Final Report for Publication

Transport Research
Fourth Framework Programme
Strategic Transport
DG Transport

Technologies for European Surveys of Travel Behaviour (TEST)

Project Co-ordinator: Leopold-Franzens-Universität, Innsbruck

Partners: Transportes, Inovaçao e Sistemas
Socialdata
Statistics Sweden
INRETS
CBS
Statistics Netherlands
Deutsche Forschungsanstalt für Luft- und Raumfahrt
Imperial College, London
TRAIL
Facultés Universitaires ND de la Paix, Bruxelles

PROJECT FUNDED BY THE EUROPEAN COMMISSION
UNDER THE TRANSPORT RTD PROGRAMME
OF THE 4TH FRAMEWORK PROGRAMME

TABLE OF CONTENTS

List of figures

PARTNERSHIP

Coordinator:

Prof. Kay Axhausen
Dipl.-Ing.. (FH) Ines Haubold
Institut für Straßenbau und Verkehrsplanung
Leopold-Franzens-Universität
Technikerstr. 13
A - 6020 Innsbruck

Partners:

Prof. Jose Viegas, Mrs. Cristina Elvas
Transportes, Inovação e Sistemas
Rua Vilhena Barbosa 11
P - 1000 Lisboa

Dr. Max Herry
Argentinierstr. 21
A - 1040 Wien

Mr. Erhard Erl
Socialdata
Hans-Grässel-Weg 1
D - 81375 München

Mr. Jon-Inge Lian, Mr. Jon-Martin Denstadli
TOI
Box 6110 Etterstad
N - 0602 Oslo

Mrs. Amie Lindeberg
Statistics Sweden
Box 24300
S - 10451 Stockholm

Dr. Jean-Loup Madre, Mr. Jean-Paul Hubert,
Mr. Olivier Flavigny, Mrs. Sophie Midenet,
Mrs. Francoise Fressant
DEST
INRETS,
2 Ave General Malleret Joinville
F - 94114 Arcueil Cedex

Mr. Ger Moritz
CBS - Statistics Netherlands
Department of Transport Statistics
P.O. Box 4481
NL - 6401 CZ Heerlen

Dr. H.G. Nüsser
Deutsche Forschungsanstalt
für Luft- und Raumfahrt
Linder Höhe
D - 51147 Köln

Mr. John Polak, Dr. Paul Jackson,
Dr. John Plaxton
Centre for Transport Studies
Department of Civil Engineering
Imperial College
UK - London SW7 2BU

Drs Mart Tacken
TRAIL
TU Delft
Berlageweg 1
NL - 2628 CR Delft

Prof. Phillipe Toint, Mr. Olivier Lothaire,
Mrs. Isabelle Reginster
Transportation Research Group
Facultes Universitaire ND de la Paix
61, rue de Bruxelles
B - 5000 Namur

EXECUTIVE SUMMARY

The cost effective collection of long distance travel behaviour data is crucial for the formulation of the European transport policy, especially at the European level. While the project MEST (Methods for European Surveys of Travel Behaviour) was developing an European benchmark survey (see e.g. Youssefzadeh and Axhausen, 1996), this project has been involved in developing innovative technologies to improve data collection, processing and dissemination. It addressed in this way some of the objectives of Task 1.9 of the Strategic Transport Research Programme "Structure and maintenance of the transport information system", while furthering the overall objectives of more efficient data collection and dissemination.

While there was a whole range of possible avenues for improvement, the project concentrated on five areas considered to be very promising:

- Using hand held computers as a tool for near real time data collection
- Offering respondents Internet based forms for the completion of long distance travel surveys.
- Using geographical information systems (GIS) to enrich travel diary data sets, in particular with respect to geocoding
- Using artificial intelligence approaches to improve data imputation and error correction
- Building a World Wide Web interface to disseminate survey results

Some of the tools developed were tested in final part of the project by a partner not involved in their development.

The work of all workpackages demonstrated not only the feasibility of the approaches chosen, but also their long term potential:

- the hand held HPC based travel diary application (TDA) was shown to be acceptable under the conditions of small scale field tests in four countries. These trials were successful in gathering data on journeys.
- the www based iTDA was also a success in two countries, in spite of the current low speeds of the Internet. The overall system architecture with its newly developed specialised HTML tags opens new avenues for travel behaviour research
- Existing sampling difficulties (in the case of the Portuguese tests: women, older respondents) limit the application of the iTDA at the moment but fast rising market penetration and greater familiarity across the population as a whole should reduce these problems in the near future.
- the GIS based work advanced knowledge about the required structure and size of place name databases and their sources substantially and provided a semi-automatic tool for geocoding of places and the derivation of shortest paths between places visited

- the AI based parsing system developed makes the classical techniques of imputation available to transport research enriching them with logical checks for the standard variables and providing for the first time an audit trail for the data correction process
- the neural net based work showed that self organising maps (SOM) can be a valid alternative to many classical imputation techniques, while also offering the possibility of detecting erroneous data
- the www based interface to travel diary data demonstrated the usefulness of this open approach to data publication with a sophisticated web site built with free software tools, making it transferable to all interested parties

The results of the project, both deliverables and software programmes, are public. The Consortium expects suitable acknowledgement in the case of citation and use (See http://www.fundp.ac.be/~grt/test for the full deliverables).

The Consortium and its members will continue with the work begun in TEST and will build on the results obtained. The www interface, for example, will be integrated into the publication of the new Belgian National Travel Survey.

The Consortium hopes that the new methods will be integrated into the detailed work for subsequent field work and analysis.

1 OBJECTIVES AND PROJECT APPROACH

Cost effective data collection is a crucial element in the development of the European Transport Policy Information System (ETIS). The companion project to this project, MEST (Methods for European Surveys of travel Behaviour) focussed on the development of a suitable benchmark survey approach (design, protocol, definition). This left TEST free to concentrate on the development and exploration of a series of new technologies, which were considered to be promising:

- Using hand held computers as a tool for near real time data collection
- Offering respondents Internet based forms for the completion of long distance travel surveys.
- Using geographical information systems (GIS) to enrich travel diary data sets, in particular with respect to geocoding
- Using artificial intelligence approaches to improve data imputation and error correction
- Building a World Wide Web interface to disseminate survey results

A final evaluation field test evaluated some of the above elements.

The development of new technologies and tools requires not only the freedom to explore without external micro management but also feedback from colleagues and interested parties to help focus on joint goals. The project tried to achieve this dual management objective by co-ordinating the independent work of the respective work package leaders through series of internal project meetings and workshops with external participants (about every 4-5 months). Internal working papers were presented, describing the current thinking of the relevant research group.

This approach worked well, has produced a group of high quality new tools and has demonstrated their feasibility. Each tool addresses a different gap in the existing survey practice. Their widespread adoption should help substantially to improve the effectiveness of data collection and the quality of the data resulting from it.

2 SCIENTIFIC AND TECHNICAL DESCRIPTION

2.1 SURVEY TECHNOLOGY REVIEW

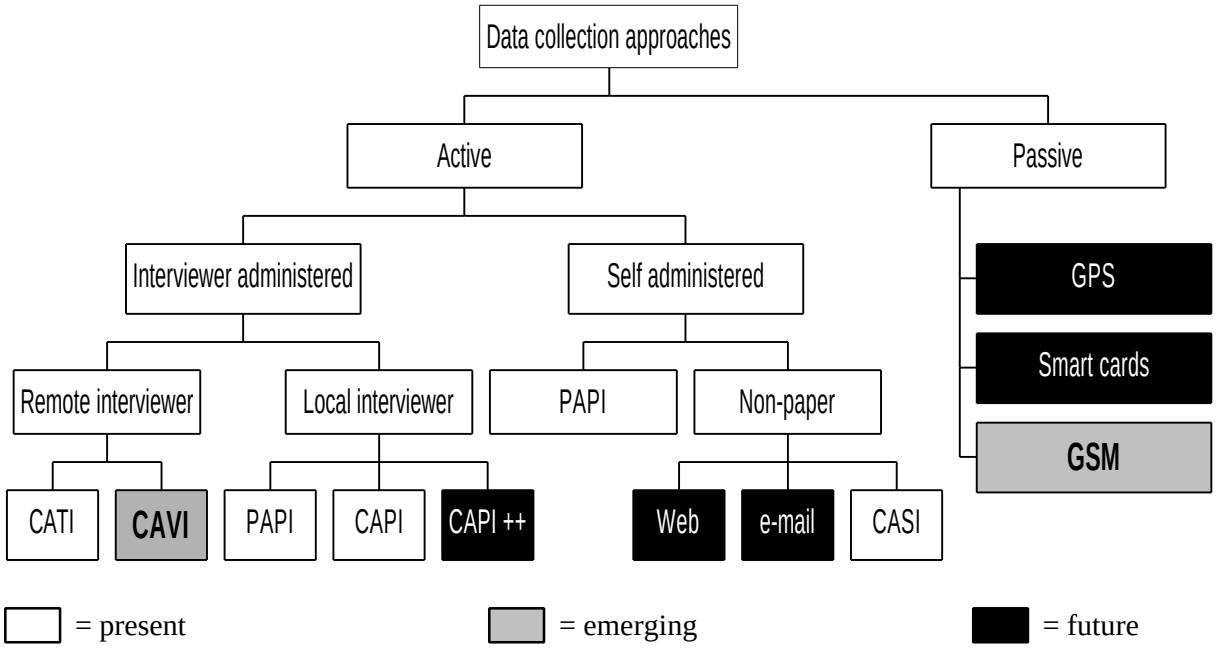
The environment of travel survey administration has been dramatically changing in recent years. A number of demographic and technical changes are profoundly altering the possible strategies for data collection. Among these changes are:

- Increasing computer literacy among the population at large
- A rapid rise in the market penetration of home personal computers (PCs)
- Falling costs of PCs and hand held computing devices
- Emergence of hand held devices combing telephony and computing
- Increasing adoption of standard interfaces across a broad range of applications and systems
- Rapidly rising penetration of Internet use
- The prospect of the Internet becoming accessible via digital TV services.

All this implies that alternatives are emerging to the three traditional forms of survey administration, which dominated the field in the mid 90s: self administered mail back surveys, face to face interviews and telephone surveys. However, the balance between the established techniques and the new technology alternatives varies widely between countries, depending on national traditions and the speed of take up of new opportunities, constraints and resource availability.

As the technology survey emphasised, the very rapid developments identified above were changing the possible approaches for both active and passive data collection techniques in the travel domain. Figure 1 classifies the different technologies used today according to the involvement of the respondent, the use of an interviewer and the employment of computers. Technologies are labelled *passive*, when data about behaviour can be collected without the direct involvement of the person observed with the exception - in the most extreme case - of gaining their permission for the observation. Typical examples are the tracing of observed persons through the location of their personal mobile phone, tracing long distance travel with the help of credit card records or the tracing of the exact route driven through a GPS based (global positioning system) location device integrated with a portable personal computer. Both mobile phones and GPS Systems have become dramatically cheaper as their market penetration has increased (See Figure 2, which shows the rapid progress across all 15 member states, but also the large variability still observable)

Figure 1 Range of data collection techniques available for Personal Travel Surveys

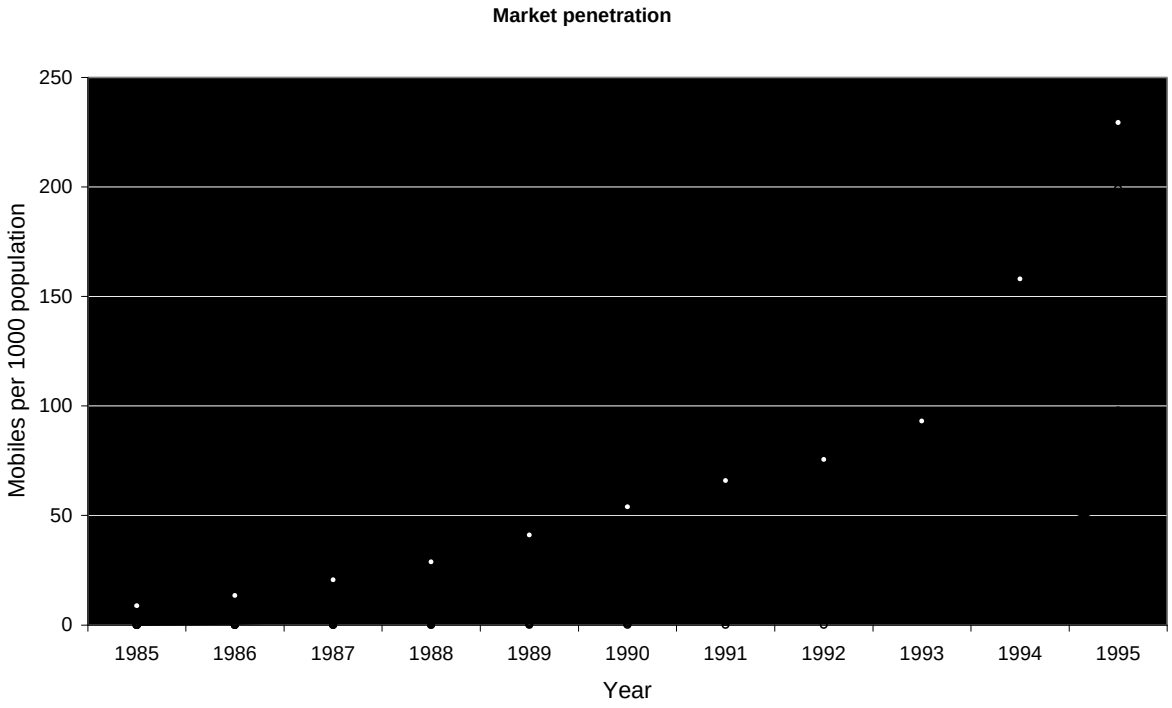


Source: Haubold, Jackson, Axhausen and Polak (1997), 5

Active survey techniques require the input of the respondent for most of the items studied. New styles of survey administration are becoming feasible through:

- Rapid market penetration of personal computers at home (Reducing the costs of survey administration and reducing potential sampling biases)
- Rapid standardisation of the user interface on the Microsoft Windows™ (Reducing the learning costs for the respondents and software development time for the survey author)
- Availability of serious hand held devices, some combining telephony and computing, other offering a scaled down version of Windows (WinCE) providing a familiar interface for mobile data collection without the weight penalty of a full portable computer
- Rapid expansion of the World Wide Web (www) in terms of hosts (See Figure 3), users and in terms of services offered through this medium. This allows delivery of surveys and collection of data in real time with the full backing of the familiar web interface and of web based data collection

Figure 2 EU member states: Market penetration of mobile telephony (1985-1995)



Based on the ITU World Telecommunication Indicators '95 (obtained from www.itu.int)

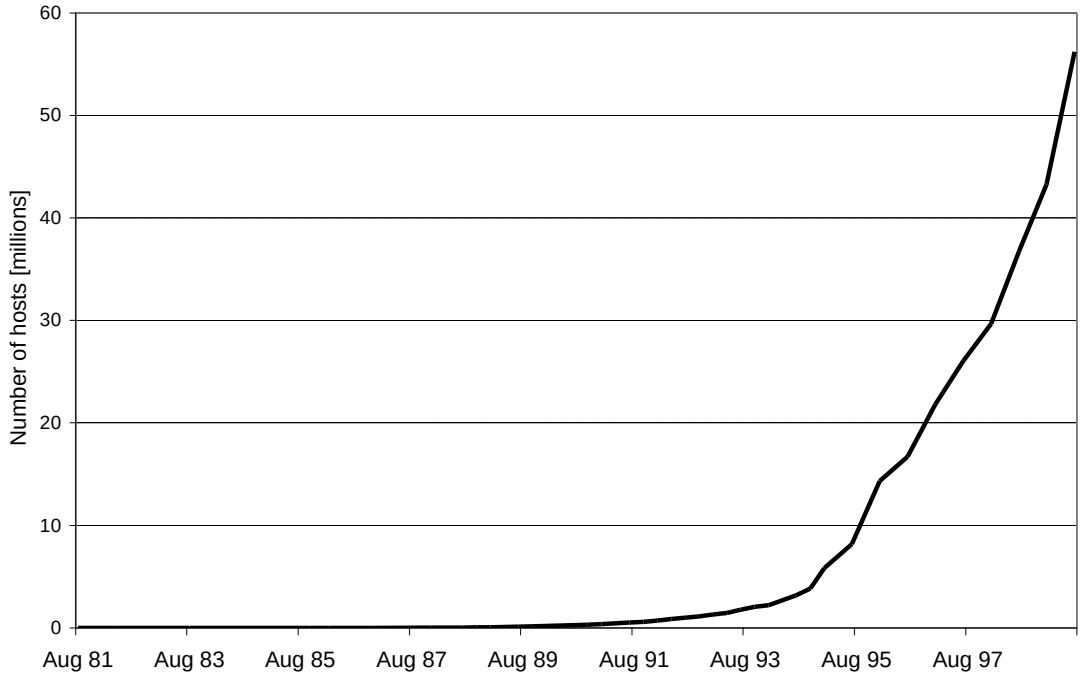


Figure 3 Internet domain survey host count

Source: http://nw.com/zone/hosts.gif (August 1999)

Based on earlier research (e.g. Kalfs and Saris, 1996) and an assessment of the options (See Haubold, Jackson, Axhausen and Polak, 1998) the project decided to focus on hand held organiser class computer running WinCE and on web based survey delivery. Some promising passive approaches, GPS supported surveys or tracing through financial records were felt too problematic at the time due to costs and accuracy problems (GPS) (See for example Ochieng and Polak, 1998 and Flavigny, 1998) or unwillingness of the intermediaries (credit card tracing) to co-operate .

2.2 HPC BASED TRAVEL DIARY APPLICATION

Travel diary surveys, in particular long distance travel surveys covering long periods of time (up to twelve weeks and even longer) (see for example Youssefzadeh and Axhausen, 1996) benefit, if the respondents can record their journeys either while they are in transit or as soon as possibly thereafter. Hand held devices that have become available in the recent years, are an elegant, reasonably lightweight platform, which can be carried by the respondents during the survey period to record their movements in *quasi real time*. The availability of WinCE, a scaled down version of Microsoft Windows for portable devices, reduces the learning costs of the respondents and increases the willingness of the respondents to participate. Against these advantages are in terms of the size and readability of the screen and the lack of the survey overview, available in paper surveys.

After a careful market survey of the available device types, their costs and their programming environments, the WinCE environment running on a Phillips Velo I was selected (Haubold and Axhausen, 1997), as it combined a familiar interface, the support of both Microsoft and a growing group of other developers with a well designed and reasonably priced machine (See Figure 4) in particular with regards to the size of the screen, its resolution, the usability of keyboard and stylus and finally its weight.

Figure 4 Phillips Velo I

Within the work package a successful travel diary application was written and tested, implementing the questions from one of the MEST surveys, being developed in parallel (Axhausen, 1996; Youssefzadeh and Axhausen, 1996; Axhausen and Youssefzadeh, 1997 and Youssefzadeh and Axhausen, 1998). A first version of the travel diary application was tested in house using a PC based screen emulation (Axhausen and Haubold, 1998) with a quota sampled group of employees of the Universität Innsbruck (by age, gender and computer literacy). A substantially revised version was field tested in Austria during the summer of 1998 (in German) and with smaller samples in Sweden (in Swedish) and in France (in French). The sample sizes were limited by the number of Phillips Velos available within the limited budget (three free machines for beta testers

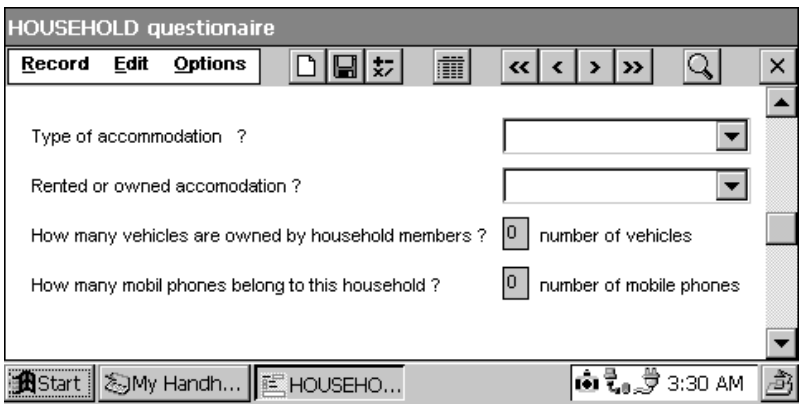
of the development environment and five purchased machines). The originally envisaged option of leasing machines could not be realised due to the lack of a supplier, despite earlier information about the availability of leased machines from Phillips. The final field tests were conducted in Portugal in the winter of 1998/99 (see below for more details) by a commercial market research firm, which had also been involved with the MEST survey work.
The travel diary application consists of three individual parts

- household questionnaire, including vehicle details
- person questionnaire
- travel questionnaire, which can be used at the stage or trip level[1].

The questions in the first two are on one long page, requiring the respondent to scroll up and down (See Figure 5) This format had been preferred in the first set of tests compared to a design based on a series of screens providing "drilling down" facilities to access the details. The use of drilling down designs was severely limited subsequent to these tests.

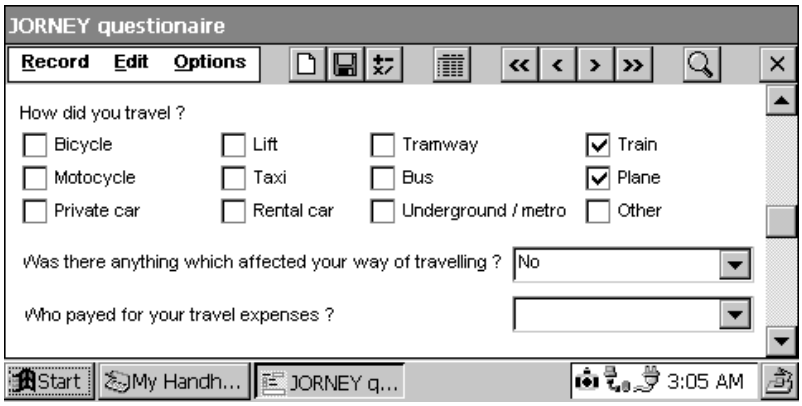
The travel questionnaire links the individual movement screens in the required sequence (See **Figure 6**). The inputs can be saved at any time, while allowing the respondents to resume their data entry later. The application provides the respondent with a summary of his/her journeys as a memory jogger. Initial difficulties with the beta version of the development environment, in particular with its memory management and size of executables, and limitations of WinCE 1.0, later replaced by WinCE 2.0, prohibited the implementation of certain desirable features, such as the provision of useful background information, such as a list of place names or maps.

Figure 5 Travel diary application: Extract of the household application (Final version)



[1] A Stage is a movement with one means of transport (mode) including the associated waiting times. A trip is a sequence of stages between two significant activities. A journey is a sequence of trips from the traveller's home ~~of the traveller~~ to a destination and back home.

Figure 6 Travel diary application: Extract from the movement screen (Final version)



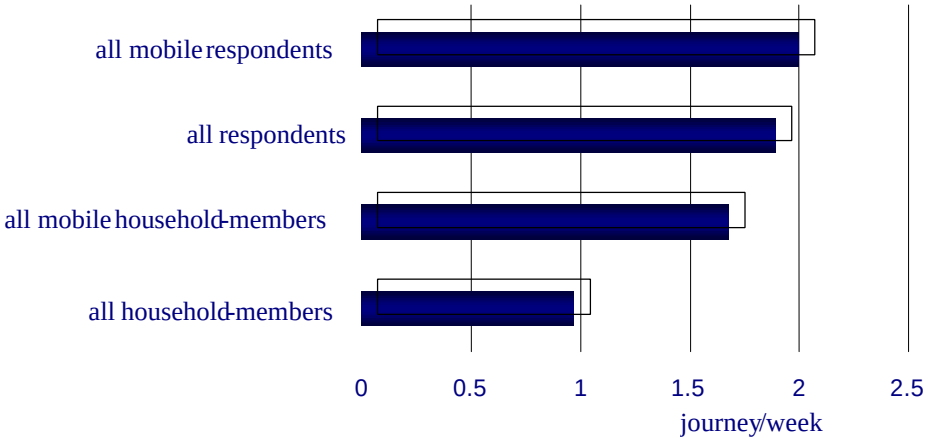
The field tests in Innsbruck involved eighteen respondents recruited by a national market research firm. The variables used in the recruitment were again age, gender and computing experience. The respondents were briefed by the project's research assistant and provided with a short user guide and a summary of the project. The research assistant was available on the phone for questions. The incentive (ATS 1000 = 73 Euro) was paid after completion of the survey. For the user guide and the briefing note see Haubold and Axhausen (1998b).

The respondents were asked to report all long distance journeys to destinations over 25km away from their home base for the next two weeks. This lower limit in contrast to the 100km minimum found in many other long distance and the MEST surveys was chosen to compensate for the shorter reporting period (two weeks compared to four and more in the MEST surveys), while trying to maintain the same probability that a respondent would report at least one qualifying journey. On average the respondents reported 1.9 journeys/week during the reporting period with more mobile respondents reporting 2.0 (see **Figure 7**). These numbers are consistent with expectations about the travel behaviour of Austrians (see Herry, Sammer, Schuster, Röschel and Russ, 1997). A more detailed discussion of the results can be found in Haubold and Axhausen (1998b).

The debriefing in Innsbruck indicated that respondents overall had coped well with the tool and had enjoyed the experience. However there had been problems in two areas:

- the initial introduction and the user guide could have been more extensive and more detailed for some respondents, in particular for the less computer literate among them
- the black and white screen of the first generation Phillips Velo I was not easily readable for a number of the respondents. They had to make special efforts to use the machine as a result.

Figure 7 Travel diary application field test: Distribution of journeys



The smaller tests in Sweden (4 persons) and France (4 persons)[2] came to the same conclusions.

The work undertaken in this workpackage has clearly demonstrated the feasibility of the HPC platform as a survey tool for long distance travel, but more generally for travel surveys. The largest remaining problems, screen readability and lack of computer literacy, will be solved by the arrival of low energy colour screens[3] and by the general increase of computer knowledge in the population.

The work package was not able to test all possible configuration of the survey interface, but it is felt that the solution implemented is a solid basis for further work.

The easy and general integration of both GSM phones and GPS receivers in HPC class devices at reasonable costs and the ensuing availability of the appropriate software development environments open new possibilities in the future. The same applies to 1) palmtop devices, the use of which is becoming more widespread as an alternatives to HPC's, 2) pagers and 3) paper diaries. These possibilities will be discussed in Section 0.

[2] The numbers were limited by the number of HPC's available and the remaining time between the field tests in Austria and the start of the test in Portugal.

[3] Already available for some of the second generation HPC devices.

2.3 A WEB DELIVERED TRAVEL DIARY SURVEY AND THE NECESSARY TOOLS

The World Wide Web (WWW) has been the computing success story of the last decade. An academic and research tool for file exchange and remote use of computing resources (ARPANET and later Internet) has been converted into a tool for interactive display, querying and submission of data, programs, text, pictures, movies and sound for the general public. The browsers, the tools giving the access to the Web, are sometimes thought of as alternatives to standard user interfaces on personal computers. See **Figure 3** for the explosive growth in hosts offering some or all of the WWW services. While the market penetration of the personal computers with an Internet connection/service at home is increasing quickly, it is not yet ubiquitous enough to make sampling an easy matter. However, the expected provision of Internet access as part of digital television services will make web access virtually universal in a relatively short period of time.

It was therefore decided, that the workpackage should concentrate on the development of a suitable architecture for the provision of web based surveys, through the development of a language to specify and create such surveys and the tools to interpret the language. Finally, the approach would be to field test the approach the outcome in the UK and Portugal.

The main design decisions for the overall architecture concerned a) the distribution of the tasks between the server, i.e. the survey researcher, and the client, i.e. the computer of the respondent and associated with it the choice of language for the implementation of the system and b) the level of platform independence desired. The second concern, in particular, ruled out Java in spite of its attractions, as the standard for this language was evolving too rapidly. It also ruled out the use of existing web based survey generators, as they are proprietary and not standardised. The solution implemented employs only stable standards. It is based on HTML 3.2 and PERL. A CGI (*Common Gateway Interface)* is constructed to create the customised HTML pages for the respondent, to interpret them and to store the results in a SQL relational database. The overall architecture of the *iTDA* (Internet Travel Diary Application) is shown in **Figure 8**. The server handles the bulk of the processing, while the client is used to display the survey and to capture the responses.

The CGI engine reads and creates the HTML pages, which are used to display the survey to the respondent. HTML (*Hypertext markup language*) allows the inclusion of non standard tags, which are ignored by the browser, but which can be detected and acted upon by the specialised CGI engine developed within TEST. The new tags implement the operations required for the survey:

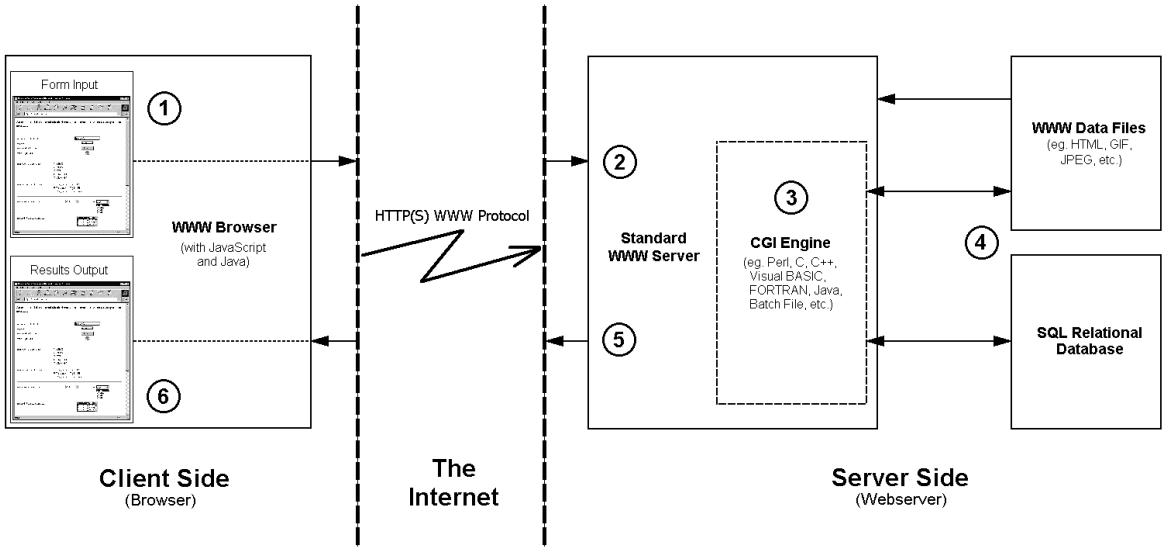
- Control of the flow: <if>, <elseif>, <else>
- Iteration: <loop>
- Movement through the form: <gotopage>
- Inserting variable text: <insert>
- Data base manipulation: <dbread>, <dbwrite>, <dbdelete>

There is also specialised treatment of existing standard tags (<form>, <textarea>, <select>, <input> and <option>). In addition the CGI engine makes it possible to give the respondent a unique identifier, a feature not normally implemented in HTML. This set of new tags and functions matches those available in commercial software. For more detail see Plaxton and Polak, 1999.

A third step of the development of the software tools is to allow the user to specify a survey in a standardised format. This in turn is translated into the appropriate scripts for the CGI engine. This removes the need for the non specialist user to learn the details of programming in this new survey specification language.

The tools developed were used to implement a web based survey covering the same topics as the MEST surveys. This survey was tested in a cognitive laboratory setting by an external market research firm in the summer of 1998.

Figure 8 iTDA: Software architecture



The survey was a retrospective survey of long distance travel with a reporting period of eight weeks. It was divided into four parts: a household form, a person form, a vehicle form and a movement form (see, for example, Figure 9 and Figure 10). In the trials, there was a mandatory tutorial and all questions in the survey had to be answered. Error checking routines were implemented in JavaScript[4]. The system generated summary pages for the household, persons, vehicles and journeys to provide additional feedback to the respondents. It also gave them the opportunity to make corrections (for an example, see Figure 11).

A sample of twenty five respondents was recruited to provide an even spread over gender, age group, frequency of Internet use and previous involvement in the various MEST surveys. The survey was undertaken either at the respondent's home, their office or the survey firm's offices. The system was simulated locally on a PC to avoid the

[4] System internal error checking routines are an extension planned

abnormally slow external telephone modem access to Imperial College. The respondents were asked to think aloud while performing the survey. Their comments were recorded and the times taken for the different questions measured. A detailed description of the survey is provided in Plaxton and Polak, 1999.

The tests were successful with - almost all - respondents completing the task and delivering clean and comprehensive data. The respondents liked the brief introductions, the summaries and the graphics used, but felt annoyed by the mandatory tutorial and the intrusive error checking. The help pages were essentially unused. The respondents disliked the inclusion of too many questions on a single page[5]. These problems can be addressed by changes in the design of the survey.

The results of field tests in Portugal, which were performed in the winter of 1998/99, are reported in Section 2.4.

The workpackage has successfully implemented and demonstrated an innovative tool for travel behaviour research and travel diary data collection. The extension of the HTML text description language, the backbone of the WWW, with a set of new tags allows the definition and conduct of web based surveys with the appropriate real time checking and identification of the respondents. The openness of the language permits the resolution of the remaining design issues. The *i*TDA points the way into a future, when universal access to the internet will be achieved and web based surveys can start to replace not only competing paper and pencil surveys but also telephone surveys for most applications.

[5] The reverse was true in the pre-tests with the HPC, where the respondents preferred this format.

Figure 9 *i*TDA example screen: Household form

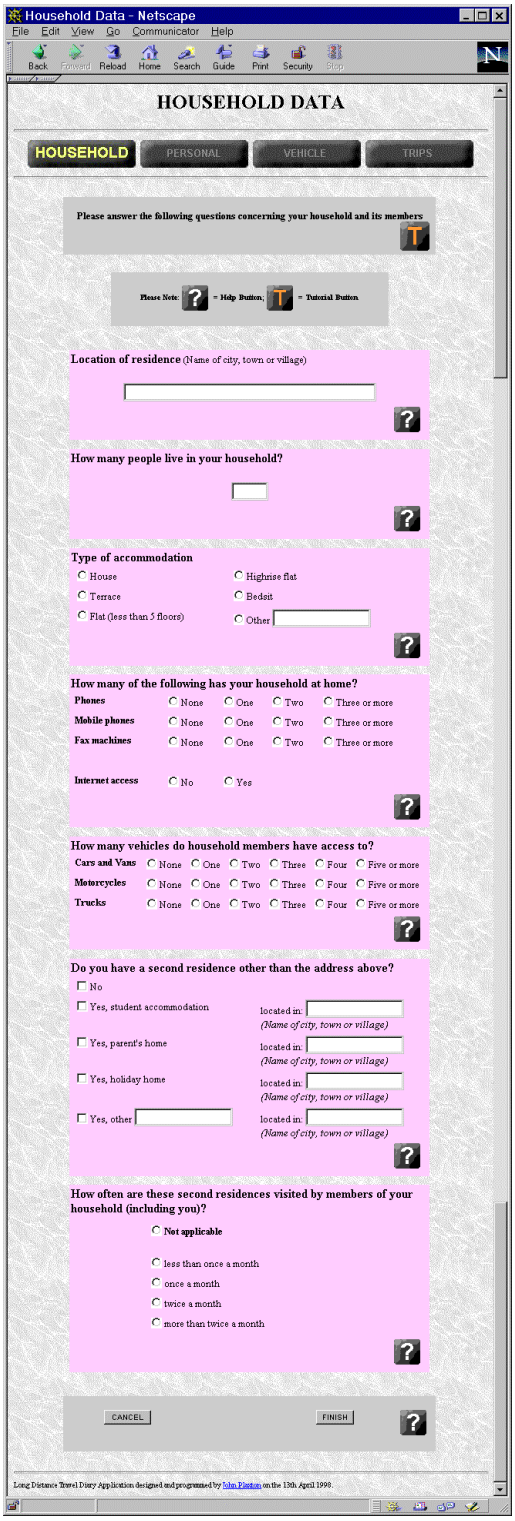


Figure 10 *i*TDA example screen: Movement form

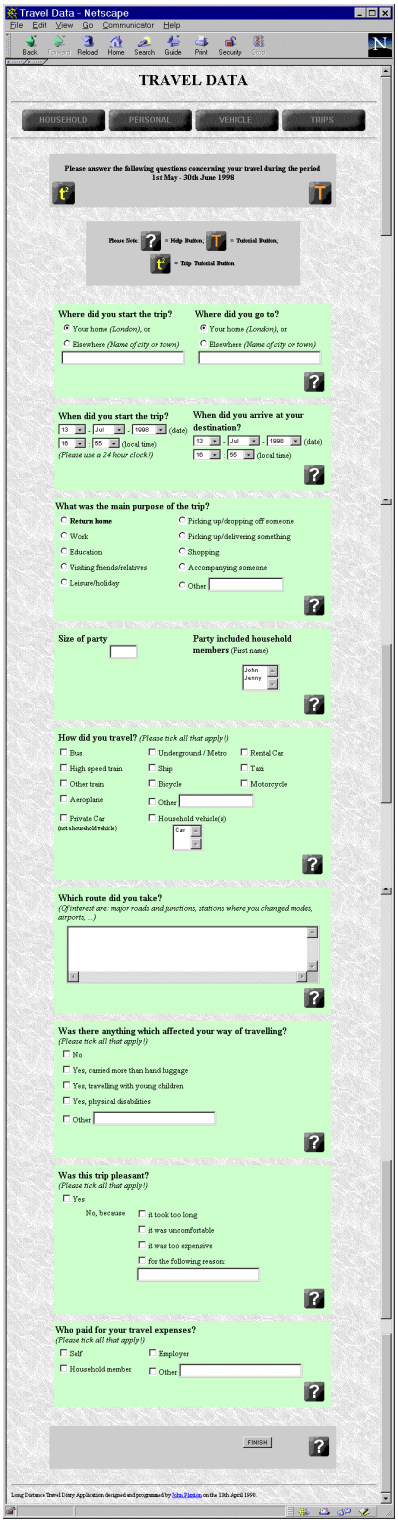

Travel Data - Netscape
File Edit View Go Communicator Help
TRAVEL DATA
HOUSEHOLD
PERSONAL
VEHICLE
TRIPS
Please answer the following questions concerning your travel during the period 1st May - 30th June 1998
Where did you start the trip?
Your home (London), or
Elsewhere (Name of city or town)
Where did you go to?
Your home (London), or
Elsewhere (Name of city or town)
When did you start the trip?
(date)
(local time)
(Please use a 24 hour clock!)
When did you arrive at your destination?
(date)
(local time)
What was the main purpose of the trip?
Return home
Work
Education
Visiting friends/relatives
Leisure/holiday
Picking up/dropping off someone
Picking up/delivering something
Shopping
Accompanying someone
Other
Size of party
Party included household members (First name)
John
Jenny
How did you travel? (Please tick all that apply!)
Bus
Underground / Metro
Rental Car
High speed train
Ship
Taxi
Other train
Bicycle
Motorcycle
Aeroplane
Other
Private Car (not household vehicle)
Household vehicle(s)
Car
Which route did you take?
(Of interest are: major roads and junctions, stations where you changed modes, airports, ...)
Was there anything which affected your way of travelling?
(Please tick all that apply!)
No
Yes, carried more than hand luggage
Yes, travelling with young children
Yes, physical disabilities
Other
Was this trip pleasant?
(Please tick all that apply!)
Yes
No, because
it took too long
it was uncomfortable
it was too expensive
for the following reason:
Who paid for your travel expenses?
(Please tick all that apply!)
Self
Employer
Household member
Other
FINISH

Figure 11 *i*TDA example summary screen: Journey

2.4 FIELD TESTS OF THE SURVEY APPROACHES AND EXTERNAL EVALUATION OF THE TOOLS

The survey approaches, the HPC TDA and the web based *i*TDA, had to be tested by someone not directly involved in their development to provide an independent assessment. This held equally for the other tools. While the survey instruments were ready early enough to perform the full tests, the other tools were delivered too late for extensive testing. This section will therefore concentrate on the debriefing interviews with the respondents involved in the tests (Elvas, 1999).

The field and laboratory tests were performed in Lisbon and conducted for TEST by the survey firm METRIS under the supervision of TEST's Portuguese partner. The test were divided in two phases:

- a prospective phase, in which the respondents used the tools (paper, HPC and web based iTDA) for a two week period, as in the usual protocol. They were debriefed using a detailed questionnaire (November-December 1998)
- a retrospective phase, in which the respondents used the tools (HPC and iTDA) in an interview situation. They were debriefed using a detailed questionnaire (December 1998 -January 1999).

The debriefing results for the first most critical phase are based on three quota samples for the paper survey, the HPC and the web based *i*TDA and on an additional sample of university students using the web based approach. The quota samples are small, as the limited number of HPC devices and the remaining time allowed only three two week reporting periods. For the paper implementation, it was possible to match the quota requirements of gender and age but men are over represented in the HPC sample, as women mostly refused the device. It was not possible to find Internet users in the age group 36 years and older, willing to participate in the tests.

The debriefing surveys offered the respondents both open and closed questions about their experience with the approaches. In general the respondents perceived all approaches as being in the range from "easy" to "neither easy, nor difficult" to use in all of their aspects. There were specific criticisms of the tools as currently implemented (HPC: small screen, difficult to read, lack of automatic record saving; *i*TDA: too many explanations, lack of short cuts through the survey and the slowness of the internet, although this was related to the internet and not the tool itself). The overall comment was positive. The time used with the instruments was below 30 min with some exceptions, in particular for the web users. While the large majority of users would have been willing to answer a paper survey, when asked about their general preference there was an interesting shift of HPC users to the Internet. While about 60% of the paper and the *i*TDA users did not want to change, 40% of the HPC users wanted to switch to the net and only 30% wanted to stay with the HPC.

The conclusion from these tests and the reactions of the participants is, that all three approaches are acceptable, meaning that the HPC and, in particular, the *i*TDA, although technologically challenging, have already found acceptance. Clearly, this conclusion needs to be verified in larger scale exercises with a more representative set of users but it is very promising given the novelty of these technologies for most citizens of Portugal and Europe.

2.5 A GIS INTEGRATED TOOL FOR GEOCODING OF PLACE NAMES AND RELATED DATABASES

Travel diary surveys provide information about the number and spatial distribution of the movements of a population. In addition the modal and temporal characteristics are provided. While the analysis of travel choices is one of the main uses of travel diary data, another is the derivation of origin-destination (OD) matrices from the data. For both purposes, the exact geographical location of the trip characteristics (origin, destination and route) is crucial. This is self evident for OD-matrices but it also applies to choice models, as these depend on the provision of network derived information for the description of the non-chosen alternatives, which in turn requires exact locations. The purpose of the TEST work in this area was to provide an improved geocoding tool for those conducting long distance travel surveys.

The capture of locations, address or other suitable descriptions is notoriously error prone, for example because of:

- A substantial share of complete item non-onse, because of
  - Lack of knowledge of the address
  - Lack of ways to describe the location alternatively
  - Poor recall of the place
  - Unwillingness to provide the address
- Partial responses, e.g. incomplete addresses (city, but no street; region, but no municipality etc.)
- Alternative responses, e.g. name of shop, colloquial name of place, name of crossroads, name of facility etc.
- Erroneous responses, e.g. non-existing street numbers, non-existing streets, misspelled names of street or municipalities etc.

When coded at the municipality level, enough for most uses of the data, there are very few item non-responses (about 10 out of 100,000 trips for daily mobility in the French NPTS). Place names are a good stimulus for remembering trips. The coding of this information to the appropriate scheme (traffic zones, NUTS codes, exact co-ordinates etc.) is therefore time consuming and, if performed manually, in itself prone to clerical error. Out of all the possibilities offered by geographical information systems (GIS) to enrich travel diary data, MEST decided to concentrate on this aspect .

The work had four themes, which interacted with each other:

- Analysis of the frequency distributions of places and populations to identify suitable search strategies
- Collation of existing lists of place names
- Development of a support tool for semi-automatic geocoding of place names and NUTS5 assignment
- Development of a GIS tool to calculate shortest paths between places

In the first phase of the work, the data from the French National Passenger Transportation Survey (NPTS) were used, while later work also employed MEST survey data .

The analysis of the interaction between a location and its population showed clearly that the bulk of the population lived in a relatively small number of the municipalities but, more importantly, that these municipalities are also the destination for the majority of journeys. At least this is true in the case of the French NPTS and probably for other countries as well. This implies that most of problems will be associated with smaller locations, visited by few people both in the surveys and generally (See Figure 12). Within France, such destinations are regionally clustered, reflecting past settlement patterns and leisure developments (Figure 13). For long distance travel surveys, these results suggest that it would be useful to add further items to questions about geographical places to help to locate these smaller ones more precisely. Possibilities are: name of region, name of nearest big town or city, duration of travel or estimated trip distances, population size.

The interaction also implies that place name lists from earlier surveys can be reused successfully. Heuristically the point where the ranked distribution of location by population distribution in the old survey no longer coincides with the known distribution provided by official statistics can be taken as an estimate of how many place names in any new survey can be geocoded with the old list of place names.

The existence of comprehensive and accurate databases of place names is a precondition for any automatic or semi-automatic geocoding. TEST explored the quality of various databases. An analysis of the list of place names from the French NPTS and the NUTS place name database showed that both have shortcomings. The NUTS database in particular had the major drawback that it only covers administrative units but not smaller units or commercial sites, such as resorts or hotels. The data base was expanded using on-line sources, such as:

- National geographical institutes files for the position of NUTS5 centres,
- NIMA (US National Imagery and Mapping Agency) *Geonet* server,
- "Grid Arendal" project for Scandinavia,
- A commercial data base from the British map publisher, Bartholomew,
- Topographic or road maps, in order to create points in the GIS environment.

Bartholomew's "Euromap" database was particularly valuable for Scandinavia and the Netherlands, where NUTS5 units are quite coarse (e.g. 288 NUTS5 units in Sweden). It was also useful for UK where the names for administrative wards may not be much used in daily life and are therefore not familiar to travel survey respondents. The number of wards decreases from 14,000 to 9,000 different names, when the extensions "West, North, East, South" are suppressed. In France, Euromap supplied many ski or sea resorts which are not municipalities. Bartholomew's set of place names is more

evenly distributed in space than NUTS56 zones in France, however, it contains only somewhat less than half of the place names in total. Bartholomew's set of French names would only have allowed about two thirds of the places appearing in the French NPTS long distance survey to be coded.

Based on this experience, a support tool for the semi-automatic geocoding of place names and the extraction of routes between two places visited in succession was developed using Visual Basic within the MapInfo System (Hubert, Flavigny and Madre, 1999). The place name database has at its core an indexed list of entries with:

- Name
- Reference spelling/name (in case of multiple spellings and variants)
- Longitude/Latitude (as simple and unambiguous location)
- Country
- NUTS5 (finest resolution of administrative system) (allow regional cross reference)
- Source (of data)
- Size (urban or rural)
- Nearest large town

The reference spellings have to be associated with the variants appearing in the normal operation of the surveys and which may vary over time.

The name parser starts from a list of names. It then:

- standardises the spellings by removing diacritics,
- translates national spellings, e.g. in the case of abbreviations,
- shortens the name to the 14 significant letters and suppressing diacritic marks,
- shortens or suppresses very common locutions such as "Saint", "Santo" or "Stadt".

The result is then compared with the central indexed list. Using any additionally available information (region or distance from last place), a list of possible places is suggested to the user. The user then associates the candidate place name from the survey with one of entries of the central database. The user may also select the reference spelling, if this is required. The cross reference selected is then added to the travel diary entry for future processing of the selected entry (See **Figure 14**).

Given the details of selected entries, the GIS calculates the shortest route between two successive places visited based on criteria defined by the user (shortest distance, shortest time or others). The routes can be overlaid to provide an image of the transport

6 NUTS5 varies from 0.06 names per 10 km² in Sweden to 7 per 10km² in Switzerland and 6 in France

system use. The relevant information can be extracted for other purposes, such as choice modelling (See Figure 15)

TEST has successfully expanded the state of the art in the automatic geocoding of place names. The analysis of the existing place name databases and of existing travel diary data has highlighted the need for additional information over and above the address, if geocoding is to be successful. The newly implemented GIS based system translates these results into a user friendly tool for geocoding at the European level given the right place name data base[7].

The project has highlighted the need for further work on place names as existing databases have substantial problems in their accuracy and comprehensiveness. This applies particularly to leisure related places (hotel complexes, sports facilities etc.) and small settlements below the NUTS5 level and is a problem which is especially acute in countries with an administrative structure, the names of which are divorced from common usage, such as the UK.

[7] This tool will be used by a survey firm with specialized GIS capabilities in order to geocode a pilot survey on freight transport for MYTHSTIC European project and border surveys.

Figure 12 Distribution of population by location

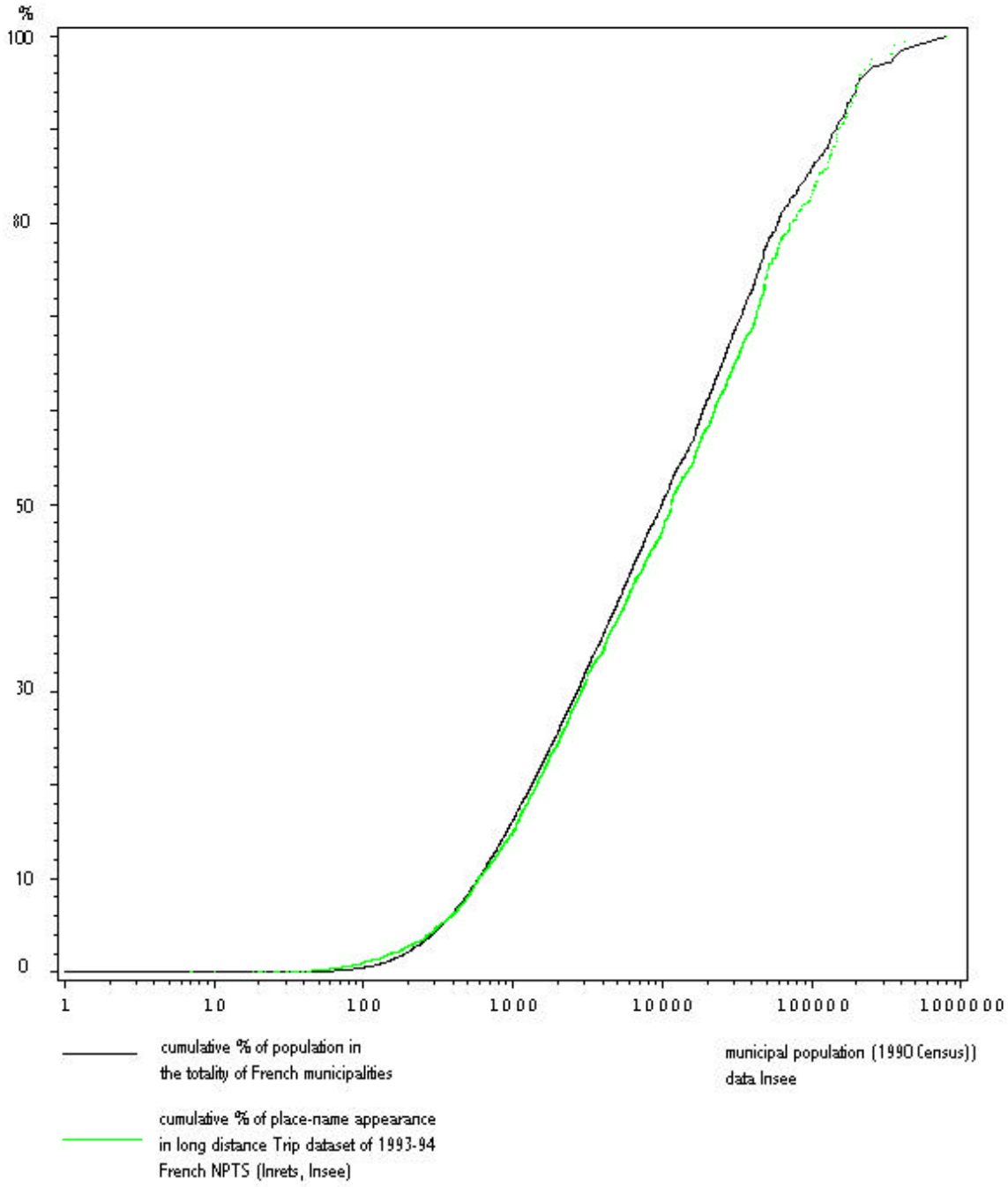
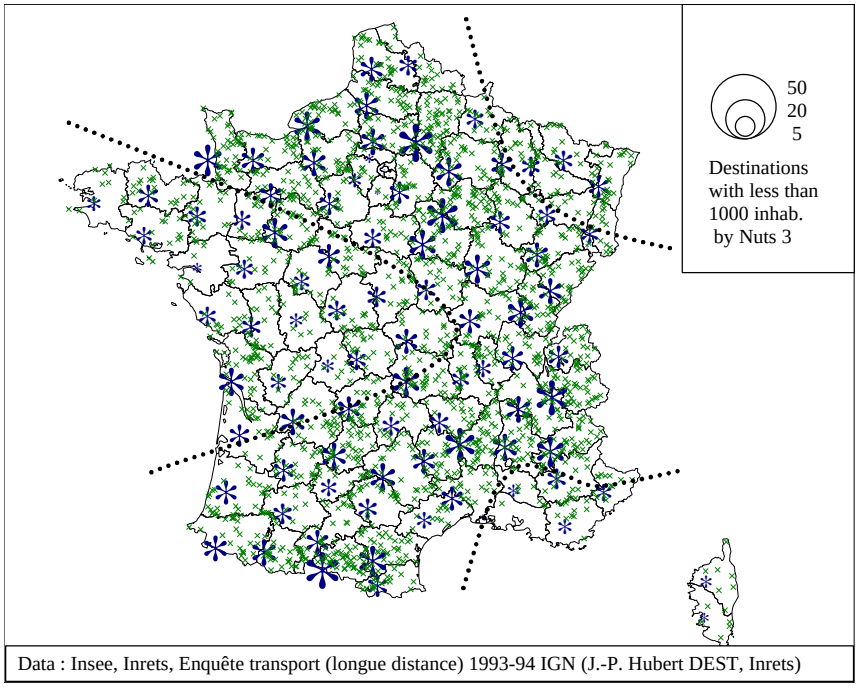


Figure 13 Distribution of difficult to identify places in France



This map shows a rather regular distribution where tourist regions stands out (Alps, Jura, Auvergne, Pyrénées, Quotentin) Champagne and Morvan, can be noticed as well.
A triangle Mont-Saint-Michel / Roanne / Bordeaux catches less destinations.
Alsace-Lorraine and Provence show also few small destinations

Spatial distribution of municipalities of less than 1 000 inhabitants, and total by *département* (nuts3)

Figure 14 Selection of correct place name

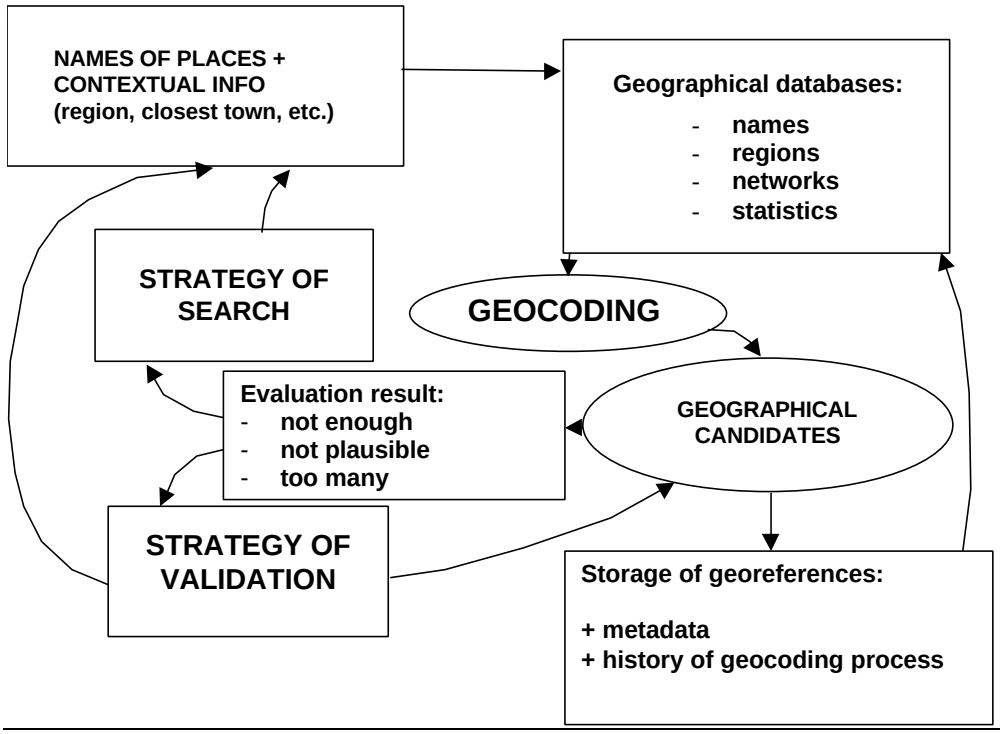
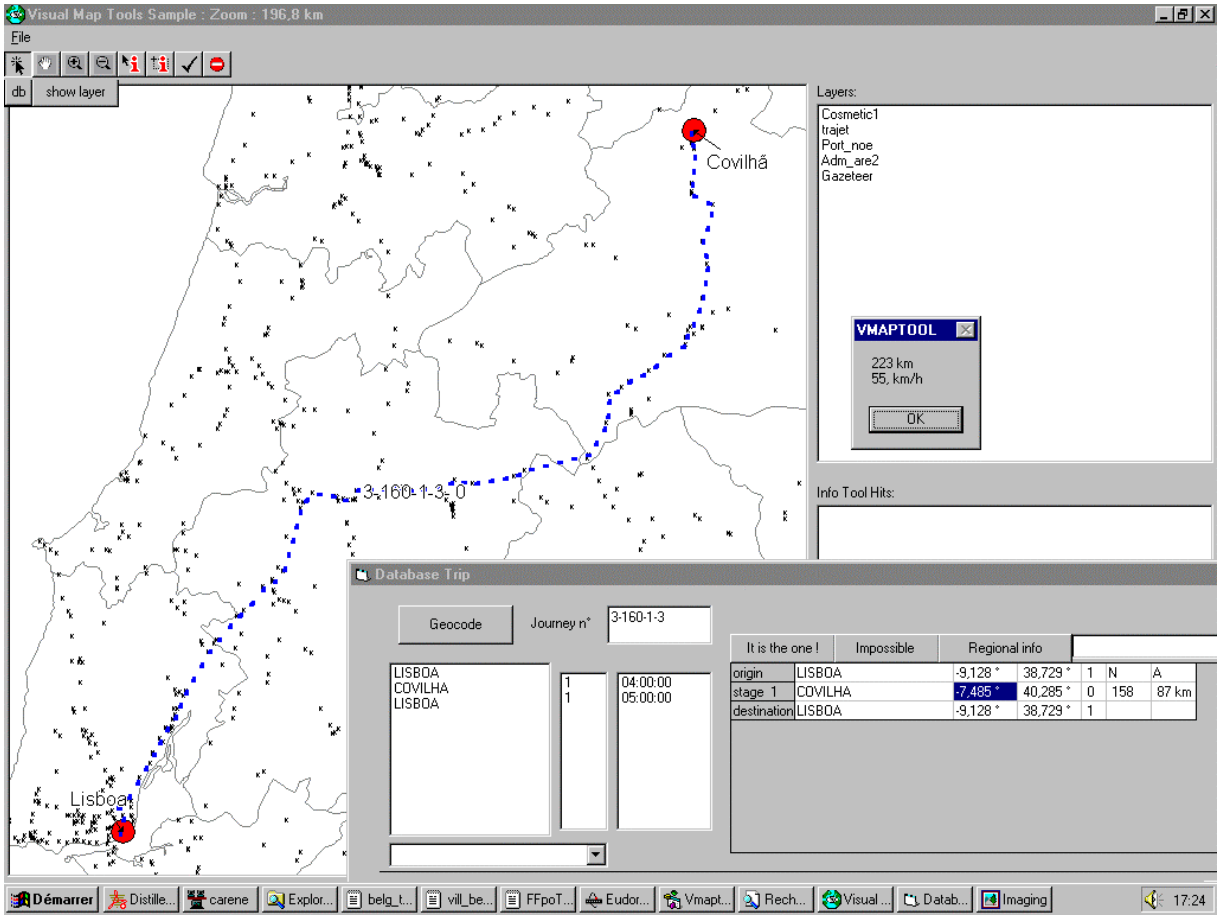


Figure 15 Calculation of routes within the GIS



2.6 APPROACHES TO DATA CORRECTION AND IMPUTATION: AN AI-BASED IMPUTATION SET OF TOOLS

Even if special care is given to the survey process, the resulting data set is likely to contain both missing and erroneous data. This applies equally to the case of travel diary surveys, but the sheer scale of these surveys with respect to sample size, number of records generated and complexity of the interactions between the different records, makes the identification and correction of erroneous data and the imputation of missing data burdensome. The TEST target here was to develop tools to support this process and to test promising new approaches for such a task.

In this area, TEST's aim was the development of an open architecture for data correction and imputation using free software tools, the work was based on an extensive review of the literature. The core of the software architecture is a parser to interpret commands specified by the user in a new language, developed to control the execution of the relevant code (See **Figure 16**). The provision of a correction audit trail is an important element for controlling the quality of the imputation process and ultimately the quality of the survey itself., Very often corrections and imputations are performed in an ad hoc manner and either poorly documented or not documented at all.

The user specifies the data and the corrections to be made in the following series of files (Figure 17):

- *start.spc* gives the names of all the other files

- *standard.spc* names the standard variables in each file, associates them with the standard file types and specifies all valid correction methods for each variable based on earlier analysis. For a subset of variables given standardised names logical error correction routines are implemented, e.g. number of household members in the household file must be consistent with the number of household members in the person file (minus persons below the relevant age threshold for the survey)

  The following imputation methods have been implemented in the software and can be selected:

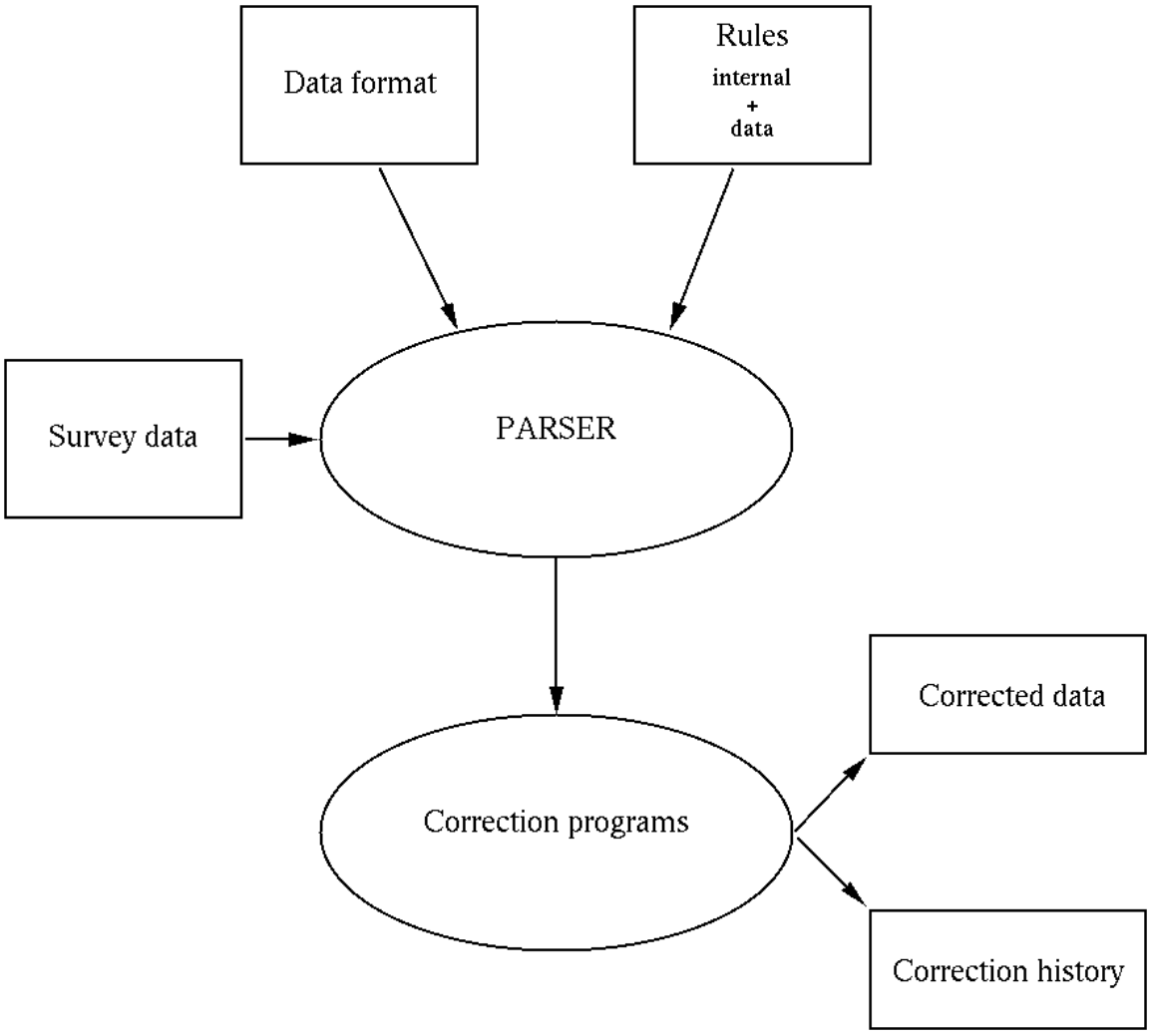
  - Mean imputation
  - Stratified means
  - Hot deck imputation
  - Regression imputation
  - Random imputation

- *problem.spc* specifies the order of execution of the imputation steps, selects the chosen imputation method and gives access to a basic set of descriptive statistics

- *fixed.spc* describes the variables which have either fixed values or are constrained to a fixed number, such as percentages which must add to 100%.

- *bounded.spc* describes the variables which have known upper and lower bounds.

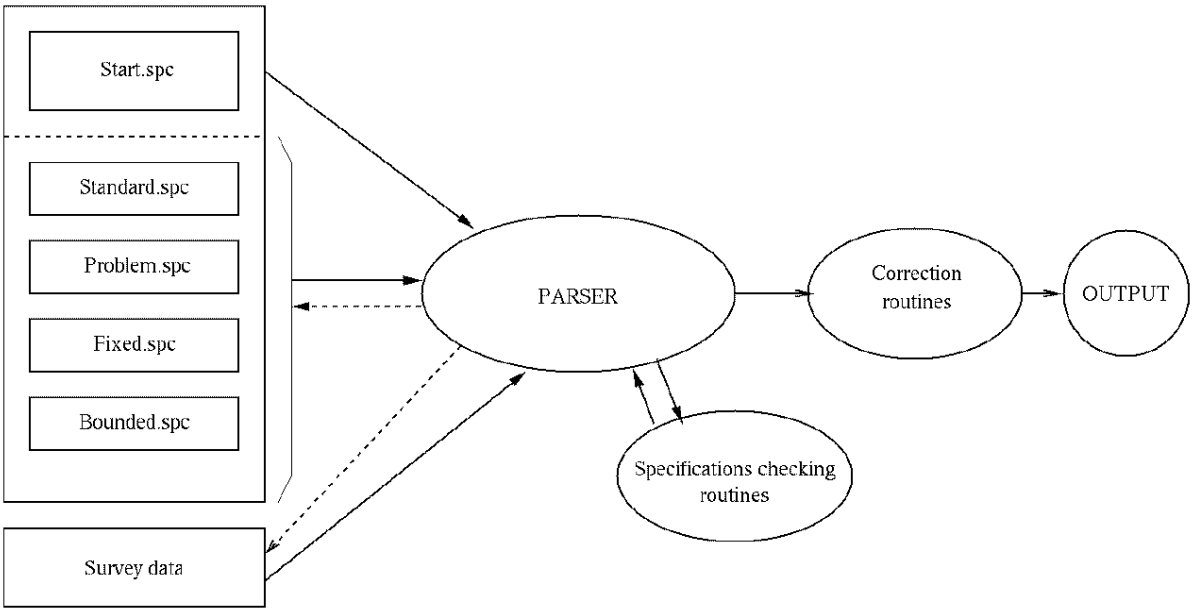
Figure 16 TEST imputation tool: Software architecture



The software was extensively tested with both MEST derived data and data from the Austrian National Travel Survey. It is currently entering more general use in the processing of the Belgian National Travel survey and other work of the consortium members. It will be made freely available to potential users.

The software makes standard tools for imputation fully available to the travel behaviour community. There will be enormous benefit from the availability of a standardised and documented approach to the imputation and correction of travel data. There are plans to integrate within the software the *Expectation Minimisation* approach developed in MEST.

Figure 17 TEST imputation tool: Specification files



2.7 APPROACHES TO DATA CORRECTION AND IMPUTATION: NEURAL-NET BASED IMPUTATION

The second element of TEST's work in the area of correction and imputation had the aim of exploring the usefulness of alternative neural network architectures. This was essentially uncharted territory and the intention was to explore new directions beyond the primarily application oriented AI element described in the previous section.

Neural networks constitute a wide class of non-linear models, which employ the notion of incremental learning to associate input variables with output variables. The review of the literature highlighted three types of neural nets, which were chosen for study within TEST:

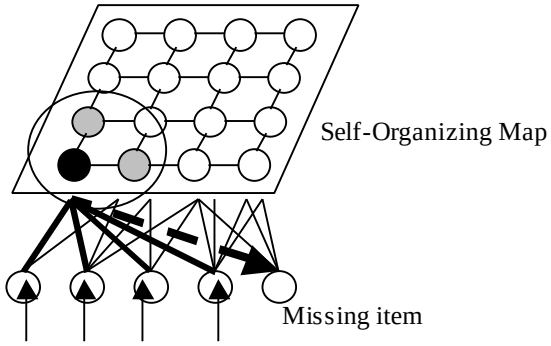
- *Self organised (Kohonen) map (SOM)* which are primarily tools to cluster observations by mapping a multi-dimensional input space to a two dimensional map of discrete nodes (points). The use of such maps as imputation tools is an innovative step (Figure 18).

- *Multi-layered perceptron (MLP)*, the most widely used type of neural net. In the context of the TEST exercise it was used as a benchmark.

- *Recurrent multi-layered neural network (RML)*, a new type of neural network in which learning takes place in both directions: from the inputs to the output, but also from the outputs to the inputs. These tests are also the first reported in the literature.

The SOM, the related visualisation tools and the other neural nets were implemented on a Windows PC with C++. This makes the code generally available (See Fessant and Midenet, 1999 for details).

Figure 18 Structure of a SOM for imputation



The SOM was tested using both a large scale data set (the long distance journey file of the most recent French NPTS) and a smaller scale data set (the vehicle description files from one of the MEST surveys) to span the range of likely sample sizes found in

practice. In both cases the complete set of survey records was identified and divided into a learning sample and a test sample with artificial missing items[8]. Both the long distance journey files and vehicle description files contain continuous variables and categorical variables.

The SOM does not require the division of the available variables into input and output variables, but it is still necessary to calibrate it properly to obtain good results. The best SOM obtained was used to compare the quality of its imputation with the results obtainable from three alternative approaches. In the case of the NPTS data with an optimised MLP, a *hot deck* imputation and the RML. For the MEST data, it was compared with the best results obtained with the Namur suite of imputation tools. The error is measured as the mean squared error for continuous variables and the error percentage for categorical variables.

The main results are summarised in. These show the performance of the approaches, measured against the original correct data set. More detail is available in Fessant and Midenet (1999). The SOM performance is comparable to that of competing classical approaches, but does so with one map (model). The other neural network approaches require the development of a different neural net/model for each variable imputed. The outcome of the tests of the RML is disappointing and the method as tested is not competitive with either the SOM or classical approaches.

In summary, the TEST work has shown that the SOM is an attractive single step and consistent approach to the imputation of missing values. While it did not perform any better than the classic alternatives in the cases studied, it performs equally well in terms of the correction achieved. However, it has two interesting by-products

- a visualisation of the underlying structure of the relationships between the variables and
- an indication of suspect observations requiring further investigation.

One drawback of the SOM approach is that the development of the network is far from automatic and requires skilled inputs to achieve a satisfactory result. Further work to explore ways of all improving this situation would seem worthwhile.

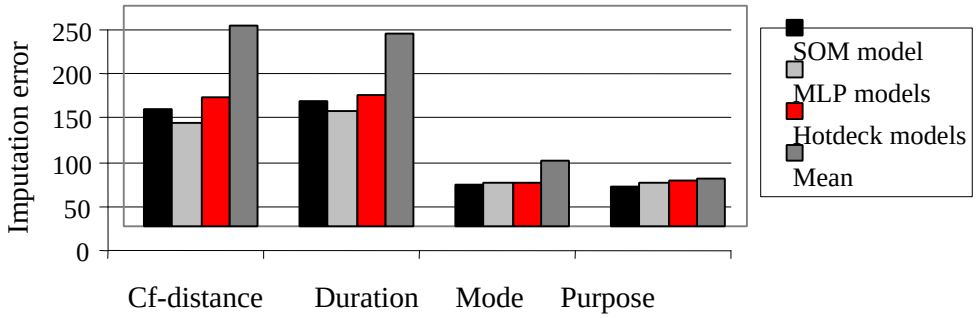
Although the RML did not produce promising results with the test data set, its inherent advantage of imputing during learning would warrant further study.

It is hoped that the innovative work performed here can be integrated into daily practice quickly through inclusion with the FUNDP suite of imputation tools.

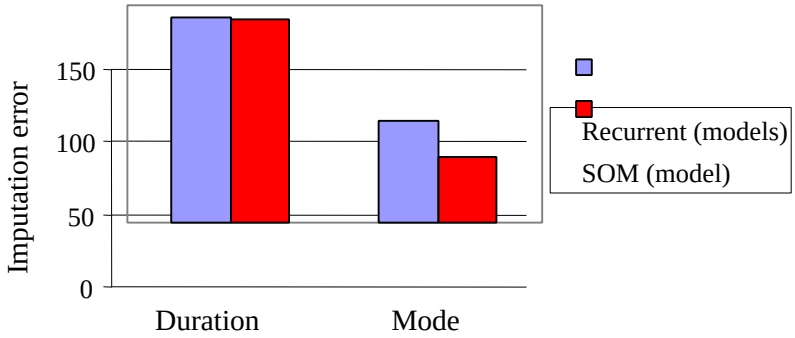
[8] Tests with the raw data file and with the remaining incomplete records (French NPTS) were consistent with those reported in this summary ~~here~~. For details see Fessant and Midenet (1999).

Figure 19 Comparison of the SOM approach

SOM against standard approaches



French NPTS against RML



MEST vehicle file against standard suite (FUNDP)

Variables tested	Selected SOM	FUNDP's (using *different* classical methods)	
Catalytic converter : percentage)	**22%**	39%	(error
Year of production : squared error)	**4.5**	4.8	(mean
Current odometer reading: squared error)	**46300**	49050	(mean
Annual mileage: squared error)	11150	**8080**	(mean

2.8 A WEB-BASED INTERFACE TO ON-LINE TABULATION OF TRAVEL DIARY DATA

The explosion in the use of the WWW and the rapid development of software, which allows the gap between the user and more sophisticated analysis software to be bridged, makes the publication of data, or more precisely the provision of supervised access to a limited set of analyses of the data, in the medium an attractive proposition for the holders of data. This applies particularly for data collected in the public sector, but there are also attractions for private sector data holders. The provision of such an access has a number of benefits for both the holder and the user of the data:

User benefits:

- No need to specify all tabulation (models) of interest beforehand
- Near real time access to the data
- No need to worry about the proper weighting and handling of the data
- Less need for expensive specialist support for easy tabulations

Owner benefits:

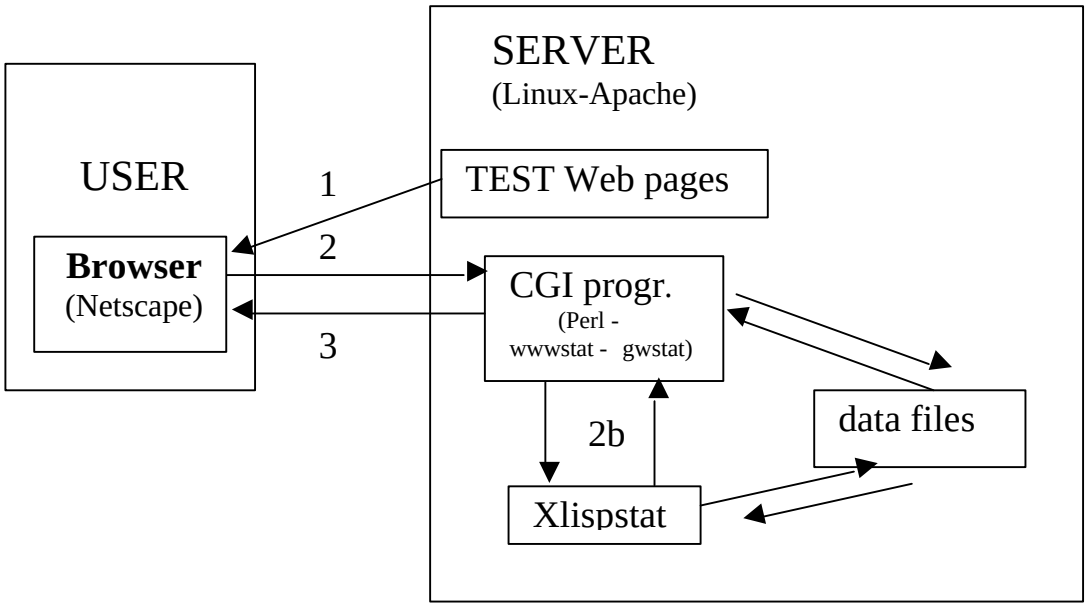
- The proper documentation of the survey and the data required for publication in this media may indirectly help to improve the quality of the data in the collection and processing collection phases
- Public access may help to justify the cost of collecting the data
- Ensures that only properly weighted tables are available in this medium
- Helps to ensures that tables with entries of doubtful significance are not published
- Implementation of data protection requirements possible, while keeping the data available
- Possibility to concentrate the resources of the data owner, in particular staff time, on more complex and difficult modelling and tabulation tasks.

TEST's target in this area was to demonstrate the feasibility of this approach, with the use of free software only, and to implement as functionally rich a site as feasible within the constraints of the project.

Designing the site on the basis of free high quality software ensured the portability of the results to other sites and users, always assuming that a WWW browser is available on the client side. The tools are:

- Linux Apache (operating system),
- CGI code written in Perl, for which a free compiler exists,
- wwwstat and gwstat for monitoring of the site and
- xlispstat as the statistical engine (See Figure 20).

Figure 20 General design of the TEST web site



The TEST web site provides the primary access to the different functions of the site and the various data sets implemented (a set of MEST surveys). The following functions have been successfully implemented in the last 18 months:

- *Welcome and general information*
- *Registration* of users to control access, if required, and to monitor individual usage
- *Selection of data sets* from the available data sets.
- *Selection of variables* from the available variables in the selected dataset
- *Construction and storage of classes* for both continuous and discrete variables
- *Calculation of one dimensional statistics* for the selected variable
- *Tabulation and cross tabulation* of the selected variables, properly weighted to the relevant control totals
- *Graphical display of one dimensional tabulations* of the selected variable
- *Storage of intermediate results* between sessions
- Specialised help pages for any of the functions

Examples of the user interface are shown in **Figure 21**, **Figure 22** and **Figure 23**.

This functionality compares favourably with some web sites with a larger budget than TEST's. Adding to this success, MEST has implemented a general language to describe data sets and their structure to make the tools more generally applicable. It also overcomes the currently strictly defined structure suitable for travel diary data sets with typically the following files: stage, trip, activity, journey, person, household and vehicle. The quality is also shown by a continuous growth in traffic at the site from people outside TEST (See Reginster, 1998 for details).

This workpackage has demonstrated the usefulness of the provision of WWW based access to real time user driven tabulation and graphing. The take up of this technology for the ongoing Belgian national travel survey is a good sign that the message has been accepted by the user community.

Figure 21 TEST web site: Welcome page

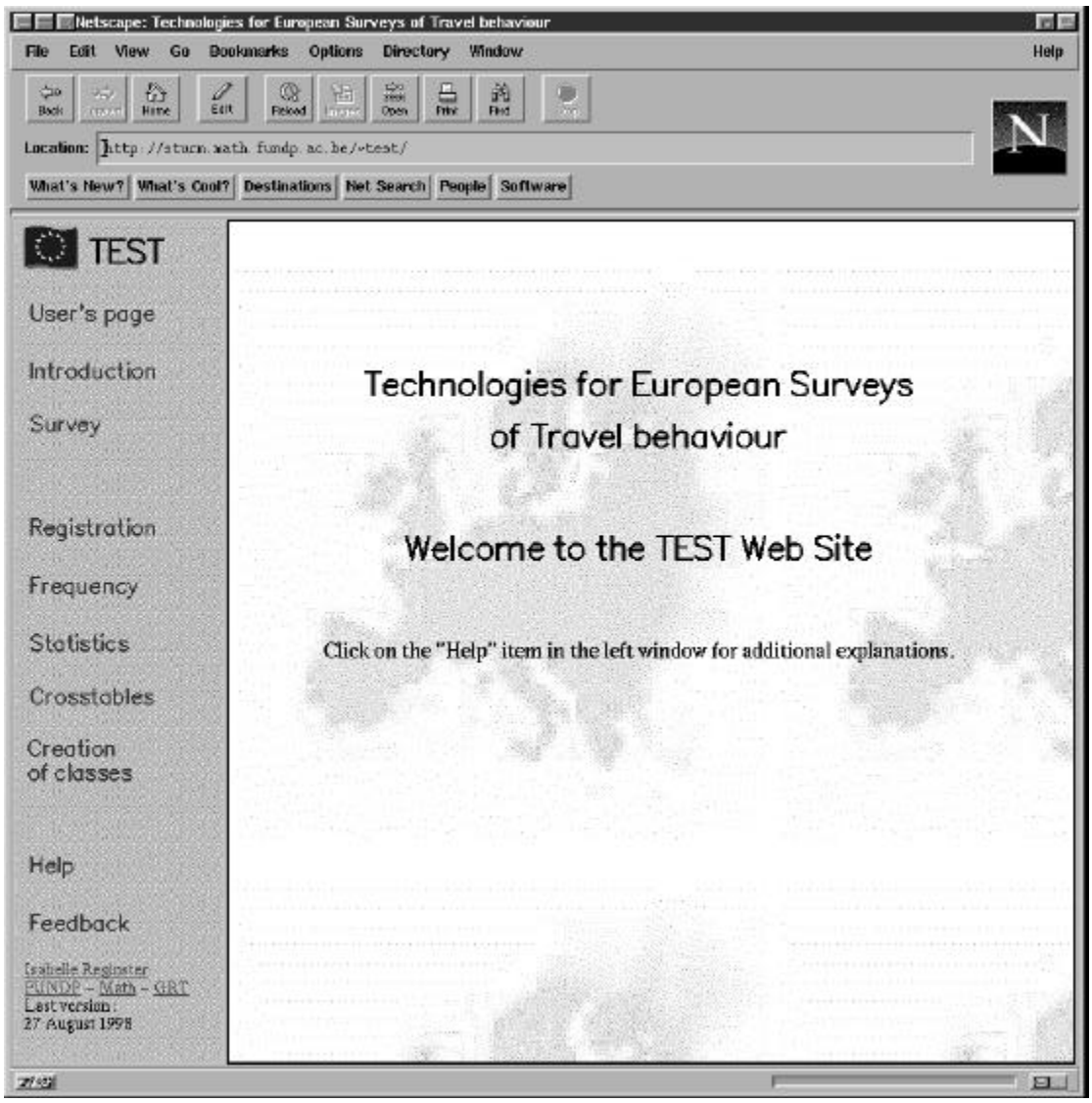


Figure 22 TEST web site: Choosing a filter

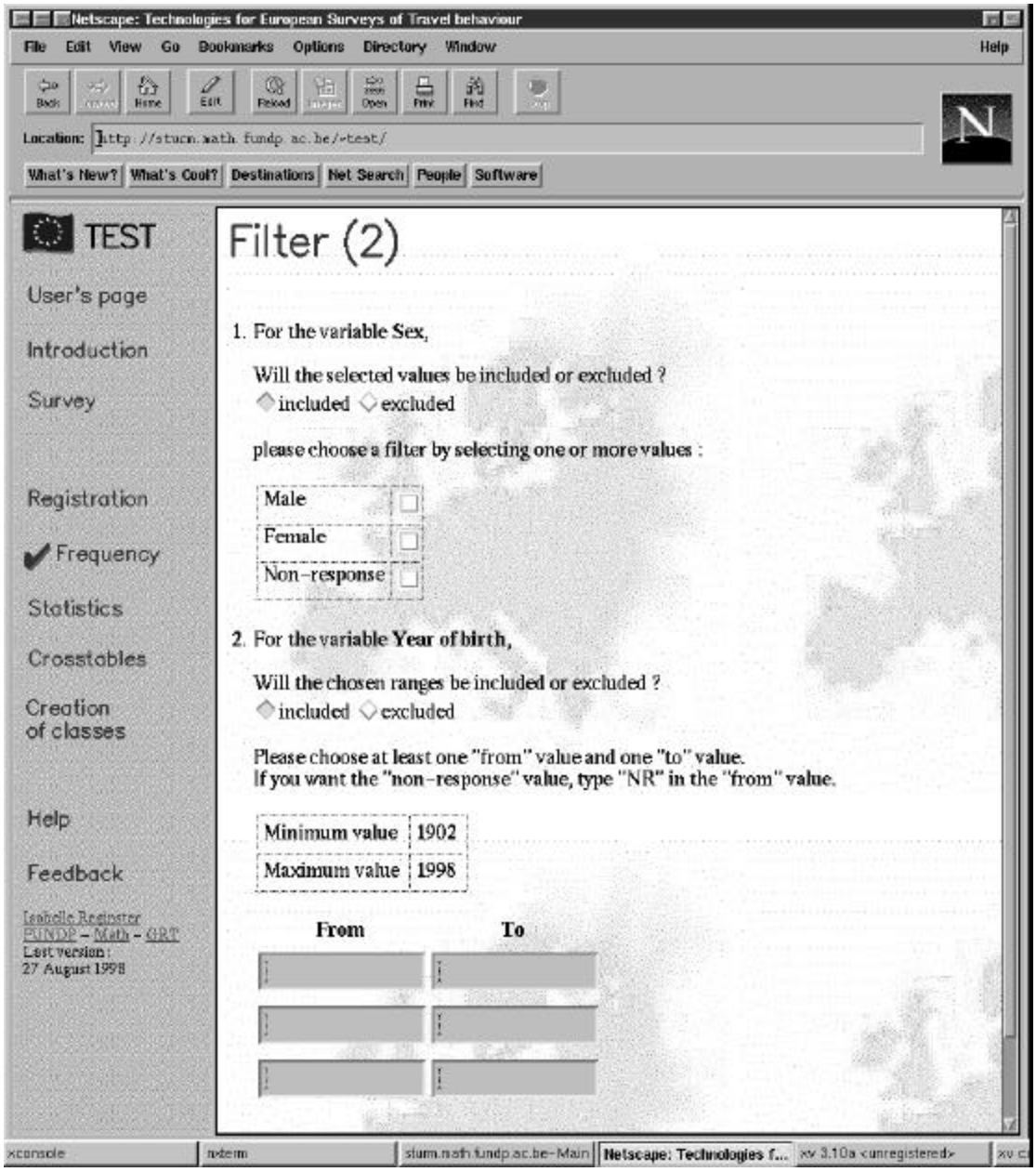
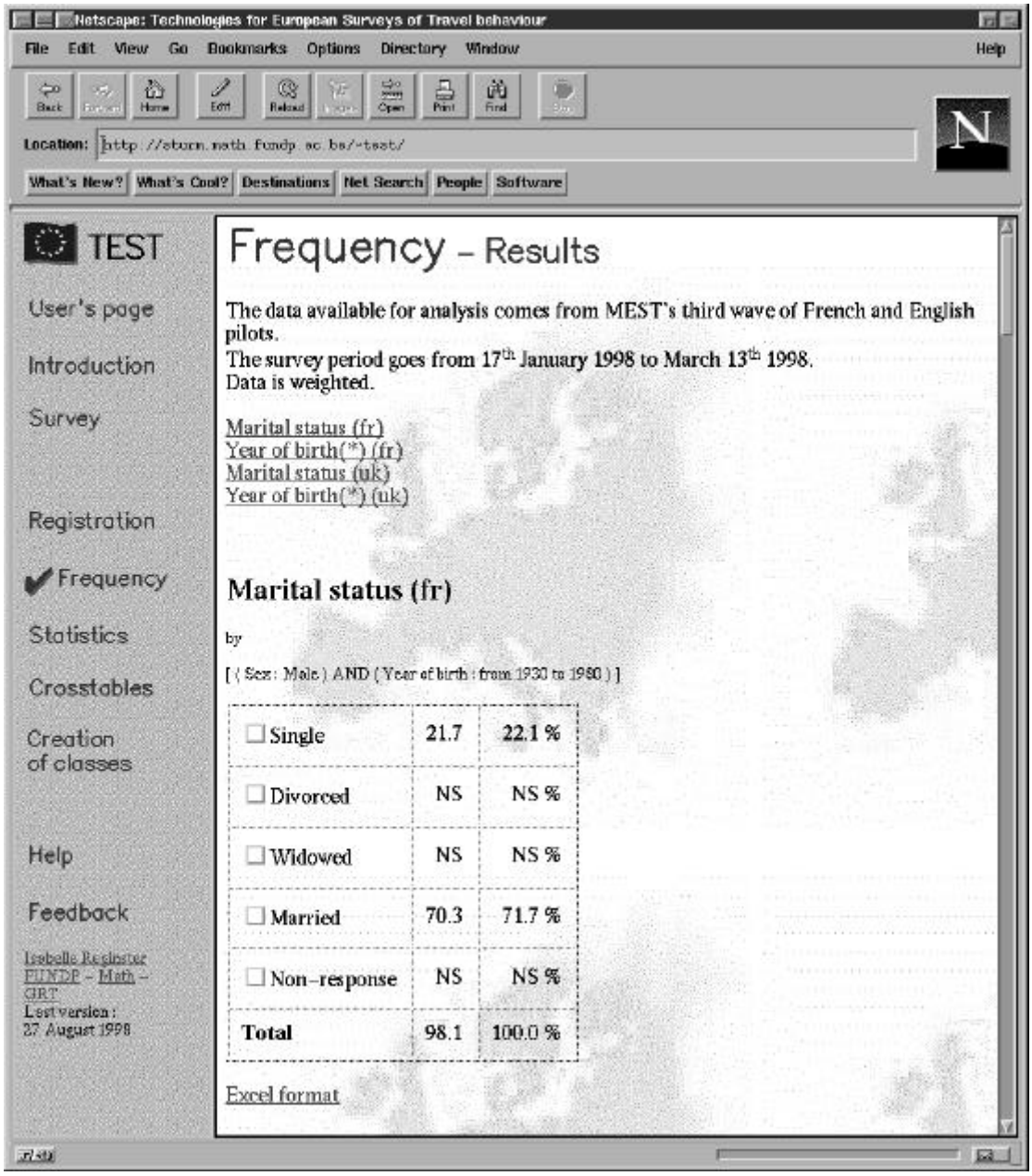


Figure 23 TEST web site: Tabulation results



3 CONCLUSIONS AND RECOMMENDATIONS

The project TEST encompassed a series of workpackages which looked at different technologies but which all aimed to improve the quality of the data obtained from travel surveys, in particular long distance travel surveys. All the workpackages demonstrated not only feasibility of the approaches chosen but also their long term potential:

- The hand held HPC based *travel diary application (TDA)* was shown to be acceptable under the conditions of small scale field tests to the respondents in four countries and collected information about the number of journeys to be expected.

- The WWW based *iTDA* was also success in two countries, in spite of the current slowness of the Internet. The overall system architecture with its specialised HTML tags opens new avenues for travel behaviour research

- Existing sampling difficulties (women, older respondents) limit its application at present, but the sharply rising market penetration and new more user friendly means of access to the Web should reduce these problems in the near future.

- The *GIS based work* substantially advanced the knowledge about the required structure and size of place name databases and their sources and provided a semi-automatic tool for the geocoding of place names and the derivation of shortest paths between places visited

- The *AI based parsing system* developed in Namur makes the classical techniques of imputation available to transport research enriched by logical checks for the standard variables and the provision of an audit trail for the data correction process

- The *neural net* based worked showed that self organising maps (SOM) can be a valid alternative to many classical imputation technique, while offering the identification of suspect data

- The *WWW based interface* to travel diary data demonstrated the usefulness of this open approach to data publication with a sophisticated web site built with free software tools, making it transferable to all interested parties

The TDA and iTDA work show that the optimal design of the survey interface depends very much on the technological environment. On the small screen HPC, respondents preferred a single page of questions, requiring scrolling, to give them an idea of the complexity of the task. In contrast, on the large PC screen, *iTDA* respondents preferred a dedicated one page one question format, which minimises data transfer across the rather slow Internet connection. The design has to be determined afresh for each new survey medium. The success of both approaches opens the way for truly near real time data collection of travel behaviour, as both allow respondents to be prompted at regular intervals in a non-threatening way.

The work on geocoding, parsing and imputation of travel surveys, imputation and error detection using neural networks has shown how current computing technologies and algorithms can support the survey manager and designer by providing very powerful tools to enrich and correct the data received by the respondents. The main conclusion derived from these efforts is the need to learn continuously from past surveys, either in the form of improved and expanded lists of place names, but also in the form of imputation methods properly calibrated. These should be integrated these into ongoing surveys, at the earliest possible moment, preferably while the survey is still in contact with the respondent: the CATI system, which prompts for the clarification of place names or the SOM, which queries outliers defined in terms of speed, mode and distance simultaneously.

The web based interface to travel data, successfully implemented here, and the parsing work has drawn our attention to the need for a specification language for surveys in general and travel surveys in particular, both with regard to the content and the logic of the data storage. The investment in the web site, which is a core component of both a general executive information system (European Transport Policy Information System) for decision makers and professional users as well as a citizens information system, will only show large dividends, if a wide variety of surveys can easily be defined in the system without the need for extensive reformatting and change.

The experience gained indicates a variety of new and interesting research directions, some specific to the technologies used and others directed at the integration of the approaches. This integration will be crucial for speedy take up in ongoing survey work. It is also clear that the methodological developments need to be seen in parallel with ongoing field work, as only the dialectic between large scale application and ongoing technological work can yield the desired results in terms of improved data for European transport policy.

The HPC class of devices is evolving rapidly to include mobile telephones as a standard feature, while their market is attacked by the even smaller palm top class of devices. GPS receivers are becoming available as PCMIA cards. Important research possibilities therefore are:

- Using the *iTDA* approach on the HPC devices by using the GSM phones integrated in such devices thereby reducing the programming effort and speeding up data retrieval

- Experiments with the new colour screens of HPC devices to overcome some of the problems encountered with the grey scale screen currently available

- Experiments with palmtop devices to establish their feasibility as survey instruments and their acceptability in use

- Integration of GSM, GPS or RDS/TMC based tracking and establishing its acceptability to survey respondents

- A large scale field test with the approaches to establish their potential as new survey tools for large scale survey work

Besides the important possibilities opened up in the integration of geocoding into other software, especially CATI and data coding software, automatic geocoding requires further work in its own right:

- Improvement and maintenance of official and public place name databases
- Development of comprehensive public map files of Europe, cross referenced against the place name data bases
- Integration of public and private data bases, especially for leisure and tourism locations
- Improvement of automatic geocoding search strategies to reduce the number of possible candidates in the case of ambiguities

The increasing computing power available at home and/or on the road, the rapid development of geographical databases and electronic maps and the work on error detection in TEST indicates that the following approaches are fruitful avenues for further work:

- Integration of automatic geocoding into survey software either in the framework of a standard commercial CATI systems, the web based system developed here or - in reduced form - into the TDA software.
- Integration of automatic detection of suspect data into either standard commercial CATI systems, the web based system developed here or the TDA software.
- Integration of further external databases, such as timetable databases, gazetteers of places and attractions and similar travel relevant data, preferably via the internet

The work on imputation in MEST and TEST has demonstrated the usefulness of a standard suite of such tools, as well as the potential of some new approaches. It is clear, that

- Integration of the new approaches into the suite of imputation tools would be an important contribution to professional practice, in particular in conjunction with
- Tools for the automatic selection of SOM for imputation and identification of suspect data

The work undertaken on the WWW interface to travel diary data opens up a road to the efficient provision of data access and comparative information for the assessment of individual travel diary surveys. It would therefore be desirable to:

- Establish a European electronic archive of travel diary surveys including a web based interface to them

- Develop further web accessible tools for the analysis, mapping and graphical display of the travel diary data

- Develop tools for the comparison between surveys

- Extend the data description language of the web site to become general to make other transport related data web accessible.

In summary, TEST has demonstrated the contribution that a selected set of computing technologies and approaches can make to the improvement of travel data quality. The work reported here is clearly only at the beginning of the development phase. The consortium is convinced that field work with the techniques described as well as further development is needed to advance the state of the art in this field.

4 REFERENCES

4.1 DELIVERABLES

Elvas, C. and J. Viegas (1999) Assessment of the technologies, report to the CEC, DG VII, *TEST Deliverable,* **D7,** TIS, Lisbon.

Fessant, F. and S. Midenet (1999) A knowledge-based parser: Neural-network-based approaches, report to CEC, DG VII, *TEST Deliverable,* **D5B**, INRETS, Arcueil.

Haubold, I. and K.W. Axhausen (1999) A real-time recording system, report to the CEC, DG VII, *TEST Deliverable,* **D2**, Fakultät für Bauingenieurwesen und Architektur and Centre for Transport Studies, Innsbruck and London.

Haubold, I., P. Jackson, K.W. Axhausen and J.W. Polak (1997) Technology assessment, report to the CEC, DG VII, *TEST Deliverable,* **D1**, Fakultät für Bauingenieurwesen und Architektur and Centre for Transport Studies, Innsbruck and London.

Hubert, J.P., P.O. Flavigny and J.L. Madre (1999) Tools to enrich travel diary data sets, report to CEC, DG VII, *TEST Deliverable,* **D4**, INRETS, Arcueil.

INRETS (1998) GIS-based enrichment of travel diary data sets, report to CEC, DG VII, *TEST Deliverable,* **D9**, INRETS, Arcueil.

Lothaire, O. and P. Toint (1999) A knowledge-based parser: Implementation of a tool-box, report to CEC, DG VII, *TEST Deliverable,* **D5A**, GRT, FUNDP, Namur.

Plaxton, J. and J.W. Polak (1999) A respondent support system, report to the CEC, DG VII, *TEST Deliverable,* **D3**, CTS, Imperial College, London.

Reginster, I. (1998) A www-interface to travel diary results, report to CEC, DG VII, *TEST Deliverable,* **D6**, GRT, FUNDP, Namur.

TEST Consortium (1999) Final recommendations, report to the CEC, DG VII, TEST Deliverable, D8, Fakultät für Bauingenieurwesen und Architektur, Innsbruck.

4.2 WORKING PAPERS

Haubold, I. and K.W. Axhausen (1997a) Analysing the options, *TEST Internal Working Paper,* **IWP13**, Fakultät für Bauingenieurwesen und Architektur, Innsbruck.

Haubold, I. and K.W. Axhausen (1997b) Initial design of the technologies, *TEST Internal Working Paper,* **IWP14**, Fakultät für Bauingenieurwesen und Architektur, Innsbruck.

Haubold, I. and K.W. Axhausen (1998) First experiences with a real-time recording system, *TEST Internal Working Paper*, **IWP07**, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.

Fessant, F. and S. Midenet (1998) First experiences with a neural-net parser, *TEST Internal Working Paper,* **IWP11**, INRETS, Arcueil.

INRETS (1997a) Design of a tool to enrich travel diary data sets, *TEST Internal Working Paper*, **IWP03**, INRETS, Arcueil.

Midenet, S. and F. Fessant (1997) Design of a neural-net parser, *TEST Internal Working Paper*, **IWP05**, INRETS, Arcueil.

Lothaire, O. and P. Toint (1997) Design of a knowledge-based parser, *TEST Internal Working Paper*, **IWP04**, FUNDP, Namur.

Lothaire, O. and P. Toint (1997) First experiences with a knowledge-based parser, *TEST Internal Working Paper*, **IWP10**, FUNDP, Namur.

Plaxton, J. and J.W. Polak (1997) Design of respondent support systems, *TEST Internal Working Paper*, **IWP02**, Centre for Transport Studies, Imperial College, London.

Plaxton, J. and J.W. Polak (1998) First experiences with a respondet support system, *TEST Internal Working Paper*, **IWP08**, Centre for Transport Studies, Imperial College, London.

Reginster, I. and P. Toint (1997) Design of a WWW-interface, *TEST Internal Working Paper*, **IWP06**, FUNDP, Namur.

Reginster, I. and P. Toint (1997) First experiences with a WWW-interface, *TEST Internal Working Paper*, **IWP12**, FUNDP, Namur.

TEST Consortium (1997) Proceedings, First TEST workshop, April 1997, Innsbruck, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.

TEST Consortium (1997) Proceedings, Second TEST workshop, June 1997, Innsbruck, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.

TEST Consortium (1998) Proceedings, Third MEST/TEST workshop, September 1998, Pörtschach, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.

4.3 ADDITIONAL WORKING PAPERS

Denstadli,J.-M. and J.-I. Lian (1998) Memory effects in long distance travel surveys, *TEST Working Paper*, TOI, Oslo.

Denstadli, J.-M. (1998) Non-response in travel surveys using CATI, *TEST Working Paper*, TOI, Oslo.

Denstadli, J.-M. (1998) Analysing air travel: A comparison of different survey methods and data collection procedures, *TEST Working Paper*, TOI, Oslo.

Lindberg, K. (1998) Border surveys: State of the art, *TEST Working Paper*, TOI, Oslo.

Lindberg, K. (1998) Method issues in road and airport surveys: Variation and coverage, *TEST Working Paper*, TOI, Oslo.

4.4 RELATED PAPERS AND PRESENTATIONS

Axhausen, K.W. (1997) Towards an European benchmark survey of long-distance travel, paper presented at the EC Concerted Action Meeting on Transport Information Systems, Brussels, December 1997, *Arbeitsberichte Straßenverkehrswesen*, **56**, Institut für Straßenbau und Verkehrsplanung, Leopold-Franzens-Universität, Innsbruck.

Axhausen, K.W. (1997) Long-distance travel diaries: European experiments, presented at "Colloque déplacement à longue distance, Ministère de 'lEquipement, du Logement, des Transports, Paris, October 1997.

Axhausen, K.W. (1997) Verkehrsbefragungen als Entscheidungsprozess: Erfahrungen mit Befragungen zum Fernverkehr, presented at the ETH Zürich, September 1997, *Arbeitsberichte Straßenverkehrswesen*, **46**, Institut für Straßenbau und Verkehrsplanung, Leopold-Franzens-Universität, Innsbruck.

Barbieri, A., Ph. Barette, C. Durand, E. Heylen, M. Mouchart, S. Saelens, Ph. Toint and P. Vansevenant (1998) Pilot-study for a national travel survey in Belgium , paper presented at the *8th World Conference on Transport Research*, Antwerpen, July 98.

Haubold, I. and K.W. Axhausen (1998) Handheld PC based travel data collection by self-administered interviewing, presented at the ASC Conference on Statistical Computing, August, London.

Fessant, F. and S. Midenet (1998) Imputation of partial non-response in surveys with a self-organizing map-based model, presented at Approches Connexionnistes en Sciences Economiques et en Gestion, 5ième Rencontre Internationale, Louvain-la-Neuve, November 1998.

Plaxton, J, P.G. Jackson and J.W. Polak (1998) Using the Internet to collect data on travel behaviour, paper presented at the *26th European Transport Forum*, PTRC, London .

Plaxton, J, P.G. Jackson and J.W. Polak (1999) Using the Internet to improve long distance travel data collection', paper presented to the *78th Meeting of the Transportation Research Board*, Washington, January 1999.

Polak, J.W. and P.G. Jackson (1997) Collecting and disseminating travel behaviour data using the World Wide Web, paper presented to the *Association for Survey Computing*, Imperial College, London, September 1997.

Reginster, I. (1999) The TEST Web Site, presentation given at the Department of Environment, Transport and the Regions, London, January 1999

Reginster, I. (1999) The TEST Web Site, presentation given at the Institut National de Statistique, Bruxelles, February 1999.

4.5 PROJECT WEBSITES

The following websites provide access to project results:

www.fundp.ac.be/~grt/test	Project deliverables
www.ic.ac.uk/jmplax/test	WWW-based surveys: tols and examples
sturm.math.fundp.ac.be/~test	WWW-based access to travel diary data

4.6 OTHER REFERENCES

Axhausen, K.W. (1996) Possible contents and formats for long-distance travel diaries, report to the CEC, DG VII, Deliverable D2 - Project MEST, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.

Herry, M., G. Sammer, M. Schuster, G. Röschel and M. Russ (1997) Allgemeine Mobilitätserhebung der österreichischen Haushalte, report to Bundesministerium für Wissenschaft und Verkehr, Wien and Graz.

Youssefzadeh, M. and K.W. Axhausen (1996a) Long-distance diaries today: Initial review and critique, report to the CEC, DG VII, MEST Deliverable, D1, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.

Youssefzadeh, M. and K.W. Axhausen (1996b) Tender documents: First MEST pilots, report to the CEC, DG VII, MEST Internal Working Paper, IWP01 and IWP02, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.

Youssefzadeh, M. and K.W. Axhausen (1998) Tender documents: Third MEST pilots, report to the CEC, DG VII, MEST Internal Working Paper, IWP05 and IWP06, Fakultät für Bauingenieurwesen und Architektur, Leopold-Franzens-Universität, Innsbruck.