



## Exploration of correlation between environmental factors and mobility at signalized intersections



Rui Guo, Yu Zhang\*

Department of Civil and Environmental Engineering, University of South Florida, 4202 E. Fowler Ave, ENB118, Tampa, FL 33620, United States

### ARTICLE INFO

#### Keywords:

Sustainable transportation  
Environmental factors and traffic metrics  
MOVES  
Multivariate multiple linear regression

### ABSTRACT

Characterizing the relationship between environmental factors and mobility is critical for developing a sustainable traffic signal control system. In this study, the authors investigate the correlation of the environmental impacts of transport and mobility measurements at signalized intersections. A metamodeling-based method involving experimental design, simulations, and regression analysis was developed. The simulations, involving microscopic traffic modeling and emission estimation with an emerging emission estimator, provide the flexibility of generating cases with various intersection types, vehicle types, and other parameters such as driver behavior, fuel types, and meteorological factors. A multivariate multiple linear regression (MMLR) analysis was applied to determine the relationship between environmental and mobility measurements. Given the limitations of using the built-in emissions modules within current traffic simulation and signal optimization tools, the metamodeling-based approach presented in this paper makes a methodological contribution. The findings of this study set up the base for extensive application of simulation optimization to sustainable traffic operations and management. Moreover, the comparison of outputs from an advanced estimator with those from the current tool recommend improving the emissions module for more accurate analysis (e.g., benefit-cost analysis) in practical signal retiming projects.

© 2014 Elsevier Ltd. All rights reserved.

### Introduction

Transportation is a major contributor to energy consumption, greenhouse gas (GHG) emissions, and other environmental pollutants that can cause serious human health issues. For instance, transportation accounts between 20 and 25% of the total energy being consumed among developed countries (WEC, 2007) and it accounts for 27% of GHG emissions in the U.S., which is the second largest source after electricity generation (34%) (EPA, 2013; DOT, 2013). Traditionally, transport planning and operations mainly aim to improve access, mitigate traffic congestion, and smooth traffic flows, but they rarely consider the environmental impacts of transport for sustainable development explicitly. Among many elements of the surface transportation system, signalized intersections along urban arterials are often “hot spots” for fuel consumption and air pollution because of higher traffic density, longer vehicle idling time, and deceleration and acceleration of the driving cycles through the intersections. Existing signal timing optimization tools, including fixed-time, coordinated actuated, traffic responsive, and adaptive control (FHWA, 2008), mainly focus on capturing an optimal cycle length and green-time split to improve mobility (i.e., reducing delays and stops) (Sun et al., 2003; Lv, 2012). Although some of these tools have built-in emission

\* Corresponding author. Tel.: +1 813 974 5846; fax: +1 813 974 2957.

E-mail address: [yuzhang@usf.edu](mailto:yuzhang@usf.edu) (Y. Zhang).

estimation modules when calculating measurements of effectiveness, they are imprecisely estimated by assuming a drive cycle that consists of constant fractions of free-flow and congestion conditions rather than realistic traffic characteristics. To better understand the environmental factors associated with different traffic conditions and control strategies, the ability to adequately model and quantify fuel consumption and emissions at a microscopic level is of high importance.

The built-in emission estimation modules within current traffic simulations and signal optimization tools are relatively under-developed and have a very limited function. The available macroscopic optimization software for traffic signal timing includes Synchro (Husch and Albeck, 2006), TRANSYT-7F (Hale, 2008), PASSER (Chaudhary and Messer, 1993) and SIDRA INTERSECTION (Akçelik and Besley, 1984). Delay and its derivatives are commonly used as objective functions in most optimization software. For example, Synchro optimizes signal settings using a percentile delay, which considers cycle-by-cycle traffic variations. TRANSYT-7F optimizes signal settings using a disutility index, which is based on a combination of delay and stops (Hale, 2008). However, mobility-based optimization is usually insufficient to characterize the fuel consumption and emission levels of real-world driving behavior due to the nature of the macroscopic simulation model. Estimation of fuel consumption in Synchro and TRANSYT-7F is a linear combination of total travel distance, total delay, and total stops, without explicit considerations such as traffic congestion, vehicle type mix, and geometric and environmental factors. Only three types of emissions (i.e., carbon monoxide [CO], nitrogen oxides [NO<sub>x</sub>], and Volatile Organic Compounds [VOCs]) in Synchro are roughly estimated based only on fuel consumption with fixed rates. There is no component of emission estimation in TRANSYT-7F. Besides macroscopic optimization tools, microscopic traffic simulation tools such as TSIS-CORSIM, VISSIM, and Transportation Analysis and Simulation System (TRANSIMS) have been developed to model and evaluate transportation networks in various traffic conditions (PTV AG, 2011; Stevanovic et al., 2009). Microscopic simulation modeling is a faster, safer, and cheaper way to test actual field implementations. Basic input parameters for microscopic simulation models, such as geometry, number of cars, and traffic signal setting, are easily obtained. However, similar to signal optimization tools, microscopic simulation software cannot adequately estimate environmental impacts of a traffic network. Although VISSIM has an add-on module related to emissions, the estimation method of the emission module is simplified without detailed considerations.

As air pollution problems attract more and more attention around the world, many researchers have attempted to incorporate traffic emission factors into traffic control strategies. Dating back to the 1970s, concerns with energy considerations and emissions at intersections were proposed (EPA, 1975; Christian, 1975), and there have been many studies on intersection emissions since then (Patterson, 1976; Tarnoff and Parsonson, 1979; EPA, 1992; Rouphail et al., 2001; Frey et al., 2001). The major limitation of most of these early studies is the lack of a sophisticated way of estimating energy consumption and emissions. Recently, several advanced emission estimation models have been developed, such as Comprehensive Modal Emission Model (CMEM) (Barth et al., 1999), the VT-Micro model (Ahn et al., 2002), and Motor Vehicle Emission Simulator (MOVES) (EPA, 2012). These microscopic models estimate vehicle pollutants at a second-by-second level of resolution using either vehicle engine or vehicle speed/acceleration data. In particular, the emerging model MOVES surpasses previous emission estimation tools. This new emission modeling system is the most sophisticated to date and is being applied at a number of different modeling scales, from the micro-scale (project-level, e.g., parking lot) to the macro-scale, where national-scale inventories are being generated for precursor, criteria, and GHG pollutants from on-road mobile sources (EPA, 2012). The embedded database and project level emission analysis in MOVES provide great opportunities for more accurate emission estimations in the traffic performance analysis.

Recent research and studies also have noted the importance of integrating traffic simulation modeling and advanced emission estimators (Li et al., 2004a,b; Chen and Yu, 2007). For example, Coelho et al. (2005a,b, 2009) explored the relative impact of traffic interruptions (e.g., pay tolls, roundabouts, and traffic signals within the corridor) on traffic performance and emissions (in terms of ratios or percentages). In their study, the research priority was given to relative values of emissions based on European driving behaviors of vehicle fleets. Park et al. (2009) proposed an optimization approach by integrating a CORSIM microscopic traffic simulation, the VT-Micro model, and a Genetic Algorithm (GA). Their study demonstrated that the proposed framework is effective in minimizing fuel consumption and emission with moderate trade-offs in delay and stops. Similarly, Stevanovic et al. (2009) presented an integration of VISSIM, CMEM, and VISGAOST to optimize signal timings. Findings of these studies show that a formula commonly used to estimate fuel consumption in traffic simulation tools inadequately estimates fuel consumption and cannot be used as a reliable objective function in signal timing optimizations. Kwak et al. (2012) quantified the impact of direct traffic signal timing optimization aimed at minimizing fuel consumption based on TRANSIMS, the VT-Micro emission estimator, and a GA-based traffic signal optimizer. Although the GA worked well in these studies, the GA-based optimization consumed significant time and computational loads. More efficient computational techniques should be sought and implemented in the direct optimization way. Lately, Lv (2012) investigated the relationship between emissions and control delay to formulate the optimization problem. Although his study demonstrated the air quality benefit by reducing vehicle emissions under different scenarios, the dataset of vehicle trajectories is quite small, and he considered only control delay as mobility measurement when exploring the relationship between mobility and emissions, with the selection of the parabolic function. Similarly, Zhang et al. (2013) and Osorio and Nanduri (2014) developed a surrogate model or metamodel for the traffic signal optimization with environmental concerns. However, there is short of studies investigating the comprehensive relationship between different objectives (e.g., environmental factors and traffic metrics) at different intersections and different relations for various types of emissions.

Given the limitations of current practice and gaps in existing research, this study aims to explore how the environmental impacts of transport are related to mobility measurements at signalized intersections based on microscopic simulation and multivariate regression analysis. By developing two comparable intersection types, this study also investigated the role of

geometrical characteristics in shaping differences in environmental externalities related to mobility measurements. Such a study could be integrated into the existing signal timing optimization software for a sustainable traffic signal control system that can simultaneously improve mobility and reduce emissions. It will save time and relieve computational loads when compared to direct optimization. In addition, from the practical implementation aspect, the findings from this study can provide more accurate estimation when conducting benefit-cost analyses, where benefits usually include reductions from delays and fuel consumptions, in current signal re-timing projects.

## Proposed framework and methodology

As concluded from the introduction, this study focuses primarily on exploring how environmental externalities are related to mobility measurements at signalized intersections. Although some of the mobility and environmental measurements (e.g., travel time, emission rates) can be collected in the field, it is difficult to collect all factors associated with traffic management operations practically (Golob and Recker, 2004), especially when considering different traffic demand levels with various geometric types and driver behaviors. However, the powerful simulation tools and emerging emission estimator provide the flexibility of using various intersection types, vehicle types, and other characteristics such as driver behavior, fuel types, and meteorological factors.

Thus, in this study, a framework based on a metamodeling technique was developed to analyze the comprehensive relationship between mobility and environmental externalities at signalized intersections. A metamodeling technique involves experimental design, simulation modeling, and regression analysis (Kelton and Law, 2000; Wang and Shan, 2007). The experimental design is used for sampling, and the regression model is developed from the outcomes of simulation modeling. In the existing literature, some studies applied simulation optimization to simultaneously optimize the mobility and environmental impacts of traffic signal timing at intersections (Stevanovic et al., 2009; Kwak et al., 2012). However, due to complicated on-line simulation and tedious computations, the direct optimization method consumes significant time and computational loads. Given the popular coordinated traffic control of corridors and major arterials, methods that can solve multiple intersection problems in an efficient way are urgently needed. The metamodeling-based method proposed in this study will provide a tool for use in simulation optimization and can reduce the complexity and computation load such that it can be used to solve large-scale sustainable traffic management problems.

As shown in Fig. 1, a traffic signal optimization tool is used to provide optimal signal timing for some basic inputs. With the timing and basic inputs, traffic micro-simulation software is applied to generate the detailed information needed for MOVES. Given the mobility and emission measurements, econometrics tools are used to unveil the relationship. The same process is applied to different intersection types. The details of each step of the framework are discussed in the subsections below.

### Traffic signal optimization

The traffic signal optimization tool Synchro, which is used by more than 4000 agencies and consultants throughout North America and the world, was selected to develop mobility-based signal timings for different levels of traffic demand. Three types of data are required for optimization and calculation: geometric information, traffic volumes, and initial signal timings. Measures of effectiveness that are calculated by Synchro software after the optimization process include vehicle delay, fuel consumption, and emissions.

### Traffic micro simulation

Micro simulation models generate a significant amount of detail on vehicle performance that is critical for determining emissions and air quality impacts. In this study, VISSIM was used to develop second-by-second resolutions of individual

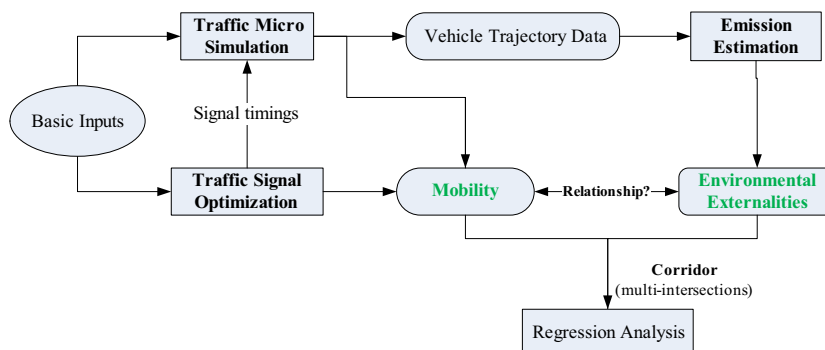


Fig. 1. Proposed framework for the study.

vehicular data (speed/acceleration profiles). The accuracy of a traffic simulation model is mainly dependent on the quality of modeling driver behavior, such as car following and lane changing. In contrast to less complex models that use constant speeds and deterministic car-following logic, VISSIM applies the psychophysical driver behavior model developed by Wiedemann (PTV AG, 2011). Two kinds of data are required for establishing a VISSIM network: (1) static data, representing the roadway infrastructure, and (2) dynamic data, required for traffic simulation applications, which include (a) traffic volumes for all links, (b) vehicle routing, departure times, and dwell times, and (c) priority rules and signal timing plans at intersections. All of these data can be collected from the field. Note that a multi-run for each scenario was conducted in this study to reflect different driver behaviors.

### Emission estimation

The emission estimator MOVES was used to model project-level emissions. There are three approaches/options for describing vehicle activity in MOVES: (1) link average speed, (2) link drive schedules, and (3) operating mode distribution. The link drive schedules and the operating mode distribution approaches are more accurate and widely used in project-level modeling. One of the most important parameters in MOVES is Vehicle Specific Power (VSP), the primary metric used to determine operating modes and estimate emissions. VSP is an estimation of engine load based on vehicle type, vehicle speed and acceleration, and road grade:

$$VSP = (A/M) \times v + (B/M) \times v^2 + (C/M) \times v^3 + (a + g \times \sin \theta) \times v \quad (1)$$

where:  $v$ : velocity;  $a$ : acceleration;  $g$ : road grade;  $M$ : weight;  $A$ : rolling resist;  $B$ : rotating resist;  $C$ : aerodynamic drag. The coefficients  $A$ ,  $B$ ,  $C$ , and  $M$  vary among vehicle types. For example, for a passenger car,  $A = 0.1565$  kW-s/m,  $B = 2.002 \times 10^{-3}$  kW-s<sup>2</sup>/m<sup>2</sup>,  $C = 4.926 \times 10^{-4}$  kW-s<sup>3</sup>/m<sup>3</sup>, and  $M = 1.479$  tons.

The outputs from VISSIM provide the necessary details to calculate the operating mode distribution of the simulated traffic volume. Except for braking and idling, the operating mode bins are stratified by speed ranges (<25 mph, 25–50 mph, and >50 mph) and by VSP. The operating mode bins are weighted by time spent in each bin to represent any driving cycle. Note that one advantage of the emission estimator MOVES is the default data of the contributing factors in the simulator database. MOVES can adjust the default emission rates to represent user-specific values of these factors. Therefore, default data were used for generating the emission rate look-up tables, with the exceptions of link drive schedules, meteorology, vehicle types, and emission types.

### Multivariate regression analysis

Regression analysis is commonly used in the field of air pollution (Vlachogianni et al., 2011). After the complex simulation modeling, correlation and regression analyses were conducted to approximate the environmental responses given microscopic simulation databases. The outcomes of the traffic simulation and the advanced emission estimator are two training sets consisting of Set 2, the evaluations of the environmental response vector  $\{y^m; m = 1, 2, \dots, n\}$  corresponding to Set 1, the mobility measurements  $\{x^m; m = 1, 2, \dots, n\}$  (detailed comments are included later in the illustrative example on forming a database). Suppose we have  $p$  variables in Set 1,  $\mathbf{X} \in \mathbb{R}^{n \times p}$  indicating mobility measurements, and  $q$  variables in Set 2,  $\mathbf{Y} \in \mathbb{R}^{n \times q}$  indicating environmental externalities:

$$\text{Set1 : } \mathbf{X} = [X_{m1}, X_{m2}, \dots, X_{mp}]^T \quad \text{Set2 : } \mathbf{Y} = [Y_{m1}, Y_{m2}, \dots, Y_{mq}]^T \quad m = 1, 2, \dots, n$$

In this study, the first data set (Set 1) related to mobility was measured by delay (e.g., control delay, total delay), stops, average speed, total travel time, and total distance traveled. A dummy variable was used to test if there is a significant difference between two intersection types. The second data set (Set 2), related to environmental factors, includes carbon dioxide (CO<sub>2</sub>) (major GHG emission), CO, NO<sub>x</sub>, particulate matter (PM), and sulfur dioxide (SO<sub>2</sub>) from the U.S. Environmental Protection Agency (EPA) criteria pollutants. Considering the multidimensional characteristics of both sets of variables, a multivariate multiple linear regression (MMLR) was conducted to determine a formula that can describe how elements in a vector of variables respond simultaneously to changes in others. Multivariate statistics encompass the simultaneous observation and analysis of more than one outcome variable. This regression is “multivariate” because there is more than one outcome variable and a “multiple” regression because there is more than one predictor variable (SAS, 2009). Compared to the outcomes from conducting linear regression separately for each response variable on the common set of predictor variable, MMLR can provide large gains in expected prediction accuracy by taking the correlations between the response variables into account (Breiman and Friedman, 1997). The hypothesis being tested by a multivariate regression is that there is a joint linear effect of the set of independent variables on the set of dependent variables. Hence, the null hypothesis is that the slopes of all coefficients are simultaneously zero. The statistical model for MMLR is:

$$\begin{aligned} [y_1 \cdots y_q] &= [x_1 x_2 \cdots x_p] (\beta_1 \cdots \beta_q) + \mathbf{E}_{n \times q} \\ \mathbf{Y}_{n \times q} &= \mathbf{X}_{n \times p} \mathbf{B}_{p \times q} + \mathbf{E}_{n \times q} \end{aligned}$$

where  $\mathbf{Y}$  represents  $n$  observations of a  $q$ -dependent variable,  $\mathbf{X}$  represents the design matrix of rank  $p$  with its first column being the vector 1,  $\mathbf{B}$  is a matrix of parameters to be estimated, and  $\mathbf{E}$  represents the matrix of residual.

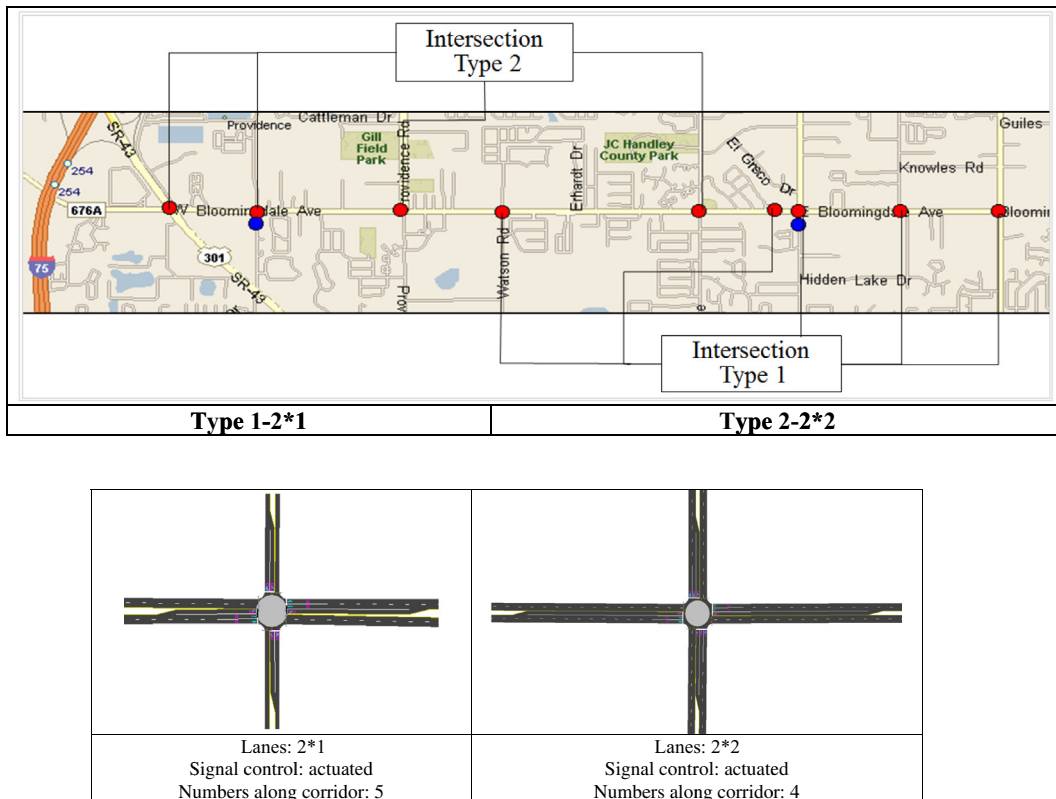
**Illustrative example and results**

This illustrative example explores the relationship between mobility and environmental externalities at signalized intersections. Based on the proposed framework, two typical intersection types along the sample corridor were examined with different levels of traffic volume. The sample corridor, Bloomingdale Avenue, is a four-lane, divided roadway in Hillsborough County, Florida. The Average Annual Daily Traffic (AADT) volumes range from 29,100 vehicles per day (vpd) (east end) to 42,600 vpd (west end) for weekday travel. Fig. 2 illustrates the locations of traffic signals (red) and the placement of BlueTOAD™ (blue) units along the corridor. The field travel time data were collected by the BlueTOAD™ units, which is very useful in the model calibration and validation.

*Sampling for microscopic simulation*

The illustrative intersections in the traffic simulation models were developed based on the two intersection types along the sample corridor, as shown in Fig. 2. They are both four-leg intersections with actuated signal control and urban unrestricted access links. The major difference in these two types of intersections is the number of lanes on the minor street, which determines the capacity of the approaches on the minor street. The speed limits are 45 mph for major roads and 30mph for minor streets. For turning vehicles, the speeds are reduced to 15 mph for a left-turn movement and 9mph for a right-turn movement, respectively.

Thirty scenarios for each type of intersection were designed to input into the traffic simulation models and emission estimator. The high-fidelity outputs were then used for the purpose of multivariate regression analysis. Thirty scenarios for each type are reasonable, given that a sample size of 25–30 is generally considered sufficiently large for most situations in statistical analysis (Howell, 2011). To assess the impact of various levels of traffic demand, different scenarios were generated based on five groups, as shown in Table 1. Group 1, with five scenarios, was based on the collected turning-movement traffic data from the field. Groups 2, 3 and 4 were developed based on the flow ratios with the consideration of different geometric configurations, with or without exclusive turning lanes. It was assumed that the base saturation flow rate was 1900 pc/h/ln in this study. The flow ratio of critical lane groups was calculated by  $v/s$ , where  $v$  is adjusted flow rate in lane group and  $s$  is adjusted saturation flow. The last group was developed to represent various percentages of turning vehicles on major and minor roads.



**Fig. 2.** Map of sample corridor, Bloomingdale Avenue in Tampa, FL.

**Table 1**  
Scenarios for different levels of traffic volume demand.

Scenario groups	Group feature	Number of scenarios
1	Base (average real traffic volume)	Scenarios 5: (0.5, 0.75, 1, 1.25, 1.5) × Base
2	Major Rd: exclusive-left, shared-right lanes Minor St: exclusive-left, shared-right lanes	Scenarios 6–10: (0.1, 0.15, 0.2, 0.25, 0.3) × Saturated Flow
3	Major Rd: exclusive-left, exclusive-right lanes Minor St: exclusive-left, exclusive-right lanes	Scenarios 11–15: (0.1, 0.15, 0.2, 0.25, 0.3) × Saturated Flow
4	Major Rd: exclusive-left, shared-right lanes Minor St: exclusive-left, exclusive-right lanes	Scenarios 16–20: (0.1, 0.15, 0.2, 0.25, 0.3) × Saturated Flow
5	Different left and right turn percentages	Scenarios 21–30

Note: Scenarios 4, 5, 9, 10, 13, 14, 15, and 20 represent congested conditions.

### Quantification of mobility and environmental measurements

Based on the proposed framework and developed scenarios, the traffic signal optimization software Synchro was used to develop mobility-based signal timings for different levels of traffic volume. Two vehicle types, a typical passenger car and a heavy vehicle, were modeled, which corresponds to MOVES vehicle types 21 (passenger car) and 62 (combination long-haul truck with diesel engine). Based on heavy vehicles percentages in the AADT dataset along the studied area, heavy vehicles were set to be 5 percent on major roads and 2 percent on minor roads in all models. Measures of effectiveness, including mobility, fuel consumption, and emissions, were calculated and exported as Synchro outputs.

Within the VISSIM model, exclusive turning lanes were coded with appropriate storage lengths obtained from Google Earth, and vehicle reports were generated for a specified start and end time. Trajectory files generated in VISSIM were configured to output vehicle speed, acceleration, and location within the network on a second-by-second basis. The data were stored in an .FZP file. Note that 10 model runs for each scenario were conducted for this analysis to reflect the stochastic nature of traffic flow and driving behaviors. Calibration issues accompany the use of any microscopic traffic simulation models. Here, calibration means the process of adjusting and fine-tuning some parameters to match local traffic conditions. Since a link in the macroscopic software usually comprises several links in a microscopic simulation network, the links in VISSIM were matched with links in Synchro first. Then, the calibration for basic parameters was conducted to match the real situation. In the base model, the average travel time from Synchro and VISSIM was compared with that from field data collected by the BlueTOAD™ units through the studied corridor. A good agreement was found between measured and simulated travel times. Following previous literature, we determined to keep the default parameters for further analysis. Moreover, the quality of the vehicle trajectory data collected by VISSIM is verified with extensive error checking, which flags any data values outside of conventional ranges (e.g., acceleration greater than  $3 \text{ ms}^{-2}$  or deceleration greater than  $-5 \text{ ms}^{-2}$ ). For congested scenarios, the inputs and outputs of vehicles in the simulation were checked and the missing vehicles were treated as unserved vehicles.

The detailed micro simulation outputs enabled a direct quantitative linkage with MOVES. All links were modeled with zero percent gradients, as no nominal grade changes exist at the intersection. Within MOVES, the vehicle operating modes are stored as an operating mode distribution, which is the percent of all vehicle-hours for a specific link, pollutant, and vehicle type that fall within each operating mode. In our study, links were defined by each segment, including links for traffic approaching and departing at the signalized intersections. In addition, the connections for different turning movements in the center of the intersection were designed as small links. The link drive schedule approach, using user-defined drive cycles in MOVES, was adopted in this research. Based on the user-defined input and default data in the MOVES model, the emission rates were generated after running the MOVES.

In existing macro and micro simulation software (e.g., Synchro, TRANSYT-7F, CORSIM, and VISSIM), the fuel usage and emissions of the unserved vehicles—vehicles that are denied entry (essentially queued at the entry points to the intersections) in the over-saturated/congested scenarios—are ignored. These vehicles use fuel, so in our study they were treated as vehicles idling on the off-network link with appropriate assumptions in MOVES. The start fraction and parked vehicle fraction parameters were both set to zero, which means no vehicle is restarted or parked during congestion. The extended idle parameter was set to 0.95, which reflects the fact that 95 percent of the total vehicle-hours (only 1 h by definition) in the off-network link are spent in an extended idle mode. As in Synchro, the time domain for the emissions estimation is one hour.

### MMLR analysis for model fitting

After the microscopic simulation modeling, statistics software, SAS, was used for MMLR analysis to explore the relationship between environmental externalities and mobility measurements. The data to be analyzed came from the quantification of the last step, with a sample of  $n = 60$ . Two collections of variables were measured, as listed in Table 2. There were seven mobility performance variables in the first group and seven environmental externality variables in the second group. A dummy variable, indicating two comparable intersection types, was used to test if the environmental-mobility relationship

**Table 2**  
Summary of descriptive statistics.

Variable	Mean	Standard deviation	Min	Max
<i>Dependent variables (Environmental measure)</i>				
Y1: CO <sub>2</sub> (ton)	1.00	0.578	0.263	2.92
Y2: CO (kg)	9.23	4.35	2.514	22.249
Y3: NO <sub>x</sub> (kg)	3.72	5.08	0.422	21.987
Y4: PM10 (kg)	0.13	0.09	0.028	0.434
Y5: PM25 (kg)	0.12	0.09	0.024	0.426
Y6: SO <sub>2</sub> (kg)	0.01	0.01	0.002	0.03
Y7: Total energy (Gigajoule)	13.83	7.96	3.64	40.30
<i>Explanatory variables (Mobility measure)</i>				
X1: Control delay (s/veh)	64.06	73.57	8.9	320.3
X2: Total delay (10 <sup>2</sup> h)	1.03	1.58	0.04	6.67
X3: Stops (/veh)	0.68	0.08	0.5	0.8
X4: Total stops (#)	2798.27	1252.20	698	6072
X5: Average speed (mph)	14.42	7.59	2	29
X6: Total travel time (h)	121.57	164.81	10	711
X7: Total distance traveled (veh-mi)	807.55	373.06	245	1985
Type: Dummy variable for intersection type	0.50	0.50	0	1

statistically depends on intersection types. Arguably, differences in locations, road geometries (capacities), and land-use policies can help explain demographic dissimilarities in mobility and environmental performance.

To avoid the multicollinearity problem, the interrelationship among independent variables was computed first, as shown in Table 3, with the bold values indicating the strong relations. Table 3 shows that X3 (*Stops/vehicle*) has relatively weak relation to all other X values (mobility measurements). X5 (*Average speed*) shows negative relation to other variables and not as strong as others' relation. X1 (*Control delay*), X2 (*Total delay*), X4 (*Total stops*), X6 (*Total travel time*) and X7 (*Total distance traveled*) are strongly related, which means only one of them will be selected for regression. Thus, X2, X3, X5 and Type are selected for MMLR analysis. To test if the relationships are statistically non-linear, (X2)<sup>2</sup> and (X3)<sup>2</sup> are also included in the model.

The results of multivariate statistics and *F* approximations for MMLR, shown in Table 4, indicate that all of the equations, taken together, are statistically significant. The *F*-ratios and *p*-values for four multivariate criteria are given, including Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace, and Roy's Greatest Root (SAS, 2009). The tests for the overall model for our study indicate that the model is statistically significant, regardless of the type of multivariate criteria used.

Table 5 summarizes the regression results for MMLR. The adjusted *R* square values, adopted to test how good the model fit to the sample data, show that all models are appropriate and good fits. The signs of the coefficients show if the mobility measurements have the positive or negative impact on the environmental factors. *T*-values show if the coefficients of independent variables are statistically significant. The regression results and *t*-values demonstrate the following: (1) For Y1 (CO<sub>2</sub>), Y2 (CO), Y3 (NO<sub>x</sub>), Y4 (PM10), Y5 (PM25) and Y7 (*Fuel*), the coefficients of X2 (*Total delay*) are significant at 1% level and the coefficients of X3 (*Stops per vehicle*) are significant at 5% or 1% level; (2) While for SO<sub>2</sub>, the dataset is more scattered, and other models may be needed to find a better-fitting curve (the adjusted *R* square value is relatively smaller); (3) Y3 (NO<sub>x</sub>), Y4 (PM10) and Y5 (PM25) show strongly relationships with not only X2 (*Total delay*), but also with the quadratic term (X2)<sup>2</sup>, statistically showing that they are not just linearly related. Moreover, the negative signs of quadratic terms indicate that the Y-X2 linear slope is getting less positive as X2 increases; (4) Similarly, Y1 (CO<sub>2</sub>), Y2 (CO), Y3 (NO<sub>x</sub>), Y4 (PM10), Y5 (PM25) and Y7 (*Fuel*) show the significant relations with the quadratic term (X3)<sup>2</sup>, with the Y-X3 linear slope getting less positive as X3 increases; (5) For X5 (*Average speed*), only the coefficients for model Y4 (PM10) and Y5 (PM25) are not significant at 5% level, and the sign of the coefficient for Y3 (NO<sub>x</sub>) is different from the others, which needs detailed investigation (e.g., eco-driving study); (6) For the dummy variable (i.e., intersection type), the coefficients for model Y1 (CO<sub>2</sub>), Y2 (CO), Y4 (PM10), Y5 (PM25), Y6 (SO<sub>2</sub>) and Y7 (*Fuel*) are also significant at 5% level, which means these six environmental factors would be statistically different for different signalized intersection types. On the other hand, the coefficient of Type for nitrogen

**Table 3**  
Correlations for independent variables-mobility measurements.

	X1	X2	X3	X4	X5	X6	X7
X1	1.00						
X2	<b>0.97</b>	1.00					
X3	0.32	0.20	1.00				
X4	<b>0.88</b>	<b>0.87</b>	0.49	1.00			
X5	-0.83	-0.76	-0.48	<b>-0.89</b>	1.00		
X6	<b>0.97</b>	<b>1.00</b>	0.21	<b>0.88</b>	-0.77	1.00	
X7	<b>0.83</b>	<b>0.86</b>	0.33	<b>0.95</b>	-0.78	<b>0.87</b>	1.00

**Table 4**  
Multivariate statistics and *F* approximations.

Multivariate statistics and <i>F</i> approximations					
<i>S</i> = 7 <i>M</i> = −0.5 <i>N</i> = 22					
Statistic	Value	<i>F</i> value	Num DF	Den DF	Pr > <i>F</i>
Wilks' Lambda	0.00004365	30.23	49	237.96	<0.0001
Pillai's Trace	3.15511872	6.10	49	364	<0.0001
Hotelling-Lawley Trace	228.28867490	207.97	49	139.57	<0.0001
Roy's Greatest Root	207.73050599	1543.14	7	52	<0.0001

Note: *F* Statistic for Roy's Greatest Root is an upper bound.

**Table 5**  
Results of MMLR with coefficients and *t*-values.

	Environmental factors						
	Y1 (CO <sub>2</sub> )	Y2 (CO)	Y3 (NO <sub>x</sub> )	Y4 (PM10)	Y5 (PM25)	Y6 (SO <sub>2</sub> )	Y7 (Fuel)
Adj. <i>R</i> <sup>2</sup>	0.911	0.924	0.990	0.977	0.978	0.774	0.910
Constant	−1.737 (−1.17)	−14.004 (−1.35)	−11.417 (−2.57) <sup>b</sup>	−0.275 (−2.26) <sup>b</sup>	−0.282 (−2.41) <sup>b</sup>	−0.015 (−0.62)	−23.939 (−1.16)
X2 (Total delay)	0.321 (3.57) <sup>a</sup>	2.025 (3.23) <sup>a</sup>	4.061 (15.11) <sup>a</sup>	0.069 (9.34) <sup>a</sup>	0.064 (9.13) <sup>a</sup>	0.0008 (0.54)	4.410 (3.53) <sup>a</sup>
X2×X2	−0.014 (−1.17)	−0.080 (−0.94)	−0.113 (−3.10) <sup>a</sup>	−0.003 (−2.87) <sup>a</sup>	−0.002 (−2.48) <sup>b</sup>	0.00002 (0.09)	−0.198 (−1.17)
X3 (Stops per vehicle)	9.312 (2.18) <sup>b</sup>	77.163 (2.59) <sup>b</sup>	31.705 (2.49) <sup>b</sup>	1.091 (3.12) <sup>a</sup>	1.094 (3.25) <sup>a</sup>	0.116 (1.67)	128.768 (2.17) <sup>b</sup>
X3×X3	−7.388 (−2.32) <sup>b</sup>	−58.085 (−2.62) <sup>b</sup>	−24.050 (−2.53) <sup>b</sup>	−0.814 (−3.12) <sup>a</sup>	−0.808 (−3.22) <sup>a</sup>	−0.094 (−1.80)	−102.192 (−2.31) <sup>b</sup>
X5 (Average speed)	−0.021 (−2.60) <sup>b</sup>	−0.189 (−3.38) <sup>a</sup>	0.069 (2.87) <sup>a</sup>	−0.0008 (−1.22)	−0.0008 (−1.28)	−0.0005 (−3.59) <sup>a</sup>	−0.290 (−2.61) <sup>b</sup>
Type (Intersection type)	−0.236 (−4.60) <sup>a</sup>	−2.222 (−6.23) <sup>a</sup>	0.139 (0.91)	−0.014 (−3.45) <sup>a</sup>	−0.014 (−3.56) <sup>a</sup>	−0.004 (−5.18) <sup>a</sup>	−3.273 (−4.61) <sup>a</sup>

N=60.

<sup>a</sup> Significant at 1%.

<sup>b</sup> Significant at 5%.

dioxide is not significant, meaning that the relationship between NO<sub>x</sub> with mobility would be statistically similar for different signalized intersection types. Thus, we can conclude that the current practice of mobility-based optimization (e.g., Synchro as an example in this study) is not good enough to reduce SO<sub>2</sub> emission. The relationships for certain pollutants (e.g., NO<sub>x</sub>, PM10, and PM25) are not just linear and the effect of mobility measure on environmental metric gets less positive as it increases. Moreover, environmental factors (except NO<sub>x</sub>) statistically show different distributions at different types of signalized intersections, since the coefficients of Type are significant for these pollutants. Thus, for a corridor-level optimization, it is not trivial to incorporate environmental factors into the formulation of control strategies given their different distributions. The trade-off between different types of intersection in the same corridor need to be specifically considered in designing control strategies. Furthermore, eco-driving capabilities and connected vehicle technologies that allowing communication between vehicles and control devices are highly recommended for reducing these pollutants.

#### Comparisons of results from MOVES and Synchro

The developed model was compared with the model used in present practice. Currently, VISSIM, TRANSYT, and SYNCHRO share the same fuel consumption formula, which is based on a linear combination of total travel, delay, and stops, as shown in Eq. (2).

$$F = K_{i1} \times TT + K_{i2} \times D + K_{i3} \times S \quad (2)$$

where, *F* = fuel consumed in gallons per hour; *TT* = total travel in vehicle miles (veh-mi) per hour; *D* = total delay in veh-h/h; *S* = total stops per hour;  $K_{i1} = 0.075 - 0.0016 \times V_i + 0.000015 \times V_i^2$ ;  $K_{i2} = 0.7329$ ;  $K_{i3} = 0.0000061 \times V_i^2$ ; *V<sub>i</sub>* = cruise speed on each link *i* (mph). Note that gallons were converted to MetaJoules in the comparison (1 gallon automotive gasoline = 131.76 MetaJoules).

It is worth pointing out the limitations of the current model. First, the model parameters were determined from studies conducted with only one test vehicle; second, no explicit consideration was given to factors such as traffic congestion, vehicle type mix (i.e., trucks and diesel engines), and geometric and environmental factors such as road gradient, curvature, surface quality, temperature and other factors; third, the model coefficients have not been adjusted for vehicle fleet mix since



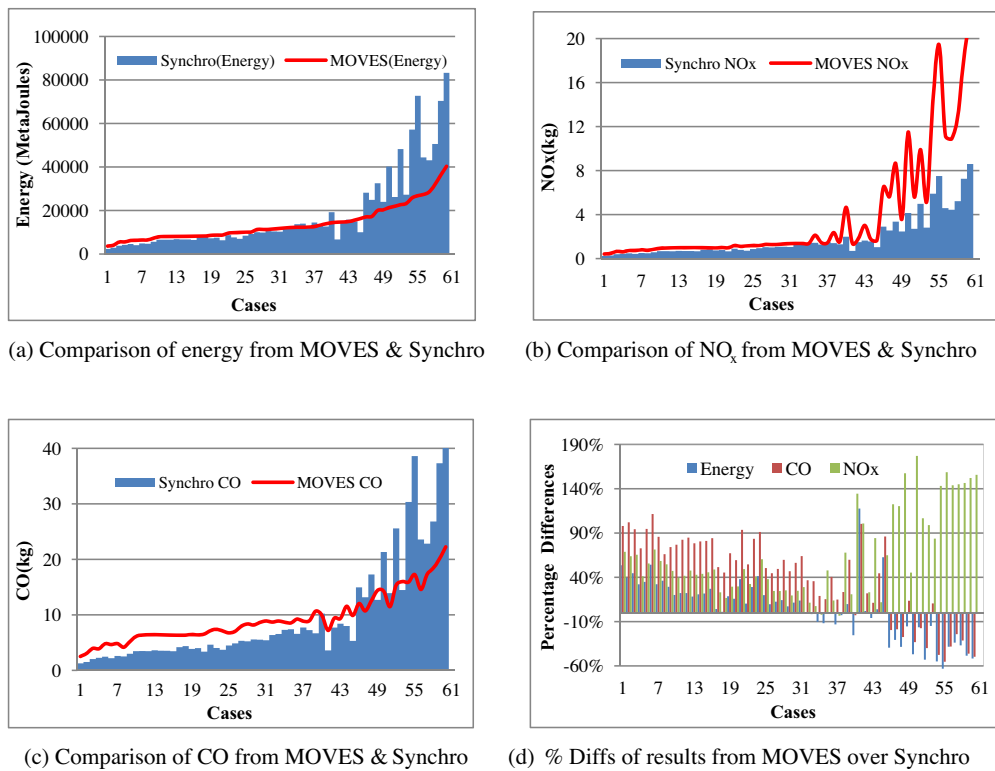


Fig. 3. Comparisons of energy, CO, and NO<sub>x</sub> from MOVES and Synchro.

1983 (Hale, 2008). In contrast, the proposed metamodel developed from the microscopic simulation database is based on various intersection types, traffic volume levels, vehicle types, driver behaviors, and detailed environmental factors.

Currently, Synchro calculates only CO, NO<sub>x</sub>, and VOCs emissions as the environmental performance measurements. When this study was conducted, at the MOVES project level, it could not model evaporative emission processes yet. It is indicated that such a capability will be added to future model upgrades (EPA, 2012). Thus, the comparison conducted in this study only look into energy consumption, CO, and NO<sub>x</sub> emissions, as shown in Fig. 3. Different from the scenarios' sequence in Table 1, the cases' sequence in Fig. 3 are re-ordered according to congestion conditions. Specifically, the 60 cases (sample size = 60 for two intersections) are ranked based on low to high energy consumption estimated by MOVES. From Fig. 3(a) and (c), it is clearly shown that Synchro overestimates energy consumption and CO emission in most cases. For over-saturated scenarios, Synchro significantly underestimates energy consumption and CO emission. From Fig. 3(b), for almost all scenarios, Synchro significantly underestimates NO<sub>x</sub> emission, which could be because Synchro has no way of calculating the emission from the un-served vehicles. Fig. 3(d) also tells us that in most uncongested scenarios, the percentage difference of CO between two tools is the highest, and in over-saturated situations, the percentage difference of NO<sub>x</sub> is the highest. Thus, it is recommended that the current signal optimization tool, Synchro in this study, should improve its evaluation methods of environmental impacts from transport to get a more accurate analysis (e.g., benefit-cost analysis) in practical signal re-timing projects.

## Conclusions and future work

Characterizing the relationship between environmental impacts from transport with mobility is critical for sustainable development. In this study, the authors developed a framework to determine how environmental externalities are related to mobility measurements during the same time period at signalized intersections. A metamodeling-based framework, involving experimental design, microscopic simulation (i.e., a traffic signal optimization tool, a microscopic simulation model, and an instantaneous emission estimator), and multivariate regression analysis were developed to explore the environment-mobility relationship at signalized intersections. Given the microscopic simulation databases, MMLR analysis was conducted to approximate the environmental responses to the mobility measurements. The results showed good fits for multiple-responses. However, *t*-values, which indicate if the coefficients of independent variables are statistically significant, showed various conclusions for different response variables (i.e., energy and emissions). The regression outcomes show that, to reduce SO<sub>2</sub>, the mobility-based optimization is not good enough. The relationships for certain pollutants (e.g., NO<sub>x</sub>, PM10,

and PM<sub>2.5</sub>) are not just linear. Furthermore, the relationships between these emissions and mobility measurements considered in the study are different for various types of intersections, which requires the consideration of trade-offs between different intersections in a coordinated arterial while pursuing eco-friendly traffic control. The results of quantitative assessment from the microscopic emission estimator were compared with the estimation from the current signal optimization tool Synchro. The comparison results recommend the improvement of the current emissions module in the tool for more accurate analyses (e.g., benefit–cost analysis) in practical signal retiming projects.

In future research, other types of regression models such as radial basis functions (RBF), multivariate adaptive regression splines (MARS), Kriging, and support vector machine (SVM) can be used and compared with MMLR used in this study. It should be noted that traffic simulation models may not accurately represent vehicle dynamics and the speed and acceleration distributions could be different from field data depending on how the parameters for human behaviors in VISSIM are calibrated (Song et al., 2013). If the readers intend to apply our proposed framework, they are encouraged to carefully calibrate the simulation model and obtain the specific correlation between mobility and emissions for their study regions. We also recommend incorporating NGSIM vehicle trajectory datasets and field travel time data using advanced technologies (i.e., BlueTOAD™ technology) to validate the simulation model. Validation is a confirmation process to justify whether the calibrated simulation network reliably replicates real traffic conditions with a new set of field data that are not used in the calibration process. Moreover, emissions data can be collected in the field to validate the estimation results in the future. Another extension of this study could be metamodeling-based optimization for a sustainable traffic signal control system that can simultaneously improve mobility and reduce emissions. Metamodeling-based optimization will save much time and relieve computational loads when compared to direct optimization.

## Acknowledgements

This research was funded by the Florida High Tech Corridor program and Albeck Gerken Inc. We appreciate the traffic data provided by Albeck Gerken Inc. and are very grateful to Jeff Gerken, Patrick O'Connor, and Transportation Research Board (TRB) anonymous peer-reviewers for their valuable comments and advice. Moreover, we would like to thank Srisri Pendem and Patricia Ball for their help on formatting and technical editing.

## References

- Ahn, K., Rakha, H., Trani, A., Van Aerde, M., 2002. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. *J. Transp. Eng.* 128 (2), 182–190.
- Akçelik, R., Besley, M., 1984. *Sidra-2 User Guide*. ARRB, Victoria.
- Barth, M., An, F., Younglove, T., Scora, G., et al., 1999. User's guide: Comprehensive modal emissions model, Version 2.0. Research report prepared by the University of California, Riverside, CA.
- Breiman, L., Friedman, J.H., 1997. Predicting multivariate responses in multiple linear regression. *J. Roy. Stat. Soc.* 59 (1), 3–54.
- Chaudhary, N.A., Messer, C.J., 1993. PASSER-IV: A program for optimizing signal timing in grid networks. In 72nd Annual Meeting of the Transportation Research Board, Washington, DC, 82–93.
- Chen, K., Yu, L., 2007. Microscopic traffic–emission simulation and case study for evaluation of traffic control strategies. *J. Transp. Syst. Eng. Inf. Technol.* 7 (1), 93–100.
- Christian, S.B., 1975. Some energy considerations in traffic signal timing. *Traffic Engineering*, Washington, DC, pp. 19–25.
- Coelho, M., Farias, T., Roupail, N., 2005a. Impact of speed control traffic signals on pollutant emissions. *Transp. Res. Part D* 10 (4), 323–340.
- Coelho, M., Farias, T., Roupail, N., 2009. Measuring and modeling emission effects for toll facilities. *Transport. Res. Rec. No. 1941*, 136–144.
- Coelho, M., Farias, T., Roupail, N., 2009. A numerical tool for estimating pollutant emissions and vehicles performance in traffic interruptions on urban corridors. *Int. J. Sustainable Transp.* 3 (4), 246–262.
- Department of Transportation (DOT), 2013. Transportation's Role in Climate Change. <<http://climate.dot.gov/about/transportations-role/overview.html>>.
- Environmental Protection Agency (EPA), 1975. Guidelines for air quality maintenance planning and analysis. Volume 9: Evaluating indirect sources.
- Environmental Protection Agency (EPA), 1992. User's guide to CAL3QHC Version 2.0: A modeling methodology for predicting pollutant concentrations near roadway intersections.
- Environmental Protection Agency (EPA), 2012. Motor Vehicle Emissions Simulator (MOVES) user guide for MOVES2010b. Assessment and Standards Division, Office of Transportation and Air Quality.
- Environmental Protection Agency (EPA), 2013. Sources of Greenhouse Gas Emissions. <<http://www.epa.gov/climatechange/ghgemissions/sources.html>>.
- Federal Highway Administration (FHWA), 2008. Traffic signal timing manual. U.S. Department of Transportation. FHWA-HOP-08-024.
- Frey, H.C., Roupail, N.M., Unal, A., Colyar, J.D., 2001. Emissions reduction through better traffic management: An empirical evaluation based upon on-road measurements (No. FHWA/NC/2002-001).
- Golob, T.F., Recker, W.W., 2004. A method for relating type of crash to traffic flow characteristics on urban freeways. *Transp. Res. Part A: Policy Pract.* 38 (1), 53–80.
- Hale, D., 2008. *Traffic Network Study Tool—TRANSYT-7F User Guide, United States version*. McTrans Center, University of Florida, Gainesville.
- Howell, D.C., 2011. *Statistical Methods for Psychology*. Cengage Learning.
- Husch, D., Albeck, J., 2006. *Synchro Studio 7 user guide*. Trafficware Ltd.
- David, Kelton W., Law, Averill M., 2000. *Simulation Modeling and Analysis*. McGraw Hill, Boston, MA.
- Kwak, J., Parkand, B., Lee, J., 2012. Evaluating the impacts of urban corridor traffic signal optimization on vehicle emissions and fuel consumption. *Transp. Planning Technol.* 35 (2), 145–160.
- Li, X.G., Li, G.Q., Pang, S.S., et al., 2004a. Signal timing of intersections using integrated optimization of traffic quality, emissions and fuel consumption: a note. *Transp. Res. Part D* 9, 401–407.
- Li, X., Li, G., Pang, S.-S., Yang, X., Tian, J., 2004b. Correlation between land-use change and greenhouse gas emissions in urban areas. *Int. J. Environ. Sci. Technol.* 9 (5), 401–407.
- Lv, J.P., 2012. Investigating emission reduction benefit from intersection signal optimization. ITE Daniel B. Fambro Student Paper Award Competition, Texas District.
- Osorio, C., Nanduri, K., 2014. Energy-efficient urban traffic management: a microscopic simulation-based approach. *Transport. Sci.* (accepted for publication).
- Park, B., Yun, I., Ahn, K., 2009. Stochastic optimization for sustainable traffic signal control. *Int. J. Sustainable Transp.* 3 (4), 263–284.

- Patterson, R. 1976. Air quality modeling at signalized intersections. Transportation Research Board, Special Report 167, 138–151.
- PTV AG., 2011. VISSIM 5.40-01 user manual. Karlsruhe, Germany.
- Rouphail, N.M., Frey, H.C., Colyar, J.D., Unal, A. (2001, January). Vehicle emissions and traffic measures: exploratory analysis of field observations at signalized arterials. In 80th Annual Meeting of the Transportation Research Board, Washington, DC.
- Institute, S.A.S., 2009. SAS/STAT User's Guide: Version 9.2, 2nd ed. SAS Institute Inc., Cary, NC, USA.
- Song, G., Yu, L., Xu, L., 2013. Comparative analysis of car-following models for emissions estimation. *Transport. Res. Rec.* 2341 (1), 12–22.
- Stevanovic, A., Stevanovic, J., Zhang, K., Batterman, S., 2009. Optimizing traffic control to reduce fuel consumption and vehicular emissions. *Transport. Res. Rec.* No. 2128, 105–113.
- Sun, D.Z., Benekohal, R.F., Waller, S.T., 2003. Multi-objective traffic signal timing optimization using non-dominated sorting Genetic Algorithm. *Proceedings IEEE Intelligent Vehicles Symposium*, 198–203.
- Tarnoff, P., Parsonson, P., 1979. Guidelines for selecting traffic signal control at individual intersections. National Cooperative Highway Research Program, American Association of State Highway and Transportation Officials, United States. Federal Highway Administration.
- Vlachogianni, A., Kassomenos, P., et al, 2011. Evaluation of a multiple regression model for the forecasting of the concentrations of NO<sub>x</sub> and PM10 in Athens and Helsinki. *Sci. Total Environ.* 409, 1559–1571.
- Wang, G.G., Shan, S., 2007. Review of metamodeling techniques in support of engineering design optimization. *J. Mech. Des.* 129 (4), 370–380.
- World Energy Council (WEC), 2007. Transport Technologies and Policy Scenarios. World Energy Council. Retrieved 2009-05-26.
- Zhang, L., Yin, Y., Chen, S., 2013. Robust signal timing optimization with environmental concerns. *Transport. Res. C: Emer.* 29, 55–71.