Demand modelling by combining disaggregate and aggregate data: an application to a displacement scenario for the Mount Vesuvius area

S. de Luca

Department of Civil Engineering, University of Salerno

Abstract

The event of an eruption of Mount Vesuvius (dormant since 1945) represents an emergency involving approx. 700,000 persons. In order to face the risks of its catastrophic awakening, Region Campania and the University of Naples have cofinanced a study aiming at simulating the impacts that a possible displacement of approx. 350,000 residents may produce both on the territory and the transport system. To this purpose, a model system has been drawn up to simulate the demand of mobility according to different displacement scenarios in terms of housing and business activities as well as in relation to a different hypothesis of reorganization of the regional public transport system. The following paper introduces a disaggregate/aggregate mixed calibration methodology, where different typologies of data are put together thus obtaining demand models which result “inexpensive” and efficacious on simulating the phenomenon.

Keywords: demand modelling, discrete choice models, aggregate/disaggregate calibration.

1 Introduction

Demand analysis is usually based on discrete choice models that can be calibrated using disaggregate or aggregate data. The former approach requires specific surveys in order to specify and calibrate the models but allows more detailed utility specifications, including disaggregate level of service attributes and socio-economic ones. The latter uses aggregate information such as origin-destination flows or traffic counts, it is cheaper but leads to simple
specifications. The choice between the two approaches depends on funds available. Such a problem is much more relevant for long term prediction and when the study area involves many municipalities, with different dimensions, transport supply services and geographical morphologies. A useful compromise may be obtained by resorting to both calibration methods: disaggregate calibration for the drawing up of demand models that allow a preliminary interpretation of the phenomenon in terms of parameters values and cross-relations among the same parameters; aggregate calibration for adjusting disaggregate models in order to allow their extension to the whole area concerned. The aim of this paper is twofold:
(a) to present a methodology to approach demand model calibration, the aim is to demonstrate how different types of data can be combined in order to obtain efficient and cheaper demand models;
(b) to apply such methodology to a displacement scenario, in order to show that its impacts are sustainable both as concerns the capacity of the transport system and the environment.

Three data sources have been used: revealed preference survey on a sub area of the study area, origin-destination (O-D) flows and road traffic counts. The revealed preference survey has been used to calibrate preliminary mode choice and distribution models (MultiNomial Logit), the O-D flows and the traffic counts have been used for a double correction of the parameters of both models.

2 The case study and the displacement strategy

Mount Vesuvius overlooks Naples and its suburbs and, though a dormant volcano for about 60 years, is expected to resume its activity sooner or later by erupting as ruinously as happened in 79 AD when it destroyed the city of Pompeii. In the past 60 years, despite such a huge risk, the slopes of the mount have been densely built-up thus leading to an illogical growth of resident population. Approx. 700,000 inhabitants are currently expected to be evacuated in the event of an eruption and to that purpose emergency evacuation plans have been drawn up.

However, in order to relieve pressure in the area, long-term decongesting policies are required displacing people and economic sectors. A reasonable measure to reach that target might consist in gradually moving part of the population to the neighbouring areas (Greater Naples). That would minimize (i) the temporary transfer of the residual population to Naples hinterland (not to other Italian regions as provided for by current evacuation plans), (ii) the global risk and (iii) would allow the evacuation to be carried out more rapidly. From this point of view, by focussing our attention on Greater Naples (see Figure 1), we have assumed to displace 350,000 persons to 29 municipalities to the north of Naples, excluding the municipalities bordering the cities of Naples and Caserta and by taking for granted that the future socioeconomic (ratios) framework remains unchanged with respect to the latest available census (details in de Luca and Ricci [4]). Along with the socioeconomic reorganization, we have also assumed the expansion of the public railway transport supply by (a) putting into
service three regional underground lines; (b) doubling the existing railway service frequency; (c) finishing the construction of the High Speed railway line.

![Map of Greater Naples, traffic zones and displacement area.](image)

Figure 1: Greater Naples, traffic zones and displacement area.

## 3 The methodology

As mentioned before we have developed a model system apt to simulate the transportation demand generated from each traffic area, its geographic distribution and its market shares through the transport modes available.

If $d_{o,d}[s,h,m]$ is the average number of round trips undertaken by the generic individual $i$ between the zone of residence $o$ and the destination $d$, for purpose $s$ in the reference period $h$, with mode $m$, the demand model has been implemented following the classical partial share approach:

$$d_{o,d}[s,h,m]=n'_i[o] \cdot m'[osh] \cdot p'[d/osh] \cdot p'[m/oshd]$$

where $n'_i[o]$ is the number of individuals belonging to category $i$ in zone $o$, $m'[osh]$ is the mean number of trips undertaken by the individual in category $i$, $p'[x/osh]$ is the trip frequency model, $p'[d/osh]$ is the distribution model, $p'[m/oshd]$ is the mode choice model.

Each partial share (except emission share) has been modelled by random utility theory (Cascetta [2]), and its parameters have been calibrated by sequentially combining disaggregate and aggregate data. The aim of such procedure has been to obtain a system of models effective in reproducing trip distribution and mode choices, and based on different existing classes of data, individually not sufficient to calibrate any statically significant model (Cascetta [2], Cascetta and Russo [3] as regards parameters calibration through traffic counts).
As regards disaggregate data, the results of two preceding surveys on residents’ mobility have been used, focussing on the people’s behaviour when choosing their destination and transport mode. In particular, 650 interviews (ITER/650) have been taken from a survey made in the city of Naples (1,100,000 inhabitants) in the 1996, and 1,000 interviews (ITER/1,000) have been taken from a survey made in ten cities of Naples suburbs (2,400,000 inhabitants). Aggregate data have been extracted from three classes: (i) origin-destination matrices as result from ISTAT decennial census about Italian socioeconomic framework and concerning people’s mobility within each Italian municipality; (ii) Naples municipality transport plan matrices relating to its outgoing and incoming mobility (MTP [1]); (iii) surveys of vehicle flows along specific links of the road network, aiming at testing the validity of origin-destination matrices (FLOWS).

The area surveyed has been divided into two sub-areas, the former corresponding to the city of Naples, and the latter to its suburbs. As regards the city of Naples, the survey ITER/650 has been utilized to calibrate a destination choice model and a mode choice model (towards Naples suburbs) as adopted by Naples’ residents.

As regards Naples suburbs, the data resulting from the interviews (ITER/1,000 partially represents the whole area) have been used for a preliminary calibration of mode and distribution models. In particular, the 1,000

Figure 2: Calibration methodology (for Naples suburbs).
interviews have been used for a preliminary calibration of distribution choice models, while the same data, added to the 650 interviews mentioned above, have been used for a preliminary calibration of mode choice models (ITER/1,650).

Starting from the results of the previous calibrations we have updated the model parameters by means of the aggregate data as above mentioned (see Figure 2). As regards the mode choice models we have used the modal split resulting from the demand matrices included in the MTP. As concerns the trip distribution model we have carried out a double update, trying to reproduce the percentage of trip distribution resulting from the origin-destination matrices (ISTAT for constrained purposes and MTP for all the other purposes) and the vehicle link flows as surveyed along the road network (FLOWS).

In applying the methodology home-based trips have been simulated, and the supply-demand interaction has been simulated with respect to the morning peak hour (7.30-9.30). The demand has been analyzed by identifying two trip purposes: commuting trips and other purposes trips. The general layout of the model system and the methodology adopted are detailed in the paragraphs below.

4 The demand models

4.1 Mode choice models

The transport mode choice models are MultiNomial Logit one providing three transport mode alternatives: car, transit, motorbike. The procedure of mixed (disaggregate/aggregate) calibration can be summed up as follows.

1st step – parameters calibration through disaggregate data (data: ITER/1,650)

For each trip purpose a MultiNomial Logit model has been calibrated. As a subsequent updating of parameters was expected by means of aggregate data, level of service and geographic variables have been used.

2nd step – parameters update by aggregate data (data: MTP or ISTAT)

The parameters resulting from the previous models have been updated by means of transport mode demand flows provided by ISTAT (for home-work trips) and by MTP (for all other purposes). The problem has been addressed minimizing the following Non-Linear Generalized Least Square estimator (NLGLS):

\[ \beta^* = \arg\min_{\beta \in S_\beta} \left[ \sum_i \frac{(b_i - \beta_i)^2}{\text{Var}[\sigma_i]} + \sum_{odm} \left( \%_{odm}(b) - \%_{odm} \right)^2 \gamma \right] \]

where:
- \( \beta_i \) is the parameter \( i \) estimated in the 1st step, \( b_i \) is the corresponding unknown parameter;
- \( b \) is the vector of parameters and \( S_\beta \) is the set of feasible parameters value;
- \( \%_{odm}(b) \) is the percentage of trips of individuals using mode \( m \), from zone \( o \) to zone \( d \) as estimated by the model; \( \%_{odm} \) is the percentage of \( odm \) trips resulting from ISTAT and/or MTP data;
\( \text{var}[\sigma_i] \) is the variance of parameter \( i \), whose \( \beta_i \) is the mean; \( \gamma \) is coefficient that measures (following a Bayesian approach) the analyst confidence in the estimates of known variables or parameters. Here follow the expressions for the systematic utilities related to each transport mode as well as the results of the calibrations (Table 1), the statistical tests refer to the models calibrated on disaggregate data.

\[
\begin{align*}
V_{\text{car}} &= \beta_1 T + \beta_2 C + \beta_3 \text{CAR} \\
V_{\text{motorbike}} &= \beta_1 T + \beta_2 C + \beta_4 \text{MOTO} \\
V_{\text{transit}} &= \beta_1 T + \beta_2 C + \beta_5 \text{TRAIN} + \beta_6 \text{NA\_dest} + \beta_7 \text{NA\_or}
\end{align*}
\]

where:

\([T]\) : trip time, \([C]\): monetary trip cost; \([\text{CAR, MOTO}]\): alternative specific attributes; \([\text{TRAIN}]\): binary attribute equal to 1 if the user takes the train; \([\text{NA\_dest}]\): binary attribute equal to 1 if the trip destination is the city of Naples; \([\text{NA\_or}]\): binary attribute equal to 1 if the trip origin is the city of Naples.

<table>
<thead>
<tr>
<th>attributes</th>
<th>home-work ( \rho^2 = 0.44 / \rho^2_{\text{correct}} = 0.43 )</th>
<th>other purposes ( \rho^2 = 0.39 / \rho^2_{\text{correct}} = 0.37 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta ) (t student)</td>
<td>Value of beta ( \beta ) (t student)</td>
</tr>
<tr>
<td>( T/TV/TP ) (h)</td>
<td>(-0.84 (-2.0))</td>
<td>5.25 (€/h) (-0.90 (-1.6))</td>
</tr>
<tr>
<td>( C ) (€)</td>
<td>(-0.16 (-3.1))</td>
<td>-</td>
</tr>
<tr>
<td>( \text{MOTO} )</td>
<td>(-2.96 (-3.8))</td>
<td>3h30 min</td>
</tr>
<tr>
<td>( \text{CAR} )</td>
<td>1.80 (4.7)</td>
<td>2h8 min</td>
</tr>
<tr>
<td>( \text{TRAIN} )</td>
<td>0.87 (2.3)</td>
<td>1h</td>
</tr>
<tr>
<td>( \text{NA_dest} )</td>
<td>0.12</td>
<td>9 (min)</td>
</tr>
<tr>
<td>( \text{NA_or} )</td>
<td>0.60</td>
<td>25(min)</td>
</tr>
</tbody>
</table>

4.2 Trip distribution models

1\(^{st}\) step – parameters calibration on disaggregate data (data: ITER/1,000)

Starting from ITER/1,000 data, from the socioeconomic framework of the territory and from the mode choice models (utilized to estimate “logsum” attributes), a destination choice model has been set and calibrated for each trip purpose. The resulting models allow one to simulate the phenomenon of distribution of extra-urban transfers from the city of Naples and, at the same time, provide a preliminary estimation of the parameters of the trip distribution model from all the other zones of study area. The model adopted is a MultiNominal Logit model.

2\(^{nd}\) step – parameters update by aggregate data (data: ISTAT and MTP).

The parameters of the models have been updated by means of ISTAT origin-destination matrices (as regards the home-work purpose) and of origin-destination matrices provided by the MTP (as regards the home-other purposes). The update has been carried out by minimizing the following target function:
where:
- \( \beta_i, b_i, b, S_B, \text{var}[\sigma] \) and \( \gamma \) refer to the distribution model. They have the same meanings proposed for the mode choice model;
- \( \%_{od}(b) \) is the percentage of trips of individuals from zone \( o \) to zone \( d \) as estimated by the model; \( \%_{odm} \) is the percentage of \( od \) trips resulting from ISTAT and/or MTP data;

3\textsuperscript{rd} step – parameters update by aggregate data (data: FLOWS)
The calibrated models, supported by mode choice models, have been adopted to the purpose of estimating the origin-destination matrices referring to the car mode, home-work purpose (main purpose to be considered in the simulation time period considered) and to the morning peak hour (7.30-9.30). Those matrices, once assigned to the private transport network, have allowed the O-D updating procedure by means of traffic counts (Cascetta 2001). The new O-D matrices, in its turn, have been used to further correct the coefficients of the model according to the procedure previously described [2\textsuperscript{nd} step].

Here follows the expression of systematic utilities related to each destination as well as the results of the calibrations.

home-work \( \rightarrow V_d = \beta_1 \ln(adj) + \beta_2 Y_d + \beta_3 SD + \beta_4 SZ + \beta_5 NA \)

home-other purposes \( \rightarrow V_d = \beta_1 \ln(adj) + \beta_2 Y_d + \beta_3 SD + \beta_5 NA \)

where:
- \([\ln(adj)]\) is the variable which measures the attractiveness of a transport zone; it is obtained from the logarithmic transformation of the number of employment (depending on trip purpose); \([Y_d]\) is the logsum variable which measures the transport accessibility to a generic destination \( d \); \([SD]\) = (same district) is a binary variable equal to 1 if the trip occurs within the same traffic district (the area has been divided into 6 districts see figure 1); \([SZ]\) = (same traffic zone) is a binary variable equal to 1 if the trip occurs within the same area; \([NA]\) = (Naples) is a binary variable equal to 1 if the destination falls within the municipality of Naples.

The parameters resulting from the calibration/updating process are reported in Table 4. All the coefficients resulting from the disaggregate calibrations have appeared statistically significant (with prob. <5%), the mean squared error (MSE) has resulted equal to 0.11 as concerns the home-work purpose and 0.19 as concerns home-other purposes.

Table 2: Calibration results.

<table>
<thead>
<tr>
<th>trip purpose</th>
<th>ln(add)</th>
<th>Y_d</th>
<th>SB</th>
<th>SZ</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>home-work</td>
<td>0.42</td>
<td>2.18</td>
<td>1.02</td>
<td>2.50</td>
<td>0.8</td>
</tr>
<tr>
<td>other purposes</td>
<td>0.68</td>
<td>2.15</td>
<td>2.42</td>
<td>-</td>
<td>0.25</td>
</tr>
</tbody>
</table>
5 Simulation results

The model system implemented has been applied to the displacement scenario as shown in paragraph 2. As concerns the simulated scenario and the no-plan scenario four kinds of impacts have been estimated: (1) impacts on trip distribution, (2) impacts on modal split, (3) impacts on the network performances (vehicles×kilometre, vehicle×hour, flow/capacity ratio along the network), (4) impacts on the environment (air pollution and fuel consumption).

In order to make the phenomenon easier to detect, the different traffic zones have been aggregated into 6 traffic districts plus the city of Naples regardless of the trips occurring within the single traffic zones (see Figure 1). Table 3 shows the percentage changes of the inter-district trips compared with the No-Plan (NP) scenario.

Table 3: O-D matrix between districts (% changes compared with NP scenario).

<table>
<thead>
<tr>
<th></th>
<th>Aversa</th>
<th>Caserta</th>
<th>C.Volturno</th>
<th>Naples*(city)</th>
<th>Naples</th>
<th>Nola</th>
<th>South</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aversa</td>
<td>+40.2%</td>
<td>+42.9%</td>
<td>+8.2%</td>
<td>+27.0%</td>
<td>+17.0%</td>
<td>-10.0%</td>
<td>+19.7%</td>
</tr>
<tr>
<td>Caserta</td>
<td>+28.6%</td>
<td>+31.0%</td>
<td>-3.2%</td>
<td>+20.6%</td>
<td>+11.9%</td>
<td>-11.1%</td>
<td>+15.9%</td>
</tr>
<tr>
<td>C.Volturno</td>
<td>+106.1%</td>
<td>+111.1%</td>
<td>+50.2%</td>
<td>+78.5%</td>
<td>+65.0%</td>
<td>+25.6%</td>
<td>+74.8%</td>
</tr>
<tr>
<td>Naples*(city)</td>
<td>+12.1%</td>
<td>+15.5%</td>
<td>-8.2%</td>
<td>+1.7%</td>
<td>-7.0%</td>
<td>-23.6%</td>
<td>-3.6%</td>
</tr>
<tr>
<td>Naples (prov.)</td>
<td>+31.9%</td>
<td>+32.7%</td>
<td>+7.9%</td>
<td>-1.9%</td>
<td>-4.4%</td>
<td>-25.9%</td>
<td>-29.2%</td>
</tr>
<tr>
<td>Nola</td>
<td>+42.9%</td>
<td>+46.9%</td>
<td>+12.6%</td>
<td>+12.4%</td>
<td>+19.8%</td>
<td>-9.5%</td>
<td>+26.9%</td>
</tr>
<tr>
<td>South</td>
<td>-13.3%</td>
<td>-10.3%</td>
<td>-30.9%</td>
<td>-19.1%</td>
<td>-24.9%</td>
<td>-46.4%</td>
<td>-18.1%</td>
</tr>
</tbody>
</table>

In order to give a comprehensive outline of trips distribution, a comparison between the percentage changes in the total volume of outgoing trips and in the incoming ones is proposed vis-à-vis the NP scenario. Trips have been classified by taking into account the whole of them, excluding intra-zone and intra-district ones (Table 4). The analysis of the resulting percentages shows the increasing power of attraction of some districts and, accordingly, their central position in the territorial hierarchy.

The results have been analyzed by identifying the following three situations: (a) districts where part of the population has been displaced or relocated: Castelvolturno, Aversa, Caserta (only relocation), Naples (traffic district) and the Southern district (only displacement). (b) districts not included in any territorial reorganization plan: Nola (c) the city of Naples

The district of Castelvolturno (30,266 people displaced) has noticeably increased both its links with the neighbouring districts of Aversa and Caserta and its outgoing trips have consequently risen too; the increasing number of people there assigned results from the increasing intra-district trips does not depend on a risen centrality of the town in the territory. Therefore, while in the NP scenario the district of Castelvolturno appears to be self-sufficient, it is not such in the relocation scenario to the detriment of an increased exchange with the surrounding areas.
The district of Aversa, where 65,711 people have been displaced, produces and attracts 32% and 36% more trips respectively (see Table 3). An increase of outgoing trips can be observed towards the districts of Caserta and Aversa (approx. +40%), and only secondarily towards the district of Naples (+27% to the city of Naples, +17% to the rest of the district). The same is true for the district of Caserta, where the largest number of residents (94,491) has been displaced. Caserta district emits and attracts 23% and 25% more than the NP scenario respectively (see Table 4).

Table 4 shows that, disregarding the intra-district trips, the balance of the percentages of emitted and attracted trips increase thus confirming a largest polarizing capacity of the district of Caserta. As regards the outgoing trips (Table 3), the increase mostly concerns the districts of Caserta and Aversa (approx. +30%), and secondarily the district of Naples (+21% as regards Naples, +12% for the remaining district). As regards the incoming trips, the district of Caserta increases its ties with all the remaining districts, except for the Southern one.

Table 4: Trips generated and attracted by each districts (% changes compared with NP scenario).

<table>
<thead>
<tr>
<th>districts</th>
<th>towards all destination</th>
<th>towards all destination</th>
<th>towards all destination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>emitted</td>
<td>attracted</td>
<td>emitted</td>
</tr>
<tr>
<td>Aversa</td>
<td>28%</td>
<td>29%</td>
<td>32%</td>
</tr>
<tr>
<td>Caserta</td>
<td>23%</td>
<td>25%</td>
<td>26%</td>
</tr>
<tr>
<td>C.Volturno</td>
<td>66%</td>
<td>43%</td>
<td>75%</td>
</tr>
<tr>
<td>Napoli* (city)</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Napoli(province)</td>
<td>-5%</td>
<td>-7%</td>
<td>-3%</td>
</tr>
<tr>
<td>Nola</td>
<td>0%</td>
<td>-12%</td>
<td>7%</td>
</tr>
<tr>
<td>Sud</td>
<td>-16%</td>
<td>-14%</td>
<td>-20%</td>
</tr>
</tbody>
</table>

The district of Naples, marked by an increase of 159,572 residents and a parallel decrease of 221,438 (balance of –61,866), is affected by a turn-about of the incoming and outgoing flows: flows to the North increase, flows to the South decrease. Table 3 shows that outgoing trips dropping because of the decreasing population and the same occurs for incoming trips as a result of diminishing business activity in the area because the district of Caserta neighbourhood are getting more and more competitive. Naples close ties with the districts of Nola and the city of Naples are weakening. Accordingly, the centre of gravity of mobility is clearly shifted towards Caserta neighbourhood thus confirming a probable outstanding role it might have in this scenario in the near future. As expected, the Southern district is affected by a decrease both in trips emitted and attracted. Trips towards the district of Nola are noticeably dropping (-16%) while outgoing trips are increasing.

The city of Naples is affected by a turn-about of demand flows. Incoming and outgoing flows to the North are increasing, flows to the South are decreasing.
Naples is further strengthening its ties with Caserta neighbourhood, by reducing its interactions with Nola and the South (see Table 3).

Table 5 shows the main impacts simulated. In particular, the percentage of usage of private car shifts from 78% to 76% and the percentage of transit users shifts from 19% to 22%. Although the total number of cars circulating in the peak hour shifts from 348,465 to 336,683, it should be noted that the vehicle×Km and vehicle×h increase, meaning that the length and the duration of trip are much larger than in the NP scenario (+14% as regards trip length, 15% as regards trip duration). Nevertheless, such results do not lead to worse network performances. In fact, average criticalities (flow/capacity) along the network and the mean speed do not observe relevant decrease. Moreover, the emissions of polluting substances if compared to fuel consumption along the network appear to be slightly improved as reflected by the network indicators reported below.

Table 5: Main impacts (simulation interval 7.30-9.30 a.m.).

<table>
<thead>
<tr>
<th></th>
<th>% car</th>
<th>% transit</th>
<th>vehic×h (×1,000)</th>
<th>vehic×km (×1,000)</th>
<th>mean speed</th>
<th>f/C</th>
<th>CO</th>
<th>NOx</th>
<th>HC</th>
<th>fuel</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-Plan</td>
<td>79%</td>
<td>19%</td>
<td>4,754</td>
<td>234,371</td>
<td>49.3 (km/h)</td>
<td>0.67</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Scenario</td>
<td>76%</td>
<td>22%</td>
<td>5,504</td>
<td>269,365</td>
<td>48.9 (km/h)</td>
<td>0.69</td>
<td>-8%</td>
<td>-11%</td>
<td>-8%</td>
<td>-8%</td>
</tr>
</tbody>
</table>

6 Conclusions

The method as a whole has produced a model system that, besides well reproducing the current scenario as regards the demand of mobility, shows a good flexibility in respect of the territorial variables and the transport supply variables (accessibility, times and costs). The analysis of the results has proved that: (a) the displacement policy modifies substantially the trips patterns, (b) such modifications do not increase the congestion of Naples city and of its suburbs, (c) the traffic districts not involved in the displacement strategy observe a relevant decrease of attracted travel demand, finally, (d) the network performances do not decrease and satisfactory results are achieved as regards modal split, road link criticalities and pollutant emissions. In conclusion, the relocation/displacement strategy can be easily supported by the transport system, and the provided infrastructural interventions let the displacement strategy to be sustainable. Nevertheless some efforts should be spent to evaluate the impact on the economy of those areas that decrease the travel demand attracted.

References