The 6th International Conference on Construction Engineering and Project Management (ICCEPM 2015)
Oct. 11 (Sun) ~ 14 (Wed) 2015 • Paradise Hotel Busan • Busan, Korea
www.iccepm2015.org

Spatiotemporal Impact Assessments of Highway Construction: Autonomous SWAT Modeling

Kunhee Choi[1] and Junseo Bae[2]

Abstract: In the United States, the completion of Construction Work Zone (CWZ) impact assessments for all federally-funded highway infrastructure improvement projects is mandated, yet it is regarded as a daunting task for state transportation agencies, due to a lack of standardized analytical methods for developing sounder Transportation Management Plans (TMPs). To circumvent these issues, this study aims to create a spatiotemporal modeling framework, dubbed "SWAT" (Spatiotemporal Work zone Assessment for TMPs). This study drew a total of 43,795 traffic sensor reading data collected from heavily trafficked highways in U.S. metropolitan areas. A multilevel-cluster-driven analysis characterized traffic patterns, while being verified using a measurement system analysis. An artificial neural networks model was created to predict potential 24/7 traffic demand automatically, and its predictive power was statistically validated. It is proposed that the predicted traffic patterns will be then incorporated into a what-if scenario analysis that evaluates the impact of numerous alternative construction plans. This study will yield a breakthrough in automating CWZ impact assessments with the first view of a systematic estimation method.

Keywords: Spatiotemporal Modeling Framework; Transportation Infrastructure Improvement; Construction Work Zone Impact Assessments; and Traffic Demand Prediction

I. Highway Infrastructure Rehabilitation

Most state highways in the United States, constructed between the 1950s and 1980s, have been deteriorated by exceeding their original 20-year design lives and increasing traffic demands (*1-3*). The aging of highway infrastructure systems has resulted in numerous highway rehabilitation projects that strive to be long lasting and require less maintenance. Nevertheless, repairing the deteriorated infrastructure systems is challenging because it can cause costly traffic delays and disruptions to the traveling public and surrounding communities during construction. Given the significant economic impact, the FHWA recently launched the Every Day Count initiative to emphasize the importance of shortening project duration (*4*). Also, in 2012, MAP-21, the Moving Ahead for Progress in the 21st Century Act, addressed key challenges that State Transportation Agencies (STAs), daily commuters and business sectors face ─ accelerating project delivery and reducing traffic congestion during lane closures (*4*).

II. Significance of Research: Motivating Case

Highway infrastructure improvement projects in heavily trafficked metropolitan areas frequently cause severe traffic congestion, resulting in the average driver burning an extra 38 hours and 19 gallons of fuel for a congestion cost of $818 in 2011 alone (*5*). To address this issue, changes that the Federal Highway Administration (FHWA) made to "23 Code of Federal Regulations (CFR) 630 Subpart J" in 2003 included the "Work Zone Safety and Mobility Rule" mandated that Construction Work Zone (CWZ) impact assessments be completed for all federally-funded highway infrastructure improvement projects (*6*).

A major focal point of the rule is on enforcing STAs to develop and implement Transportation Management Plans (TMPs) for each of their projects. All state and local governments that receive federal-aid funding are now required to comply with the provisions of the rule, which went into effect in October, 2007 (*7*).

However, developing a systematic TMP has typically been a daunting task for STAs, often impeding timely delivery of a project while adding cost. Most STAs have resorted to ad hoc estimation methods to carry out a work zone impact analysis, due to the lack of standardized methods and available tools, and so continue to be challenged by time delays that occur during the mandated implementation of the CWZ impact analysis.

III. Gaps in Current Knowledge

As a way to control congestion problems and evaluate facility performance, Highway Capacity Manual (HCM) is the most widely used traffic analysis techniques that are rapid and reliable for predicting whether a facility would be operating above or below its capacity (*8*). A number of unknown traffic demand data sets are an essential to evaluate the operating status of roadway segments. However, the HCM procedures still have constraints on achieving realistic and reliable travel demands that vary from city to city and from state to state.

To overcome this difficulty, several studies have reported that existing methods of travel time prediction based on the shortest-path algorithm produce biased results (*9; 10*). Instead of using real-world temporal data, they are dependent upon the start time at a source node, using simplistic models with synthetic datasets to represent the temporal aspect of the road network. Obtaining proprietary real-world traffic data is costly and, therefore, an obstacle to developing accurate modeling.

[1] Assistant Professor, Department of Construction Science, Texas A&M University, 3137 TAMU, Francis Hall 305, College Station, TX 77843-3137, kchoi@tamu.edu
[2] Ph.D. Candidate, Department of Construction Science, Texas A&M University, 3137 TAMU, Francis Hall 328, College Station, TX 77843-3137, junsbae07@tamu.edu (*Corresponding Author)

Further, most of the traffic simulators developed in the recent decade adopts microscopic simulation models (*11; 12*). However, the microscopic simulations demonstrate a critical limitation as it cannot capture global descriptions of the traffic flow-rate, density and velocity and often are restricted to synthetic or simplified data (*13*).

In addition, most of the current data summarization and archival techniques are generic and are therefore not designed to leverage the unique characteristics of the traffic data for effective data reduction. For example, the freeway Performance Evaluation Monitoring System (PeMS) that aims to convert freeway sensor data into intuitive tables and graphs that show historical and real-time traffic patterns on highways in California. The PeMS collects and stores data from numerous real-time streaming sensors operated by California Department of Transportation (Caltrans), with no capability to predict future traffic demand (*14*).

IV. Research Objective and Methods

The key objective of this study is threefold: 1) test the validity whether a set of spatiotemporal signature traffic patterns can represent the ongoing traffic flow on segments of the highway network with a fairly accurate approximation; 2) create and validate an Artificial Neural Networks (ANN)-driven model that predict future traffic patterns in a set of geo-spatial clusters; and 3) develop a SWAT framework to automate CWZ impact assessments.

In this study, a spatiotemporal pattern can be translated into the long-term (i.e., season) and short-term (i.e., 24/7) temporal patterns responding to a set of geo-spatially classified clusters. These particular objectives were achieved through the following methodology:

1) Data collection ─ this study drew a total of 43,795 traffic sensored data that were collected from major highways in the City of Los Angeles (LA), California in 2014. Census track data that reveals socio-demographic characteristics were collected from the U.S. Census Bureau, and facility information, such as number of lanes and lane widths, were obtained from the Caltrans PeMS. Subsequently, these two additional datasets were then matched with traffic sensor locations to capture socio-demographic and highway facility characteristics. Furthermore, daily precipitation data during the study year were acquired for the research phase of traffic demand prediction, through AccuWeather, Inc.
2) Multilevel-cluster-driven traffic characterization ─ the collected data were classified into five geo-spatial characteristics to capture a distinct traffic pattern in multilevel clusters that encompass socio-demographic, highway facility, and temporal characteristics. The characterized traffic patterns were then scientifically verified to qualify the accuracy and reliability, using the Measurement System Analysis (MSA).
3) Future traffic demand/pattern prediction ─ using spatiotemporally characterized existing traffic patterns, an ANN-driven prediction model was created to predict future traffic demand/pattern in a distinct set of geo-spatial clusters automatically, for incorporation into the mandated CWZ impact assessments. Lastly, the predictive power of the proposed model was statistically validated.
4) SWAT framework design ─ it is proposed that the derived around-the-clock future traffic patterns can be then incorporated into a CWZ demand-capacity model to perform a what-if scenario analysis that evaluates the impact of various alternative construction plans on the CWZ assessment components.

This study focuses on discovering signature spatiotemporal traffic patterns for incorporation into the required impact assessments in heavily trafficked metropolitan areas where Annual Average Daily Traffic (AADT) volumes are over 250,000 (*15*). This research was performed under the following assumptions and limitations:

- This study assumes that five types of geo-spatial classification, such as commercial and industrial areas in downtown, residential area, attraction and green space, can be applicable to other similar metropolitan areas having over 250,000 AADT.
- It is assumed that any incident that has the higher uncertainty does not occur.
- The focal point of this study is confined only to highway systems, and the scope of highway network is also limited to its mainline, not ramps or HOV lanes, where the traffic flow is much simpler and more predictable than that of a local road network.
- The scope of weather condition is limited to the existence or nonexistence of daily precipitation during the study year.

V. Data Collection

To be successful in discovering the most representative traffic pattern in heavily trafficked urban areas, AADT served as a baseline to sort heavily trafficked urban areas where the AADT volumes are over 250,000 (*15*). Among various heavy AADT clustered areas, the City of LA in California was selected to capture the most feasible signature traffic pattern for incorporation into the CWZ impact assessments because LA has been one of the U.S. cities with the most traffic congestion (*16*). Subsequently, a total of 43,795 traffic sensor reading data was collected from the City of LA in 2014, utilizing the Caltrans PeMS (*17*). Among collectable data from real-time streaming traffic sensors, a key traffic parameter to assess the traffic impact is traffic demand. Traffic demand is defined as hourly traffic volume at a certain point of interest, which needs to be measured (*8*). Census track data encompassing socio-demographic and commute mode characteristics and highway facility information obtained from the PeMS, such as number of lanes and lane width, were then matched with traffic sensor locations to identify the unique traffic characteristics at geo-spatial and socio-demographic cluster levels.

VI. Spatiotemporal Data Characterization

Existing data summarization techniques cannot leverage the unique characteristics of the collected traffic data, which would often result in biased results of traffic patterns when incorporating into a traffic pattern analysis, within a scope of CWZ impacts assessments. To overcome this issue, it is a pressing need to develop data sets that are representative, unbiased, and pertinent to this study. This research phase aims to test the research hypothesis that a set of signature traffic patterns can collectively represent the ongoing traffic flow on a particular segment of the highway network with a fairly accurate approximation. To achieve this goal, data classification was conducted to create a general model that can effectively handle the aforementioned limitation by conducting a multilevel-cluster-driven traffic data classification, as shown in Figure 1.

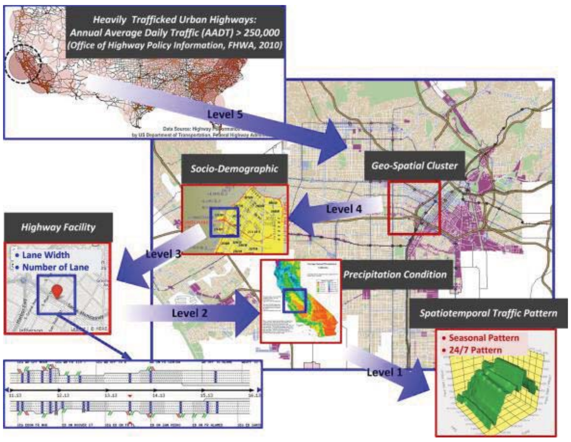

FIGURE 1 Multi-Level-Cluster Driven Classification

Most urban areas include typical geo-spatial clusters, such as downtown, residential and attraction areas, and green spaces. Especially, a study area of Downtown LA is characterized by commercial and industrial uses which are designated by zoning regulations of the City of LA (*18*). A geo-spatial cluster has each unique socio-demographic characteristic (e.g., population density, households, and commute mode) and facility information (e.g., lane width and number of lanes), which would impact its signature traffic pattern for the corresponding road network.

When it comes to the temporal classification for each geo-spatial cluster, a three-level temporal hierarchy process was implemented based on its temporal representation. The first level could use a single signature for each sensor. At the second level, temporal granularity was increased by providing a set of signatures with each one representing a unique day in a week (i.e., Monday to Sunday, and ten holidays). At the third level, temporal granularity was increased by introducing seasonal information.

A. *Multilevel-Cluster-Driven Spatiotemporal Characterization*

The multilevel-cluster-driven classification approach was applied to explore spatiotemporal traffic patterns. Figure 2 shows the average 24/7 traffic patterns responding to the geo-spatial classification. The trend of discovered geo-spatial traffic patterns (i.e., maximum average hourly volume) resembles the magnitude ordering of households (dwelling units) parameter, rather than that of other socio-demographic characteristics, such as population density and commute mode.

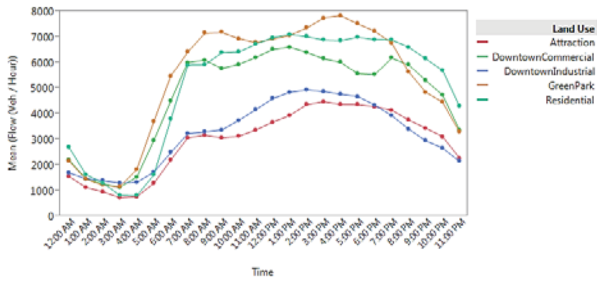

FIGURE 2 Traffic Characterization Responding to Geo-Spatial Clusters

It is noticeable that season and 24/7 temporal traffic patterns differ from geo-spatial classification to geo-spatial classification, which means that it is needed to develop realistic models for traffic flows in spatiotemporally combined networks. Specifically, long-term temporal traffic patterns in each geo-spatial cluster hold its similar pattern during four seasons, considering the weather condition in the state of California. When it comes to 24/7 temporal patterns, traffic patterns in each geo-spatial cluster were very similar during weekdays, while these patterns in each cluster during weekends and holidays show the lower traffic volume. For example, as seen in Figure 3, in the downtown commercial use area, there was no significant difference in seasonal traffic patterns, having over 6,000 hourly volumes during morning and evening peak hours and lunch hours. In terms of date in a week level, the average hourly volume was over 7,000 during weekday morning peak hours, which is much higher than other time periods.

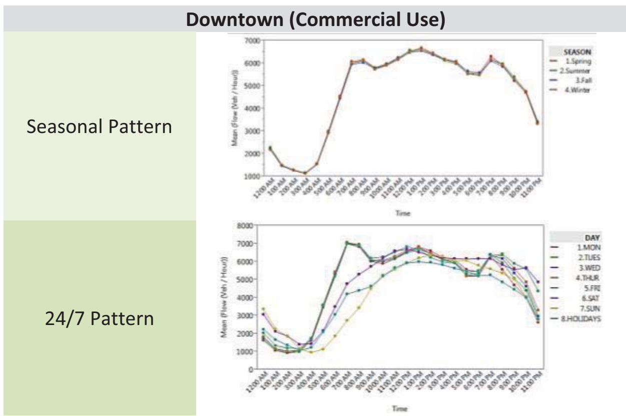

FIGURE 3 Spatiotemporally Discovered Traffic Pattern: Downtown (Commercial Use)

B. *Validation*

To scientifically verify the quality of discovered traffic patterns, a Measurement System Analysis (MSA) was conducted. In general, MSA qualifies a measurement system for use by quantifying its accuracy, precision, and stability (*19; 20*). A key fundamental of MSA lies in the determination of the number of samples and repeat readings (*21*). Larger numbers of parts and repeat readings give results with a higher confidence level. A huge amount of traffic sensor readings employed in this study is well fitted into the MSA approach. Specifically, Gauge Repeatability and Reproducibility (Gauge R&R) and Wheeler's Evaluating the Measurement Process (EMP) methods were used to conduct a MSA.

The results confirmed that all the discovered traffic patterns according to the geo-spatial classification were qualified for use by meeting with the guidance criteria of EMP and Gauge R&R approaches. Except for the measurement in the downtown industrial area, all the classified traffic measurements were involved in the First Class with the R&R percentage of less than 10%, which means the measurement system is acceptable. Meanwhile, the measurement system in the industrial area resulted in the classification of Second Class with R&R of 22.6%, which represents that the measurement system may be acceptable without data improvement.

VII. Future Traffic Demand/Pattern Prediction

One of the main problems frequently delaying project commencement and increasing project overhead cost is that existing methods are limited to collection and analysis of the historical and real-time data with no capability to predict traffic flows (*17*).

To overcome these obstacles, the main objective of this research phase is to test the validity of the research hypothesis that an ANN-driven learning algorithm can predict future traffic patterns reliably and automatically. The proposed prediction model was developed with multi-layer feed-forward networks, which is the most popular and widely-used network approach (*22; 23*). The ANNs learned the relationship between the variables by studying previously recorded data even if the relationship were hidden or hard to describe.

A. ANN-Driven Traffic Demand/Pattern Prediction

A structure of ANN model to predict the future signature traffic patterns in a set of clusters encompasses an input layer (i.e., geo-spatial cluster, households, number of lanes, truck flow, precipitation, seasonal and daily time variation, 24 hours), a hidden layer and an output layer (i.e., hourly traffic volume). Each single neuron is connected to other neurons of a previous layer through adaptable synaptic weights. When each pattern is read through the connection, the network uses the input data to produce an output, which is then compared with the training pattern. Training is the process of adjusting the connection weights employing an appropriate learning method (*23*). When the training arrives at a satisfactory level, the network holds the meaningful weights constant and uses the trained network to identify patterns in new input datasets that were not used in the training process. After the training, the weights contain meaningful learning information whereas they have no meaning and random before training.

The conducted analysis proves the effectiveness of ANNs to predict future traffic demand in a particular spatiotemporal cluster. Nevertheless, the accuracy and reliability of the proposed model still remain uncertain. To verify the accuracy, a comparative analysis was conducted to test the difference between actual and ANN-driven predicted traffic demand/patterns. This comparative analysis produced the following illustrative results, as shown in Figure 4. Most predicted traffic patterns are almost overlapped with the corresponding actual patterns, except for patterns during holidays that happen in spring that were slightly underestimated. In the following validation section, these illustrative comparison results were statistically validated.

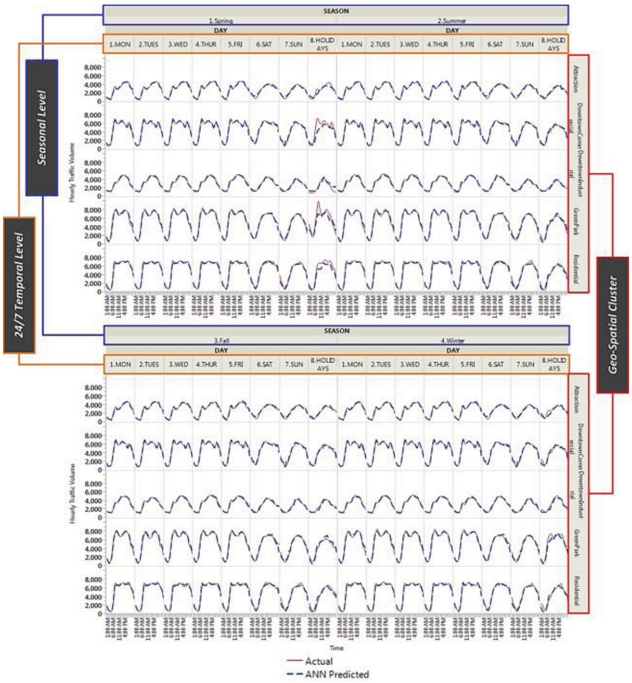

FIGURE 4 Comparisons of Predicted and Actual Spatiotemporal Traffic Patterns (Hourly Traffic Volume)

B. Model Validation

The proposed model's predictive power on future observations can be validated by comparing the predicted values with data not used in the training process. The robustness of the proposed ANN-driven traffic prediction model was statistically validated by using the coefficient of determination (R^2). The R-Square statistic for the validation was 0.964, which signifies that the model is predicting very well on data not used to train the model, with its 96.4% confidence.

VIII. Integration of Predicted Traffic Patterns and Highway Construction Planning

The objective of this task is to quantify the future traffic impacts produced by four construction alternatives such as nighttime, weekday, weekend, and around-the-clock (working 24/7) construction scenarios performed under one of the three typical lane closure schemes (i.e., single lane closure, double lane closure, and full lane closure with counter-flow traffic). Using this approach, the model considers 12 what-if scenarios (i.e., any combination of 4 construction and 3 closure alternatives), which should produce results that fit into one of the modeled distinct signature patterns (by the studied clusters) identified in the previous research phase.

As the future work, the SWAT model will integrate lane closure scenarios as inputs in the algorithm, including 1) duration and occurrence of closures (e.g., 55-hour weekend), 2) starting and ending time (e.g., from 10:00 p.m. Friday to 5:00 a.m. Monday), 3) signature traffic pattern for the given clusters stratified by AADT, 4) closure type (e.g., full closure with counter-flow traffic), and 5) number of affected lanes. The traffic-quantifying model will be used based on predicted temporal traffic patterns and spatial characteristics of highway networks that are developed in the initial research task. The end result will quantify several critical components for a sounder TMP, such as accurate 24/7 traffic patterns during lane closures, queue delays, and road user costs to capture the unique regional characteristic. The result and findings are expected to efficiently and accurately predict certain spatiotemporal traffic patterns and estimate the work zone impacts.

IX. Conclusion

CWZ impact assessments are required by a federal rule, enforcing all STAs to conduct a traffic impact assessment study in a viable way to improve safety and mobility during construction. However, performing the mandated CWZ assessments imposes additional overhead costs, and frequently produces project delay, and STAs have struggled to respond to a difficult mandate since the inception of the rule in 2007.

To overcome these urgent challenges, this study attempted to classify, characterize, and predict, and validate spatiotemporal signature traffic patterns in heavily trafficked metropolitan areas, which specifically aiming at supporting the betterment of CWZ impact assessments. The defining advantage of this study lies in its capability to reliably and realistically automate the prediction of future traffic demands both spatially and temporally. The predicted 24/7 traffic patterns can be then integrated with numerous construction alternative scenarios of transportation infrastructure improvement projects. This study will create significant new knowledge in fully automatic and optimized impact assessments of construction work zones, which contributes to improved safety and mobility by capturing its potential impact ahead of lane closures and by expediting the completion of the required assessment. This study will greatly benefit STAs, the general traveling public, and society in general by significantly improving safety and mobility in and between CWZs, and positively impacting regional development.

References

[1] Lee, E.-B., K. Choi, and S. Lim. Streamlined strategies for faster, less traffic-disruptive highway rehabilitation in urban networks. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2081, No. 1, 2008, pp. 38-45.

[2] Napolitan, F., and P. C. Zegras. Shifting Urban Priorities?: Removal of Inner City Freeways in the United States. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2046, No. 1, 2008, pp. 68-75.

[3] Federal Highway Administration. Life-Cycle Cost Analysis Primer.In, FHWA, Office of Asset Management, Washington D.C., 2002.

[4] Federal Highway Administration. *MAP-21: Moving Ahead for Progress in the 21st Century.* Federal Highway Administration. https://www.fhwa.dot.gov/map21. Accessed 07.05, 2014.

[5] Hasley, A. I. *Washington Rated the Worst for Traffic Congestion — Again.* The Washington Post. http://www.washingtonpost.com/local/trafficandcommuting/washington-rated-the-worst-for-traffic-congestion--again/2013/02/04/125be724-6ee3-11e2-8b8d-e0b59a1b8e2a_story.html. Accessed Dec. 2, 2013.

[6] Choi, K., and Y. H. Kwak. Decision support model for incentives/disincentives time–cost tradeoff. *Automation in Construction,* Vol. 21, 2012, pp. 219-228.

[7] Federal Highway Administration. *Rule on work zone safety and mobility 23 CFR 630 Subpart J*, 2007.

[8] Federal Highway Administration. *6.0 Comparison of Highway Capacity Manual (HCM) and Simulation.* United States Department of Transportation. http://ops.fhwa.dot.gov/trafficanalysistools/tat_vol1/sect6.htm. Accessed 6/30, 2014.

[9] Ding, B., J. X. Yu, and L. Qin. Finding time-dependent shortest paths over large graphs.In *Proceedings of the 11th international conference on Extending database technology: Advances in database technology*, ACM, 2008. pp. 205-216.

[10] Kanoulas, E., Y. Du, T. Xia, and D. Zhang. Finding fastest paths on a road network with speed patterns.In *Data Engineering, 2006. ICDE'06. Proceedings of the 22nd International Conference on*, IEEE, 2006. pp. 10-10.

[11] Hourdakis, J., P. G. Michalopoulos, and J. Kottommannil. Practical procedure for calibrating microscopic traffic simulation models. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 1852, No. 1, 2003, pp. 130-139.

[12] Ben-Akiva, M., M. Bierlaire, H. Koutsopoulos, and R. Mishalani. DynaMIT: a simulation-based system for traffic prediction.In *DACCORS Short Term Forecasting Workshop, The Netherlands*, Citeseer, 1998.

[13] Van Lint, J. W., S. Hoogendoorn, and H. J. van Zuylen. Freeway travel time prediction with state-space neural networks: Modeling state-space dynamics with recurrent neural networks. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 1811, No. 1, 2002, pp. 30-39.

[14] Demiryurek, U., F. Banaei-Kashani, and C. Shahabi. TransDec: A spatiotemporal query processing framework for transportation systems.In *Data Engineering (ICDE), 2010 IEEE 26th International Conference on*, IEEE, 2010. pp. 1197-1200.

[15] Federal Highway Administration. *Travel Monitoring.* Federal Highway Administration. http://www.fhwa.dot.gov/policyinformation/tables/02.cfm.

[16] Gorman, S. *McDonnell says Northern Virginia has the nation's worst traffic congestion.* PolitiFact, Florida. http://www.politifact.com/virginia/statements/2012/dec/20/bob-mcdonnell/mcdonnell-says-northern-virginia-has-ther-nations-/201505/07.

[17] Caltrans. *PeMS.* Caltrans. http://pems.dot.ca.gov/. Accessed 09/15, 2013.

[18] Dept. of City Planning, L. A. *Genralized summary of zoning regulations.* City of Los Angeles, CA. http://planning.lacity.org/. Accessed 02/10, 2014.

[19] Larsen, G. A. Measurement system analysis in a production environment with multiple test parameters. *Quality Engineering,* Vol. 16, No. 2, 2003, pp. 297-306.

[20] Awad, M., T. P. Erdmann, Y. Shanshal, and B. Barth. A measurement system analysis approach for hard-to-repeat events. *Quality Engineering,* Vol. 21, No. 3, 2009, pp. 300-305.

[21] MoreSteam. *Measuremt System Analysis (MSA).* MoreSteam. https://www.moresteam.com/toolbox/measurement-system-analysis.cfm. Accessed 04/27, 2015.

[22] Zhang, G., B. E. Patuwo, and M. Y. Hu. Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting,* Vol. 14, No. 1, 1998, pp. 35-62.

[23] Kalogirou, S. A. Applications of artificial neural-networks for energy systems. *Applied Energy,* Vol. 67, No. 1, 2000, pp. 17-35.