Data Requirements for a Province-wide Travel Demand Model

David Kriger, P.Eng., MCIP
David Kriger Consultants Inc., Ottawa

Paper prepared for presentation at the
"Best Practices in Transportation Planning" Session
of the 2014 Conference of the
Transportation Association of Canada

Montréal, Québec

May 2014

ABSTRACT

Saskatchewan is in the midst of a population and economic boom, driven by resources such as oil and gas. At the same time, the distribution of other industries, such as agriculture, is changing, as the consolidation of rail transfer points forces farmers to truck their wheat longer distance; and the catchment areas of the Province's two largest urban activity centres – Regina and Saskatoon – continue to expand.

One result is that the Province must consider rapidly growing demands for private and commercial travel on its highways, and the resultant impacts for improvements, maintenance and rehabilitation. To help in the planning decision-making, the Saskatchewan Ministry of Highways and Infrastructure (MHI) is considering the development of a Province-wide travel demand forecasting model. The model would focus on highway traffic, and would be integrated with the existing urban models.

As a first step, MHI studied the data requirements for a Province-wide model. The study began by defining potential modelling approaches. Based on this definition, an inventory was prepared of existing data sources among Provincial, Federal and local agencies that potentially could be used to develop the model. Over 30 sources were found, comprising both travel data, such as traffic counts and origin-destination surveys, and demographic and socio-economic data. The inventory considered completeness, geographical coverage, representativeness, levels of spatial and temporal detail and ease of use, among other attributes. It also examined the ability to forecast the data at the necessary geographies.

Next, the available data were assessed against the needs of the potential modelling approaches. From this, data gaps were identified as the basis for the specification of several data collection options. The options, and combinations of these options, are designed to allow the Province-wide model to evolve as different needs arise and as funding allows.

The paper describes the inventory and the proposed data collection options. Importantly, the options focused on the feasibility of emerging electronic data collection technologies, including cellphones, GPS traces and loggers. The paper also speaks to the practical challenges and opportunities from the use of these data, especially the electronic sources.

Although the paper describes a rural / inter-urban application, it should be of interest to urban as well provincial transportation planners, many of whom are faced with similar data needs and opportunities in developing forecasts for their own jurisdictions. In particular, this paper speaks to the following considerations:

- Many of the actual data sources examined in this study, and certainly many of the data types, are applicable to urban, regional and provincial jurisdictions across the country.

- Interest in the nation’s transportation community is growing in the potential deployment of electronic technologies to collect travel data.

- Most attention in the literature is given to data for large urban areas. However, there is still a need to conduct surveys in smaller urban communities, whose requirements can differ from those of the large areas.

1. INTRODUCTION

1.1 Overview

Saskatchewan has recently shifted from low growth to rapid growth Province. Accordingly, the Saskatchewan Ministry of Highways and Infrastructure (MHI) is now preparing for growth in passenger and freight movements on the provincial transportation system, commensurate with the Province's demographic and economic growth.

As part of these preparations, MHI is now investigating the development of a multi-modal, passenger and freight travel demand model (TDM). Other provinces have also deployed province-wide models. The most recent is Ontario's 2009 auto/truck highway basic model and Québec's truck-only highway assignment model. Alberta Transportation and the Ministry of Transportation of Ontario are both now investigating multi-modal models for their respective provinces. Transport Canada has developed multi-modal forecasts of freight flows (although this is not a network model). Similarly, the cities of Regina and Saskatoon – along with other urban areas across the country - are also well along in the development of urban travel demand models; and, as a result, they have developed data that could be important to the TDM.

Following an initial study that investigated the feasibility of developing a province-wide model, MHI commissioned a follow-on study that reviewed and assessed the data that are required for calibrating and running the TDM. This paper summarizes the outcomes of that assessment.

The study began by defining potential modelling approaches. Within this context, an inventory was prepared of existing data sources among Provincial, Federal and local agencies that potentially could be used to develop the model. Over 30 sources were inventoried, comprising both travel data, such as traffic counts and origin-destination surveys, and demographic and socio-economic data. The inventory considered completeness, geographical coverage, representativeness, levels of detail and ease of use, among other attributes. It also examined the ability to forecast the data at the necessary geographies.

Next, the available data were assessed against the needs of the potential modelling approaches. From this, data gaps were identified as the basis for the specification of several data collection options. The options, and combinations of these options, are designed to allow the Province-wide model to evolve as different needs arise and as funding allows.

The paper describes the inventory and the proposed data collection options. Importantly, the options focused on the feasibility of emerging electronic data collection technologies, including cellphones, GPS traces and loggers. The paper also speaks to the practical challenges and opportunities from the use of these data, especially the electronic sources.

Following this introductory section, the rest of the paper is organized as follows:

- Section 2 provides a conceptual definition of province-wide travel demand models.
- Section 3 inventories the available data.
- Section 4 describes emerging data collection technologies and their potential roles.
- Section 5 assesses the available data and identifies needs for new data.
- Section 6 summarizes the findings.

1.2 Acknowledgements

This paper describes the findings of the *Province-wide Travel Demand Model Data Assessment* study, prepared for MHI in September 2013. (1) The study was directed by Miranda Carlberg, Executive Director of MHI's Systems Planning and Management Branch; Ben Liu, Director, Systems Planning; and Stephen Cook, Systems Management Analyst. The study was prepared by David Kriger Consultants Inc. in association with Cambridge Systematics (whose contributions were led by Dr. Arun Kuppam) and Associated Engineering (whose contribution were led by Monique Kealey).

The views expressed in this paper are those of the author alone, and do not necessarily reflect the official policies of MHI or any other government agencies consulted or referenced herein or of any other organization.

1.3 Guiding Principles

Three 'guiding principles' defined the project. These principles were:

- This study focused upon identifying the data needs for a travel demand model. However, there are several types of TDMs and the data needs for these vary accordingly. Accordingly, a first step was to define two broad categories of TDMs as the basis for determining data requirements.

- The feasibility of using cellphone GPS traces and other existing or emerging electronic sources of travel data was of considerable interest to MHI, and this was a key focus of the study.

- The data collection options had to be flexible, in order to accommodate available budgets and resources, which may also change over time.

2. TYPES OF PROVINCE-WIDE MODELS

The structure of the Province-wide TDM informs the data requirements. There can be different forms, which are related to planning applications, data availability and quality. Potential planning applications of the Province-wide TDM include highway twinning and bypasses, capital investment and priorities, property protection, system planning, design, operational and rehabilitation needs, emergency response, traffic impact analyses for new residential developments, resource development and P3 infrastructure.

It was also understood that the TDM would focus on autos and trucks – that is, on road and highway traffic. Public transit would be estimated in the urban models. However, inter-urban bus service and potential commuter bus service to serve the commutersheds of the two main cities could be considered in the Province-wide TDM.

Although MHI's desired model specification had not yet been determined, with these requirements in mind for the purposes of the data assessment project, it was possible to define two categories of models:

- Basic approach, which comprises the estimation of an origin-destination matrix from traffic counts. This approach yields a base (current) year trip table that can be assigned to the base year network. Growth factors based on population or employment projects are applied to the OD matrix in order to develop forecasts. The forecast OD matrix can then be assigned to a future year network. The advantage of this approach is that it uses available count data and can be developed relatively quickly. Its main disadvantage is that it cannot account for any of socio-economic or policy determinants that might impact travel behaviour.

- Four-step model, which accounts for trip generation, distribution, mode choice and assignment. The approach can be simplified if it is assumed that mode choice is not a consideration (i.e., all users will stay in with their current mode), or if observed factors are used to determine mode share. This approach has the advantage of being able to account for socio-economic or policy determinants of travel behaviour. However, it requires an investment in data, especially those describing trip origins and destinations.

Some states and provinces have elected to use one or the other approach. It is also possible to see the two approaches as sequential, with the Stage 1 approach providing the basis to which Stage 2 components are added over time. Either way, it is important to note that the definition of both a TAZ system and a road/highway network for trip assignment is needed. In addition, the TAZ system must incorporate and be integrated with the much more detailed TAZs that are associated with urban TDMs.

3. INVENTORY OF AVAILABLE DATA

3.1 Data Requirements

Travel demand models require a significant amount of demographic, socio-economic, travel and transportation system data. The data requirements for the two types of Province-wide TDM can be grouped into four categories:

- Demographic and socio-economic data - e.g., household- and population-level characteristics. These are the independent variables that are used to estimate travel.

- Transportation networks and traffic analysis zones (TAZs).

- Model estimation data, which are used to estimate the model formulations. Key among these are origin-destination (OD) surveys for people and trucks, and travel time data.

- Model validation data. These data are used to verify the model formulations, notably including cordon classification and occupancy counts.

Table 1 summarizes how the requirements vary between the two types of models. It can be seen that the requirements for the basic model are much less than for the four-step model. The most significant difference between the two is the need for origin-destination data for the four-step approach.

Table 1. Data Requirements for the Four-Step and Basic Approaches

Category	Four-Step Approach	Basic Approach
1. Demographic and socio-economic data	• The household is the key forecasting basis. Of interest are household size, income, workers per household and the number of vehicles per household. • Population, including transient and seasonal. • Employment, by type. • Students.	• Population, ideally including transient and seasonal. • Employment.
2. Transportation networks and TAZs	• Attributes of the transportation network, including highways (classification, lanes, capacities, speed limits, restrictions) and public transit (itineraries, headways, fares, travel times). • TAZ definitions, including shape files, zone attributes (area, type, centroid connectors) and aggregations (districts).	• Attributes of the transportation network, including highways (classification, lanes, capacities, speed limits, restrictions). • TAZ definitions, including shape files, zone attributes (area, type, centroid connectors) and aggregations (districts).
3. Model estimation data (OD surveys) / other calibration data (used to fit the model equations)	• Person OD surveys capturing trip origin and destination, trip purpose, mode, and departure and arrival times. • On-board public transit surveys, if that mode is modelled. • Truck OD surveys, describing vehicle characteristics, origins and destinations (itinerary), activity at each stop, commodity carried and fullness of vehicle. • Travel time surveys. • Volume-delay parameters, time penalties, transit time (if applicable).	
4. Model validation (used to verify how well the model represents actual conditions)	• Vehicle classification counts – auto / trucks separate along the transportation network and at external boundaries.	• Vehicle classification counts – auto / trucks separate along the transportation network and at external boundaries.

3.2 Inventory of Available Data

Governments in Saskatchewan already generate and procure several of the types of data that the models could use. To learn more about the available data, a workshop was held among the data 'owners.'

The workshop was important for three reasons: it explained what a TDM does, as the basis for identifying data requirements; it described the types of data that are needed for a TDM; and, it sought participants' inputs on the types of data they have that could be used for the model. Participants included staff from MHI and several other Provincial ministries and agencies and urban areas. Detailed interviews then were conducted with each participant, in order to understand more about the nature and availability of the data. Sources at Transport Canada, Statistics Canada and selected other provincial agencies also were contacted. A comprehensive inventory of over 30 parameters was prepared. The inventory described each parameter, its source, how the data were collected, how often, and the level of spatial and geographical detail, among other parameters and attributes.

4. EMERGING DATA TECHNOLOGIES AND THEIR POTENTIAL ROLES

4.1 Introduction

In recent years, there have been initiatives in the transportation planning community to use electronic means for gathering data in place of other methods, such as paper surveys and manual observations. This is driven by cost effectiveness (potentially lower unit costs translate into broader coverage), the desire to reduce respondent burden (potentially less intrusion), and the potential to increase accuracy and precision. Accordingly, MHI emphasized its interest in understanding the potential role of electronic data collection technology for this study.

This section provides a context for the ensuing inventory and assessment of existing data, by elaborating on the use of selected emerging electronic technologies.

4.2 Potential Use of GPS

Of particular interest is the use of GPS technologies. Despite the significant improvement in the quality of telephone and web surveys in recent years, there is evidence of some under-reporting of travel. This happens in two ways: respondents sometimes forget the 'bypass' trip that they make as part of the primary trip purpose – a common example is the commuter travelling from home to work who forgets that he stopped along the way to take his daughter to day care (this matters because such stops can impact route, mode and time of day decisions); and respondents who are reporting for the rest of the household sometimes miss the details for the non-respondents (and experience with other recent Canadian surveys suggests that this often manifests itself through higher daily trip rates for respondents compared with non-respondents within the same socio-economic and demographic profiles).

Some household travel surveys in the US employ a subsample GPS travel behaviour study for additional data capture for travel modelling needs. The GPS subsample data show that respondents consistently under-report trip data either due to respondent burden of diary

reporting, or respondent misunderstanding of what a "trip" is. Short and non-vehicle trips (e.g., the noon-time walk to lunch) are often underreported, particularly in urban areas. An improved understanding of the under-reporting of trips can significantly enhance modelling efforts. The use of GPS allows for precise data collection as a comparison with other methods, such as telephone or web surveys.

4.3 GPS Data Loggers

Although GPS can capture where a person (or vehicle) is at a particular point in time and space, alone it cannot record other information about the trip. Transport Canada's Canadian Vehicle Use Survey (CVUS) provides one means of addressing this gap.

Transport Canada has surveyed a sample of light-duty vehicles for several years. Since 2011, the quarterly CVUS has been conducted with GPS units and loggers that allow the driver to record pertinent trip information such as trip purpose and the number of vehicle occupants. Saskatchewan is one of several provinces that is partnering with Transport Canada on this initiative.

The CVUS uses the OttoView custom data logger to survey light-duty vehicles (having a gross vehicle weight under 4,500 kilograms). The Canadian-made OttoView logger fits directly into the vehicles On-Board Diagnostic device (OBD II), which has a standard format for light-duty vehicles. The OttoView-CVS42 records the information described above, as well as fuel consumption, other vehicle engine data and – within municipal boundaries - the posted speed limits. The survey participant can also enter vehicle occupancy, trip purpose, driver and passenger demographics and refuelling costs.

The CVUS is now well established. Its key advantages are that it covers the entire Province and provides precise depictions of origins, destinations, routes taken and travel times and delays. There is an urban / rural stratification, with vehicles registered in Regina and Saskatoon dominating the sample.

However, it comprises only a very small sample: in Saskatchewan, Transport Canada and SGI reported a raw sample of 1,000 vehicles and a net (i.e., participating) sample of approximately 300 vehicles, all quarterly. This equates to a net annual sample of 0.18% of the 683,000 light-duty vehicle registrations in the Province, or 1,200 net (completed) surveys annually. (Note that all values are estimates only.) Other modes - transit, walking and cycling - are not covered.

Commercial providers such as TomTom and INRIX also can provide GPS travel time traces. The advantage of these sources is that they can provide coverage over extensive time periods. However, these data are sourced from actual users – i.e., they represent only a selection of travellers, representing an unknown sample of travellers, and only on selected routes. For example, INRIX can provide travel time data for Saskatchewan as a whole. However, the sources are predominantly drivers on certain highways, and so the actual geographic and network coverage is limited.

GPS units also have been used to transmit background data – for example, they are now commonly used for travel time surveys. The GPS units provide a significant improvement in quality and reliability over previous travel time methods.

4.4 Cellphone GPS Traces

Cellphone GPS traces provide another alternative. Cellphone traces have become popular in congested urban areas, such as Toronto and Montreal, as a means of providing real-time information to drivers on the location of delays and bottlenecks. The proliferation of cellphones means that a large number of data points can provide a 'passive' basis (meaning that drivers do not have to be pulled over or complete an interview) for origin-destination, routing and travel time information. On the other hand, sample sizes cannot be determined reliably, and no information can be collected on trip purpose or the traveller's demographic or socio-economic data. Moreover, the data can be collected only where cellphone coverage exists, and – in Saskatchewan – coverage may vary in remote areas.

Cellphone traces have been used in the United States as the basis for estimating origin-destination characteristics. At the time of the study, the Province of Alberta had begun to investigate the feasibility of using cellphone data for its proposed province-wide model.

Cellphone trace data could be used in lieu of roadside intercept surveys to provide data on trip origins, destinations and routes (although not on trip purpose or vehicle occupancy). Leveraging the CVUS data - which do record trip purpose and vehicle occupancy – then, by 'force of numbers,' the cellphone traces could provide a reasonable profile of external activity from (and to) Regina and Saskatoon. However, the availability of cellphone data from SaskTel Mobility is still under investigation.

4.5 CVUS Heavy Truck

In 2011, Transport Canada successfully tested a version of the CVUS for medium- and heavy-duty vehicles, meaning that the CVUS could accommodate a range of trucks. The OttoView logger was adapted to these vehicles, whose OBDs are not standardized. In addition to the data that are recorded automatically, as with the light-duty CVUS, the CVUS Heavy Trucks allows the driver to enter fuel cost, activity at each stop, cargo type and fullness of the vehicle.

As of the last quarter of 2013, the CVUS Heavy Trucks had been deployed in the partnering provinces. The current plans are to sample 500 vehicles each quarter, equating to an annual raw sample of 4% of the 50,000 medium and heavy truck registrations in the Province. Transport Canada's pilot tests have suggested compliance by truckers of the same order as that of the CVUS; that is, approximately one in three respondents.

The response rates also suggest a net sample of 1.33%, or of the order of 650 vehicles annually in Saskatchewan. This should yield a good basis for modelling trucking activity. However, the distribution by vehicle type is not yet known; and the existing MHI surveys cover only heavy vehicles.

5. ASSESSMENT AND GAP ANALYSIS

5.1 Assessment

With the inventory in place, the available data were assessed in terms of their usability for a future Province-wide TDM. The assessment considered data quality and possible biases, level of detail, timeliness (how up-to-date are the data?) and 'forecastability' (can the data be projected?). **Table 2** synopsises the assessment. In sum:

- A key consideration is the availability of certain data at the Traffic Analysis Zone (TAZ) level. Given that the TAZ system for the TDM has yet to be developed, the appropriateness of the data in terms of level of detail could only be estimated; nonetheless, the ensuing discussion can be considered to be reasonable and generally indicative.

Table 2. Data Assessment

Category	Status of Data
1. Demographic and socio-economic data	• Household: Census only, at Census Subdivisions (CSD); needs detailing in urban areas. There are no forecasts. Some data available from the Census, and others from the NHS (although the usability of the latter at small geographies must be determined) • Population: Census at CSD; by age and gender – needs detailing in urban areas. The Ministry of Health does annual updates, but for 13 Health Districts. The Bureau of Statistics (Ministry of the Economy) is preparing a population model and forecasts, but at MHI District level. The possibility exists for further geographical detailing. • Employment: Some labour force data are available from the NHS (although their usability must be determined). Other data on labour demand and supply may be available from Employment and Social Development Canada. The Ministry of the Economy and SaskBuilds (an agency of the Province) are developing economic forecasts. • Students: Estimates are available from population age cohorts and from the Ministry of Education • Synopsis: o ✓ Population data are in best shape – but need detailing to TAZ o Households – needs detailing, estimates. Some data are available (e.g., size), the availability of other data is to be determined (e.g., workers from the NHS) while other sources are needed for other data (e.g., income). ▪ ✓ Vehicles – can calculate from Saskatchewan Government Insurance (SGI) o Employment – needs estimates: the SaskBuilds and Ministry of the Economy models might help. o Students – needs estimates.
2. Transportation networks and TAZs	• Transportation networks: Data for preparing road / highway networks generally available at required level of detail. Transit networks are available for Regina and Saskatoon (although the need for modelling transit had not been determined). • TAZ: Jurisdictional and other geographies exist in Provincial (Information Services Corporation) / MHI cadastral maps/GIS. However, the actual TAZs must be defined. • Synopsis: o ✓ Data available for defining transportation road / highway network; also transit networks if necessary. o ✓ TAZ must be defined, but can be done in GIS.

Table 2. Data Assessment

Category	Status of Data
3. Model estimation data (OD surveys) / other calibration data (used to fit the model equations)	• Origin-destination data: Origin-destination surveys have been conducted only in Regina (2009) and Saskatoon (2013). There is no Province-wide survey. The Canadian Vehicle Use Survey (CVUS) is a GPS/logger survey that has been conducted quarterly since 2011; however, it samples vehicles as opposed to households. The Census Place of Work / Place of Residence (POW/POR) linkages provide a reasonable base for home-work commute; however, the availability of small-geography data from the NHS is not yet known. Cellphone GPS traces could be a potential source, although their availability is not yet known. • Truck OD surveys: The MHI weigh station truck OD surveys are comprehensive roadside intercept surveys of inter-urban heavy trucks at selected points on the Provincial highway system. Other sources include the National Roadside Survey (but only at selected points and for certain years), the new CVUS Heavy Trucks, and GPS fleet tracking data (routing and ODs). • Synopsis: o For passengers, could estimate from existing sources, but these are limited. One option is the CVUS, augmented with additional samples. ✓ Traditional household or roadside surveys are other options. o For trucks, ✓ the existing MHI surveys provide excellent basis. Other potential options include the ✓ CVUS Heavy Trucks, with add-on sample and ✓ commercial fleet tracking data (GPS only).
4. Model validation (used to verify how well the model represents actual conditions)	• Vehicle classification counts: ✓ MHI has comprehensive program, using FHWA classifications; but some sites are not up to date hence the cited volumes may be estimates only. Some of the municipalities also have counts. Screenlines can be defined on the MHI network, subject to the definition of the TAZ system. • Synopsis**:** main need is to update and maintain counts.

- The checkmarks (✓) in the table indicate either that a parameter is available and is usable for the TDM (albeit with some modifications, as described in the table and below), or that the parameter has a potential application for the TDM.

- Up-to-date data for several key parameters exist at appropriate levels of geographical detail. These parameters are population, households, households by size and households by number of vehicles; road / highway network definitions; and traffic counts and travel time studies. Subject to how the TAZ system is eventually defined, the aforementioned TAZ-level population and household data can be derived from the existing Census of Population data, subject to suppression of data for confidentiality. Similarly, the network definitions also would fit an anticipated TAZ definition. It is *likely* that the existing traffic count locations and travel time studies would provide sufficient coverage. The eventual TAZ definition would determine the need for additional counts: in any event, the more important need is to ensure that existing count are up-to-date.

- The National Household Survey (NHS) is expected to be the source of data on household income, workers, occupation and the Place of Work / Place of Residence linkages. However, the quality of the NHS is still being determined. The voluntary NHS was distributed to 1 in 3 households, and achieved a 65% response rate. The resultant

20% coverage of Canadian households is equivalent to the 1 in 5 distribution that the predecessor mandatory Census long form had. However, the distribution of the NHS responses has not been revealed and, accordingly, geographical, demographic or socio-economic biases are unknown. As well, Statistics Canada suppresses data if it has concerns about quality (in addition to maintaining minimum numbers of respondents so as to ensure respondent confidentiality). Statistics Canada publishes a global non-response rate for each of the data summaries it issues. This reflects both complete non-response (i.e., the household did not respond at all) and partial non-response (some questions were not answered). However, no further details are provided. Note that Statistics Canada will not publish data for a given geography of the global non-response rate exceeds 50%. (2) The resultant unknowns represent a potential significant impact on these sources of information.

- Forecasts are available only for population. Although the current efforts are for district-level geographies, the ability exists to develop projections for smaller geographies (i.e., at the TAZ level). Even though TAZ-level population estimates can be developed from the existing Census data, some analysis would be required to develop forecasts of TAZ-level population (and household and other demographic and employment information): This is because forecasts typically are prepared as 'top-down' estimates (that is, Province-wide forecasts are estimated first according to broad indicators such as birth and deaths rates, immigration, etc.; and smaller-area estimates are then derived from these). No forecasts exist for population by age or gender, for households, or for household-related data (size, income, workers, vehicles).

- Other than the NHS, employment data – demand and supply – are available only at high-level geographies, although they are available by economic sector. Forecasts are being developed for district-level geographies, but assumptions and further analyses would be required to translate these into jobs and into smaller geographies. Current and projected land use data provide a means of estimating jobs; however, the translation is not precise.

- Road / highway network data are readily available from various sources within MHI and the Provincial government, at a high level of detail and quality. Transit routes can be added, if needed. The capability exists to define TAZ boundaries within the existing GIS systems.

- There is no single Province-wide origin-destination passenger survey. However, several components of passenger travel exist, notably: the two urban surveys; the CVUS (although this covers vehicles only, not other modes); the Place of Work / Place of Residence linkages (a proxy for inter-urban commuting); and selected visitation surveys.

- MHI has a comprehensive set of truck origin-destination survey data. The quarterly CVUS Heavy Truck survey, which commenced in autumn 2013, provides the potential for filling in some of the gaps in the existing data (notably, capturing different types of trucks, and not just heavy trucks). Other sources, such as commercial truck GPS traces, also have a potential use as the basis for a Province-wide truck origin-destination trip table. There is a lack of data on urban trucking activity, which the CVUS Heavy Truck survey could partially address.

5.2 Data Gaps and Options for Addressing Gaps

Three data gaps were identified. The gaps, and options for addressing them, are summarized below.

5.2.1 Province-wide Origin-Destination Survey

The principal data gap was a Province-wide passenger origin-destination survey. There were different options for meeting this gap:

- Conduct a Province-wide origin-destination survey. This would be a household telephone survey, similar to the Regina and Saskatoon surveys. Because the two cities have (or will have) recent surveys, the Province-wide survey would exclude residents (households) within Regina and Saskatoon, but would cover the remaining cities and rural parts of the Province outside these two urban regions.

- Increase the number of vehicles included in the CVUS in order to capture a larger sample. This would increase both the representativeness of the data, and allow for a stronger representation of vehicles outside the two large cities.

- Combinations of the above two surveys. For example, an augmented CVUS could be complemented by a household origin-destination survey that captures multi-modal travel in a smaller community and its surrounding rural area; and the latter survey could be used as a proxy for other similarly sized communities. Roadside intercept surveys also could be used.

- Cellphone GPS traces. Again, the first consideration is to determine the availability and feasibility of these data for use in Saskatchewan.

5.2.2 Truck Origin-Destination Surveys

For truck O-D data, the existing MHI data provide a strong basis for developing truck origin-destination matrices. These provide a comprehensive profile of truck activity on Saskatchewan highways. However, they comprise only heavy trucks, they do not cover all highways and, although many locations are surveyed over several years (in some cases annually), MHI's surveys do not capture seasonal variations. They also do not include urban truck activity.

Going forward, the CVUS Heavy Truck provides a potential source of data that would provide broader geographic coverage and would include a full range of vehicles. An augmented sample similar provides an opportunity to enhance the number of survey.

GPS traces from commercial fleet managers provide another source of origin-destination routing information, though only for heavy trucks and likely not including small fleets.

5.2.3 Traffic Counts (Vehicle Classification Counts)

The need to update existing traffic counts also was identified. The purpose was to fill in gaps at existing count locations, in order to record the changes in traffic volumes that have occurred due to the recent rapid population and economic growth. The object here was to update the existing annual MHI and municipal count programs within their respective jurisdictions.

Maintaining current count data has several benefits, in addition to assisting the development of a TDM. These include applications for pavement management and rehabilitation, traffic operations and safety (and, it follows, for helping MHI set funding priorities and establishing capital and operational budgets).

5.2.4 Other Gaps

The usability and availability of the National Household Survey data at small geographies could be considered as a potential fourth gap. However, this cannot be determined until (or unless) the data are actually mined for a potential modelling application. In the meantime, addressing the first gap (a Province-wide origin-destination survey) could obviate potential problems with small-geography NHS data.

5.3 Summary - Data Collection Options

The data collection options described above were developed into various combinations of survey methods, technologies, sampling sizes, scopes and budgets. The intent was to provide flexibility for MHI, in order to adapt if Ministry budgets and priorities vary.

In addition to varying budgets, some other possible needs were noted:

- The ability to forecast key parameters such as population, household and jobs is tied directly to the definition of the TAZ geographies. In other words, although current data are available, MHI will need to consult with the Bureau of Statistics, the Ministry of the Economy and SaskBuilds to understand how their new forecasting models can be applied or modified to work with the desired TAZ system. Assuming that any collected OD data would be geocoded, this would not inhibit the collection of data; however, it could slow down the post-survey analysis and tabulation.

- The uncertain availability of some data sources at the required geographies might require changes to the data collection options – notably, cellphone GPS traces and Statistics Canada's 2011 NHS data at small geographies.

To some extent, the options mitigate these possibilities by allowing some degree of substitution between traditional methods – such as household travel surveys – and GPS-based survey technologies.

6. SUMMARY

The Saskatchewan Ministry of Highways and Infrastructure proposes to develop a Province-wide travel demand model. A necessary first step in developing the model was the identification

and assessment of available data that could be used to develop and apply the model. This paper summarizes the findings of that assessment.

Three main findings serve to summarize the main points of this paper.

- Several demographic, socio-economic and travel data sources were available, although some of them are not fully up-to-date and others require detailing at small geographies. MHI's data inventory and assessment were predicated by the newness of the Ministry's proposed modelling initiative, and other agencies that already have more experience with models might not feel the need for such an analysis. However, the inventory served to provide a comprehensive catalogue of both existing and potential sources, as well as ways to forecast these data. In many cases, potential new uses for long-established data were identified (notably, for the CVUS). This in turn broadens the potential applicability and cost-effectiveness of collecting a particular data set. It also identifies potential needs for existing data sets – for example, broadening the usability of important data sets, such as the NHS, at smaller geographies.

- Three main data gaps were identified: Province-wide passenger origin-destination data, expanded truck O-D data, and updated traffic counts. Alternate options for collecting these, comprised of different combinations of traditional and electronic methods, were proposed. This flexibility allows a data collection program to move forward, even in the face of changing priorities and available budgets.

  It also should be noted that, as an ancillary product of this exercise, the study established that the collected data would be of benefit to other Government functions, such as safety, design, construction, corridor planning, Official Community Plans, environmental assessments, land use planning and economic development planning.

- Electronic, GPS-based data collection technologies have some potential for addressing the data gaps. However, their availability and actual usability in Saskatchewan is not yet fully determined. It is the author's understanding that there are similar interest and implementation experiences elsewhere in the country, both for urban and inter-urban situations. As a result, then, although these technologies offer much promise, some issues must be addressed in order to promote their use in both urban and inter-urban model development and transportation planning applications. These include:

  - Availability of commercial data sources
  - Geographic and modal coverage
  - Spatial and temporal precision
  - Accuracy
  - Statistical representation
  - Data processing burden
  - Privacy and confidentiality
  - Cost

7. REFERENCES

(1) Province-wide Travel Demand Model Data Assessment Study, Final Report. Prepared for Saskatchewan Ministry of Highways and Infrastructure, Regina, September 2013.

(2) See http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/fogs-spg/Pages/GNR.cfm?lang=E&level=3&GeoCode=433