during night and off-peak periods. This can be attributed to the fact that high levels of congestion builds up during peak-hours, and traffic flows in a congested but steady manner. The inclement weather conditions do not make the already crawling traffic flow worse, while, during off-peak hours, the traffic flow is not congested and there is more room for external effects.

TABLE 1Weekday Mean Travel Time C&RT Results for Different Weather Conditions

TOD Period	Weather	Monday		W	Tuesday		Wednesday		Thursday		Friday	
		Mins per mile	% Change		Mins per mile	%	Mins per mile	%	Mins per mile	%	Mins Per mile	%
12AM - 1AM	1 – 5	4.23	0%	1 – 5	4.23	0%	4.65	0%	4.65	0%	4.65	0%
	6 – 7	6.72	59%	6 – 7	6.72	59%	6.99	50%	6.22	34%	6.22	34%
4AM - 5AM	1 – 5	3.73	0%	1 – 5	3.73	0%	3.73	0%	3.73	0%	3.73	0%
	6 – 7	6.72	80%	6 – 7	6.72	80%	5.62	51%	5.62	51%	5.62	51%
8AM - 9AM	1	5.79	0%	1	6.97	0%	6.80	0%	6.97	0%	6.80	0%
	2 – 4	6.40	10%	2	6.87	-1%	6.87	1%	6.87	-1%	6.87	1%
	5	5.79	0%	3 – 4	7.59	9%	7.59	12%	7.59	9%	7.59	12%
	6	6.40	10%	5 – 7	6.97	0%	6.49	-4%	6.97	0%	6.49	-4%
	7	6.22	7%									
12PM - 1PM	1	5.92	0%	1	6.97	0%	6.80	0%	6.97	0%	6.80	0%
	2	6.61	12%	2	7.76	11%	7.76	14%	7.76	11%	7.76	14%
	3 – 4	7.17	21%	3 – 4	7.98	15%	8.49	25%	8.49	22%	7.98	17%
	5	5.92	0%	5 – 7	6.97	0%	5.77	-15%	6.97	0%	5.77	-15%
	6	7.17	21%									
	7	6.22	5%									
4PM - 5PM	1	5.92	0%	1	6.97	0%	6.80	0%	6.97	0%	6.80	0%
	2 – 4	6.40	8%	2	6.87	-1%	6.87	1%	6.87	-1%	6.87	1%
	5	5.92	0%	3 – 4	6.48	-7%	7.23	6%	7.23	4%	7.23	6%

	6	6.40	8%	5 – 7	6.97	0%	5.77	-15%	6.97	0%	5.77	-15%
	7	6.22	5%									
8PM - 9PM	1	5.79	0%	1	5.34	0%	5.34	0%	5.34	0%	5.34	0%
	2	6.12	6%	2 – 4	6.10	14%	6.10	14%	6.10	14%	6.10	14%
	3 – 4	5.92	2%	5 – 7	5.34	0%	5.34	0%	5.34	0%	5.34	0%
	5	5.79	0%									
	6	7.42	28%									
	7	6.22	7%									

TABLE 2Weekday Mode Travel Time C&RT Results for Different Weather Conditions

Period	Weather	Monday		Tuesday		Wednesday		Thursday		Friday	
		Mins per mile	% Change	Mins per mile	% Change	Mins per mile	% Change	Mins per mile	% Change	Mins per mile	% Change
12AM - 1AM	1 – 5	3.29	0%	3.29	0%	3.29	0%	3.95	0%	3.95	0%
	6 – 7	6.13	86%	6.13	86%	6.27	91%	6.27	59%	6.27	59%
4AM - 5AM	1 – 5	3.09	0%	3.09	0%	3.09	0%	3.09	0%	3.09	0%
	6 – 7	6.13	99%	6.13	99%	5.24	70%	5.24	70%	5.24	70%
8AM - 9AM	1	5.12	0%	5.91	0%	5.87	0%	5.91	0%	5.87	0%
	2 – 4	5.66	11%	5.83	-1%	5.83	-1%	5.83	-1%	5.83	-1%
	5 – 6	5.12	0%	5.91	0%	5.87	0%	5.91	0%	5.87	0%
	7	5.66	11%	5.83	-1%	5.83	-1%	5.83	-1%	5.83	-1%
12PM - 1PM	1	5.12	0%	5.91	0%	5.44	0%	5.91	0%	5.44	0%
	2 – 4	5.66	11%	6.30	7%	6.30	16%	6.30	7%	6.30	16%
	5 – 6	5.12	0%	5.91	0%	5.44	0%	5.91	0%	5.44	0%
	7	5.66	11%	6.30	7%	6.30	16%	6.30	7%	6.30	16%
4PM - 5PM	1 – 2	5.21	0%	5.21	0%	5.21	0%	5.21	0%	5.21	0%
	3 – 4	5.56	7%	5.56	7%	5.56	7%	5.56	7%	5.56	7%
	5 – 6	5.21	0%	5.21	0%	5.21	0%	5.21	0%	5.21	0%
	7	5.56	7%	5.56	7%	5.56	7%	5.56	7%	5.56	7%
8PM - 9PM	1 – 2	4.61	0%	4.61	0%	4.61	0%	4.61	0%	4.61	0%
	3 – 4	5.56	20%	5.56	20%	5.56	20%	5.56	20%	5.56	20%
	5 – 6	4.61	0%	4.61	0%	4.61	0%	4.61	0%	4.61	0%
	7	5.56	20%	5.56	20%	5.56	20%	5.56	20%	5.56	20%

Thus so far, the findings regarding the impact of weather on travel time have made no reference to *variability*. The coefficient of variation provides interesting insights on this issue. As shown in TABLE 3 and TABLE 4, CoV stays the same or decreases as weather conditions change, for both weekdays and weekends. Rain conditions do not affect the CoV and it decreases only for snow conditions. In other words, rain generally increases the travel time but does not affect variability, whereas snow also increases the travel time but it results in *less* variability in travel time distribution. Besides, during weekdays, the weather-related decrease in CoV is generally lower during business hours. As for weather impacts on mean and mode travel times, there is both increased travel time and reduced variability due to inclement weather during off-peak hours when there is no congestion.

TABLE 3Weekday CoV C&RT Results for Different Weather Conditions

Period	Weather	Monday		Tuesday		Wednesday		Thursday		Friday	
		CoV	% Change	CoV	% Change	CoV	% Change	CoV	% Change	CoV	% Change
12AM - 1AM	1 – 5	0.539	0%	0.539	0%	0.539	0%	0.539	0%	0.539	0%
	6 – 7	0.400	-26%	0.400	-26%	0.400	-26%	0.400	-26%	0.400	-26%
4AM - 5AM	1 – 5	0.612	0%	0.612	0%	0.563	0%	0.612	0%	0.563	0%
	6 – 7	0.400	-35%	0.400	-35%	0.400	-29%	0.400	-35%	0.400	-29%
8AM - 9AM	1 – 4	0.460	0%	0.460	0%	0.460	0%	0.498	0%	0.460	0%
	5	0.453	-2%	0.453	-2%	0.453	-2%	0.453	-9%	0.453	-2%
	6 – 7	0.438	-5%	0.438	-5%	0.360	-22%	0.438	-12%	0.438	-5%
12PM - 1PM	1 – 4	0.487	0%	0.487	0%	0.487	0%	0.498	0%	0.487	0%
	5	0.453	-7%	0.453	-7%	0.453	-7%	0.453	-9%	0.453	-7%
	6 – 7	0.438	-10%	0.438	-10%	0.360	-26%	0.438	-12%	0.438	-10%
4PM - 5PM	1 – 4	0.487	0%	0.487	0%	0.487	0%	0.498	0%	0.487	0%
	5	0.453	-7%	0.453	-7%	0.453	-7%	0.453	-9%	0.453	-7%
	6 – 7	0.438	-10%	0.438	-10%	0.360	-26%	0.438	-12%	0.438	-10%
8PM - 9PM	1 – 4	0.460	0%	0.460	0%	0.460	0%	0.498	0%	0.460	0%
	5	0.453	-2%	0.453	-2%	0.453	-2%	0.453	-9%	0.453	-2%
	6 – 7	0.438	-5%	0.438	-5%	0.360	-22%	0.438	-12%	0.438	-5%

TABLE 4Weekend CoV, Mean and Mode Travel Time C&RT Results for Different Weather Conditions

	Mean Travel Time						Mode Travel Time					Coefficient of Variation				
Period	W	Saturday		W	Sunday		W	Saturday		Sunday		W	Saturday		Sunday	
		Mins/ mile	% Change		Mins/ mile	% Change		Mins/ mile	% Change	Mins/ mile	% Change		CoV	% Change	CoV	% Change
12AM - 1AM	1 – 5	4.65	0%	1 – 5	4.81	0%	1 – 5	3.95	0%	3.95	0%	1 – 5	0.539	0%	0.539	0%
	6 – 7	9.41	102%	6 – 7	6.99	45%	6 – 7	8.23	108%	6.27	59%	6 – 7	0.400	-26%	0.400	-26%
4AM - 5AM	1 – 5	3.73	0%	1 – 5	3.73	0%	1 – 5	3.09	0%	3.09	0%	1 – 5	0.563	0%	0.563	0%
	6 – 7	9.41	152%	6 – 7	5.62	51%	6 – 7	8.23	166%	5.24	70%	6 – 7	0.400	-29%	0.400	-29%
8AM - 9AM	1	5.05	0%	1 – 7	4.76	0%	1 – 7	4.07	0%	4.07	0%	1 – 4	0.498	0%	0.498	0%
	2 – 4	5.47	8%									5	0.453	-9%	0.453	-9%
	5	5.05	0%									6 – 7	0.360	-28%	0.438	-12%
	6	5.47	8%													
	7	6.22	23%													
12PM - 1PM	1	5.92	0%	1 – 2	5.52	0%	1 – 5	4.99	0%	4.82	0%	1 – 4	0.498	0%	0.498	0%
	2	6.61	12%	3 – 4	6.29	14%	6	6.45	29%	6.45	34%	5	0.453	-9%	0.453	-9%
	3 – 4	5.62	-5%	5	5.52	0%	7	4.91	-2%	4.91	2%	6 – 7	0.360	-28%	0.438	-12%
	5	5.92	0%	6 – 7	5.85	6%										
	6	5.62	-5%													
	7	6.22	5%													
4PM - 5PM	1	5.92	0%	1 – 2	5.52	0%	1 – 5	4.99	0%	4.82	0%	1 – 4	0.498	0%	0.498	0%
	2 – 4	5.47	-8%	3 – 4	6.29	14%	6	6.45	29%	6.45	34%	5	0.453	-9%	0.453	-9%
	5	5.92	0%	5	5.52	0%	7	4.91	-2%	4.91	2%	6 – 7	0.360	-28%	0.438	-12%
	6	5.47	-8%	6 – 7	5.85	6%										
	7	6.22	5%													
8PM - 9PM	1	5.05	0%	1 – 7	4.76	0%	1 – 5	4.99	0%	4.26	0%	1 – 4	0.498	0%	0.498	0%
	2	6.12	21%				6	6.45	29%	6.45	51%	5	0.453	-9%	0.453	-9%
	3 – 4	6.54	30%				7	4.91	-2%	4.91	15%	6 – 7	0.360	-28%	0.438	-12%
	5	5.05	0%													
	6	7.42	47%													
	7	6.22	23%													

CONCLUSIONS

In this study, the impact of weather conditions on travel time variability in NYC is investigated using Classification and Regression Trees (C&RT). For this purpose, taxi GPS data provided by the New York City Taxi and Limousine Commission (TLC) is merged with historical weather data gathered from www.wunderground.com website. The new data set is used to obtain the travel times in New York City for all day-of-week (DOW), time-of-day (TOD) and weather condition categories. Overall, a computationally extensive analysis of more than 370 million records is conducted and the results are presented.

It is found that travel time variability, just like travel time mean and mode, varies among DOW-TOD periods and is also affected by weather conditions. In this respect, the study results confirm the findings in the previous literature. However, the present analysis also reveals some intriguing and counter-intuitive facts that shed light on travel time variability in urban contexts. It is found that the impact of weather is not consistent among TOD periods. Inclement weather causes more increases in travel time mean and travel time mode for periods that are not congested, such as early mornings. The effect of weather conditions decreases as the network gets more congested. In terms of travel-time variability, it is shown that inclement weather conditions *reduce* variability. Moreover, the reduction in variability is less during peak hours and more during off-peak or less congested hours. On the one hand, heavy snow during peak hours *does* increase travel times, but the relative travel time increase would be higher if the same vehicles were caught under the same heavy snow conditions at off-peak hours. On the other hand, the same heavy snow causes less variation in travel times throughout the network, though it takes longer for drivers to reach their destination. This is mainly because during peak hours the network is highly congested, and the traffic flows in a slow but steady manner. This results in higher mean travel times but less variability as well. On top of the existing high level of congestion, the marginal impact of an external factor, such as weather, is limited. During night hours, the network is relatively empty, and drivers can choose their travel speed (and, therefore, travel times), and this causes travel time variations. Under inclement weather conditions, the drivers mostly slow down, increasing their travel times but at the same time reducing the speed differentials with other vehicles in the network, thus reducing travel time variability.

Although there are plausible explanations for these findings, they partially contradict the existing literature, which reports that inclement weather conditions increase *both* travel time *and* travel time variability. There are two caveats, though: First, the variability in the literature is mostly measured by standard deviation of the travel time distribution, whereas in the present study coefficient of variation (which is standard deviation over the distribution mean) is used to measure the variability. As a result, the measures of travel time variability do not match. Second, researchers mainly study freeway travel time variability, which is different from urban networks a fact recognized in the literature.

Considering that the travel time variability (and its sibling, *reliability*) is now becoming a measure for level of service (18), a factor in policy making and investment decisions (19,20), a component of cost-benefit analysis (12,21), and a parameter in various other transportation considerations (22), a clear understanding of the factors that affect travel time variability is needed. By using a rich dataset and applying the appropriate analytical methods, the present study provides valuable insights for the discussion of travel time variability in urban networks.

REFERENCES

1. Yaron Hollander. *The Cost of Bus Travel Time Variability*. Ph.D. Dissertation, The University of Leeds, 2007.
2. Martchouk, M., F. L. Mannering, and L. Singh. Travel Time Reliability in Indiana. Publication FHWA/IN/JTRP-2010/08. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2010
3. Cambridge Systematics Inc. and Texas Transportation Institute "Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation" Technical Report Prepared for Federal Highway Association (FHWA), 2005.
4. Tilahun, N. Y., and D. M. Levinson. A Moment of Time: Reliability in Route Choice Using Stated Preference. Transportation Research Board 87th Annual Meeting, Washington, D.C., 2006.
5. Elefteriadou L., and C. Xiao. Review of Definitions of Travel Time Reliability, Draft Report 2005. Available at: http://www.dot.state.fl.us/planning/statistics/mobilitymeasures/define-ttr.pdf (Last accessed July 30th 2012)
6. Lomax Tim, Schrank David, Turner Shawn, Margiotta Richard. Selecting Travel Reliability Measures", Report, Texas Transportation Institute and Cambridge Systematics, 2003. Available at: http://tti.tamu.edu/documents/474360-1.pdf (Last accessed July 30th 2012)
7. Fosgerau, M., Hjorth K., Brems, C., Fukuda, D. "Travel Time Variability Definition and Valuation", Technical Report, Denmark Technical University, 2008.
8. Noland, R. B., Polak, J. W. "Travel Time Variability: A review of Theoretical and Empirical Issues" *Transport Reviews*, Vol.22(1), 2002, pp.39-54.
9. Yazici, M. A., Kamga, C., Mouskos, K., "Analysis of Travel Time Reliability in New York City Based on Day-of-Week Time-of-Day Periods" *Transportation Research Record: Journal of the Transportation Research Board*, 2012. (Paper in Press)
10. Chien S. I., Liu, X. "An Investigation of Measurement for Travel Time Variability" in *Intelligent Transportation Systems*, InTech, 2012.
11. Chien, S. I., Kolluri, K. K., "Evaluation of Freeway Travel Time Variability and Reliability under Adverse Weather with TRANSMIT Data" Presented in Transportation Research Board's 89th Annual Meeting, Washington D. C., 2010.
12. Peer, S., Koopmans, C., Verhoef, E. T. "Predicting Travel Time Variability for Cost-Benefit Analysis" Tinbergen Institute Discussion Paper, 2010.
13. Tu, H., Van Lint, J. W. C., Van Zuylen, H. J. "The Impact of Adverse Weather on Travel Time Variability of Freeway Corridors" Presented in Transportation Research Board's 86th Annual Meeting, Washington D. C., 2007.
14. Kwon, J., Barkley, T., Hranac, R., Petty, K., Compin, N. "Decomposition of Travel Time Reliability into Various Sources: Incidents, Weather, Work Zones, Special Events, and Base

Capacity" Presented in Transportation Research Board's 90th Annual Meeting, Washington D. C., 2011.

15. Franklin, J. P. and A. Karlstrom.Travel Time Reliability for Stockholm Roadways Modeling Mean Lateness Factor. In *Transportation Research Record: The Journal of the Transportation Research Board*, No. 2134, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 106-113.
16. Breiman L., Friedman J., Olshen R. A., and Stone C. J. *Classification and Regression Trees*. Chapman and Hall, 1984.
17. MATLAB Help, Version 7.11.0.584 (R2010b), 2010.
18. Chen, C., Skabardonis, A., and P. Varaiya. Travel Time Reliability as a Measure of Service. In *Transportation Research Record: The Journal of the Transportation Research Board*, No. 1855, Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 74-79.
19. Lyman, K., and R. L. Bertini. Using Travel Time Reliability Measures to Improve Regional Transportation Planning and Operations. In *Transportation Research Record: The Journal of the Transportation Research Board*, No. 2046, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 1-10.
20. Van der Waard, J. Role of Reliability in Policy, System Planning and Investment Decisions. *International Meeting on Value of Travel Time Reliability and Cost-Benefit Analysis*, Vancouver, British Columbia, October 15, 2009.
21. Taylor, M. Reliability and cost-benefit analysis in Australia and New Zealand.*International Meeting on Value of Travel Time Reliability and Cost-Benefit Analysis*, Vancouver, British Columbia, October 15, 2009.
22. Online Proceedings of International Meeting on Value of Travel Time Reliability and Cost-Benefit Analysis. Available at: http://www.internationaltransportforum.org/Proceedings/reliability/index.html