

**QUANTIFYING THE EFFECTS OF INPUT AGGREGATION AND MODEL
RANDOMNESS ON REGIONAL TRANSPORTATION EMISSION INVENTORIES**

Timothy Sider^a, Gabriel Goulet-Langlois^b, Naveen Eluru^c, Marianne Hatzopoulou^d

^a M.Eng. Candidate

Department of Civil Engineering and Applied Mechanics, McGill University
817 Sherbrooke St. W., Room 492
Montréal, Québec, H3A 2K6, Canada
Tel: 514-398-6935, Fax: 514-398-7361
E-mail: timothy.sider@mail.mcgill.ca

^b B.Eng. Candidate

Department of Civil Engineering and Applied Mechanics, McGill University
817 Sherbrooke St. W., Room 492
Montréal, Québec, H3A 2K6, Canada
Tel: 514-398-6935, Fax: 514-398-7361
E-mail: gabriel.goulet-langlois@mail.mcgill.ca

^c Associate Professor

Department of Civil, Environmental and Construction Engineering, University of Central Florida
12800 Pegasus Drive, Room 301D,
Orlando, Florida 32816, USA
Tel: 407-823-4815, Fax: 407-823-3315
E-mail: naveen.eluru@ucf.edu

^d Assistant Professor (Corresponding Author)

Department of Civil Engineering and Applied Mechanics, McGill University
Macdonald Engineering Building
817 Sherbrooke St. W., Room 492
Montréal, Québec, H3A 2K6, Canada
Tel: 514-398-6935, Fax: 514-398-7361
E-mail: marianne.hatzopoulou@mcgill.ca

ABSTRACT

Accurate road-traffic emission inventories are of great interest to metropolitan planning agencies especially in the appraisal of regional transport policies. Integrated road transport emission models are an effective means of establishing emission estimates, yet their development requires significant investments in data and resources. It is therefore important to investigate which data inputs are the most critical to inventory accuracy. To address this issue, an integrated transport and emissions model is developed using the Montreal metropolitan region as a case-study. Daily regional hydrocarbon (HC) emissions from private individual travel are estimated, including the excess emissions due to engine starts. The sensitivity of emission estimates is then evaluated by testing various levels of input aggregation common in practice and in previous research. The evaluated inputs include the effect of start emissions, ambient weather conditions, traffic speed, path choice, and vehicle registry information. Inherent randomness within the integrated model through vehicle selection and path allocation is also evaluated. The inclusion of start emissions is observed to have the largest impact on emission inventories, contributing approximately 67% of total on-road HC emissions. Ambient weather conditions (season) and vehicle registry data (types, model years) are also found to be significant. Model randomness had a minimal effect in comparison with the impact of other variables.

Keywords: Emission modeling, traffic assignment, start emissions, model sensitivity, emission inventories

1. INTRODUCTION

The importance of accurate transport emission inventories has never been greater given the current impact that transportation systems have on the local and global environment. The current state of practice primarily involves the development of inventories through integrated transportation and emission models. Increases in model complexity have generally resulted in improved estimates, yet the improvements often come at the cost of investing in resource-intensive data inputs. In addition, not all investments in more accurate data inputs yield a similar increase in estimate accuracies. For practitioners involved in transportation, environmental, and public health policy, this issue becomes of utmost importance.

In developing transport emission inventories, detailed input data are often non-existent or difficult to gather therefore simplifying assumptions must be made. Yet which simplifying assumptions are reasonable, and which result in estimation errors, are still not entirely understood. Kioutsioukis et al. (2004) were among the first in assessing the uncertainty and sensitivity of emission models of the late 1990s and early 2000s. Their summary of previous research concluded that accurate activity data was as important as accurate emissions data in generating better inventories.

This study aims to explore the different inputs within an integrated transport and emission model and identify those that contribute the most to the accuracy of the final emission inventory. For this purpose, we estimate total daily hydrocarbon (HC) emissions from passenger travel at the metropolitan level through the development of an integrated regional traffic assignment model and emission simulator. The emission simulation contains sub-models for both running emissions (generated throughout the trip during driving and idling) and start emissions (generated at the beginning of a trip). The primary inputs of both running and start emission models are then varied in order to test the effects of different levels of aggregation in model inputs on the final estimates. We also test the effects of prevailing assumptions frequently made when emission inventories are conducted within government agencies. Our study is set in the Montreal metropolitan region, Canada, an area further described in the methodology section.

2. CONTEXT

Integrated transport and emission models often aim to estimate traffic-related emissions by combining transportation data inputs and emission simulators. The inputs from the transportation side of the ledger have ranged from aggregate methods such as commute distances from travel survey data (Ko et al., 2011; Brand and Preston, 2010; Frank et al., 2000), to mesoscopic modeling using improved travel behaviour data (Borge et al., 2012; Barla et al., 2011; Xia and Shao, 2005; Borrego et al., 2004; Anderson et al., 1996), and even all the way to complex methods involving fully integrated agent-based models (Hülsmann et al., 2014; Kickhöfer and Nagel, 2013; Lefebvre et al., 2013; Hao et al., 2010; Beckx et al., 2009a). Meanwhile, emission models have also varied regarding the level of detail of vehicle characteristics, ambient conditions, and speed aggregation. Aggregate emission models typically use constant emission factors while ignoring congestion-related speed impacts (Barla et al., 2011; Ko et al., 2011; Brand and Preston, 2010). Others have employed more detailed emission factors yet do not account for congestion in their speed estimates (Frank et al., 2000). However, the majority of emission models favour a congestion-related approach in combination with more accurate vehicle registry data (Borge et al., 2012; Hao et al., 2010; Beckx et al., 2009a; Borrego et al., 2004; Anderson et al., 1996). Start emissions have also been incorporated in some models. Previous research that had employed the MOBILE platform implicitly account for start emissions through the generated running emission factors (Hao et al., 2010). COPERT models account for starts in a similar fashion (Waked and Afif, 2012). The current generation of emission software favours a distinct separation between running and start emissions, with start emission factors used to dictate the relationship between vehicle characteristics, soak-time and amount of pollutants emitted during engine ignition.

Recent research has focussed on the development of truly integrated transport and emission models incorporating detailed road-traffic simulations. Road-traffic networks are modeled at a macroscopic or mesoscopic scale, and then trips are assigned to the network on an hourly basis. Output data include vehicle kilometers travelled and average link speeds, which are both used directly in conjunction with the emission factors. This process has garnered international popularity and has been employed in cities and regions such as Hamilton, Canada (Anderson et al., 1996), Helsinki, Finland (Karppinen et al., 2000), Antwerp, Belgium (Mensink et al., 2000), Lisbon, Portugal (Borrego et al., 2004), Hong Kong, China (Xia and Shao, 2005),

Norwich, England (Nejadkoorki et al., 2008), and Madrid, Spain (Borge et al., 2012). Beckx et al (2009a) and Hao et al (2010) even incorporated activity-based models into their integrated modeling frameworks applied respectively to the Netherlands and the Greater Toronto Area, Canada. The advantage of integrated transport and emission models is that they are able to combine congestion-related speed effects and detailed vehicle data, often at the individual trip level. This provides significant detail regarding spatial and temporal variation, and allows for connections to socioeconomics or even to dispersion models that portray how meteorological conditions and the built environment influence pollutant concentrations (Hatzopoulou and Miller, 2010; Beckx et al., 2009b).

Despite the spate of recent studies, there is an on-going debate as to whether emission inventories from integrated models have actually become more accurate over time. In a meta-analysis of 50 studies, Smit et al (2010) found that although models have become more complex, their prediction accuracy has not necessarily increased noting an inadequate understanding of uncertainties in modeling estimates.

Regarding transportation inputs, the key components from the transport model are the path selection type (i.e. how an assignment model allocates a trip between an OD pair to the road network) and the level of speed aggregation (instantaneous speed, link speed, shortest path speed, average network speed). The level of detail in the base vehicle registry data is also vital, specifically whether it contains individual/zonal details on vehicle type, age, or fuel composition. Components of the emission model play an important part as well, principally which engine phase (cruising, accelerating, idling, ignition) is included in estimates, and which atmospheric conditions (seasonal variation, air conditioning use) are considered. Therefore, the objective of this study is to build on the research by Smit et al (2010) by testing the impacts of input data on emission estimates using an integrated emissions model that incorporates a module for vehicle allocation and specifically accounts for start emissions.

3. METHODOLOGY

The study methodology comprises two main tasks in the estimation and evaluation of total daily regional emissions (HC per day): 1) development of the integrated regional vehicle emissions model; and 2) testing the effects of model randomness and variability in input data.

The Montreal metropolitan region is the study area for this exercise and consists of the island of Montreal, the cities of Laval and Longueuil, and municipalities on the North Shore and South Shore (Figure 1). The region covers an area of approximately 7,000 km² and has a population of 3.8 million (Statistics Canada, 2011). The island of Montreal contains 47 percent of the region's inhabitants and 71 percent of the region's employment opportunities (AMT, 2008). The central business district is located on the island of Montreal and is primarily surrounded by dense mixed-use neighbourhoods. The region has an average vehicle ownership of 1.23 vehicles per household and a mode split of 68% car, 17% public transit, and 12% active transportation (AMT, 2008). Meanwhile, the island of Montreal is connected to the other sub-regions through a system of bridges. Five bridges connect the island to the north and five to the south, while two bridges at either end of the island connect the peripheral eastern and western edges. With the very high proportion of off-island and on-island commuters, bridges linking the island to the rest of the region have become the salient element of the road network.

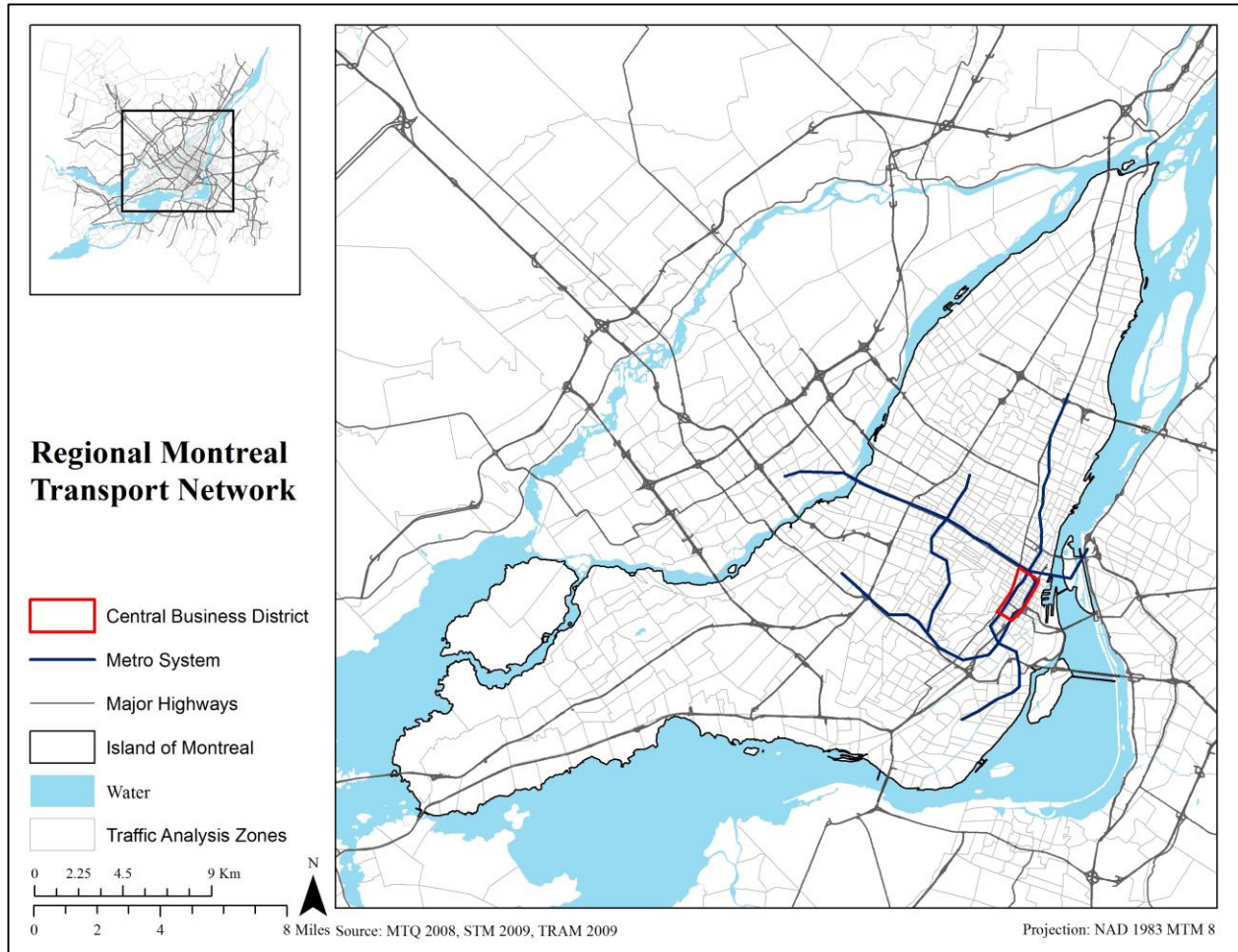


FIGURE 1 Montreal metropolitan region

3.1 Model development

A regional vehicle emission modeling framework was developed for the Montreal metropolitan region. The model uses Origin-Destination (OD) survey data in order to accomplish the following tasks: 1) allocating vehicles per household and trip, 2) estimating emissions on the road network for all trips in a typical day (24 hours), and 3) estimating emissions for vehicle starts over the entire day. The modeling framework is presented in Figure 2 and illustrates all the interactions between the major model components and the base data sources. Detailed descriptions of each sub-model, specifically the vehicle allocation, running emissions estimation, and start emissions estimation, are provided over the remainder of this subsection. Further details on the integrated model are documented in Sider et al. (2013).

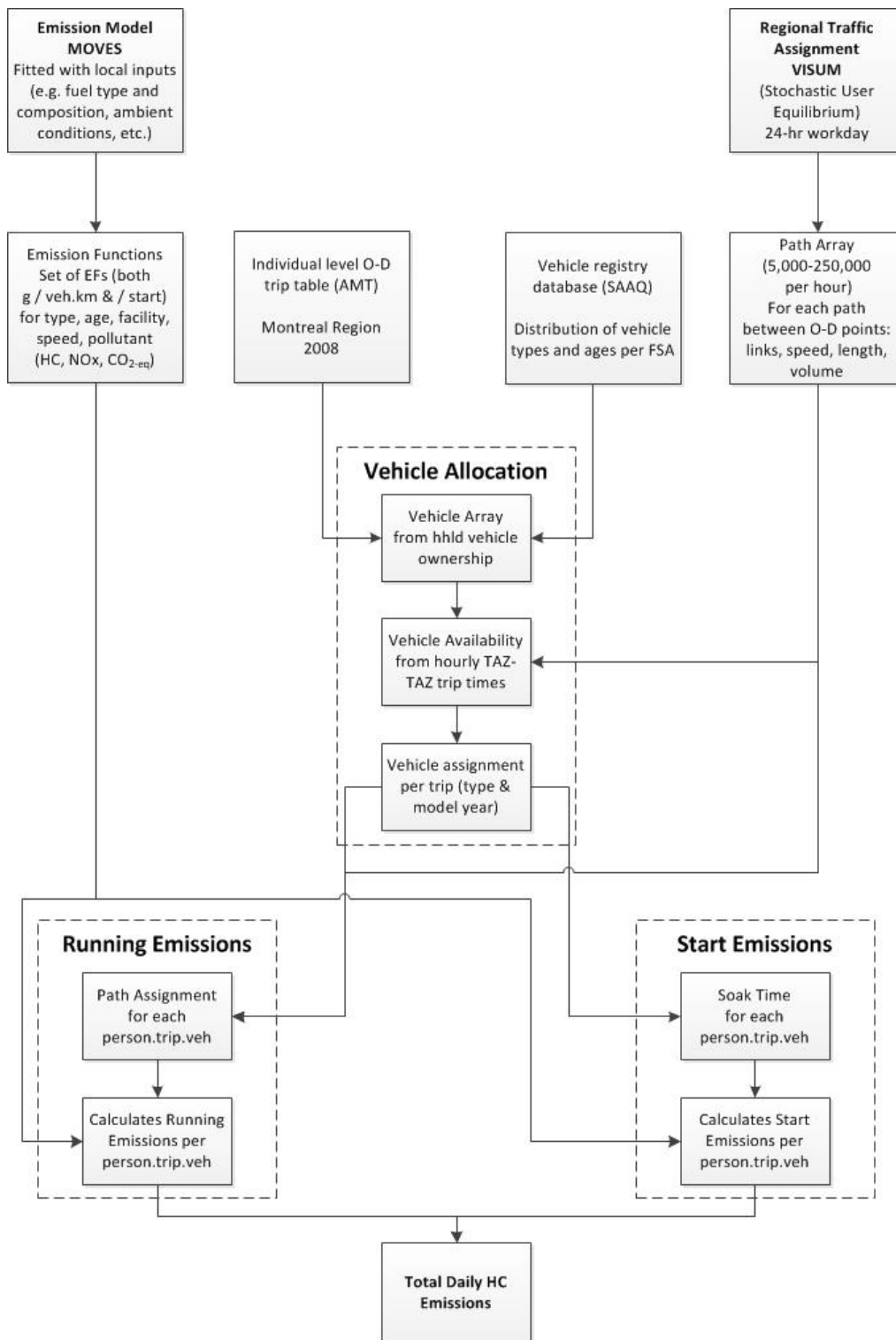


FIGURE 2 Integrated emissions modeling framework

Vehicle Allocation

The base travel data for this study encompasses the 2008 origin-destination survey conducted by the Agence Métropolitaine de Transport (AMT), the regional transit authority in the Montreal metropolitan region. Every 5 years, AMT collects detailed travel diaries from 5 percent of the region's households, resulting in a sample of about 43,000 households from the most recent survey (AMT, 2008). The survey results are then released in a trip-by-trip database format with details such as origin and destination coordinates, mode choice, start time, trip purpose.

In order to maintain vehicle consistency between the estimate methods for running and start emissions, the first step in this study involved allocating a vehicle to each of the driving trips in the 2008 OD survey (162,364 trips). Working at the household level, the main elements involved with vehicle allocation are the number of vehicles owned and each vehicle's time of availability and geographic coordinates. Trip availability/end times are approximated using a regional traffic assignment model of the Montreal metropolitan region implemented using the VISUM platform. The driving trips from the OD survey are then expanded to the full population and allocated to the road network using a stochastic user-equilibrium assignment. Hourly travel times were estimated for travel between the 1552 traffic analysis zones (TAZs), and were then added to the start times for each trip in order to approximate trip end times.

In a second step, an array was created for each household that had a number of elements equal to the number of vehicles owned. Each vehicle in the array was initialized at the household's geographic coordinates. An algorithm then ordered all household trips chronologically and assigned to every trip an index based on vehicle availability (time and geographic coordinates). Finally, each vehicle index was randomly allocated a vehicle type and model year based on the cumulative distribution function of the vehicle fleet composition of the household's residential area. The vehicle registry database was obtained from the provincial registry at the Société de l'Assurance Automobile du Québec (SAAQ), and contained information on vehicles registered at the forward-sorting area (FSA) level (ie. the first three digits of an area's postal code). Vehicle age was included in the registration data and has been conserved in our model. The fuel type was assumed to be exclusively gasoline, which represents over 95% of the fuel used in passenger cars (of varying sizes) in Canada. The characteristics of gasoline sold in the province of Quebec (in which the Montreal region is located) were used as input in the emission model. The regional distribution of all vehicles by vehicle age and category

is shown in Figure 3. In this study, we have grouped all passenger cars under one category (light-duty gasoline cars) while mini-vans and sports utility vehicles (SUVs) are grouped under another category (light-duty gasoline trucks). This collapse into two categories is required by the emission simulator, which includes one set of emission factors for each category and doesn't differentiate between passenger cars or passenger trucks of different sizes. Therefore, every driving trip in the OD survey was allocated a vehicle type and model year that remained constant over a day's worth of trip chains.

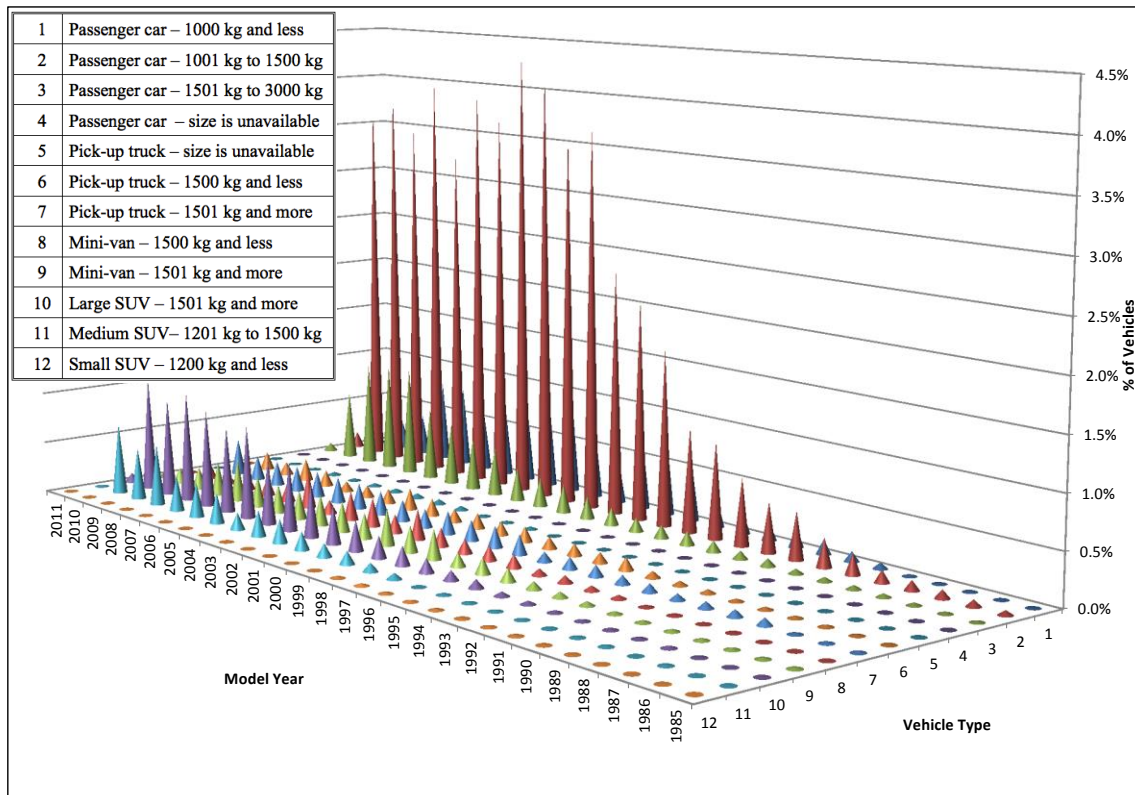


FIGURE 3 Vehicle age and type distribution for the Montreal metropolitan region

Running Emissions

Total daily running emissions (which include emissions released during a trip: cruising, acceleration, deceleration, idling) in the Montreal metropolitan region were estimated through the integration of a regional traffic assignment model and an emission model. The regional traffic assignment model was input with hourly OD matrices generated from the AMT's 2008 survey. The OD matrices were created by expanding the driving trips in the OD survey to the full

population based on expansion factors provided by the AMT. Traffic was assigned to the regional network using a stochastic user-equilibrium assignment, and detailed hourly OD-level path information was output, including congestion-related link speeds and volumes.

Concurrently, a set of emission factors (EFs) expressed in grams of pollutant per vehicle-kilometer were generated using the Mobile Vehicle Emissions Simulator (MOVES2010b) platform developed by the United States Environmental Protection Agency (USEPA). Montreal-specific data were input in the form of climate trends, vehicle fleet registry, and fuel composition, and individual EFs were output that account for vehicle type (passenger car and passenger truck), vehicle model year (1978-2008), vehicle speed (17 speed bins), facility type (restricted versus unrestricted), pollutant (total hydrocarbons (HC), nitrogen oxides (NO_x), and greenhouse gases (CO_{2-eq})) and season (winter and summer). Also, 2008 was chosen as the base year for vehicle model year estimation in order to match the OD survey year.

An emission post-processing algorithm was then used to iterate trip-by-trip through the OD data and estimate trip level emissions. An individual trip is first randomly allocated to a path between an origin and destination based on a cumulative distribution function of the volumes on all paths from O to D. Link-level emissions are then calculated by linking the vehicle type and model year from the vehicle allocation with its specific EF (g/veh.km) at the congested link-speed, and by multiplying the EF by the link length. All link emissions along the path were then summed in order to calculate total trip emissions. Full population running emission estimates were then developed by expanding the individual trip emissions with statistically derived expansion factors.

Start Emissions

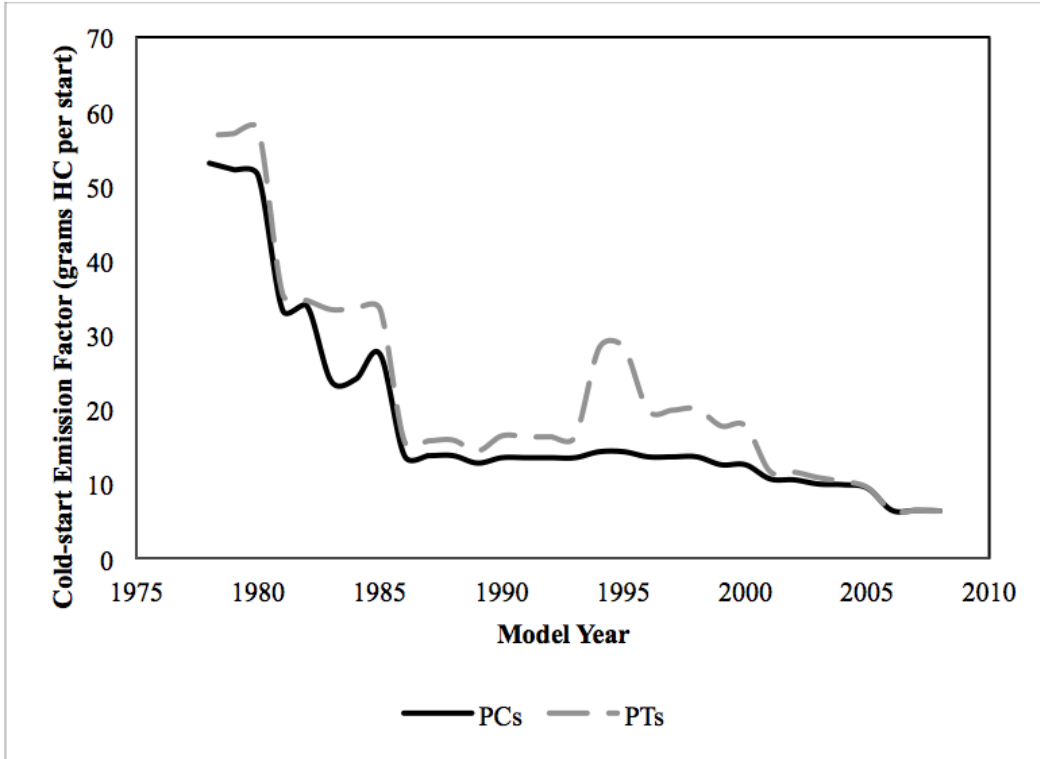
Start emissions occur at engine ignition and are considered separate from those that are emitted during an engine's running phase. They are mostly caused by: (1) excess gasoline due to higher fuel enrichment and (2) poor catalytic converter performance due to the large gap between engine temperature at ignition and optimum temperature for catalytic conversion. The gap between engine and optimum temperatures is dependent on the ambient temperature, which varies by season, and the engine's ignition temperature, which is related to the engine soak-time, i.e. amount of time since the engine was turned off prior to ignition. Therefore, an engine that

had been turned off recently would have a smaller soak-time and thus would have an ignition temperature that is closer to the optimum temperature of the catalytic converter.

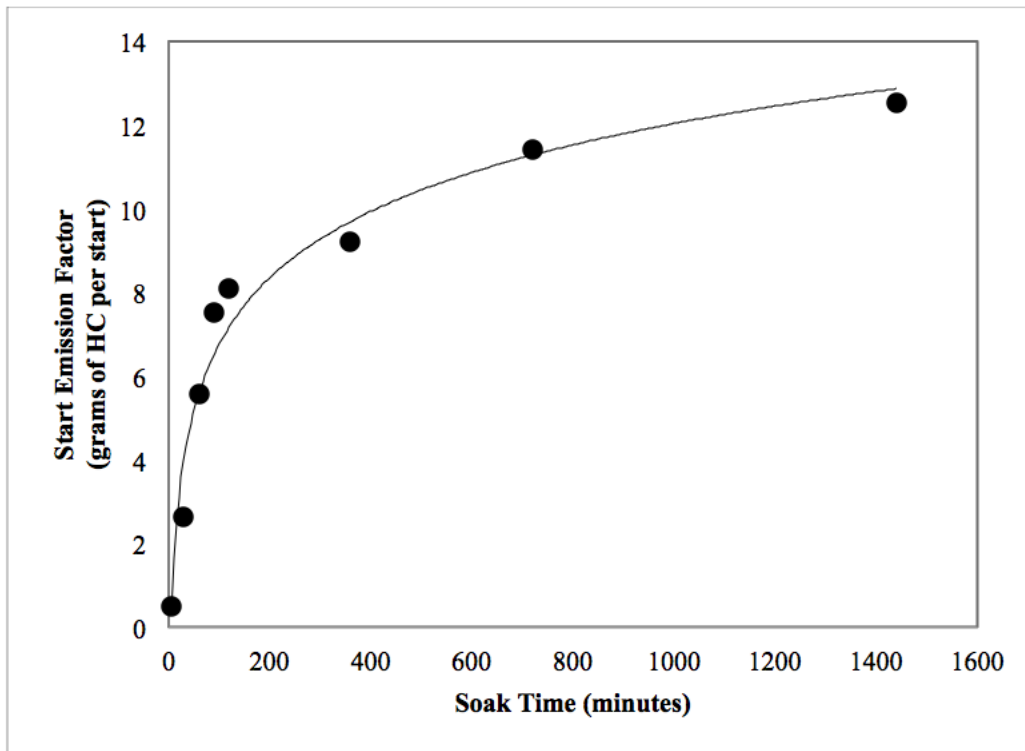
In order to estimate trip-level start emissions, each trip needed to have a soak-time and a vehicle-specific start EF. Soak-times were estimated once the vehicle allocation process was completed. The entire trip chain was tracked for every vehicle during the day and a sequence of engine ignitions and shutdowns was used to calculate soak-times. Specifically, the soak-time for a given trip is calculated as the start time of the current trip minus the end time of the previous trip. The maximum soak-time was capped at 1440 minutes (24 hours), and if a vehicle only had one trip in its trip-chain then it was given the maximum soak-time. The vehicle-specific start EFs were generated using the same platform (MOVES2010b) as the running emissions with a similar Montreal-specific context. EFs were generated in order to consider the effect of soak-time (8 bins), weather (2 seasons), vehicle type (2 types), vehicle model year (30 model years) and pollutant type, independently. Given that Montreal has significant seasonal variability and that meteorological conditions can have a significant effect on start emissions, we simulated EFs for both a summer (21.1 °C) and winter (-6.5 °C) scenario.

EFs for vehicle model years between 1978 and 2008 were generated for passenger cars and passenger trucks independently. Increased vehicle age is associated with increased start emissions (Figure 4A). A similar relationship is observed for all pollutants and atmospheric conditions considered. Vehicles manufactured before 1985 are associated with considerably higher start EFs.

Start EFs were developed for eight different soak-time bins in line with the eight operating modes used for start emission calculations in MOVES2010b. The eight operating modes corresponded to the following time bins (ranges in minutes): (1) 0-6; (2) 7-30; (3) 31-60; (4) 61-90; (5) 91-120; (6) 121-360; (7) 361-720; and (8) 721-1440. The relationship between soak times and EFs is nearly logarithmic (Figure 4B) and compares well to previous findings into the connection between start EFs and soak-times (Favez et al., 2009). Engine starts with soak-times greater than 12 hours (720 minutes) are considered to be cold-starts. Trip-level start emissions were estimated by assigning a start EF to each vehicle/trip based on its soak-time bin and vehicle characteristics (type and model year). The trip-level start emissions from the OD survey were then expanded to the full population based on survey expansion factors to produce total daily start emissions.



A) Start emission factors for vehicle types and model years under winter conditions



B) Start emission factors for passenger cars (model year 2000) as a function of soak time under winter conditions

FIGURE 4 Development of start emission factors (EF)

3.2 Testing the effects of model randomness and input variability

Following development of the integrated traffic assignment and emission model, a series of model runs were undertaken in order to evaluate the effects of input data precision on the final emission inventory. Our strategy was to identify specific input data aggregations that are commonly applied in practice in the evaluation of transport plans. These common aggregations include assumptions based on vehicle ownership, route selection, engine soak-times, etc. The goal was to then identify the ‘cost’ that such simplifications entail in terms of the final regional emission inventories. In addition, the effect of randomness we built into the model and which pertains to the vehicle allocation as well as path allocation processes is evaluated through multiple model runs leading to the generation of a standard error associated with every daily regional emission estimate.

A total of 7 sets of model inputs were varied; these include: 1) the inclusion or exclusion of start emissions; 2) ambient temperature (winter vs. summer); 3) level of speed aggregation extracted from the assignment model (average network speed, trip speed, link speed); 4) vehicle age distribution (average vehicle age in the province vs. real distribution obtained from registry data); 5) vehicle type distribution (assuming all passenger cars vs. real distribution obtained from registry data); 6) soak-time distribution (assuming all starts are cold starts, assuming that all starts are warm starts, using default soak time distributions, and deriving the real soak time distribution from trip start and end times); and 7) path selection (shortest path vs. stochastic assignment).

1) Inclusion of start emissions – Excess emissions during engine starts have been estimated to account for nearly 28-31 percent of total on-road volatile organic compounds (VOC) emissions under summer conditions (Borrego et al., 2004; Houk, 2004). Kioutsioukis et al. (2004) also found that cold start effects in relation to average trip lengths were one of the most important elements in accurate VOC estimates. While earlier emission inventories have accounted for starts through three different types of running EFs (cold-transient, hot-transient, and hot-stabilized), the new generation of emission simulators tend to separate running and start EFs. Therefore, in order to evaluate the effect of start emissions on daily regional emission inventories in isolation from other factors, we ran the model while including and excluding start emissions.

2) Season – For regions and urban areas with high temperature differentials between summer and winter, season becomes an important consideration in developing emission inventories. Emission rates, especially the excess from starts, tend to significantly change under winter conditions, with start contributions rising to over 50 percent of total on-road emissions (Houk, 2004). Therefore, total emissions (start and running) were evaluated under summer and winter conditions in order to assess the isolated effect of ambient temperature.

3) Travel speed – There are varying levels of detail with regard to travel speeds used in emission modeling. The simplest assumption involves applying the average network speed (daily or peak-period) to all trips (Ko et al., 2011; Brand and Preston, 2010). Certain models have increased the level of detail by assuming average trip speeds for emission calculations (Borge et al., 2012; Barla et al., 2011; Frank et al., 2000; Hao et al. 2000). Adding another layer of detail, the majority of emission models tend to use average link-speeds (Borge et al., 2012; Hao et al., 2010; Beckx et al., 2009a; Xia and Shao, 2005; Borrego et al., 2004; Anderson et al., 1996). The detail in travel speed has been repeatedly shown to be highly significant with regards to emission estimates. Smit (2006) concluded that congestion and its impact on travel speeds was the second most important element in HC estimates after vehicle kilometers travelled. Anderson et al. (1996) found that HC estimates increased by 56 percent when using congested versus free-flow speeds. Therefore, three different travel speed assumptions were tested. The first was the most detailed and involved estimating link emissions using the congested speed of every link in a trip's path. The second involved calculating an average trip speed for every OD trip using the trip length and time, and then estimating link emissions using this speed. The third speed assumption was to apply an hourly network-wide average speed to each link/trip. Despite recent research into the added value of incorporating mean speed distributions into link speed estimates (Smit et al., 2008), deterministic link speeds were used in all analysis due to limited traffic speed data at the regional level.

4) Vehicle age – Vehicle age is a significant factor in emission modeling, as it typically captures many elements related to emission generation. These include advancements in engine technology (increased efficiency, etc.), emission-capture technology (catalytic converters, etc.) as well as the effects of ageing and deterioration on vehicle performance. Previous studies have made three common assumptions: (1) all trips use the same vehicle age (Ko et al., 2011; Hao et al., 2000); (2) all trips are assigned a single EF representing the actual distribution of vehicles

(obtained from fleet registry data) (Borge et al., 2012; Xia and Shao, 2005; Borrego et al., 2004; Frank et al., 2000; Anderson et al., 1996); or (3) each trip is allocated a model year that is related to a unique EF (Barla et al., 2011; Hao et al., 2010; Brand and Preston, 2010). Therefore for this study two different vehicle age assumptions were tested. One involved creating a set of EFs from Montreal-specific fleet information for 30 different model years (1978-2008). When each trip was assigned a unique vehicle age, it was also associated with a single EF for that same model year. The second involved using an EF for one model year and assuming that all trips were made using that same vehicle age. The vehicle age that was used was the average for the Montreal fleet, which at the time of simulation (2008) was eight-years old (model year 2000). While this second scenario is not particularly common considering that most agencies have access to vehicle age distributions nowadays, it represents an extreme that has been employed in the past.

5) Vehicle type – A common misconception in operational emission modeling frameworks is to assume that all household vehicles are of the same type, often assumed to be passenger cars (Ko et al., 2011; Brand and Preston, 2010; Anderson et al., 1996). For this reason, we evaluated the effects of two different vehicle type distributions: a) the actual household vehicle types existing in Montreal including passenger cars (PC), and passenger trucks (PT), a category that covers light-duty trucks and sports utility vehicles; and b) the assumption that every trip was made using a passenger car.

6) Soak-time – The time between turning an engine off and its successful re-ignition is known as the vehicle soak-time, and is a primary determinant of start emissions. Previous generations of emission models employed default soak-time distributions in order to account for start emissions. Over a certain period, all the trips would be randomly allocated soak-times based on that hour's distribution. Nair et al. (2000) recommended replacing the default distribution with regional-specific data if possible. For this study, individual soak-times for each vehicle/trip were estimated based on the vehicle allocation algorithm. Three other soak-time configurations were also tested. The second and third involved the assumptions that every start was a cold start (largest soak-time bin), or a warm start (smallest soak-time bin). The fourth configuration is based on the assumption that all vehicles in the survey data are randomly assigned a soak-time based on default cumulative distributions from MOBILE6 (USEPA, 1998). The MOBILE series is the previous generation of emission models developed by the USEPA, with MOVES2010b being its successor. Given that numerous studies have used either MOBILE6 (Hao et al., 2010)

or a previous generation of the MOBILE model (Frank et al., 2000; Anderson et al., 1996), it is interesting to assess its impact compared to emission estimates generated from locally derived soak-times. In the Montreal case, the differences across several hours seem to be relatively consistent between the MOBILE6 distribution and the travel diary distribution (Figure 5), however the overall differences between the two distributions are fairly large. The travel-diary distribution favours larger soak-times (greater than 12 hours), whereas the MOBILE6 distribution has a more relatively balanced range. Initially, this suggests that the MOBILE6 soak-time distribution might underestimate the total start emissions given that the soak-times are smaller on average.

7) Path selection – Not to be confused with path allocation (where vehicles are allocated to different paths linking their origin and destination in the emissions processor), path selection involves the choice of path used for any trip between an OD pair that is determined by the assignment type employed in the regional traffic model. Various assignment types used in conjunction with emission models include: (1) shortest path (Barla et al., 2011); (2) deterministic user equilibrium (Beckx et al., 2009a; Borrego et al., 2004; Anderson et al., 1996); (3) stochastic user equilibrium (Hao et al., 2010); and (4) dynamic user equilibrium (Borge et al., 2012; Xia and Shao, 2005). Alternatively, VKT data are sometimes gathered directly from travel surveys (Ko et al., 2011; Brand and Preston, 2010; Frank et al., 2000; Hao et al., 2000). For computational reasons, both deterministic and dynamic user equilibrium assignments were ignored in this study, resulting in two path selection algorithms being tested. One involves a stochastic path selection, which randomly allocates a trip to a path within a probability-based path set for every OD pair. The other involves the simplification that all trips select the shortest path.

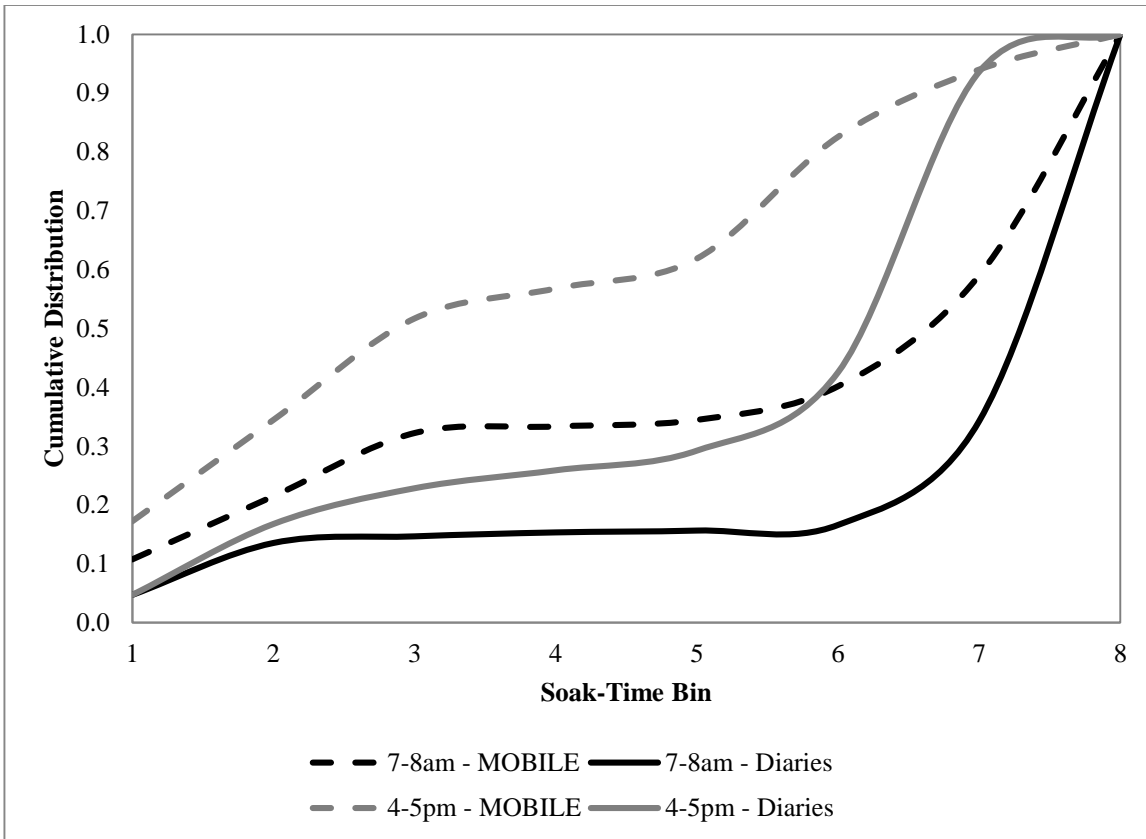


FIGURE 5 Soak-time comparison between local travel diary information and MOBILE6 defaults

The final element of interest in the framework is the variance within the emission processor. Assuming all other inputs remain constant, the variance within the emission output is caused by randomness in both the vehicle allocation and the path allocation steps. The vehicle allocation step entails randomly assigning each vehicle a model year and type based on the cumulative vehicle fleet distribution characterizing the household’s residential zone. The cumulative distribution function was created using vehicle registry data, broken down into number of vehicles owned by type and model year. Therefore, the individual vehicles change between model runs, however the aggregate makeup of each zone remains consistent with the actual fleet distribution. Meanwhile, the path allocation step entails randomly assigning a trip to a certain path between its origin and destination TAZs based on a cumulative path volume distribution function. The cumulative distribution function is created for every path set and is based on the volumes assigned to each path from the regional traffic assignment model. Therefore, individuals will have a higher chance of taking the most popular (ie. most congested) paths, yet lengthier alternatives are also available.

In order to account for this variance, the model was run under three different vehicle allocation simulations as well as three different path allocations combined to create three model iterations. We use the same path file generated by the regional traffic assignment model when varying the path allocations. These three iterations will form the basis for error estimation, and every input (belonging to the seven categories presented) is evaluated three times thus leading to three values for the total emissions output. The three values are averaged and a standard error (SE) is calculated. Figure 6 details the process used in testing the seven sets of data input. The results are presented in the same format with the mean and SE from the three iterations for each inventory estimated.

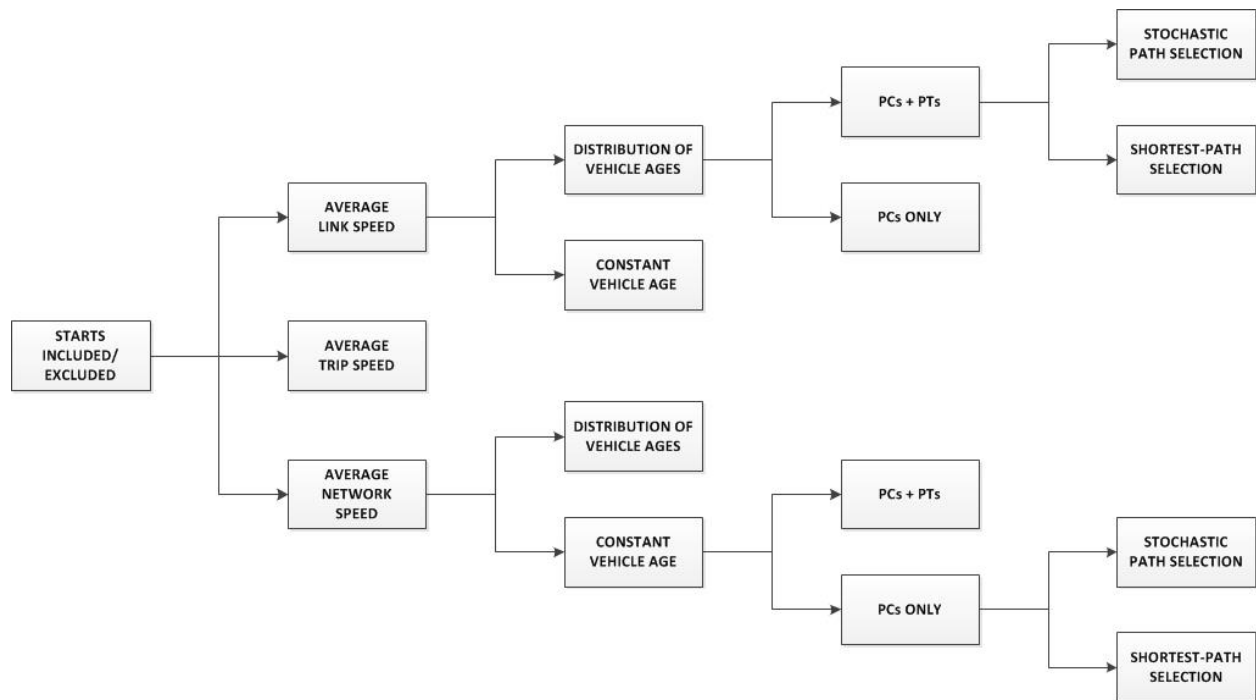


FIGURE 6 Flowchart of changes to data inputs and model simulations

(Note: the tree for average trip speed is similar to those for average link/network speeds yet is omitted for clarity)

4. RESULTS AND DISCUSSION

4.1 Validation of base-case model results

The main output of the integrated model that was validated includes traffic volumes on selected links, and these were compared with count data for the same time period. The hourly traffic counts used for volume validation were collected on bridges and major arterials, and represented

averages over five days of data collection. The hourly correlation between simulated volumes and field counts at 160 data points ranged from 0.62 to 0.88 over the 24-hr period.

4.2 Variability analysis

Results of the data input permutations that were tested are presented in the two trees that show daily regional emission estimates in tons per day, along with standard error values as they propagate across various input aggregations (Figure 7). Additionally, a summary of results is presented in Table 1 whereby the regional emission estimates are ordered in terms of percent change from base-case.

The two trees in Figure 7 serve to differentiate two sets of emission models with one involving a common omission: start-emissions. The figure shows an estimate that ignores start emissions (bottom tree) versus a more accurate one that includes starts (top tree). The uppermost branch of the top tree reveals the most detailed emission estimate of about 18 tons per day (0.36 grams per VKT), a quantity comparable to previously estimated daily HC inventories for similar metropolitan areas (Hao et al., 2010; Anderson et al., 1996). Meanwhile, it is important to note that the standard error of all estimates is due to randomness in the vehicle and path allocation steps and accounts for at most 0.36 percent of the daily regional emission estimate. This means that the variance in allocating vehicles to individual trips based on zone-level registry data is insignificant given the size of the region and the fact that we are only looking at total daily emissions. More specifically, vehicle ownership trends (type and age) at the zonal level are respected during each iteration resulting in relatively consistent emission estimates, even though vehicle allocation at the individual level may be varying drastically. The same conclusion can be reached for the randomness due to the path allocation process. Individuals are likely taking different paths over different iterations, however at the aggregate level the variance is small. These results are contrasted with the substantial inaccuracy when solely making assumptions based on vehicle ownership data. For instance, estimates involving a simplification in vehicle age actually result in overestimations of 0.24-5.73 tons per day. Larger overestimations occur in scenarios where start emissions are included, owing to the fact that they are highly dependent on vehicle age (Figure 4A).

The impact of various assumptions in combination was tested, for instance, the lowest branch of the lower tree shows the multiplying effect of simplifications regarding start-exclusion,

speed aggregation, vehicle type and age, and path selection. Together, those impacts lead to an underestimation of 1.69 tons per day compared to the detailed estimate excluding starts, and an underestimation of 13.7 tons per day compared to the most accurate estimate including starts.

Analyzing the effect that starts have on daily estimates, it is clear that ignoring start emissions will result in significant underestimations. Daily regional emission estimates increase threefold when including starts under summer conditions and this increase can rise to as much as eight times if winter conditions are considered (Figure 8A). These results show that starts alone can contribute 67 to 86 percent of total daily HC emissions. These results also highlight two important points with regards to season, in that: 1) substantial differences in weather conditions do not have a significant effect on running emissions, at least for HCs; and 2) that any impact seasonality has on running estimates pales in comparison to its pivotal role in start estimates.

The effect of using average trip speeds in emission calculations reduces the total daily estimate by a small amount from the baseline, which employs more accurate average link speed data (Figure 8B). Applying the more simplistic assumption that all drivers travel at average hourly network speeds results in a similar-sized underestimation. The fact that both simplifications result in underestimation comes as no surprise given that congestion effects tend to be lost in speed aggregation. However, the impacts are much smaller than anticipated. This is possibly caused by aggregating hourly running emissions, which typically vary significantly in peak versus non-peak traffic, into one daily estimate. The reasons for this are likely twofold: 1) the EFs are not as sensitive to speeds since the average network speeds observed from the traffic assignment model are located on a flatter segment of the speed-EF curves, and 2) the congestion algorithm employed by the traffic model is underestimating the impact of congestion on link speeds.

Regarding the impacts of vehicle age, the assumption that all trips are made with an average model year (8 years old) results in a consistent overestimation across each of the speed aggregation scenarios. The error bars in Figure 8B represent the range of estimate totals from model randomness. The smaller error bars of the constant vehicle age scenarios clearly indicate that the bulk of model randomness lies in the vehicle allocation step. While this scenario is perhaps less plausible than the others tested, the effect that it illustrates is nevertheless interesting when combined with input data assumptions on vehicle type.

If the effect of vehicle type in combination with vehicle age is then isolated, the resulting outputs show that assuming all trips are made with passenger cars leads to a large underestimation in emission estimates (Figure 8C). The underestimate from this assumption is 1.21-1.33 tons per day and is consistent whether applying the full fleet age distribution or assuming an average vehicle age. When including starts the underestimation is 3.97-6.34 tons per day, a percent decrease similar to that seen when ignoring starts. While the relationships between vehicle age/type and emission rates have been well established through previous research, this finding is interesting nevertheless in that it parallels previous findings even when applied in combination with other assumptions at a regional scale.

Additionally, the type of path selection algorithm employed in the transportation model has a significant impact on the daily emission estimate. The assumption that all trips were taken on the shortest path results in a small drop in total regional emissions from the baseline, in which the path set is created through a stochastic distribution (Figure 7). It is important to note that these results are based on 24-hour emissions (estimated from running 24 hourly assignments) and therefore include peak and non-peak periods. The difference between the emissions resulting from shortest path selection and stochastic path selection vary by time of day. During congested hours, we observe that shortest path emissions are lower than stochastic emissions by 44-57 kg per hour. During uncongested hours and at night, the difference drops to 1-2 kg per hour. In smaller cities or cities with lower levels of congestion than Montreal, this difference between stochastic and shortest path selection is likely to decrease. However, for cities with high congestion levels, having a stochastic path set for assignment purposes is vital in order to capture the impact of trips that take lengthier, but less congested, route alternatives.

The results of analyzing the effects of soak time on starts can be seen in Figure 8D. The error bars shown represent the range in values from different vehicle allocations. The soak-time distribution estimated from detailed travel diaries resulted in 12.00 tons of start-based HCs. Assuming all trips began under warm-start conditions resulted in a drastic decrease of 11.3 tons per day, whereas the total start emissions rose by 4.0 tons per day under the assumption that all trips were cold-starts. Given that start emissions are responsible for the majority of total HC emissions, it is important that soak time distributions are accurate due to the wide range of values seen in Figure 8D. Another finding of note is that the total daily start emission estimates are relatively similar when using the local soak-time distributions derived from travel diaries versus

the default distribution in MOBILE6. Although the contribution from starts was higher under the soak-times derived from travel diaries (12 tons per day vs. 11.5 tons per day), the differences are smaller than expected. Note however that when hourly start emissions are compared, it is evident that the discrepancy in estimates is primarily due to differences in the morning and afternoon peak-periods.

Meanwhile, the modeling results for NO_x are qualitatively comparable to those for HC. Both atmospheric conditions and vehicle age had similarly significant impacts on regional NO_x emission estimates, even to a greater extent than with HC estimates (note: start emissions were omitted due to their relative insignificance in NO_x emission rates). At the same time, input data assumptions on vehicle type, speed aggregation and path selection all resulted in decreases to estimate accuracy, paralleling findings from the HC analysis.

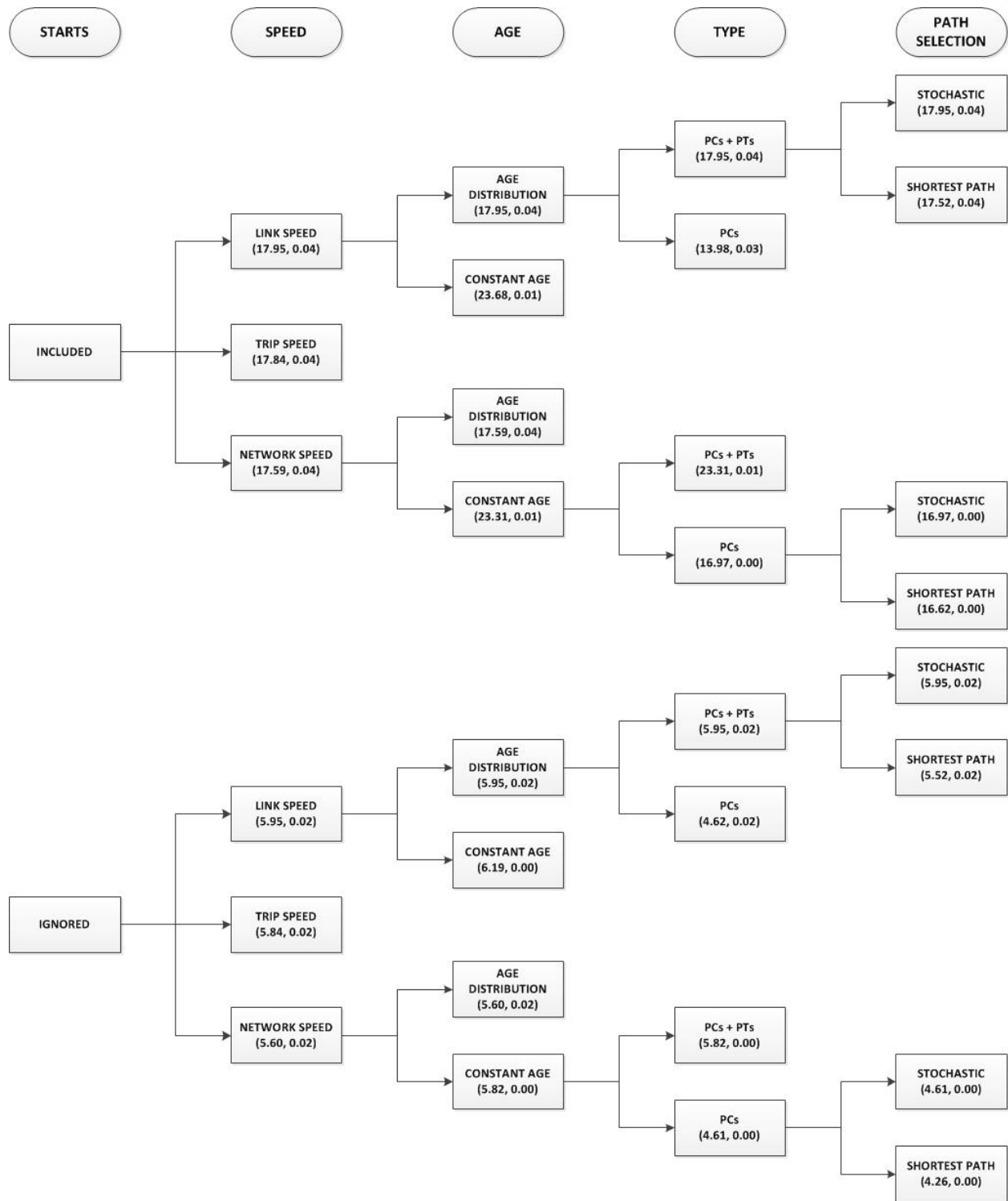


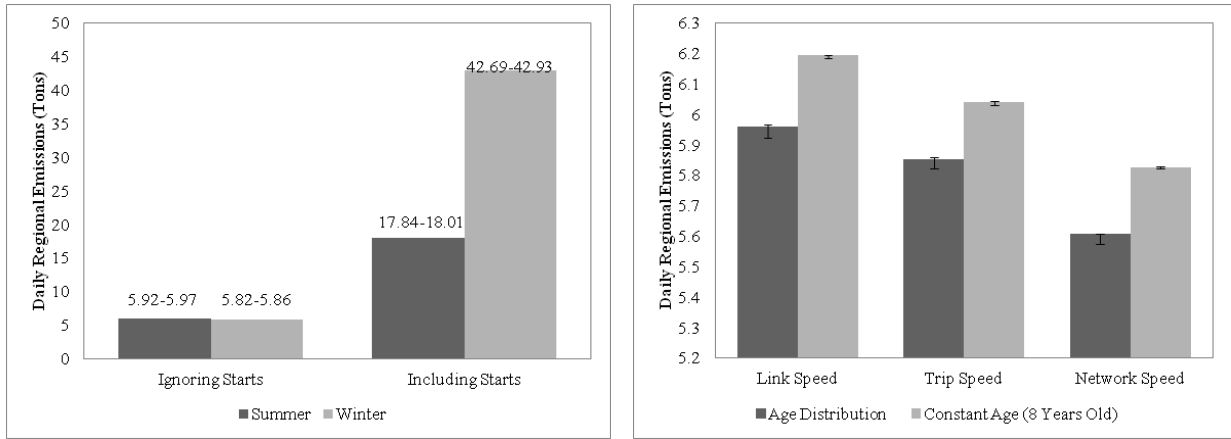
FIGURE 7 Sensitivity results for daily regional emission estimates (values in parentheses include daily mean HC amount followed by its standard error)

TABLE 1 Summary of regional emissions across input distributions considered

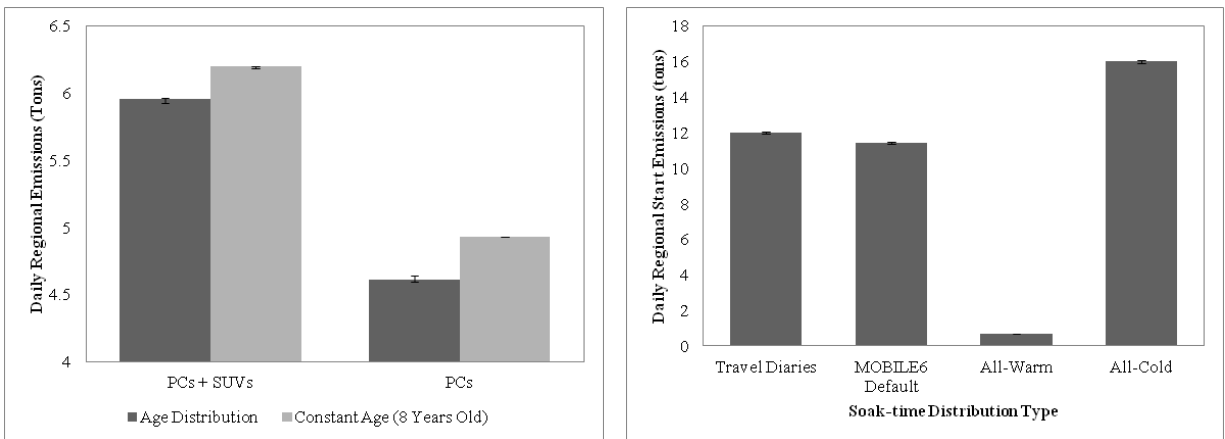
| Atmospheric Conditions | Start Emissions | Speed Aggregation | Vehicle Age | Vehicle Type | Path Selection | Mean HC Emissions (tons/day) | % Change From Base |
|------------------------|-----------------|-------------------|--------------|--------------|----------------|------------------------------|--------------------|
| Summer | Included | Link Speed | Distribution | PCs+PTs | Stochastic | 17.95 | ^a |
| | | Link Speed | Distribution | PCs | Stochastic | 13.98 | -22.1% |
| | | Network Speed | Constant Age | PCs | Shortest | 16.62 | -7.4% |
| | | Network Speed | Constant Age | PCs | Stochastic | 16.97 | -5.5% |
| | | Link Speed | Distribution | PCs+PTs | Shortest | 17.52 | -2.4% |
| | | Network Speed | Distribution | PCs+PTs | Stochastic | 17.59 | -2.0% |
| | | Trip Speed | Distribution | PCs+PTs | Stochastic | 17.84 | -0.6% |
| | | Network Speed | Constant Age | PCs+PTs | Stochastic | 23.31 | 29.9% |
| | Ignored | Link Speed | Constant Age | PCs+PTs | Stochastic | 23.68 | 31.9% |
| | | Link Speed | Distribution | PCs+PTs | Stochastic | 5.95 | ^b |
| | | Network Speed | Constant Age | PCs | Shortest | 4.26 | -28.4% |
| | | Network Speed | Constant Age | PCs | Stochastic | 4.61 | -22.5% |
| | | Link Speed | Distribution | PCs | Stochastic | 4.62 | -22.4% |
| | | Link Speed | Distribution | PCs+PTs | Shortest | 5.52 | -7.2% |
| | | Network Speed | Distribution | PCs+PTs | Stochastic | 5.60 | -5.9% |
| Network Speed | | Constant Age | PCs+PTs | Stochastic | 5.82 | -2.2% | |
| Winter | Included | Link Speed | Distribution | PCs+PTs | Stochastic | 42.83 | - |
| | Ignored | Link Speed | Distribution | PCs+PTs | Stochastic | 5.84 | - |

^aThis 'base' emission estimate is used for comparison with all other 'summer' + 'starts included' scenarios

^bThis 'base' emission estimate is used for comparison with all other 'summer' + 'starts ignored' scenarios



A) Daily emissions ignoring and including starts and season B) Daily emissions for different speed aggregations and vehicle ages (ignoring starts)



C) Daily emissions for different ages and types (ignoring starts) D) Daily start emissions for different soak-times

FIGURE 8 Illustrating the effects of selected variables on regional emissions

5. CONCLUSION

The level of accuracy in travel behaviour data is highly linked to accurate emission inventories (Kioutsioukis et al., 2004). However, the question remains as to which inputs contribute the most to the accuracy of emission estimates. To address this question, an integrated traffic and emission model was developed in order to estimate daily regional emissions for the Montreal metropolitan region. The detailed model was set as the baseline, and then seven different data inputs were altered in isolation, as well as several in combination, to assess their impact on daily emission inventories. Overall, start emissions had the largest impact. Even under summer conditions, start

emissions are estimated to make up two thirds of total on-road hydrocarbon emissions. The impact of starts would likely be less if inventories were being made for other pollutants such as nitrogen oxides. Vehicle registry data is the next element with the highest impact, specifically data on vehicle type distribution. SUVs and light-duty trucks emit more running and start emissions than passenger cars with similar model years, and so regions with high levels of passenger truck use certainly need to account for them. Meanwhile, speed aggregation and path selection type both have significant impacts on regional emission estimates, however nowhere close to the extent of including starts or accurate fleet composition data. However, the results on speed aggregation likely mask their true impact to a certain extent, especially given that deterministic link speeds were used instead of mean speed distributions (Smit et al., 2008). It is likely that this effect would also have a crossover impact on the significance of path selection type. Aside from that, our results do seem to support the argument that vehicle distance travelled remains a significant element in determining trip emissions at the regional level. For practitioners, reaching the modeling optimum that balances input accuracy with model accuracy likely means investing resources in start emissions and accurate vehicle ownership data.

It is important to note that our sensitivity results are pollutant-specific. Other major traffic pollutants (e.g. particulate matter, sulfur oxides) and greenhouse gases are likely to be impacted differently by the various input parameters as their emission rates react differently under congestion, ignition conditions, atmospheric conditions, etc. Furthermore, the HC estimates in this study do not incorporate evaporative emissions, a significant portion of total urban emissions. Evaporative emissions are strongly influenced by ambient conditions, and have been estimated to account for up to 15% of total urban HC emissions under summer conditions (Hao et al., 2010). Meanwhile at a more macroscopic level, owing to the complexity and resource-intensiveness of developing integrated transport and emissions models with specific transport and emission simulators, the impact of model/software choice was not assessed in our study. This omission extends to the analysis of various congestion algorithms used in regional traffic modeling.

Further work should initially focus on including evaporative emissions within the analysis framework. From there, it would be interesting to extend the sensitivity analysis to other major traffic pollutants. Additionally, the impact of using different traffic simulations/algorithms

and emission models on regional emission inventories should be studied, although the time and cost investments of such an endeavour are significant.

REFERENCES

- Agence Metropolitaine de Transport (AMT) (2008). La mobilite des personnes dans la region de Montreal: Faits Saillants. *Enquete Origine-Destination 2008*.
- Anderson, W.P., Kanaroglou, P.S., Miller, E.J., and Buliung, R.N. (1996). Simulated automobile emissions in an integrated urban model. *Transportation Research Record*, 1520, 71-80.
- Barla, P., Miranda-Moreno, L.F., and Lee-Gosselin, M. (2011). Urban travel CO2 emissions and land use: A case study for Quebec City. *Transportation Research Part D*, 16, 423-428.
- Beckx, C., Int Panis, L., Vankerom, J., Janssens, D., Wets, G., and Arentze, T. (2009a). An integrated activity-based modeling framework to assess vehicle emissions: approach and application. *Environment and Planning B: Planning and Design*, 36, 1086-1102.
- Beckx, C., Int Panis, L., Uljee, I., Arentze, T., Janssens, D., and Wets, G. (2009b). Disaggregation of nation-wide dynamic population exposure estimates in the Netherlands: applications of activity-based transport models. *Atmospheric Environment*, 43 (34), 5454-5462.
- Borge, R., de Miguel, I., de la Paz, D., Lumbreras, J., Perez, J., and Rodriguez, E. (2012). Comparison of road traffic emission models in Madrid (Spain). *Atmospheric Environment*, 62, 461-471.
- Borrego, C., Tchepel, O., Salmim, L., Amorim, J. H., Costa, A. M., and Janko, J. (2004). Integrated modeling of road traffic emissions: application to Lisbon air quality management. *Cybernetics and Systems: An International Journal*, 35, 535-548.
- Brand, C., and Preston, J.M. (2010). '60-20 emission' - The unequal distribution of greenhouse gas emissions from personal, non-business travel in the UK. *Transport Policy*, 17, 9-19.
- Crouse, D.L., Goldberg, M.S., and Ross, N.A. (2009). A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in Montreal, Canada. *Atmospheric Environment*, 43, 5075-5084.
- Favez, J.-Y., Weilenmann, M., and Stilli, J. (2009). Cold start emissions as a function of engine stop time: evolution over the last 10 years. *Atmospheric Environment*, 43, 996-1007.

- Frank, L.D., Stone Jr., B., and Bachman, W. (2000). Linking land use with household vehicle emissions in the central puget sound: methodological framework and findings. *Transportation Research Part D* , 173-196.
- Hao, J., He, D., Wu, Y., Fu, L., and He, K. (2000). A study of the emission and concentration distribution of vehicular pollutants in the urban area of Beijing. *Atmospheric Environment*, 34, 453-465.
- Hao, J.Y., Hatzopoulou, M., and Miller, E.J. (2010). Integrating an Activity-Based Travel Demand Model with Dynamic Traffic Assignment and Emission Models. *Transportation Research Record: Journal of the Transportation Research Board* , 2176, 1-13.
- Hatzopoulou, M., & Miller, E. J. (2010). Linking an activity-based travel demand model with traffic emission and dispersion models: Transport's contribution to air pollution in Toronto. *Transportation Research Part D: Transport and Environment* , 15 (6), 315-325.
- Houk, J. (2004). Making use of MOBILE6's capabilities for modeling start emissions. *Proceedings of the A and WMA's 97th Annual Conference and Exhibition (1884)*, 5115-5130.
- Hülsmann, F., Gerike, R., and Ketzel, M. (2014). Modelling traffic and air pollution in an integrated approach – the case of Munich. In Press. *Urban Climate*.
- Karppinen, A., Kukkonen, J., Elolahde, T., Konttinen, M., Koskentalo, T., and Rantakrans, E. (2000). A modelling system for predicting urban air pollution: model description and applications for the Helsinki metropolitan area. *Atmospheric Environment*, 34, 3723-3733.
- Kickhöfer, B., and Nagel, K. (2013). Towards high-resolution first-best air pollution tolls. *Networks and Spatial Economics* 13.
- Kioutsioukis, I., Tarantola, S., Saltelli, A., and Gatelli, D. (2004). Uncertainty and global sensitivity analysis of road transport emission estimates. *Atmospheric Environment*, 38, 6609-6620.
- Ko, J., Park, D., Lim, H., and Hwang, I. (2011). Who produces the most CO2 emissions for trips in the Seoul metropolis area? *Transportation Research Part D* , 16, 358-364.
- Lefebvre, W., Degrawe, B., Beckx, C., Vanhulsel, M., Kochan, B., Bellemans, T., Janssens, D., Wets, G., Janssen, S., de Vlieger, I., Int Panis, L., and Dhondt, S. (2013). Presentation

- and evaluation of an integrated model chain to respond to traffic- and health-related policy questions. *Environmental Modelling & Software* 40, 160-170.
- Mensink, C., de Vlieger, I., and Nys, J. (2000). An urban transport emission model for the Antwerp area. *Atmospheric Environment*, 34, 4595-4602.
- Nair, H.S., Bhat, C.R., and Kelly, R.J. (2000). Modeling soak-time distribution of trips for mobile source emissions forecasting: techniques and applications. *Transportation Research Record*, 1750, 24-31.
- Nejadkoorki, F., Nicholson, K., Lake, I., and Davies, T. (2008). An approach for modelling CO₂ emissions from road traffic in urban areas. *Science of the Total Environment*, 406, 269-278.
- Sider, T., Alam, A., Zukari, M., Dugum, H., Goldstein, N., Eluru, N., and Hatzopoulou, M. (2013). Land-use and socio-economics as determinants of traffic emissions and individual exposure to air pollution. *Journal of Transport Geography*, 33, 230-239.
- Smit, R., (2006). An examination of congestion in road traffic emission models and their application to urban road networks, PhD dissertation, Griffith University, Brisbane, Australia.
- Smit, R., Brown, A.L., and Chan, Y.C. (2008). Do air pollution emissions and fuel consumption models for roadways include the effects of congestion in the roadway traffic flow? *Environmental Modelling and Software*, 23, 1262-1270.
- Smit, R., Poelman, M. and Schrijver, J. (2008). Improved road traffic emission inventories by adding mean speed distributions, *Atmospheric Environment*, 42 (5), 916-926.
- Smit, R., Ntziachristos, L., and Boulter, P. (2010). Validation of road vehicle and traffic emission models – a review and meta-analysis. *Atmospheric Environment*, 44, 2943-2953.
- Statistics Canada. (2011). *2011 Census of Population*. Ottawa: Statistics Canada.
- United States Environmental Protection Agency (USEPA) (1998). USEPA Assessment and Modeling Division Report on ‘Modeling Hourly Diurnal Emissions and Interrupted Diurnal Emissions based on Real-Time Data’.
- Waked, A., and Afif, C. (2012). Emissions of air pollutants from road transport in Lebanon and other countries in the Middle East region. *Atmospheric Environment*, 61, 446-452.

Xia, L., and Shao, Y. (2005). Modelling of traffic flow and air pollution emission with application to Hong Kong Island. *Environmental Modelling and Software*, 20, 1175-1188.