AM-SP1110

Travel speed prediction using machine learning techniques

Maha Gmira[1*,2], Michel Gendreau[1,3], Andrea Lodi[1,2], Jean-Yves Potvin[3,4]

1. Département de mathématiques et de génie industriel, Polytechnique Montréal, C.P. 6079, succ. Centre-Ville, Montréal, Canada H3C 3A7 ; Phone : 1 514 340-5121 # 3392 ; Email : maha.gmira@polymtl.ca
2. Chaire d'excellence en recherche du Canada sur la science des données pour la prise de décision en temps réel, Polytechnique Montréal, Canada
3. Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise la Logistique et le Transport (CIRRELT), Université de Montréal, Canada
4. Département d'informatique et de recherche opérationnelle, Université de Montréal, Canada

Abstract

Given that the travel time required to get from one location to another in a road network is an important performance measure in intelligent transportation and advanced traveler information systems, accurate predictions in this regard are of foremost importance. In an urban environment, the travel time calculation depends on the vehicle speed, which can be highly variable due to congestion caused, for example, by accidents or bad weather conditions. At another level, one also observes daily patterns (e.g., rush hours), weekly patterns (e.g., weekday versus weekend), and seasonal patterns. Thus, capturing these features when modeling travel speed can have an immediate impact on commercial transportation companies that distribute goods by allowing them to optimize their routes to be more efficient and reduce their environmental footprint.

This paper presents a travel speed prediction methodology that mine big data collected from mobile location devices installed inside delivery vehicles to feed machine learning models. It is expected that the added information obtained through (big) data mining, should help these models to better predict travel speeds and travel times.

KEYWORDS:

Machine learning, Neural network, Travel speed, Prediction, GHG emissions.

Travel speed prediction using machine learning techniques

Introduction

There is a need for models that can derive future travel times from observed trends, in order to provide accurate predictions under recurrent and non-recurrent congestion. With the increasing amount of available data collected from probe vehicles, smartphone's applications and other location technologies, the challenge is no longer the quantity of data but rather the modelling and extraction of useful information from these data.

Being able to predict very precise traffic information can be of great value for transportation companies operating in urban areas. In particular, home delivery services can be greatly affected by congestion. In this industry, there is a great potential for the development of models based on real data aimed at reducing travel times and thereby minimizing GHG emissions.

According to a recent survey made in 2012 by the *Ministère de l'Environnement du Québec* (see Table 1), the sector that produced the most GHG emissions in the Province of Québec is transportation, accounting for 44.7% of the total, i.e., 40 million CO_2-equivalent tons (Québec, 2015). It should be noted that the transportation sector includes road transport, civil aviation, off-road motor vehicles, rail and water transport. However, on its own, *road transportation* accounted for 78.3% of all emissions in the transport sector, or 35.0% of the total GHG emissions.

Table 1 : GHG emissions in Québec between 2008 and 2012, (Québec, 2015).

Source	GFG Emissions (Mt CO_2-eq.)					Change in GHG Emissions 2008 - 2012		Share of Emissions in Québec 2012	
	2008	2009	2010	2011	2012	Mt CO_2-eq	%	Mt CO_2-eq	%
Transportation	**35.72**	**35.65**	**35.09**	**35.75**	**34.84**	**-0.88**	**-2.5**	**-0.91**	**-2.5**
Road	27.44	27.38	27.49	27.33	27.29	-0.15	-0.5	-0.04	-0.1
Civil aviation	0.73	0.67	0.66	0.63	0.63	-0.10	-13.4	0.00	0.6
Rail	0.90	0.93	0.85	0.90	0.94	0.04	4.6	0.04	4.4
Water	1.57	1.79	1.35	0.99	0.84	-0.73	-46.5	-0.15	-15.4

Thus, road is the dominant mode of transportation for people and for goods. According to the annual report of the *Minister of Public Works and Government Services*, 45.1% ($149 billions) of exports and 73.5% ($162 billions) of imports in 2011 were transported by truck across the border between U.S. and Canada, which represent 56.5% of all Canada-U.S. trade.

The aim of this paper is to predict more accurately travel speeds to better optimize the routes of delivery vehicles. For this purpose, a big database of GPS traces of the delivery activities of major Canadian retailers has been made available to us. From a methodological point of view, the proposed approach first asks for data cleaning, missing data pre-processing and geomatics validation. Then, unsupervised learning will be used to reduce data dimensionality. Finally, future speed values in the road network will be predicted through a recurrent Long Short-Term Memory neural network (LSTM). At the end, the impact of using more accurate travel speeds on the delivery routes will be evaluated based on greenhouse gas (GHG) emissions savings. Our approach is distinctive by its combination of unsupervised and supervised machine learning techniques, and by its interaction with a previously developed vehicle routing optimization algorithm.

Travel speed prediction using machine learning techniques

Literature review

There is a huge literature related to traffic prediction methods due to the development and use of intelligent transportation systems. These systems are highly dependent on accurate and real-time traffic information and at the same time, they collect large amount of real-time data (positions, speeds and individual itineraries). Methods to predict traffic information are classified as parametric, non-parametric or a combination of both, called hybrid methods. Parametric methods correspond essentially to Autoregressive Integrated Moving Average based models (ARIMA), smoothing techniques and Kalman filters. Non-parametric methods include non-parametric regression and different types of neural networks. Non-parametric methods are also known as data-driven methods, because they require a large amount of data. Although this can be seen as a disadvantage, they do not need to integrate any extensive expertise about traffic theory.

1. Parametric methods

The Kalman filter algorithm is a very popular short term traffic flow prediction method because it allows the state variable to be updated continuously, which is very important in time-dependent contexts. Some research works related to traffic state estimation are based on the original Kalman filter, as described in (Kalman & Bucy, 1961), and its extensions for nonlinear systems (Julier & Uhlmann, 1997) called the *extended Kalman filter*. In (Wang & Papageorgiou, 2005), the authors developed a freeway state estimator using an extended Kalman filter to solve a macroscopic traffic flow model. This traffic state estimator approach was tested on several scenarios and real cases. (Antoniou, Ben-Akiva, & Koutsopoulos, 2007) solved a dynamic traffic assignment model, which is a nonlinear state-space model, by applying three different extensions of the Kalman filter. The author in (Van Lint, 2008) used a new extended Kalman filter (EKF) based on online-learning to provide predictive travel time information on freeways. This EKF was applied in this case because the travel time depends on traffic conditions that are highly dynamic and nonlinear, changing over time and space. In (Jula, Dessouky, & Ioannou, 2008), the aim was to estimate arrival times at the nodes of a route. The authors first predicted travel times on the arcs of a transportation network by feeding the Kalman filter with historical data. Then, this prediction was corrected and updated with real time information using the Kalman filter's corrector-predictor form.

To predict short-term traffic characteristics such as speed, flow or travel time, time series models using ARMA techniques (which is a combination of autoregressive AR and moving average MA) have been widely used. ARIMA models generalize ARMA models for non-stationary time series. They rely on stochastic system theory since the processes are non-deterministic. (Smith, Williams, & Oswald, 2002) compared a non-parametric regression model using a specific state space, where the forecast generation method and neighbor selection criteria were heuristically improved, with a seasonal ARIMA model, called SARIMA. The tests showed that the SARIMA model performed better than the improved non-parametric regression.

In (Hamed, Al-Masaeid, & Said, 1995), the authors used the Box-Jenkins approach to develop a forecasting model based on ARIMA using data collected from urban streets (i.e., 1-min traffic-volume on each street during peak periods). After comparing several ARIMA models, the one of order (0,1,1) yielded the best results in terms of forecasting traffic volume, where the values 0, 1 and 1 refer to the order of the autoregressive, differencing and moving-average terms, respectively.

2. Non-parametric methods

These methods are called non-parametric because the number and types of the parameters are unknown a priori. In this category, Artificial Neural Networks (ANN) are the most popular prediction methods.

Travel speed prediction using machine learning techniques

Among neural networks, backpropagation networks are widely applied because of their ability to model complex nonlinear relationships among continuous variables. These multilayer networks perform supervised learning. That is, the error between the network's response and the desired response is propagated backward using the back-propagation algorithm (Rumelhart, Hinton, & Williams, 1988) to modify the weights on the connections between successive layers, so as to obtain a gradient descent along the error function.

In (Moniruzzaman, Maoh, & Anderson, 2016), a multilayer backpropagation neural network was used to predict the time needed to cross the Ambassador bridge, one of the busiest between Canada and US. For this study, a database of a yearlong GPS records was used to train and validate the neural network. The same type of neural network was also used in (Smith & Demetsky, 1997) to forecast freeway traffic flow, using data collected by an operational ATMIS (Advanced Traffic Management and Information System). However, this neural network model produced significant underestimates of future traffic volumes because the training data were not representative of the general traffic flow.

Another class of non-parametric methods is Non-Parametric Regression (NPR) where the objective is to estimate the regression function directly without specifying its form explicitly. This is to be opposed to traditional parametric regression models, where the form is given but the parameters must be estimated. NPR is a forecasting technique that exploits past observations of the system under study, and uses a search procedure to find observations that are geometrically similar to the current conditions. Then, it feeds those neighboring observations to the forecast function to estimate the future state of the system. NPR is suitable for short term traffic prediction because it can deal with traffic flow uncertainty. The *k*-nearest neighbor *k*-NN-NPR is a class of non-parametric regression, made of two components:

- the search procedure: the nearest neighbors (historical data most similar to the current input) are the inputs to the forecast function aimed at generating an estimate. The nearest neighbors are found using a dissimilarity measure, which is usually based on the Minkowski distance metrics, as defined below:

$$\mathrm{L}_m = \left[\sum_{i=1}^{n} |p_i - q_i|^m\right]^{1/m}$$

  where n is the dimensionality of the state vector, p_i is the i^{th} element of the current historical observation, q_i is the i^{th} element of the current conditions, and m is a parameter with values between 1 and 2. The most common implementation uses a sequential search procedure. However, as the number of historical observations increases, the sequential search can become very time-consuming.

- the forecast function: the most general approach to generate a forecast is to compute an average of the dependent variable values over all nearest neighbors. However, this approach ignores the information provided by the distance metric (i.e., past observations closer to the current input should have more impact on the forecast).

(Davis & Nihan, 1991) were among the first to use NPR to estimate short-term traffic flows. The authors highlighted the importance of a large and representative amount of data. (Smith, 1995) used NN-NPR to estimate traffic volumes from two locations on the Beltway using five months of observations. The results showed that this method produces better predictions. (Smith, 1995) compared historical averages, time series, back-propagation neural networks and non-parametric regression models using a performance index that included absolute error, error distribution, ease of model implementation and model portability. The results showed that NPR was better than the others models and was also easier to implement.

Travel speed prediction using machine learning techniques

General context

Our travel speed prediction method is developed within the framework of a larger project aimed at the reduction of GHG emissions from commercial delivery operations in large urban areas. We want to exploit accurate travel speed predictions under various traffic conditions to reduce GHG emissions by minimizing the total travel time of commercial delivery routes. More precisely, the evolution of travel speeds over a given day (aggregated in intervals of 15 minutes) has to be predicted based on historical data. To this end, GPS traces collected from mobile positioning devices installed inside delivery vehicles are first processed through data mining techniques.

The information obtained on the travel speed evolution over a day for each road segment is then fed to a neural network to produce travel speed predictions. After optimizing vehicle delivery routes based on these predicted travel speeds, an evaluation of the gain in GHG emissions can be performed. For example, the emission function from the MEET report (Hickman, Hassel, Joumard, Samaras, & Sorenson, 1999) can be used to obtain the emission rate for a given travel speed. It should be noted that other works concerned with the reduction of GHG emissions in the vehicle routing domain can be found in (Jabali, Woensel, & de Kok, 2012), (Bektaş, Demir, & Laporte, 2016), (Bektaş & Laporte, 2011), among others.

Figure 1 shows the various tasks involved in our research. The most important tasks are described in the following.

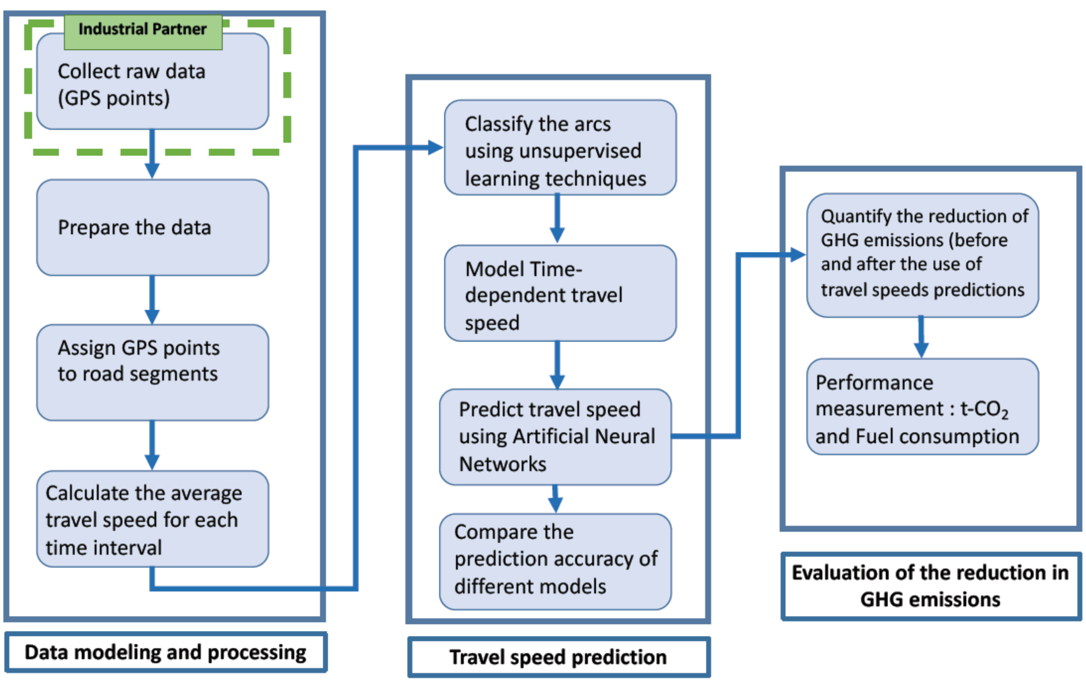

Figure 1 : **Description of the global research project**

Travel speed prediction using machine learning techniques

1. Data modeling and processing

We work with a company involved in the development of vehicle routing optimization software for large Canadian retailers that deliver furniture, appliances, electronics and other products to their customers. Overall, the delivery routes contain more than 2,500,000 delivery points over the country, served by nearly 200,000 routes. The software developed by our partner combines mathematical algorithms with GPS technology to automatically plan the home delivery process. The company has provided us with a database of GPS traces (around 50 GB) for the city of Montreal and its surroundings (Figure 2), which spans two years. Data were collected through Automatic Vehicle Location (AVL) systems where GPS receivers are interfaced with Global System for Mobile Communications (GSM) modems installed inside the vehicles. The system records point locations (as latitude-longitude pairs), vehicle speed, date and time. The attributes of the given database are described in Table 2.

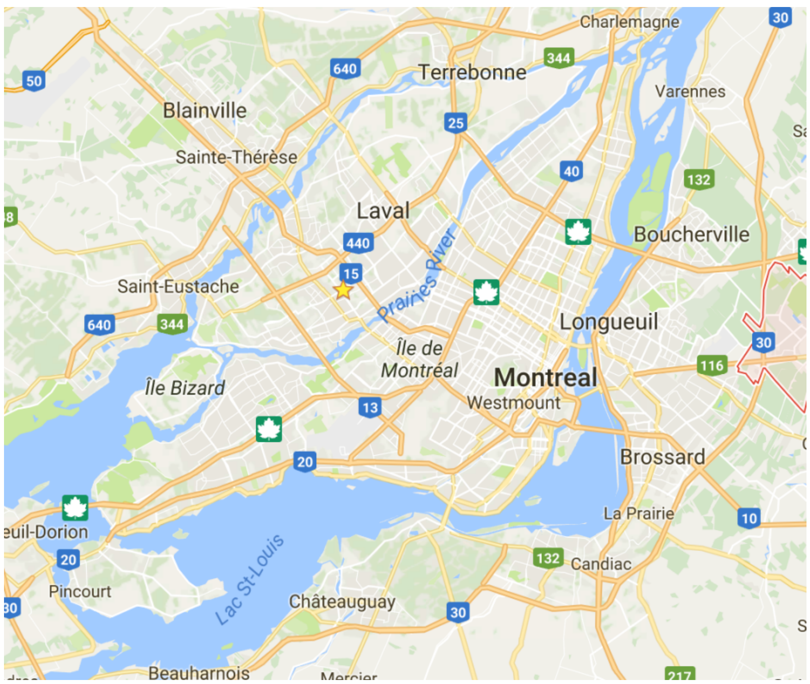

Figure 2 : **Metropolitan area of Montreal, Google Map 2017**

Table 2: **Description of the attributes of the database**

Attributes	Formats
Idposition	character
Latitude	Numeric(10) with a precision 6
Longitude	Numeric(10) with a precision 6
Speed	Numeric(10) with a precision 6
Idmobile	character
Iddriver	Numeric(10) with a precision 0
Date	Timestamp without time zone

Travel speed prediction using machine learning techniques

Because real data usually contain missing, erroneous and inconsistent values, data validation is a crucial step. Specifically, some data collection technologies, such as sensors, are inherently inaccurate because of reading errors, transmission limitations, etc. Once the data are cleaned, a geographical information system (GIS) Is used for edition and analysis purposes. This first phase of the project can be summarized as follows:

- Prepare and clean the data
- Assign GPS traces to road segments
- Determine the traveling direction of GPS points
- Calculate the average speeds on each road segment for intervals of 15 minutes in each direction over a day. These travel speeds account for speed limits, direction of the trip, one-way streets, restricted turns and all forms of limited access to specific roads.

2. *Clustering*

To aggregate the data, the arcs or road segments in the network that exhibit similar speed evolution patterns are first grouped together through unsupervised learning. The clustering algorithm takes a dataset of *N* vectors as input, where each vector contains speed values for a given arc within time slots of 15 minutes (for two years of recordings, which represents 1,008,000 records). The algorithm creates a set of *K* cluster centroids C_k as output, where each observation is assigned to a particular cluster. The objective is to minimize the following distance metric between the inputs $\mathbf{x_n}$ assigned to each cluster *k* and the corresponding centroid $\boldsymbol{\mu_k}$ *:*

$$Min \sum_{k=1}^{K} \sum_{x_n \epsilon C_k} ||x_n - \mu_k||^2$$

To find a solution to this problem, which is NP-hard, we use an iterative method called the *Lloyd's algorithm* that converges to a local minimum. In this algorithm, the following steps are performed repeatedly: 1) update the clusters to include the observations that are the closest in distance from their centroid and 2) recalculate each centroid as the mean of all observations assigned to the corresponding cluster.

$$C_k = \left\{ x_n : \left||x_n - \mu_k|\right| = \min_{l=1,\dots,K} \left||x_n - \mu_l|\right| \right\}$$

$$\mu_k = \frac{1}{C_k} \sum_{x_n \in C_k} x_n$$

At the end of this process, each arc is associated with a cluster or class. This information is then used for travel speed prediction, as it is explained below.

3. *Travel speed prediction*

In the last decades, several researchers have proposed the use of artificial neural networks for different transportation problems. They have been applied in particular for the prediction of traffic characteristics such as the flow, travel speed or travel time because these networks are well suited to complex non-linear dynamic problems. Basically, artificial neural networks can deduce existing relationships in large input-output database during their learning phase.

Travel speed prediction using machine learning techniques

Accordingly, we propose a LSTM recurrent neural network to make travel speed predictions. This choice is motivated by the predictive power of these models which relies on their ability to learn from historical data and to catch hidden and strongly non-linear dependencies, even when there is significant noise in the training data. Here, the neural network will be able to predict the evolution of the travel speed for the next day for any given road segment, by providing as input some aggregation of the evolution of the travel speed on the same road segment in the past, plus the class of that road segment as well as other exogenous variables that can have an impact on the prediction, like meteorological conditions.

The LSTM neural network model is suitable for variables that evolve over time. Due to its structure, it is able to take into account short-term dependencies while avoiding the problem of loss memory encountered with basic recurrent neural networks. The network structure chosen can discover correlations by looking at the independent variables with the most significant impact on the travel speed. It will therefore assign larger weights to the connections related to the most important independent variables.

To perform supervised learning with the LSTM neural network to predict traffic speeds, a training process is needed. This process includes three phases:

- Training phase: use a set of data to train the model by pairing the input speed with the expected output speed for all the arcs of the road network. This training phase is performed via backpropagation.
- Validation phase: select the best performing parameter values (optimal number of hidden layers, stopping point for the back-propagation algorithm, etc.)
- Testing phase: use a set of data to estimate the accuracy of the selected approach. In this case, we use the testing phase to estimate the error rate after we have chosen the final LSTM size and actual weights.

To evaluate the performance of the neural network, we use the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{l=1}^{n}(s_{p,l} - s_{r,l})^2}$$

$$MAPE = 100 * \frac{1}{n}\sum_{l=1}^{n}\left|\frac{s_{p,l} - s_{r,l}}{s_{r,l}}\right|$$

where $s_{p,l}$ is the predicted travel speed of the lth arc and $s_{r,l}$ is the real travel speed of the lth arc.

4. Evaluation of GHG emission reductions

Using the predicted travel speeds, the software of our industrial partner will be used to optimize delivery routes. These routes will then be compared with those used in practice to quantify the gains obtained in total travel time. These gains will finally be translated into fuel consumption and GHG emission gains. Our method for GHG accounting will be consistent with ISO 14064-2 (ISO, 2006).

Conclusion

Because travel speeds are of foremost importance for intelligent transportation systems, the aim of this research is to analyze traffic data through machine learning techniques to accurately predict travel speeds in urban networks.

Travel speed prediction using machine learning techniques

We propose a Long Short-Term Memory neural network to predict the travel speed between any two locations at a given time of a given day. This method exploits historical data (GPS traces) gathered from delivery vehicles. The LSTM model is trained on a large database of speed records obtained from an industrial partner who develops vehicle routing solutions for home delivery companies. The output of the neural network will then be fed to a previously developed vehicle routing optimization algorithm to show the benefits of more accurate travel speed predictions on the obtained delivery routes.

References

1. Antoniou, C., Ben-Akiva, M., & Koutsopoulos, H. N. (2007). Nonlinear Kalman filtering algorithms for on-line calibration of dynamic traffic assignment models. *IEEE Transactions on Intelligent Transportation Systems, 8*(4), 661-670.
2. Bektaş, T., Demir, E., & Laporte, G. (2016). Green vehicle routing. *Green Transportation Logistics* (pp. 243-265): Springer.
3. Bektaş, T., & Laporte, G. (2011). The pollution-routing problem. *Transportation Research Part B : Methodological, 45*(8), 1232-1250.
4. Davis, G. A., & Nihan, N. L. (1991). Nonparametric regression and short-term freeway traffic forecasting. *Journal of Transportation Engineering, 117*(2), 178-188.
5. Hamed, M. M., Al-Masaeid, H. R., & Said, Z. M. B. (1995). Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering, 121*(3), 249-254.
6. Hickman, J., Hassel, D., Joumard, R., Samaras, Z., & Sorenson, S. (1999). *MEET-Methodology for calculating transport emissions and energy consumption. European Commission, DG VII*.
7. ISO (2006). 14064-2: 2006 Greenhouse gases-Part 2: Specification with guidance at the project level for quantification, monitoring and reporting of greenhouse gas emission reductions or removal enhancements. *International Organization for Standardization, Geneva, Switzerland*.
8. Jabali, O., Woensel, T., & de Kok, A. (2012). Analysis of travel times and CO^2 emissions in time-dependent vehicle routing. *Production and Operations Management, 21*(6),1060-1074.
9. Jula, H., Dessouky, M., & Ioannou, P. A. (2008). Real-time estimation of travel times along the arcs and arrival times at the nodes of dynamic stochastic networks. *IEEE Transactions on Intelligent Transportation Systems, 9*(1), 97-110.
10. Julier, S. J., & Uhlmann, J. K. (1997). New extension of the Kalman filter to nonlinear systems. Presented at AeroSense'97.
11. Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Journal of basic engineering, 83*(3), 95-108.
12. Moniruzzaman, M., Maoh, H., & Anderson, W. (2016). Short-term prediction of border crossing time and traffic volume for commercial trucks: A case study for the Ambassador Bridge. *Transportation Research Part C: Emerging Technologies, 63*, 182-194.
13. Québec (2015). Inventaire québécois des émissions de gaz à effet de serre en 2012 et leur évolution depuis 1990. *Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques*.
14. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning internal representations by error propagation. *Neurocomputing: Foundations of Research* (pp. 673-695): The MIT Press
15. Smith, B. L. (1995). Forecasting freeway traffic flow for intelligent transportation systems application. Doctoral dissertation. Department of Civil Engineering, University of Virginia, Charlottesville.

Travel speed prediction using machine learning techniques

16. Smith, B. L., & Demetsky, M. J. (1997). Traffic flow forecasting: comparison of modeling approaches. *Journal of Transportation Engineering, 123*(4), 261-266.
17. Smith, B. L., Williams, B. M., & Oswald, R. K. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies, 10*(4), 303-321.
18. Van Lint, J. (2008). Online learning solutions for freeway travel time prediction. *IEEE Transactions on Intelligent Transportation Systems, 9*(1), 38-47.
19. Wang, Y., & Papageorgiou, M. (2005). Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transportation Research Part B: Methodological, 39*(2), 141-167.