How Do You Decide When To Repeat Transport Surveys?

Marcus Wigan[1], Nariida Smith[2] and Ben Timmis[3]
[1] Principal Oxford Systematics, Melbourne, Victoria, Australia
[2] Sydney Manager, Meyrick and Associates, Sydney, NSW, Australia
[3] Consultant, Meyrick and Associates, Melbourne, Victoria, Australia

1 Introduction

1.1 Sustainable Data for Sustainable Transport

Sustainable transport requires a perspective that extends beyond the present and beyond the transport system itself. To view sustainability properly requires more sophisticated models than we currently employ, and more data than we currently collect. As we extend the capabilities of models, and require better monitoring of transport impacts, the scope and scale of data needs increases significantly. Scope increases as environmental data items, covering issues such as air pollution or greenhouse gas emissions, are needed. Scale increases as both temporal and spatial detail expands. For example, morning weekday peak commuter travel data to key destinations was once sufficient for planning transport capacity. However to monitor the environmental impacts of transport, or assess measures to reduce them, data covering travel throughout the day and the week is needed. Moreover the impacts of travel for all purposes to all destinations need to be assessed, requiring fine spatial detail.

There are two basic characteristics required for sustainability assessments and monitoring which are particularly data hungry:

- *Taking account of interactions:* Interactions between different transport, activity and physical systems are necessary components of any serious sustainability information system. One dimensional assessments of, for example, a single mode, can ignore the effects on all the others.
- *Extending over time to ensure that sustainability is maintained:* Impacts of measures which seem successful in the short term need to be tracked in case benefits unwind over time. Sustainability demands a longer term perspective, and the data and analysis framework must follow the same path.

Requirements for extra data, for data update for tracking over time, plus data management systems to manage large and diverse data sets, place strains on transport data budgets. In this situation the data manager asks *"when is it worth – or when is it necessary – to invest in updating this information?*" This is particularly an issue for updates of transport surveys which are often costly to conduct. Yet the type of information collected in transport surveys, providing insights into revealed and stated behaviour and preferences, is particularly important when moving to a more sustainable transport system depends upon behavioural change. Regular updates are likely to be needed and a continuing horizon for data review and refreshment cannot be avoided.

This paper, which builds on work done for the Queensland Department of Transport (Meyrick and Oxford Systematics, 2006), addresses a process model for dealing with this crucial question. It develops a way of ensuring that it is problem based rather than an attempt to upgrade existing data wholesale, or fill gaps for the sake of it. Many different types of data are required, and to be able to prioritise where the value is greatest in securing fresh information is critical. A better system than we currently have to address this in a timely and balanced manner is needed.

1.2 Practical needs

It is simply not (yet) practical to gather data continuously on all – or even most – of the factors that one might need in order to monitor, assess and predict changes. Continuous surveys have been designed more to spread the load of gathering enough data to accumulate to a statistically reliable total, rather than to detect significant changes in what is inevitably a noisy background.

The sheer diversity of customers and users of transport survey data also complicates the decisions as to how frequently, how much and with what level of statistical reliability is needed – often for the self-same variable! It is clearly not possible to satisfy all potential users at the same time, on either practical or budget grounds. So, what alternatives are there?

A four pronged approach was adopted which considered: a national and international survey practice via a scan of document sources, review of the research literature, direct international experience via international colleagues; and views of Australian transport survey practitioners and researchers.

A key finding of the literature scan was the thinness of available literature at the conceptual level regarding the setting of survey update frequency. While most decisions seem to be made on pragmatic grounds it was, however, possible to derive key issues relevant to survey update from reports on both survey theory and practice. The literature review confirmed that the questions that one has in terms of transport data, such as what processes could be used to determine when to update, have seldom been formally addressed.

It is clearly an important question. Major surveys are expensive, even with modern data capture aids, and there are two separate needs:

- Is there a basis for determining when - and how much - data (in terms of variety, and statistical and geographical reliability) is needed;
- Is there a process for determining this in any given case?

Thus, at the same time as developing a conceptual decision framework for survey update, we sought the advice of national and international experts about the state of practice. This was in order to find if or when survey intervals were explicitly considered as a policy issue, and also to find more detail about how they were set. This elicited an enthusiastic response, often citing how significant and poorly addressed these questions were.

1.3 Survey types, characteristics and contexts

For the purposes of this paper ‘Surveys’ are defined as activities based on questioning or observation, and ‘Transport’ is defined as the operation and use of transport. These definitions exclude automated data collection processes by equipment, such as Intelligent Transport Systems, and surveys of technical characteristics of transport vehicles or transport infrastructure, such as road pavement conditions.

There are numbers of options for categorisation and sub-categorisation relating to survey operations and purposes. Three categorisations relevant to update frequency are:

- Application domain or area;
- Survey approach or style;
- Types of data collected.

Table 1 Survey domains and selected areas of applicability

	Sub Types	Urban/ Conurbation	Regional Cities	Intercity & Regional
Household travel surveys	▪ resident/location and person based	✓✓	✓	-
	▪ activity surveys	✓	-	-
Freight and Commercial vehicle surveys	▪ Freight cost monitoring	✓	-	✓
	▪ Weight	✓	-	✓✓
	▪ O/D and routing	✓	✓	✓
	▪ Commodity flow	✓	-	✓✓
	▪ Cost/fleet monitoring	✓	-	✓
Public Transport surveys	▪ O/D	✓✓	✓	-
	▪ Boarding /interchange capacity	✓✓	✓	✓✓
	▪ Quality crowding/reliability/variability/	✓✓	-	-
	▪ Perceptions of service	✓✓	-	✓
Visitor surveys	▪ Inbound destinations	✓✓	-	✓
	▪ Length of stay	✓✓	-	✓
	▪ Expenditure	✓✓	✓	✓
Traffic surveys	▪ Classified/Unclassified vehicle counts	✓✓	✓✓	✓✓
	▪ Travel time	✓✓	-	✓
	▪ Vehicle occupancy	✓✓	-	-
Non-motorised transport	▪ Pedestrian and Bicycle O/Ds	✓	-	-
	▪ Volumes	✓✓	-	-
Safety	▪ Compliance: seatbelts/helmets	✓✓	✓✓	✓✓
	▪ Attitudes	✓✓	-	-
	▪ Incident investigations	✓	✓	✓✓

Table 1 presents the first of these grouping transport surveys, with example sub-categories, in terms of their 'usual' practical applicability to travel in urban areas; travel in regional cities; and regional travel; which are shown as *often applicable* ✓✓, *sometimes applicable* ✓, and *seldom* or *never applicable,* respectively.

Table 2 shows categories of survey approach, with some comments on the frequency issues involved in finding a sufficiently representative sample to provide data 'fit for purpose' in each case. We use the term *approach* to reserve the term *method* to apply to the *survey collection method*, such as by face to face interview, mail out or computer assisted telephone.

Table 2 Categories of survey approach

Survey Approach	Description	Frequency Issues
Continuous surveys	Collect data on an ongoing basis building up sample size over time data pooled across periods	Explicit: none, as survey is continually updated Implicit: assumes data collected over a period is sufficiently stable for it to be pooled
Incrementally updated surveys	Periodic update of selected items to 'refresh' data	Frequency issues include both how often and how much
Panel surveys	Tracking same cohort of individual households or firms over time, usually monitoring behavioural change	Explicitly: can provide information about change frequency and hence required update frequency Implicitly make assumption about likely change rates in setting panel interview frequencies
Cross sectional surveys/ updated cross sectionals surveys	Survey at one point in time When update is required, the entire survey is repeated, preferably in a manner comparable to earlier collections.	Since these are one off rather than part of survey program update, issues are not considered until the decision to survey again, usually due to perception of the data being out of date.

There are many typologies possible, but in the sustainability context rates of change are especially critical.

Table 3 Change drivers potentially influencing data update

- *Land use changes* such as suburban development in rural areas;
- *Social changes* such as changes in demographic profiles;
- *Economic changes* effecting shipment of goods and employment and thus travel to work;
- *Technological change*s such as in information communication technology;
- *Performance reporting changes* such as additional performance measurement and reporting requirements by governments and operating agencies.

Typical Rates of Change: The type of data being assembled by survey also affects update frequency. The signal for update will usually come for external information suggesting change, but certain types of data have a natural pace of change. Thus we can suggest five major areas for such a categorisation of transport and traffic database assembly and maintenance.

In approximate order of volatility and rate of change, these are:

- Direct information on infrastructure elements;
- Indirect information on planning;
- Direct information on transport demand for passenger and freight;
- Indirect information affecting traffic movements and transport demands;
- Direct information on traffic movements.

Variation with Location and Epoch: Within, over and above these typical rates of change are variations with place and time. For example, in urban fringe areas land use and road infrastructure may have been constant for 20 years and then change in the space of two years with the establishment of a new development in one area, while surrounding areas remain unchanged.

Total Database: In a wide range of cases, the data from surveys is a minor part of the total operational and planning databases. However, the proportion of time and resources spent considering and commissioning the collection of these types of data is usually high compared with that spent in other data assembly. The place of survey data in the total data needs of the organisation could be an important consideration in determining the need for data update. While primary data collections are often 'one off', there are both effectiveness and efficiency benefits in including an overall program of surveys within a more general data update framework.

2 International experience

2.1 Survey of experts

We found no explicit treatment of survey frequency determination in the literature. Moreover, national and international practitioners and experts, contacted in an initial survey about usual practice, reported survey update in reaction to triggers rather than as part of a proactive data maintenance strategy. Thus as the matter clearly required further investigation, we went back to the experts with questions about principles as well as practice. The e-mail survey instrument, intentionally informal and open ended, is shown in figure 1.

Email Survey

As a result of some policy related work I have been doing, I've been searching for anything in the literature addressing the methodologies used to determine the appropriate frequency to undertake transport (and to some extent traffic) surveys. The results have been a large number of survey design papers, but none (literally) on the methodologies for determining what intervals should be left between surveys.

As this is a question that many of us will have considered (if not encountered) over the years, it seemed sensible to check on a few questions to see if there was any common experience on this question.

If you had a few minutes that you could spare to comment on the following couple of questions and email the response back to me, I undertake to collate the responses and send them back to you in a summary form.

Q1 Has the question of the appropriate interval between a transport (or traffic) survey been raised with you so that you had to answer it?

IF NO to Q1, Have you considered the matter and published anything?

IF YES to Q1, Any details/reference you might care to add?

Q**2 What are your views on the key reasons for the intervals at which surveys you know about have been repeated?**

1 Following a policy for repeats at specified intervals?
2 A new major project?
3 A new policy strategy?
4 Simply the ability to secure a budget at the time?

Any general comments would be very much appreciated.
Many thanks.

Figure 1 Email Preamble and Questions

The survey drew a much more enthusiastic response than our initial e-mail about current practice. Thirty eight replies were received from Australia, USA, Belgium, France, UK, Canada, Chile and Switzerland. Not everyone answered all the specific questions but all offered extensive general comments. Several respondents indicated that this was a question of considerable importance and interest and had not, to their knowledge, been addressed directly before.

While the number of respondents is not large, the parties responding have also been involved in, or advised upon, many surveys and so their responses represent a large number of surveys. Thus it is quite significant that nearly half of the respondents had never been asked to consider appropriate survey intervals. Figure 2 shows the percentage of 'yes' responses to survey questions.

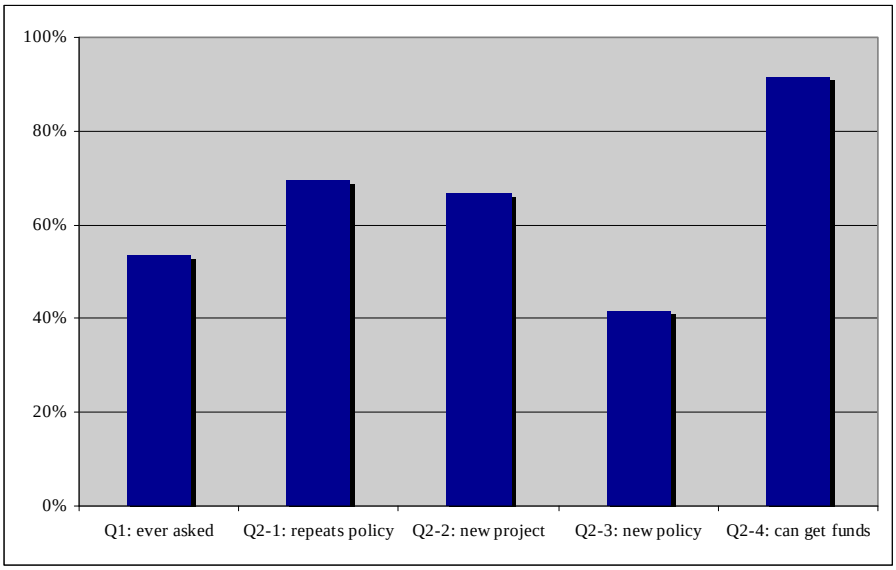

Figure 2 International Experts Responses to Survey Questions

Availability of funds is the most important factor. It seems that in general most people agree with a response to our initial survey: "*Transport survey update should be carried out whenever funding is available, that is unlikely to be too often*".

While more than half the respondents cited a frequency policy as a determinant for update, the comments accompanying the responses showed that these updates where generally for panel or continuous surveys rather than update of cross sectional data. Respondents provided references to these types of studies:

- The London Area Transport Study reports (LATS, 2001a and LATA 2001b) considered the options for the current cycle of London surveys for passengers and freight. As it addresses the mixes of cross-sectional updates to be handled within the rolling process, it has some direct relevance to the current issue.
- A paper on continuous commodity flow survey methodology (Southworth, 2003) again includes the context of update intervals but within an argument for a continuous survey process.

Respondents also advised the only direct treatment of methodologies for timing survey updates known to them was by Slavik and Stopher (1979). Another paper by Stopher reported actual survey frequencies in the US (Stopher and Metcalf, 1996) although, like a later German paper, Kunert, Kloas and Kuhfeld (2002), examining 'the state of practice,' the question of optimal frequency was not addressed.

Thus it is valuable to address both data specific issues and pragmatic issues in setting a framework for survey updates.

2.2 The international literature

In the absence of literature directly addressing principles for survey update frequency it was necessary to draw lessons from related literature.

Survey methodology literature, while extensive, predominantly treats survey frequency in the context of longitudinal surveys or continuous surveys (Ruiz, 2004). It shows that an interval, even where it is a logistical consideration rather than data accuracy, usually dictates a frequency of repetition. The most useful finding was the concept of using coefficients of variation for detecting change (Richardson, 2003). Although Richardson, the author, did not apply it to frequency determination, it might be applied in this way.

Modelling literature (Wigan and Southworth, 2005) contributes primarily in our context by emphasising the importance of clarity about the purpose and accuracy of the model before data collection is commenced. This is doubly important in the case of updating data where the benefit of the update will be closely tied to the accuracy and intended application of the model.

Performance measurement literature and related asset management literature address a set of survey applications to track operations and investments (Harrison, 2002). Checks are to confirm status rather than to seek change. Therefore frequency is set by when that confirmation is needed, whether of operations, performance, or cost. It is a management decision related to maintaining confidence in operations. While that confidence may be reduced by fluctuating performance, and lead to more frequent checks, steady performance will not necessarily lead to fewer checks.

Indirect references to survey update in these assorted literature sources support the experience of international experts that decisions are seldom made on the basis of data collection principles. This means that there is rarely any use for a quantitative scoring framework using a multi-criteria analysis framework, where users rate the likelihood of meeting the criteria and assign an importance rating to each criteria to come up with a 'go/no go' decision. Where external triggers strongly influence update decisions, it is more valuable to have a qualitative assessment process, which takes the user step by step through a set of assessment questions.

The literature yielded a set of factors and principles that affect how often transport data collections need to be updated to ensure that valid and usable data is available for strategic planning and decision making. These ranged from pragmatic context issues to statistical validity measure. When setting frequency, it should be valuable to consider some of the following issues:

- Technical assessments of survey updating processes, avoiding the need for a full fresh survey (Bayesian methods);
- The levels of information required to be able to transfer and adjust findings from other locations;
- The technical basis for determining frequencies versus monitoring triggers that will demand a fresh survey or survey update;
- Effective use of external, non-transport collections from planning, administration or nation census dates or other sources, to determine when the next survey is required;
- Exploitation of new technologies for data acquisition, and reducing and processing the types of data that can be secured this way to meet client requirements;
- The ability to link different data collections together reliably and effectively, which has a similar effect, once the outcomes have been communicated.

These issues can affect frequencies of collection and most of them also have the potential to increase the use of incremental updating due to the larger range of data available to bear on any given issue. Better metadata can make this possible, both technically (Westlake and Wigan, 2007) and practically.

Timing Extent Duration: There are a range of durations and design strategies. All surveys take some time to execute, but execution time extends too many months or even years in rolling surveys. Data collected continuously, in one location after another (Garham and Maxwell, 1990) or in small samples, are drawn from all areas/locations to build a full data set. They face problems in areas of rapid change so need to be tied to extent and sampling requirements, including spatial factors. Major metro areas can have the greatest complexity of transport and traffic issues, as well as rapid rates of change. Inter-city movements are generally simpler. The reliance on external surveys, and transfers of findings and values for other areas, is therefore greater on some levels of scale than others, increasing reliance on such sources.

External Triggers: Similar sets of events were reported as survey triggers in the USA, Europe and Australia. They varied from one off unexpected special events to periodic events, such as service contract monitoring. Triggers mentioned included:

- A major project or event such as the Olympics, a toll road proposal or a new bridge requirement;
- Sudden funding availability, often as part of an electoral cycle, or, as in the US, a new five year transport bill such as ISTEA, SAFETEA, etc;
- Timing and availability of a major complementary government survey;
- A policy requirement for a performance indicator or contract requirements where performance evaluation is tied to the contract.

It is worth noting that these are mostly triggers for seeking **current data** rather than survey triggers per se. 'Let's conduct a survey' may be the right response, but consideration of the likely currency of existing data and/or alternative options to a survey might be equally effective in some cases. Community dissatisfaction with road or public transport performance can also trigger demands for data. Garling and Friman (2001) discuss the influence of negative incidents perceived by customers in determining the frequency required from public transport surveys, generally as part of quality assessments (Friman, 2004). It is therefore valuable to address both data specific issues and pragmatic issues in setting a framework for survey update.

3 Decisions and framework

3.1 Three decision groupings

A generic decision framework was proposed to be used to either decide if it is time for data update by survey or, with suitable adaptation as shown in section five, to set the timing for a program of data updates. The more usual case of deciding whether 'now is the time for an update survey' is treated first. The decision framework needs to cover three broad groupings of the possible decisions:

- *Issues specific to the data itself:* The need for update will depend on factors such as the degree of change in the data since last collection, and the ability to detect or make use of such change. These issues should always be reviewed, even if external factors dictate data update.
- *Update trigger issues:* A range of triggers exists, from administrators or legislators calling for new data, to new policies or projects needing that data. The best means of data update should be considered, even if the policy specifically dictates data collection by survey, and especially if it doesn't.
- *Data collection by survey:* Are there alternative options to data collection by survey? Is a full survey needed? Questions of both cost and accuracy need to be addressed.

Our generic framework encompasses a step by step set of questions, which enables the user to assess the necessity and value of a survey or alternative data collection strategy and to feed information into a business case for the survey.

While the framework is generic, it is designed to be applied to particular surveys in particular contexts. It begins with *the premise that the data required is of value and considers whether update of the data is of value*. Thus, it enables the user to specify the data required and the type of survey envisaged. It also provides for initial cost estimates of the process so that these may be compared with both the cost of alternative options for data collection and to the benefits of having updated data compared with using existing data.

3.2 Data specific steps

The proposed framework steps through the issues that need to be considered in deciding on survey update. The process would provide inputs to a business case for a project-specific survey and would begin with the assumption that there is a specific intended use for the data to be collected, rather than just a general intention to build a database of information to be available in case of need.

Questions address data needs, leading to an evaluation of the case for new data. This includes existing data, likely changes since collection, proposed collection method, and the advantage of new over existing data. The fact that existing data is 'old' is not a sufficient case for data update. It is more important to establish that the new data will be more useful than the existing old data.

The value of data depends upon more than its technical specification. All evidence from the study points towards influences beyond technical data specification that influence updates. A decision framework thus needs to include pragmatic, as well as data specific, influences on updates.

3.3 External influences and triggers

The range of external influences on update can be grouped into three generic categories:

- Perceptions about data currency
- Events or policies
- Opportunity

Although more than one of these triggers might be present, usually one is sufficient to lead to data update. Thus the decision framework is not a step by step choice. Instead, it is a dichotomous choice where mere presence of the factor will signal time for data update. However for consistency it is desirable to present these yes/no decisions in a similar format to that used above.

3.4 Survey alternatives

A growing number of alternative means of data update exist for replacing or augmenting survey data. The major options to consider are Intelligent Transport Systems (ITS); partial survey update; and parameter or data transfer. The range of available procedures and technologies is very large. Hence, a comprehensive study of all options is outside the scope of this paper. However, a brief description of their potential is needed before considering their place in a decision framework to consider frequency of data update by survey.

Intelligent Transport Systems: A large amount of data is now available from automated systems. Potential data sources and their uses recognised early by authors such as Margiotta (1998) offer numbers of alternatives to direct surveys, as shown in table 4.

Table 4 ITS data sources and data elements (adapted from Margiotta 1998)

ITS DATA SOURCE	POTENTIAL DATA ELEMENTS
freeway and arterial surveillance sensor data	vehicle volume, vehicle speed, vehicle classification / weight, estimated travel time
video surveillance data	vehicle occupancy, vehicle classification, vehicle stops, queue length
Transit systems	passenger boardings, trip patterns, trip distances, schedule adherence, rideshare/paratransit requests
incident management logs	incident response times, cause, type, extent, and duration of incidents
commercial vehicle operations, administrative and clearance systems	commercial vehicle counts, commercial vehicle identification, type/quantity of freight
probe vehicle data	travel time/travel time reliability, travel distance and patterns

Data and Parameter Transfer: Techniques are now available to transfer data from other places or times and recalibrate them to current local conditions, using available secondary data such as traffic counts and small, specific target surveys (Greaves, 2001). Related options simulate survey data, such as household travel, entirely from secondary sources (Greaves and Stopher, 2000; Stopher, Bullock and Rose, 2002).

Update Sample: Methodological advances such as the application of Bayesian statistical methods are now available to allow data update with small targeted samples. Similar methods may also be used to transfer data from other places. Aggregate data was used to transfer parameters from the US 1995 National Personal Travel Survey to update the 1997 Baton Rouge Personal Transportation Survey. Bayesian updating was also used to update samples from the 1995 NPTS, with values from a small 1998 survey in Baton Rouge which yielded a slight improvement (Wilmot & Stopher, 2001).

4 Framework application

4.1 Should we update?

The sequential assessment proposed in the framework steps applies to a wide range of survey update issues. The three major areas of application are:

- Reactive response: Is data update needed in response to a particular call for data?
- Routine action and review: Should scheduled updates go ahead?
- Proactive forward planning: Setting a future program of regular data updates.

When a new development requires a data update, the lead times will not always provide enough time for an informed response. This increases the policy and application risk either from using data that is not ‘fit for purpose’ or committing resources to collection when there are other options for providing data.

The Framework developed provides an opportunity to give a considered response to a call for data. The key questions that should be asked are:

Is it necessary to use data at all? Available information may suffice, and survey data may be found not to be appropriate or needed once the requirements have been checked through the Framework. If data is required at intervals for a specified or contractual purpose as a matter of record, then it must be collected. However if new data is required, but a survey is not, the data could in many cases be estimated by means other than measurement.

Is what we have enough? The adequacy of the available data needs to be checked using the Framework. It is quite possible that the data already held may suffice for the particular purpose.

Do we need to update our data holdings? While the data held may not be very current, the robustness of the desired results may be sufficient to avoid the need for updating.

Is there a way of avoiding an update? The data and its accuracy and age may be insufficient to give an adequate response to the requirements, but alternative types of data sources may either complement the available data, or indeed replace it. This issue is addressed in the Framework: for example, ITS data may be used to deliver the mean network speeds from continuous flow collections, instead of undertaking or updating special travel time surveys.

4.2 Regular updates

Policies and processes in place may call for the periodic update of data. Indiscriminate updating with an overuse of resources may be avoided if the update is questioned when it falls due. The Framework developed here may be applied to answer these questions:

Is a scheduled survey update needed? The Framework specifies attention to the effects on outcomes by data errors or shortfalls. If little has changed since the last update, it may not be necessary. The rates of change are also required to be considered, as are the effects of taking no updating action at all.

Should a scheduled update be undertaken? The rate of change of factors of interest has a major impact on the decision to proceed with a scheduled routine update. This question is the mirror of the previous one.

Is a routine schedule of updates required? The Framework focuses attention on the adequacy of available information for a specific purpose – or range of purposes – and the rates of change then become the key factor in determining the schedule. This implies that a **regular review process** rather than a regular survey update process will often be appropriate in a process of **proactive forward planning.**

What is the role of regular updates? While the Framework leads in the direction of intervals set by rates of change, the lead times for response when trigger events occur means that a regular update program should be considered as well as a case by case approach. Regular use of the Framework for individual cases again leads to the recognition that a combination of case by case response decisions and a basic timetable is needed to ensure appropriate responses.

What is the role of complementary data? Increasing flows of data from automated and ITS sources are identified as essential aspects of update requirement reviews and decisions. These are many excellent reasons to time major updates in order to coincide with Census collections, and these intervals provide a basis reference timetable. In the US, the five-yearly Economic Census provides a powerful motive for five-yearly updates, as it also comes with funding. In Australia the decennial Census collections set the pace as the major relevant national collection by the ABS.

Organisations holding consolidated collections of transport data, as the TISB does, are well placed to develop a planned program of data review. This is to ensure that data relevance, availability, access and currency are maintained for use when needed. The description of the process is encapsulated in the formal process table shown in Table 5.

Table 5 Survey decision framework data specific steps

1.requirements	2.existing data	3. likely changes	4. proposed survey	5. data benefits

Step 1 DATA REQUIREMENTS

CASE STUDY:
Update of accessibility to a regional CBD

1.1 Key features required

- Accessibility to the facility by different socio-demographic groups in a range of locations
- Transport travel time and distance by mode
- Spatial locations and networks in greater detail close to the facility
- Decide if the modelling capacity of a destination choice model is required

1.2 Data needed

- Population segment knowledge by location
- Distance and travel times by travel mode by day of week/time of day
- Socio-demographic segment information in the catchments

1.3 Associated data

- Census data
- housing starts and
- development data can assist

1.4 Sensitivity of Output

- socio-demographic or environmental change may have occurred

1.5 Sensitivity of Outcome

Results that do not reflect the current or projected situation will lead to inappropriate investment in traffic and transport service and facilities provision. Where projections are needed, the use of 'old' data will amplify this risk.

1.1 Intended Application

- What are the key features required for this application?
- Is the data being used to measure or monitor performance?
- What are the key areas required for the data?
- What level of aggregation is appropriate for this application?
- How will it be analysed?

Applications have their own special characteristics and needs. If a new model is to be estimated then some aspects (such as destination choice behaviour) will often require some additional data collection.

1.2 Data Requirements Audit

- Which data elements are required?
- Locations and level of detail required?
- Timing?

Applications arise from an issue or location specific problem or concern. Each has different demands, some requiring time series and historical data, others a detailed coverage at one point in time.

1.3 Associated Data Required

- Other data sets needed for the application?
- Availability and Date/s?

Few applications find existing data available exactly matching their requirements. Traffic surveys may need traffic counts for area wide estimates, travel surveys may need ABS Census Journey to Work counts. Freight surveys may need Census of Production data. The availability of fresh association data of this sort supports the case for new application specific survey data.

1.4 Sensitivity of Application OUTPUT to data

- How do changes in input data change output results?
- What order of change in data is needed before changes are detectable in the output?

Where the application requires a fresh picture of a situation, unless the sample size is large enough to detect changes, updated survey work will be ineffective.

1.5 Sensitivity of Application OUTCOME

- How will changes in the application output affect policy planning or operation?
- What threshold of change needs to be detected, for what factors?

In some cases changes in average travel time of 5 minutes or more might be considered important, in others traffic level changes of 5% may be critical.

1.requirements	2.existing data	3. likely changes	4. proposed survey	5. data benefits

Step 2 EXISTING DATA

CASE STUDY:
Network Origin-Destination freight movements for a range of truck sizes

2.1 Existing Data?

- A freight vehicle O-D (origin-destination) surveys are not frequent.
- OD matrices might be synthesised from traffic counts but need a start point.
- Freight vehicle flows on specific roads are used reasonably frequently, so traffic count data is usually available.

2.2 Accessibility?

- Is data still held on a database? (if collected some years ago)
- Is classified count or weigh in motion data available?
- Is location coverage sufficient?

2.3 Content

- Did survey have coverage of vehicle types required?
- What areas did it cover?
- Who did the survey and how?
- Previous data might be a combination of classified counts, screen line counts, WIM data collected by a mix of methods

2.4 Applicability

- Tracking of existing automated counts provides a basis for assessing if updated information is needed.
- If major industrial development has occurred, some counts are needed in the road nearby

2.1 Existence of appropriate data

- Are relevant data sets available?
- Where and how are they held?
- When were they collected?
- What collection methods were used?
- Has it been used for other purposes and is therefore well known?

Data often appears to be available at first sight, until examined more closely when gaps in coverage, inadequate sample sizes, and problems with quality and documentation emerge. The oft-used phrase ‘it’s too urgent to collect fresh data’ must await such checks.

2.2 Accessibility

- Is the data available at the level of disaggregation required?
- Are there any intellectual property, privacy or organisational constraints?
- Is the data fully documented?
- Is it in a format that can be used for this application?

It is often assumed that if a dataset exists actually exists and is available, and that the interpretation of the variables will be straightforward. This is not always the case.

2.3 Content

- When was it collected using what method?
- Who has used it previously with what experience?
- Does the existing data set(s) include the key variables needed?
- Are the variables adequately described and specified?
- Does the data cover the appropriate location(s)?
- Is the spatial resolution adequate for the purpose?
- Is the accuracy of the data known and adequate?
- Is the segmentation of the population adequate for purpose?

It is not unknown for variables in a survey collection to have been omitted when coding the survey results: a detailed review of the content of datasets and surveys is an essential step.

2.4 Applicability

- Can the general levels of change since collection be estimated?
- Is the data held in a form that can be used?
- Was it collected in a form suitable for the present application?
- Is there evidence that key variables or locations are missing?

Given that the dataset is apparently suitable for the purpose of the application being considered, can it actually be accessed, and is it in a usable form.

1.requirements	2.existing data	3. likely changes	4. proposed survey	5. data benefits

Step 3 LIKELY CHANGES

CASE STUDY:

Network Origin-Destination freight movements for a range of truck sizes (continued)

3.1 What sorts of Change

- Freight is closely associated with economic activity and its proxy, urban employment.
- Might also be change in logistic practice

3.2 Checking change

- ABS including fine area level industry employment data from journey to work question
- Ask industry association about logistics changes
- Traffic counts also provide a key means of assessing the levels of change since the last survey or OD matrix estimation

3.3 Sensitivity

- Will model aggregate change?
- Can model accommodate new vehicle classes?
- What degree of shift in freight patterns is needed to trigger planning changes?

3.1 What sorts of change would be suspected

- Would land use have changed?
- Has the population changed/grown?
- Is there economic change? new industry?
- Are there infrastructure changes new roads, dedicated bus lanes?
- Are there new technologies such as B-doubles?
- Has there been significant traffic increase

Clearly the purpose of update is to measure change but there needs to be 'likely change' and that change must relevant to survey population and purpose.

For instance new dedicated bus lanes will impact transit commuting but not the composition of freight traffic. An aging population may affect pedestrian behaviour but not B-double uptake.

3.2 How can change be checked?

- Census data for demographic change and industry/employment data?
- Council land use data?
- Traffic counts and classified traffic counts for fleet numbers and composition

Large and obvious changes such as an urban development on previously semi rural land but if less obvious may check

- Has there been change in similar localities?

3.3 Sensitivity of Application to data changes

Returning to the question in step 1.3 and 1.4, what level of changes will affect OUPUTS and OUTCOME:

- What thresholds of change are needed to change analysis?
- What level of change will affect planning or policy?

For example: New update data, which finds a change of 5 minutes in travel time, will not be useful if the model has a10 minute threshold. Would a 10 minute change in travel time trigger planning or operation changes?

1.requirements	2.existing data	3. likely changes	4. proposed survey	5. data benefits

Step 4 PROPOSED SURVEY

CASE STUDY:

Update survey of bicycle helmet wearing by school children

4.1 What type of survey

- Observation survey
- Schools in a regional town
- Sample size all schools in town

4.2 Detecting change

- Sampling strategy is comprehensive
- But sensitivity also depends on estimated bicycle use
- If use is very low may have difficulty checking whether there was change from earlier survey

4.3 Indicative Costs

- This is a simple observational survey
- Costs per school could be based on earlier survey factored to today prices

4.1 What type of survey

- Proposed scope and scale?
- Proposed collection method?
- Sample size? / Likely response rate if relevant?
- Timing and duration?
- What accuracy is expected?
- Is there an option for larger sample/better accuracy?

This is a 'first pass' specification rather than the full specification needed for the business case.

4.2 Will the update detect change?

- Will the accuracy of the new data be sufficiently sensitive to detect the suspected level of change?
- Will it detect extra useful information?

The first question is predominantly related to sample size but also is dependent on survey method error and response bias.

Update if it can detect change or gain more information.

4.3 Indicative Costs

This is a 'first pass' estimate only based on options such as:

- Costs of similar types of survey carried out recently in Australia
- Costs of original survey for which update is planned factored to current prices
- Rough estimates of survey stage costs

Should include the full cost to the department commissioning survey

- Opportunity cost of planning and managing survey and dealing with data

1.requirements	2.existing data	3. likely changes	4. proposed survey	5. data benefits

Step 5 DATA BENEFITS OF UPDATE

CASE STUDY:
Update survey of bicycle helmet wearing in regional town (Continued)

5.1 Updated data
- Accuracy of old data not known
- New data will get extra information

5.2 Benefits to cost
- If for education campaign decision
- May be better to direct cost to campaign

5.1 Will updated data be better?

Is existing data fit for purpose
- Sufficiently accurate?
- Sufficient information?
- Is it likely to have changed?

Will new survey data
- Show change from old?
- Provide more accurate information?
- Provide extra information?

5.2 How do these benefits relate to cost?
- Will policy, planning or operation benefit from the update?
- Will benefit from the extra information?
- Does cost seem reasonable in terms of added value?

5 Conclusions

5.1 Suited to changing circumstance

Changes in data demand, data availability and data management are all increasing the importance of effective and efficient survey processes:

- Concerns about sustainability lead to a desire for more comprehensive data.
- A broad spectrum of users relies on transport data to feed applications in administration, policy, safety, planning, and operations. The FHWA manual (Cambridge Systematics, 1996) provides an extensive list of these users. Relevant applications locally are performance measurement and contract monitoring.
- Richer data is needed as budget constraints force planners to consider effective use of infrastructure.
- Data from intelligent transport systems and data or parameters transferred from other collections is available but must be adapted to use.
- Data information and sharing systems are developing rapidly.
- Funding for surveys often comes from the same budget as that used for data management, providing both incentive and opportunity to apply data management to ensure more efficient survey processes.
- There is increased recognition that while data is costly, ignorance can be even more costly when important decisions are to be made.

5.2 A timely response

It has become clear that there is no internationally accepted methodology to determine update frequencies in traffic and transport surveys, and that addressing this issue is breaking new ground. As there are business and policy risks in working without reliable and relevant information in transport and traffic areas, with large sums and community resources involved, actions to improve the quality of data are worthwhile. This becomes more important as transport sustainability becomes more urgent in an era of human activity induced climate change. Prioritisation of mitigation and adaptation measures needs to be based on complete information.

This paper has taken a needs-based approach to determining if and when updates are required. This is a rather different approach to the practice widely accepted across the world of undertaking updates if and when resources permit. The Framework provides a baseline for developing and applying a needs-based approach, and thereby prioritising and making better use of available data and associated sources. This is based on methodological standards as well as pragmatism. As Werner Brög wrote, 'Methodological standards are … necessary, and especially necessary when many do not heed them' (Brög, 1997).

6 Acknowledgements

This paper builds on work done in this area for the Queensland Department of Transport (Meyrick and Associates and Oxford Systematics, 2006), and we wish to acknowledge our gratitude for their permission to make use of this sound foundation.

References

Brög, W 1997, 'Raising the Standard! Transport Survey Quality and Innovation Keynote Paper', *Proceedings of an International Conference on Transport Survey Quality and Innovation*, Grainau, Germany, May, pp. 24-30.

Cambridge Systematics 1996, *Travel Survey Manual Travel Survey Manual,* prepared by Cambridge Systematics, Inc. for U.S. Department of Transportation, Federal Transit Administration, Federal Highway Administration, Office of the Secretary, U.S. Environmental Protection Agency, June.

Friman, M 2004, 'Implementing quality improvements in public transport*', Journal of Public Transportation*, 4(7).

Garham, B and Maxwell, P 1990, 'Rail Passenger Survey in Sydney', *14th Australasian Transport Research Forum*, Australia, September.

Garling, T and Friman, M 2001, 'Frequency of negative critical incidents and satisfaction with public transport services'*, Journal of Retailing and Customer Services*, 2(8), pp.105-114.

Greaves, S P 2001, 'Local Sample Updates for Synthetic Household Travel Survey Data', *Proceedings of the 23rd Conference of Australian Institutes of Transport Research*, Monash University, December.

Greaves, S P and Stopher, P R 2000, 'Creating a Synthetic Household Travel/Activity Survey- Rationale and Feasibility Analysis', *Transportation Research Record,* 1706, pp. 82-91.

Harrison, F D 2002, *Performance Measures and Targets for Transportation Asset Management,* NCHRP Report 551, Cambridge Systematics for FHWA.

Kunert, U, Kloas, J & Kuhfeld, H 2002, 'Design Characteristics on National Travel Surveys: An International Comparison for Ten Countries', *Papers of the 81st Annual Meeting of the Transportation Research Board*, January.

LATS 2001a, 'Consideration of a continuous survey process', *London Area Transport Study*, April.

LATS 2001b, 'Consideration of a continuous survey process- Technical Annex', (Neffendorf, Ashley, Bates, Gunn, Jones, Fearon, Ramsey, Wigan, McCaig), *London Area Transport Study,* May.

Margiotta, R 1998, *ITS as a Data Resource: Preliminary Requirements for a User Service*, Federal Highway Administration, Washington, D C, April.

Meyrick and Associates and Oxford Systematics (2006), *Framework for Establishing the Frequencies of Update for Transport Surveys: Final Report* Queensland Transport: Integrated Transport Planning: Transport Information Systems Branch, June

Richardson, A J 2003, 'The temporal variability of public transport usage', *Proceedings of the European Transport Conference*, Strasbourg, France, PTRC London.

Ruiz, T 2004, 'Attrition in transport panels: a survey', *Proceedings of the Seventh International Conference on Travel Survey Methods*, Costa Rica, August, viewed April 2006, <http://www.its.usyd.edu.au>.

Slavik, M M and Stopher, P R (1979), A Note on the Optimum Planning Interval, T*ransportation Planning and Technology*, 5(4), 227-245.

Southworth, F 2003, *A Preliminary Roadmap for the American Freight Data Program* (DRAFT) November 26, (Modified December 2004 for TRB), ORNL for Bureau of Transportation Statistics USDoT.

Stopher P, Bullock P & Rose J 2002, 'Simulating Household Travel Data in Australia an Adelaide Case Study', *Proceedings of the 25th Australasian Transport Research Forum*, Canberra, Australia, September.

Stopher, P R and Metcalf, H M A 1996, 'Methods for Household Travel Surveys', *Synthesis of Highway Practice*, no. 236, Transportation Research Board, Washington, DC.

Westlake, A and Wigan, M R 2007, Integrating Information about Complex Systems: the Role of Meta-Data in the Acceptability of Results from Models, *Transportation Research Record* 1972, 60-68

Wigan, M.R. and Southworth, F.A. (2005) What's wrong with freight models? *Proceedings: European Transport Conference*, Strasbourg, France. PTRC London.

Wilmot, C.G. and P.R. Stopher. 2001. Transferability of Transportation Planning Data. *Transportation Research Record* 1768:36–43