Ontwikkeling van een multimodaal goederenvervoermodel voor Vlaanderen

Fase 1: Tripgeneratie- en Distributiemodel

September 2006

Opdrachtgever:
Vlaamse overheid
Mobiliteit en openbare werken
Verkeerscentrum
Vuurkruisenplein 20
2020 ANTWERPEN

Opdrachtnemer:
K&P Transport Consultants
Habsburgerstrasse 5
79104 FREIBURG

Tritel
Transport Infrastructure & Telematics
Louizastraat 40b
2800 MECHELEN
1. GENERATION MODEL

1.1. Purpose and principle

This generation step aims at calculating the global volume of goods that is emitted and attracted by each European zone, differentiated by kind of goods (10 NST categories):

- The traffic emitted by one zone corresponds to the sum in tons of all the shipments transported from the different sites of production or warehouses located in the regarded zone.

- The traffic attracted by one zone corresponds to the sum in tons of all the shipments transported to the different sites of production or warehouses located in the regarded zone.

The zonal system that is used for the generation model is the base zoning, that is to say the NUTS3 level in Belgium (43 arrondissements).

Figure 1.1 : base zonal system
The general principle of the modelisation consists in considering that the emissions and attractions are directly linked to the intrinsic economic characteristics of the regarded zones expressed in socio-economic figures.

The purpose of the generation model is to explain the interdependence between socio-economic data and emitted/attracted volumes with a mathematical function. Obviously, this function is not the same for the productions and for the attractions, and it also differs for each commodity. Consequently, the model must be differentiated by direction (emission/attraction) and by commodity.

1.2. Methodology

Since it is very difficult to obtain socio-economic data for the whole Europe, these equations are only used to calculate the Belgian domestic emissions/attractions (i.e. from Belgian zones to Belgian zones). The other flows are calculated by using annual growth rates that differ between the kind of goods and the kind of traffic.

The following table sums up the methods that have been used to calculate all the emissions and attractions of the model.

Table 1.1: Methods for calculating emissions and attractions

<table>
<thead>
<tr>
<th>Zone</th>
<th>Type of flow</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>Domestic</td>
<td>Linear formulations</td>
</tr>
<tr>
<td>Belgium</td>
<td>Import and Export</td>
<td>Use of average annual growth rates</td>
</tr>
<tr>
<td>Abroad</td>
<td>Import, Export and Transit</td>
<td>Use of the BIP’s growth rates of the destination countries and use of attraction elasticities</td>
</tr>
</tbody>
</table>

These three methods are described in the following sections.

The model structure enables more than just the pure generation. It actually enables to:

- grow the socio-economic data (population and employment) until the selected forecast year.
- take the evolutions of the productivity and the consumption of households into account.
- correct some localised errors after having modelled the Belgian emissions and attractions, either to correct some figures in the base situation or to take specific forecasted changes into account.
1.3. **Domestic emissions and attractions**

**PRINCIPLE**

The modelling of the domestic emissions and attractions is based on linear formulations that take the socio-economic context into account. Thus, the evolution of the emission and attraction figures is directly linked to the evolution of the socio-economic data, at least after the first step of the modelisation. Some other parameters like the productivity or the households consumption are included to balance this close link and hereby avoid some misinterpretations. Eventually, a multiplicative pivot is applied to the base emissions and attractions.

**EVOLUTION OF THE SOCIO-ECONOMIC DATA**

As far as a future situation is concerned, the first step of the forecast calculation in the generation model consists of estimating the evolution of the socio-economic data. This evolution is based on a function using the Average Annual Growth Rate (AAGR) of each variable. The following equations are used:

\[
\text{Evolution of the population}: \\
\text{Population}_{\text{Scenario}} = \text{Population}_{\text{Base}} \cdot \left(1 + \frac{T_{\text{pop}}}{100}\right)^n
\]

\[
\text{Evolution of the employment (per employment category)}: \\
\text{Employment}_{\text{Scenario}} = \text{Employment}_{\text{Base}} \cdot \left(1 + \frac{T_{\text{empl}}}{100}\right)^n
\]
where:

\[ Population_{\text{Scenario}} = \text{Population in scenario situation} \]

\[ Employment_{\text{Scenario}} = \text{Employment in scenario situation per employment category} \]

\[ Population_{\text{Base}} = \text{Population in base situation} \]

\[ Employment_{\text{Base}} = \text{Employment in base situation per employment category} \]

\[ n = \text{Number of years between base situation and scenario situation} \]

\[ T_{\text{pop}} = \text{AAGR of the national population} \]

\[ T_{\text{empl}} = \text{AAGR of the employment per employment category} \]

**PRODUCTIVITY AND HOUSEHOLDS CONSUMPTION**

The socio-economic data must be corrected to take into account evolutions that could affect the volumes of goods emitted and attracted. Actually, a diminution of the number of employees in a sector does not necessarily result in a diminution of the corresponding volumes, even on the contrary in some cases. As we cannot get the average annual growth rates of the volumes emitted and attracted by each employment category – which would be optimal for the forecast side of the model – the productivity and the households consumption are introduced. The correction functions are the following:
Socio-economic data linked to the population:

\[
Var(\text{Pop})_{\text{Scenario}} = Population_{\text{Scenario}} \cdot \left(1 + \frac{T_{\text{cons}}}{100}\right)^n / \left(1 + \frac{T_{\text{ratio}}}{100}\right)^n
\]

Socio-economic data linked to the employment (per employment category):

\[
Var(\text{Empl})_{\text{Scenario}} = Employment_{\text{Scenario}} \cdot \left(1 + \frac{T_{\text{prod}}}{100}\right)^n / \left(1 + \frac{T_{\text{ratio}}}{100}\right)^n
\]

where:

\(Population_{\text{Scenario}}\) = Population in scenario situation

\(Employment_{\text{Scenario}}\) = Employment in scenario situation per employment category

\(Var(\text{Pop})_{\text{Scenario}}\) = Socio-economic variable linked to the population in scenario situation

\(Var(\text{Empl})_{\text{Scenario}}\) = Socio-economic variable linked to the population in scenario situation per employment category

\(n\) = Number of years between base situation and scenario situation

\(T_{\text{cons}}\) = AAGR of the households consumption

\(T_{\text{prod}}\) = AAGR of the productivity per employment category

\(T_{\text{ratio}}\) = AAGR of the ratio value/ton:

- per employment category for the employment
- global for the population
LINEAR FORMULATIONS OF THE EMISSIONS AND OF THE ATTRACTIONS

The socio-economic data that have been calculated in the last step are now used in the generation model.

The generation model results of the estimation of the following generation functions:

\[
\text{Emitted traffic : } EM_{[\text{zone } i, \text{ NST } x, t_0]} = f_{x-e}[\text{socio-eco } (i,k,t_0), \beta_{e}(x,k)] \\
\text{Attracted traffic : } AT_{[\text{zone } i, \text{ NST } x, t_0]} = f_{x-a}[\text{socio-eco } (i,k,t_0), \beta_{a}(x,k)]
\]

where:

\[
EM_{[i, x, t_0]} \quad (AT_{[i, x, t_0]}) = \text{Annual volume in tons emitted (attracted) by the zone } i, \text{ for the category of goods } x, \text{ at the instant } t_0
\]

\[
\text{socio-eco } (i, k, t_0) = \text{Socio-economic variables (k variables) standing for transport generating activities in the zone } i, \text{ at the instant } t_0
\]

\[
\beta_{e}(x,k) \quad (\beta_{a}(x,k)) = \text{Parameter applied to each variable (k variables) in the equation of generation, for the category of goods } x, \text{ for the emission (attraction)}
\]

\[
f_{x-e}[...] \quad (f_{x-a}[...]) = \text{Function of generation of the volume of goods emitted (attracted) per commodity}
\]

The equations of emission and attraction are different for each commodity because they obviously do not depend on the same socio-economic variables (e.g. the agricultural products generated by one zone are linked to the agricultural figures of this zone, but the emission of metal products depends on the metal industry of the concerned zone).

The following table shows the socio-economic data that have been used in the model. These are population and employment figures split in different economic activities (NACE typology), available for Belgium on the NUTS3 level (arrondissements) for base year 2004.
Table 1.2: Socio economic base data used in the generation model

<table>
<thead>
<tr>
<th>POPULATION 2004</th>
<th>EMPLOYMENT 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP</td>
<td>Population</td>
</tr>
<tr>
<td>A0</td>
<td>Agriculture</td>
</tr>
<tr>
<td>B0</td>
<td>Agricultural and food industry</td>
</tr>
<tr>
<td>C1</td>
<td>Clothing industry</td>
</tr>
<tr>
<td>C2</td>
<td>Printing industry</td>
</tr>
<tr>
<td>C3</td>
<td>Pharmaceutical industry</td>
</tr>
<tr>
<td>C4</td>
<td>Household equipment</td>
</tr>
<tr>
<td>D0</td>
<td>Automotive industry</td>
</tr>
<tr>
<td>E1</td>
<td>Shipbuilding, aircraft industry, railway equipment industry</td>
</tr>
<tr>
<td>E2</td>
<td>Mechanical industry</td>
</tr>
<tr>
<td>E3</td>
<td>Electric and electronic equipment</td>
</tr>
<tr>
<td>F1</td>
<td>Minerals</td>
</tr>
<tr>
<td>F2</td>
<td>Textile industry</td>
</tr>
<tr>
<td>F3</td>
<td>Wood and paper industry</td>
</tr>
<tr>
<td>F4</td>
<td>Chemical industry</td>
</tr>
<tr>
<td>F5</td>
<td>Metal industry</td>
</tr>
<tr>
<td>F6</td>
<td>Electric and electronic industry</td>
</tr>
<tr>
<td>G1</td>
<td>Production of combustible and fuels</td>
</tr>
<tr>
<td>G2</td>
<td>Water, gas, electricity</td>
</tr>
<tr>
<td>H0</td>
<td>Construction industry</td>
</tr>
<tr>
<td>J2</td>
<td>Wholesale</td>
</tr>
<tr>
<td>K0</td>
<td>Transport</td>
</tr>
<tr>
<td>R0</td>
<td>Other service industry</td>
</tr>
<tr>
<td>S0</td>
<td>Administration</td>
</tr>
</tbody>
</table>

The parameters $\beta(x,k)$ that are used in the generation function have been calculated by using the least squares method. They fit the observed figures of emission and attraction and correspond to logical links between fields of activities and commodities. They are differentiated by commodities and by direction. These parameters are available in the annex.
USE OF THE MULTIPLICATIVE PIVOT

Once these steps have been achieved, a pivot is applied on the base emissions and attractions, that is to say on the observed emissions and attractions for base year 2004, which leads to:

- In current situation (2004), the pivot results in the observed emissions and attractions.
- In forecast situation (scenario), the pivot multiplies the observed emissions and attractions by the evolutions calculated between the current synthetic emissions/attractions and the forecast synthetic emissions/attractions of the corresponding scenario. This can be accounted for by the following scheme:

Figure 1.2: principle of the multiplicative pivot
1.4. Importations and exportations of the Belgian zones

The importations and the exportations of the Belgian zones in scenario situation have been calculated by introducing annual growth rates which are based on the following parameters:

- Average Annual Growth Rate of the importations or exportations per NST category of goods.
- Average Annual Growth Rate of the ratio value/ton of the importations or exportations per NST category of goods.

1.5. Emissions and attractions of the foreign zones

The generation step does not take into account the emissions and attractions between foreign zones (i.e. transit). These flows are only calculated in the distribution model, which directly follows the generation model.

However, the growth of the attraction of the foreign zones is taken into account in the generation model, so as to limit the flows proportionally to the incoming capacity of their destination. The annual growth rates of the BIP of the principal destination countries are used in conjunction with elasticities on the attractions of the latter.

The emissions of the foreign zones do not need to be restricted of even increased as the distribution model bases the concerned flows on the attractions of Belgium.

1.6. Correction of punctual anomalies

As it is based on the socio-economic data, the generation model assumes that these data can result in a realistic transport activity.

Even if this assumption is very often true, it may happen that some anomalies appear in some zones. In this precise case, the transport activity cannot be completely accounted for by socio-economic data. The following reasons can for instance lead to this problem:

- Concentration of industries with high productivity rates (storage centers, logistic activities...)
- Unique big structures in few populated zones / zones with low employment figures (metal factory, productive industry...)
- Implantation of new resident areas or enlargement of cities / industries

These anomalies can be directly compensated by integrating additional volumes to the original emissions or attractions of the model (not only in Belgium but also abroad), either for the base case (2004) or in scenario situation (one separate file for each scenario).
1.7. Calibration

The following graphics give the results of the calibration of the attractions (figure 1.3) and of the emissions (figure 1.4). One can clearly see that the results are satisfying, even before the application of the multiplicative pivot.

Figure 1.3: Comparison between estimated and observed attractions (in million tons)

![Figure 1.3](image)

\[ y = 1.3172x + 815672 \]
\[ R^2 = 0.6966 \]

Figure 1.4: Comparison between estimated and observed emissions (in million tons)

![Figure 1.4](image)

\[ y = 1.1301x + 2E+06 \]
\[ R^2 = 0.6208 \]
2. DISTRIBUTION MODEL

2.1. Purpose and principle

The distribution step aims at modeling all the flows between the different zones of the study, and not only the emission and the attraction of each zone. Actually, these data that have been calculated thanks to the generation model will now be used to get the origin-destination (O-D) flows on the NUTS3 level (43 arrondissements in Belgium).

A gravity model is used to distribute these flows on the whole zoning. The results are provided for each kind of goods by O-D matrices in tons and for the sum of all modes (the mode differentiation follows in the modal split model). More information about the gravity model is given further.

Eventually, the transit flows are also taken into account in this step, even if they are not calculated with a gravity model.

2.2. Distribution of the different kinds of traffic

The domestic flows and the international flows are distributed in a different way:

- The domestic flows are distributed thanks to a gravity model. This type of model uses an exponential resistance function depending on the squared distance in our case. It can also depend on the times of transport, on the costs or on the generalised costs.

- The import and export flows are elaborated in the same way with a gravity model that distributes the emitted and attracted volumes of Belgium to/from the foreign zones.

- The transit flows are calculated with growth rates applied to the current observed flows.

As a result, the distribution model provides matrices of flows per kind of goods.

2.3. Modeling process

The modeling of the distribution flows takes place in the following steps:

- distribution of the domestic flows
- distribution of the exportations
- distribution of the importations
- growth of the transit flows thanks to growth rates
- use of a pivot on the base matrices (observed situation 2004)

DISTRIBUTION OF THE DOMESTIC FLOWS, THE EXPORTATIONS AND THE IMPORTATIONS
For each kind of flow (domestic, import or export), the distribution is calculated with a gravity model corresponding to the following formulation:

\[ T_{ij, x} = \alpha_i, x \cdot \beta_j, x \cdot E_i, x \cdot A_j, x \cdot F(D_{ij}) \]

where:
- \( x \) = NST category of goods
- \( T_{ij, x} \) = Flow of goods, in tons, from the zone \( i \) to the zone \( j \)
- \( \alpha_i, x, \beta_j, x \) = Coefficients enabling to equilibrate the emissions/attractions in the end of the process
- \( E_i, x \) = Emissions in tons of the zone \( i \) *
- \( A_j, x \) = Attractions in tons of the zone \( j \) *
- \( F() \) = Resistance function
- \( D_{ij} \) = Distance on the road between the zones \( i \) and \( j \) *

(* calculated by the model)

The proportionality coefficients \( \alpha_{NST} \) et \( \beta_{NST} \) enable to respect the balance of the attractions and the emissions in the end of the iterative process.

The resistance function is an exponential formula which looks like following:

\[ F(D_i) = \exp(-p_x.(1+\frac{x}{100})^\gamma \cdot D_{ij}) \]

where:
- \( x \) = NST category of goods
- \( F() \) = Resistance function
- \( D_{ij} \) = Distance on the road between the zones \( i \) and \( j \) *
\[ \rho \] = Fitting parameter per NST category of goods
\[ \gamma \] = Annual growth rate applied to the fitting parameter per NST category of goods
\[ y \] = Number of years between base situation and scenario situation

(*) calculated by the model

The parameter \( \gamma \) enables to model an evolution of the average length of the goods flows. However, it is possible to keep it equal to zero. In this case, one considers that the distribution of the volumes in base situation and in scenario situation goes the same way.

Once all the kinds of flows have been distributed, the model provides matrices of flows that include the domestic exchanges as well as the importations and the exportations for each NST category of goods.

GROWTH OF THE TRANSIT FLOWS

The transit flows in scenario situation are calculated thanks to annual growth rates which are estimated from:

- BIP elasticities on the attractions for each NST category.

USE OF THE PIVOT

This step consists in applying a multiplicative pivot to the base matrices. The process is quite similar to the one of the generation model:

- In base situation, the base matrices can directly be used.
- In forecast situation (scenarios), we apply to the base matrices the evolutions calculated between the current synthetic matrices and the forecast synthetic matrices.

2.4. Calibration of the distribution model

The model has been calibrated through a comparison between the results of the model and the data of the base matrices on the following points:

- Difference between the desired and the modeled attractions for each zone.
- Distribution of the flows in distance classes.

ANALYSIS OF THE CONVERGENCE (RMSE)
The analysis of the convergence of the iterative process does also enable to validate the calibration. The method consists in calculating the value of the Root Mean Square Error (RMSE) for each iteration until the best result has been reached.

\[
RMSE_i = \sqrt{\frac{\sum (F_i - F_{\text{arget}})^2}{i-1}}
\]

where:
- \( RMSE \) = Root Mean Square Error
- \( i \) = Number of iterations that have been run
- \( F_i \) = Modeled flow at iteration \( i \)
- \( F_{\text{arget}} \) = Observed flow

The RMSE enables to analyse how the attractions differ between the model and the base matrices, as the distribution process is based on variations of the attractions. Actually, the emissions are considered as basis and the attractions are modified at each iteration to distribute the flows while keeping a symmetrical matrix.

In our case, the distribution results in very good values of RMSE in so far as it is equal to zero for each NST and for each kind of flow (domestic, import and export). That means that the matrices converge well and that the final attractions correspond to the observed ones.

DISTRIBUTION OF THE FLOWS IN DISTANCE CLASSES

The values given by the RMSE enabled to test the convergence of the model and the quality of the modeled emissions and attractions in comparison to the observed ones. The second phase of the calibration does now consist in testing the distribution in terms of separate flows, and no more in terms of emissions and attractions.

The comparison of the observed and modeled flows by distance classes is a good indicator to test if the flows have been correctly distributed and correspond to the reality. The following graphic shows this comparison:

Figure 2.1 : Comparison between observed and estimated flows per distance classes, before calibration of the intra-zonal traffic
The difference between the observed and modeled figures is quite low, except for the distance class 0-10 kilometers, which can be accounted for by the fact that the intra-zonal traffic has been increased by a factor of 2.5 in the observed matrices. If the modeled figures are multiplied by the same factor, the difference gets much lower (figure 2.2).
Figure 2.2: Comparison between observed and estimated flows per distance classes, after calibration of the intra-zonal traffic.