IMPROVING EMISSIONS ESTIMATES IN THE MEXICO CITY METROPOLITAN AREA

Julia A. Gamas
Massachusetts Institute of Technology
Cambridge, Massachusetts, United States of America, jgamas@mit.edu

Dan Amano
Massachusetts Institute of Technology
Cambridge, Massachusetts, United States of America, danamano@mit.edu

William Anderson
Boston University
Boston, Massachusetts, United States of America, bander@bu.edu

Raymond Dezzani
Boston University
Boston, Massachusetts, United States of America, dezzani@crsa.bu.edu

Joseph Sussman
Massachusetts Institute of Technology
Cambridge, Massachusetts, United States of America, sussman@mit.edu

ABSTRACT

The serious air pollution problem in the Mexico City Metropolitan Area has led policy makers to ask what measures could be taken to improve air quality. There is data scarcity regarding transportation sources of air pollutants; current estimates of how much pollution is generated by transport are based on the assumption that traffic flows at an average speed of 32 kilometers per hour. This assumption is problematic because amounts of emissions of different pollutants vary with speeds. Therefore, it is important to know the actual traveling speeds on the road network, in order to get the best possible estimates of emissions. An urban transportation model was developed for the MCMA, to find congestion levels, travel times and emissions. Using available information from the 1994 Origin Destination Survey, as well as Census data, relevant parameters were found. Then these were used to model a future scenario for 2025. When speeds are close to 32 km/h, emissions are almost the same as when the 32 km/h average speed assumption is used. However, when travel speeds are low, emissions are underestimated. Because of the data scarcity problem during its construction, the model should not be used for policy analysis. Yet, it serves to underscore the importance of such a tool of analysis.

KEY WORD: Urban transportation modeling, emissions estimates, travel demand

1. INTRODUCTION[i]

Transportation is a major source of air pollution in the Mexico City Metropolitan Area (MCMA). The contribution of transportation related emissions to overall emissions is very high and has increased over the years (see Molina 2002 and SMA 2000). As population increases and the economy grows, demand for travel continues to increase. As households become wealthier they will also purchase more private vehicles. As a consequence, congestion on the road network and pollution levels will also increase.

In order to decide on measures that will help reduce air pollution, it is important to know in what way traffic congestion contributes to such emissions. A simple approach has been to calculate emissions by using the following formula for a given type of vehicle and pollutant:

[illegible] (1)

where VKT is vehicle kilometers traveled per vehicle, and EF is the emissions factor measured in grams per VKT. In other words, the level of activity (vehicle kilometers traveled) by transportation mode is multiplied by an emissions factor (grams/vehicle kilometer) that reflects the characteristics of the vehicle stock, ambient temperature and driving cycle characteristics. Driving cycle refers to the sequence of accelerations, decelerations, cruising and idling in a typical vehicle trip. A vehicle that travels ten kilometers in congested conditions, where there are frequent starts, stops and idling, will produce significantly more emissions of most pollutants than the same vehicle driving the same distance at a steady, moderate speed.

Since detailed information on driving cycles is not generally available, the average speed of travel is often used as a proxy for traffic conditions. Low average speeds are assumed to reflect stop-and-go traffic conditions while high average speeds are assumed to reflect highway driving. The practice thus far (Molina, 2004) in generating emissions estimates for the MCMA has been to estimate total vehicle kilometers per hour (VKT) for different categories of vehicles and multiply those numbers by emissions factors based on assumption of a single average speed of 32 kilometers per hour (km/h).

The weakness of this approach is that choosing an average speed does not reflect the variability in driving cycles with different pollutant profiles. For example, an average speed of 32 km/h may arise from a large number of cars traveling at low speeds and a small number traveling on highways at very high speeds. If this is the case, emissions are systematically underestimated because slow and fast driving cycles produce more emissions than intermediate driving cycles.

The US EPA's MOBILE series of emissions models have been applied in Mexico for the creation of emissions inventories of cities at the border (Radian International, 1996). Figure 1 shows the relationship between average speed and emissions factors generated by that model for Mexican cities at the border. As we can see in figure 1, different pollutants have different speed profiles. Note that two of the three pollutants have the highest emissions factors at low speeds, but one (NOx) has higher emissions factor at high speeds. Thus different driving cycles are intensive in different pollutants. At low speeds, more carbon monoxide and total organic gases are produced, as compared to their production at medium speeds (such as 32 km per hour). At low speeds the emissions of NO_X will be lower than emissions of the same pollutant at higher speeds.

Furthermore, the combination of the emissions profile with distance traveled will make the emissions pattern more complex. Even if it is assumed that all cars travel at the same speed, there will be lower emissions per kilometer if a given vehicle has traveled a short distance as compared to a trip by the same vehicle that covers a long distance. This is because the car engine will be running for a longer period of time during a long distance trip, as opposed to a short distance one.

In a large metropolitan area as the MCMA, congested conditions are frequent. These congested conditions occur when demand exceeds the capacity of the roadway. The variability in speeds and distances per trip are needed to estimate the emissions caused by traffic more precisely. However, so far there has been little available information regarding these speeds and distances. There also has been an absence of any means of modeling emissions using this information.

For this reason, we have developed an Urban Transportation Model (UTM) for the MCMA that will allow us to estimate the speeds, distances and emissions more accurately than when using the assumption of average speeds of 32 km per hour. The results from the model are the congestion levels and speeds on different parts of the road network. This information can then be used together with an emissions model to find the emissions levels for traffic, given our assumptions made about the road network and land use patterns. There was a data scarcity problem in the construction of the model. Even so, the model underscores why it is important to have it as an analytical tool. Improvements to the model are ongoing. They will render the model a useful tool for policy analysis in the future.

The UTM for the MCMA is presented in the section 2. A discussion and general conclusions are presented in the last section.

2. THE UTM FOR THE MCMA

2.1. Introduction to the Urban Transportation Model

In the UTM socio-demographic information about the population is used to explain its decisions to travel (travel demand) and information about the roadway network is used to explain the route taken by each traveler (travel supply). The ultimate results obtained from such a model are travel speeds and distances of those trips. In travel demand individual travel decisions are analyzed as a function of their economic and demographic characteristics (income levels and age, for example). In travel supply trip routes and congestion patterns are modeled on the network (as a function of the road infrastructure conditions).

The UTM is broken down into several stages. Travel demand estimation is conducted in the first three stages: trip generation, trip distribution and mode choice. Travel supply modeling, in which trips are assigned a particular route on the roadway network, is conducted in a fourth stage. The first four stages are also known as a Four Stage Model (FSM). Congestion and emissions are modeled in

Stage 5 of the UTM, by taking the travel speeds and using them together with emissions factors, as described in the introduction. A model is developed for each stage. Each stage uses as inputs the output of the previous stage.

Land use and road infrastructure availability are taken as exogenous. In reality there are feedback loops between travel and the development of land use and road infrastructure. An improvement in road infrastructure will improve travel conditions on that road. However, such conditions may worsen as travelers adjust to these new improved conditions by switching from another mode to driving, changing their schedules to travel on the new road, travel on this road when they hadn't done so before, or change their route to take the new un-congested route. In the short term, the road will soon fill up and may reach previously congested levels. This phenomenon of induced travel demand, in the short term, may consequently reduce travel speeds on the new improved road.

Additionally, in the long run, improvements in road infrastructure and changes in land use may lead to more economic growth, will lead to an increase in wealth, which together with population growth may result in more trips. Once the UTM model is operational, land use changes can be modeled to see their effect on travel, but it will not be possible to see the opposite effect (of travel on land use). These feedback loops should be added as extensions to the model in future research.

An UTM was constructed for the MCMA. Travel demand models were constructed using information about the population of the city (from Population Census information for 1990 and 2000), information about travel behavior (from the 1994 Origin Destination Survey), and data about trip attractors (1994 Economic Census). INEGI was the source of all of this information (see INEGI 1990 to 2000). However, there was missing information that had to be assumed or estimated. For travel supply, there was no information. Specifically, a digitized road network for the MCMA that could be used in TransCad modeling software was unavailable and was therefore constructed. Additional information about the road network itself (such as number of roads, classification into major or minor roads, capacity, length, traffic counts) was also necessary. Additionally, a layer of zone (*distrito)* boundaries had to be digitized in manually as well.

In each step, the data was used to estimate a series of parameters. Once these parameters were estimated for 1994, they were then used with 2025 future scenario information to illustrate the use of the model. Because of the data scarcity involved in its construction, the model was not used to conduct policy analysis, but only for purposes of establishing its utility in improving emissions estimates for the MCMA.

2.2. Stage 1: Trip Generation

In this stage, models were constructed to calculate parameters by type of trip. Then these parameters were used with data from a future scenario to calculate how many trips would be generated by and attracted to each zone of the MCMA?

For each trip type, the relationship between the total number of trips entering and leaving a zone with the variables responsible for producing and attracting those trips was established. Socio-demographic variables were used for trip generation out of each zone, and economic variables were used for trip attractions into each zone.

Total trips leaving and entering each zone or *distrito* were recorded during the 1994 ODS. These trips were broken down into the following categories: 1) trips to work, 2) trips to go shopping, 3) trips to school, 4) trips for entertainment and recreation, 5) work-related trips, 6) trips to pick someone up or drop someone off, 7) trips to lunch, 8) other trips, 9) trips going back home.

The relationship between explanatory variables and total trips was modeled by using a series of regression equations where the dependent variable was the total number of trips per distrito for a given trip category, as reported by the 1994 ODS and the independent variables were the trip generator and attractor variables per distrito. The general specification for each model was of the following form:

$$T_i = \alpha + \beta_1 income + \beta_2 population + ... + \upsilon \tag{2}$$

where _= _W_ + _, _ ~ (0,_I_n), W is a matrix of adjacency, _ is a statistical parameter and _ is a "white noise" residual. (Anselin, 1988). A summary of the model results is found in Gamas Buentello (2003).

In general, explanatory variables included, for trip productions: incomes per household, number of households, total population that works, population that goes to school, population density. In general, the higher the household income, the more trips can be taken of different types. If the number of households, total population and number of people who work are high within a zone, then that zone will generate more trips. For trip attractions, the explanatory variables were: total number of jobs and numbers of services, commercial, manufactures and establishments of different kinds. The larger the value of these variables, the more trips would be attracted to a zone.

2.3. Stage 2: Trip Distribution

In this stage, the relationship between the zones that generate trips and the zones that attract them is established. A parameter of friction or impedance between zones is estimated. It can then be used with data from a future scenario to find the origin and destination of each trip. For the estimation of this parameter, we needed an Origin Destination (OD) matrix. Each entry of such a matrix would contain the number of trips going from one zone to another, for each pair of zones in the metropolitan area. If used in conjunction with the number of total trips entering each zone (vector of OD matrix column sums), the total trips leaving each zone (vector of OD matrix row sums) and the distance between zones, the relevant parameter could be found by using a gravity model (Fotheringham and O'Kelly 1989). The parameter in question is _ in the following gravity model equation:

$$T_{ij} = A_i O_i B_j D_j d_{ij}^{-\beta} \tag{3}$$

where O_i represents the number of trips originating in zone i, D_j represents the number of trips entering j. T_{ij} is the total trips going from zone i to zone j. The variable d_{ij} represents the distance between zone i and zone j. Ai and Bj are "balancing terms":
$Ai = 1 / \Sigma_j B_j D_j d_{ij}^{-\beta}$, and $Bj = 1 / \Sigma_j A_j D_j d_{ij}^{-\beta}$.

For the case of Mexico City, however, a complete OD matrix was not available. Therefore, the gravity model parameter was estimated using the available fifty entries from the OD matrix. Therefore, the resulting parameter could include a bias from assuming the fifty available observations to be representative of the rest of the city[ii]. The resulting parameter was equal to –0.363 (see Gamas Buentello 2003). This parameter can be interpreted as the cost of traveling from one zone to another. If the cost or distance between two given zones were to increase, the total number of trips between these two zones would decrease by a multiple of this parameter as specified in the gravity model equation.

For this calculation a matrix containing the distance between every distrito pair was used. The calculation of this parameter could have also included such information as travel times, although such information was not available at the time this study was conducted.

2.4. Stage 3: Mode Choice[iii]

In this stage, all trips were assigned a mode of transportation. Usually the mode of transportation chosen for each mode is the one that a given traveler is most likely to take given his or her socio-demographic and economic profile, together with specific characteristics from each mode of transportation such as travel times and costs involved in its use.

For the MCMA, however, there was no information about travel times or costs. In fact, one of the reasons for developing the UTM for the MCMA was to obtain travel timesso that costs of delays in traffic could also be estimated.

Therefore, a simple rule was developed where car use, in comparison to public transportation, was made dependent on income levels. An average income per household in the MCMA was found. Then, average income per household was calculated for each distrito. After this, the percentage deviation from the MCMA average was found for each distrito. This percentage was applied to the average car share reported in the 1994 ODS. For example, if for a given distrito, average income fell below the average household income for MCMA by 2%, then the average percentage of total trips made by car was decreased by 2% for that distrito. The new mode shares were an input to the fourth stage of traffic assignment. The results of this exercise can be found in Gamas Buentello (2003).

2.5. Stage 4: Network Assignment and Congestion Modeling

Once the transportation demand side of trip generation and mode share calculation was complete, those trips were distributed throughout the roadway network. Using the transportation modeling software Transcad (by Caliper), a roadway network was built covering the entire MCMA. Due to the lack of a pre-digitized roadway network, a skeletal network consisting of highways and primary arterial roadways was made (see Amano 2004).

Important roadway characteristic data required for the model included the roadway design capacity (vehicles per hour) and the un-congested free-flow speed. Combining the demand side OD matrix with the supply side roadway network, the model was used to distribute the trips throughout the city using a user equilibrium approach. A link performance function for each segment of roadway is defined as follows (U.S. Bureau of Public Roads, see Transportation Research Board, 2000):

$$TT = FF \times [1 + \alpha \times (V / C)^{\beta}] \qquad (4)$$

Where TT is travel time for a vehicle, FF is the travel time of a vehicle when there is free flow on the roadway, V is the roadway volume (vehicles/hour), C is the roadway capacity (vehicles/hour), _ and _ are calibration parameters that relate the travel time on the roadway to the volume to capacity ratio. If these parameters have large values then the travel time increases, meaning that that road segment is more sensitive to congestion. Traffic flows are assigned to the links of the network in such a way that no traveler can reduce his or her travel time by switching to a different origin-destination path. The results obtained from the model included the vehicle flow and the congested vehicle speed on each roadway segment. A summary of data requirements and model outputs can be found in table 1.

The volume-to-capacity (V/C) ratio for each roadway segment is a key result in the network model. At a V/C ratio of 1 or less, the vehicle flow on a roadway is below its design capacity and traffic moves at or near the free flow speed. However, as the ratio exceeds 1 and the vehicle flow exceeds the roadway design capacity, congestion begins to have an impact on vehicle speed. The vehicle speed decreases significantly as the V/C ratio increases. When the volume of the roadway exceeds the design capacity by a factor of 3 or more, the vehicle speed slows to a crawl.

Using the Bureau of Public Roads equation and the V/C ratio on each roadway segment, we calculated the speed distribution for all vehicles traveling throughout the city. Transcad allowed us to determine the number of vehicle-kilometers traveled across the entire speed range, allowing us to calculate a more accurate emissions estimate.

The network model was run under two scenarios assuming peak hour flows- once with the 1994 OD data and once with the 2025 OD data. The vehicle speed distribution and spatial analysis of congestion was analyzed for each scenario (Gamas Buentello 2003). For the 2025 scenario, data was obtained from assumptions made for the “Divided City” future story developed by MIT (Dodder, 2004).

For the 1994 and 2025 comparison, it was assumed that the roadway network would remain the same, and no new infrastructure improvements would be built. Figure 2 shows the distribution of vehicle-kilometers traveled at varying speeds for both the 1994 and 2025 OD data.

Due to projected population and economic growth, the number of trips in the 2025 OD data is significantly higher than in 1994 OD data. In 2025 there are projected to be XXX, while they were

XXXX in 1994 (an increase in total trips of XX%). This resulted in an increase in vehicle flow and shift towards slower vehicle speeds.

In 1994, roadways in the city center predominantly face moderate levels of congestion, but for the most part were moving at a smooth pace. By 2025 most of the roadways in the city center would be severely congested. Congestion has also spread outward from the city center into the State of Mexico, due to population growth and sprawl. We can no longer pinpoint one bottleneck since the network has become saturated to the point that all roadways are congested (see Amano, 2004).

2.6. Stage 5: Emissions Modeling

The next stage of the model uses the vehicle speed distribution data gathered in the Transcad network to calculate vehicle emissions. Using speed-emissions relationships derived from the MOBILE6 emissions model as seen in figure 1, we can calculate the total emissions for each of the three scenarios and compare them with the emissions derived from the previous method assuming a fixed speed of 32 km per hour.

Table 2 is a summary of the comparison between the old "one speed" method and the network model method for the 1994 OD data, for traffic during a peak hour of the day. The results obtained are surprising; emissions estimated using the network model for hydrocarbons and carbon monoxide are actually lower compared to the old method, while emissions estimates for nitric oxides are higher. Because a large number of vehicles are still traveling at higher un-congested speeds where emissions for HC and CO are lowest, the congested vehicle emissions are offset by these lower emissions vehicles.

However, if we were to run the network model again using the 2025 data, the results show that the old model significantly underestimates emissions (see table 3). The level of congestion on roadways in the 2025 scenario has become a significant factor in an increase in emissions. Emissions are now significantly underestimated in the old method, by up to a factor of two for HC and CO. As congestion increases in the MCMA, the old method for calculating emissions becomes much less accurate. Details of emissions estimation can be found in Amano (2004).

3. CONCLUSION AND DISCUSSION

The UTM model created here has been used to illustrate the effect on congestion of the Divided City scenario. If actual speeds deviate from the 32km/h assumption, emissions can be underestimated if an analysis tool such as the UTM is not used. While before this study no complete UTM existed for the MCMA, there is much work to do to improve the model so that it can, additionally be used for transportation policy analysis.

The model requires improvement on several fronts. First, data is now available from INEGI and the Ministry of Transportation in D.F. (SETRAVI) that will allow us to incorporate missing information such as the actual digitized road network, disaggregated census data and a complete OD matrix for 1994.

Currently there is ongoing work at improving the model with the new information. In trip generation, a refinement of some of the variables used (such as population attending school and total services establishments) will lead to better parameter estimates. For trip distribution, any new information regarding the OD matrix will be incorporated. For modal choice, new data will be used to improve on the current estimates. Finally, a more detailed analysis of how the Divided City scenario will impact land use in the MCMA will be conducted.

The results of the model can also be used for valuation of delays in traffic (see Gamas Buentello 2003). Future research will include, not only putting a valuation on improved speed estimates, but also, incorporating differences in time valuation according to different trip types.

There is work to be done on the UTM in order to incorporate induced travel effects. Currently such effects can only be assumed and incorporated into the model via an assumption about land use. However, this topic requires in depth study. This is because if an addition is made to the roadway network in the UTM as is, this will show the instant benefit of reduced congestion, but the feedback loop whereby people incorporate information about the new improved network performance is not considered. The long term effect of land use and network changes are currently being modeled by using future scenarios that are constructed as projections of what might happen in the future. These scenarios are not intended as predictions. Feedback loops that help explain the long term relationship between transportation and land use changes are an important subject of future research.

Future work on the network model will improve our understanding of the relationship between congestion, mobility and emissions. In particular, we hope to improve the model by conducting discrete choice analysis in the third stage in order to obtain more information about the behavior of different modes of transportation. Also, freight has been assumed here to be a fixed percentage of total traffic. In future research we hope to give a broader picture of freight and analyze it in its own right, as it is a major contributor to economic growth.

The creation and use of the model has pointed out not only that it is a useful tool for improving emissions estimates, but also, that it has the potential to be a powerful policy analysis tool. The need for a new OD survey becomes evident, when we consider the need to use our parameters in the 2025 scenario, by assuming that they are fixed throughout time. Hopefully this data will become available in the future.

Acknowledgements

Comisión Ambiental Metropolitana (CAM), in Mexico, has provided funding for this research.

References:

Amano, D. (2004) *Macroscopic Modeling on a Microscopic Scale: A Road-Based Network Model of Mexico City.* MSc Thesis, Center for Transportation and Logistics, Massachusetts Institute of Technology.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models:* Kluwer Academic Publishers: Dordrecht, The Netherlands.

Ben-Akiva, M. and W. Lerman (2000) *Discrete Choice Analysis: Theory and Application to Travel Demand:* The MIT Press, Cambridge, Mass.

Dodder, R. (2003) Evolving City: Reducing Emissions Under Uncertainty. In Molina, M. and Molina, L. (eds), *Reducing Emissions in the Mexico Megacity*. Kluwer Academic Publishers, Boston, Mass. –in print-

Fotheringham, S. and O'Kelly, M. (1989) *Spatial Interaction Models, Formulations and Applications:* Kluwer Academic Publishers: Norwell, Mass.

Gamas Buentello, J. (2003a) *Impacts of Urban Transportation Congestion and Pollution Reduction Policies: Analytical Studies of Two Latin American Cities*. PhD Dissertation, Center for Transportation Studies, Boston University.

Gamas Buentello, et al. (2003b) Adding Spatial Resolution to Emissions Estimates in the MCMA. In Molina, M. and Molina, L. (eds), *Reducing Emissions in the Mexico Megacity*. Kluwer Academic Publishers: Boston, Mass. -in print-.

INEGI (1990) *XI Censo General de Poblacion y Vivienda (SCINCE 1990)*. Instituto Nacional de Estadistica Geografia e Informatica, Aguascalientes, Mexico.

INEGI (1994) *Censo Economico 1994*. Instituto Nacional de Estadistica Geografia e Informatica, Aguascalientes, Mexico.

INEGI (1994) *Encuesta Nacional de Origen y Destino 1994*. Instituto Nacional de Estadistica Geografia e Informatica, Aguascalientes, Mexico.

INEGI (2000) *XII Censo General de Poblacion y Vivienda (SCINCE 2000)*. Instituto Nacional de Estadistica Geografia e Informatica, Aguascalientes, Mexico.

Molina, M. and L. Molina eds. (2002) *Air Quality in The Mexico City Megacity, An Integrated Assessment*: Kluwer Academic Publishers: Boston, Mass.

Molina, M. and L. Molina eds. (2004) *Reducing Emissions in the Mexico Megacity*: Kluwer Academic Publishers: Boston, Mass. -in print-.

Radian International (1996) Final. *Mexico Emissions Inventory Program Manuals,* Volume VI: Radian International: Sacramento, Ca.

SAM (2000) *Inventario de Emisiones a la Atmósfera Zona Metropolitana del Valle de México*. Secreataría de Medio Ambiente, México, D.F.

Train, K. (1986) *Qualitative Choice Theory*: The MIT Press, Cambridge, Mass.

TRB (2000) *Highway Capacity Manual*: Transportation Research Board, Washington D.C.

Table 1: Transcad Model Inputs and Outputs

Model Inputs	**Model Output**	**Derived Outputs**
Roadway capacity (veh/hr)	Vehicle flow (veh/hr)	Volume/Capacity ratio
Free-flow speed (km/hr)		Congested vehicle speed (km/hr)
Number of trips between each distrito (veh/hr)		Vehicle travel time (hr)

Table 2: Emissions Data for 1994

1994 OD Data	**HC emissions (Tonnes/hr)**	**CO emissions (Tonnes/hr)**	**NOx emissions (Tonnes/hr)**
Old method (32 km per hour)	28.67	290.8	23.76
Network model method	26.62	269.9	25.12

Table 3: Emissions Data for 2025

2025 OD data	**HC emissions (Tonnes/hr)**	**CO emissions (Tonnes/hr)**	**NOx emissions (Tonnes/hr)**
Old method (32 km per hour)	47.31	480.2	39.60
Network model method	82.59	877.3	45.36

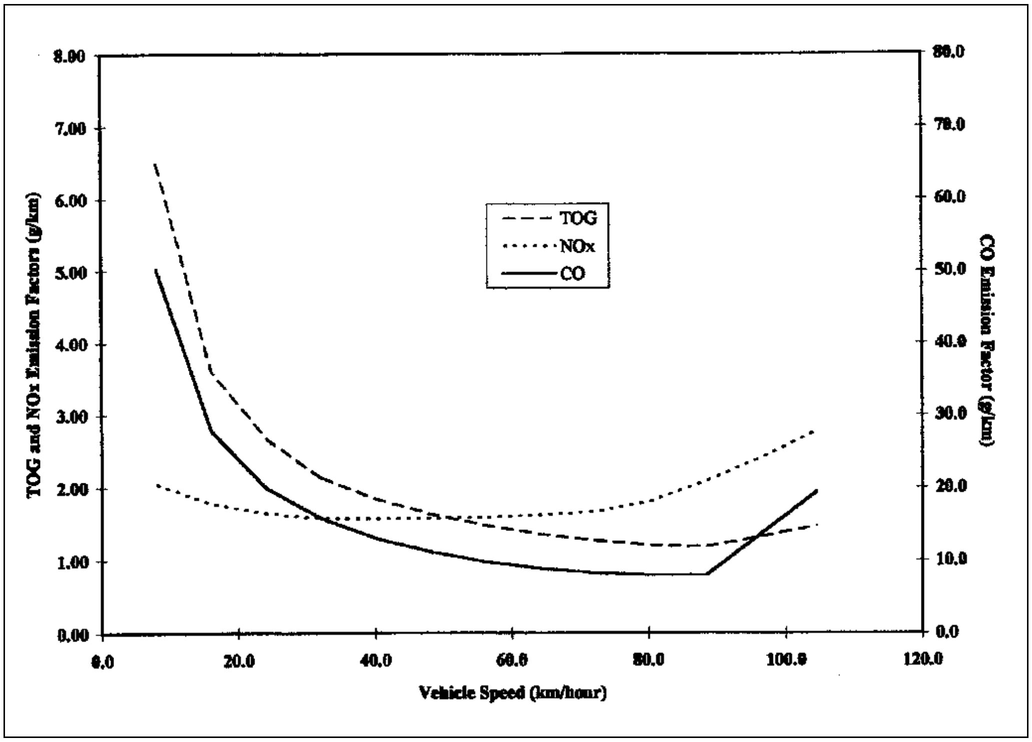

Figure 1: Average Vehicle Emissions Factors for TOG, CO and NO$_X$ Varying Speeds

Source: Mexico Emissions Inventory Program Manuals, Volume VI (Radian International, 1996)

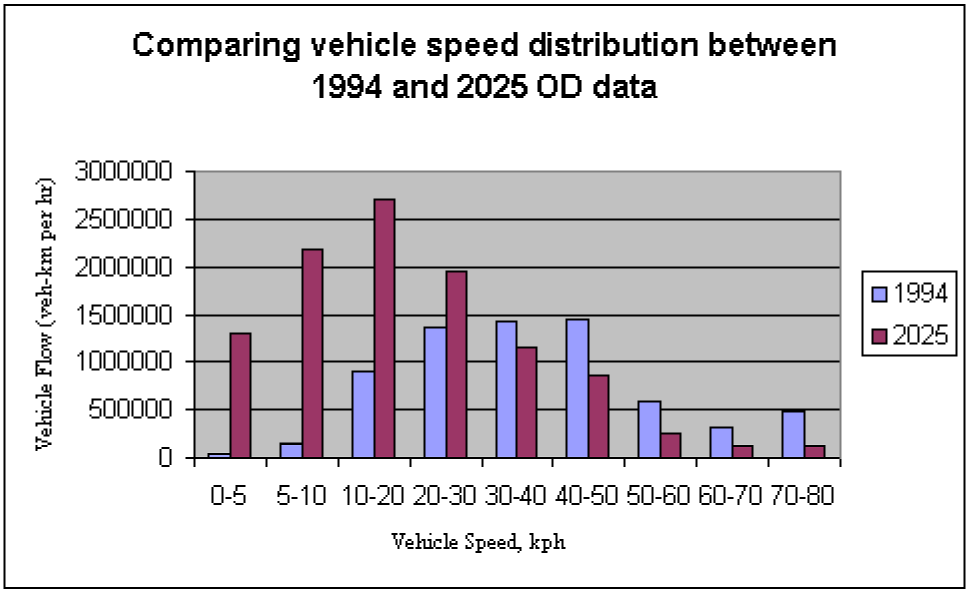

Figure 2: Speed Distribution for 1994 and 2025 OD Data

Notes:

---

[i] This paper is based on a chapter, Adding Spatial Resolution to Emissions Estimates in the MCMA, in a forthcoming book edited by Mario and Luisa Molina (2004).

[ii] To this effect, once the parameter was estimated, an analysis was conducted using the parameter plus one standard deviation and the parameter minus one standard deviation to re-estimate the OD matrix. No bias was apparent from this operation.

[iii] For discrete choice methods, which would typically be used in this case see Ben-Akiva and Lerman (2000), and Train (1986). Details of the calculations conducted here can be found in Gamas Buentello (2003).