A NEW MODEL FOR DISAGGREGATE TRAFFIC ASSIGNMENT MAKING EXPLICIT THE SPATIAL DISTRIBUTION OF TRIP EXTREMITIES

Vincent Benezech
Université Paris-Est – Ecole des Ponts ParisTech

1. INTRODUCTION AND LITERATURE REVIEW

In all traffic assignment models, one of the central assumptions is that all the activity of a Traffic Analysis Zone (TAZ) is concentrated at the centroid. Trips originate and end at a single point; in the simple case of all or nothing assignment, all trips between any two given zones use the same pair of connector. The spatial diversity of trip extremity points is completely ignored: OD matrices are taken as traffic flows between centroids. Connectors, which link centroids to the network, are dealt with just like any other network link would be, although they have no physical reality.

Historically, the problem takes roots in the Modifiable Area Unit Problem (MAUP), introduced in the 30s by geographers. In their paper, Gehlke and Bieahl (1934) showed that regional partitioning according to which data are aggregated have very strong influence on the measures of certain quantities and on the correlations between these quantities. This phenomenon has been observed in many other papers, leading to abundant literature; in particular Amrhein (1994) shows that correlation coefficients are very sensitive to data aggregation. Openshaw (1977) was the first one to link the MAUP with transport issues. His approach has led many authors to investigate the ways TAZs are designed and to elaborate rules and algorithms for TAZ delineation. Those algorithms generally work iteratively, by aggregating basic spatial units into bigger ones until the desired number of zones is reached. Baass (1981) gives an example of how such an algorithm might work; his algorithm is quite complete in the sense that it allows to take account of many factors, such that zone completeness, zone shape or zonal homogeneity in terms of certain socioeconomic characteristics.

Another aspect of research focuses on centroid positioning and the assumption that it is representative of the whole zone at the centre of an area. Ortuzar and Williamsen (2001) enunciate the following rule: “Zone size must be such that the aggregation error caused by the assumption that all activities are concentrated at the centroid is not too large”. Since, for practical reasons, it is not desirable to have too many TAZs, many zones fall behind this rule, especially in rural areas where activity is very diffuse. As a remedy, some have tried to act on centroid positioning. Chang et al. (2002) introduce a population weighted centroid and a household weighted centroid. They remark that these new centroids generate better assignment results than the ordinary (geometric) ones for large zones but that it is not systematic for smaller zones. Their estimates, at the US state level, use cities as basic units; results would certainly be different on smaller scaled problems. Martinez et al. (2007) deal with urban areas. The authors perform a two dimensional nonparametric regression on the distribution of trip extremities obtained from mobility surveys, thus identifying highest peaks of demand that can be used as centroids and constructing the zones around. They find that their method significantly decreases the length between the centroid and trip extremities, indicating that their centroids represent better whole zones.

More recently, Constantin and Florian (2010) argue that the stochastic assignment framework is more adapted to deal with zonal heterogeneity. Their model is inspired by the work by Dial (1971): at each node, the distribution of flows among competing lines is not directly proportional to line frequencies but result from a *logit* model. When the node considered is the origin centroid, this model divides the flow originating from a zone among the possible connectors. However, Constantin and Florian do not give suggestions as to the form of the *logit* model and the explanatory variables to be used.

In this paper, the emphasis is placed on modelling the diversity of trip extremities inside zones or, equivalently, the distribution of the distance from trip extremities to network access points. We introduce a discrete choice model, based on the assumption that connector travel times are not deterministic but stochastic, to account for the heterogeneity of terminal travel times. This new model belongs to the family of stochastic assignment models where the discrete choice is made between routes (and not at each node). These models are well documented (Sheffi (1985) for instance).

Leurent et al. (2011) contains a technical description of the model with an application to the whole Paris region. In this paper, emphasis is put on the assumptions behind the model and on some justification of these assumptions. The remaining of the paper is organised in three parts. Section 2 recalls the foundations of the mode. In Section 3, some justification is given about the underlying assumptions. Finally, Section 4 describes the application of the new model to an area of the Paris region.

2. MODEL DESCRIPTION

2.1 Network representation

It is supposed that two descriptions of the road network are available. The first one, from now on the main network, consists of all the main roads. It is coarse enough to allow heavy calculations of shortest paths. In the examples that follow, it consists of the network that is currently in use for traffic assignment by the Paris region metropolitan agency. The second one, from now on the detailed network, includes as much as possible of the roads that cars can use. The Open Street Map network is a good candidate because it contains all roads but it has two major drawbacks: it might contain paths that can only be travelled by foot and, when introduced into a GIS, link connectivity is sometimes flawed. In this paper, we use the IGN network compiled by the French national institute of geography, for which the two former points do not apply.

Due to its coarseness, the main network cannot be used to describe disaggregate trip ends in details, which is why the detailed network is needed. The main assumption on which our model is built is as follows:
(H) *Small roads are used around origin and destination points on a scale that is comparable to zone size.*
This means that, outside origin and destination zones, only the main network is used and trip disaggregation can be described at the zone level.

Using (H), trips can be divided into three parts: the first one, on the detailed network, links the trip starting point to the main network; the second one, on the main network, links the origin and destination zones; the last one links the main network to the end point. In order to be able to perform an assignment procedure with the main network, it is supposed that there are only a few nodes at which the detailed network branches onto the main one. Following the terminology of Leurent et al. (2011), those points are called anchor points. To simplify notation even further, it is assumed that the number of anchor nodes per zone is fixed to a certain number, n_a. In the standard centroid based assignment procedure, n_a access/egress centroid connectors with deterministic travel times would be created, effectively linking the centroids to the network.

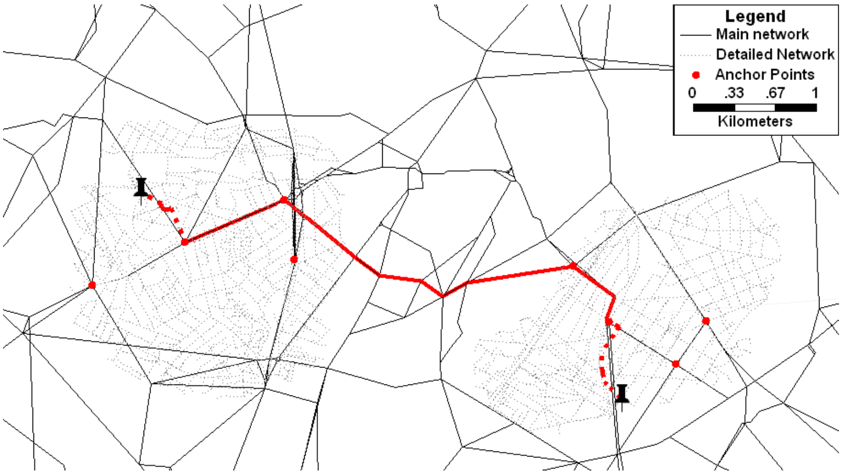

Figure 1: Path between two points (pins on the map) decomposed in a main part and two access/egress parts.

2.2 Assignment procedure

In this paper, we limit ourselves to all-or-nothing assignment, leaving aside congestion. The assignment consists of two parts. The first part searches for the shortest paths between any two anchor points. The second one assigns, by OD pair, trips to pairs of anchor points (one for the origin zone and one for the destination zone). Since congestion is ignored, shortest paths can be found easily using a Dijkstra-type algorithm. It remains to render operational the second part of the assignment, that is to calculate the proportions according to which the pairs of anchor nodes are used.

Once the shortest paths are found, there are n_a^2 possible routes linking two specific zones. The problem is rewritten as a discrete choice model that pertains to the whole route. Assuming that the time it takes to reach one anchor node of the origin zone does not depend on the destination zone (and *vice-versa*), we write the total travel time from origin zone *o* to destination zone *d*, using anchor node α at origin and anchor β at destination, as:

$$T_{\alpha\beta}^{od} = \theta_{\alpha}^{o} + t_{\alpha\beta}^{*} + \theta_{\beta}^{d}$$

where θ_{α}^{o} is a random variable describing the access time to anchor point α from zone o (similarly for θ_{β}^{d}) and $t_{\alpha\beta}^{*}$ is the optimal time between α to β. Due to the strong correlations between the θ_{α}^{o}, and between the θ_{β}^{d}, a probit model is preferred to a logit model: it is assumed that the vectors $(\theta_{\alpha}^{o})_{\alpha}$ and $(\theta_{\beta}^{d})_{\beta}$ follows a multivariate normal distribution. Section 2.3 describes how the statistical inference is carried out.

Denote by μ^{o} and Σ^{o} the vector of means and the variance-covariance matrix of $(\theta_{\alpha}^{o})_{\alpha}$; by μ^{d} and Σ^{d} those of $(\theta_{\beta}^{d})_{\beta}$. Then, the mean travel time from *o* to *d* via anchor nodes α and β is:

$$\overline{T}_{\alpha\beta}^{od} = \mu_{\alpha}^{o} + t_{\alpha\beta}^{*} + \mu_{\beta}^{d}$$

and the covariance between two routes (with indices 1 and 2) between *o* and *d* is :

$$\chi_{12} = \Sigma_{\alpha(1)\alpha(2)}^{o} + \Sigma_{\beta(1)\beta(2)}^{d}$$

The probabilities coming from the probit model with the inputs as above can be calculated using formulas by Clark (1961) that approximate the maximum of normal variables by another normal variable and give the mean and the variance of that maximum (also Maher and Hughes (1997) for a practical application). Depending on the size of the region, the value chosen for n_a and the computing power available, it is also possible to calculate these probabilities by simulation. This is the solution chosen here.

2.3 Distribution of θ_{α}^{o} and θ_{β}^{d}

In order to statistically infer the distribution of the vector $(\theta_{\alpha}^{o})_{\alpha}$ (and, in a similar fashion, that of $(\theta_{\beta}^{d})_{\beta}$), a sample of trip extremities within each zone is needed. It is then easy to calculate the travel time from each point in the sample to all the anchor points and to estimate, with Maximum Likelihood Estimators for instance, the vector of means and the variance-covariance matrix. Assuming the $(\theta_{\alpha}^{o})_{\alpha}$ follow a multivariate normal distribution, these two pieces of information characterise fully the probability density of $(\theta_{\alpha}^{o})_{\alpha}$.

For the Paris region, there exists a map fully describing land use, cutting the area into small zones, called MOS, with similar land use and thus similar population and job densities. With views of predicting traffic load during the evening peak hour, job weighted estimates were used for access times while population weighted estimates served for egress times. The reverse should be used in the case of prediction concerning the morning peak hour.

3. MODEL JUSTIFICATION

3.1 Assumption (H)

The rationale behind (H) is that small roads are mostly used close to the origin and destination points (Bovy and Sterne (1990)). However, the numeric validity of (H) depends highly on the coarseness of the main network and the way zones were designed. In our case, the main network and the zoning system were designed by the same organisation, so that zones are more or less sized inversely proportionally to the level of details of the network. Furthermore, if the main network contains all roads of significant capacity, the share of traffic that is overlooked is inherently negligible.

3.2 Local traffic does not influence traffic load on main network

The number of anchor points being limited, not all connection nodes from the detailed network onto the main network are included. As a consequence it is possible that the access/egress part of trips also use main roads while being treated as using the detailed network. To avoid this local traffic influencing the traffic load on the main network, it is necessary that the access/egress traffic is negligible compared to the main traffic. This will be the case if the following conditions are fulfilled:

- The number of anchors is sufficient so that the distance or time travelled before joining the main network is small. Anchor positioning can also play a role.
- Zone population is small compared to traffic flows expected on the main road arcs that go through the zone.

On the test case presented in Section 4, travel time on connector represents on average less than one tenth of total travel time. However, for some zones, zone population is of the same order of magnitude than the sum of traffic flows going through them, so that the second condition is not always fulfilled. All in all, in terms of vehicle time, the overlooked traffic flow on the main network is still relatively small.

3.3 Number and construction of connectors

The number of anchor points associated to each zone can be fixed and chosen in advance. It does not matter whether some are unlikely to be used by car drivers as a null probability will be assigned to those paths that go through those anchors. However, this number should be high enough in order to account for the diversity of possible ways of getting onto the main network. The number four chosen in the practical cases comes from the four cardinal directions reflecting four totally ways of accessing or leaving a zone (if it close enough to a square). However, there is not any more rationale to it.

The anchor points were constructed using the TransCAD point to network connecting procedure. There has to be variety in the directions towards which the connectors go, however once this condition is met, they are chosen as close as possible to the zone centroid. This is not optimal in our case. It would be better to have anchor nodes that cover as much activity location as possible inside every zone, thus minimising average access/egress trip length. We are currently working on a better anchor node selection procedure.

3.4 Distribution of access/egress times

There is no tractable analytical formula giving the distribution from a random set of points (the extremities of trips) to a fixed point (the anchor point), and so regardless of the probability density describing the spatial distribution of the random points. If population density is normally distributed, the squared distance to a fixed point is distributed as the sum of two noncentral chi-square distributions. This means that the normal assumption might not be too far fetched, especially if the number of observations is high.

The assumption of normality of the access/egress times is further checked on one of the zones from the Paris region, chosen randomly (see Figure 2 below).

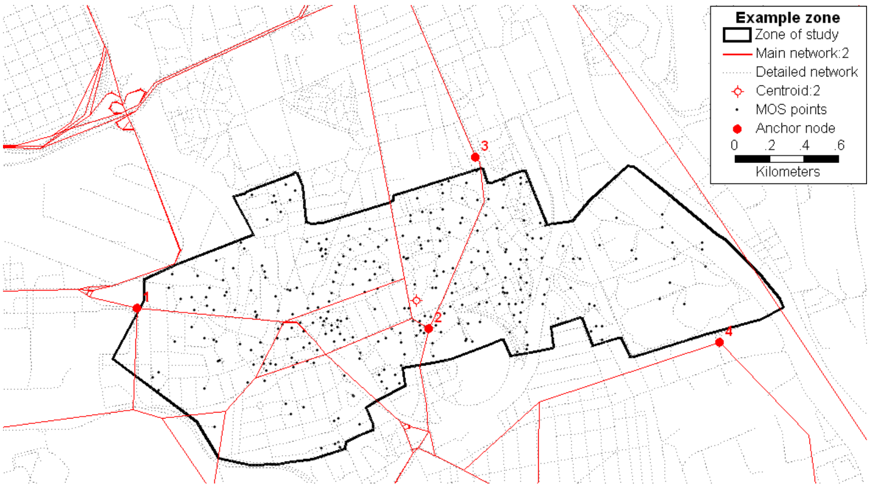

Figure 2: Map of the zone studied (delimited in brown). The main network with the anchor points are in red. The detailed network is in black dotted lines.

Access times to all the anchor points are evaluated, using all MOS points that have positive population. Results from Figure 2 show that the normal assumption holds broadly, except maybe at the tails.

We would now like to compare the *probit* and the *logit* models. To simplify, we assume that all passengers have the same destination, somewhere North, via the motorway (in red, upper left corner on Figure 1). From this zone, entrance on the motorway is possible via two access roads; both are considered. Table 1 shows the proportion of traffic going through each anchor for this simple assignment problem. For the exact assignment, all MOS points that have positive population are once again considered. The results from the probit method agree better with the exact results than those from the logit method.

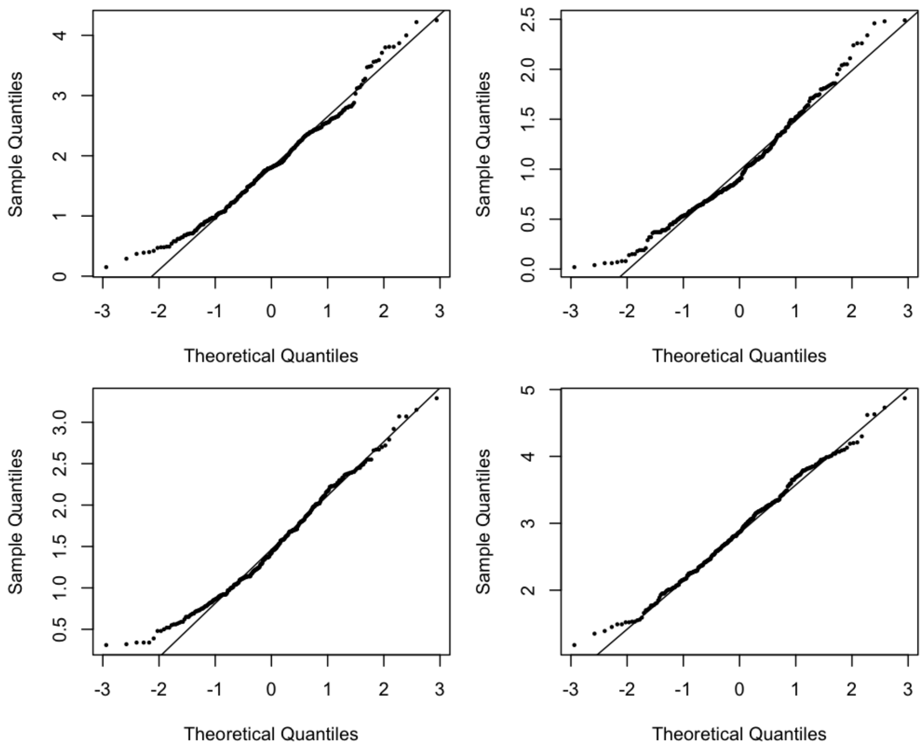

Figure 3: Nomal Q-Q plots of the distribution of travel times to the four anchor points.

Anchor number	Exact	Probit	Logit	Centroid
1	0.44	0.46	0.34	1
2	0.22	0.20	0.34	0
3	0.34	0.34	0.32	0
4	0	0	0.01	0

Table 1: Proportion of traffic going through each node, four different computing method.

The scale parameter is arbitrarily fixed to one; since the three first options are assigned very similar probabilities, a different scale parameter would not change much. In this case, it could also be argued that a nested logit model would be more appropriate; however, the underlying nesting structure does not appear clearly before calculation and certainly depends on the destination, which renders the model enormously more complex.

The difference between a logit and a probit model seems significant. As another indicator, the difference in terms of flow proportion between those two models was calculated for the 60 zones introduced in Section 4. This difference was bigger than 0.1 for more than a third of the anchor nodes and bigger than 0.25 for approximately a tenth of the anchor nodes.

This comparison relies on the assumption that trip extremities are distributed similarly to population or job location. The data from mobility survey in the Paris region is not geographically precise enough, nor is it detailed enough (only a couple of observations per zone) to allow estimation of access/egress times and the population and job locations are the best proxy to trip extremity locations that can be easily accessed, at least for commuting trips; this estimation method can be adapted for other trip purposes, using shop locations for leisure trips, for instance. It is difficult to assess the validity of this approximation, but mobility surveys (Dreyfus (2005)) suggest, at least in the Paris region, that it is not too big; it is also possible to use the coefficients linking mobility to socioeconomic characteristics that are given in those surveys to further improve the model.

4. APPLICATION

We now apply out model to a portion of the Paris region. This was preferred to an application to the whole metropolitan area mainly because, the detailed network being only available by *départements* (counties), the shortest paths from some MOS points to anchor nodes may be miscalculated around *départements* boundaries. The study region was delineated taking this problem into account and is integrally part of the Seine-Saint-Denis *départment* (North-East of Paris). It consists of 60 zones and there are 194 external points, i.e. nodes that are outside extremities of links going across the study region boundaries. The OD matrix for the study region was determined *via* the sub-area analysis of the TransCAD software.

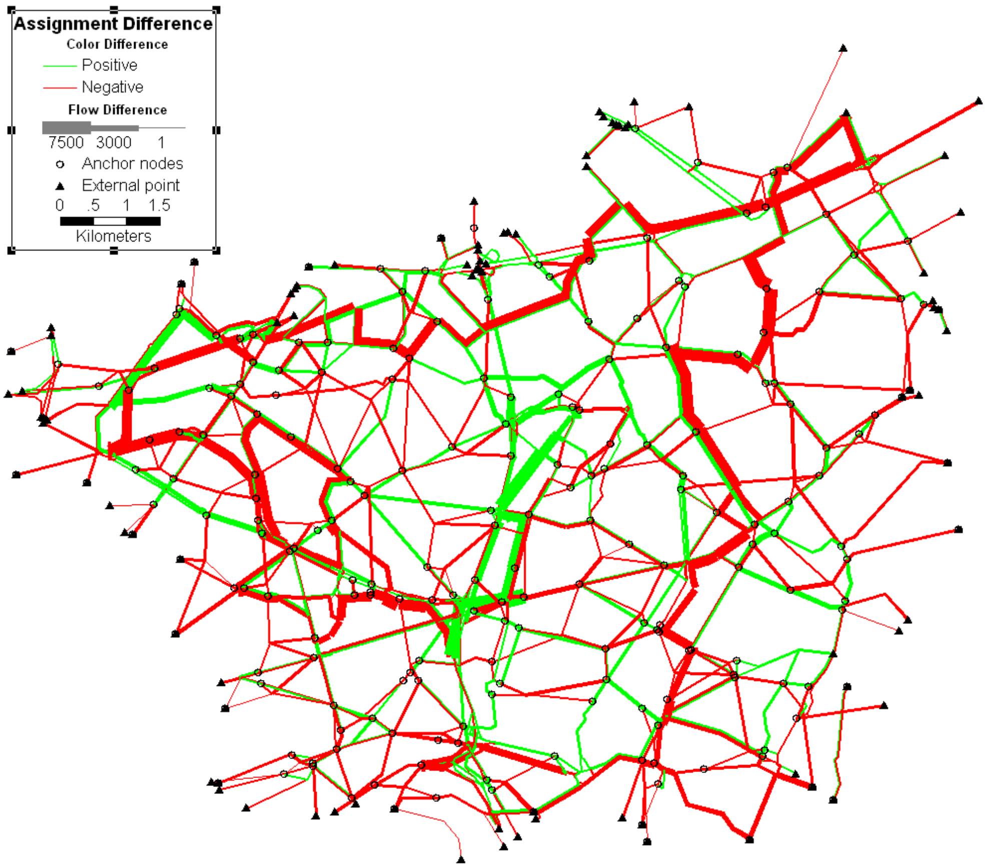

Figure 4: Differences in car flows between the standard assignment procedure and the new disaggregate method. Green lines correspond to links where the standard assignment procedure yields more traffic.

Traffic assignment was done twice on the study region, once using the standard assignment procedure, and a second time using the disaggregate procedure described above. The results are shown in Figure 4. There are no clear patterns in the differences. They can be positive and negative on any type of link, except on those links that connect the network to external nodes, which is normal since their traffic is fixed by external conditions (those links appear coloured on Figure 4, but the actual difference in car flow is negligible, less than one unit in general). The differences can be significant in relative value, even on main links with big capacities. This difference represents more than half of the maximum flow on some links (see Figure 5), indicating a complete rerouting of a sensible part of the traffic flow. A possible explanation lies in the fact that jobs and dwellings tend to concentrate on areas that are near some kind of major road, but in proportions that can vary greatly. While centroids are totally blind to this fact, treating equally anchor nodes that are in deserted and densely populated areas, the disaggregate procedure will assign flows to anchors in proportions to the known activity surrounding those anchors.

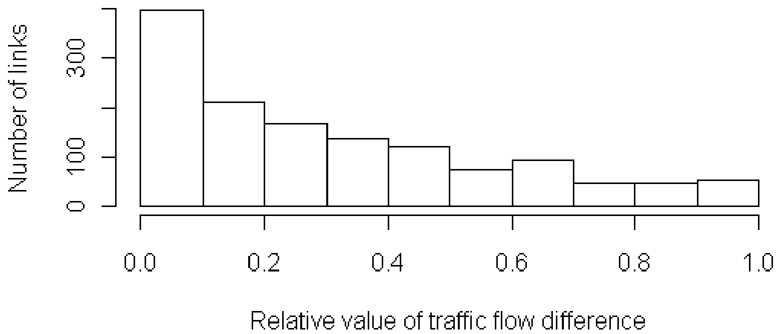

Figure 5: Histogram of the ratio between traffic flow difference between the two assignment procedures and the value of the traffic flow obtained with the standard assignment procedure.

The average trip time, without access/egress time, is 19.3 min for the standard method and 18.3 min for the new method, hence a difference of 5.5%, indicating a better repartition of vehicles among the various options. With the new model, the average time spent on centroid connectors is 1.8 min and the variance of travel times for trips from one OD pair is 2.2 min^2 on average, showing a real variability in the anchor node chosen for just one OD pair.

It appears clearly that the new disaggregate procedure modifies significantly the repartition of traffic on the network and that, in terms of travel times, trip end dispersion is not a negligible phenomenon. The new procedure allows to deal with zones that are not monocentric or that spans on huge areas, for instance in the cases where the centroid connector that is used is located at one end of a zone and the trip starts at the other end. However, it is not yet possible to be conclusive about the results, since only all or nothing assignment has been tested.

5. CONCLUSION

In this paper, a new traffic assignment method that takes account of the distribution of trip extremities is introduced and tested on a real network coming from the Paris area. Under this new method, connector link travel times of zones are not deterministic but stochastic and their joint distribution is assumed to be normally distributed, thus accounting for some part of the dependency among the access/egress times. A method is given to estimate those normal variables when some geographically precise data are available. If congestion is left apart, the model is then equivalent to a number of *probit* models, one per OD pair, for which there is no computing problems.

Under the assumption that the distribution of trip extremities is similar to that of population (or, equivalently, to that of jobs), the test case shows that the *probit* model is the most accurate. Combining job-based estimates and population-based estimates allow to model morning peak-hour or evening peak-hour, as in the application to the Paris region. The results from this application are promising. They also demonstrate that possible errors resulting from bad modelling of connectors do not naturally tend to compensate and so should be taken into account. However, to analyse the results fully, the model should be able to deal with congestion. This is one of our areas of research at the moment, as well as the link between connector positioning and assignment precision.

BIBLIOGRAPHY

Amhrein, CG (1994) Searching for the elusive aggregation effect: Evidence from statistical simulation. *Environment and Planning A* **27**:105-119.

Baass, KG (1981) Design of zonal systems for aggregate transportation planning models. *Transportation Research Record* (807):1–6.

Bovy, PHL and Stern, E (1990) *Wayfinding in transport networks*. Kluwer Academic Publishers.

Chang, KT, Khatib, Z and Ou, Y (2002) Effects of zoning structure and network detail on traffic demand modeling. *Environment and Planning B: Planning and design* (29):37–52.

Clark, CE (1961) The greatest of a finite set of random variables. *Operations Research*, **9**:145-162.

Constantin, I & Florian, M (2010). Some variants of the optimal strategy transit assignment method. *Proceedings of the European Transport Conference*.

Dial RB (1971) A Probabilistic Multipath Traffic Assignment Model which Obviates Path Enumeration. *Transportation Research* (5):83-111.

Dreyfus, J (2005) *Le profil des déplacements journaliers en transport en commun et voiture particulière*. Les cahiers de l'EGT, numéro 2, DREIF.

Gehlke, C and Biehl, H. (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association Supplement,* **29**:169–170

Leurent, F, Benezech, V and Samadzad, M (2011) A Stochastic Model of Trip End Disaggregation in Traffic Assignment to a Transportation Network. *Procedia – Social and Behavioral Sciences*, Elsevier.

Maher, MJ and Hughes, PC (1997) A probit-based stochastic user equilibrium assignment model. *Transportation Research B* **31** (4):341-355.

Maher, M.J. and Hughes, P.C. (1995) The attractions of SUE—a practicable approach to stochastic user equilibrium assignment. *23rd European Transport Forum, Proceedings of Seminar E*, pp. 103–115.

Openshaw, S (1977) Optimal zoning systems for spatial interaction problems. *Environment and Planning A*, **9**:169-184.

Ortuzar, J D and Willumsen, L (2004). *Modelling Transport*. Wiley, 3rd edition.

Sheffi, Y (1985) *Urban Transportation Networks*. Prentice-Hall.