The Cost of Congestion in the Greater Toronto Area

By

ALI MEKKY, Ph.D., P.Eng.
Transportation Planning Branch
Ministry of Transportation Ontario.
1201 Wilson Ave., Building B, 3rd Fl
Downsview, ON, M3M 1J8.
Tel (416) 235-4131, Fax (416) 235-4936
e-mail: ali.mekky@Ontario.ca

Paper prepared for presentation at the Session:
Best Practices in Urban Transportation Planning-An Economic Enabler

of the 2007 Annual Conference of the
Transportation Association of Canada
Saskatoon, Saskatchewan

October 2007

ABSTRACT

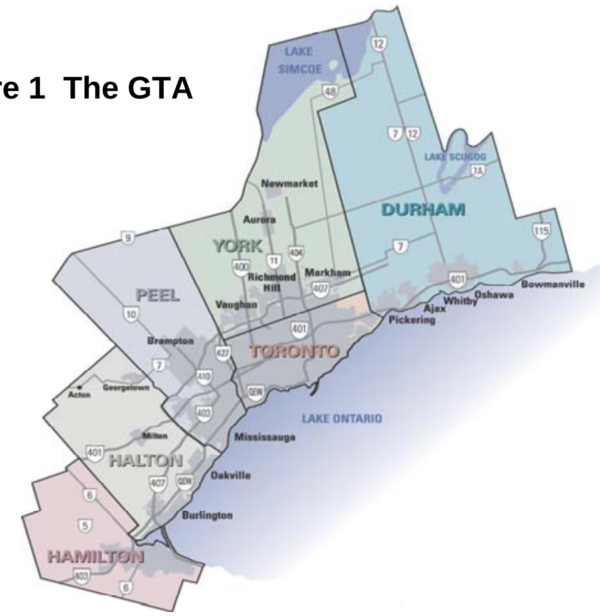
In large urban areas, congestion has become one of the most important problems facing commuters and economic growth. Congestion results in wasted time and energy (fuel consumption), more gas and particle emissions, increased accident frequency as well as additional costs to truck and passenger transportation. Governments have to be aware of the magnitude of the cost of congestion in their territories as it affects economic prosperity and the welfare of the population. Congestion cost should be one of the important performance measures in large urban areas as systems of infrastructure, land-uses and economic activities. The Greater Toronto Area (GTA), in southern Ontario, Canada, is defined as the City of Toronto and the Regional municipalities surrounding it (Durham, York, Peel, Halton and Hamilton). The objective of this paper is to study congestion in the (GTA) and estimate its cost. The approach developed in this study is based on a combination of field measurements and mathematical modelling. The developed procedure should be of help to many other cities that need to implement a similar exercise.

The study estimates the cost of congestion in the GTA to be about $5.5 billion in the year 2006. As population and employment continue to grow, the increasing congestion trend is also expected to continue unless significant steps are taken. Managing congestion should be, like the phenomenon itself, area-wide and multi-jurisdictional.

1. INTRODUCTION

For the last three decades, in most urban areas, traffic congestion has become one of the most important problems that affects not only commuters but also the economic growth and prosperity of those areas. In most of the large cities around the world congestion is one of the main reasons for increased gas emissions and accident frequency. Congestion also results in wasted time and energy (fuel consumption) that represent a significant additional cost to truck and passenger transportation.

Figure 1 The GTA



Every driver knows what congestion is. However, to study congestion and to quantify its impacts, one needs to define it in a clear and precise way.

The main objective of this study was to develop an estimate of the cost of congestion in the Greater Toronto Area (GTA). The GTA includes the six regional municipalities of Toronto, Durham, York, Peel, Halton and Hamilton (Figure 1). The cost estimate can be updated at regular intervals (every 2-3 years) to assess growth management and investment impacts on congestion in the GTA. The estimated cost should be conservative.

There are several reasons for the need to estimate the cost of congestion in the GTA. The government has to be aware of the magnitude of that cost which affects the economic prosperity in the GTA and in the entire country. The estimate should be one of the important performance measures of the GTA as a system of infrastructure, land-uses and economic activities. The government should monitor the change in the cost of congestion over time, instead of relying on the general "common perception", which could be faulty. This monitoring would result in avoiding a major deterioration in the system's performance, if left untreated, for several years. As a result of the obtaining the estimate, the government can compare the congestion cost in the GTA with that of other major cities in North America and around the world. Also, the government investment decisions related to road improvements, transit investments and land-use policies, should be assessed, partly on their contribution to the decrease in the cost of congestion.

An early attempt to estimate the cost of congestion in the GTA, for goods movements, was carried out in the mid 1980's. A study entitled the Metropolitan Toronto Goods Movement

Study (1987) investigated goods movements to, from and within Metropolitan Toronto. The study area included Metropolitan Toronto, the Regional Municipality of York as well as urbanized portions of Peel (Mississauga and Brampton) and Durham (Ajax, Pickering). Therefore, the study area was smaller than the GTA as we know it today. The Study concluded that "… Assuming the continuation of past trends, and without offsetting measures, the projected increase in congestion over the next decade will cost society more than $(1986) 15 billion within the Greater Toronto Area in terms of goods movement alone." (page 1-14). In another section of the same report (page 5-12) it states that "… The estimate of congestion cost suggests that these had risen to $1.9 billion for the total commercial vehicle costs in 1986 or 30 percent of the total costs." The latter refers to the commercial vehicles operating costs.

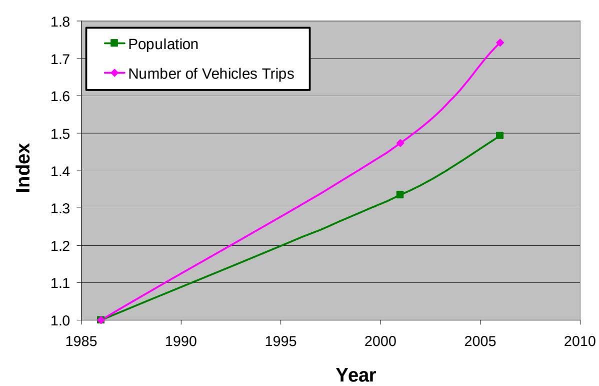
No serious attempt to quantify the cost of congestion in the GTA was made during the last 20 years.

This report is composed of 13 sections. Section 2 gives a brief account of the fast growth in the GTA during the 1986-2006 period. Sections 3 and 4 deal with the definition of and the reasons for congestion respectively. In Section 5 discusses the factors that affect congestion. Section 6 presents the impacts of congestion. The important issue of how to measure congestion is discussed in Section 7, which also includes the methodology adopted for that purpose. The main assumptions and the results of the analysis are given in Sections 8 and 9 respectively. Section 10 gives some examples of the cost of congestion exercises carried out in other jurisdictions. The importance of transit in the GTA is discussed in Section 11. Strategies to mitigate the effects of congestion are discussed in Section 12. Conclusions are given in Section 13.

2. GROWTH IN THE GTA

Between 1986 and 2006, growth in the GTA was one of the fastest in North America. During that period, the GTA grew by almost 100,000 persons/ year. Most of that growth took place in the 905 area.

Figure 2 Relative Growth Indicators Population and vehicle Trips in the GTA

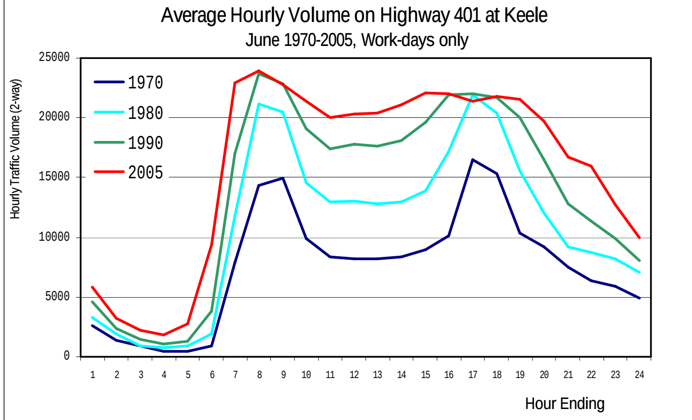


As Figure 2 shows, from 1986 to 2006 the number of vehicle trips in the system grew faster than the population. During the same period, the GTA benefited from the completion of the construction of highway 407 *ETR* (105 km of 6-lane freeway), the first fully electronic toll road in the world. Currently, the highway serves more than 300,000 vehicles per day and has resulted in a significant relief to the transportation network in the area.

Within the GTA, Highway 401, one of the busiest highways in North America, serves more than a million vehicle daily in the City of Toronto alone. At its busiest section (between Keele Street and Highway 400), the highway carries more than 400,000 vehicles per day. In the same location, the traditional morning and afternoon peak periods are becoming one long peak period extending 13 hours from 7:00 am to 8:00 pm (Figure 3).

During the 1986-2001 period, auto dominance was increasing while local transit share of total travel dropped from 16% in 1986 to 11% in 2001. On the positive side, this fact indicates that the LOS on the highways was not too bad, that is why people shifted to auto. This happened partly because of the opening of highway 407 *ETR*. During the same period, GO-Train trips increased by 126%, relieving congestion in some corridors. One GO Train carries the equivalent of one lane of traffic on the Don Valley Parkway or Queen Elizabeth Way into the downtown core.

Figure 3 The Flattening of the Peak Periods



During the same period (1986-2001), nearly 75 percent of all trip growth into Toronto across its borders with Peel was CBD bound. In 2001, travel from the 905 areas made up 35 percent of all CBD-bound travel, up from 20 percent in 1986.

Truck volumes on the highways have also been growing significantly. In some areas, commercial vehicle traffic grew faster than auto traffic.

3. WHAT IS TRAFFIC CONGESTION?

There are many definitions of congestion (see Mekky 2004). Traffic congestion is basically a high density of traffic (number of vehicles per kilometer) moving relatively slowly. The high density means short headway between vehicles. This usually happens (but not always) when the travel demand arriving at a facility almost equals to or exceeds the capacity of that facility. The reason for that could be either the traffic demand is too high or the road capacity is too low (e.g. partial road closure, weather conditions, merging points, lanes drop), or both. Congestion results in time loss and increased fuel consumption and emissions. It also contributes to driver's frustration and an increase in the number of accidents.

When a road operates near capacity, a small increase in vehicle flow could result in a large decrease in travel speed (large delay). This is the case in most of the GTA main highways during the peak periods.

It is worth noting that not all the workforce in the GTA suffer directly from driving under congested conditions. During the peaks periods, about a quarter of the trips (TTS, 2001) are made by transit or walk. Not all of those trips are directly affected by congestion.

4. WHY DOES IT HAPPEN?

To discuss the reasons for traffic congestion, one first need to differentiate between two type of congestion; namely, recurrent and non-recurrent congestion. Recurrent congestion is the congestion that happens on a daily basis, when the system is free from incidents or accidents and the weather is relatively good. It is mainly caused by inadequate road network capacity to handle the travel demand. On the freeway system, this type of congestion usually occurs at merging points, lane drop areas, and exit and entry points.

Non-recurrent congestion occurs because of several factors. Traffic accidents/ incidents blocking certain parts of the network or drivers slowing down to take at a look at incidents/ accidents could cause non-recurrent congestion. Weather conditions (e.g. rain or snow fall) could also result in a decrease in the capacity of roads which lead to non-recurrent congestion if the demand is high. Demand may occasionally increase due to special events (e.g. sport or national events). Supply (number of lanes) may occasionally decrease due to the presence of work zones.

Fluctuations in traffic volumes, which regularly happen on a daily basis, could also lead to traffic congestion, particularly, when the highway is operating near capacity.

5. FACTORS AFFECTING CONGESTION

Three main factors affect congestion, namely, land-use/ demographic factors, transportation demand and transportation supply factors. This is discussed below. Data was derived from the TTS for the various years.

5.1 Land Use/ Demographic Factors

- Increase in Urbanization. As mentioned earlier, most of the 1.4 million growth in population (1986-2001) occurred in the 905 area, which grew by 57% compared to the City of Toronto that grew by only 11%.
- Increase in population not supported by transit.
- Increase in employment not supported by transit.
- Increase in land/ housing prices that force people to live further away from their places of work.

- Increase in female participation in the labour force made it harder for families, where both spouses work, to optimize their home location
- Decrease in the family size (number of people).

5.2 Transportation Demand Factors

Daily trips have increased from 8,213,000 trips in 1986 to 11,515,300 trips in 2001, i.e. a 40.2% increase. This is caused by the following factors:

- An increase in daily trips per household from 5.6 trips in 1986 to 5.8 trips in 2001, i.e. a 3.6% increase.
- An increase in the number of households from 1,466,100 in 1986 to 1,975,200 in 200, i.e. a 34.7% increase.
- An increase in population from 4,062,900 to 5,386,100, i.e. a 32.6% increase.
- An increase in the number of adults per population
- A decrease in auto occupancy
- Increase in car ownership/ availability.
- Increase in income per capita.

The above factors (TTS, 2006-2001, see Data Management Group Website) result in both an increase in the number of autos and trucks and in the trips made by those vehicles. This would also result in an increase in the number of accidents leading to more congestion.

5.3 Transportation Supply Factors

Usually the increase in the extent and efficiency of a network doesn't match the increase in the demand. However, over the last several years, the ministry was successful in implementing major expansion to the highway network, mainly Highway 407 mainly built in the late 1990's and early 2000's.

Uncoordinated road closures could result in traffic congestion. Insufficient improvement to the transit network could also result in traffic congestion.

6. COMPONENTS OF CONGESTION COST

When a vehicle enters a congested highway, it imposes and bears congestion delay and its costs. Traffic congestion costs consist of the following components:

- Cost of Delay, at main road sections as well as at intersections, imposed on users of private autos and trucks as well as public transportation modes. This cost represents loss of income/ opportunities. The FHWA (Cambridge Systematics, 2005) estimates the sources of delay because of recurrent and non-recurrent congestion as shown below.
  - Bottlenecks (40%)
  - Traffic Accidents/ incidents (25%)
  - Bad Weather (15%)
  - Work Zones (10%)

  - o Poor Signal Timing (5%)
  - o Special events (5%)
  - o Fluctuation in traffic demand.
- Cost of Operating a Vehicle, including the cost of fuel;
- Cost of Environmental Pollution, including cost of emitted gases, particles and noise;
- Cost of Accidents and the change in their patterns;
- Cost of Drivers' Stress; and
- External costs imposed on non-motorized travel (pedestrians and cyclists, also called Barrier cost or severance).

The above costs are costs to both the individuals and society at large.

As indicated above, non- recurrent congestion sources represent more than half of the total delay. To adopt a conservative approach, the delay for non-recurrent congestion was assumed to be equal to that of the recurrent congestion. This assumption is not only supported by FHWA work sited above but by other works in the literature as explained below.

7. HOW CAN WE MEASURE CONGESTION AND ITS COST?

7.1 Measures of Congestion

According to Meyer (1994), a measure of congestion should
- have a strong relationship to the cost of congestion;
- be theoretically and functionally related to other predictable measures;
- be based on statistically sound measurement techniques.
- be acceptable to transportation professionals; and
- be comparable across time and space

Congestion measures should be able to quantify all the dimensions of congestion, namely:
- Extent (or breadth, e.g. number of persons affected)
- Intensity (or depth, e.g. minutes of delay per person, the average speed).
- Duration (number of hours the system is congested)

It could also be argued that the variability in travel time, from day to day, is an important dimension of congestion. However, it is not easy, without extensive measurements to quantify that dimension.

Since there is no one indicator that would describe congestion fully, several indicators were used in this study that are listed below.

- Number Of Congested Trips
- Percentage of Congested Trips

- Total Veh.km For Congested Trips

- Congested Veh.km For Congested Trips
- Percentage Veh.km Congested For Congested Trips

- Total Veh.min For Congested Trips
- Congested Veh.min For Congested Trips
- Percentage Veh.min Congested for a Congested Trip

- Average Trip Length For Congested Trips
- Average Trip Length Congested For a Congested Trips
- Percentage Trip Length Congested For Congested Trips

- Average Trip Time For Congested Trips
- Average Trip Time Congested for a Congested Trip
- Percentage Trip Time Congested For a Congested Trip

- Cost of congestion with respect to Free-flow
- Cost of congestion with respect to LOS D/E
- Percentage Difference between Free-Flow & Congested Bases

- Travel Time Index

- Annual Delay per peak traveller
- Annual Congestion cost per peak traveller

Most of the indicators that are listed above for this study are travel time-based. The indicators selected satisfy all the criteria discussed earlier and are easy to understand.

7.2 Measures of the Cost of Congestion

Recurrent Congestion
Before implementing any procedure one has to adopt a working definition of what *congestion delay* is. One way of doing this is to assume that, for a trip, congestion delay is the difference between the congested travel time and the free-flow travel time. This is the approach adopted by the Texas Transportation Institute (TTI) in the largest Urban Mobility Study done in the US to date (Schrank and Lomax, 2004, also Lomax *et. al*. 1997). That study (and the following one of the same kind) used travel volume data for highway and arterial sections that are in the USA major cities that is available in the Federal Highway Administration's Highway Performance Monitoring System.

The TTI's definition of the travel time delay due to congestion as the difference between the observed travel time and the free-flow travel times is unrealistic. Driving during the peak periods in a large urban area as the GTA, one can't reasonably expect to drive under at LOS A. Also, driving at a free-flow speed should not be the ideal norm, as it is associated with higher gasoline consumption and emissions. Therefore, the TTI approach would lead to an over-estimation of the cost of congestion.

Earlier in the mid 1980's, Lindley (1986) used the congestion threshold as the border of LOS C and D (i.e. v/c = 0.77). A more realistic threshold would be the border between LOS D & E, which is the LOS that highway drivers would expect. This is the threshold adopted for this study.

The Highway Capacity Manual 2000 (HCM 2000) was used as the LOS guide. According to HCM 2000, the LOS for freeways and multilane highways is a function of traffic density or volume/ capacity ratio. For other types of roadway sections it could be different, e.g. it is the combination of speed & percent time-spent-following for two-lane highways.

Non-recurrent Congestion
Non-recurrent congestion is usually difficult to estimate and needs relatively long time to study. Knowing that a full study of non-recurrent congestion cost could take more than year to do, in the interim, one can try to utilize the relationship between recurrent and non-recurrent congestion delay from other major studies. Literature was reviewed to find such relationships.

Lindley (1986, 1987, 1989) studies the subject extensively and came to the conclusion that non-recurrent congestion delay represents about 60% of the total congestion delay. This means that the non-recurrent congestion delay is about 1.5 that of the recurrent delay. Lindley's work has been extensively cited in the literature.

Using simulation experiments, Hall (1993) supported Lindley's findings, stating that "… In many respects the findings of this paper are consistent with prior studies. Non-recurrent congestion delay exceeded recurrent congestion delay in most simulations."

To adopt a conservative approach, the cost of non-recurrent congestion for this study was taken to equal to that of the recurrent congestion.

The total Annual cost of recurrent congestion is composed of delay cost as well as fuel cost for both passenger and commercial vehicles. The cost of non-recurrent congestion includes the cost of delay due to traffic accidents/ incidents, bad Weather, work zones poor signal timing, special events, and fluctuations in traffic Demand.

7.3 Methodology

One of the important cost-of-congestion studies in the USA was carried out by the Texas Transportation Institute (TTI). As a result of those studies, the TTI published several annual reports under the title "Urban Mobility" (see also Lomax 1997). Some of the contributions of the TTI study will be used in this study, as described below.

The TTI methodology is based on sampling the network sections and measuring travel times on these sections. To do so requires a significant amount of time (several months to a year) to categorize network links, develop a reliable, statistically sound sampling scheme and do the measurements.

An alternative method would be to try to work with the data available. For the GTA, this includes traffic counts collected over most of the important road sections. Since the available sample of traffic counts was not selected in any valid statistical way, it was not possible to solely rely on that source.

A second alternative way was to do the exercise using the GTA mathematical model that is available and covers all the study area. However, one then needs first to validate the model in a comprehensive way.

A third alternative was to combine the above two alternative approaches. The way this was done is to ensure that the model reflects the available traffic counts in a reasonable way. If that is the case, then the assigned volumes on the links that had no counts would hopefully reflect reality reasonably well. In this way, the issue of "sampling" was avoided and the approach would have a good chance of producing realistic estimates of recurrent congestion.

The cost of congestion in this document is based on the 6-hour peak period of work-days for a year. It includes the delay cost as well as fuel cost for both passenger and commercial vehicles. It also includes the cost of non-recurrent congestion covering the cost of delay due to traffic collisions/ incidents, bad weather, work zones, poor signal timing and special events. The cost estimated doesn't include the

- cost of congestion during the off-peak periods, weekends and holidays;
- cost of congestion inflected on transit riders;
- external costs imposed on non-motorized travel (pedestrians and cyclists, also called Barrier cost or severance);
- cost of accidents/ incidents and the change in their patterns;
- cost of environmental pollution due to congestion, including gases, particles and noise; and
- cost of drivers' stress.

7.4 Mathematical Modelling

The model used is a traditional sequential four-stage model. The main data input was the Transportation Tomorrow Survey (TTS) for the year 2001, as well as the GTA networks and the volume-delay functions.

The TTS is a comprehensive telephone travel behaviour survey conducted in the Greater Toronto Area (GTA) once every five years since 1986. The 2001 TTS sample (5%) included 137,000 completed interviews. The 2001 survey area covered the GTA as well as some of the surrounding areas. For details about the TTS, please visit the Data Management Group Website: http://www.jpint.utoronto.ca/dmg/tts.html.

The demand matrix for the base-year model was extracted from the TTS 2001. The transportation Planning package used in this exercise was Emme, a Canadian product by the Montreal firm INRO. The demand matrix was fine-tuned to the 2001 traffic counts using the "Gradient-based Approach" developed by Spiess (1990). It is very important not to excessively adjust the demand matrix as this might result in a distortion to the original demand matrix and a loss of information (Mekky *et. al*. 2003). Three iterations of adjustment were carried out.

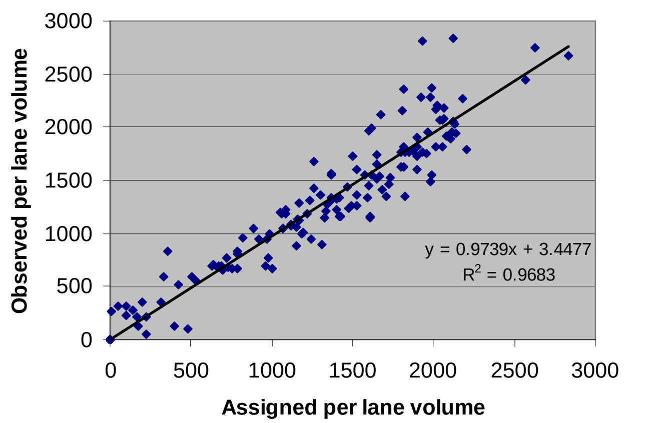
Figure 4 shows the observed (counted) volumes versus the model-estimated (assigned) volumes. The line of best fit happens to almost coincide with the 45-degrees line of ideal fit. R^2 is quite high (0.97) indicating that the model is representing the observed volumes in a good way. The scatter around the ideal 45 degrees line is reasonable. Other checks were also carried out to make sure that the model represents the base-year (2001) in a good way.

The study covered the years 1986, 2001 and 2006. Three types of runs were needed for each year studied.

- Equilibrium assignment, which represent the main traffic assignment approach.
- Free-Flow Assignment, to derive indicators under free-flow conditions.
- Select link Assignment to study the congested trips.

Emme macros to estimate congestion indicators were prepared and several scenarios were structured.

Figure 4 Volumes Assigned vs Observed



8. MAIN ASSUMPTIONS

In this report, the GTA is defined as the urban area that includes the City of Toronto, and the Regional Municipalities of Durham, York, Peel, Halton and Hamilton.

For the analysis discussed below, it is assumed that the road and transit networks remain the same as in the base year (do nothing scenario). Future work will include networks with planned improvements.

8.1 Congested Hours

Although some network components do experience congestion during some of the off-peak hours, the bulk of congestion is known to take place during the morning and the after-noon peak hours; six hours in total. Therefore, in this study, off-peak congestion was not included.

On the weekends and during holidays, congestion occurs on certain routes. However, due to the lack of detailed origin-destination data for those no-work days, congestion on those days was not quantified.

In summary, the estimated congestion cost in this report relates only to the peak hours of work-days for a year.

8.2 Vehicle Occupancy

Vehicle occupancy was observed to decline over the last several years. However, that trend is not expected to continue for long and should start to reverse as oil prices and the cost of car ownership/ lease increase and more HOV lanes on the 400-highway series are constructed. Accordingly, vehicle occupancy was taken to be 1.25 persons per vehicle.

8.3 Value of Time

One of the important assumptions in this exercise is the Values of Time (VOT) for auto travellers and for truck drivers. One way of producing an estimate is to equate it with the average income per person per hour in the GTA. For the year 2001, the average income per person per hour in the GTA was slightly over $20 (Statistics Canada, 2001 Census).

To adopt a conservative approach, the VOT for the year 2001 was taken as $15 (2001 dollars). It was also assumed that the VOT for trucks is five folds that for autos.

8.4 Modal Split

During the morning peak period between 1986 and 2001, the share of local transit trips decreased from 21% in 1986 to 14% in 2001. This trend is not expected to continue. For this work, it was assumed that the future transit modal share will not change in any substantial way.

8.5 Price of Gas

It is very difficult to forecast the price of gas 25 years into the future. An assumption was made that it will increase by 4% per year above the inflation level. This modest increase is definitely much less than what happened over the last several years.

9. ANALYSIS RESULTS

9.1 All Trips

Average weighted network speed in the year 2001, for congested trips was 52 km/h. This figure drops to 46 km/h in 2006.

Figure 5 shows several congestion indicators for all the trips. The dramatic increase in the cost of congestion after the year 2001 is very clear.

9.2 Congested Trips

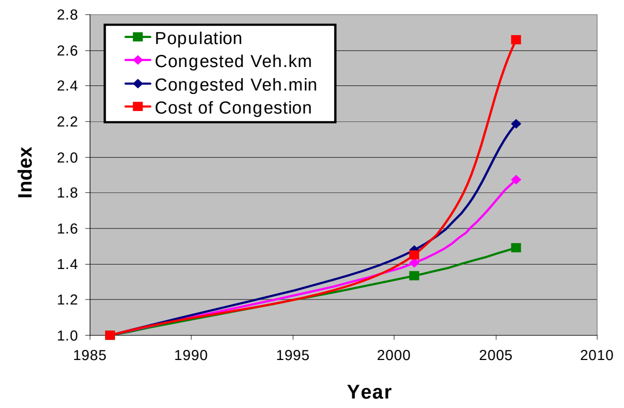
It is important to differentiate between two types of trips:

- Non-congested trips; trips that experience no congestion (at any part of the full trip)
- Congested trips; trips which has some part congested, not necessarily the entire length. Accordingly, these trips are composed of two parts:
    - Congested parts
    - Non-congested parts

Analysis results show the following patterns:

- Figure 5 shows several congestion indicators for all the trips
- Congested VKM increase faster than population
- Congested VMIN increase faster than VKM and population
- The cost of congestion increases faster than VMIN, VKM and the population
- The dramatic increase in the cost of congestion after the year 2001 is very clear.

Figure 5 Change in Population and Congestion; Relative Indicators



Other results indicate that trips experienced congestion during the peak period of the year 2001 represents about 85% of all the trips in that year. This number increases to 88% in 2006.

Table 1 shows that for the congested trips, on the average, the percentage of their trip length (in km) that is congested is 44% in 2001. This number increases to 51 in 2006. For the congested trips, on the average, the percentage of their trip duration (in minutes) that is congested is 60% in 2001. This number increases to 68% in 2006. Average weighted network speed in the year 2001, for congested trips is 38 km/h. This figure drops to 34 km/h in 2006.

Table 1 Main Congestion Indicators

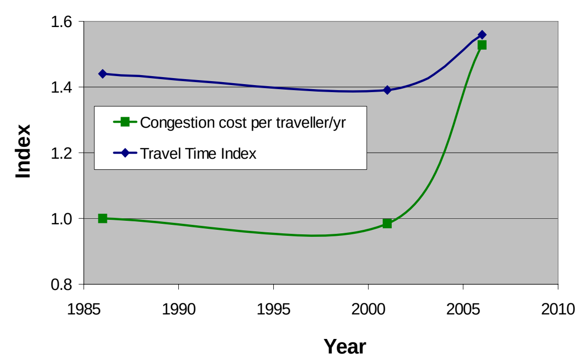
Description	1986	2001	2006
Population	**4,062,900**	**5,424,384**	**6,064,229**
Number of Vehicles Trips	439,521	647,590	765,466
% Vehicle Trips Congested	100%	85%	88%
Travel Rate Index	1.44	1.39	1.56
%Veh.km Congested, of all trips	47%	44%	50%
%Veh.km Congested, of congested trips	47%	44%	51%
%Veh.min Congested, of all trips	62%	59%	68%
%Veh.min Congested, of congested trips	62%	60%	68%
% Trip Length Congested For a Congested Trip	47%	44%	51%
% Trip Time Congested For a Congested Trip	62%	60%	68%
Cost of congestion wrt Free-flow millions $	2,288	3,306	6,002
Cost of congestion wrt LOS D/E millions $	2,084	3,023	5,542
Difference between FF & Congested Bases	10%	9%	8%
Annual Delay / pk traveller Hrs/ yr	32	31	46
Congestion cost per traveller/yr w.r.t border of LOS D/E	705	694	1,076

Figure 6 (below) shows that after a drop in the congestion cost per traveler and the travel time index (the ratio of the average free-flow speed to the average speed) just around the year 2000 (because of the 407ETR effect), they started to increase significantly thereafter.

9.3 Congestion Cost

Analysis results indicate that the cost of congestion in the GTA in the year 2001 was about $3.0 billion. This figure increases to $5.5 billion in the year 2006.

Figure 6 Relative Congestion Cost Per Traveller and the Travel Time Index



If the cost of congestion was estimated with respect to an ideal situation (i.e. with respect to free-flow conditions), the cost of congestion figures would have increased by about 8-10%.

If the value of time was assumed as $20/hr, rather than $15, for the year 2001, the cost of congestion (2006 dollars) would be:

- for the year 1986: $2,737 billion rather than $2,084 billion;
- for the year 2001: $3,968 billion rather than $3,023 billion;
- for the year 2006: $7,184 billion rather than $5,542 billion;

10. COST OF CONGESTION IN OTHER URBAN AREAS

Over the last decade, several cities in Canada and the USA carried out studies to measure the cost of congestion. One has to be careful in comparing the cost of congestion for different urban areas, as there could be differences in the elements of cost considered, the price of gasoline, the costs of accidents, the methodologies, etc. A few examples are given below.

Montreal
In 2001, the Ministere des Transports du Quebec (MTQ) conducted a study to estimate congestion using a newly calibrated model. The total cost of congestion for the Montreal region was estimated to be about $779 million per year (in 1998 dollars). This estimate doesn't include the non-recurrent congestion cost.

Vancouver
The British Columbia Regional office of Transport Canada used Translink travel demand forecast model to estimate the cost of congestion for commuters in the Lower Mainland/ Vancouver. The estimate was approximately $680 million (in 2001 dollars). This doesn't

include the cost to commercial vehicles. It is reported, by Transport Canada (2006), that including both truck and general traffic, the cost would reach $1.2 billion (in 1999 dollars).

The Very Large Urban Areas in the USA

In “The 2004 Urban Mobility Report” Schrank and Lomax (2004) report the following estimates for the Year 2002, for the Very Large Areas (larger than 3 million (Exhibit 11):

- The average cost of congestion per traveler is $1104. The equivalent number estimated for the GTA is about $840 for the year 2002 and about $1165 for the year 2006 (costs compared to free-flow).
- The average annual delay per traveler is 62 hours while the corresponding figure for the GTA is 38 hours.

It is hard to compare the estimates above from different cities as they are based on different assumptions, include different components of cost, and are based on different methodologies. However, one thing remains clear, that the studies confirmed the long-held view that the cost of congestion is very substantial. It runs into the billions of dollars each year for a larger city (larger than 3 million people).

11. THE IMPORTANCE OF TRANSIT

Transit has a very important effect on congestion. According to the Transportation Tomorrow Survey (2001), about 16% of the trips in the GTA are made by transit. Although most of these trips are still on the roads, they do not suffer directly the frustration of driving under congested conditions.

If there were no transit in the GTA this year (2007), the number of auto trips during the peak periods would have increased by 17%. This increase is high enough to bring the whole transportation network to a grounding halt. There are at least two reasons to explain that. First, during the peak periods, the transportation system in the GTA operates at capacity in most of its sections. This happens at the very non-linear part of the system’s performance function. Second, the GTA demand for transit is concentrated in the City of Toronto with the Central Business District (CBD) as its focal point. Therefore, a relatively small change in demand, but concentrated in a part of the network that is at capacity, could result in a large change in delay.

Simulation experiments show that if all the transit in the GTA stop today, we would have had a situation that is much worse than that if we leave the GTA network without any further improvements until the year 2031.

Therefore, without transit, the GTA would not have been able to function with the current distribution of activities (e.g. residential, commercial, industrial). This means that, without transit, the GTA would have developed in a radically different way.

12. MITIGATING THE EFFECTS OF TRAFFIC CONGESTION

There are several strategies that help mitigate the effects of traffic congestion. They are briefly listed below.

12.1 Improve Transportation Network

- Add new road Capacity:
  - As General Purpose Lanes (GPL); or
  - As High Occupancy Vehicle (HOV) lanes. These lanes intend to serve more people with less lanes with shorter and more reliable travel times.
- Improve local and regional transit network
  - Add new local and Interregional Transit Services
  - Reduce travel times
  - Integrate fares
  - Improve safety
  - Improve system connectivity
  - Improve and expand Park and Ride facilities
  - Improve accessibility to transit
- Encourage other Passenger Modes, e.g. bicycle and pedestrian facilities
- Encourage freight modes, other than trucks (e.g. the forgotten mode “rail”).

12.2 Decrease Peak-Period Demand

- Plan land-uses that promote less driving
  - Encourage land use forms conducive to transit use (high density developments).
  - Encourage land use distributions that result in a decrease in trip length (e.g. mixed land uses).
  - Encourage higher densities near main arterials.
  - Restrict new developments that result in the deterioration of the LOS.
- Eliminate some of the auto trips:
  - Encourage people to work from home (telecommuting)
  - Encourage shorter work week
- Spread auto demand over time
  - Encourage flexible work hours
- Spread auto demand over other modes (encourage shift to other modes, e.g. transit, bikes, walk).

- Encourage ridesharing by:
  - Increasing the number and capacities of commuter car-pool lots.
  - Encouraging access to the computer ride-sharing programs that match commuters that need to share rides.
- Discourage trucks during peak periods

Many of the opportunities above are now incorporated into the government's Growth Plan.

12.3 Improve System Management

- Implement ramp metering where appropriate
- Improve communication system for travel and commute information
- Expand Changeable Message Signs
- Implement Changeable Speed Limits
- Improve incident management
- Optimize the scheduling of road Improvements to minimize delay
- Improve goods movement option

12.4 Decrease Number Of Potential Accidents

- Improve the public's awareness of traffic safety.
- Invest in ITS collision avoidance systems.

12.5 Decrease Accident/ Incident Detection Time

- Implement automatic accidents detection system at a wide scale.
- Improve emergency notification and reporting procedures.

12.6 Decrease Accident/ Incident Duration

- As discussed above, almost half of congestion delay on urban freeways is caused by accidents/ incidents (collisions, disabled vehicles, spills,…,etc). Any decrease in accident duration can result in substantial benefits.
- The common perception is that for every minute a lane remains blocked, four to 10 minutes of congestion may result, depending on the state of congestion before the accident happened.

12.7 Help drivers select the optimal route

- Enable drivers to view traffic conditions from home
- Encourage highway advisory radio stations
- Implement "changeable message signs" on all the main arterials (e.g. freeways and expressways).

12.8 Implement Congestion Pricing

For any commodity or service, the price is the instrument that brings to balance, the quantity produced and that consumed. If the price of a scarce commodity is fixed (no matter how much the demand grows), shortages might occur and become more severe as the demand increases.

Demand for road capacity is somehow similar. Although drivers pay a price for their usage of the road (travel time + fuel consumption + cost of maintenance + etc.), they don't pay the full price. When a driver enters a traffic stream, s/he inflicts delay on all the drivers on that facility. Most of this external social cost is not borne by that driver. Accordingly, inefficiencies in using transportation facilities would result. Making drivers pay for the delay they inflict on others is the main idea behind congestion pricing.

Faced with the prospect of paying a congestion toll, some drivers would try to avoid the payment by shifting to other routes, other times of the day, other destinations, other modes, or simply decrease the frequency of their travel. In most of these cases, congestion decreases.

The idea of congestion pricing is not new. It was first introduced 80 years ago by Knight (1924). The first well know application was in Singapore in 1975. Many papers have been published on the issue, yet implementation took place in a handful of cities around the world. Over the last ten years, a number of large cities have implemented or are working on implementing congestion pricing.

13. CONCLUSIONS

The estimate of the cost of congestion in the GTA is approximately $5.5 billion for the year 2006. The above estimate is conservative since it doesn't include the cost of congestion during the off-peak periods, weekends and holidays, the cost of congestion inflected on transit riders, the external costs imposed on non-motorized travel (pedestrians and cyclists), the cost of accidents/ incidents and the change in their patterns, the cost of environmental pollution due to congestion, including gases, particles and noise, and the cost of drivers' stress.

Analysis shows that the cost of congestion is increasing over time. More people are experiencing more congestion for longer time periods. As population and employment continue to grow, the congestion trend is expected to continue unless significant steps are taken to improve congestion levels.

There are many ways to alleviate congestion and reduce its cost. Since almost half of the congestion cost is due to accidents and incidents, efforts to prevent, detect and clear accidents/ incidents could be very cost effective.

Congestion, to some extent, is a symptom of a non-transportation problem; namely land-use decisions that impact transportation choices.

Managing congestion should be, like the phenomenon itself, area-wide and multi-jurisdictional. It can be implemented by controlling both the supply and demand. Increasing the supply is expensive and can't by itself improve the congestion condition. However, selective widening of facilities in certain areas and other low-cost geometric improvements could be very effective.

Controlling demand is the most promising opportunity. This is done through decreasing the overall demand level and/or spreading it over space, time and modes. Changes to land use planning such as those included in the Growth Plan for the Greater Horesshoe area will contribute to a better balance of supply and demand.

When all the demand & supply instruments of mitigating congestion are exhausted, then congestion pricing might be an alternative worth considering.

ACKNOWLEDGEMENT

Mr. Kenny Ling, P.Eng. from the Ministry of Transportation Ontario carried out the EMME/2 runs. The opinions and the figures included in this paper are the author's. They do not necessarily represent the point of view of the Ministry of Transportation, Ontario.

REFERENCES

Cambridge Systematics (2005) Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation, prepared for Federal Highway Administration, US Department of Transportation, September.

Data Management Group Website: http://www.jpint.utoronto.ca/dmg/tts.html.

Gordon, P. and Richardson H.W. (1994) Congestion Trends In Metropolitan Areas, University Of Southern California, Los Angeles.

Gordon, P., Richardson H.W., and Kumar, A. (1990) Peak Spreading: How Much? Transportation Research A, Vol. 24a, Pp. 165-175.

Hall, R.W. (1993) Non-recurrent congestion: How big is the problem? Are traveler information systems the solution? Transportation research (c), vol 1 no 1, pp89-103.

Knight, F. (1924) Some Fallacies In The Interpretation Of Social Cost. Quarterly Journal Of Economics, 582-606.

Lindley, J. A. (1986) Quantification of Urban Freeway Congestion and Analysis of Remedial measures. Report FHWA/RD-87/052. FHWA, US. Department of Transportation.

Lindley, J. A. (1987) Urban freeway congestion: Quantification of the problem and effectiveness of potential solutions. ITE Journal, January, 27-32.

Lindley, J. A. (1989) Urban freeway congestion problems and solutions: an update ite journal, Dec. 21-23.

Lomax, T., Turner, S. and Shunk, G. with Levinson, H.S., Pratt, R.H., Bay, P.N. and Douglas, G.B. (1997) Quantifying Congestion, NCHRP Report 398, Transportation Research Board, National Research Council, National Academy Press, Washington, D.C.

Mekky, A. (2004) Measuring Congestion In The Greater Toronto Area, Transportation Research Board (T.R.B.) 2004 Annual Conference, January 2004, National Research Council, Washington D.C., (Paper #04-2771).

Mekky, A. (2004) Measuring Congestion In The Greater Toronto Area, Traffic Engineering and Control, Vol 45 (2), February 2004, 57-61, Hemming Group, London, England.

Mekky, A., Tai, A., Khan, M.I. (2003) Fine Tuning Mathematical Models for Toll Applications, 18th Annual International EMME/2 User's Group Conference, Calgary, Alberta, October 2003.

Metropolitan Toronto Goods Movement Study (1987) Technical Report, Metropolitan Toronto Roads and Traffic Department.

Meyer, M.D. (1994) Alternative Methods For Measuring Congestion Levels, Transportation Research And Education Center, Georgia Institute Of Technology.

Schrank, D. and Lomax, T. (2004) The 2004 Urban Mobility Report, Texas Transportation Institute, Sept.

Spiess, H. (1990) A Gradient Approach for the O-D Matrix Adjustment Problem. EMME/2 Support Center, Haldenstrasse 16, CH-2558 Aegerten, Switzerland.

Statistics Canada (2001 Census). Average Earnings of the popultion 15 years and over by highest level of schooling, by census metropolitan areas.

Transport Canada (2006) The Cost of Urban Congestion in Canada, Environmental Affairs, March.