

# THE EUROPEAN TRANSTOOLS TRANSPORT MODEL

Otto Anker Nielsen

Professor, Ph.D.

Center for Traffic and Transport (CTT)

Technical University of Denmark (DTU)

Building 115, st.tv

2800 Lyngby, Denmark

Tel.: +45 45 25 15 14, Fax.: +45 45 93 64 12

E-mail: [oan@ctt.dtu.dk](mailto:oan@ctt.dtu.dk) - Homepage: <http://www.ctt.dtu.dk/personal/oan/index.php>

## ABSTRACT

The paper presents the use of ArcGIS in the Transtools Transport Model, Transtools, created by an international consortium for the European Commission. The model describe passenger as well as freight transport in Europe with all medium and long distance modes (cars, vans, trucks, train, inland waterways, ships and air transport).

ArcGIS was not only used for combining, improving and quality control of data, but by using the Geoprocessing Framework also as the backbone for the model complex itself. This includes a regional economic model, freight demand and trade models, a freight logistic model, passenger demand modelling, an airport choice sub-model, network assignment models for the different modes, and impact models for economics, safety, environment, etc.

Covering the entire Europe, the model is most likely the largest transport model ever built in terms of number of countries, population and area being modelled. It is also one of very few models, which both describe freight and passenger transport.

## 1. INTRODUCTION

The Transtools model is the new European Transport Model. It build heavily on ArcGIS since ArcGIS is both used for all the data handling, and the Model Builder and Geoprocessing framework are used to run the transport model.

### 1.1 Background

The Transtools<sup>1</sup> project was tendered by the European Commission to produce a new European transport network model covering passengers and freight, as well as intermodal transport, which overcomes the shortcomings of current European transport network models.

Main shortcomings in prior transport models include an unsatisfactory representation of mix of traffic (short/long distance and freight/passenger), a (partly) missing presence of intermodality and freight logistics in models, differences in implementation of Origin-Destination base year for freight traffic in some models, outdated character of some models (included both software approach, theoretical foundation and the data), no sufficient linkage of network based transport models with socio-economic effects and external effects, and limited feedback between network loads and travel speeds, and transport volumes.

As on the European realm different models for different options and with different IPR (Intellectual Property Rights) settings were anticipated, it was chosen to construct an IPR free instrument on the basis of the best available knowledge (i.e. notably at partners that have been involved in building models that involve European policy questions).

The aim with the project was to develop a European network-based transport model starting from the ideas consolidated in the modelling experience of the consortium partners. This means that some of the features of the current available EU models were used, considering that while the model cannot be a tool for every purpose, the selection of the model features should be essentially on the basis of the policy needs addressed by the European Commission services.

### 1.2 Model approaches and innovations

Approaches and components from prior models were to some extent adapted into Transtools. This included the SCENES model approach that provided good suggestions for the treatment of passenger transport and the interaction of local and long distance traffic, that NEAC provided the information for proper description of freight transport, and that the SCENES model constituted a reference for the treatment of intermodal transport, as well as SLAM for logistics. Furthermore, the CGEurope was directly included and improved to describe regional economic impacts.

It was decided to replace prior route choice and assignment models with new models developed by the Technical University of Denmark ([www.ctt.dtu.dk](http://www.ctt.dtu.dk)) together with Rapidis ([www.rapidis.com](http://www.rapidis.com))

This led to the following clear innovations obtained from Transtools:

---

<sup>1</sup> <http://www.inro.tno.nl/transtools/index.html>

- New set up of a demand/supply model;
- Intermodality for passenger/freight (as National and European transport policies promote intermodalitet through different measures);
- Inclusion of intercontinental flows (mainly for freight), as some models do not cover this segment;
- Full coverage of Central and Eastern Europe (Accession Countries and the countries at the borders of the enlarged European Union);
- Integration of the new Member Countries States at a level similar to those of EU 15 (former member countries);
- Feedback infrastructure development economy (as the question of indirect effects in the economy and on network level is important, especially where investment has a substantial influence - notably for Accession Countries);
- Logistics/freight chain explicitly included;
- Coupling method with local traffic in order to address the effect of congestion on long-distance traffic;
- The consortium provided access to all relevant experience concerning EU and national modelling;
- A software approach was chosen which results in a software modelling tool on network level.

The homepage of the project (<http://www.inro.tno.nl/transtools/Consortium.htm>) describe the consortia and background in more detail. The partners based in 4 European countries are;

| <b>Participation Role (*)</b>   | <b>Participation name</b>                                       | <b>Short name</b>    |
|---|---|----------------------|
| Coordinator   | <u>Netherlands Organisation for Applied Scientific Research</u> | <u>TNO</u>           |
| Responsible for the trade model   | <u>NEA Transport Research and Training</u>                      | <u>NEA</u>           |
| Responsible for the air network   | <u>TRT Trasporti e Territorio SRL</u>                           | <u>TRT</u>           |
| Responsible for the passenger models  | <u>Universitaet Karlsruhe (TH)</u>                              | <u>UniKarl / IWW</u> |
| Responsible for the regional economic model   | <u>Christian-Albrechts-Universität zu Kiel</u>                  | <u>CAU</u>           |
| Responsible for user perspectives   | <u>Joint Research Centre</u>                                    | <u>JRC</u>           |
| Responsible for the overall model system implementation and for the assignment models | <u>Technical University of Denmark</u>                          | <u>DTU / CTT</u>     |
| Responsible for the impact models   | <u>Istituto di Studi per l'Integrazione dei Sistemi</u>         | <u>ISIS</u>          |



includes several economic sectors and goods categories. The spatial resolution is NUTSII (294 zones). The model is created by CAU (refer to partner list in the beginning of the paper)

Secondly a freight transport model calculates changes in 1) demand, 2) distribution of trade flows, 3) logistical chains, and 4) mode chains choices. These flows are then split into uni-modal matrices used in the assignment procedures. The main model approaches are nested logit models and gravity models, although the trade model is a simultaneous model. The freight models work on a zonal structure with 1,286 zones.

The passenger model similarly calculates changes in 1) demand (total number of trips), 2) choice of destination, and 3) mode choice. The model design consisted of nested logit models and gravity models. The utility functions are non-linear reflecting the wide distribution of trip lengths in a model that covers the entire Europe. The passenger model works on a structure with 1,270 zones.

## **2.2 Assignment models**

The assignment models load the traffic onto the network, and calculate congestion levels. The road assignment models simultaneously load passenger cars and trucks in a Multi-Class Mixed Probit Stochastic User Equilibrium Procedure. Different procedures are used to distinguish between peak hours and non-peak hours. Each user class has a separate utility function, within which the coefficients follow distributions (“Mixed model”) which reflect variation of preferences and Value of Times (which are assumed Logarithmic Normal distributed). The error term follows Gamma Distributions, which avoid the truncation problem of normal distributions, yet being able to model overlapping routes and approximate the results of a Probit model due to the central limit theorem.

The rail model assigns freight trains as well as passenger trains. Since no information was available on the exact schedule, a pure stochastic (Mixed Multi-class Probit) assignment procedure was used.

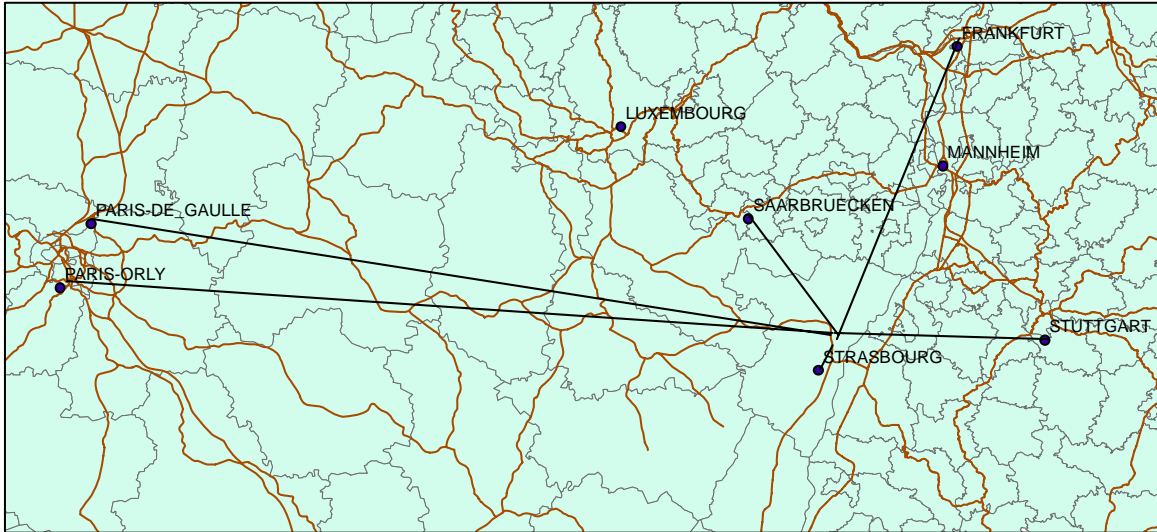
Inland waterway freight transport is assigned onto the inland waterway network by a shortest path procedure which takes ship sizes (and size of the rivers/canals) into account.

## **2.3 Air model**

The air model assigns both the air journeys in the air network, but it also calculates the choice of feeder and access link, and choices of the airports that are used for the journey.

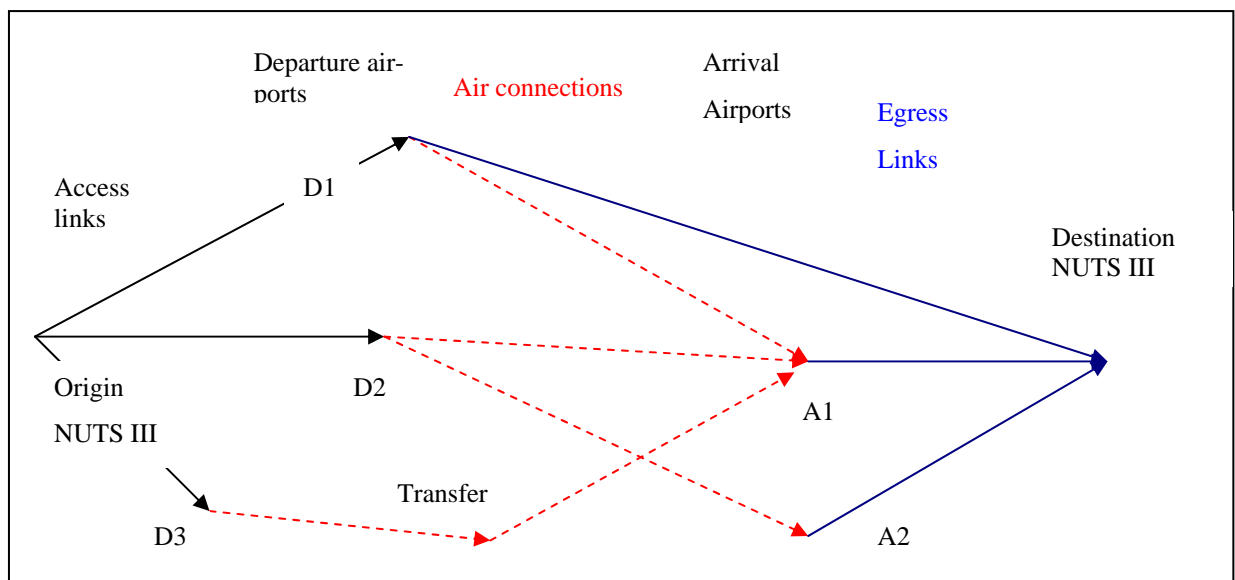
The choice process is pretty complex for airports (refer to the example in Figure 2 below). More remote regions may e.g. have access to a nearby local airport with few direct connections, and maybe high prices, other regional airports further away, but maybe with direct connects to other locations, larger hubs even further away, however maybe in a neighbour country with a more difficult land journey, and finally an even more distant national airport. The choice process is complicated due to the fact, that distance may not be proportional to travel times, e.g. whether there are motorway or not, or whether there is a direct high-speed train or a chain of slow trains must be used. The travelling times for these are calculated by the car traffic assignment models and rail models respectively, and then imported into the air model.

The choice of journeys between airports is then based on number of transfers, fare and travelling time.



**Figure 2** Example on choice-set of airport from a given NUTSII zone – in this case in East France.

The overall air model is solved by a multi-modal multi-class traffic assignment model. The value of times differ between trip purposes (business traveller have the highest VoT, vacation the lowest), logarithmic normal distribution of these (e.g. to reflect income variation and variation of willingness to pay) and an error term (again approximating a Probit model to describe overlapping routs). An example of a choice set is shown in the Figure 3 below.



**Figure 3** Example of a joint choice-set of airports and air-routes.

## **2.4 Impact models**

Finally, a set of impact models calculates transport impacts such as time savings, costs, traffic accidents, pollution, energy use, etc. Most of these were implemented in the ArcGIS Geoprocessing framework.

## **2.5 Modes and trip purposes**

As indicated above, the model describes all long- and medium distance models. Each of these is split into different user classes, i.e. trip purposes.

Freight transport is split into a number of goods categories. Available modes are trucks, rail (including rail ferries), inland waterways, and ships (Sea-born transport).

Passenger transport is split into 3 trip purposes; business, private and holiday trips. Available modes are cars, rail (including rail ferries), and air.

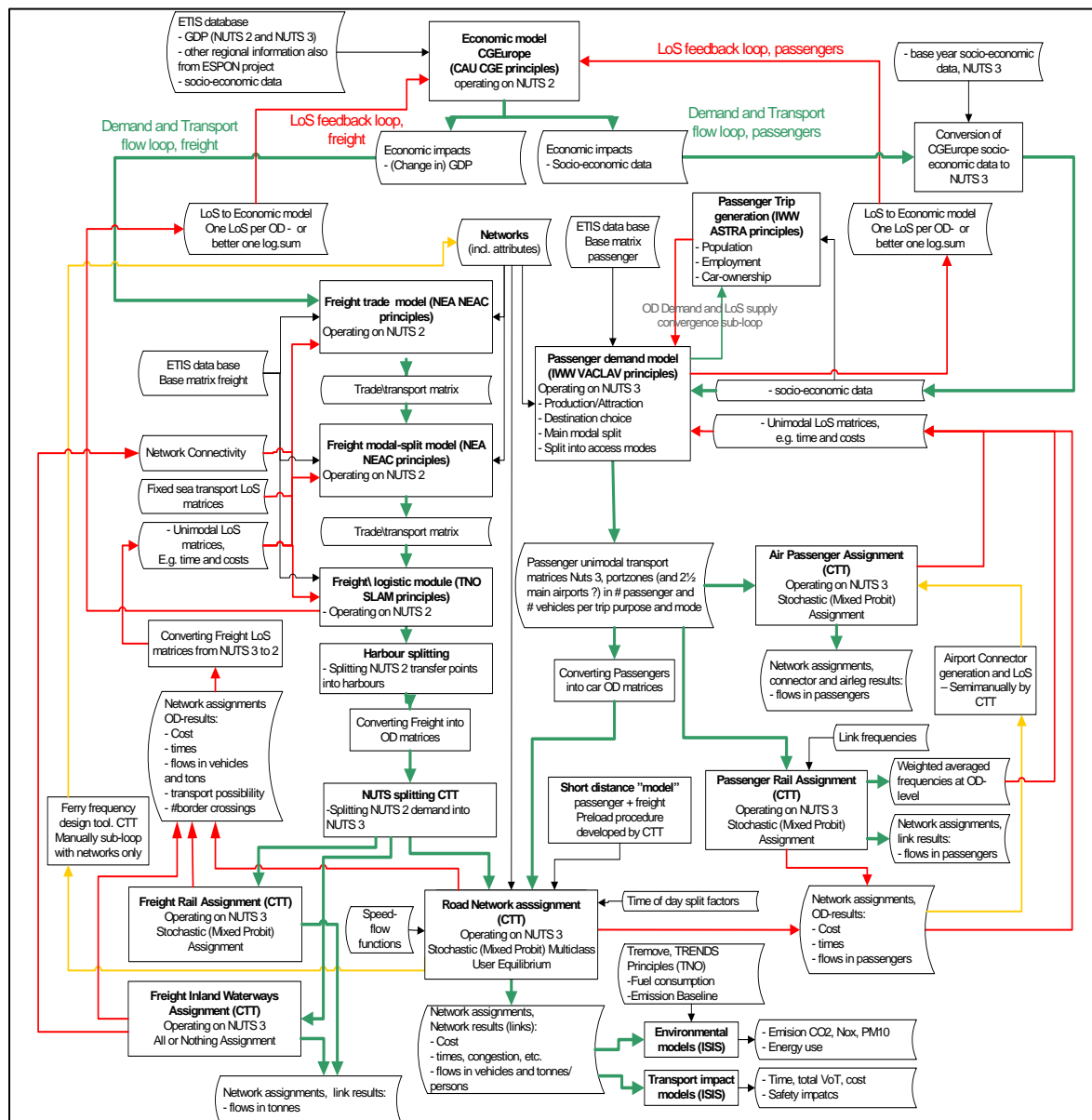
## **2.6 Linking models**

The resulting diagram of the overall model system (Figure 4), is more complex than the muck up indicated above. This is mainly because 1) the freight and passenger models include a number of sub-models, 2) the assignment models are different for each mode, and 3) the air model includes choices of airports and mode chains.

Furthermore, it is necessary with a number of data conversion steps (some of which were implemented directly by ArcGIS Geoprocessing framework). This includes disaggregation of flows from the economic model to the transport models'. And visa-versa aggregation of transport costs and times for the economic model.

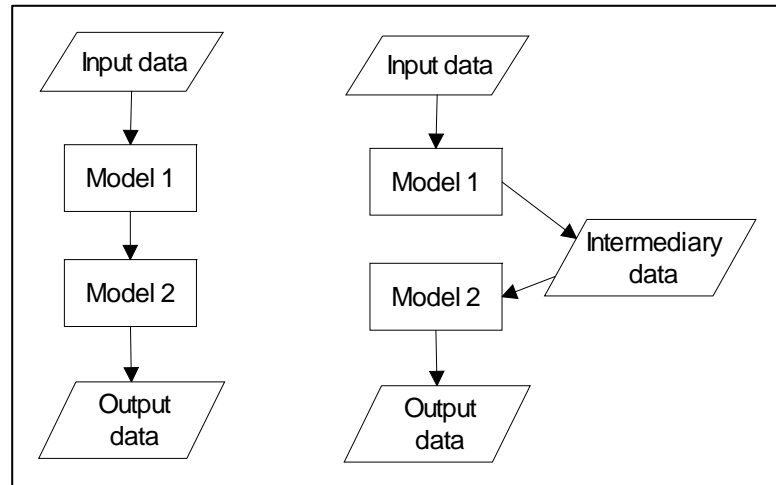
The Economic model and trade model works on monetary values (e.g. Euro), while the remaining freight models works with freight volumes (Tons). These are then remodelled to truck units before the assignment. Similarly are passengers remodelled to passenger car units before assignment.

A further complication is that trips are split up into rush-hours and non-rush hours before assignment, and that local road traffic is also assigned onto the network.





The main philosophy when building this rather large model complex was NOT to link the models directly (Figure 5, left), but to build in intermediate data storage steps (the figure on the right). This does not only allow for quality control of the results, but also makes it possible to replace and improve sub-models independently of the other models in future model updates. This is opposite many standard transport modelling software packages, where one is more or less locked to the same software for every steps.



**Figure 5** Model approach storing intermediary data

Therefore, it was also decided to store all data and parameters in databases as well. The database thereby documents the settings for model use, and it is easily accessible through the Model Builder framework. It also makes it easier to run scenarios, to validate sub-models, to speed up calculation time and to make quality control and validation easier.

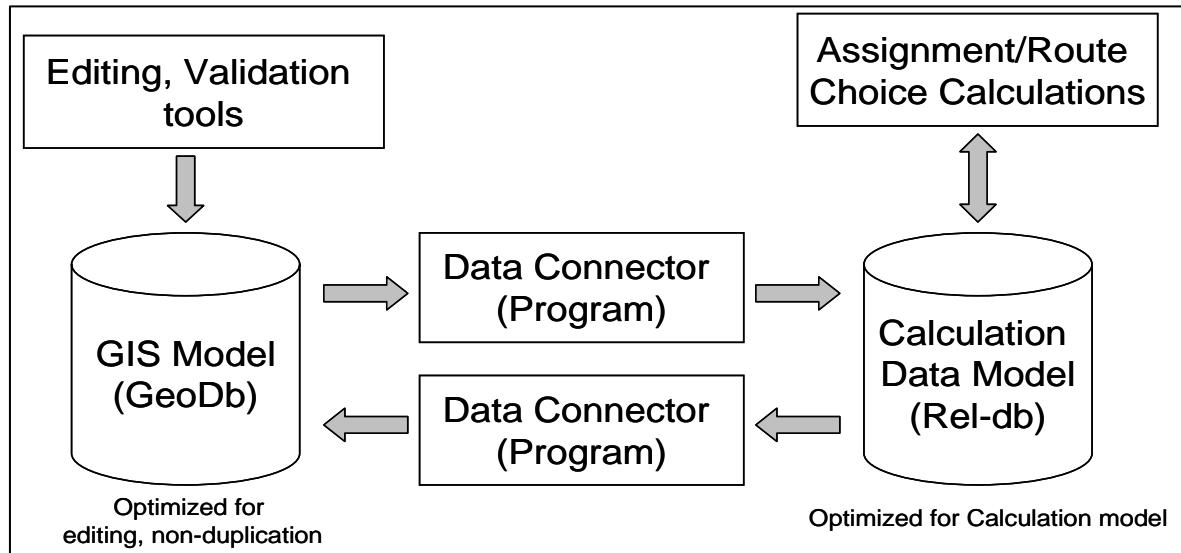
## 2.7 Implementation

In general the model is implemented with ArcGIS Model Builder as the front end interface. However, different sub-models were used, re-used, modified or built to build the whole model system. To a large extent prior models were improved and linked to save resources. Therefore, the overall model consists of sub models that are linked to the overall framework by various approaches, i.e.

- ArcGIS Geoprocessing tool
  - ▶ The data format is open, and besides the ESRI formats, it can use a wide array of other data sources, including MS Access, Oracle Database, SQL Server Database, Text files etc.
  - ▶ Most import/export routines and processing of results were implemented in ArcGIS,
- Models using the .NET component for use with C#, VB.NET or C++.NET
  - ▶ The assignment models, that were built by DTU
- COM components for use with C++ or VB 6 (or other COM-compliant languages) – some sub-modules
- Stand-alone executable programs (the SCGE model, freight model and passenger model)
  - ▶ Run by control file

- Input and output as text-files (SCGE), databases (demand models)

As it is very slow dynamically to access to data in databases, each sub-model typically imports data and builds a more flexible internal pointer structure to speed up calculation time (refer to Figure 6).



**Figure 6** Interaction between the GIS-model and the calculation model.

### 3. DATA FUSIOPN AND QUALITY CONTROL OF DATA

The preconditions for the project were, that it should utilise (reuse) data from previous modelling projects that have been made for the European Union (EU). These datasets where anticipated to have been quality controlled in prior projects. However, transferring the data to ArcGIS and quality controlling these in a GIS-environment soon revealed many problems and bugs in data. In addition to this, the assignment modelling tools revealed further data problems, especially concerning network connectivity. The experiences from this quality control are outlined in the following.

#### 3.1 Zonal database

The model covers the entire Europe, except Russia and Turkey that are treated as one-country zones. The regional economic model and trade model operates at a so-called NUTSII level, which includes 294 zones. This equals regions, i.e. a sub-division of each country.

The freight transport model works with 1,270 European zones as well as so-called port-zones, 1,286 zones in total, while the passenger transport model and assignment models operates with 1,270 zones. These are the so-called NUTSIII Zonal-level (similar to Counties).

#### 3.2 Road network

Examples of problems in the road network (CISCO version II TEN\_STAC) are:

- Important roads were missing. These were to a large extent located using alternative road maps, and by comparing these by programming a macro tool in ArcGIS.
- New motorways were missing, as well as many toll-motorways run by private firms (this is especially a problem in some European countries). These were located mainly by using the IPR-free ESRI map.
- Outdated ferry lines and new ferry lines were missing<sup>2</sup>. In some countries, ferries play a quite important role (see Figure 7).
- Dangling links and disjoint relationships at intersections. The amount of this problem was so large, that different ArcTools and ET Geowizards were used to locate and fix these errors semi-automatically. A total of 1,731 dangling links were found.



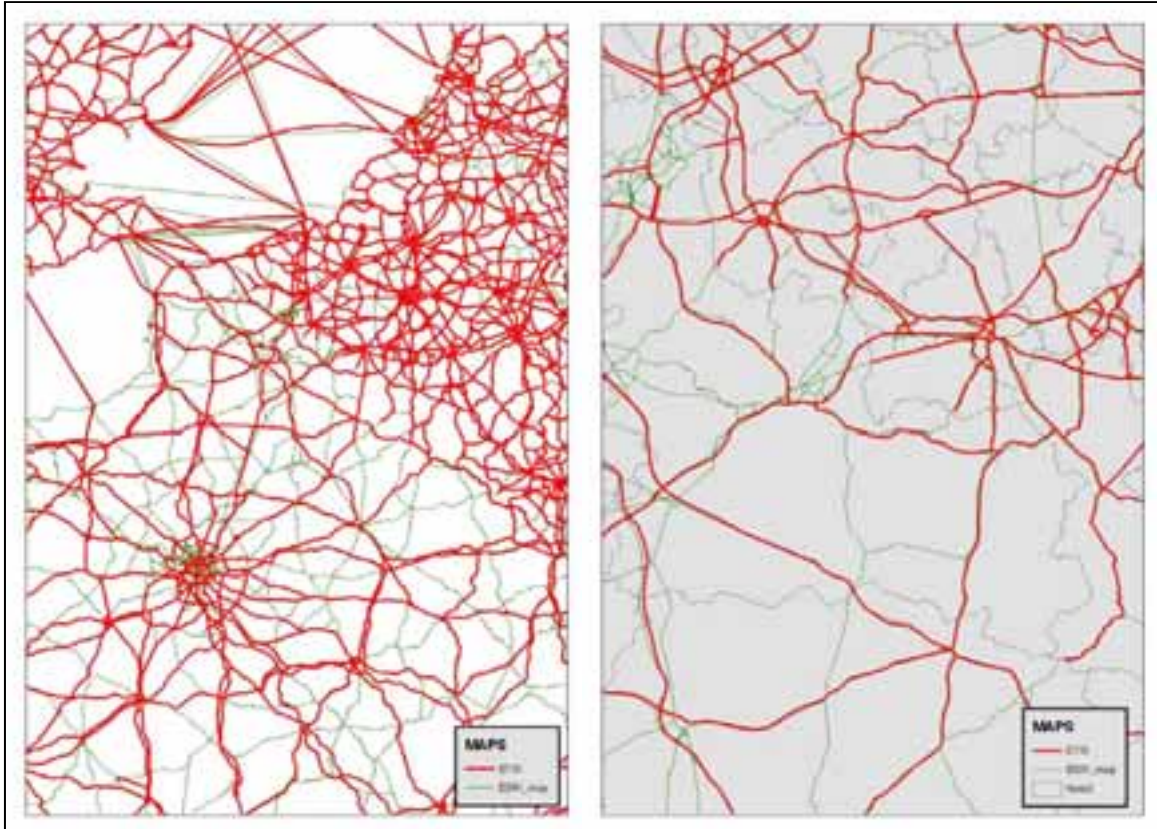
**Figure 7** Ferry links around Denmark

- Unexplained variable values. Different queries were implemented to locate these.
- Missing information on whether roads were in urban or rural areas. An GIS area overlay analyses was carried out to create this data using the ESRI sample datasets.
- Missing cost and speed.
- Strange values e.g. no. of lanes, e.g. -9.9.
- Network details varied from country to country resulting in dead ends. These were located by creating buffers along borders and analyzing non-fixable dangling nodes here (see Figure 8).

In total about 5% road links were added to the database - recall that these were mainly new motorways, and thereby major links. About 25% ferry lines were removed, and about 25% added. About 5% of the links had some problems concerning connectivity or attribute data that were corrected.

---

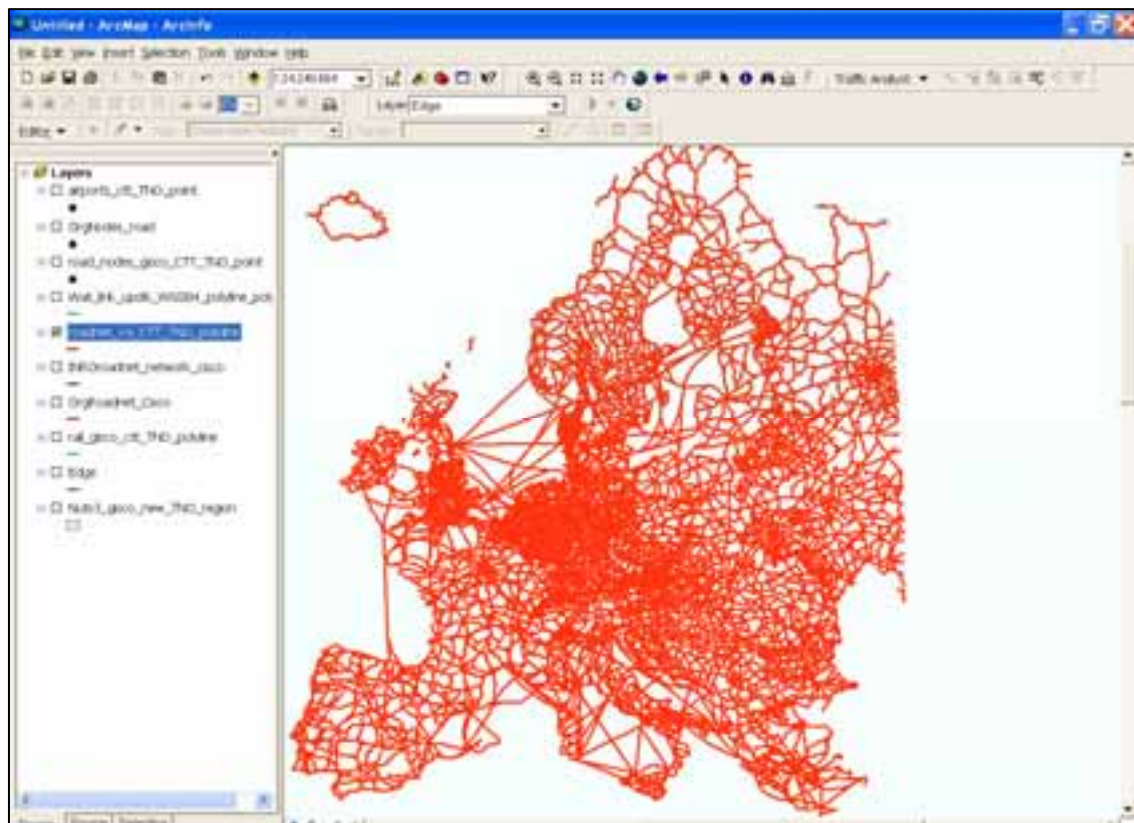
<sup>2</sup> The model considers car-ferries as part of the road network.



**Figure 8** Example on different levels of road network details along a border (here between France and Belgium + Luxemburg)

Zonal connectors were then generated from each NUTSIII zone to road intersections within the NUTSIII by some GIS-procedures. Illogical sea crossings were removed.

The road network consists of 35,079 nodes, 47,373 links (see Figure 9) and 1,265 Zonal connectors.



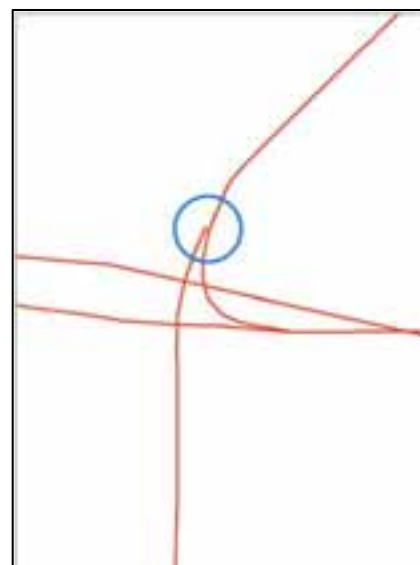
**Figure 9** Screenshot of the road and ferry network.

### 3.3 Rail network

The main problem in the rail network (GISCO version II from TEN-STAC) was that it was spatially insufficient. Therefore another network was used as source for the remaining rail links. This - however - had fewer link attributes. It was assumed, that the secondary network primarily consists of side-line, and attributes therefore were estimated. The two networks were then merged (after modifying their projections) by use of ArcGIS-procedures. Due to this, as well as due to the problems in the two networks themselves, many dangling nodes were found - again using GIS-procedures (see example in Figure 10).

After this “jumps” in rail-line classifications were pinpointed by GIS-procedures as well, whereby connectivity was secured not only in the overall network, but also in the main rail network (major lines).

Finally, rail ferries were added to create connectivity in the network. Islands with no rail were connected by normal ferries/ships to the mainland to create connectivity of the network.



**Figure 10.** Example on missing connectivity in the rail network.



Zonal connectors were then generated from each NUTSIII zone to road intersections within the NUTS by use of GIS-procedures. If no rail existed within the zone, longer connectors were created. Illogical sea crossings were removed.

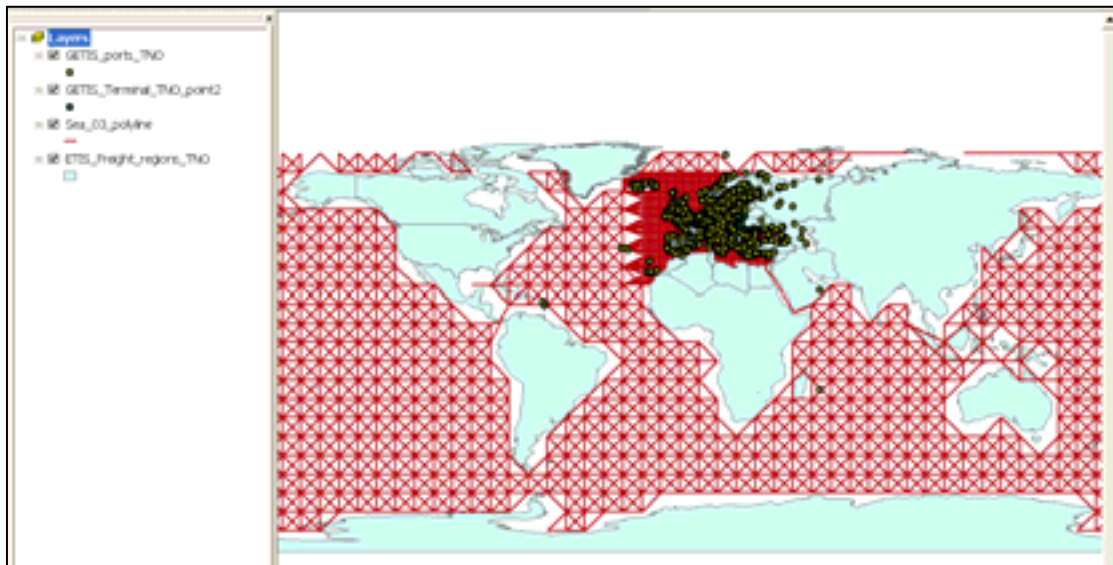
The combined network was about twice the size of the original network. About 10% of the links had some problems (like lacking connectivity), that were modified.

The rail network consists of 18,851 nodes, 19,867 links and 1,269 Zonal connectors.

### 3.4 Maritime network

The maritime network consists of a sea network (“grid-type” of network), ports and inland waterways. The “grids” in the sea network increased accuracy in the European seas as indicated on Figure 11. The main problem with the maritime network was that the network contained far too many ports. As an example in Denmark, some very small “ports” like the lake “Furesøen” lake canoe “Port” with space for 40 canoes was included in the database. Clearly in a European model, this was far too many ports. A huge reduction of the number of ports (about 90%) was then carried out mainly by using information from the freight transport model, as well as analysing ferry lines and ports with access by major roads and/or by railway. Hereby a more realistic network was created.

Only zones with harbours or access to inland waterways were connected to the maritime network, i.e. not all zones could use this mode.



**Figure 11** Grid-network for Sea transport.

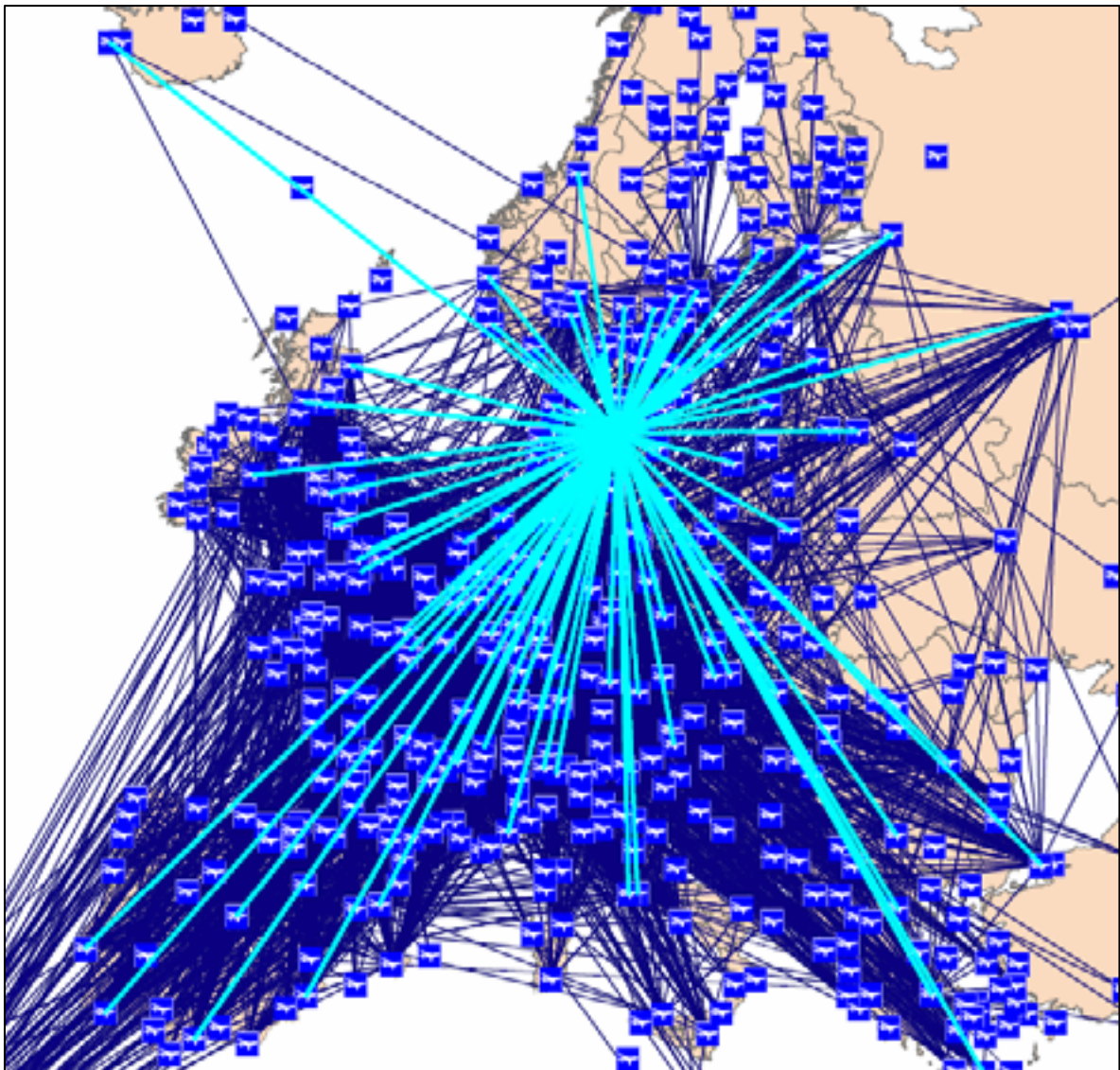
The inland waterway network consists of 747 nodes and 812 links, i.e. a very small network compared to the road and rail transportation networks.

### 3.5 Air network

The GIS-database contained a high quality map of airport locations, including the 3-figure IATA codes. This basically included all European airports with scheduled flights. I.e. both local and international airports, but not military airports and not airports with taxi-flights only. A list of air-connections (legs) was also received from a prior European transport

model. This however, did not include low-cost airlines at all, and there was missing a large amount of “conventional” flights as well.

The Transtools partner ISIS coded all low-cost airlines manually based on information from their web-pages. CTT located the 50 main air hubs in Europe (after a set of criterion). Then the partner IWW searched all possible air links to these. The assumption here was that most minor airports are connected to one or more of the major airports. Focusing on the 50 largest airports would therefore secure reasonable network connectivity by a reasonable work effort compared to a work process where all airports should be manually be quality controlled. All in all, the number of air legs was about doubled due to this quality control. Figure 12 shows an example on a selection of air links to Copenhagen Airport (Kastrup).



**Figure 12** Air transport network – with a selection of legs to/from Copenhagen Airport, Denmark (Kastrup),

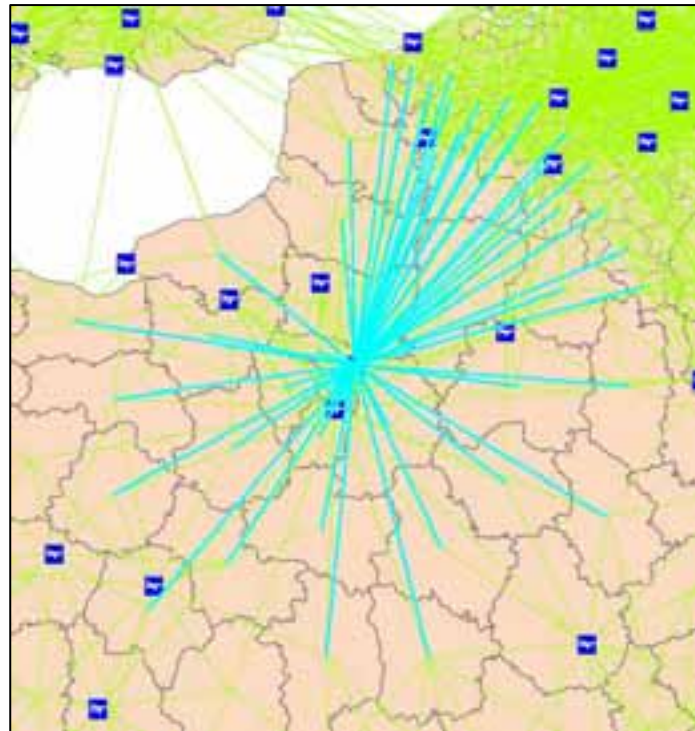
The zones were then connected to airports by various distance-based overlay analyses. It was assumed that all airports within a NUTSIII are relevant for these NUTS. However,

since the NUTSIII's are of very different sizes in Europe (e.g. very detailed in Germany, quite large in France) NUTS were also connected to all airports within a certain buffer distance. Furthermore, airports were connected to major airport hubs within large distance. The assumptions here were that passengers may be willing to travel further for a direct or cheap leg than to an airport with no direct leg, i.e. with a two-leg journey. Finally, it was assumed that travellers are willing to travel longer to national airports than to airports in a neighbour country. After this, connectors that had illogical sea crossings were removed by GIS overlay analyses.

After this process, some airport didn't have Zonal connectors, some NUTSIII were not connected to airports, and some airports did not have any flights. These – fairly few - problems were then corrected manually.

Figure 13 shows an example on connectors to an Airport – in this case an international hub (Charles de Gaulle in Paris). Note that it is assumed that some neighbour countries may use this airport using rail or car as feeder mode (but in competition with national airports).

The air network consists of 522 airports, 8,507 links and 7,962 Zonal connectors.



**Figure 13** Land transport access links to Paris Charles de Gaulle Airport.

### 3.6 Summary

The Transtools model was proposed to build exclusively on existing supposedly quality controlled traffic modelling datasets. When these were transferred into ArcGIS and visualised it soon became clear, that massive problems existed concerning coherency, connectivity and quality of network topology, as well as concerning the attribute data.

By use of existing ArcGIS tools, as well by use of some self-created tools linked together in the ArcGIS Geoprocessing Framework and Model Builder, it was possible to locate most problems automatically and also to fix most of the located problems automatically or semi-automatically. Still, the work effort was large (about a man-year) to quality control the model. But the work effort would have been ten-fold without the ArcGIS-tools. Concerning the existing Zonal connectors, they were found mainly so low quality that new connectors were created from scratch by various set of overlay analyses, and afterwards manually quality controlled.

All in all, this improved the quality of the network drastically.



#### 4. SUMMARY AND CONCLUSION

The Transtools model is according to the knowledge of the author the largest existing transport model in terms of number of countries covered, population covered, geographical scale, as well as the complete coverage of both freight and passenger transport, and both cars, trucks, trains, canal ships, sea ships and air transport.

The modelling complex links a regional economic model with demand models for freight and passenger transport. The freight model itself is especially ambitious, since it both model trade (as the model covers many countries) and since it also model logistic chains (warehousing) and mode-chains (e.g. truck-rail-truck, or truck-ship-rail).

The assignment models are also very complex, as they describe various trip purposes, equilibrate the routes concerning congestion, and describe multiple routes due to variation of preferences.

The Transtools model built on existing data, but ArcGIS has played a major role to merge different dataset, linking data and especially for quality control. The workload pin-pointing bugs and afterward correcting them would have been overwhelming – and impossible - within the budget and time-frame – if a GIS have not been used. An additional challenge was that data originated from many different countries making the merge of data more complicated.

As a new innovation in transport modelling, ArcGIS has also been used to link all the models using the ArcGIS model builder and ArcGIS Geoprocessing framework. The core models are external .exe-files. The assignment models are more closely linked with ArcGIS, since they are a modified version of the Rapidis Traffic Analyst software (refer to [www.rapidis.com](http://www.rapidis.com)). Using the ArcGIS Model Builder also helped the work process building such a complex transport model over a relatively short time-period.

Using the ArcGIS as a model interface made it much easier to illustrate, analyse and process data than in most transport models. Furthermore, the Model Builder interface can be used to run calculations as well as different scenarios. This potentially also enlarge the possible number of users. As an example, different offices and directorates of the European Commission use ArcGIS for other purposes than transport – or to analyse transport related data. They – as well as national Government and other potential users – may more easily learn Transtools if they already are ArcGIS users, than if a proprietary modelling software packages were used.

The Transtools model is now becoming the key decision support model concerning transport impact analyses conducted by the European Commission (European Union). Examples of analyses –and scenarios – are use to evaluate impacts of changes of taxation or road pricing, analyses of infrastructure improvements, analyses of future congestion in “do-nothing” scenarios, analyses of changes in GDP and population, etc,

## REFERENCES

The following list of reference documents can be downloaded from (<http://www.inro.tno.nl/transtools/Deliverables.htm>).

| Del. nr.           | Deliverable name   | WP nr             | Lead participant        |
|--------------------|--|-------------------|-------------------------|
| <a href="#">D1</a> | <a href="#">Report on policy requirements and selected models</a>  | <a href="#">1</a> | <a href="#">TNO</a>     |
| <a href="#">D2</a> | <a href="#">Report on describing the datasets needed for model estimation and calibration, supported by databases</a>  | <a href="#">2</a> | <a href="#">TRT</a>     |
| <a href="#">D3</a> | <a href="#">Report on specification and calibration results</a>  | <a href="#">3</a> | <a href="#">NEA/IWW</a> |
| D4                 | TRANS-TOOLS software Tool with user and technical documentation  | 4                 | CTT                     |
| D5                 | Report on Scenarios Model Outcome of Scenarios and Policy Assessment   | 5                 | TNO                     |
| D6                 | Final report, derived from the WP reports  | 6                 | TNO                     |
| D7                 | Final Conference (in conjunction with another project: FP 5 TN/FP6 TA)   | 6                 | TNO                     |
| D8                 | Final plan for using and disseminating knowledge   | 6                 | TNO                     |
| D9                 | Website / information leaflet<br>- <a href="#">Logo's TRANS-TOOLS project (19kb)</a><br>- <a href="#">Newsletter 1 on the project - October 04 (367kb)</a><br>- <a href="#">Leaflet 2 on the project - June 05 (1528kb)</a><br>- <a href="#">Leaflet 3 on the project - February 06 (1916kb)</a> | 6                 | TNO                     |