EUROPEAN COMMISSION
DG RESEARCH

SIXTH FRAMEWORK PROGRAMME

THEMATIC PRIORITY 1.6

SUSTAINABLE DEVELOPMENT, GLOBAL CHANGE & ECOSYSTEMS

INTEGRATED PROJECT – CONTRACT N. TIP5-CT-2006-031315

CityMobil

Towards advanced transport for the urban environment

Evaluation Framework

Deliverable no.	5.1.1
Dissemination level	Public
Work Package	5.1
Author(s)	Dr. Greg Marsden
Co-author(s)	Dr Astrid Gühnemann, Dr Adriano Alessandrini, Daniele Stam, Giovanni Ruberti
Status (F: final, D: draft)	F.14.02.07
File Name	D5.1.1-Public Evaluation Plan-FINAL.doc
Project Start Date and Duration	01 May 2006 - 30 April 2011

SIXTH FRAMEWORK PROGRAMME

Executive Summary

The objective of the CityMobil project is to achieve a more effective organisation of urban transport, resulting in a more rational use of motorised traffic with less congestion and pollution, safer driving, a higher quality of living and an enhanced integration with spatial development. This objective is brought closer by developing integrated traffic solutions: advanced concepts for innovative autonomous and automated road vehicles for passengers and goods, embedded in an advanced spatial setting. The CityMobil project therefore requires an evaluation framework capable of capturing the social, environmental, economic, legal and technological impacts of Advanced Transportation Systems. The framework is required to operate at spatial scales ranging from laboratory and test-track interventions, through computer modelling to real-world implementations on a large scale. It is an aim of the evaluation framework is to facilitate mainstreaming of the assessment of ATS within local and national evaluation processes.

A review of the state-of-art in evaluation frameworks for ATS systems and, more broadly, for integrated transport and land use strategies demonstrates the need to establish the following key evaluation elements:

- Objectives – a series of key evaluation categories relating to the achievement of sustainability goals and capable of capturing the practical implementation of new technologies;
- Framework – a specific interpretation of the evaluation categories developed from the objectives which sets out the key impacts that relate to each category and the more specific indicators that can be used to assess achievement;
- Assessment methods – a clear understanding of how the framework should be applied (for example using social cost-benefit, MCA or mixed approaches) and to which parts of the CityMobil project
- Implementation – the establishment of clear procedures for the ex-ante and ex-post evaluation of the different parts of the project.

The objectives which form the outline of the evaluation framework for CityMobil are:

- Acceptance
- Quality of service
- Transport patterns
- Social Impacts
- Environment
- Financial Impacts
- Economic
- Legal impacts
- Technological success

Indicators were selected to allow monitoring against each of these key objectives. For the evaluation of passenger transport systems a list of 64 indicators was generated. This can be viewed as the complete envelope of overarching indicators which could form part of the evaluation of each of the different elements of CityMobil.

We do not anticipate that it will be appropriate or necessary for all of these indicators to be measured at all sites nor that it will be feasible for the modelling tools to necessarily produce all of the indicators. A sub-set list of core indicators that each demonstration site and the city-wide simulations should report on has therefore been generated as shown below.

In addition to the generation of a framework for the evaluation of passenger transport systems a framework for freight systems has also been generated. This will be able to guide the freight related elements of the project although no freight demonstration is yet planned.

Evaluation category	Indicator
Acceptance	User willingness to pay
Quality of service	Perceived level of safety
	Perceived performance
Transport Patterns	Mode Share
	Delay per passenger trip
	Journey time variability
Social Impacts	Access times for mobility impaired users
	Accident levels
Environment	Energy Use
	Toxic Emissions[1]
	Total CO_2
	Total Land-Use Change
Economic	Net Present Value
	Internal Rate of Return

This document also seeks to provide a practical guide to how the framework should be applied at each of the different spatial scales to ensure that the evaluation provides information which is both targeted at the different users (technology developers through to politicians) but closely linked. This is summarised in the following table.

Application	Method
SP1 - Demonstration sites	Full social Cost-Benefit Analysis + tables + financial assessment; Local cost values
SP2 - Scenario modelling	Multi-Criteria Analysis for land-use transport interaction models Cost-Benefit Analysis + tables
SP3 & SP4 – Technology Developments	Quantitative Measures, including indications of costs and impacts of technologies

[1] The most relevant toxic emissions for the local circumstances should be selected. It is expected that NO_x and PM_{10} will be the most common problem emissions.

Advice is provided on the application of the framework, the definition of indicators and the flow of data within the project. This document forms the start point for all of the project workpackages to reflect on and then specify how their outputs connect to the overall project aims. The ex-ante evaluation is the next stage in this process where the information flows highlighted in this report have to be realised. This document has been prepared with the input of members of a project wide evaluation working group. The outcomes therefore represent a project-wide commitment to adhering to the principles and processes identified within.

TABLE OF CONTENTS

TABLES

FIGURES

1 Introduction

1.1 Background

The objective of the CityMobil project is to achieve a more effective organisation of urban transport, resulting in a more rational use of motorised traffic with less congestion and pollution, safer driving, a higher quality of living and an enhanced integration with spatial development. This objective is brought closer by developing integrated traffic solutions: advanced concepts for innovative autonomous and automated road vehicles for passengers and goods, embedded in an advanced spatial setting. The city of tomorrow is in need of integrated traffic solutions that provide the required mobility in an efficient, safe and economic manner. It is inevitable that automation, in all possible forms between providing information at one end of the spectrum and fully autonomous driving at the other, will play a major role. We wish to make significant steps forward that will, on the short to medium term, support a sustainable development of European cities.

The long term solution may be found in a more balanced split between high-capacity transport modes like buses; trains and metros, which have the advantage of space and energy efficiency and more individual transport modes, offered by innovative vehicles with the advantage of flexibility and availability. A good mix of these two modes will enable cities to offer improved mobility to all citizens while improving the quality of life and the sustainability of the environment.

The CityMobil project will build on the results of recent European and national projects and will validate and demonstrate the capabilities of new mobility solutions in different European cities. In 5 horizontal sub-projects the issues that still prevent full scale implementation of innovative transport systems will be investigated and solutions will be developed. At three sites: Heathrow, Castellón and Rome, large scale demonstrators will be set up to supply proof of concept of innovative transport systems integrated in the urban environment.

1.2 Project Objectives

The global objective of CityMobil, described above will be achieved by:

- *Developing advanced concepts for advanced road vehicles for passengers and goods*

  The introduction of the various concepts will be largely based on work done before in other European projects (e.g. CYBERMOVES, CYBERCARS, EDICT, STARDUST). Most of these projects addressed isolated aspects of the mobility problems of cities, where CityMobil focuses on the overall urban transportation problem. CityMobil will integrate the results of many of these projects in its deliverables and will bring automated transport one step forward by solving problems that were not addressed before and by identifying and solving problems that are directly related to the integration of innovative transport systems in an existing environment.
- *The introduction of new tools for managing urban transport* CityMobil will develop tools that can help cities to cross the thresholds that are preventing them from introducing innovative systems. For instance the absence of certification procedures makes it difficult for city authorities and operators to judge whether or

not a system is safe, so certification procedures will be one of the key deliverables. Another threshold is the question as to whether or not a system can be justified and operated profitably. In order to improve the capabilities of cities and operators to judge the profitability tools for assessing business cases will be developed. Traffic circulation and mode assignment models will be used as tools to predict the improvements that innovative systems will supply for passengers and goods. All of these individual tools will have their merits but the major innovation of CityMobil is the integration of all of these tools, methods and concepts in the city infrastructure.

- *Taking away barriers that are in the way of large scale introduction of automated systems*

  A number of barriers that are in the way of large scale introduction of automated systems must be identified and taken away. Some of these barriers are of a technological nature, for instance the lack of adequate systems for obstacle detection and avoidance. Some are of a legal or administrative nature, for instance the legal requirement for vehicles using public roads that the driver is responsible for the vehicle at all times, which effectively prohibits driverless vehicles from using public roads. There are also barriers on the operational field, where present day traffic management systems are not geared towards dealing with automated vehicles. One of the main barriers is the present lay-out of European cities. Most cities are exclusively laid out for automobiles with a smaller or bigger contribution for public transport, bicycles and pedestrians. The CityMobil consortium will provide an overview of all of these barriers and will develop tools and strategies to remove them.

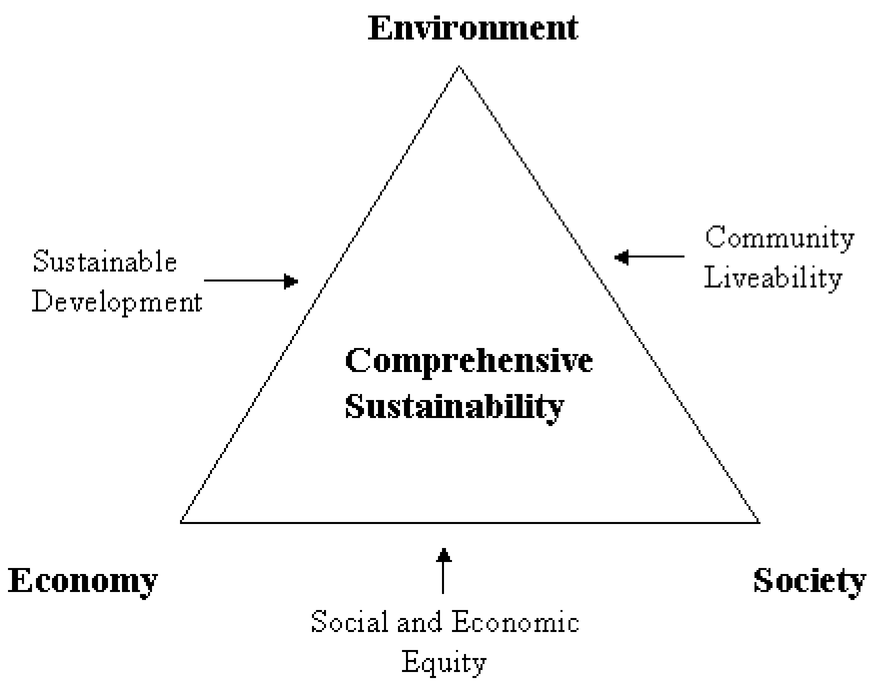
- *Demonstrating the concepts, methods and tools developed in CityMobil in a number of different European cities under different circumstances through a series of demonstrations*

  Three of the demonstrations (Heathrow, Rome and Castellón) will be real implementations of innovative new concepts. In practice these will be the first stages of automated transport systems that are really integrated in an urban environment. In a number of other cities studies will be carried out to show that an automated transport system is not only feasible, but will also contribute to a sustainable solution for the city's mobility problems, now and in the future.

1.3 Definition of sustainability

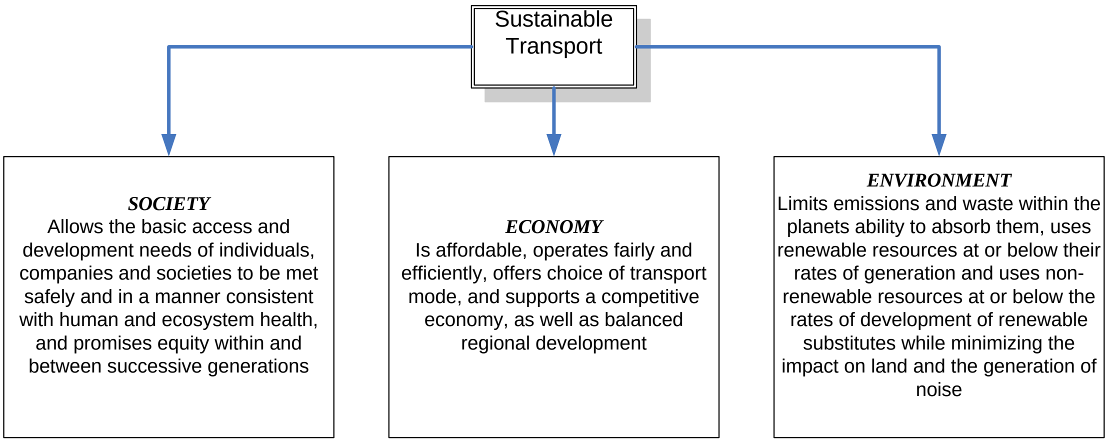
Sustainability or Sustainable development has been commonly defined as “Economic and social development that meets the needs of the current generation without undermining the ability of future generations to meet their own needs" (WCED, 1987). This definition brought together what is now known as the three pillars of sustainable development; economic development, social development and ecological development under one societal goal of sustainability. The TRANSLAND project conceptualised the tensions and interactions between the pillars as shown in Figure 1.

Figure 1: Comprehensive Sustainability - Source: TRANSLAND (2000)



Whilst many definitions of sustainable development and sustainable transport exist, the Council of the European Union (2001) has adopted a working definition that guides the development of much of the evaluation framework presented in this document (Figure 2).

Figure 2: Sustainable Transport Definition (Source: Council of the European Union 2001)



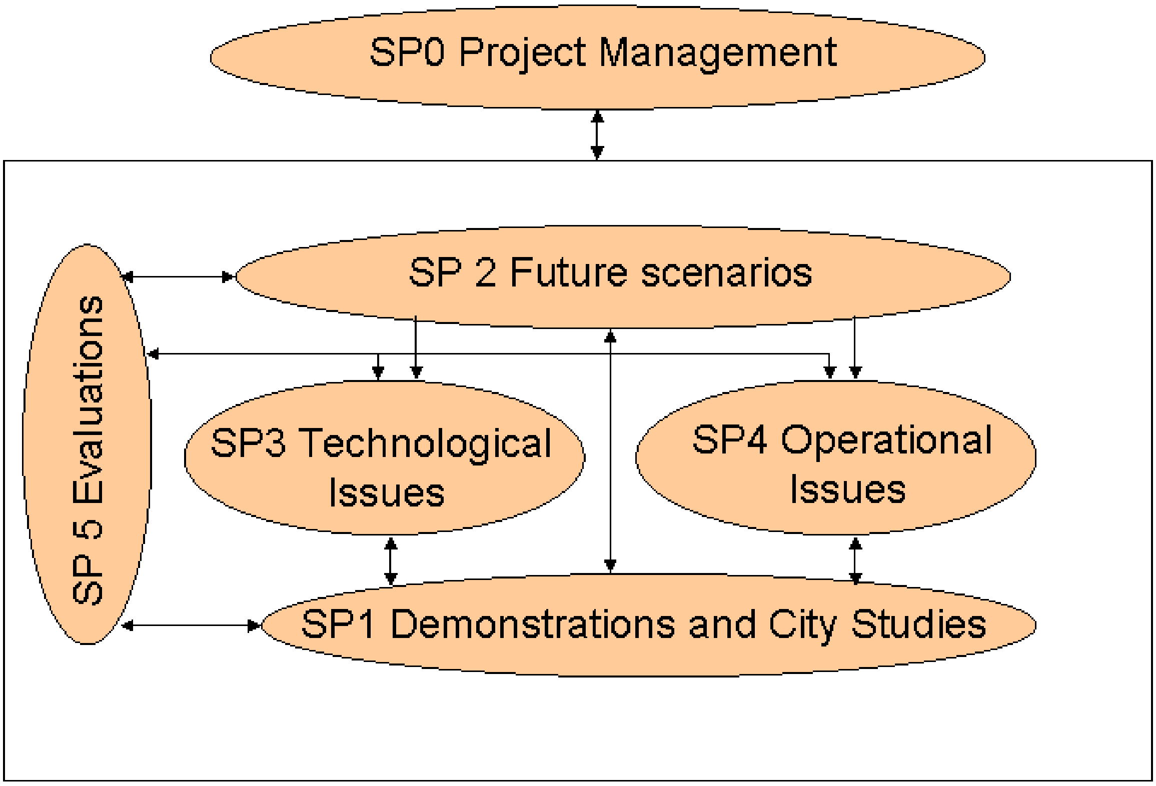
1.4 Project Structure

CityMobil is organised in six workpackages. SP0 oversees the project management and dissemination function. SP1 covers the implementation and monitoring of the three demonstration sites and the series of smaller city studies. There are three ‘technical’ workpackages (SP2, SP3, SP4) each with a specific function to further the capabilities and understanding of the impacts of ATS systems. SP5 is the evaluation workpackage which provides a framework around which the flows of information for the project are linked.

Figure 3 shows the basic relationships between the Sub-projects and the demonstrations. The future scenarios study SP 2 gives the direction in which urban transport is developing which will be input for SPs 3 and 4 on Technological issues and Operational issues respectively. SP 5 Evaluations will provide standards and indicators to SP 2. The results of SP 2 will also be useful for the City studies and the demonstrations. The evaluative framework that is being developed in SP 5 will supply performance indicators to allow the work in SPs 2, 3, 4 and 5 to be assessed for financial, operational and sustainability outcomes. The results of SPs 2, 3 and 4 will be evaluated in SP 5.

The demonstrations and city studies will generate a number of requirements and research issues to be addressed in SPs 2, 3 and 4 and these SPs will provide answers, solutions and options to the demonstrations and city studies.

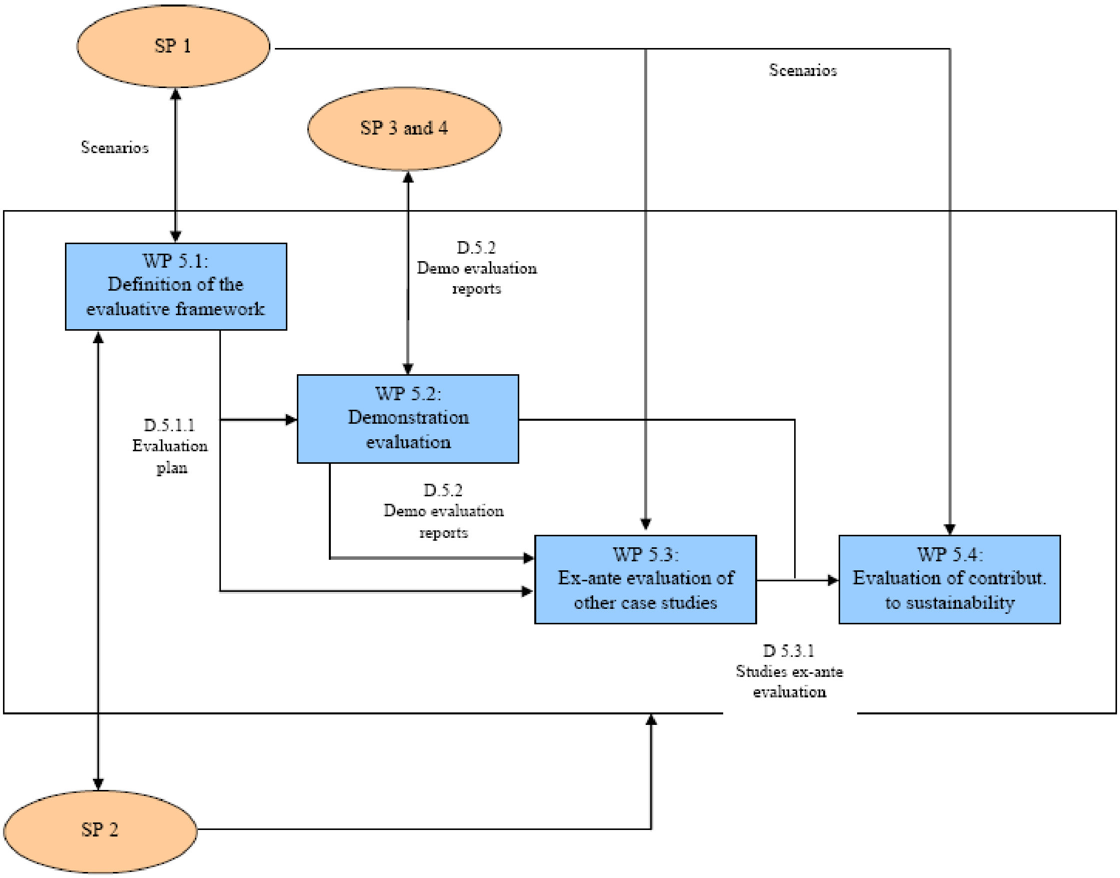
Figure 3: Overall CityMobil Project Structure



Central to the success of the evaluation of CityMobil is the effective understanding, communication and monitoring of the required flows of information between the different sub-projects. In setting out a framework for evaluation, this document also indicates the nature of the flows between the different SPs and the Evaluation task. This understanding has been developed in conjunction with the demonstration sites and the SP leaders through an Evaluation Working Group which has been established for the CityMobil project (see Annex 1 for membership).

Concerning the SP5, in Figure 4 the sub-project structure and its links with the other sub-projects are reported.

Figure 4 Sub-project 5 structure and links with other sub-projects



As shown in Figure 4, the major part of the works, concerning the SP5 are done inside Work package 5.2 and Work package 5.3.

Work package 5.2, demonstration evaluation, is directly linked with SP3 and SP4 through the deliverable 5.2 about demo evaluation reports (which is also the link between WP5.2 and WP 5.3). It is also linked with WP5.1 and consequently with SP1 and SP2 through this deliverable 5.1.1, and it is linked with WP5.4 through the deliverable 5.3.1 concerning the evaluation of ex-ante studies.

Work package 5.3 is directly linked with SP1 and WP5.4, is linked with SP3 and SP4 through WP5.2 and with SP1 through WP5.1.

1.5 Purpose of this deliverable

This deliverable is the first in the evaluation workpackage and its goal is to establish the methodology for all the following evaluation activities. It provides a common platform for the integration of the trials and horizontal work packages to enable a comprehensive assessment of the contribution of advanced road transport systems to sustainability. The nature of CityMobil – an integrated project requiring evaluation of long-term strategic impacts, through practical monitoring of implementation to laboratory simulation of new technologies – requires a new approach to integrating existing evaluation methodologies. This deliverable explains how the evaluation framework was derived and how it should be applied to enable the project to be able to:

- Quantify and qualify the benefits of advanced road transport systems
- Monitor the progress of the demonstrations and provide feedback
- Generalise the evaluation results of trials and studies and transfer them to other case studies
- Identify how advanced road transport systems can contribute to sustainability

In developing the framework, we placed great emphasis both on the role of the framework in promoting effective and transparent internal communications of project findings but also that the project has to report its findings to a wide range of audiences:

- The European Commission
- Decision-makers – what are the impacts and costs of such systems and how do they differ from existing systems?
- End-users – what are their reactions to the system?
- Operators – how does the system work in practice?
- Non-users – what impacts does it have on other people and vehicles?
- Technology developers and producers – what opportunities exist for development?

1.6 Structure of Deliverable

This deliverable is divided into three main sections, of which the introduction is the first. Section two reviews the state of art in evaluation in the areas of new technology assessment, demonstration assessment and long-term sustainability assessment. This provides the basis for the decisions taken on the approach to the evaluation framework. The framework is presented in Section three including the objectives, evaluation categories and more detailed advice about the application of the framework and the necessary flows of information. The deliverable is deliberately presented as a concise document. Further details about the composition of the evaluation working group, the definitions of indicators and sources of data to assist in the evaluation task are all included as separate annexes.

2 Evaluation State-of-Art

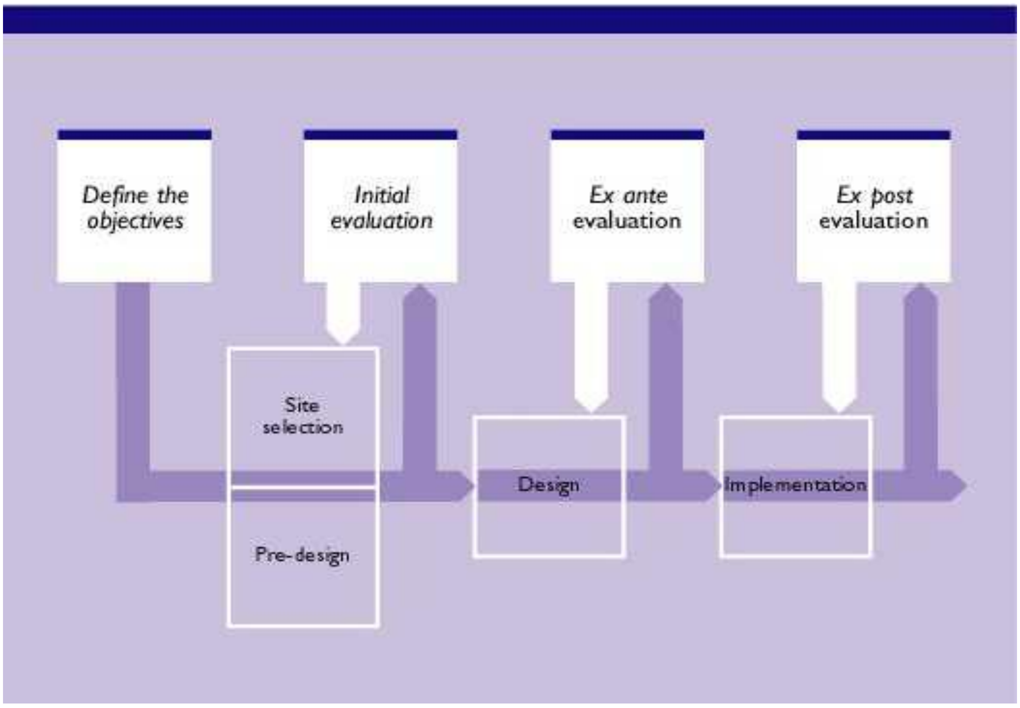
This chapter reviews the state-of-art in evaluation methods. The review builds on the existing well established and accepted methods for evaluating new technologies (CONVERGE) and demonstration projects (MAESTRO) and examines recent leading evaluation advancements in projects such as PROPOLIS and SPECTRUM, reflecting the need to further integrate advances in Automated Transport Systems with larger-scale land-use & transport evaluation techniques and those employed by governments across Europe.

2.1 Goals of Evaluation

"Policy analysis or evaluation, in general, aims at rationalizing planning and decision-problems by systematically structuring all relevant aspects of policy choices."(Vreeker and Nijkamp, 2005, p509). This information can include economic (e.g. costs, job creation), social (e.g. safety, distributional effects) and environmental (e.g. toxic pollution, climate change, land take) impacts.

Evaluation occurs at many different stages of the decision-making process as shown in Figure 5.

Figure 5: MAESTRO methodology flow (source: MAESTRO Guidelines)



Even when the range of different options to be evaluated has been narrowed down, there remains a series of different approaches that could be applied to conduct the analysis:

- "Checklist approaches;
- Cost Benefit studies;
- Planning balance sheet techniques;

- Goals achievement methods;
- Multi-criteria analysis; and
- Multiple objective programming" (Ibid., p509)

An evaluation framework might therefore be used to examine questions ranging from "what are the long-term sustainability implications of two proposed packages of policies?" to "is this technology safe and, if so, in what circumstances?". The scale of the issue (e.g. vehicle or city) and the point in the evaluation process (e.g. pre (ex ante) or post (ex post) implementation) will help to determine the information available and relevant to informing the decision.

The role of the evaluation framework within CityMobil is to establish one overarching framework which will allow the different scales of analysis to connect together in a meaningful manner such that the contribution of different elements of the project to answering the key aims of the project is clear.

2.2 Previous ATS evaluation approaches

The state-of-art in advanced personal mobility system development, testing and assessment was recently brought together through the NETMOBIL cluster (www.netmobil.org). NETMOBIL covered four different types of EU research project ranging from technology development and testing to hypothetical city scenario testing. The different projects and their evaluation approaches are shown below in

Table 1: Summary of evaluation approaches from previous ATS research

Project	City Modelling	Demonstration	Technology Development	Evaluation Approach
EDICT	✓	✓	✓	MAESTRO
CYBERCARS		✓	✓	CONVERGE
CYBERMOVE	✓	✓		MAESTRO
STARDUST	✓			CONVERGE/bespoke

The principal approaches to evaluation have previously been based on the MAESTRO and CONVERGE guidelines and these are therefore reviewed below.

2.2.1 MAESTRO

MAESTRO was a Fourth Framework project conducted for DGVII. The following description summarises the key principles of the MAESTRO approach and draws heavily on material taken directly from the MAESTRO guidelines (MAESTRO, 2000).

The main purpose of the MAESTRO Guidelines is to aid decision-making for the selection, design and evaluation of pilot and demonstration (P/D) projects for transport in Europe. The Guidelines are intended to connect the different decision moments and evaluation phases throughout the entire lifecycle of a P/D project.

MAESTRO proposes three main evaluation phases (see Figure 5):

- Before the project begins, when users define their specific transport problem and decide whether a P/D project is the best way to try to solve the problem;
- During the project (the methodology), when users address the issues associated with setting up the project, specifically defining the objectives, site selection, pre-design and initial evaluation, as well as considering the design and ex ante evaluation and concluding with implementation of the P/D and ex post evaluation;
- After the project, when users consider how best to use the project results, and whether to proceed to full-scale implementation.

The first stage of the methodology involves the definition of objectives. General transport and sector-specific objectives are first identified to set the context for the definition of the project objectives. The definition of the project objectives is vital: all other parts of the methodology depend on the accurate definition of the project objectives.

The site selection and pre-design are the two components of the same project stage and MAESTRO recommends running them simultaneously with information exchange between them. Within the site-selection process, important selection parameters are highlighted, together with issues to consider to ensure the process is transparent and that as far as possible, conflict is minimised. Pre-design is the process of specifying the functionality of the applications or systems to be demonstrated, based on the project objectives, the user needs and requirements (identified within this stage) and the site characteristics.

After the site selection and pre-design, the first of the three evaluation phases may begin. This initial evaluation is based on expected impacts of the project and is mainly qualitative. The identification of expected impacts, and indicators to describe the impacts, is addressed within this section. MAESTRO divides impacts into four categories: (1) transport system performance; (2) economic efficiency; (3) environmental; and (4) safety and security. During the initial evaluation, the choice of evaluation method for assessing the impacts in all three evaluation phases should be made. MAESTRO proposes four alternatives, to be used alone or in combination: cost-benefit analysis; cost-effectiveness analysis; multicriteria analysis methods; and goal achievement matrix.

The design stage will follow a successful initial evaluation. In this stage, the pre-design of the demonstration is refined and hardware, software and support items for the demonstration are defined. The design must be consistent with the project objectives and the user needs/requirements and budget and time constraints should be taken into consideration as important issues for the development of the design. Once the detailed design of the project is in place, an ex ante evaluation should be carried out. This evaluation reviews the results of the initial evaluation following refinement of the pre-design. The detailed design of the P/D project is compared with the 'do-nothing' scenario. This evaluation will enable more accurate estimation of the detailed impacts of the project, as more data will be available by this time. The question of whether implementation of the P/D will meet the project objectives should be considered. Where unfavourable impacts are estimated by this evaluation, the design may be refined in an attempt to reduce these impacts.

Implementation of the P/D project will commence only when those involved in the project are content with the design of the demonstration and happy that the estimated impacts are promising. The demonstration should be implemented as designed.

The ex post evaluation phase completes the three evaluation phases. The success of the P/D project is evaluated in terms of how far it achieved its objectives and the actual impacts measured in the four impact categories. The outputs from this phase will feed into the evaluation method chosen during the initial phase. The outcome of this final phase will influence the decision whether or not to proceed to full-scale implementation.

2.2.2 CONVERGE

CONVERGE is a common methodology established to validate and demonstrate transport telematics applications from the DGXIII fourth framework programme of research. The guidelines were subsequently updated in 1998 (Zhang et al., 1998) from which this section draws heavily.

The methodology is based around a seven stage process:

- Determination of user needs;
- Classification and description of telematics applications;
- Formulation of assessment objectives;
    - Technical (system performance, reliability)
    - Impact assessment (safety, environment etc.)
    - User acceptance (users' opinions, preferences, willingness to pay)
    - Socio-economic evaluation (benefits and costs)
    - Market assessment (demand and supply)
    - Financial assessment (initial and running costs, rate of return, etc.)
- Pre-assessment of expected impacts;
- Choice of assessment methods;

  Assessments methods - includes the following elements: indicators, reference case, data collection, conditions of measurement, statistical considerations, measurement plan and the integrity of measurements. CONVERGE provides a general procedure for developing the assessment methodology, suitable for the majority of transport telematics projects, whether R&D, verification or demonstration. To ensure a consistent approach, CONVERGE recommends that all the above steps of the assessment methodology are summarised.
- Methods available for data analysis; and
- Reporting of results.

2.2.3 Evaluation techniques used by demonstration sites

The three large scale city demonstrations (Rome, Castellon and Heathrow) provided an initial evaluation, each one with a different technique.

Even if a standard framework for the evaluation of the demonstration sites has not yet been built, in the three different techniques used two common points can be however observed, both linked to the feasibility of the systems designed.

The first regards the foreseen demand: in each of the three demonstrations the demand has been studied in order to provide a realistic estimation of the people going to use the systems. On the basis of such demands and the infrastructures needed to satisfy them, each system has been dimensioned.

The second common point, directly linked with the first one, concerns the economic considerations made in the systems studies according to the funds allocated to finance the project. The dimensioning of the three systems have been done both considering the foreseen demands and the funds available to make systems capable to satisfy them.

Concerning the single demonstration evaluation activities and techniques, in Rome a system dimensioning and a financial evaluation based on the financial cost-benefit analysis of the designed system have been done. With such analyses, the system has been dimensioned in order to satisfy the foreseen demand and it has furthermore been proved to be financially good. On the basis of the dimensioned system, the Rome Municipality is calling for tenders for the system management (including the car-park inside which the system will be operating). Once the call for the tenders has been awarded, the works will start.

Similarly to Rome technique, in Castellon a socio-economic evaluation of the designed system has been made, in order to provide a strong indicator of the viability of the demonstration.

The evaluation technique adopted in Heathrow was different from that of Rome and Castellon. The new terminal T5 of the airport is far from its car-park and needs a connection to it. This PRT system is conceived to be such connection and a first step, a pilot project, aiming at analysing in depth such automatic transport systems for future wider application in the same airport and in other BAA airports. The evaluation done consisted in selecting the most appropriate system in a call for tenders. The initial economic evaluation has been postponed and will be done in the course of CityMobil project.

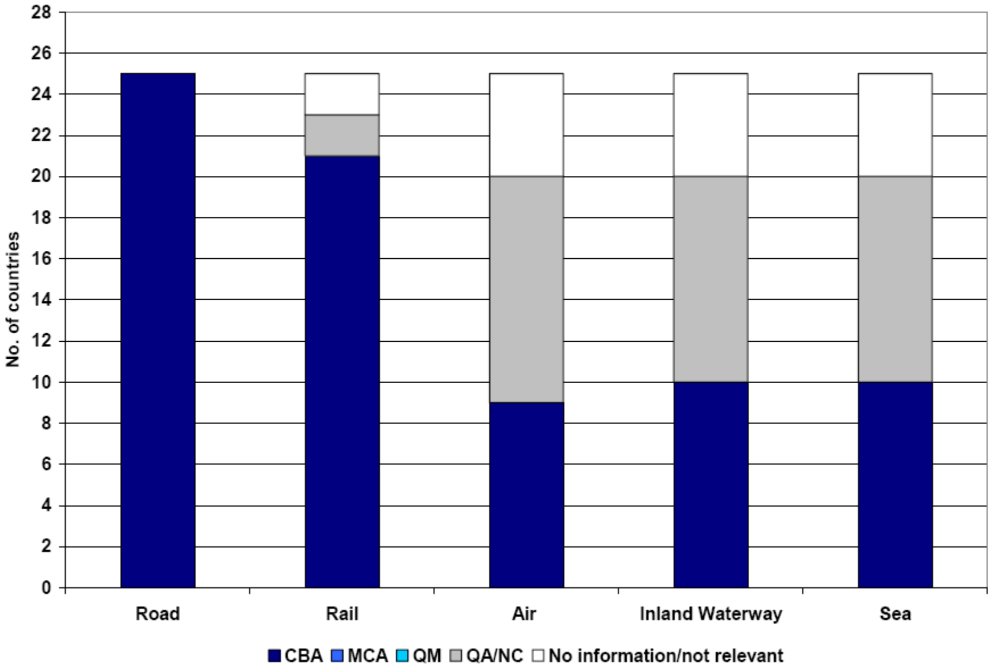
2.3 Advances in Evaluation Methodologies

Both MAESTRO and CONVERGE are now over five years old and there has been some progress made in the field of evaluation within the EU over this period of time. This section will review the state-of-art in evaluation applied across the EU.

Overall, a distinction can be made whether evaluation methodologies are applied in the appraisal of investment projects or in the assessment of new technologies and policies, mainly in research projects. The appraisal of investment projects generally has to follow more formalised procedures since financial and planning decisions are based on the results. In this respect, CityMobil needs to draw from experiences beyond those found in the more demonstration and technology oriented guidance from CONVERGE and MAESTRO. Therefore, in the following we present an overview of recent applications of evaluation methodologies in research and transport planning. Comprehensive introductions into evaluation methodologies can e.g. be found in the PROPOLIS and PROSPECT project reports (Lautso et al., 2004, Minken et al., 2002).

With the aim to develop a harmonized framework for transport project appraisal in Europe, the HEATCO project analyzed the current practice in the European Union (Odgaard et al., 2005) and the state-of-the-art in project appraisal (Bickel et al., 2005a) with a special focus on cost-benefit analysis. The methods applied in Europe are being classified as cost-benefit-analysis (CBA), multi-criteria analysis (MCA), quantitative measurements without weighting of indicators (QM) and qualitative measurement or not covered in a formalized method (QM/NC). The highest degree of formalisation of appraisal methods can be found for road and rail projects. It shows that in particular in road and rail project appraisal, cost-benefit analysis is the dominant type of method applied. However, several countries such as the UK and Germany apply a mix of methods which is not reflected here.

Figure 6: Types of analysis by mode (No. of countries using relevant type of analysis by mode) (Odgaard et al., 2005, p. 14)[2]



Based on the review, HEATCO recommends the use of full social CBA analysis as preferable method for transport investment appraisal taking into account value of time and congestion, value of changes in accident risks, environmental costs as well as costs and indirect impacts of infrastructure investment (Tavasszy et al., 2005). In order to cope with distributional issues, additional information can be provided in a "winners and losers" table. Of importance for CityMobil is also the recommendation to use local (country specific) cost values in order to reflect local preferences and willingness to pay. HEATCO (Tavasszy et al., 2005) as well as the ExternE (Bickel, Friedrich 2005) and UNITE project (Nash, 2003 and Bickel et al., 2006) provide a basis for a harmonized framework for deriving cost values for externalities, which are necessary for the full social CBA. The evaluation period and the choice of discount rates vary strongly across Europe; the median values are around 30 years and 5-6%. An overview of cost values, discount rates and evaluation periods applied in Europe can be found in the annex.

Since the focus of CBA is the assessment of the single criterion of economic efficiency as a means of maximising welfare and due to limitations in the valuation of impacts, MCA methodologies are often used exclusively or in addition to CBA in the long-term assessment of the sustainability of new technologies and transport policies. The MAESTRO guidelines also describe non-monetary discrete evaluation methods as an alternative method when a majority of indicators is deemed inadequate to be transferred to monetary values. They distinguish between MCA methods which apply weighting to indicators and survey table methods without weighting such as goal achievement matrices (GAM).

[2] The ranking of types of appraisal is as follows: CBA - MCA - QM - QA/NC - No information/ not relevant, i.e. if for example both CBA and MCA is used the figure reflects CBA.

In recent applications of evaluation applications to technology demonstration and policy assessment projects, CBA and MCA are often used in combination. CyberMove adopted more of a CBA approach whilst the EDICT project used a mix of CBA and MCA (without deriving weightings). Another example for a mixed approach is the PROPOLIS project (Lautso et al., 2004) within which two separate assessments have been carried out, an economic evaluation and an assessment of sustainability. The economic evaluation comprised of a full cost benefit analysis including transport indicators as well as indicators describing regional distribution. In the assessment of sustainability, a multi-criteria approach has been applied. Indicators are valued according to value functions and combined by using a weighting technique. The preferred technique of weighting in PROPOLIS is an outranking approach called Analytic Hierarchy Process (AHP), published by Saaty in 1977. Additionally, a software tool has been developed which allows users to apply their preferred weighting technique and user specific / local weights. AHP has also been used by Tudela et al. (2006) demonstrating that it was better capable of reflecting public acceptance in the decision making process for urban transport infrastructure than CBA.

A recent example for the application of a discrete utility-based multi-criteria method is the STEPS project (Lopez, Monzón, 2006), which evaluated energy supply scenarios for the transport sector according to performance criteria. For each scenario a performance score for four main criteria was calculated from their impact indicators. These scores are not aggregated further in order to allow decision makers to apply their specific preferences in the choice of scenarios. To dispense with the weighted aggregation of evaluation criteria is common practice in several applications of MCA in European transport planning, for example the Swiss NISTRA approach (Walter et al., 2003).

Other common state-of-the-art outranking techniques are PROMETHEE and ELECTRE which are capable of dealing with indifferences in preferences as well as uncertainties and non-compensation. Based on the results of the PROGR€SS project, Ukkusuri et al. (2004), for example, evaluate different congestion pricing methods using the multi-criteria decision-making algorithm ELECTRE IV which is capable of dealing with uncertainties and lack of weights (see for example the proposed application in the REORIENT project described by Hedel, Vance, 2005). However, there usually is a trade-off between the capabilities of methods in dealing with issues like uncertainty of results or fuzziness in decision makers' preferences and the complexity of methods. A crucial issue for the application of multi-criteria methods is the determination of the value functions for indicators and the weights. These can be obtained by means of a survey of decision makers as e.g. in the STEPS project. Finally, a sensitivity analysis needs to be carried out in order to determine the stability of results.

In the next section, we draw together the findings of the review of evaluation methods and propose an approach to evaluation that allows a common framework to be applied across the whole project.

2.4 CityMobil evaluation approaches

The review of evaluation frameworks demonstrates the need to establish the following key evaluation elements:

- Objectives – a series of key evaluation categories relating to the achievement of sustainability goals and capable of capturing the practical implementation of new technologies;

- Framework – a specific interpretation of the evaluation categories developed from the objectives which sets out the key impacts that relate to each category and the more specific indicators that can be used to assess achievement;
- Assessment methods – a clear understanding of how the framework should be applied (for example using social cost-benefit, MCA or mixed approaches) and to which parts of the CityMobil project
- Implementation – the establishment of clear procedures for the ex-ante and ex-post evaluation of the different parts of the project.

Based on the review of developments in evaluation methodologies it is proposed to follow a mixed approach in CityMobil, combining the strengths of CBA and MCA. Besides the validity and robustness of the methods, the selected evaluation methodology has to fulfil the requirements of transparency for decision makers and be sensitive to data availability and model capabilities. In particular, the latter will be differing between demonstration sites (SP1), modelling of city-wide scenarios (SP2) and technical developments (SP3 and SP4). The approach is summarised in Table 2 and explained further below. Further details about the application of the methods follows in Section 3.3 after the objectives and indicators are described.

There are some issues which need to be addressed across the different approaches: firstly, the spatial boundaries of the system need to be defined. This should not only include the corridor of the ATS application itself but at least also surrounding origination and destination zones. Secondly, all applications need to be compared to a reference case (without the ATS application). This should be a do-minimum approach, i.e. a trend development without investments into new transport system infrastructure.

Table 2: Evaluation approaches for CityMobil

Application	Method
SP1 - Demonstration sites	Full social CBA + tables + financial assessment; Local cost values
SP2 - Scenario modelling	MCA for land-use transport interaction models CBA + tables
SP3 & SP4 – Technology Developments	QM, including indications of costs and impacts of technologies

Since the demonstrations (SP1) can deliver quite robust data on the costs and benefits of the projects (see chapter 2.2.3) and economic efficiency plays an important role in their assessment of future feasibility, it is suggested to use a methodology based on the HEATCO recommendations, comprising of a full social CBA plus additional non-weighted tables to deal with issues of public acceptance and distribution / equity. Additionally, a financial assessment should be carried out. Details on the methodology can be found in chapter 3.3.1.

The future scenario modelling in SP2 should pursue a mixed methods approach related to the types of model available. An MCA approach should be pursued for strategic land-use transport interaction models. Such an approach overcomes any potential difficulties in incorporating the economic impacts of larger-scale land-use responses within the assessment. Since local preferences and objectives for the implementation of ATS might differ, in the first step, weightings will be defined locally by a survey of participants from the relevant reference group city in the ex-ante evaluation stage. In a second step, average values will be applied to achieve comparability across the scenarios including a sensitivity analysis on the range of values. In the optimal case, it will be possible to derive a typology of

ATS applications and corresponding value functions and weights (depending for example on city characteristics, income levels, existing alternatives, type of operator etc).

Where models are applied with no endogenous land-use modifications we recommend an approach based on Cost-Benefit Analysis with additional non-weighted tables of indicators.

In the case of the technology developments in SP3 and SP4, data will only be available on specific subsets of the evaluation criteria (e.g. workload and safety or integration of operations). A less formalised Quantitative Measure approach is more appropriate. However, indications of costs and impacts of technologies need to be delivered and where this is available a cost-effectiveness analysis should be undertaken.

This section has reviewed past approaches to evaluation of ATS systems and the current state-of-art in harmonising European evaluation techniques. Some requirements for the evaluation framework with regards to establishing clear evaluation categories, impacts and indicators have been established. The techniques for applying this framework were then discussed and selected. The next section outlines the evaluation framework.

3 Evaluation Framework

3.1 Scope of evaluation framework

The evaluation framework covers the objectives, impacts and indicators of relevance to all of the activities within the CityMobil project. It also sets out how the framework should be applied to the different activities undertaken within CityMobil (SP1 Demonstrations, SP2 Future scenario simulation and SP3 & 4 vehicle and operational issues) and what information flows are required within the project to facilitate the task of project evaluation. Section 4 provides more detailed recommendations on standardised sources of information for evaluation (e.g. valuation studies) and on measurement standards for indicators.

Not all of the activities to be undertaken within the CityMobil programme are yet specified. This document provides the framework under which evaluation should be conducted. Each sub-project must therefore apply the guidelines to ensure that its evaluation is consistent with the project as a whole. Support is available from within SP5 to assist in this process and to manage the ex-ante and ex-post evaluation activities.

This document has been developed in consultation with all the actors involved at different levels in evaluation. To this aim an Evaluation Working Group has been created. Other than the project co-ordinator and the partners involved in SP5, the leaders of SP2, SP3 and SP4 as well as the local evaluation managers in WP1.2 (Heathrow demo), WP1.3 (Rome Demo) and WP1.4 (Castellón Demo) and the leader of WP 1.5 (showcases and other city studies) are members of the Evaluation Working Group. The EWG, devoted to manage and steer the evaluation action, has been important in reviewing how and when the different WPs and SPs affected by the evaluation will exchange which information.

3.2 The framework

3.2.1 Evaluation categories

The overall framework for evaluation is broken down by a series of evaluation categories. The evaluation categories are steered by both the objectives of transport policy as a whole and the objectives of stakeholders wishing to understand the success of actual implementations.

“The objectives, or goals, to be pursued form a key dimension of transport policy. When those responsible for transport policy state a set of objectives, they define the direction to be taken by transport policy, which includes imposing constraints with regard to what transport policy should avoid, and they provide a benchmark against which to measure the success of transport policy, both in terms of appraisal and monitoring” (Kallionen et al., 2004, 11).

The TIPP project recently compared policy objectives at a local, national and European level concluding that the following set of objectives capture those most commonly applied within the EU for transport policy making:

- “Efficiency
- Environment
- Safety
- Equity

- Economic development
- Future generations

Other objectives such as regional development, urban development, accessibility, health, integration and explicit objectives for freight were either only used sporadically, inconsistently applied or considered by inference through other objectives" (May et al., 2005, 14).

This overarching set of objectives is consistent with the sustainability goals of the project. However, these broad policy objectives do not serve to directly evaluate the acceptance, operational performance and implementation impacts that are required for CityMobil evaluation and recommended in MAESTRO and CONVERGE. The full list of evaluation categories is therefore:

- Acceptance
- Quality of service
- Transport patterns
- Social Impacts
- Environment
- Financial Impacts
- Economic
- Legal impacts
- Technological success

The coverage of the evaluation categories has been cross-checked to ensure that this list is comprehensive.

3.2.2 Framework for Passenger Systems

The full framework of evaluation categories set out above is populated by a series of key impacts (e.g. for under the social impact evaluation, road safety is an impact) and indicators (e.g. the number of people killed is an indicator of road safety). The list of impacts and indicators shown in the framework below has been selected from a full review of the MAESTRO urban indicators list and evaluation frameworks from the PROPOLIS (Lautso et al., 2004) and SPECTRUM (2004).

Indicators were selected first based on their importance to understanding their importance to the evaluation and secondly to their likely practicability from either a measurement or modelling perspective. Further indicators were then reviewed and added in consultation with project partners to ensure a full and balanced coverage. The indicators are listed in below in Table 3. There are 64 indicators in the list. This can be viewed as the complete envelope of overarching indicators which could form part of the evaluation of each of the different elements of CityMobil. We provide further details on how the framework should be applied for each of the SPs in Section 3.3.

We do not anticipate that all indicators will be measured at all sites neither that it will be feasible for the modelling tools to necessarily produce all of the indicators. A sub-set list of core indicators that each demonstration site and citywide simulation should report on is given below in Table 4. Where local differences or constraints make variations from this list necessary this must be agreed between project participants and the leader of SP5 as part of the ex-ante evaluation.

Table 3: Passenger evaluation categories, impacts, indicators and where indicators can be measured/modelled

Evaluation Category	Impact	Indicator	SP1 Demonstrations (Selection depends on nature of demonstration)	SP2 Future Scenarios	SP3 Vehicles and Technological Issues	SP4 Operational Issues
Acceptance	User acceptance	Usefulness	✓			
		Ease of use	✓		3.2	
		Reliability	✓			
		User satisfaction for the on demand service	✓			
		Integration with other systems	✓			4.5
	Willingness to pay	User willingness	✓			4.3
		Authorities willingness	✓			
Quality of service	Information	Availability	✓			4.3
		Comprehensibility	✓			4.3
	Ticketing	User satisfaction	✓			4.3
	Cleanliness	Perceived cleanliness	✓			
	Comfort	Perceived comfort	✓		3.2	
	Privacy	Perceived level of privacy	✓			4.3
	Perception of safety and security	Perception of safety	✓			
		Fear of attack	✓			
Transport patterns	Modal change	Induced mode changes in the other segments of the journey	✓	2.3, 2.4		
		System modal share	✓	2.3, 2.4		
	System use	Total passenger·km travelled	✓	2.3, 2.4		

Evaluation Category	Impact	Indicator	SP1 Demonstrations (Selection depends on nature of demonstration)	SP2 Future Scenarios	SP3 Vehicles and Technological Issues	SP4 Operational Issues
		Total N° of trips	✓	2.3, 2.4		
		Vehicle occupancy	✓	2.3, 2.4		
	System performances	Average Journey time per OD pair	✓	2.3, 2.4		4.4
		Journey time variability	✓			4.4
		Total delay per/trip	✓	2.3, 2.4		4.4
		Average Waiting time	✓	2.3, 2.4		4.4
		Waiting time variability	✓			4.4
		Interchange time	✓			
		Effective system capacity	✓	2.3, 2.4	3.4	4.4
Social Impacts	Spatial Accessibility	Change in range of key activities accessible within time thresholds	✓	2.3		
		Distribution of accessibility changes by social group	✓	2.3		
	Service Accessibility	Access times for mobility impaired users	✓			
	Safety	Accident levels		2.3, 2.4	3.2	4.4
		Incidents	✓		3.3, 3.4	
		Driver workload			3.2	
Environment	Energy	Daily consumption (KWh)	✓			

Evaluation Category	Impact	Indicator	SP1 Demonstrations (Selection depends on nature of demonstration)	SP2 Future Scenarios	SP3 Vehicles and Technological Issues	SP4 Operational Issues
		Energy Efficiency (KWh/pkm)	✓	2.3, 2.4		4.4
	Toxic emissions	NO_x	✓	2.3, 2.4		4.4
		PM_{10} and/or $PM_{2.5}$	✓	2.3, 2.4		4.4
		CO	✓	2.3, 2.4		4.4
	Climate Change	CO_2	✓	2.3, 2.4		4.4
	Noise	L_{DEN} and L_{night}	✓			
	Land take	Loss of green space from construction	✓	Secondary input		
		Total land use change		2.3		
Financial impacts	Start up costs	Track construction and civil works	✓	2.4		
		Vehicle acquisition/construction	✓	2.4		
		Control systems and apparatus	✓	2.4	3.3	
	Operating costs	Personnel	✓	2.3, 2.4		
		Vehicle maintenance	✓	2.3, 2.4		
		Track and civil infrastructures maintenance	✓	2.3, 2.4		
		Control system maintenance	✓	2.3, 2.4	3.3	
	Revenues	Operating revenues	✓	2.3, 2.4		
	Subsidies	Perceived public subsidies	✓	2.3, 2.4		

Evaluation Category	Impact	Indicator	SP1 Demonstrations (Selection depends on nature of demonstration)	SP2 Future Scenarios	SP3 Vehicles and Technological Issues	SP4 Operational Issues
Economic	Temporary job provided by installation and demonstration	Jobs provided at the demonstration site	✓			
		Jobs increase induced at the manufacturers	✓			
	Long terms effects on jobs	Local effects on employment		2.3		
		Non local effects on employment		2.3		
	Vitality	Footfall within defined area	✓			
		Vitality index		2.3		
	Efficiency	Net Present Value	✓	2.3, 2.4		
		Internal Rate of Return	✓	2.3, 2.4		
Legal impacts	Impacts on legal and regulatory framework	Induced regulation procedure changes	✓	2.5		
Technological success	Performance	Response time	✓		3.3, 3.4	
		Accuracy	✓		3.3, 3.4	
		Data updating delay	✓		3.3, 3.4	
	Reliability	Failure rate	✓		3.3, 3.4	

Table 4: Core indicators

Evaluation category	Indicator
Acceptance	User willingness to pay
Quality of service	Perceived level of safety
	Perceived performance
Transport Patterns	Mode Share
	Delay per passenger trip
	Journey time variability
Social Impacts	Access times for mobility impaired users
	Accident levels
Environment	Energy Use
	Toxic Emissions[3]
	Total CO_2
	Total Land-Use Change
Economic	Net Present Value
	Internal Rate of Return

We accept that even within this table there will be elements that are not common to all SPs. For example, it may be that SP2, with its focus on the strategic level will not produce estimates of willingness to pay. Likewise, the technology assessments in SP3 will have nothing to report on total land-use change. This lists acts as the principal point of focus. The full list in Table 3 then allows the SPs to tailor their selection around the most appropriate core.

3.2.3 Framework for Freight Systems

There is currently no plan for the implementation of freight systems either as a demonstration or a showcase. This greatly simplifies the examination of freight systems compared to passenger systems.

The evaluation framework will be applied within a simulation-based workpackage only. We have identified, using a similar approach as adopted for the passenger systems, a series of key indicators for each evaluation category. Many of the definitions (e.g. costs, environment) are the same as for the passenger evaluation framework and definitions are as per Annex 2. Given the comparatively tight control over where in the project freight issues will be addressed, it is felt preferable to allow the detail of the freight specific indicators to be defined by the project participants responsible for the freight simulations. These indicators should be clearly defined for the ex-ante evaluation and consistently applied for the ex-post evaluation. The approach to evaluation for freight simulations should the same as for passenger simulations (MCA or CBA plus tables).

---

[3] The most relevant toxic emissions for the local circumstances should be selected. It is expected that NO_x and PM_{10} will be the most common problem emissions.

Table 5: Freight evaluation categories, impacts, indicators and where indicators can be measured/modelled

Evaluation Category	Impact	Indicator
Acceptance	Operator acceptance	Ease of management
		Reliability
		Ease of use
		Flexibility
		Perceived performances
	User acceptance[4]	Reliability
		Delivery performance
	Willingness to pay	Operator willingness
		User willingness
Quality of service	Information	Accuracy
		Comprehensibility
		System response time
		Number of actors in chain
	Integration	Integration with other freight systems
		Number of transfers
Transport patterns	Modal change	Induced mode changes in the other segments of the journey
		System modal share*
	System use	Total tonne·km traveled
		Vehicle utilisation
	System performances	Average Journey time per OD pair
		Journey time variability
		Total delay per/trip
Environment	Energy consumption	Daily consumption
		Average kW/pkm*
	Toxic emissions	NO_x*
		PM_{10} and/or $PM_{2.5}$
	Air Quality	Number of excedences of limit guidelines
	Climate Change	CO_2*
	Noise	L_{DEN} and L_{night}
	Land take	Loss of green space from construction
Social	Safety	Accident levels
Financial impacts	Start up costs	Track construction and civil works
		Vehicle acquisition/construction

[4] Here the user is the recipient of goods.

Evaluation Category	Impact	Indicator
		Control systems and apparatus
	Operating costs	Personnel
		Vehicle maintenance
		Track and civil infrastructures maintenance
		Control system maintenance
		Transfer costs
	Revenues	Operating revenues
Economic	Temporary job provided by installation and demonstration	Jobs provided at the demonstration site
		Jobs increase induced at the manufacturers
	Long terms effects on jobs	Local effects on employment
		Non local effects on employment
		Internal Rate of Return*
Legal impacts	Impacts on legal and regulatory framework	Induced regulation procedure changes

3.2.4 Indicators

Table 3 and Table 5 list a series of indicators that could be measured as part of the evaluation. To ensure greater comparability of the results across the project it is essential that the results are compared using common metrics for each of the indicators. Energy use, for example, could be measured in Kilowatt hours, Megajoules or in tonnes of oil equivalent. Delay could be defined relative to traffic free conditions or off-peak conditions. Annex 2 provides a complete list of the indicators proposed for the passenger transport system evaluation with a brief description of the metric to be used and, where this is felt appropriate, guidance on measurement and modelling techniques. This is based on forerunner projects (such as CyberMove, EDICT and CyberCars) and EU projects specialising in the evaluation of transport and land use policies (e.g. PROPOLIS) and economic instruments (e.g. SPECTRUM).

We expect the guidance to assist partners in harmonising their approach to collecting and presenting information. We do not, within the evaluation framework specify for the partners how the indicators will actually be measured and modelled. Many of the decisions about the exact site budgets for monitoring demonstrations, the location and nature of showcases, the specification of exactly what technological developments will be assessed within SP3 and 4 and the models to be deployed in SP2 have not yet been taken and such an approach would be unnecessarily restrictive for the project and ultimately impractical.

In providing a large envelope of 64 indicators we believe that we have covered the range of important evaluation issues. Nonetheless, it is likely that partners will wish to include additional indicators within their own SP or WP which are of local value. The framework does not preclude this from happening but such indicators should not be measured as well as, rather than instead of, the relevant framework indicators. For example, consider the evaluation of dual-mode operation within the simulator environment in SP3.2. It would be standard practice for the following measures to be taken:

- Speed, variation in speed, 85th percentile of speed
- Lateral variation, variability of lane position
- Time to Collision, time headway, gap acceptance
- Performance under system failure
- Driver workload (heart rate or post-hoc survey)
- Stress
- Acceptability

The first three bullet points relate to measures to classify the safe operation of the system and would connect to the accident indicator as a qualitative measure. Driver workload is a direct measure of safety within the framework. Stress and acceptability would connect to the user acceptance of the system. These indicators are highly specific to the type of test being undertaken and therefore not suitable to the whole project evaluation framework. It is the responsibility of each SP leader, in discussion with the ex-ante evaluation co-ordinator to ensure that these links are clearly established before the testing process begins. Within the evaluation framework we have made, with the SP leaders, an initial attempt to identify where these links are most likely exist. Such an approach allows us to connect the more specific technological assessments to the larger evaluation of the sustainability impacts of Advanced Transportation Systems. This provides connectivity for decision-makers without diluting the technical information available to technology developers. This is a key aspect of our evaluation approach as the outcomes of the project need to be used at different levels.

3.3 Application of the framework

Given the complexity of the different roles that the evaluation framework will be expected to be put to use in, this section provides some further guidance about the application of the framework to the different elements of the project.

3.3.1 Demonstrations and showcases

According to the MAESTRO methodology flow reported in Figure 5, the evaluation steps are four: the definition of the objectives plus initial, ex-ante and ex-post evaluations, each one followed and preceded by one of the three project design/implementation phases.

CityMobil has completed the objective selection, thus the phases to do are three:

1. Initial evaluation;
2. Ex-ante evaluation;
3. Ex-post evaluation.

As shown in Figure 5, the initial evaluation is directly linked with site selection and pre-design phases. According with MAESTRO the steps to be done toward the evaluation plan are:

- to derive a list of impacts to observe and monitor with the necessary associated indicators;
- to choose appropriate evaluation methods for the next evaluation phases;
- to establish an evaluation plan for the remaining of the project;
- to forecast impacts on the basis of the local sites functional specifications and user expectations assessed during the pre-design stage.

CityMobil is now doing the initial evaluation. The three demonstrations has yet provided a first not common evaluation, as reported in the sub-section 2.2.3.

At this stage, the works to be made are mainly two: 1) the definition of an evaluation plan, through the selection of the set of core indicators, which have to be used in the other phases of the evaluation to measure the satisfaction with the service provided in the sites and to make a comparison between the different sites, and 2) the selection of the showcases to be included in the project, by a sort of ballot made by the partners.

The output produced in the initial evaluation has to be made of three parts:

1. The core indicators and their number;
2. The way in which such indicators will be measured;
3. The definition of sample sizes required to quantify the indicators.

Concerning the next phase, ex-ante evaluation, two main tasks have to be done.:

1. The construction of a "reference case" (a sort of before measurement) for the indicators selected in the previous evaluation phase (initial); it quantifies the actual situation by measuring the indicators at this stage. Some of the indicators will have a reference value

to be individuated, whereas other indicators will not provide reference values at this stage.

2. The first quantitative evaluations of the system done on the basis of the results of the system dimensioning and the first impact forecasting. Such evaluations allow to consider the design well working and technically good. Furthermore they provide quantitative measures concerning the expectations linked with each indicator, thus setting the reference-case for the ex-post evaluation.

Concerning the indicators to be measured, three main data collection activities will take place: 1) transport and socio-economic indicator data, 2) energy and environmental indicator data, and 3) safety indicator data.

The output produced in this phase, other than the reference case to be constructed, has to be linked with parameters as the O/D matrix representative of the case which is studied and the kind of evaluation to be done in the next phase.

Therefore this phase will assess which local project is feasible and, eventually, provide feedback to the site designers.

The ex-post evaluation is the evaluation phase that concludes the project.

The systems have to be implemented and, during the field-trials of such systems, field data measurements and surveys will be done.

The new, more detailed simulations carried out in the previous phases will give new data for the indicators, as well as some of the outcomes from the field trials will be transferred.

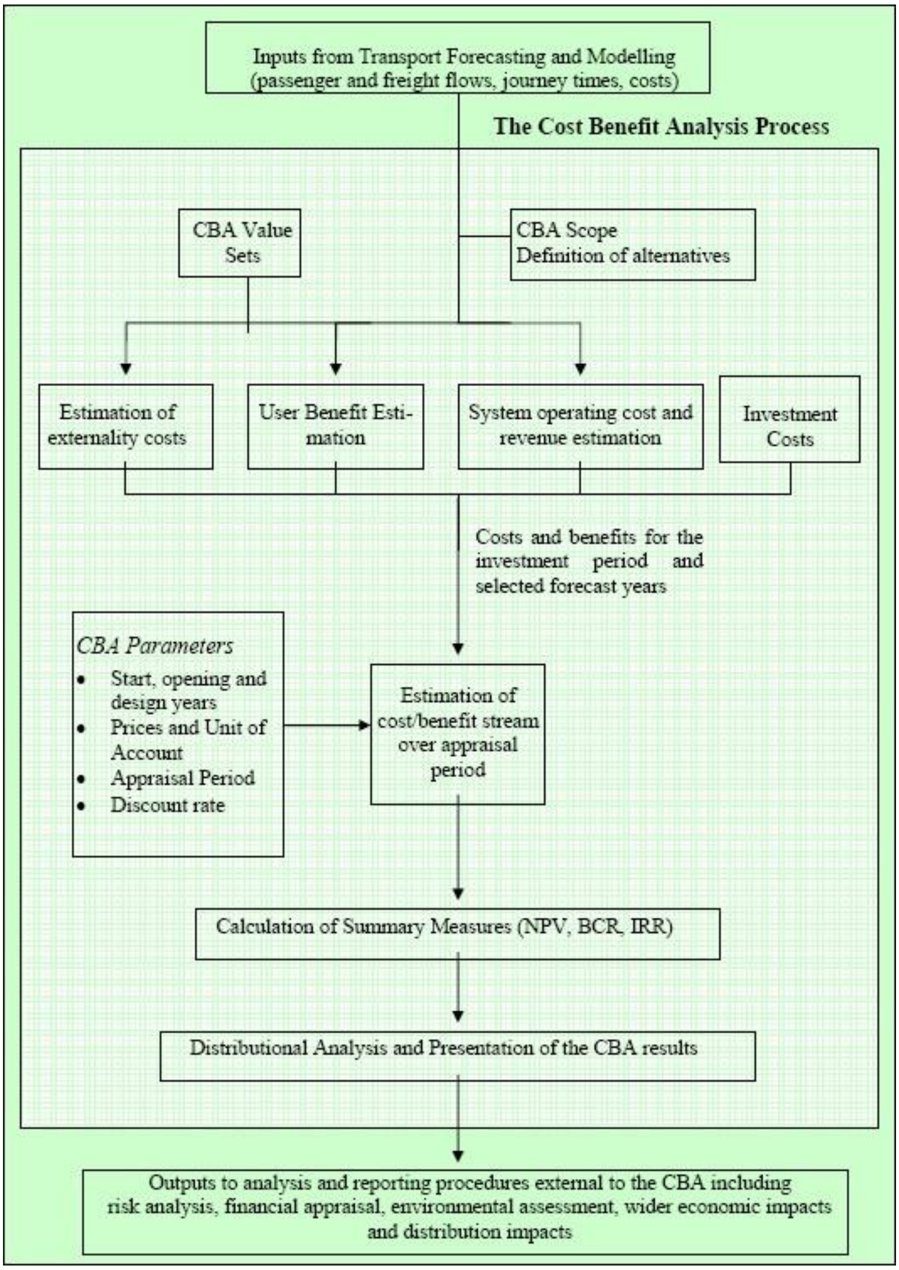
The measurements of the indicators have to be done several times, in order to provide a strong database on the basis of which the values of all the indicators can be obtained. In particular. As reported in the Technical Annex of the CityMobil project, three ex-post phases will be done in the Workpackage 5.2

In this phase the evaluations methods have to be made by an approach based on full social Cost-Benefit Analyses with additional non-weighted tables of indicators, as reported in section 2.4. Such CBAs will be the final output of the ex-post phase.

The steps to do in a CBA in order to analyse a transport system from the initial input to the final output are reported in an extended form and explained in the deliverable 5 of HEATCO, called Proposal for harmonised guidelines, especially in Section 2. The final outcome of a CBA is a summary measure such as NPV (Net Present Value) or IRR (Internal Rate of Return). Crucial factors determining the outcome of the CBA are the evaluation period / time horizon as well as the discounting rate. These issues are discussed in detail in Annex 3. It shows that their treatment is varying across Europe, median values are around 30 years and 5-6%. In order to achieve a comparability of results between the results of the demonstrations and showcases, harmonised values should be applied and, unless otherwise agreed these values should be 30 years and 5%.

In Figure 7 the Cost-Benefit Analysis process according to the deliverable 5 of the HEATCO project is reported.

Figure 7 The Cost-Benefit Analysis process for HEATCO (Bickel et al., 2005b, p.10)



3.3.2 Future scenarios

The principal application of the framework within SP2 is in Workpackage 2.3 and Workpackage 2.4.

It is standard modelling practice in assessing the benefits accruing from a transport intervention to establish calibrated current year conditions for the city and to develop alternative future policy tests for comparison. The strategic city impact assessment modelling work to be conducted in WP2.3 should include the core list of indicators and other indicators from the framework where available and applicable across the sites being modelled. Where any additional indicators from outside the framework are included, they should be reported against an impact category established within the framework.

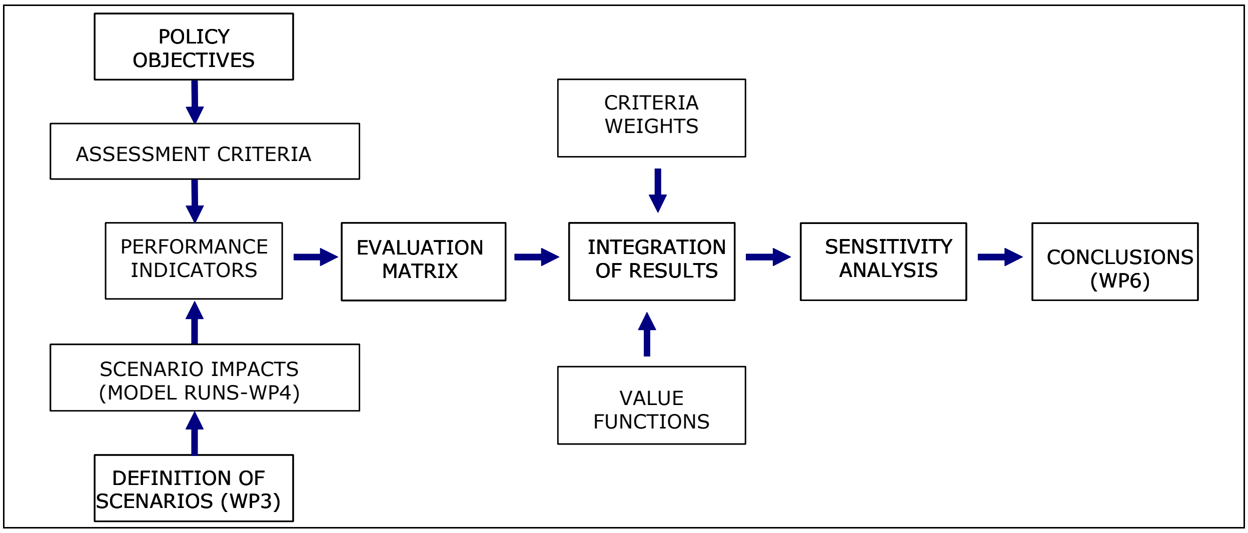
A 'do-minimum' scenario should be developed consisting of the continuation of current approaches to system development and levels of expenditure. Interventions and packages of interventions with ATS systems should be compared to the 'do-minimum'. The outputs of the modelling assessments are to feed into the ex-ante evaluation process. It is also understood that where the CityMobil project develops new understandings of the likely response of travellers to the new systems that these will be used to enhance the relationships used within the models and further model outputs will be generated.

Such a task is in Workpackage 5.3, where other advanced system applications will be evaluated in the light of the experience gained from the on-going demonstrations, thus providing ex-ante evaluations of other case studies as final output.

For the models we suggest a more data intensive multi-criteria approach comparable to the STEPS project. Figure 8 illustrates the outline of the MCA methodology.

To adopt this to the CityMobil project, the evaluation framework described in this chapter provides the evaluation categories and performance indicators, while SP2 will deliver the scenarios and impact results from the modelling runs. The result will be presented in an evaluation matrix. Subsequently, each indicator result is transformed into a score (or partial utility value) on a harmonized scale by a value function. In the STEPS project, linear value functions with scores ranging between 0 (least preferred) to 100 (most preferred) were applied. For the economic efficiency criterion, alternatively to the application of a value function for each indicator, a combined value for money indicator could be the outcome of a cost-benefit analysis and subsequently transformed to a score.

Figure 8: Outline of the STEPS MCA methodology adapted to CityMobil (based on Lopez, Monzón, 2006, p. 71)



For the aggregation of the scores to a utility value for scenarios, a weighting function needs to be applied to all criteria. In the linear, additive model used in STEPS, the weights add up to 100%. Figure 9 illustrates the valuation and weighting procedure in the linear additive utility based MCA with m potential scenarios S_1 ... S_m and n assessment criteria C_1 ... C_n.

Figure 9: Valuation and weighting procedure in the linear additive utility based MCA approach

Target Achievement — Target Criteria

$$\begin{array}{cccc} & c_1 & \cdots & c_n \\ S_1 & e_{11} & \cdots & e_{1n} \\ \vdots & \vdots & \ddots & \\ S_m & e_{m1} & \cdots & e_{mn} \end{array}$$

Valuation $\varepsilon_{ii} = f(e_i)$ $\longrightarrow$

Scores — Target Criteria

$$\begin{array}{cccc} & c_1 & \cdots & c_n \\ S_1 & \varepsilon_{11} & \cdots & \varepsilon_{1n} \\ \vdots & \vdots & \ddots & \\ S_m & \varepsilon_{m1} & \cdots & \varepsilon_{mn} \end{array}$$

Weighting $n_{ii} = g(\varepsilon_{i1}, \ldots, \varepsilon_{i1})$ $\longrightarrow$

Utilities — Utility

$$\begin{array}{cc} & U \\ S_1 & u_1 \\ \vdots & \vdots \\ S_m & u_m \end{array}$$

The weights as well as the value functions need to be determined in future tasks of the CityMobil project. They can differ between ATS applications as much as the specific objectives of these applications might be different. These weights will be obtained in work package 5.3 with the weights being defined during the second half of 2007. through a survey which involves relevant decision makers from the demonstration sites and other cities. Finally, a sensitivity analysis will be carried out in which the stability of the evaluation results is tested under the variation of valuing and weighting functions.

The evaluation period should be equal to the policy test horizon chosen and value functions and weightings will then reflect the decision makers' preferences with regard to intergenerational aspects.

The business case assessment will focus primarily on the balance of costs and revenues for different system configurations. Whilst such a 'business case' might be seen to be relatively narrow in focus, it sits within the context of a complete sustainability assessment for the systems functioning within the whole city dynamic (discussed above). Local values of costs and benefits should be used to calculate benefit:cost ratios and financial rates of return. Data on values of time, emission rates and externality costs should also be local values where applicable. Further guidance is provided on this in Annex C.

3.3.3 Technological Assessments

The technical evaluation focuses on the question whether and to what extent the systems tested fulfil their technical objectives.

A technical assessment is usually the first kind of evaluation which is performed in the research and development process, the installation process, and the verification process of a technical system. In the verification process the actual operation of the system and its functionalities at the demonstration sites are tested. Questions here are: is the functionality there? Does it work? Does it work according to the specifications? Is it really used? After that, the fully integrated functionality of the demonstrator is verified and its performance is tested.

The criteria which are used for this type of assessment are technical indicators of the functions of the tested system. This makes the technical assessment one of the most specific stage of the overall evaluation methodology. Usually these criteria are specified and measured by the technical experts developing and/or using the system(s).

Systems to be evaluated technically in the context of CITYMOBIL can be distinguished according to the following categories:

- Obstacle detection and avoidance systems;

- Cooperative vehicles and navigation systems;
- Vehicle management inside the operating areas.

In all cases the installed systems should be compared with the zero-state, i.e. the situation in which no system is installed, or in which the “old” system was functioning. For example: the telephone is replaced by a mobile data communication system to contact the driver, a trip planning package is utilised instead of manual planning.

It will be the responsibility of the demonstration site manager to carry out the technical evaluation, together with people at the test site, under supervision of the expert from SP5. It is important that the technical evaluation is conducted in a full operational environment.

Three important steps form the technical evaluation:

1. Verification of functionalities
2. Verification of the performance of the full-working system
3. Justification of the solution chosen

This last step should assure that the alternatives, or the “old” solution, was not a better one after all. This involves an analysis of the cost-effectiveness of the interventions.

Four indicators have to be considered significant in assessing the technical functionality of the applications to be tested: response time, accuracy, data updating delay and failure rate.

Failure rate and response time are used to quantify service reliability; data updating delay and accuracy were chosen to quantify the system performances.

Response time is measured from the user point of view, through manual tests. Such tests have to be carried out in different conditions, all of them however representative of the reality, as reported in the next section 3.4.

Accuracy is defined as the number of times in which the service provided was right (for example people arrived to the destination without delay times) against the number of times it was tested.

Data updating delay refers to the time elapsed since the last information data updating. For example, it can be considered when a user is waiting for a vehicle at a stop and the time foreseen before vehicle arrive is provided.

The failure rate is measured as the number of times in which the service is not working properly over the number of tests.

3.4 Data requirements

Ideally assessment should be based on measurements on an application under full-scale, real-life operation.

However, often this is impractical or prohibitively expensive. In practice, small-scale but representative measurement or simulation may be required to provide the necessary data.

The conditions surrounding (and influencing) the data collection should as far as possible be controlled and homogeneous. Thus the time of day, traffic and weather conditions etc. must be chosen so that a group of measurements or simulations take place under more or less the same conditions (blocking).

As reported in the deliverable 2.3.1 of the CONVERGE project, three considerations have to be done in the data collection phase, in order to have representative data:

1. The application may perform differently from the reference case for particular conditions of measurements;

2. Specification and calibration of a simulation model may vary in adequacy over the range of conditions being simulated;
3. Measured indicators may be strongly correlated with parameters which describe the measurement conditions.

The usual response to such considerations is to measure or simulate indicators for conditions which are as well-defined as possible (that is, as homogeneous as possible).

Concerning statistical considerations strictly linked with the sample of measurements needed, for individual indicators is necessary to:

- Determine the expected improvement due to the application;
- Associate this improvement with a level of statistical confidence if a statistical approach is possible.

The linking of an assessment objective with overall definition(s) of success for its related indicator(s) is a fundamental step in the design of validation methods.

If a statistical approach is possible, it is necessary to associate the expected improvement with a level of statistical confidence.

The main considerations are:

- In general terms, a larger sample is needed:
  - the smaller the expected improvement in performance
  - the greater the variation between individual measurements of the indicator
  - the greater the required statistical accuracy
  - if the results are to be compared with those at other sites.
- A larger sample is needed for questionnaire surveys as the level of required analysis becomes more disaggregate (simply because fewer responses become available at higher levels of dis-aggregation).
- it is reasonable to attach a higher level of credibility (as distinct from statistical confidence) to results based on objective or 'hard' measurement than to results based on subjective or 'soft' measurement such as by interview/questionnaire methods.

It is important to plan the timing of measurement of indicators for the trial application and the reference application in such a way that the measurement plan does not itself contribute to bias in the comparison of the performance of the application and the reference application.

The last consideration to do in data collection is about the integrity of measurements.

Three characteristics linked with the measurements have to be taken into account:

1. Completeness: limited resources have not to influence the measurements, thus not only easiest indicators have to be considered, but all the indicators evaluated as important;
2. Insularity: measurement or simulation of indicators for reasonably well-defined conditions may still not deal adequately with all the factors which might influence the performance of the application and the reference application;
3. Disturbance of the validation process: there may be accidental or intended bias introduced into the measurement process by factors as respondent fatigue, policy response bias and justification bias, which have to be taken into account in the validation of data.

Once collected the data, two main different techniques can be used for the data analysis. Such techniques, reported in the extended form in the deliverable D2.3.1 of the CONVERGE project, are the interval data technique and the categorical data technique.

4 Summary

The CityMobil project requires an evaluation framework capable of capturing the social, environmental, economic, legal and technological impacts of Advanced Transportation Systems. The framework is required to operate at spatial scales ranging from laboratory and test-track interventions, through computer modelling to real-world implementations on a large scale.

This document sets out the key objectives that the project should monitor progress towards and the main measures that should be used to assess this. It also seeks to provide a practical guide to how the framework should be applied at each of the different spatial scales to ensure that the evaluation provides information which is both targeted at the different users (technology developers through to politicians) but closely linked.

The framework has been built around previous experience in the evaluation of ATS and through an understanding of the state-of-art in land-use and transport evaluations. It is an aim of the evaluation framework is to mainstream the assessment of ATS within local and national evaluation processes.

This document forms the start point for all of the project workpackages to reflect on and then specify how their outputs connect to the overall project aims. The ex-ante evaluation is the next stage in this process where the information flows highlighted in this report have to be realised. This document has been prepared with the input of members of a project wide evaluation working group. The outcomes therefore represent a project-wide commitment to adhering to the principles and processes identified within.

5 Sources

5.1 Reference List

BICKEL P., BURGESS A., HUNT A., LAIRD J., LIEB C., LINDBERG G., ODGAARD T. (2005a) HEATCO Deliverable 2 State-of-the-art in project assessment

BICKEL P., FRIEDRICH, R., BURGESS A., FAGIANI, P., HUNT A., DE JONG, G., LAIRD J., LIEB C., LINDBERG G., MACKIE, P., NAVRUD S., ODGAARD T. RICCI, A., SHIRES, J., TAVASSZY, L. (2005b) HEATCO Deliverable 5 Proposal for Harmonised Guidelines

BICKEL, P., FRIEDRICH, R. (Editors) (2005) ExternE - Externalities of Energy: Methodology 2005 Update. European Commission DG Research, EUR 21951

BICKEL P., FRIEDRICH R., LINK H., STEWART L., NASH C. (2006) Introducing Environmental Externalities into Transport Pricing: Measurement and Implications. Transport Reviews, Vol. 26, No. 4, 389–415, July 2006

CONVERGE (1998), *Guidebook for Assessment of Transport Telematics Applications: Updated Version,* Deliverable 2.3.1; Telematics Applications Programme - Transport Sector, Project TR 1101

HEATCO (2005): *Proposal for harmonised guidelines*, Deliverable 5; European Commission research contract number FP6-2002-SSP-1/502481

HEDEL, RALF; VANCE, COLIN (2005): On the application ELECTRE III to the analysis of rail freight options: Advances in assessing the sensitivity of rankings. In: 2005 MCDA62 Workshop, 2005 MCDA62: 62nd Meeting of the European Working Group on "Multiple Criteria Decision Aiding", Borlänge

HELMREICH W. (2000) RECONNECT "Reducing Congestion by Introducing New Concepts of Transport". Final Report for Publication

KALLIOINEN, J., LINDSEY, R., MATTHES A., MATTHEWS, B., MAY, A.D., MOILANEN, P., DE PALMA, A., SEIDEL, T. AND WIELAND, B., 2004, *Theoretical and practical approaches to institutional analysis, Transport Institutions in the Policy Process*, Deliverable 2, European Commission DG TREN, Available from: http://www.strafica.fi/tipp/TIPP-D2.pdf, last accessed 11th October 2006

LAUTSO, K., WEGENER, M., SPIEKERMANN, K., SHEPPARD, I., STEADMAN, P., MARTINO, A., DOMINGO, R. AND GAYDA, S., 2004, *Planning and research of policy for land use and transport for increasing urban sustainability* (PROPOLIS), Final Report to the European Commission, Version 2.0, Contract No: EVK4-1999-00005

LOPEZ, E., MONZÓN, A. (2006) The assessment methodology. In: Monzón, A, Nuijten, A. (Eds.): Transport strategies under the scarcity of energy supply. Published by Buck Consultants International, The Hague

MAESTRO CONSORTIUM (2000), *Guidelines for transport in the 21st century*, MAESTRO Project, European Commission Research contract number PL-97-2162, Brussels.

MAY, A.D., MARSDEN G., ZOGRAFOS, K., TSANOS, C.,WIELAND, B., MATTHES A., PETER, B., TEGNER, H., MOILANEN, P., AND DE PALMA, A, 2005, *Institutional Issues of Transport Policy Implementation: Synthesis and Recommendations*, Deliverable 7, European Commission DG TREN, Available from: http://www.strafica.fi/tipp/TIPP-D7.pdf, last accessed 11th October 2006

MINKEN, H, SHEPHERD, S., JÄRVI-NYKÄNEN, T., MAY, T., PEARMAN, A. PAGE, M., ARJAMAA, H., VOLD, A. JONSSON, D., PFAFFENBICHLER, P. (2002) PROSPECTS Procedures for Recommending Optimal Sustainable Planning - of European City Transport Systems: Evaluation Tools. Deliverable No. 2. European Commission, Funded by 5th Framework RTD Programme

NASH C. (2003) with contributions from partners: UNITE (UNIfication of accounts and marginal costs for Transport Efficiency) Final Report for Publication, Funded by 5th Framework RTD Programme

ODGAARD T., KELLY C., LAIRD J. (2005) HEATCO Work Package 3: Current practice in project appraisal in Europe. Deliverable 1/Volume 1 (main text)

SPECTRUM, (2004) Study of Policies regarding Economic instruments Complementing Transport Regulation and the Undertaking of physical Measures, Deliverable 6, PROJECT N° : GMA2/2000/32056-S12.335673, European Commission.

TAVASSZY L., JONKHOFF W., BURGESS A., RUSTENBURG M., HUNT A. (2005) HEATCO Deliverable 3 - Key Issues in the Development of Harmonised Guidelines for Project Assessment and Transport Costing

Transland (2000) Deliverable 4, Final Report, Integration of Transport and Land Use Planning, Available from: http://www.inro.tno.nl/transland/ last accessed 11th October 2006

TSAMBOULAS D., MIKROUDIS G. (2000) EFECT - evaluation framework of environmental impacts and costs of transport initiatives. Transportation Research Part D 5 (2000) 283-303

TUDELA A., AKIKI N., CISTERNAS R. (2006) Comparing the output of cost benefit and multi-criteria analysis – An application to urban transport investments. Transportation Research Part A 40 (2006) 414–423

UKKUSURI S., KAROONSOONTAWONG A., WALLER S. T., KOCKELMAN K.M. (2004) Congestion Pricing Technologies: Synthesis and an Evaluation Framework. Submitted for publication in Transportation Research Part C December 2004

VREEKER, R. and NIJKAMP, P. (2005) Multi-criteria evaluation of transport policies, in Eds (Button, K.J. and Hensher, D. A.) *Handbook of Transport Strategy, Policy and Institutions*, Handbooks in Transport Vol. 6, Elsevier, UK, 507-525

VREEKER, R. AND NIJKAMP, P. (2005) Multi-criteria evaluation of transport policies, in Eds (Button, K.J. and Hensher, D. A.) Handbook of Transport Strategy, Policy and Institutions, Hanbooks in Transport Vol. 6, Elsevier, UK, pp507-525

WALTER, F.; GUBLER, F.; SOMMER H. (2003) NISTRA: Nachhaltigkeitsindikatoren für Strasseninfrastrukturprojekte Ein Instrument zur Beurteilung von Strassen¬infrastrukturprojekten unter Berücksichtigung der Nachhaltigkeitsziele. Methodenbericht, Bern

WCED (1987) Our Common Future, World Commission on Environment and Development, Oxford University Press

ZHANG, X., KOMPFNER, P., WHITE, C., AND SEXTON, B. (1998), Guidebook for Assessment of Transport Telematics Applications: Updated Version, Fourth Framework Telematics Application Programme, Transport Sector, CONVERGE Project TR 1101, Deliverable 2.3.1.

5.2 Databases

http://www.eltis.org