# OPUS

## *Optimising the use of Partial information in Urban and regional Systems*

### Project IST-2001-32471

IST Programme

| | |
|---|---|
| **Title :** | **Evaluation plan** |
| **Author(s) :** | Axhausen, K.W. (IVT, ETH Zürich) |
| **Deliverable No. :** | D12.1 |
| **Version :** | 1 |
| **Date :** | Final version |
| **Dissemination Level :** | LI — Limited to programme participants |
| **Deliverable Nature :** | RE — Report |
| **Deliverable Type :** | PD — Programme Deliverable |
| **Project Coordinator :** | Imperial College London |
| **Contact Person :** | Prof. John Polak |
| **Address :** | Centre for Transport Studies<br>Department of Civil and Environmental Engineering<br>Imperial College London<br>London SW7 2AZ<br>United Kingdom |
| **Telephone :** | +44-(0)20-7594.6089 |
| **Fax :** | +44-(0)20-7594.6102 |
| **e-mail :** | j.polak@imperial.ac.uk |
| **Consortium :** | CTS, DEPH, TfL, KATALYSIS, ETHZ, FUNDP, PTV, SYSTEMATICA, WHO.<br>MINNERVA, SURVEY AND STATISTICAL COMPUTING |

# TABLE OF CONTENTS

# TECHNICAL ABSTRACT

The evaluation plan (deliverable D12.1) describes how the OPUS consortium will evaluate and document the results of its case and feasibility studies and through this the progress delivered by the OPUS methodology.

# EXECUTIVE SUMMARY

This Evaluation Plan is Deliverable D12.1 of the Fifth-Framework project OPUS. Against the background of the general OPUS methodology and the specific case and feasibility studies it sets out the steps undertaken to evaluate the performance of the methodology in the specific case studies, but also the overall progress made by the project.

The approach has four elements. The first element requires the case and feasibility studies to document the quality of their estimates in comparison with independent controls and the input data. The specific contribution of a generally Bayesian approach is an insight into the variance of the estimates. The second element requires the case and feasibility study to document these estimates and to assess the improvements possible in comparison to alternative approaches or through incremental addition of data sources. The project has advanced the state-of-the-art in the archiving and documentation of data and models through its work on advanced meta-data standards. The third element asks the case and feasibility studies to archive their data, models and results in the spirit of these advances and to comment on the efforts involved.

As the feasibility studies were originally not scheduled to produce numerical results, steps 1 to 3 will only be possible for those, which were able to advance well beyond their original task list. In the same spirit, case and feasibility studies might not be able for time and budget reasons to obtain alternative estimates from competing or naïve approaches. In this case, they will provide qualitative assessments of the contributions of the OPUS approach.

The final element sketches a qualitative approach to obtain an overall evaluation of the project through the joint assessment of the case and feasibility studies and the methodological advances made.

# 1. PURPOSE

The OPUS project is developing a general statistical framework to improve the combination of complex spatial and temporal data from survey and non-survey sources. This approach is Bayesian in principle and acknowledges the structural relationship between the variables of interest while accounting sampling and non-sampling errors. The framework will be applied in a series of case and feasibility studies of increasing complexity drawn from the transport and health fields.

The technical annex describes the specific objective of this work package and therefore deliverable as:

- This work package will provide an independent evaluation of the effectiveness of the methods in a practical application. It will concentrate on validation of results and examine the extent to which the results represent an improvement on previous methods based on input from all project partners.

The changes in the timetable and the results obtained since the start of the project have led to a readjustment of these objectives, so as to maximise the contribution of this report to the further progress of the project. The deliverable will focus now on:

- Sketch of the case and feasibility studies
- Outline of the evaluation strategy for the individual contributions
- Outline of the joint evaluation of all case and feasibility studies

The next section will briefly describe the case and feasibility studies. The main section will describe the four elements of the evaluation. The deliverable concludes with a summary and suggestions for future wider evaluation approaches.

# 2. BRIEF DESCRIPTION OF THE CASE STUDIES

Evaluation is not context independent. While there are general objectives, an evaluation procedure should account for the specific case to be able to capture the essence of it. To set the scene, this section will briefly describe each of the case and feasibility studies.

## 2.1 Depart of Epidemiology

The Department of Epidemiology and Public Health at Imperial College will integrate transport and time activity data, which in turn with airborne pollution and hospital admissions data will lead to improved estimates of the impact of transport use on cardiovascular disease in Sheffield..

## 2.2 Imperial: Tebaldi case study

The Centre for Transport Studies (CTS) will expand the work started in Deliverable 4.2 to substantially larger areas of London, incorporating further data sources.

## 2.3 Imperial: Dominici case study

In its second case study, CTS will focus on the updating of complex contingency tables.

## 2.4 Systematica: OD-estimates

Systematica will also be addressing an origin-destination problem in the context of the transport plan for the province of Lombardy. Here substantial origin-destination survey information needs to be integrated with transport models and traffic counts to obtain detailed flow estimates for a potential new river crossing.

## 2.5 FUNDP: Agent generation

The FUNDP will engage in a comparative study of a traditional parametric approach and the Bayesian approach of the project. The problem at hand is the generation of a large scale artificial sample of agents for later micro-simulation work. Using census information FUNDP will carefully compare the conceptual advantages of the approaches.

## 2.6 WHO: Health impact assessment

The World Health Organization will define potential areas of application for OPUS methods within the health domain. An area of relevance to the development and utilization of OPUS, that is particularly important to the health sector, is the implementation of health impact assessment models in a Bayesian perspective addressing all the form of uncertainty. Particular attention will be given to particulate matter exposure and health impact in terms of attributable deaths and years of life lost.

## 2.7 IVT: Leisure excursions

This case study addresses the combination of various data sources to obtain a consistent and richer estimate of the total amount of travel. Currently three sources provide information about the number of leisure excursions and day trips: the national travel survey (Mikrozensus Verkehr 2000), the national income and expenditure survey (Einkommens und Verbrauchserhebung 2000) and the privately funded survey Schweizer Reisemarkt (Swiss Travel Market 2000). These three are inconsistent in their scopes, partially in their object definitions and their sampling processes. The aim of this case study is to integrate the partial information available in each to obtain a joint estimate of the distribution of trip making for excursions.

## 2.8 IVT: Updating the regional origin-destination matrix using multiple data sources

The IVT has access to a range of data which provide information about the traffic volumes in the Zürich area. These sources look at different aspects of the traffic. The automated count data of the Federal Office for Roads (ASTRA) and the automatic and manual counts of the Canton Zürich provide estimates of volumes in each direction at a cross section. At some locations, these are supplemented by speed measurements. The third data source are speed estimates for each link in the regional network derived from a spatial-lag regression model (Hackney, Bernard, Bindra, Axhausen, 2006) of floating car data collected in 2003 by the IVT.

The case study employs the expanded Tebaldi and West sampler in an incremental fashion. Given its size it is an important contribution in establishing the stability and speed of the overall approach.

# 3. EVALUATION PROCESS

The overall evaluation process needs to be both systematic and adaptable, given the differences among the case studies. The subsections below will detail the suggested approach.

The evaluation approach is predicated on certain assumptions with regards to data availability, time and resources for the implementation of the OPUS and alternative approaches. This best case scenario is beyond the scope of the feasibility studies as originally conceived. It might also be beyond the scope of the case studies due to unforeseen data problems or software limitations. In this case a more qualitative assessment is appropriate, which still should integrate the elements of the full evaluation methodology were possible.

## 3.1 Case study assessment

The case studies have different degrees of complexity with regards to their size (number of elements; number of interactions modelled etc.), their modelling structure (the number of submodels, number of parameters etc.) and their temporal range. In the presence of submodels integrating different data elements, the case study will perform the tests described below for each submodels.

The initial models (data sets) and the information to be added each have errors, or better biases and variances, if compared with each other and, in particular, third data sets, which have been kept apart from the others so far. Even before the first model step, the case studies will document

a) the moments and co-variances of the elements (results of the initial models; variables from the additional data). The selection of these elements should be representative, rather than comprehensive. The considerations behind their selection need to be discussed in the respective reports.

b) the elements and results will be compared with known, independently derived aggregate indicators and their distributions. Examples from the transport domain are mean trip lengths, speeds and durations, which are available from both a modelled origin-destination matrix and a travel diary survey; for health one could consider mean morbidity or mean exposure to a particular risk or pathogen. In addition to such point estimators, the case studies will compare the moments and distributions of the underlying variables or processes. Again, in the transport case, one should compare trip length, duration and speed distributions.

c) finally, where possible, a comparison with spatially or temporally defined controls will be defined. In the transport case, for example, independent traffic counts by hour or time period would provide such an opportunity. In the health domain, one might wish to look to spatially or socially defined subgroups at particular times.

The comparisons will be conducted using appropriate tests of difference between the modelled and control data. These comparisons will be conducted for the results of each submodel, which is part of the overall modelling process.

In some cases, there are alternative methods available to perform the task of the case study, in full or in parts. Sometimes these methods will be naïve, such as a weighted mean of the three data sources in the first ETH case study, sometimes they will be mathematically sophisticated. Were possible the case studies will implement such alternative approaches and will compare the results, as sketched in the second and third bullet point above.

## 3.2 Contribution of the Bayesian approach

The generally Bayesian approach aims to integrate and fuse information in a systematic way assuming that additional information helps to improve an estimate. In addition, the OPUS methodology wants to provide consistent estimates of the variances of its estimates. This implies that the case studies always generate at each step an artificial sample of the estimates and variables in the usual way. Point estimates are not enough.

In the spirit of the discussion above, the simulated results distributions can be compared with the previous data or results at each step of the way. For the second Zürich case study one would provide these statistics when each of the three data sources has been integrated.

Where such an opportunity does not exist for a case study, it will be replaced by the repeated application of the model based on partial additional data sets. The reduction of the data sets will be varied by 0% and 30% to obtain an idea of the value of additional data. This approach allows the comparison of the model results with the artificially removed items. These will also be reported.

## 3.3 Archiving the case studies

The case studies will be documented and archived using the DDI standard, as well employing the approaches developed here for modelling meta-data (StatModel). The case studies will report on the effort involved (labour and other resources, such as software acquisition and training) and will assess qualitatively what benefits this process had for the research results and quality.

## 3.4 Qualitative approach for the first three steps

As mentioned above the case and feasibility studies had different tasks within the project as a whole and the degree of implementation of the OPUS differed by design. It is therefore necessary to provide for a qualitative assessment of the contributions of the OPUS approach. If possible, the quantitative elements described above should be integrated.

The qualitative assessment should address the following points:

> What were the objectives of the study and how do they relate the OPUS objective of bringing various data together to produce a better result ?
>
> What was the quality and characteristics of the prior information ? How will the additional priors, e.g. distributional assumptions or meta-parameters, be determined ?
>
> What is the contribution of the OPUS approach to these objectives ? (documented through a discussion of the expected results or of the results obtained; see above)

What was learned and what would be the next steps ? (In particular comments on the role that the synthetic data derived from the joint posterior distributions might offer to analysts and modellers, especially in comparison with the input information)

## 3.5 Overall assessment

In the final stage of the assessment, the case studies will compare their progress against the original list of objectives, as set out in the technical annex and cited above, as well integrating the numerical and qualitative assessments and comparisons into an overall judgement.

The project co-ordinator will base his global assessment on these case study assessments.

# 4. SUMMARY AND OUTLOOK

The deliverable has set out the evaluation approach for the case and feasibility studies. It provides for both a quantitative and a qualitative approach, respecting the different tasks of the case and feasibility studies, as well as the differing levels of implementation, which were feasible during the project.

# 5. LITERATURE

Axhausen, K. W. and M. R. Wigan (2003) Public use of travel surveys: The metadata perspective, in P. Stopher and P. M. Jones (eds.) Transport Survey Quality and Innovation, 605-628, Pergamon, Oxford.

Chalasani,V. S., S. Schönfelder and K. W. Axhausen (2002) Archiving travel data: The Mobidrive example, Arbeitsberichte Verkehrs- und Raumplanung, 129, Institut für Verkehrsplanung, Transporttechnik, Stassen- und Eisenbahnbau (IVT), ETH, Zürich.

Tebaldi, C. And M. West (1998) Bayesian inference on network traffic using link count data, *Journal of the American Statistical Association*, **93** (442) 557-576.