



CIRRELT

Centre interuniversitaire de recherche
sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre
on Enterprise Networks, Logistics and Transportation

Measuring Fuel Consumption in Vehicle Routing: New Estimation Models Using Supervised Learning

Hamza Heni
S. Arona Diop
Leandro C. Coelho
Jacques Renaud

March 2019

CIRRELT-2019-08

Document de travail également publié par la Faculté des sciences de l'administration de l'Université Laval,
sous le numéro FSA-2019-004.

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Measuring Fuel Consumption in Vehicle Routing: New Estimation Models Using Supervised Learning*

Hamza Heni**, S. Arona Diop, Leandro C. Coelho, Jacques Renaud

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Department of Operations and Decision Systems, 2325 de la Terrasse, Université Laval, Québec, Canada G1V 0A6

Abstract. In this paper we propose and assess the accuracy of new fuel consumption estimation models for vehicle routing. Based on real-world data of instantaneous fuel consumption, time-varying speeds observations, and high-frequency traffic data related to a large set of shipping operations we propose effective methods to estimate fuel consumption and greenhouse gas emissions. By carrying out nonlinear regression analysis using supervised learning methods, namely Neural Networks, Support Vector Machines, Conditional Inference Trees, and Gradient Boosting Machines, we develop new models that provide better prediction accuracy than classical ones. We correctly estimate consumption for time-dependent point-to-point routing under realistic conditions taking into account freight transportation operations during peak hour traffic congestion, stop-and-go driving patterns, idle vehicle states, and the variation of vehicle loads. Extensive computational experiments on real datasets show the effectiveness of the proposed machine learning emissions models, clearly outperforming the Comprehensive Modal Emissions Model (CMEM) and the Methodology for Estimating air pollutant Emissions from Transport (MEET). Advanced econometric analysis using Monte Carlo simulation are used to calibrate CMEM parameters, significantly improving its quality. Based on sensitivity analysis we show that MEET underestimates real-world consumption by 24.94% and CMEM leads to an overestimation of consumption by 7.57% (13.18% before calibration) according to observed data, while our best machine learning model (Gradient Boosting Machines) exhibited superior estimation accuracy and is off by only 1.70% considering real-world driving conditions. A detailed analysis of the relative importance of input variables confirms the efficiency of our models.

Keywords: Consumption models, time-dependent routing, traffic congestion, machine learning, Monte Carlo simulation.

Acknowledgements. This research was partly supported by grants 2014-05764 and 0172633 from the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Centre d'Innovation en Logistique et Chaîne d'Approvisionnement Durable (CILCAD). The CILCAD receives financial support from the Green Fund under Priority 15.2 of the 2013-2020 Action Plan on Climate Change, a priority implemented by Transition Énergétique Québec (TEQ). We also thank Mr. Jean-Philippe Gagliardi, President of Logix Operations Inc., for providing us with real data from an important wholesaler partner in Québec City. This support is highly appreciated.

*Revised version of CIRRELT-2018-15

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

** Corresponding author: hamza.heni@cirrelt.ca

1 Introduction

Freight transportation is known to be an important source of greenhouse gas (GHG) emissions [46]. GHG emissions are proportional to the fuel consumption which in turn, depends on several factors including speed, acceleration, distance, weight of the vehicle, backhauls and roadway slope [9].

Accurate emissions estimation is a valuable information for transportation experts in making effective decisions that improve routing operations. The current literature on GHG emissions for road freight transportation offers different models for estimating emission and fuel consumption, the more well-known being the Comprehensive Modal Emissions Model (CMEM) [2] and the Methodology for Estimating air pollutant Emissions from Transport (MEET) [25]. Over the past few years, CMEM and MEET have been integrated into various routing models, with a focus on environmental impacts in addition to economic implications.

The CMEM is designed for light, medium and heavy duty vehicles. It computes fuel consumption of route plans considering the traveled distance, vehicle speed, carried load and roadway gradient. Relevant studies on green vehicle routing calculating the amount of GHG emissions following CMEM are those of Bektaş and Laporte [4], Demir et al. [8], Franceschetti et al. [14], and Heni et al. [23] in which the objective is to minimize a function comprising fuel consumption and driver costs. Pathak et al. [38] used CMEM to estimate fuel consumption under real-world driving patterns. Androutsopoulos and Zografos [1] and Huang et al. [29] integrated path selection decision on the vehicle routing problems considering a multigraph representation [19, 43] for the road network that incorporates the set of candidates paths between all pairs of key-destinations.

Figliozzi [13], Jabali et al. [30], Qian and Eglese [39], and Ehmke et al. [12] derived emissions from the MEET model, which allows the conversion of speeds into emissions based on fuel consumption rates that have been derived from engine test-bed measurements. MEET considers the impact of load and roadway gradient through error-corrective parameters.

Real-time traffic congestion information, the behavior of drivers, timely fuel consumption, and GHG emissions data collected by various sensors, and Global Positioning System (GPS) devices are becoming more present in commercial operations [24]. With such rich amount of traffic-related data much attention is now accorded to the computation of emission-minimizing paths on very large road networks based on time-dependent speed observations provided by logistics companies using Intelligent Transportation System (ITS) technology [5]. However, different fuel consumption models exist and they are based on very distinct assumptions and yield contrasting results. Making an accurate estimation of fuel consumption is an important aspect of a firm's decision-making process as realized emissions and fuel price affect the profitability [10].

Demir et al. [7] elaborated a comparative analysis of several fuel consumption models that have been developed to compute fuel consumption associated with road freight transportation. Fuel consumption models vary in their performance according to numerous factors such as speed, acceleration, and vehicle types. Turkensteen [47] evaluated the accuracy of CMEM, indicating that fuel consumption computations assuming fixed speed are not accurate in time-dependent routing. The author observed that fixed average speed computations are likely to underestimate fuel consumption. Through

sensitivity analysis he showed that much fuel is consumed when speed fluctuates and vehicle load increases. It should be noted that the CMEM was elaborated based on a bottom-up and analytical approach according to the operated vehicle type, engine efficiency, weather and route conditions [2]. Consequently, it requires the calibration of many parameters using statistical techniques.

Jaikumar et al. [31] performed a modal analysis of vehicular emissions under real-world driving conditions. They found out that short term events such as acceleration and braking significantly affect emissions. Despite their findings that CMEM underestimates emissions they have only used average speed and acceleration for distances ranging from 1 to 10 km based on field data obtained from an on-board diagnostic tool.

It follows from previous studies that approaches based on aggregated speeds can underestimate GHG emissions and fuel consumption. Greater estimation accuracy relies on data reflecting real-world operations. The last decade has seen substantial advances in building prediction models using machine learning methods, which capture complex nonlinear relationships in the systems under study and produce accurate estimations by learning from the available data [36, 6, 32, 34]. There have been a few studies on the application of machine learning methods for establishing practical fuel consumption or emissions models that can be used in routing problems with both environmental and operational considerations. Inspired by the need of emissions estimation, Zeng et al. [50] proposed a new emission model derived from the theory of vehicle dynamics. The parameters of their model were computed with the maximum likelihood estimation (MLE), and its accuracy was validated using GPS data collected for a light duty passenger car through a comparative analysis with the Virginia Tech Microscopic Energy and Emission Model (VT-Micro) [40], Support Vector Machines (SVM) model and Neural Networks (NNET) model. Liu et al. [35] proposed an effective emissions estimation model of a diesel engine using SVM that can be used by diesel engine manufacturers to measure emissions.

Due the growing interest of accurate fuel consumption estimations, the current study follows previous streams of literature by applying Gradient Boosting Machines (GBM) in addition to NNET, SVM and Conditional Inference Trees (CIT) machine learning methods to predict consumption considering relevant variables derived from in-field consumption data considering real-world driving conditions. We also perform a detailed calibration of CMEM parameters to improve its accuracy.

From a machine learning point of view, a number of opportunities may exist with the availability of speeds observations, instantaneous fuel consumption, roadway gradient, vehicle load, and stop-and-go traffic data. GPS and on-board real-time fuel consumption measurement devices provide real-world observations of consumption of micro scale events under real-world traffic conditions. In this work, we used field data collected across the entire road network of Québec City, which contains up to 50,000 road links. The obtained GPS dataset contains 58,215 instantaneous fuel consumption and speed observations monitored over 97 days between November 2016 and March 2017. In terms of prediction accuracy, families of supervised learning algorithms are shown to be effective in fitting artificial outputs to the real one. Therefore, using supervised learning methods we build nonlinear fuel consumption models considering speeds, vehicle load fluctuations, stop-and-go driving patterns, acceleration, and breaking events. The contributions of this paper are fivefold:

- (i) we propose an effective approach for the computation of fuel consumption;
- (ii) we provide several insights concerning fuel consumption through the analysis of real-world consumption data considering shipping operations under a large road network with fluctuating traffic congestion;
- (iii) we calibrate the CMEM parameters using the in-field fuel consumption data and large-scale Monte Carlo simulation;
- (iv) we develop efficient nonlinear fuel consumption models using NNET, SVM, CIT, and GBM supervised learning methods, which are trained by applying the k -fold cross validation method on real-life fuel consumption data;
- (v) we demonstrate the effectiveness of the proposed supervised learning models at micro scale events compared to MEET and the calibrated CMEM model that incorrectly predicted fuel consumption.

The remainder of this paper is organized as follows. In the following section we review the literature on fuel consumption models for road freight transportation. Section 3 describes the data collection procedure and provides some initial analysis of the available data. Section 4 presents the calibration of the CMEM with Monte Carlo method. In Section 5 we describe our proposed approach for modeling consumption using supervised learning methods. In Section 6, we present the results of our extensive computational experiments and sensitivity analysis of several existing and newly introduced models. Conclusions and directions for future research are stated in Section 7.

2 Existing estimation models

Motivated by the need to account for traffic congestion, this section describes the existing methods to compute emissions in time-dependent networks (multigraphs) using CMEM and MEET. We present the models in their original context, thus CMEM estimates fuel consumption and MEET estimates emissions. To do so, any path p from an origin o to a destination d follows an ordered sequence of nodes on the road network (see Figure 1):

$$p_{od} = (o = v_0, v_1, \dots, v_{k-1}, v_k = d), \quad (1)$$

where v_k are road nodes, and k represents the number of nodes of the path.

For any road segment $(u, v) \in \mathcal{A}$ let l_{uv} denote the distance between nodes u and v . Let τ_{uv} and \mathcal{F}_{uv} be the time-dependent travel time and the amount of fuel consumed, respectively, related to traveling across (u, v) . With each road segment (u, v) is associated a time-dependent travel speed s_{uv} .

2.1 Time-dependent fuel consumption using CMEM

The CMEM is one of the most used fuel consumption models in green vehicle routing. It was designed by Barth and Boriboonsomsin [2] for a wide variety of vehicles. According to this model, vehicle fuel consumption depends on many environmental and traffic-related parameters, namely load, speed, roadway gradient, among others. Considering

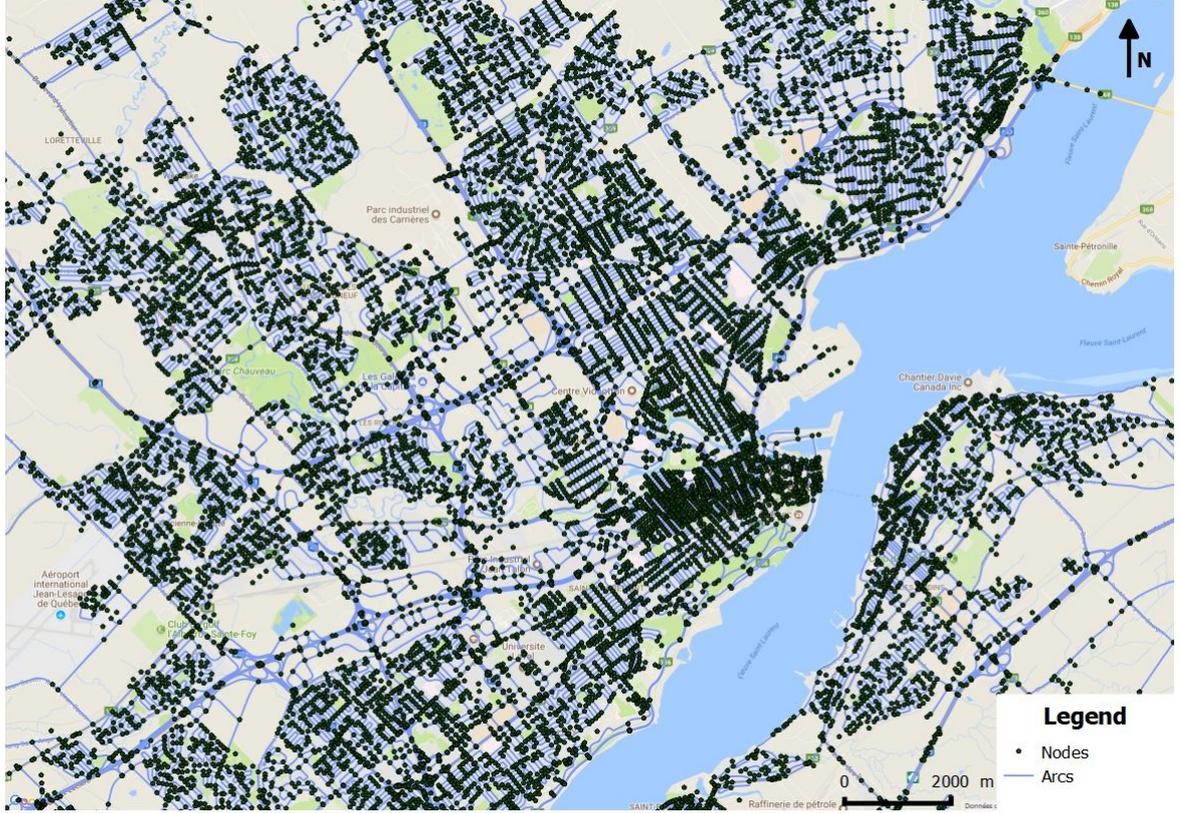


Figure 1: Illustration of a portion of the road network in Québec City

vehicle speed s (m/s), total vehicle weight \mathcal{M} and roadway gradient θ , for a given path p the corresponding fuel consumption (in liters) can be computed using CMEM based on equation (2):

$$\mathcal{F}_p = \sum_{(u,v) \in p} \mathcal{F}_{uv}^1 + \mathcal{F}_{uv}^2. \quad (2)$$

The term \mathcal{F}^1 describes the fuel consumption related to the vehicle weight and \mathcal{F}^2 represents the fuel consumption incurred by the speed:

$$\mathcal{F}_{uv}^1 = \tau_{uv} \frac{\alpha \mathcal{M} \mathcal{E}_0}{\mathcal{E}_2} s_{uv} = \frac{\alpha \mathcal{M} \mathcal{E}_0}{\mathcal{E}_2} l_{uv}, \quad (3)$$

and

$$\mathcal{F}_{uv}^2 = \tau_{uv} \mathcal{E}_0 \left(\mathcal{E}_1 + \frac{\beta}{\mathcal{E}_2} (s_{uv})^3 \right) = \mathcal{E}_0 \mathcal{E}_1 \tau_{uv} + \frac{\beta \mathcal{E}_0}{\mathcal{E}_2} l_{uv} (s_{uv})^2, \quad (4)$$

where $\mathcal{E}_0 = \frac{\zeta}{\omega \psi}$, $\mathcal{E}_1 = k N_e V$, $\mathcal{E}_2 = \frac{1}{\varepsilon 1000 \eta_{t,f}}$, $\mathcal{M} = \omega + q$, $\alpha = a + g \sin \theta + g C_r \cos \theta$, and $\beta = 0.5 C_d A \rho$ are constant parameters related to the vehicle and its engine such as inertia force, rolling resistance, and other vehicle characteristics. All parameter values used are shown in Table 1. Note that the values of CMEM parameters are determined either by manufacturer measurements or calibration process. On one hand, those related to the vehicle specifications are provided by the manufacturer. On the other hand, a

set of values of the parameters allowing calibration can be determined through Monte Carlo Simulation (see Section 4) using measured consumption data.

Table 1: Parameters used by CMEM for the computation of fuel consumption

Notation	Description	Typical values
w	Curb-weight (kg)	4500
q	Carried load (kg)	0-4350
ζ	Fuel-to-air mass ratio	1
k	Engine friction factor ($kJ/rev/liter$)	0.25
N_e	Engine speed (rev/s)	40
V	Engine displacement ($liter$)	5.12
g	Gravitational constant (m/s^2)	9.81
ρ	Air density (k/m^3)	1.2041
C_d	Coefficient of aerodynamic drag	0.7
A	Frontal surface area (m^2)	4.6
C_r	Coefficient of rolling resistance	0.01
η_f	Vehicle drivetrain efficiency	0.4
η	Efficiency parameter for diesel engines	0.9
c_f	Fuel and GHG emissions cost per liter ($\$/CAD/liter$)	1.15
ϖ	Heating value of a typical diesel fuel (kJ/g)	44
ψ	Conversion factor (g/s to $liter/s$)	737
s^l	Lower speed limit (m/s)	5.555
s^u	Upper speed limit (m/s)	22.222
s	Average speed at a portion of segment (m/s)	
a	Acceleration (m/s^2)	$[-3, 1]$
θ	Roadway gradient (degree)	0

2.2 Time-dependent emission using MEET

The MEET emission model was developed by Hickman et al. [25] for estimating vehicle emissions using a variety of polynomial functions of speed and acceleration levels. It computes GHG emissions produced by a vehicle of weights ranging from 3.5 to 32 tons according to travel speed and a wide range of input parameters related to the type of vehicle. Given an unloaded vehicle traveling at speed s (km/h) on a flat surface the MEET calculates the rate of emissions (g/km) using the following function:

$$\eta_r = \mathcal{K} + as + bs^2 + cs^3 + d\frac{1}{s} + e\frac{1}{s^2} + f\frac{1}{s^3}. \quad (5)$$

The coefficients ($\mathcal{K}, a, b, c, d, e, f$) are defined based on the vehicle type and weights. For example, if we consider the case of a vehicle weighing 3.5-7.5 tons the coefficients for the GHG emissions function for this specific vehicle category are $(\mathcal{K}, a, b, c, d, e, f) = (110, 0, 0, 0.000375, 8702, 0, 0)$.

To consider the effect of road gradient for each vehicle category, pollutant and gradient class, MEET proposes the following road gradient correction factor:

$$\eta_g = \mathcal{A}_0 + \mathcal{A}_1s + \mathcal{A}_2s^2 + \mathcal{A}_3s^3 + \mathcal{A}_4s^4 + \mathcal{A}_5s^5 + \mathcal{A}_6s^6, \quad (6)$$

where $(\mathcal{A}_0 \dots \mathcal{A}_6)$ are coefficients for CO_2 pollutant that vary according to the vehicle category and gradient class. Moreover, to consider the load, MEET applies the following load correction factor:

$$\eta_l = \kappa + n\theta + \rho\theta^2 + q\theta^3 + rs + ys^2 + zs^3 + \frac{\varrho}{s}, \quad (7)$$

where θ is the roadway gradient and $(\kappa, n, \rho, q, r, y, z, \varrho)$ are coefficients of the load correction function.

Based on MEET the amount of GHG emissions \mathcal{E}_p in grams produced by traversing path p (with time-varying speeds taken into account) is given by:

$$\mathcal{E}_p = \sum_{(u,v) \in p} \eta_r \eta_g \eta_t l_{uv}. \quad (8)$$

As GHG emissions are directly proportional to fuel consumption, the amount of fuel consumed can be derived from the amount of emissions using standard conversion factors.

3 Data collection and analysis of fuel consumption

In collaboration with an important furniture, appliances and electronics retailer from Québec City, on-road fuel consumption data collection was conducted with Toyota HINO SERIES 195 light duty vehicles [26, 27] across different time periods of each workday during shipping operations, which also covers rush hours times. The vehicles were monitored with a GPS and data logging device, which can measure the instantaneous fuel consumption between GPS points. The device incorporates a fuel analyzer sensor, an engine scanning tool, and a communication port for obtaining accurate measurements.

During 97 days between November 2016 and March 2017 up to 58,215 instantaneous speed and fuel consumption observations were collected. Real-time information includes fuel consumption, travel speed, acceleration, deceleration, GPS coordinates, and vehicle load. The average travel time between two consecutive measurements is 14.54 seconds.

We now present the characteristics and analysis of real-world, on-road vehicle fuel consumption. The main goal is to quantify and characterize the fuel consumption in a real-world road freight distribution environment regarding relevant input variables. For data validation the daily observed fuel consumption was compared to fuel invoices showing that the consumption device yields perfect accuracy. Yet, outliers analysis of fuel consumption data was done to ensure that there is no time lag between instantaneous consumption observations. Hence, this section describes how data analysis were offset to cleanup any lags in the fuel consumption sample that will be used by our machine learning algorithms.

For each observed workday the vehicles travel on average through 14 paths corresponding to shipping trips. Translating journeys into trips involves three main steps:

- geomatics and geospatial manipulations by geomatic specialists allow us to match GPS coordinates of each trip to the road segments of Québec City. It is therefore essential to combine road segments that form part of a single trip but which have been divided into individual paths according to service time at customers, refueling stops, and/or driver breaks.
- identification of whether a pause is a refueling stop or driver break, or service time at a customer where goods are picked up or dropped off, by grouping the observations according to when the vehicle ignition is turned on or off. For the

purpose of this study, a trip is defined as a combination of paths traveled across a given workday where the ends are the real location of a pickup or delivery, thus grouping subsequent journeys that include breaks at fuel stations or truck-stops.

- matching the information of GPS points, starting time and idle time with orders details from another database to identify the vehicle load at each GPS point, which is constant throughout the path connecting two customers.

A cleanup process is applied on the prepared data to remove GPS observations corresponding to engine off during breaks or delivery operations. Then, based on the obtained consumption sample composed of 46,476 observations we define five explanatory variables: travel speed, acceleration, vehicle load, stop-and-go driving pattern, and traveled distance, while the output variable is the amount of fuel consumed between two GPS points.

The frequency of link-based fuel consumption observations is displayed in Figure 2. We see that the number of observations is high for low fuel volumes. As shown in Figure 2 the mean of fuel consumption considering all observations is 0.035 liter per road link.

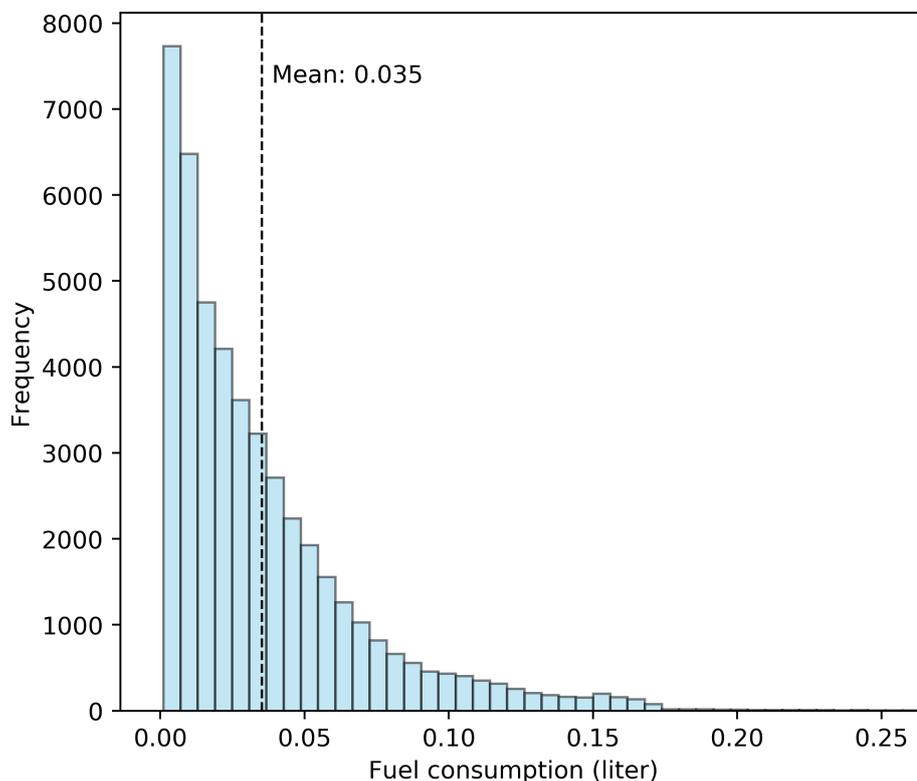


Figure 2: Fuel consumption histogram of real-world shipping trips in Québec City

A subset of data composed by observations corresponding to intervals of 11 seconds is presented in Figure 3. We see a high level of fuel consumption variability based on speed and acceleration levels. It also illustrates the nonlinear behavior of fuel consumption as a function of travel speed and acceleration. When acceleration and speed levels increase, consumption tends to increase. In deceleration, consumption values are generally low.

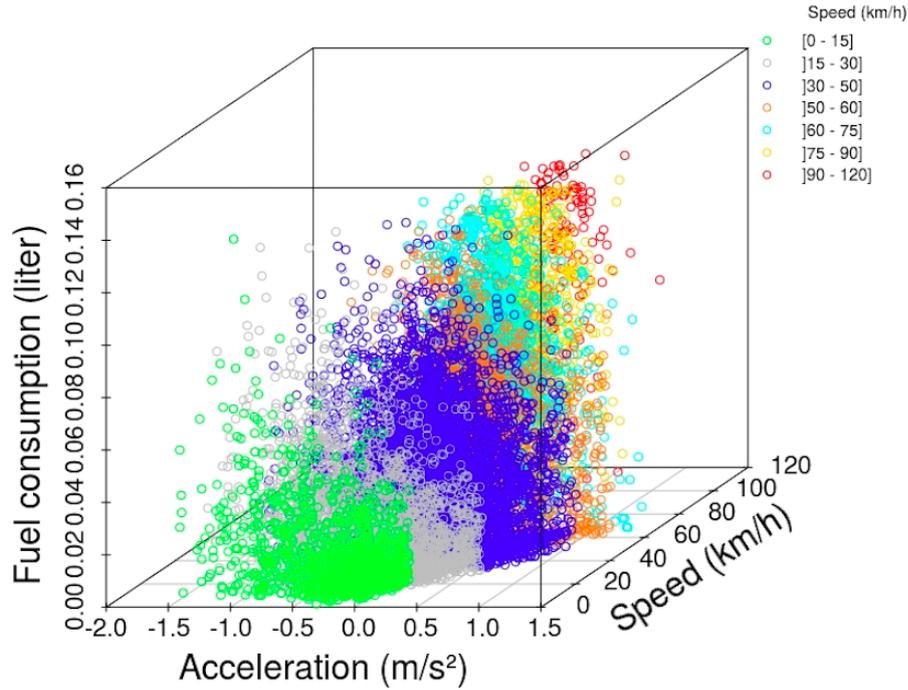


Figure 3: Fuel consumption as a function of instantaneous speed and acceleration

Figure 4 shows the trade-off between fuel consumption and travel speed over different times of a typical shipping workday. It is clear that the fluctuation of fuel consumption is directly impacted by the vehicle speed. The shape of curves has two distinct phases (or its increase). In a first stage, we observe that fuel consumption increases with speed. This phase is characterized by a regular form of speed (ascending or descending). The second phase, marked by erratic fluctuations of speed, gives a very accidental relationship between speed and fuel consumption. This situation corresponds to the different phases of acceleration and deceleration. We can see that vehicular consumption during idling and cruising are generally low compared to consumption during acceleration. We also observe that fuel consumption depends on very short term events such as rapid acceleration and braking (stop-and-go). The majority of microscopic fuel consumption models assume a constant consumption rate when a vehicle is decelerating.

To summarize, there is a clear need to perform an effective predictive modeling taking into account the specificity of fuel consumption data structures. Therefore, in this study we used model-based machine learning to estimate it.

4 Calibration of CMEM using Monte Carlo simulation

A key challenge when dealing with CMEM is the parameter setting which may impact the accuracy of the estimation of fuel consumption. In fact, even with same vehicle models, the efficiency of engines vary according to the traffic and weather conditions that differ from one place to another (Québec City in this study vs Arizona/California where the CMEM was designed) [42, 2, 31]. Therefore, upon making the comparison

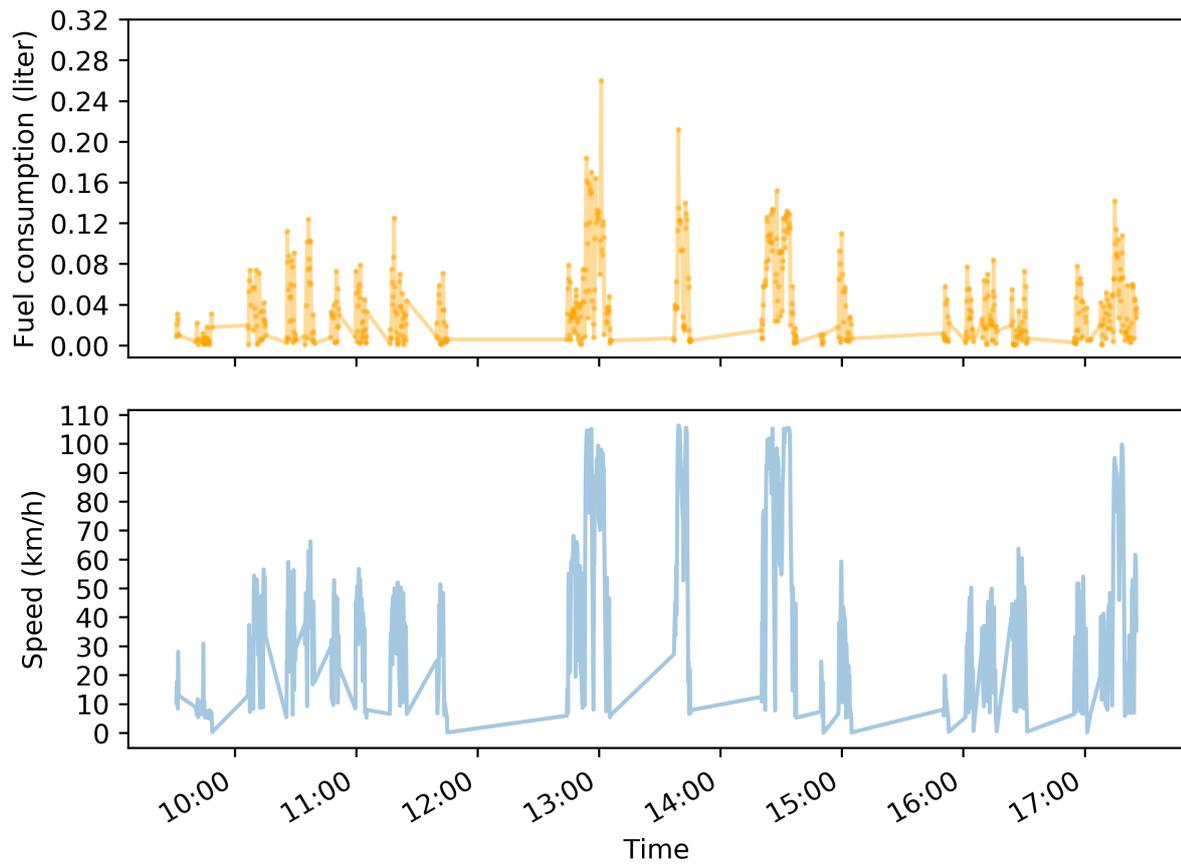


Figure 4: Instantaneous variation in fuel consumption and speed

against machine learning models, the application of statistical and simulation techniques to calibrate and optimize the default parameter setting of CMEM is crucial. To this end, we have addressed the issue of the parameterization and calibration of the CMEM model using Monte Carlo simulation (MCS), which is often used to generate effective simulations. Widely used in the sciences [20, 17] and operations research [18, 37], MCS allows the estimation of functions through stochastic search and approximation methods [41].

Vehicle and operation variables (e.g., frontal surface area, speed, acceleration, roadway gradient) and physical parameters (e.g., cold start coefficients, fuel to air mass ratio) are fixed inputs to the CMEM model and not included in the calibration process. Hence, to conduct MCS we first consider the set of calibration parameters, namely the engine friction factor, the coefficient of aerodynamic drag, the coefficient of rolling resistance, the vehicle drive train efficiency, the efficiency parameter for diesel engines, the heating value of a typical diesel fuel, and the conversion factor. Then, when running MCS we calibrate the selected CMEM parameters to minimize the modeled and measured differences through the evaluation of the average prediction error measures. Note that the values of the set of parameters requiring calibration are randomly selected with respect to search intervals and probability distributions that allow to realistically represent the uncertainty of the variables. More specifically, any candidate CMEM parameter that needs to be calibrated is optimized according to the prediction accuracy on the same training sample used to generate machine learning models (see Section 5). The composition of the training and testing samples is presented in Section 6.

Moreover, using MCS we execute a repeated sampling of data from a set of populations and a randomization process to produce a large set of combinations. MCS combines a huge set of possibilities (more than 1.5 million combinations in our study), ranging the values of the calibration parameters at the same time. Datasets are selected from this population and fed into the MCS procedure, thus providing an efficient way of producing a sufficiently large number of assessments to enable a statistically valid appraisal of the calibration process. The simulation process is repeated many times to enhance the accuracy of fuel consumption estimations. More specifically, the MCS calculates the model outcome thousands of times, each time using different randomly selected samples.

Table 2 provides the retained values of the different parameters after running the calibration with MCS.

Table 2: Calibrated parameters of CMEM obtained with Monte Carlo simulation

Notation	Description	Typical values	Calibrated values
k	Engine friction factor ($kJ/rev/liter$)	0.25	0.17
C_d	Coefficient of aerodynamic drag	0.70	0.90
C_r	Coefficient of rolling resistance	0.01	0.046
η_{tf}	Vehicle drivetrain efficiency	0.40	0.66
η	Efficiency parameter for diesel engines	0.90	1.00
ϖ	Heating value of a typical diesel fuel (kJ/g)	44.00	45.03
ψ	Conversion factor (g/s to $liter/s$)	737.00	741.46

5 Fuel consumption modeling with supervised learning methods

This section shows the development of multiple nonlinear fuel consumption models using supervised learning based on the applied econometric approach of Mullainathan and Spiess [36] guiding our design choices to:

1. pick a variety of machine learning methods according to prediction function classes, namely nonparametric predictors (e.g., Conditional Inference Trees), mixed predictors (e.g., Neural Networks), and combined predictors (e.g., Gradient Boosting Machines);
2. perform regularization based on the prediction function class;
3. and empirical tuning through cross-validation.

In this study four supervised learning methods were selected: Neural Networks (NNET), Support Vector Machines (SVM), Conditional Inference Trees (CIT) and Gradient Boosting Machines (GBM). Each model-based machine learning uses a set of regularization parameters. These determine the performance profile of each model. To choose the appropriate combination of parameters values while avoiding over-fitting we used grid search method for SVM and CIT and trial-and-error approach for NNET and GBM. For each model we define a set of candidate values for the appropriate regularization parameters according to the relevant literature, sample size and computational resources. We then perform empirical tuning to fit each model with each candidate set using the training dataset on which we apply the k -fold cross validation method [21, 49] for estimating prediction error. The k -fold cross validation works by splitting the training dataset into k roughly equal-sized subsamples or folds. Each supervised learning method performs k iterations and at each time it excludes one held-out fold in turn to evaluate their prediction accuracy once the model is estimated using the remaining $k - 1$ folds. There is no formal rule of defining the value of k , and we used $k=10$. The prediction accuracy of each model is given by the average of k obtained prediction error measures. For each candidate machine learning model, the optimal settings of tuning parameters is determined according to the obtained performance metrics. Then, we evaluate the performance of their accuracy prediction using a testing dataset (see Section 6).

5.1 Neural Networks

NNET learning methods allow the extraction of linear combinations of the inputs to produce a nonlinear fuel consumption model. NNET is composed of a set of neurons connected together [11] and uses massive interconnections to fit nonlinear models to multidimensional data [22]. Figure 5 shows a schematic diagram of the proposed NNET used to model fuel consumption. In this diagram the nodes are the neurons and the arcs are the connections. NNET is a multi-layer network composed of three layers: input layer, hidden layer and output layer. The input layer incorporates five input variables x_1, \dots, x_5 defined based on the chosen parameters affecting emissions, namely

speed, acceleration, vehicle load, stop-and-go driving patterns, and distance. Note that the driving pattern was defined as a binary variable indicating whether the vehicle is moving or not (e.g., vehicle is stopped at red light or after breaking or when waiting at customer, etc.) while the engine is running. The hidden layer incorporates a set of hidden units or unobserved variables used to model the outcome [33]. These hidden units perform intermediate computations using linear combinations of the input variables. The output layer is the combination of obtained hidden units to perform the estimation of fuel consumption.

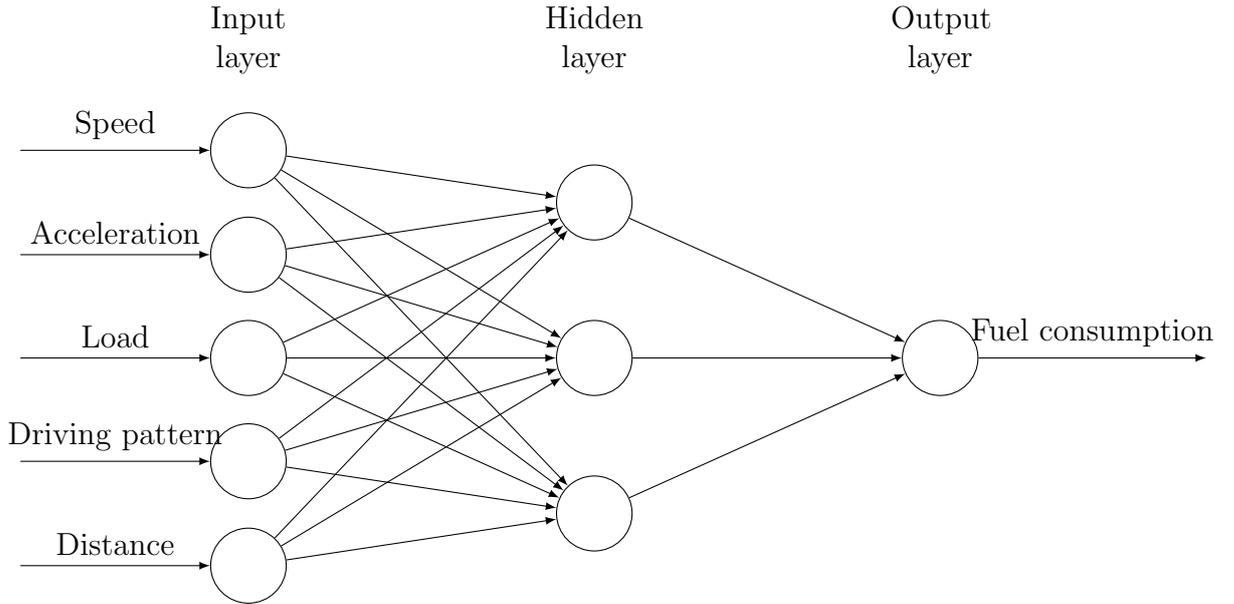


Figure 5: Schematic diagram of the NNET fuel consumption model

In this study, several NNET tasks were performed to accurately predict consumption by studying field data. We applied the quasi-Newton back propagation learning algorithm [3]. The linear combinations of the predictors are transformed by a nonlinear activation function (sigmoidal). To reduce over-fitting our NNET algorithm minimizes the following function [33]:

$$\mathcal{G} = \sum_{i=1}^{\mathcal{N}} (y_i - f_i(x))^2 + \eta \left(\sum_{k=1}^{\mathcal{H}} \sum_{j=0}^{\mathcal{P}} \beta_{jk}^2 + \sum_{k=0}^{\mathcal{H}} \gamma_k^2 \right), \quad (9)$$

where \mathcal{N} is the total number of observations, \mathcal{P} is the number of predictors, \mathcal{H} represents the number of hidden units, η is the weight decay, and y_i is the outcome. The coefficient β_{jk} represents the effect of the j th predictor on the k th hidden unit. The function f defines a linear combination that connects the hidden units to the outcome:

$$f(x) = \gamma_0 + \sum_{k=1}^{\mathcal{H}} \gamma_k h_k, \quad (10)$$

where γ_k are the regression coefficients of hidden layers.

Several combinations of NNET parameter values were investigated by trial-and-error to identify the best learning performance. Four different weight decay $\eta \in \{15^{-4}, 15^{-3}, 15^{-2}, 15^{-1}\}$ were evaluated along with one hidden layer including between one and 10 hidden units. The convergence to the best NNET model is achieved with a maximum number of iterations equal to 2000. The optimal NNET model used $\eta = 15^{-3}$ and $\mathcal{H} = 9$ hidden units.

5.2 Support Vector Machines

SVMs is a supervised learning method applied for classification and nonlinear regression [48]. SVM algorithms use a kernel function allowing this model to transform input data to the required form of relationships. There are multiple kinds of kernel functions, such as linear, polynomial, radial basis function, and sigmoid. After several trials, we used a linear kernel function defined as a simple sum of cross products, which have shown to be effective for the current study.

The SVM regression minimizes the following regularized function:

$$\mathcal{W} = \mathcal{C} \sum_{i=1}^{\mathcal{N}} \mathcal{L}(y_i, \mathcal{F}(x_i)) + \sum_{j=1}^{\mathcal{P}} \beta_j^2, \quad (11)$$

where x_i is the input space-vector, $\mathcal{L}(\cdot)$ is the loss function, β are coefficients used by the regularization term considering \mathcal{P} predictors, and constant \mathcal{C} is the error penalty factor for adjusting the complexity of the model [33]. \mathcal{F} is a prediction function defined as follows:

$$\mathcal{F}(x) = \sum_{i=1}^{\mathcal{N}} \alpha_i \varphi(x) + \beta_0, \quad (12)$$

where $\varphi(x)$ is the linear kernel function.

The tuning of regularization parameter \mathcal{C} through grid search method produced a constant with a value of 1.

5.3 Conditional Inference Trees

CIT is a machine learning method that uses unbiased tree-based models for regression and classification [28]. CIT algorithm's estimates regression relationship using a binary recursive partitioning method, which efficiently performs the exhaustive search across the predictors according to split points. A simplified description of this method is provided by the following steps:

1. perform the null hypothesis test of independence between each input variable and the outcome one. The algorithm continues until the hypothesis cannot be rejected;
2. apply a binary split to the selected input variable;
3. recursively repeat steps 1 and 2.

The p -value statistical test is applied for candidate splits by evaluating the difference between the means of two groups. On our tests with the training dataset we found that the optimal CIT model is obtained with a value of $1 - p$ equal to 0.821.

5.4 Gradient Boosting Machines

GBM is gaining considerable interest on a wide range of data driven applications such as travel time prediction [51] and the modeling of the energy consumption [44]. It is a highly adaptable supervised learning method encompassing both classification and regression in order to find an additive model that minimizes the loss function [15]. GBM iteratively investigates decision trees (basic learner) to reduce the loss function and improve prediction accuracy. The GBM model is defined as follows [15].

Let $\hat{\mathcal{R}}(x)$ be the regression function that minimizes the expectation of loss function $\mathcal{S}(y, \mathcal{R})$ over the joint distribution:

$$\hat{\mathcal{R}}(x) = \arg \min_{\mathcal{R}(x)} E_{x,y}[\mathcal{S}(y, \mathcal{R}(x))], \quad (13)$$

where $\mathcal{R}(x)$ can be formulated as a function with a finite number of parameters β estimated by selecting those values that minimize the loss function \mathcal{S} using the training sample as shown in equation 14:

$$\hat{\mathcal{R}}(x) = \arg \min_{\beta} \sum_{i=1}^{\mathcal{X}} \mathcal{S}(y_i, \beta). \quad (14)$$

To optimize the GBM model we have performed the tuning of several regularization parameters:

- d : the depth of decision trees that controls the maximum interaction order of the model;
- I : the number of boosting iterations, which also corresponds to the numbers of decision trees;
- α : the learning rate that controls the contribution of each base model or decision trees by shrinking its contribution by a factor between 0 and 1;
- δ : the subsampling rate or fraction of the training set observations, which is randomly selected to propose the next tree in the expansion.

After the training of the model, the depth of the decision trees d was selected in the set $\{2, 5, 7, 9\}$, the learning rate α was chosen from 0.01 to 0.5 with a granularity of 0.02. The number of iterations I was selected within a set spanning from 50 to 250 iterations with a granularity of 50 iterations. The minimum number of observations in trees terminal nodes φ was defined between 5 and 10. The subsampling rate δ was fixed to 0.5. The final combination of values used for the GBM model were $d = 9$, $I = 250$, and $\alpha = 0.07$.

6 Numerical experiments

It is not recommended to use the same set of observations for both training and testing [36, 32]. Hence, in this work the assessment of predictive performance has been carried out on an independent sample of field data in order to avoid over-fitting, which is the tendency of the models to fit the training sample too well, at the expense of the predictive accuracy. The preprocessed field dataset composed of 46,476 fuel consumption observations (1406 paths) was randomly split in two subsets using days as the splitting criterion:

- training sample: composed by 80% of days corresponding to 38,004 observations (1263 paths);
- testing sample: composed by 20% of days including 8,472 observations (143 paths).

Each model was trained with the same training dataset with R version 3.4.3 through R-Studio 1.0.153 using a ThinkCenter professional workstation with Intel i7 vPro (8 cores) and 32-gigabyte RAM, running Ubuntu Linux 16.04 LTS x86 operating system. The following R machine learning packages were used to generate nonlinear emission models: nnet, e1071, party, gbm, caret, and h2o. The Monte Carlo simulation (MCS) was executed using a SparkR cluster and sparklyr R package to enable distributed computations. The evaluation process was initiated by comparing the models fuel consumption prediction outcomes on in-field observations. Note that the outcome of the studied models (machine learning, CMEM and CMEM calibrated with MCS) is an estimate of the fuel consumption, while the MEET computes emissions requiring the use of a conversion coefficient (2.757 Kg of CO₂/liter for the case of Toyota HINO SERIES 195 light duty according to Transition Énergetique Québec [45]). The obtained models were then evaluated on the testing sample. Their effectiveness was validated by computing and analyzing the following accuracy measures:

- Root Mean Squared Error (RMSE): interpreted as the average distance between the observed values and the model predictions. The RMSE is then computed by taking the square root of the Mean Squared Error (MSE). The smaller the values of RMSE, the closer the predicted values are to the observed ones. The RMSE is computed by squaring the residuals, summing them up and dividing by the number of observations, and taking the square root of the result, as $\frac{1}{n} \sum_i^n (y_i - \hat{y}_i)^2$, where y_i is the observed value, \hat{y}_i is the predicted output, and n is the total number of observations;
- Mean Absolute Error (MAE): is the average magnitude of the errors in a set of predictions. It is computed as $\frac{1}{n} \sum_i^n |y_i - \hat{y}_i|$.
- Mean: is the average of the corresponding predicted outcomes. It is calculated as $\mu = \frac{1}{n} \sum_i^n \hat{y}_i$.
- Std Error: is the standard error of the mean. It is computed as $\frac{\sigma}{\sqrt{n}}$, where $\sigma = \sqrt{\frac{\sum_i^n (y_i - \mu)^2}{n-1}}$ is the standard deviation of the mean μ .

- Gap (%): reports how close the corresponding predicted outcome is to the observed value. The percentage gap values are calculated as $100(y_i - \hat{y}_i)/y_i$.

6.1 Experimental results and analysis

In this section we provide the experimental results and analysis. Table 3 shows the accuracy metrics of CMEM, CMEM calibrated with MCS (CMEM-MCS), MEET, NNET, SVM, CIT and GBM predictions. In this table, successive columns give for each consumption model the RMSE, the MAE, the Std Error, the mean value, the Gap (%) aggregated across all paths (trips) in the testing dataset, and the computational time (CPU) of training or simulation (minutes). Doing so, we estimate the fuel consumption for each road segment, then we aggregate the obtained values for each path. The results obtained for the RMSE metric show that the proposed nonlinear models, namely GBM, NNET, CIT and SVM outperform CMEM and MEET and appear to be more accurate in estimating instantaneous vehicle fuel consumption across paths. In fact, we see that the average RMSE ranges from 0.258 to 0.315 for the machine learning models, which are lower than those of the CMEM and MEET models (0.501 and 0.850), and calibrated CMEM (0.419). More specifically, it can be clearly seen that GBM model exhibited the best estimation accuracy as the fuel consumption predictions are very consistent with in-field observations, with the lowest RMSE of 0.258. As can be seen, the machine learning models clearly outperform the CMEM even when calibrated with MCS.

Table 3: Comparative performance of the proposed machine learning models against MEET, CMEM and CMEM-MCS regarding fuel consumption estimations aggregated by paths

Consumption models	RMSE	MAE	Std Error	Mean	Gap (%)	CPU* (<i>min</i>)
Real-world	0	0	0	1.539	0	–
CMEM	0.501	0.305	0.459	1.742	–13.184	–
CMEM-MCS	0.419	0.239	0.403	1.656	–7.566	11520
MEET	0.850	0.404	0.760	1.155	24.942	–
SVM	0.315	0.170	0.301	1.444	6.173	9.05
CIT	0.264	0.151	0.263	1.561	–1.453	6.74
NNET	0.271	0.155	0.270	1.569	–1.930	0.53
GBM	0.258	0.150	0.257	1.565	–1.709	0.28

*using a cluster of machines in parallel for the MCS procedure.

Table 3 shows that GBM outperforms MEET, CMEM and CMEM-MCS that were found to under- and over-predict fuel consumption. Regarding the obtained percentage gap values, we observe that the machine learning models give the best prediction results with a gap ranging from -1.930% to 6.173% when compared against MEET underestimating fuel consumption on average by 24.942% , CMEM overestimating fuel consumption by 13.184% , and CMEM-MCS overestimating fuel consumption by 7.566% . We also see that GBM and CIT yield the lowest underestimation with a gap of -1.709% and -1.453% , respectively.

We observe that the estimation accuracy of CMEM can be enhanced with the MCS reducing the RMSE from 0.501 to 0.419, and the gap from -13.184% to -7.566% . We conclude from these results that the calibrated CMEM is more effective at estimating consumption. Note that the trade-off between computation time and prediction accuracy is a key aspect when calibrating the CMEM model with MCS. It is important to highlight that the MCS is computationally demanding as it took over 8 days of execution time to calibrate CMEM which is not the case of machine learning methods that are computationally faster taking only few seconds or minutes to generate fuel consumption estimation models. Clearly, the machine learning models are more effective for real-time forecasting and decision-making considering high-frequency data and large-scale transport applications, as an efficient supervised learning achieve a better prediction accuracy.

Additional experiments are conducted to study the performance of the developed machine learning models. Figure 6 shows scatter plots that graphically illustrate the prediction accuracy of the studied models superimposed on the field data. On the vertical scale, the observed value of fuel consumption is displayed, whereas the predicted values are presented on the horizontal scale. We observe that NNET, GBM, SVM and CIT models fit similarly as their prediction outcomes are more concentrated and closer to the identity lines represented by a solid line indicating that the observed and predicted emission values are very close. This implies that the machine learning models yield more effective prediction of fuel consumption than those produced by the classical CMEM and MEET. As expected, machine learning models provide good fitting regarding observed fuel consumption as they are able to better reflect differences that result from traveling on congested areas with frequent stop-and-go events impacting the speed.

Figure 6 also illustrates the difference observed between the identity solid line and the dotted regression line, which shows the variation in prediction between each model results compared to observed data. Notably, this graphical trend was validated by the goodness of fit test. The null hypothesis of this test is performed with a slope=1 and intercept=0. This test leads to the rejection of the null hypothesis with very low p -value ($<2.2e-16$), lower than the threshold 0.05. Hence, the three classical models are not preferred candidates for predicting fuel consumption. More specifically, the best prediction accuracy belongs to the GBM model yielding the lowest p -value of 0.314, which is larger than the threshold 0.05. Therefore, the null hypothesis is not rejected which indicates that the prediction of the GBM model is statistically the same as the observed values.

In order to make further analysis on the prediction accuracy of the proposed models, the boxplots presented in Figure 7 illustrate numerical outcomes of the studied consumption models through their quartiles. When looking at the boxplots of CMEM, CMEM-MCS and MEET, we can see a difference between the medians represented by the lines in the middle of the boxes, indicating that these models tend to incorrectly predict fuel consumption. Clearly, the median of GBM, CIT and NNET models seem to be very close to the observed fuel consumption one, exhibiting superior accuracy regarding fuel consumption. Also, when looking at Figure 8 we observe that even with the calibrated econometric structure, the CMEM-MCS (dashed line) overestimates the real world fuel consumption, which is not the case of GBM and NNET that perform much better.

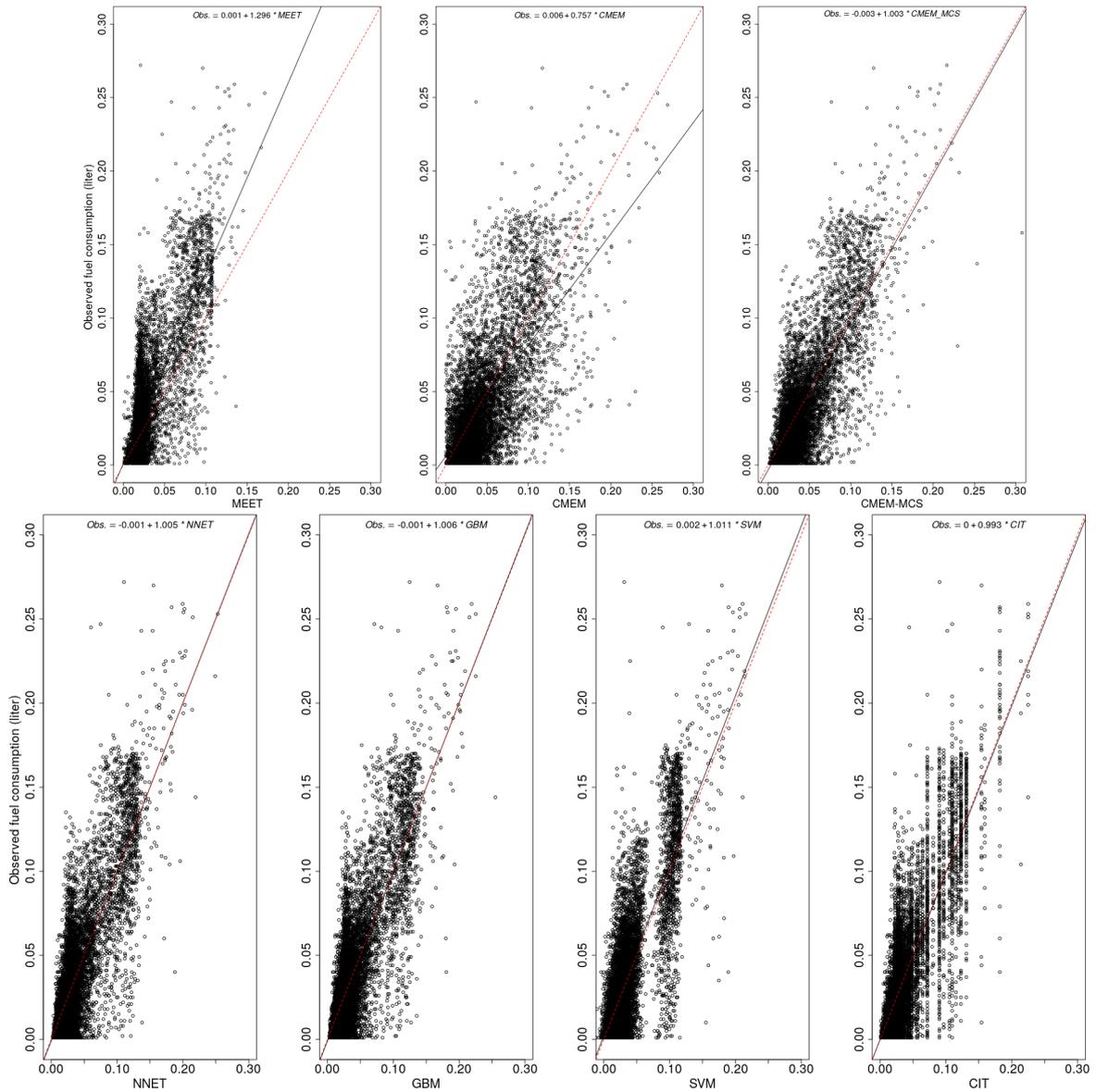


Figure 6: Scatter plots of predicted outcomes by CMEM, MEET and machine learning models against observed fuel consumption

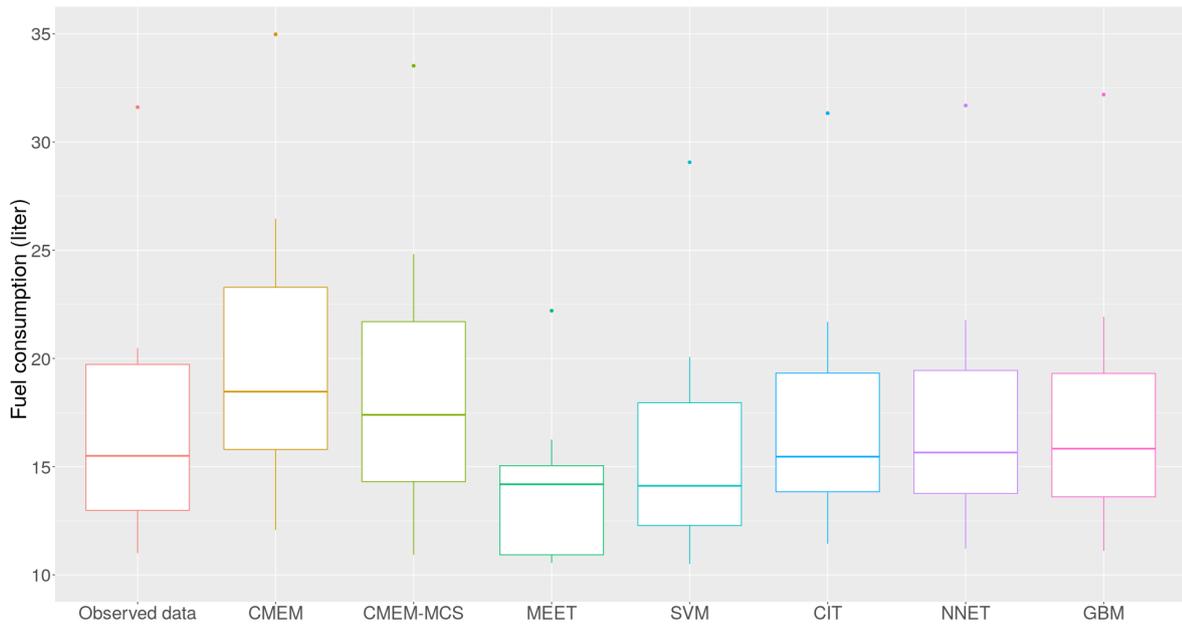


Figure 7: Boxplots of fuel consumption models estimation performance against observed fuel consumption aggregated by days



Figure 8: Sample of the consumptions produced by CMEM-MCS, NNET and GBM models against real-world observations.

To further evaluate the performance of the proposed consumption models, a sensitivity analysis is performed to compare their prediction accuracy under multiple criteria: congested (low speeds) and free flow (high speeds) situations, empty and loaded vehicle, stop-and-go driving patterns corresponding to acceleration/deceleration events, and different periods. In Table 4 the performance of the consumption estimation of CMEM, CMEM-MCS and MEET is compared against the proposed models with in-field measurements considering each criterion, which includes corresponding mean and gap for the best machine learning model, CMEM, CMEM-MCS, and MEET according to the selected subset of observations. Clearly, the degree of estimation varies for all criteria according to real-world driving conditions. We have noticed that the estimation of CMEM, CMEM-MCS and MEET are deteriorated in the case of low speeds with an overestimation of 107.032%, 56.790% and 11.800%, respectively. We see that the prediction of MEET, CMEM and CMEM-MCS are negatively affected in the case of low speed. However, GBM provides a low overestimation of only 3.776%. More specifically, it is remarkable that GBM and NNET produce accurate estimations under fluctuating speeds according to the variations of acceleration.

Regarding the driving pattern criterion, as expected machine learning models adequately handle acceleration variability as NNET has the smallest gap (-1.564%). Regarding loads, GBM gives the lowest gaps -1.666% for empty vehicles and -1.716% for loaded ones. Interestingly, the NNET model shows its performance under different periods providing a gap ranging from 1.107% to 2.783%.

Based on the results presented in Table 4 it can be argued that GBM and NNET models give the best results and are the most accurate for all aspects, exhibiting a small gap just ranging from -3.776% to 2.783%. Further, we can see that the overall performance of both models is very good not only in normal or moderate traffic conditions, but also during traffic congestion. Compared to MEET, CMEM and CMEM-MCS, machine learning models are less sensitive to these conditions and maintain superior prediction accuracy.

Predictor variables have very different influences on the output of the machine learning models. By investigating the influence of each input variable on the response, we can gain better insight. Doing so, one way to explain the outcomes of machine learning models is to compute variable importance in the fitted functions [33, 49, 36]. Figure 9 shows the partial least squares variable importance scores for the studied fuel consumption input variables considering each machine learning model. Following Kuhn and Johnson [33] we compute variable importance scores based on the difference between the class centroids to the overall centroid. Then, the relative importance measure for each variable is scaled so that they sum to 1. A larger variable importance score implies a higher influence of the input variable to the model. Note that variable importance for GBM is determined according to the reduction in squared error. Specifically, an improvement value is calculated for each pair of predictor and tree. Hence, the overall importance value is computed by averaging the improvement values for each predictor across the entire ensemble [15, 16].

It is worth noting that the variable importance scores for the set of models show that they tend to rely more on continuous predictors (speed and acceleration) rather than on discrete ones (distance and load). As shown in Figure 9 for all machine learning models the speed stands out to the top in terms of importance score followed by distance and

Table 4: Comparative performance statistics of the GBM, NNET, MEET, CMEM and CMEM-MCS models regarding multiple performance indicators

Criterion	Aspects	# Obs.	Mean real-world	Best model (RMSE)	Mean BM	Gap BM (%)	CMEM	Gap CMEM (%)	CMEM-MCS	Gap CMEM-MCS (%)	MEET	Gap MEET (%)
All observations	Testing dataset	8472	0.041	GBM	0.041	-1.709	0.046	-13.184	0.044	-7.566	0.031	24.942
	Low (0-15 km/h)	1832	0.017	GBM	0.017	-3.776	0.034	-107.032	0.026	-56.790	0.018	-11.800
	Moderate (~50 km/h)	175	0.034	GBM	0.035	-2.533	0.043	-28.210	0.041	-20.686	0.022	34.444
Speed	High (>75 km/h)	1573	0.101	NNET	0.101	0.529	0.090	11.252	0.093	8.195	0.074	27.160
	Empty vehicle	1377	0.037	GBM	0.038	-1.666	0.041	-11.414	0.036	2.113	0.028	23.101
Load	Loaded vehicle	7095	0.041	GBM	0.042	-1.716	0.047	-13.490	0.045	-9.237	0.031	25.26
	Acceleration	4703	0.044	NNET	0.044	-1.564	0.056	-28.222	0.049	-13.046	0.030	30.701
Driving pattern	Deceleration	3691	0.036	NNET	0.037	-2.081	0.033	9.002	0.036	0.438	0.031	16.255
	07h30 - 08h00	744	0.060	NNET	0.060	0.867	0.062	-2.671	0.062	-3.431	0.044	26.650
Period	08h00 - 08h30	776	0.050	NNET	0.049	1.107	0.051	-4.265	0.050	-1.013	0.036	26.630
	08h30 - 09h00	359	0.038	NNET	0.037	2.783	0.043	-13.490	0.041	-8.875	0.026	30.692

acceleration. However, the importance orders are much different between the machine learning models. Contrasting SVM importance results to CIT (Figure 9) we see that importance order are the same for all predictors (speed, distance, acceleration, driving-patterns, and load). However, we observe a different pattern for GBM and NNET (speed, distance, acceleration, load, and driving-patterns).

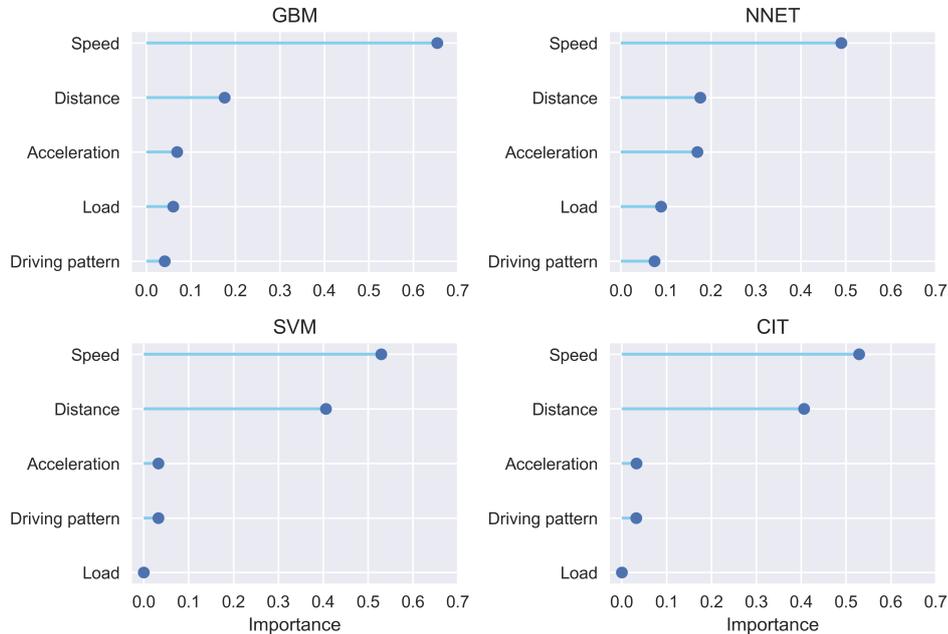


Figure 9: A comparison of partial least squares variable importance scores for the studied fuel consumption input variables for each supervised learning methods

We also see from Figure 9 that for the case of GBM and NNET the load variable has a higher influence on the model output compared to SVM and CIT, which ends up ranked 5th. Clearly, the load variable under GBM and NNET has a much steeper importance slope than under SVM and CIT. Further, the variable importance scores for the GBM show that it tends to rely more on speed and acceleration than the driving patterns. We also see that NNET captures effectively driving patterns leading to a significant increase in prediction performance. This suggests that there is significant value of considering stop-and-go events. In fact, the NNET model is able to capture this sudden change. The GBM model also adequately captures the instantaneous variations of speed. Therefore, the GBM model is able to model nonlinear characteristics of dynamic traffic systems leading to superior estimation accuracy.

To summarize, even if fuel consumption estimation is complex and challenging, it is clearly shown that machine learning models enhance fuel consumption estimation accuracy by taking into account the interactions among different combinations of input variables. In all experiments presented in this section, we conclude that the proposed machine learning models significantly outperform MEET, CMEM and CMEM-MCS. In fact, machine learning-based fuel consumption models, and in particular GBM and NNET models are able to fit complex nonlinear relationship leading to superior fuel consumption estimation accuracy.

7 Conclusions and future research

In this paper we have proposed nonlinear fuel consumption estimation models using supervised learning methods. The estimation accuracy was compared to the classical MEET and CMEM methods as well as CMEM calibrated with Monte Carlo simulation (CMEM-MCS). We used a very large database of real-world information from a logistic provider, which is shown to be very detailed and precise. In our numerical experiments, we have observed that MEET and CMEM incorrectly predicted consumption by 24.942% and -13.18% , respectively. Results revealed that the proposed NNET, SVM, CIT and GBM models outperform MEET, and CMEM as they clearly improve prediction accuracy. We have shown that GBM produces the best predictability which is off by only 1.70% according to real-world data. This indicates that we cannot take for granted that existing fuel consumption models are sufficiently accurate, requiring machine learning models that update them by applying supervised learning methods on collected real-time traffic data and on-road vehicular exhaust fuel consumption. Even with very large-scale MCS the calibrated CMEM still underestimates fuel consumption by 7.57%, performing worse than any of our machine learning models.

The results of this work show that using machine learning models and more specifically the GBM and NNET models enhance the prediction accuracy of fuel consumption estimations. A direction of future research is to evolve machine learning fuel consumption models by investigating the effects of weather, driver profiles and road-wide factors such as temperature, rain, snow, road maintenance events, etc. Another area of future work will be the integration of machine learning emission models in routing problems and practical road freight transportation applications.

References

- [1] Konstantinos N Androutsopoulos and Konstantinos G Zografos. An integrated modelling approach for the bicriterion vehicle routing and scheduling problem with environmental considerations. *Transportation Research Part C: Emerging Technologies*, 82:180–209, 2017.
- [2] Matthew Barth and Kanok Boriboonsomsin. Energy and emissions impacts of a freeway-based dynamic eco-driving system. *Transportation Research Part D: Transport and Environment*, 14(6):400–410, 2009.
- [3] Roberto Battiti. Optimization methods for back-propagation: Automatic parameter tuning and faster convergence. In *International Joint Conference on Neural Networks*, volume 1, pages 593–596, 1990.
- [4] Tolga Bektaş and Gilbert Laporte. The pollution-routing problem. *Transportation Research Part B: Methodological*, 45(8):1232–1250, 2011.
- [5] Khaled Belhassine, Leandro C Coelho, Jacques Renaud, and Jean-Philippe Gagliardi. Improved home deliveries in congested areas using geospatial technology. Technical Report CIRRELT-2018-02, Montreal, Canada, 2018.

- [6] Tsan-Ming Choi, Stein W Wallace, and Yulan Wang. Big data analytics in operations management. *Production and Operations Management*, 27(10):1868–1883, 2018.
- [7] Emrah Demir, Tolga Bektaş, and Gilbert Laporte. A comparative analysis of several vehicle emission models for road freight transportation. *Transportation Research Part D: Transport and Environment*, 16(5):347–357, 2011.
- [8] Emrah Demir, Tolga Bektaş, and Gilbert Laporte. An adaptive large neighborhood search heuristic for the pollution-routing problem. *European Journal of Operational Research*, 223(2):346–359, 2012.
- [9] Emrah Demir, Tolga Bektaş, and Gilbert Laporte. A review of recent research on green road freight transportation. *European Journal of Operational Research*, 237(3):775–793, 2014.
- [10] David F Drake, Paul R Kleindorfer, and Luk N Van Wassenhove. Technology choice and capacity portfolios under emissions regulation. *Production and Operations Management*, 25(6):1006–1025, 2016.
- [11] Gérard Dreyfus. *Neural Networks: Methodology and Applications*. Springer Science & Business Media, Berlin Heidelberg, Germany, 2005.
- [12] Jan Fabian Ehmke, Ann Melissa Campbell, and Barrett W Thomas. Data-driven approaches for emissions-minimized paths in urban areas. *Computers & Operations Research*, 67:34–47, 2016.
- [13] Miguel Andres Figliozzi. The impacts of congestion on time-definitive urban freight distribution networks CO_2 emission levels: Results from a case study in Portland, Oregon. *Transportation Research Part C: Emerging Technologies*, 19(5):766–778, 2011.
- [14] Anna Franceschetti, Emrah Demir, Dorothe Honhon, Tom Van Woensel, Gilbert Laporte, and Mark Stobbe. A metaheuristic for the time-dependent pollution-routing problem. *European Journal of Operational Research*, 259(3):972–991, 2017.
- [15] Jerome H Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29:1189–1232, 2001.
- [16] Jerome H Friedman. Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4):367–378, 2002.
- [17] Michael C Fu and Jian-Qiang Hu. Efficient design and sensitivity analysis of control charts using Monte Carlo simulation. *Management Science*, 45(3):395–413, 1999.
- [18] Michael C Fu, Shreevardhan Lele, and Thomas WM Vossen. Conditional Monte Carlo gradient estimation in economic design of control limits. *Production and Operations Management*, 18(1):60–77, 2009.

- [19] Thierry Garaix, Christian Artigues, Dominique Feillet, and Didier Josselin. Vehicle routing problems with alternative paths: An application to on-demand transportation. *European Journal of Operational Research*, 204(1):62–75, 2010.
- [20] Dwight Grant, Gautam Vora, and David Weeks. Path-dependent options: Extending the Monte Carlo simulation approach. *Management Science*, 43(11):1589–1602, 1997.
- [21] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, USA, 2009.
- [22] Simon Haykin. *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, Prentice Hall PTR, 1994.
- [23] Hamza Heni, Leandro C Coelho, and Jacques Renaud. Determining time-dependent minimum cost paths under several objectives. *Computers & Operations Research*, 105:102–117, 2019.
- [24] Stephane Hess, Mohammed Quddus, Nadine Rieser-Schüssler, and Andrew Daly. Developing advanced route choice models for heavy goods vehicles using GPS data. *Transportation Research Part E: Logistics and Transportation Review*, 77:29–44, 2015.
- [25] John Hickman, Dieter Hassel, Robert Joumard, Zissis Samaras, and S Sorenson. *Methodology for calculating transport emissions and energy consumption*. 1999. URL <http://www.transport-research.info/sites/default/files/project/documents/meet.pdf>. Available online (accessed on January 31, 2019).
- [26] Hino Canada. Growth with Toyota. Available online (accessed on January 31, 2019), 2019. URL <http://www.hinocanada.com/about-us/growing-with-toyota>.
- [27] Hino Canada. Town trucks: Hino light-duty cab-overs. Available online (accessed on January 31, 2019), 2019. URL <http://www.hinocanada.com/light-duty-trucks>.
- [28] Torsten Hothorn, Kurt Hornik, and Achim Zeileis. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3):651–674, 2006.
- [29] Yixiao Huang, Lei Zhao, Tom Van Woensel, and Jean-Philippe Gross. Time-dependent vehicle routing problem with path flexibility. *Transportation Research Part B: Methodological*, 95:169–195, 2017.
- [30] O Jabali, T Van Woensel, and A.-G. de Kok. Analysis of travel times and CO₂ emissions in time-dependent vehicle routing. *Production and Operations Management*, 21(6):1060–1074, 2012.

- [31] R Jaikumar, SMS Nagendra, and R Sivanandan. Modal analysis of real-time, real world vehicular exhaust emissions under heterogeneous traffic conditions. *Transportation Research Part D: Transport and Environment*, 54:397–409, 2017.
- [32] Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. Human decisions and machine predictions. *The Quarterly Journal of Economics*, 133(1):237–293, 2018.
- [33] Max Kuhn and Kjell Johnson. *Applied Predictive Modeling*, volume 810. Springer, New York, USA, 2013.
- [34] Hau L. Lee. Big data and the innovation cycle. *Production and Operations Management*, 27(9):1642–1646, 2018.
- [35] Bo Liu, Jie Hu, Fuwu Yan, Richard Fiifi Turkson, and Feng Lin. A novel optimal support vector machine ensemble model for NOx emissions prediction of a diesel engine. *Measurement*, 92:183–192, 2016.
- [36] Sendhil Mullainathan and Jann Spiess. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106, 2017.
- [37] David Munger, Pierre LEcuyer, Fabian Bastin, Cinzia Cirillo, and Bruno Tuffin. Estimation of the mixed logit likelihood function by randomized quasi-Monte Carlo. *Transportation Research Part B: Methodological*, 46(2):305–320, 2012.
- [38] Sunil Kumar Pathak, Vineet Sood, Yograj Singh, and SA Channiwal. Real world vehicle emissions: Their correlation with driving parameters. *Transportation Research Part D: Transport and Environment*, 44:157–176, 2016.
- [39] Jiani Qian and Richard Eglese. Fuel emissions optimization in vehicle routing problems with time-varying speeds. *European Journal of Operational Research*, 248(3):840–848, 2016.
- [40] Hesham Rakha, Kyoungcho Ahn, and Antonio Trani. Development of VT-micro model for estimating hot stabilized light duty vehicle and truck emissions. *Transportation Research Part D: Transport and Environment*, 9(1):49–74, 2004.
- [41] Christian P Robert, George Casella, and George Casella. *Introducing Monte Carlo methods with R*, volume 18. Springer, 2010.
- [42] George Scora and Matthew Barth. *Comprehensive modal emissions model (CMEM), version 3.01*. User guide. Centre for Environmental Research and Technology. University of California, Riverside 1070, 2006.
- [43] Hamza Ben Ticha, Nabil Absi, Dominique Feillet, and Alain Quilliot. Empirical analysis for the VRPTW with a multigraph representation for the road network. *Computers & Operations Research*, 88:103–116, 2017.
- [44] Samir Touzani, Jessica Granderson, and Samuel Fernandes. Gradient boosting machine for modeling the energy consumption of commercial buildings. *Energy and Buildings*, 158:1533–1543, 2018.

- [45] Transition Énergetique Québec. Rapport de consommation énergétique et d'émissions de gaz à effet de serre - Secteur institutionnel. Available online (accessed on January 31, 2019), 2017. URL <http://www.transitionenergetique.gouv.qc.ca/fileadmin/medias/pdf/institutions/TEQ-Rapport-consommation-energetique-GES-institutionnel-Partie-1.pdf>.
- [46] Transports Canada. Transportation in Canada 2016. Available online (accessed on January 31, 2019), 2017. URL https://www.tc.gc.ca/media/documents/policy/comprehensive_report_2016.pdf.
- [47] Marcel Turkensteen. The accuracy of carbon emission and fuel consumption computations in green vehicle routing. *European Journal of Operational Research*, 262(2):647–659, 2017.
- [48] Vladimir Vapnik. *The Nature of Statistical Learning Theory*. Springer Science & Business Media, USA, 2013.
- [49] Hal R Varian. Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2):3–28, 2014.
- [50] Weiliang Zeng, Tomio Miwa, and Takayuki Morikawa. Prediction of vehicle CO₂ emission and its application to eco-routing navigation. *Transportation Research Part C: Emerging Technologies*, 68:194–214, 2016.
- [51] Yanru Zhang and Ali Haghani. A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies*, 58:308–324, 2015.