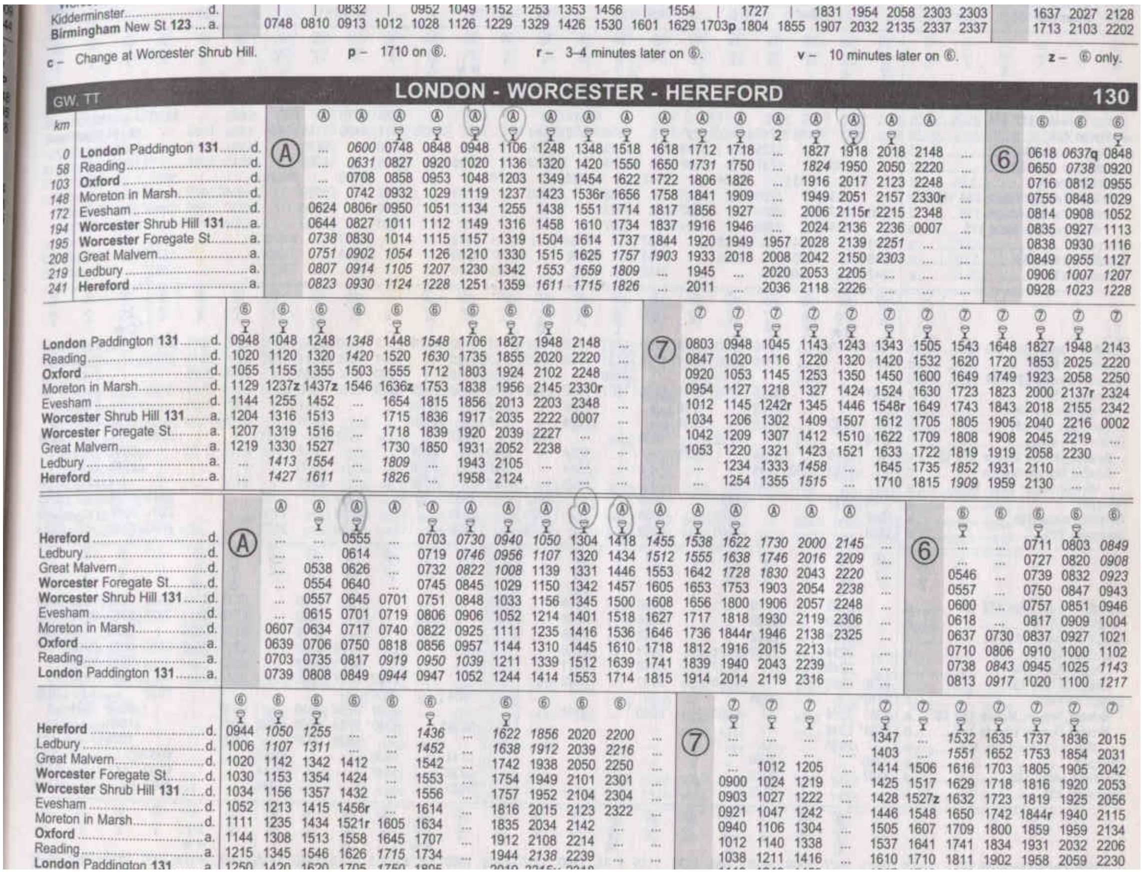


Working Paper

Measures of rail Impedance in the improved IVT European rail model

J. Hackney

Arbeitsberichte Verkehrs- und Raumplanung **262**

September 2005

Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Contents

1 Introduction 5

2 Data needs for realistic route choice modelling 5

3 Source of the timetable data 6

3.1 Option 1: Batch search with Hafas 6

3.2 Option 2: HaCon performs the search 6

3.3 Option 3: Script to automatically search in Hafas 7

3.4 Option 4: IVT model with Thomas Cook Timetable 8

4 The method for calculating rail impedances 8

4.1 The IVT European Rail Model 8

4.2 Extraction of NUTS3 centroid geocodes from GISCO 9

4.3 Conversion of geocode to IVT coordinate system 9

4.4 Introduction of NUTS3 centroids into the rail model 9

4.5 Calculation of the shortest paths 10

4.6 Spatial improvement of IVT European Rail Model for ETIS Base 10

5 Evaluation of improved IVT European Rail Model 15

5.1 Spatial suitability of the data: Uncovered TEN corridors 15

6 Post-processing of assignment results to derive distance and train type per country 17

7 Travel time, frequency, and in-vehicle time matrices 19

7.1 Some feasibility tests of the impedances 21

8 References 21

Tables

Table 1: Number of Train Lines per Country in IVT European Rail Model....................13

Figures

Figure 1: The initial map of assigned lines (IVT European Rail Model): Blue= assigned, Red = TEN (RAILNET attribute TENALIGN=1), Blue Points = NUTS3 zone centroids. 11

Figure 2: The final map of assigned lines (IVT European Rail Model): Blue= assigned, Red = TEN (RAILNET attribute TENALIGN=1), Blue Points = NUTS3 zone centroids. 14

Arbeitsberichte Verkehrs- und Raumplanung 262

Widerstände in einem verbesserten IVT Eisenbahn-netzmodell

JK Hackney
IVT
ETH
CH – 8093 Zürich

Telefon: +31-1-633 3325
Telefax: +31-1-633 1057
hackney@ivt.baug.ethz.ch

September 2005

Kurzfassung

Text

Schlagworte

European Rail Model, timetable, IVT Rail Model, NUTS, impedance, ETIS Base

Zitierungsvorschlag

Hackney, J. (2004) Measures of rail impedance in an improved IVT European rail model *Arbeitsbericht Verkehrs- und Raumplanung*, **262**, IVT, ETH Zürich, Zürich.

Working Papers Traffic and Spatial Planning 262

Measures of rail impedance in an improved IVT European rail model

JK Hackney
IVT
ETH
CH – 8093 Zürich

Telefon: +31-1-633 3325
Telefax: +31-1-633 1057
hackney@ivt.baug.ethz.ch

September 2005

Abstract

The rail network model at IVT was used to calculate measures of travel impedance between NUTS3 centroids. The model was improved to spatially cover the Trans-European Network and to ensure basic connectivity of all NUTS3 regions which are reachable in reality by train. The improvements provide realistic aggregate spatial coverage of the TEN. Significant additional improvement of the model in the form of additional train lines is necessary in order to provide realistic routes and impedances on specific OD relations. The options for obtaining improved measures of rail impedance are discussed.

Keywords

European Rail Model, timetable, IVT Rail Model, NUTS, impedance, ETIS Base

Preferred citation style

Hackney, J. (2004) Measures of rail impedance in an improved IVT European rail model, *Arbeitsbericht Verkehrs- und Raumplanung*, **262**, IVT, ETH Zürich, Zürich.

1 Introduction

The ETIS Base project requires rail impedances on all relations between NUTS3 centroids and major airports. In addition, the distance traveled by each type of train in each of 27 EU and EFTA countries are needed between all NUTS3 regional centers in order to calculate the rail travel cost in the regression models.

The relation or OD (origin-destination) travel time, service frequency, and in-vehicle time between NUTS3 regional centers and between NUTS3 centers and the airports are needed to construct the set of alternatives for the mode choice model.

2 Data needs for realistic route choice modelling

The time and distance traveled are route-dependent and special effort was placed on a realistic representation of rail route choice alternatives:

- Accurate geocoding of NUTS3 centroids.
- Realistic connector travel times between NUTS3 centroids and the rail services.
- Availability of complete timetables for all train operators.
- Timetables that are consistent with each other (e.g. from the same month).
- Realistic (if not complete) sampling of the train timetables including access/egress to NUTS3 centroids.
- Realistic assignment assumptions.

The volume of data is very high. For all 1262 NUTS3 regions in the EU and EFTA, 1262 x 1262 (about 1.6 million) relations are needed. This indicates an additional factor in the modeling, the need for batch processing.

The solutions to all these challenges and the resulting database of travel times between NUTS regional centers and the travel distance per train type per country are described here.

3 Source of the timetable data

Detailed, digitized train timetables were necessary for all train operators in the 25 EU, 3 EFTA, and the countries immediately east of Europe. Discussions with EU project partners and HaCon GmbH, which together with the Deutsche Bahn maintains the most complete digitized timetable for Europe and surroundings, revealed that a timetable search of the scope and scale of the one necessary for ETIS Base is unprecedented in their experience. By HaCon's estimates, the search through their database for the 1.6 million OD relations would require 10-14 days on 8-10 desktop PCs.

3.1 Option 1: Batch search with Hafas

A batch search of all train schedules is to our knowledge possible only with the Hafas software of HaCon. Working with Hafas would be logical since access to these timetables would otherwise have to be negotiated with each individual railway. The raw data from each railway is provided in confidence to HaCon, who translates this data into a rapidly searchable binary format whose properties are proprietary. This means that the Hafas search program has to be used to search the database.

IVT had hoped to use the batch-search extension of Hafas to search the database of the Deutsche Bahn. IVT was privileged to use this program at a transit authority in Zurich to search the local transit services. However this was the only example of this tool outside of HaCon. The software is understandably not available to parties outside the company for other databases for reasons of data privacy: it would amount to giving away the data that HaCon has collected in confidence. The transit authority in Zurich was searching its own data so there was no privacy issue in that case.

Writing a program to search the Hafas database is not realistic because of its unknown encoding. A batch search of the Hafas format is therefore not possible unless HaCon performs it.

3.2 Option 2: HaCon performs the search

HaCon could have been contracted to carry out the route search at an estimated 10,000-20,000 Euro at a time horizon of December, 2004. This was determined to be disproportionately expensive for the relative importance of this subproject within ETIS, and too late. The cost would have been for a one-time search through the base-year timetable of the Deutsche Bahn. The results are a static database of OD shortest-path impedance variables and are not part of a network model, nor are they usable in a network model with a timetable, since they only represent one instance of travel from an O to a D.

The results of the search would probably have given a selection of the correct “best” routes with the lowest generalized cost, an accurate count of train changes, train types, and accurate travel times, between train stations. But it would probably not have been suitable for the question of routes between NUTS centroids or for kilometers per train type. The latter is because there is no provision in Hafas for writing out the number of kilometers traveled per country. And since the train stations are also not expressed in geocodes, geographic analyses would require additional data handling.

The other problem is that the Hafas search engine must search from train station to train station, so the route search does not include the generalized costs of access and egress from a geocoded point such as a NUTS regional center. The simultaneous choice of ”best” train station and “best” route is therefore not automatic and would have to be done for several train stations for each NUTS3 origin and destination, whose “best” routes would then have to be compared in a post-processor to find the route with the lowest generalized cost. Searching 3 train stations for each NUTS3 zone would have increased the search volume 9 times over the initial estimate. A price for this increased effort was not discussed.

3.3 Option 3: Script to automatically search in Hafas

An alternative to letting HaCon calculate the relations would be to write a batch query script for the server or CD version of the Hafas Windows interface. The challenge therein is how to save the formatted text output of the searches, geocode them, and reformat them to render the search useful for calculation. The advantage is saving the money that would otherwise be spent at HaCon, having the ability to control search parameters, and having a program that can be used to repeat the procedure for a future dataset of updated timetables. The weaknesses of the Hafas search engine for this application are the same as if HaCon did the search.

It is likely that a search carried out by HaCon would be of the best available quality however, since Hafas is their own program and they understand the search parameters the best. It is possible for the longer term in ETIS that Hafas could be developed with tools for broader Europe-wide searches if there is continued demand for this service.

3.4 Option 4: IVT model with Thomas Cook Timetable

In the absence of a batch-searchable database, IVT instead used its European Rail Network model (Fröhlich and Bleisch, 2003) written for VISUM 8 and updated to VISUM 9.15 (PTV 2004). Timetables from the base year for high speed trains had already been digitized by hand from the Thomas Cook Timetable (September, 2002). The search algorithm simultaneously incorporates egress and access times to choose the "best" local train station as well as the route for a relation, and the geocoded nodes and links in the model mean that the geographic characteristics of the relations can be written out more explicitly than from Hafas. The disadvantage is the bias toward high-speed service and the neglect of local trains, which covers long-distance travel well, but tends to impose longer travel routes than are realistic on local travel. The waiting times between connections are also unrealistically long because of the incomplete representation of the timetables. The basic timetable and the extension of this model for ETIS to spatially cover the TEN corridors is described below.

4 The method for calculating rail impedances

4.1 The IVT European Rail Model

IVT uses the IRPUD European railway network (Schürmann, 2001), with improvements added by IVT (Bleisch and Fröhlich, 2003). The rail network geometry consists of 36,000 nodes and 78,000 links. The geometry of the network and the link attributes were coded from the RAILNET GIS network. The workday and weekend schedules for high-speed intercontinental and interregional trains were added by hand from the Thomas Cook Timetables of September, 2002. Trains to the regional express level were added in the Alps for workdays. Only the schedules of the high-value connections were included for the rest of Europe. The Thomas Cook timetables contain an estimated 100,000 trains. In the form before the changes for ETIS, there were some 5700 rail lines implemented in the IVT model, serving some 2300 stops, plus selected ferry lines.

85 European airports are also integrated in the network. Airports without a major train station (with Regional Express service) were connected with appropriate "S-Bahn" (light rail), tram, or bus link. This linkage of the airports to the network is important for the assignment of intermodal models. For those airports which do not have a direct link to the rail model, the access was approximated with a road access time as described below to simulate local bus service to/from the city center.

4.2 Extraction of NUTS3 centroid geocodes from GISCO

The geocodes (longitude, latitude) of the NUTS3 centroids were extracted from the GISCO (2003) database file 7ptl3. These are the administrative centroids for EU, EFTA, CEC, and eastern countries for 2001. ArcToolbox was used to convert the coverage to a shapefile and ArcMap was used to assign the X, Y variables to longitude and latitude and to write out the database. This ASCII file consists of the NUTS3 name, plus the longitude and latitude of the centroid.

4.3 Conversion of geocode to IVT coordinate system

The longitude and latitude were converted with the shareware map conversion program Projmap 4.4.6 (www.remotesensing.org/proj) to the coordinate system of the IVT model, which is a Gauss-Kruger conical conforming projection on a Clarke1866 spheroid with center (Longitude = 10, Latitude = 52) and reference latitudes 27 and 63 degrees North. The X,Y coordinates are in meters from this origin. The ASCII file consists of the NUTS3 code, X, and Y.

4.4 Introduction of NUTS3 centroids into the rail model

The 1354 NUTS3 centroids (EU, EFTA, CEC, Balkans and Russia) are entered into the IVT model as model zones (Bezirke) in a “network” file, a formatted ASCII file with the suffix “.net”. Each zone is first assigned a dummy identifier variable ID between 90001 and 91354, and zone type 3 to differentiate it from the other zone types that are already in the model (airports and harbors). These variables are added manually in MS Excel. An appropriate file header describing the format and data content is manually written to the file so it can be read into VISUM (an example for this is in the VISUM help files). The zones are imported to the model in VISUM using the network file import utility.

The zones are then connected to the rail network with connectors of fixed travel speed to simulate access to and egress from the train stations to and from the zone centers. Straight line connectors shorter than 20 kilometers are given the speed 25 km/h to simulate urban and intra-regional travel on secondary roads. Straight line connectors longer than this were given the speed 40 km/h to simulate regional trains not modeled and interregional travel on higher speed roadways. Only major ferry service was modeled, with speeds from actual timetables found on the internet (see references). The model was checked so that islands without rail connections to the continent are not connected to the continent artificially and that regions outside the area of consideration were not unnecessarily connected. These NUTS regions remain intentionally unconnected by rail to the rest of Europe: DK007, ES538, ES6319, ES6322, ES7013, ES7027, FI2, FR8329, FR91, FR92, FR93, FR94, GR4214, GR4220, GR4315, GR4329, GR4336, GR4349, IS, PT2, PT39, SE094, UKM46, RU.

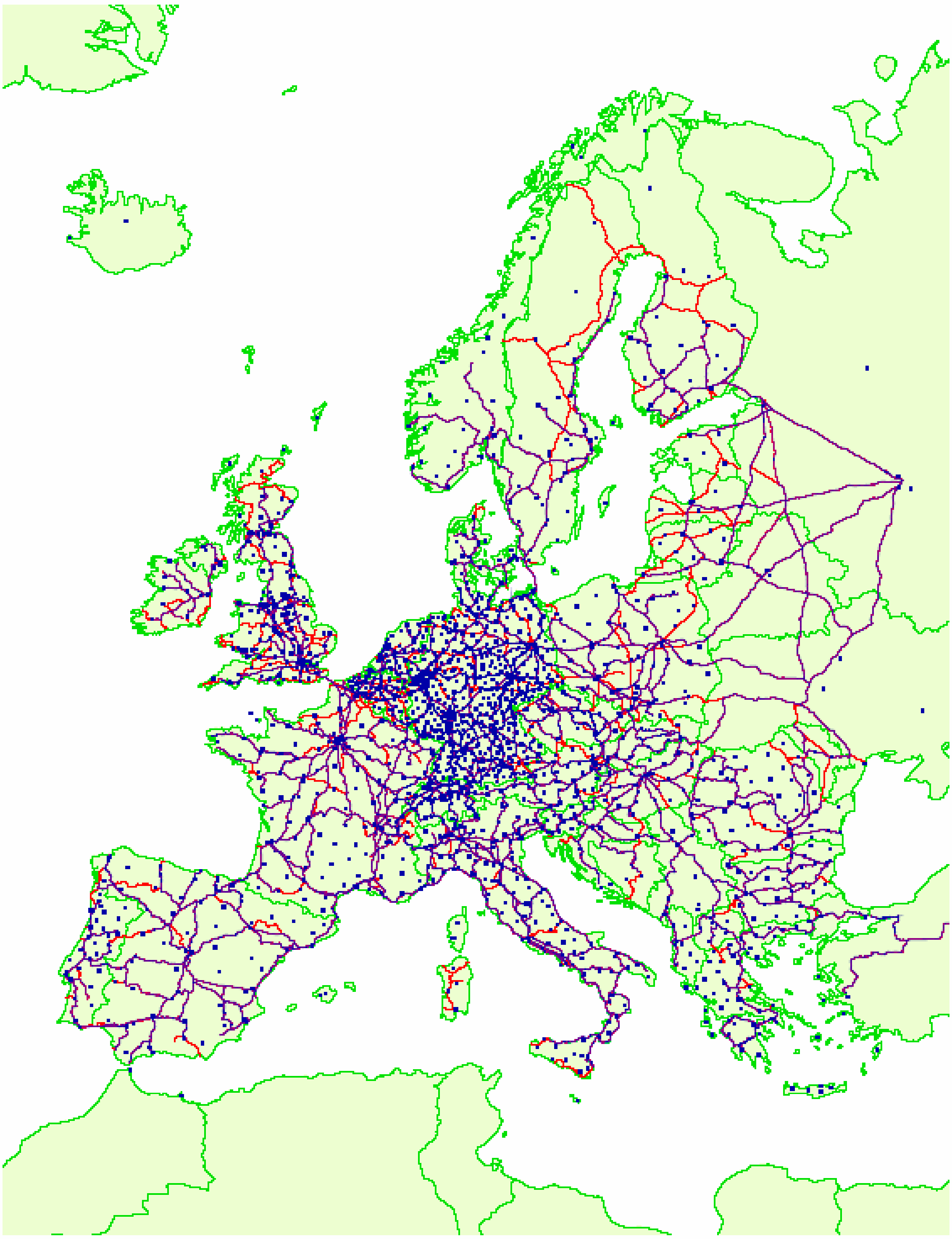
4.5 Calculation of the shortest paths

A branch and bound search was used for the shortest time path. The generalized cost was the total travel time, with equal weighting for access/egress, travel, and waiting times, with no time penalty for train changes. The assignment was made using a Kirchoff rule with β=4.0. A generous search envelope was used: maximum 10 train changes per relation und 24 hours' waiting time for each change, possible departure times between 0:00 and 24:00 (the entire timetable was searched), and a maximum interval in which the trip must be completed, of 168 hours (7 days). This large envelope was necessary to find rail connections between certain NUTS3 zones because of the poor connectedness of some parts of the network. A "demand matrix" of 1 passenger per NUTS3 origin and destination was produced in a Java program and read into the VISUM model for assignment. The assignment step is necessary in order to make VISUM write out the line-level information from the timetable. The standard matrix parameters were written out, plus two special tables for post-processing, as outlined below.

4.6 Spatial improvement of IVT European Rail Model for ETIS Base

The assignment in the original IVT rail model in Figure 1 showed insufficient spatial coverage of the TEN corridors, which is a baseline goal for ETIS. It also showed a tendency to overestimate local travel distances because of the lack of regional train service. New train service was therefore strategically added to cover the TEN.

Figure 1: The initial map of assigned lines (IVT European Rail Model): Blue= assigned, Red = TEN (RAILNET attribute TENALIGN=1), Blue Points = NUTS3 zone centroids.



The spatial weaknesses in the Baltic states, Scandinavia, UK/Ireland, Iberian Peninsula, southern Italy and eastern Europe are understandable because the original model was built to focus on central Europe and the Alps. Significant densification of the train service in the weakly represented areas was necessary. However time was not sufficient to add trains up to the regional express level in all of Europe, as had already been done in the Alps. This would have entailed an increase in the size of the model of approximately 10,000 trains, or 200%. Building the existing model had already required 6 man months of labor, so such an expansion would have meant another year of work.

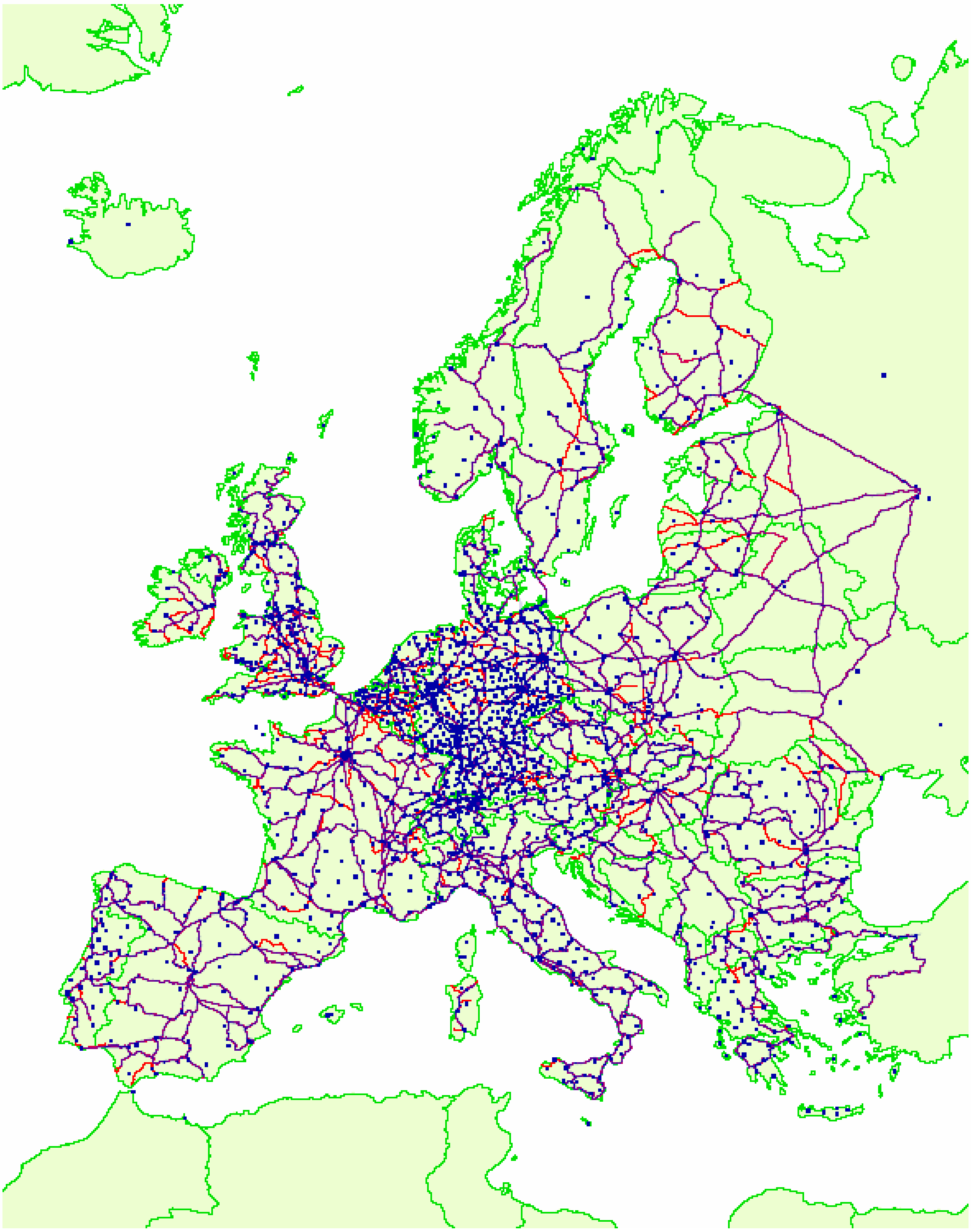
A prioritizing system was needed for adding service from the timetable since not all lines could be added. Three dimensions of service were considered: spatial density, frequency, and number of stops on a route. Experience showed that small variability in frequency of service on a corridor or to a particular train station can cause large changes in the assignment and therefore the distances traveled in each country. The largest change of course, comes when the first train is put on a route. It was therefore decided to concentrate on putting trains on unserved TEN corridors. Table 1 shows the number of lines in the original model by country, and the improvements made for ETIS in bold face.

In all, 398 new trains were added to the 5756 in model, an increase of about 7%. The improvements are seen in Figure 2. The coverage of the TEN lines (RAILNET attribute TENALIGN = 1; Route defined) is much improved.

Table 1 Number of train lines per country in IVT European Rail Model

Country	Lines	Country	Lines	Country	Lines
Albania	8	Greece	37	Slovenia	6
Austria	161	Holland	58	Spain	204 + **40**
Belarus	4	Hungary	103	Sweden	103 + **26**
Belgium	50	Ireland	77 **(see UK)**	Switzerland	664
Bulgaria	32	Italy	680 + **62**	Turkey	18
Croatia	8	Macedonia	2	Ukraine	8
Czech Rep.	53+ **8**	Norway	13 + **14**	LT/LV/ET	**+ 10**
Denmark	174	Poland	107 + **32**		
Finland	83 + **34**	Portugal	40 + **4**		
France	592 + **42**	Romania	40	International	698
Germany	653 + **61**	Russia	9	Airport Cnx.	166
Great Britain	414 + **65**	Slovakia	93		

Figure 2: The final map of assigned lines (IVT European Rail Model): Blue= assigned, Red = TEN (RAILNET attribute TENALIGN=1), Blue Points = NUTS3 zone centroids.



5 Evaluation of improved IVT European Rail Model

5.1 Spatial suitability of the data: Uncovered TEN corridors

There are several reasons why a TEN corridor would not be covered. Most obviously, it could be that there was no train information on that corridor in the Thomas Cook timetable of 2002. Either the route was not completed yet, the information was unavailable to the publishers of Thomas Cook, or the RAILNET attribute designating a TEN corridor was faulty. Alternatively, the train could be a local service which was not considered in this effort. (Regional trains were indeed included where they were the only service to a certain NUTS3 region, but in this case the links would be assigned.)

Assuming the timetable was available for a TEN corridor and it was entered in the model, there are reasons that it may not be assigned. The choice of lines favored high speed trains which have so few stops that they pass by many NUTS3 centers. These centers had to be connected to the further main train station, from which other TEN routes could also be accessed in addition to the one that runs closest to that NUTS center. That segment of the TEN corridor is then bypassed by the demand, “short circuited” essentially. This will result in unassigned corridors until the nearest station to each NUTS3 center is served with at least one train per day. But in order to get the correct train type (and therefore price structure), the relative number of stops per day at each station should roughly correspond to the real distribution, which is a very large task of data entry as mentioned.

Other TEN corridors that extend to remote areas are not assigned because there is no NUTS3 centroid near the end station of the corridor. This occurs for example in coastal areas like southern Spain, southern Ireland or western Sardinia and Sicily. In the NUTS3 zone “South-West Ireland” (IE025), for example, the NUTS3 centroid is more logically connected to Millstreet, a train station further inland, than to Tralee-Casement Station on the coast, and the remainder of the serviced route from Millstreet to Tralee-Casement is not assigned.

5.1.1 Suitability of the distance/train type calculation and bias

The failure to cover all TEN routes results therefore from a combination of the location of the NUTS3 centers and from the particular train services selected from the Thomas Cook timetables for use in the model. Increasing the sample size of trains from the timetable to spatially even out the supply on the network would provide more, but not complete, coverage of the TEN network.

More importantly, it would also decrease the bias toward high-speed service in those regions where more trains were sampled. The train-type bias is easily explained, but the effect on the cost calculation is not certain. Local travel in regions with few or no regional trains takes place in the current model on higher speed trains over longer distances than is realistic. There are, on average, 1.75 train stations per NUTS3 region with at least one train per day and 1.09 stations per NUTS3 region with 6 or more trains. The relative number of regional trains is greatly reduced compared to reality so the expected wait times are increased, as is the expected connectivity. As a result, the traveler from a NUTS3 zone traveling to another closely located NUTS3 zone must sometimes drive to a faraway train station and take a high speed train to another far away station, then drive a long distance again. An unrealistically high proportion of the journey is taken by non-rail modes. The train chosen is of too high a quality for realistic local travel. The distance traveled in the train is however perhaps not as far as that in a regional train, so the product of distance and the train-type-dependent cost might not result in higher travel costs than reality. Probably a good bit of this bias is moot when it comes to choosing a single train type to represent all travel within a country. A detailed study of this error has not been carried out (see below).

Improving local service would have system-wide effects, and if not done evenly (in space), it would distort the local bias in the choice of train type. This is already observed in the Alps region which has samples of the timetable to include regional express service. The most pronounced system effects of locally enhancing the train service is the added flexibility in departure and arrival times at the local NUTS3 centers, leading of course to more realistic local travel, but also to a much larger choice of high-quality trains for long-distance travel at the next major train station. This increases the international connectedness of the local region disproportionately to other similar regions whose local service had not been improved in the model.

However, every added train in a spatially neutral distribution would be desirable because detail really matters at the local level in getting the train type, thus the travel cost, correct. The overrepresentation of high speed trains is, apart from the Alps, relatively evenly spread throughout this model. Selective enhancement of certain corridors with better levels of service (higher frequency of regional trains) would correct the choice of train type locally, but it would alter the spatial distribution of the problem. The magnitude and spatial distribution of this change is not predictable. The Alps has already been enhanced as part of another project. However for the rest of the network, spatial bias should be avoided in favor of a spatially neutral rule about adding more trains.

This was attempted here with the spatial improvements to the network to cover the TEN map.

The resulting assignment is the best that can be accomplished for ETIS within the given constraints.

6 Post-processing of assignment results to derive distance and train type per country

The desired output is the type of train used and the number of kilometers traveled in each country on each shortest (time) path connection between an origin and destination pair. It is a challenge to unite geographic data with network data, which exists fundamentally in a time rather than spatial frame of reference. VISUM version 9.15, the newest available, does not output the desired values directly. The basic quantity needed for the calculations is the correspondence between the train (line) chosen and the links or nodes on the path traveled. This is the key since the latter are geocoded and assigned to a country in the VISUM model. The approach used here matches the train (line) to the country-coded nodes along its path, and interpolates where the country boundary lies between nodes. Between the last node of one country and the first node of the next country, the algorithm assumes that the country boundary lies in the middle of the link and assigns half the kilometers of track to each country.

6.1.1 Route table

The route table gives the train chosen for the OD pair and the stations for train changes (NUTS3_Bahn_Routen.att). This table is over 12 million lines long for this demand matrix, including access and egress trips and trips on foot or by bus to change trains. It includes a counting index for the stage of each OD journey. Lacking is the summary of how many kilometers were traveled in which country for each train. The only geocodable information in the file is the station where a train change occurred. On a shortest-time path this is an infrequent occurrence; often hundreds of kilometers are traveled between changes. Interpolating the allocation of traveled distance on one side or the other of the international border is imprecise with such long stretches. There is no information in the file about which links were traveled between the nodes at which a train was changed. This file contains the origin, destination and line (train type) chosen however, and must form the basis of the travel-distance matrix. To simplify further analysis, the file is manually manipulated as follows:

- A second counting index is introduced so that the order of the stages within the OD relations can be restored after necessary sorting steps.

- The non-train modes (access/egress and foot/bus modes used for train changes) are removed. This renders the counting indexes discontinuous, which does not diminish its usefulness.

- The train types are included in the name of the line chosen for the shortest-paths, and are derived as a variable on their own in the file. The train names (from the Thomas Cook European Timetable 9/2002) do not always correspond with the expected train names that IWW used in its regressions.

- Finally, the model zone number is converted to NUTS code using a correspondence table.

6.1.2 Train line table

A table of travel time and distance for all train lines, broken down into segments between train stations, can also be produced in VISUM (linienweg_europa_bahn_2002.att). The segments are labeled by the node number of the beginning and ending station of the segment. The country of each train station is known. The density of the train stations along the route is higher than the density of train stations where train changes occur. It can be assumed with only a few kilometers error that an international border falls exactly in the middle between two train stations which are in different countries.

6.1.3 Merging the tables: Country and train type

This table of stations per line is merged with the route table to allocate the traveled kilometers between countries. A Java program was written to parse the files and merge them together, matching the node number of the serviced stations for each OD relation and summarizing the traveled kilometers in each country for each train type. Comparing 12 million lines with some 5500 lines is a CPU- intensive step that requires an order of magnitude of a day to calculate.

Clearly, it is possible that several types of train were used on a relation within one country. Only one type of train is reported however, because the travel cost regressions are modeled based on a single train type per country. The dominant train type is defined in the regression and in the work here as the train type which traveled the most kilometers in the country, whereas the number of kilometers traveled is the total number of kilometers traveled by train in that country. There is some error in this method of reporting, but this is taken into account (in the mean) in the regression parameters, which were derived using the same method.

The train types in the Thomas Cook timetable (Thomas Cook 2002) are labeled differently in some countries than the convention used for deriving the regression. Especially the trains in Belgium and England have numbers instead of the designations R” for regional, D” for direct, and so on. The classification of the train types for their input into the regression will be done at IWW. Again, here is another possible source of error: the various train definitions in each country. Future replications of this work should take care to standardize the train definitions before beginning work.

6.1.4 Delivered distance matrix

The final matrix (cntry_dist_europa_bahn_2002.dat) is a TAB-delimited ASCII table with 1,758,486 + 1 lines and 78 columns. The first line is the column heading and the following lines are each one OD relation. The number of relations is not equal to 1262 x 1262 because some relations external to EU + EFTA were kept, and some of the internal zones are not connectable to the rail network (as indicated above). The columns correspond to each country as follows:

NUTS3_O NUTS3_D (KM1 TRAIN1) (KM2 TRAIN2)* …

Where * is repeated for each of 38 countries for a total of 76 columns: AL, AT, BA, BE, BG, BY, CH, CZ, DE, DK, EE, ES, FI, FR, GR, HR, HU, IE, IT, LI, LT, LU, LV, MD, MK, NL, NO, PL, PT, RO, RU, SE, SI, SK, TR, UA, UK, YU

Countries outside the focus of ETIS are included because the shortest-time path for many OD pairs occurs over tracks in these countries.

The units of distance are kilometers and the missing value is 999999.

7 Travel time, frequency, and in-vehicle time matrices

The rail impedance file delivered to ETIS consists of 5 measures:

1. Total trip time (min)
2. In-vehicle time (min)
3. Total trip distance (km)
4. In-vehicle distance (km)

5. Service frequency (per day)

Total trip time and distance include the access/egress on the network connectors with a generic non-rail mode, the train travel itself, and the changes between trains. In-vehicle time and distance refers only to travel with the train and does not include the generic non-rail modes for access/egress to the network.

One matrix is delivered (FINAL_TT_NUTS3.dat) with travel times from all airports and NUTS3 centroids in the model to all the other airports and NUTS3 centroids in the model. The matrix is ASCII formatted text with 11 TAB-delimited columns:

- Origin Code (NUTS3 code or airport code)
- Destination Code (NUTS3 code or airport code)
- Origin Name (NUTS3 code or airport name)
- Destination Name (NUTS3 code or airport name)
- Origin Code within model
- Destination Code within model
- Total trip time (min)
- In-vehicle time (min)
- Total trip distance (km)
- In-vehicle distance (km)
- Service frequency (per day)

There are 576,907 + 1 lines of data (OD relations). The missing value is 999999, but the data has been cleaned of lines with missing total trip time or total trip distance. The attributes of the relations is consistent with the relations used for the distance per train type per country.

If these impedance measures are realistic, they can be used for the destination and mode choice model that IVT intends to estimate as part of ETIS Base. Where they are not realistic enough, then HaCon will have to be contracted to calculate the correct shortest paths, because it is not feasible for the time deadlines in ETIS Base for IVT to enter thousands of trains from Thomas Cook into its model, or to have a subcontractor do this hand work.

7.1 Some feasibility tests of the impedances

The SPSS file Impedance_NUTS3.sav contains the impedances on OD relations between NUTS3 centroids. An additional column with the total waiting time per relation was added.

It is not clear what a good measure of quality is for service on an OD relation. Some measure must be assumed in order to begin assessing the realism of the model and its areas of weakness. A study of the aggregate data shows that the total wait time exceeds 3 hours in 22,532 cases and 6 hours in 7381 cases. In 3170 cases the trip speed is less than 40 km/h, which would be an indication that the train was slower than the connector speed. These rough measures establish the order of magnitude of the number of relations which probably do not have realistic routes, travel times, or other impedance measures in the improved IVT European Rail model

Comparison of individual relations could be made by hand via a search on Hafas, but this is not feasible for several thousand relations. Even if it is known that the service on a relation could be improved, it is difficult to tell using Hafas which lines should be added in the model in order to provide realistic service.

Additional work has to be done to establish a matrix of problematic OD relations before a mode and destination choice model can be estimated, to identify which relations are critically incorrect. But more must be known about critical values for measures of impedance beyond which a connection is not feasible for the traveler. Once the critical relations are identified, there are additional steps to take to rectify the shortcoming. It is possible to isolate these OD relations in the output data and to perform a deeper analysis, but a systematic method for identifying the steps to take is also not known. If the steps are few and simple, they may be done by hand. For example, more connectors could be linked to NUTS3 regions, or more trains could be entered into the model from the Thomas Cook timetable. However even small steps may not be feasible given the large number of unrealistic relations. It may suffice to identify the critically lacking information and have these gaps filled by HaCon in a subcontracted search through the DB timetables of 2002.

8 References

Bleisch, A. und Ph. Fröhlich (2003) Die Erreichbarkeit von Regionen, IBC Modul Erreichbarkeit Phase 1, Schlussbericht, (Hg.) BAK Basel Economics, Basel.

Ferry Services:

http://www.viamare.com/salamis-pireus-limassol-hafia-portsaid.html http://www.aferry.to/

http://www.undiscoveredscotland.co.uk/lewis/ferry/index.html

PTV AG, Stumpfstr. 1, D – 76131 Karlsruhe.

Schürmann, Carsten (2001): „Networks.txt“. Institut für Raumplanung, Universität Dortumund.

Thomas Cook's European Timetable with High Speed Trains Feature, September, 2002.

A 1 : Model Improvements in Spring, 2005 from IRDATA

IRDATA was asked to improve the IVT Rail Model by making sure the connectors between the NUTS3 centroids and the nearest train station represent the realistic choice of train station; adding names and codes to the NUTS3 centroids; adding additional train lines, especially in Eastern Europe; adding additional train stations as necessary to improve service to all NUTS3 regions; properly classifying Bus and Ship modes that had been entered as general "public transportation"; checking of the assumptions of the connectors to island NUTS regions. The work was carried out using VISUM 9.2 during May and June 2005.

The condition for connecting the NUTS3 centroid to a train station is that the station be the largest (most well-served) station within the NUTS3 region's boundaries. The model had, to this point, permitted connectors that crossed NUTS3 and even country boundaries, and some of the train stations were geographically closest to the NUTS3 centroid, but were sometimes not the main stations.

Train lines were introduced primarily in GB, I, B, NL, E, P PL,CZ, SK, and H. The network elements added include:

- 307 new train lines
- 765 new line routes
- 1031 time profiles
- 5555 new vehicle trips
- 736 new stations

The improvements to the impedances between NUTS3 regions has yet to be systematically investigated and will constitute a continuation of this working paper.