TAZ Delineation Application for Transport Planning Studies: A New Approach Applied to the Ilê-de-France (Paris)

Luis Miguel Garrido Martínez
José Manuel Caré Baptista Viegas

Technical Report for DEST-INRETS

September 2009
ABSTRACT

In most transport planning studies one of the first steps is the definition of a zoning scheme with which the study area is divided and the corresponding space is disaggregated. There are no clear rules on how to carry out this operation in an optimal way, and the dominating practice is to do it based on experience, trying to mix a certain degree of within-zone homogeneity and the convenience of using administrative borders as zone limits.

Firstly, the report present a summary of the state of the art and practice on zoning definition on transportation studies, identifying a set of constraints or guidelines for zoning delineation. After this initial assessment, the report presents a set of quality criteria for a general zoning scheme and an algorithm that constructs an initial zoning based on a sample of geo-referenced trip extreme points and improves it in successive steps according to those criteria. This kind of zoning fits perfectly well to traffic assignment purposes.

But this study will investigate an improvement of this approach in order to give a better understanding of the mobility determinants and its externalities on the environment. In doing so, the new zone is determined not only by the trips generation and distribution but also constrained by other indicators. In our case, we have selected a combination of the following: 1. air pollution emissions, 2. population density, 3. work and study density and 4. public transport accessibility. The integration of these 4 indicators allows us not only to picture the mobility within the region and to identify at the very precise level the main zones of activities and traffic exchanges. This integration relates the picture to the land use and the clustering of the economic activities location at a very discrete level. Furthermore, it relates the density of the mobility in dense, large and economically dynamic urban area to its externalities in terms of air pollution.
In order to be effective for mobility analysis and policy purposes, this kind of approach cannot only rely on the cell grid unit but a hierarchical aggregation should be set up. This aggregation allows analyses within the administrative and political boundaries but with a more disaggregated perspective.

A case study based on the Mobility Survey for the Paris Metropolitan Area in 2001 is developed to illustrate this new approach to zoning.
# TABLE OF CONTENTS

Abstract ........................................................................................................................................ i

Table of Contents .......................................................................................................................... iii

Figures .......................................................................................................................................... v

Tables .......................................................................................................................................... ix

Glossary ........................................................................................................................................ xi

1. Introduction ................................................................................................................................. 1

   1.1. Research Goals .................................................................................................................... 6

   1.2. Structure of the Report ....................................................................................................... 6

2. TAZ Delineation Algorithm ....................................................................................................... 9

   2.1. Problem Formulation .......................................................................................................... 9

   2.2. Methodology of analysis and selection of zones limits ....................................................... 9

   2.3. TAZ Delineation Algorithm Structure ............................................................................... 15

   2.4. TAZ indicators of compliance ......................................................................................... 19

   2.5. Local Improvements to the TAZ Delineation Algorithm ............................................... 21

3. Assessment of EGT Data .......................................................................................................... 23

   3.1. The region of the Île-de-France ....................................................................................... 24

   3.2. Household characterization ............................................................................................... 27

      3.2.1. Home Ownership and Dwelling Characterization .................................................... 28

      3.2.2. Household Composition and Income Distribution ..................................................... 29

      3.2.3. Characterization of the respondent ............................................................................ 32
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.4.</td>
<td>Motorisation and private car use of the household</td>
<td>35</td>
</tr>
<tr>
<td>3.3.</td>
<td>Mobility characterization</td>
<td>36</td>
</tr>
<tr>
<td>3.4.</td>
<td>Accessibility and land-use characterization</td>
<td>44</td>
</tr>
<tr>
<td>3.5.</td>
<td>Analysis of the Indicators to integrate the new TAZ Delineation Algorithm</td>
<td>47</td>
</tr>
<tr>
<td>4.1.</td>
<td>New Problem Formulation</td>
<td>58</td>
</tr>
<tr>
<td>4.2.</td>
<td>Changes to the Original TAZ Delineation Algorithm</td>
<td>59</td>
</tr>
<tr>
<td>4.2.1.</td>
<td>Vicinity Rules</td>
<td>59</td>
</tr>
<tr>
<td>4.2.2.</td>
<td>Objective Function and Constraints</td>
<td>61</td>
</tr>
<tr>
<td>4.3.</td>
<td>Application to the Region of the Ilê-de-France</td>
<td>62</td>
</tr>
<tr>
<td>5.</td>
<td>Conclusions and Future Developments</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>69</td>
</tr>
</tbody>
</table>
FIGURES

Figure 2.1 - Total origins and destinations in Lisbon Municipality 3D view .................... 10
Figure 2.2 – Search example ........................................................................................................ 11
Figure 2.3 – Explanation of the variables used in the spreading rules equations .............. 11
Figure 2.4 – Decision tree for determining the optimum of the TAZ algorithm objective function .................................................................................................................. 14
Figure 2.5 – Delineation algorithm functioning and data flow ............................................. 18
Figure 3.1 – Map of Ilê-de-France location within France ................................................ 24
Figure 3.2 – Map of the Ilê-de-France communes ................................................................. 25
Figure 3.3 – Trip origins of the EGT 2001 mobility survey .................................................. 27
Figure 3.4 – Spatial distribution of the percentage of home owners within Ilê-de-France region ........................................................................................................................................ 28
Figure 3.5 – Spatial distribution of the percentage of households living in apartments within Ilê-de-France region ........................................................................................................ 29
Figure 3.6 – Spatial distribution of the average household size within Ilê-de-France region ................................................................................................................................. 30
Figure 3.7 – Spatial distribution of the average number of children under 6 years old within Ilê-de-France region ........................................................................................................ 30
Figure 3.8 – Spatial distribution of the average income of the households within Ilê-de-France region ................................................................................................................................. 31
Figure 3.9 – Spatial distribution of the average of active persons in the household within Ilê-de-France region ................................................................................................................................. 32
Figure 3.10 – Spatial distribution of the average age of the respondents within Ilê-de-France region ........................................................................................................................................ 33
Figure 3.11 – Spatial distribution of the average percentage of retired respondents within EGT 2001 database ................................................................................................................................. 34
Figure 3.12 – Spatial distribution of the average percentage of liberal professional respondents within EGT 2001 database .............................................................................................................. 34
Figure 3.13 – Spatial distribution of the average number of motorised vehicles per household within Ilê-de-France region ........................................................................................................ 35
Figure 3.14 – Spatial distribution of the average number of km travelled with the motorised vehicles of the household within Ilê-de-France region .................................................... 36
Figure 3.15 – Trips start time distribution of EGT 2001 survey ................................................... 37
Figure 3.16 – Trip length distribution of the EGT 2001 survey ..................................................... 38
Figure 3.17 – Trip duration distribution of the EGT 2001 survey ................................................ 38
Figure 3.18 – Mode share distribution of the EGT 2001 survey .................................................... 39
Figure 3.19 – Trip purpose distribution of the EGT 2001 survey .................................................. 40
Figure 3.20 – Spatial distribution of destinations of trips in the EGT 2001 survey ................. 40
Figure 3.21 – Distribution of weekly emissions of households of the EGT 2001 survey .... 44
Figure 3.22 – Public transport accessibility distribution of the households surveyed on the EGT 2001 ................................................................................................................................. 45
Figure 3.23 – Public transport accessibility spatial distribution on the Ilê-de-France region ................................................................................................................................. 45
Figure 3.24 – Population density spatial distribution of the Ilê-de-France region ............... 46
Figure 3.25 – Work and study density spatial distribution of the Ilê-de-France ..................... 47
Figure 3.26 – 1500 zones scheme of the Ilê-de-France region (TAZ Delineation Algorithm) ................................................................................................................................. 48
Figure 3.27 – Percentile assessment of the Emissions factor indicator for the Ilê-de-France region ................................................................................................................................. 50
Figure 3.28 – Percentile assessment of the Population density indicator for the Ilê-de-France region ................................................................................................................................. 50
Figure 3.29 – Percentile assessment of the Public transport indicator for the Ilê-de-France region ................................................................................................................................. 51
Figure 3.30 – Percentile assessment of the Work/study density indicator for the Ilê-de-France region ................................................................................................................................. 51
Figure 3.31 – Cluster analysis with the 81 profiles for the Ilê-de-France region ............... 52
Figure 3.32 – IUARIF zoning of the Ilè-de-France region............................................................... 53
Figure 3.33 – Cluster analysis of the Ilè-de-France (Emissions factor as main indicator) ... 54
Figure 3.34 – Cluster analysis of the Ilè-de-France (Population density as main indicator) ... 54
Figure 3.35 – Cluster analysis of the Ilè-de-France (Public transport accessibility as main indicator) .................................................................................................................................................. 55
Figure 3.36 – Cluster analysis of the Ilè-de-France (Work/study density as main indicator) .................................................................................................................................................. 55
Figure 4.1 – New search rules example.......................................................... 60
Figure 4.2 – Decision tree for determining the optimum of the Hierarchical TAZ algorithm objective function.......................................................................................................................................... 62
Figure 4.3 – TAZ 50 zones (Application of the Hierarchical TAZ Delineation Algorithm) .... 64
Figure 4.4 – TAZ 75 zones (Application of the Hierarchical TAZ Delineation Algorithm) .... 65
<table>
<thead>
<tr>
<th>2010</th>
<th>TAZ Delineation Application for Transport Planning Studies: A New Approach Applied to the Ilê-de-France (Paris)</th>
</tr>
</thead>
</table>

viii  Figures | IST
TABLES

Table 1.1 – Constraints to the definition of TAZ .......................................................... 3
Table 2.1 - Microscopic and Macroscopic Indicator Summary ................................... 20
Table 3.1 – Aggregate statistics of the profession distribution or activity status of EGT 2001 respondents........................................................................................................ 33
Table 3.2 – Coefficients of TREMOVE models for different transport modes and vehicles 42
Table 3.3 – Weekly assessment of trips emissions of the different pollutants for EGT 2001 ........................................................................................................................................... 43
Table 3.4 – Summary of the 1500 zones scheme indicators (TAZ Delineation Algorithm). 48
GLOSSARY

Error – A geographical error is a difference between a computed, estimated, or measured value and the true, specified, or theoretically correct value.

Enquête Globale de Transport de 2001 (EGT 2001) – Mobility survey performed in 2001 in the Île-de-France region with 23,656 respondents who describe their daily individual trips, resulting in a total of 81,386 trips (0.23% of total trips). The trip ends are geocoded using a 300 square meters grid cell.

Geographic information system or geographical information system (GIS) – GIS is a system for creating and managing spatial data and associated attributes. In the strictest sense, it is a computer system capable of integrating, storing, editing, analyzing, sharing and displaying geographically-referenced information. In a more generic sense, GIS is a ‘smart map’ tool that allows users to create interactive queries (user created searches), analyze the spatial information, and edit data. Geographical Information Science is the science underlying the applications and systems, taught as a degree program by several universities.

Modifiable Areal Unit Problem (MAUP) – Occurs when spatial zoning systems used to collect and/or analyze data is ‘modifiable’ or arbitrary. It can have two effects: scale effects, that result from spatial aggregation of the data, and zoning effects relate to changes in the spatial partitioning at a given level of spatial aggregation.

Noise Level – Sum of the percentage of intra-zonal trips and the percentage of trips in non statistically significant O/D matrix cell for a given zoning system. This indicator excludes the overlapping of both components (subtraction to the sum of both percentages).

TAZ (Traffic Analysis Zone, Transportation Analysis Zone or Traffic Assignment Zone) – A TAZ is the unit of geography most commonly used in conventional transportation planning models. The size of a zone varies, but for typical metropolitan planning software,
a zone of under 3000 people is common. The spatial extent of zones typically varies in models, ranging from very large areas in the exurbs to as small as city blocks or buildings in central business districts. There is no technical reason why zones cannot be as small as single buildings, however additional zones add to the computational burden.

**Transport Demand models** – Within the rational planning framework, transportation forecasts have traditionally followed the sequential four-step model or urban transportation planning (UTP) procedure, first implemented on mainframe computers in the 1950s at the Detroit Area Transportation Study (DATS) and Chicago Area Transportation Study (CATS).

Land use forecasting sets the stage for the process. Typically, forecasts are made for the region as a whole, e.g., of population growth. Such forecasts provide control totals for the local land use analysis. Typically, the region is divided into zones and by trend or regression analysis of the population and employment.

The four steps of the classical urban transportation planning system model are: trip generation, trip distribution, modal choice and traffic assignment.

**Transportation Planning studies** – Is the field involved with the sitting of transportation facilities (generally streets and highways and public transport lines). Transportation planning historically has followed the Rational Planning model of Defining Goals and Objectives, Identifying Problems, Generating Alternatives, Evaluating Alternatives, and Developing the Plan. Other models for planning include Rational actor, Satisficing, Incremental planning, Organizational process, and Political bargaining.
1. INTRODUCTION

TAZs have been pointed out as one of the keystones on Transportation Planning studies. Nevertheless, defining a good set of TAZ is still one of the transportation unsolved problems.

Transportation Planning studies have used transport demand models over the past four decades to forecast travel demand for short and long term planning. Transport Demand models typically follow a four-step process of trip generation, trip distribution, modal split, and network or trip assignment (Ortúzar and Willusen 2001). The steps are chained in a sequence, and the outputs of each step become inputs of the following step.

The key elements of a transport demand model are a study area divided into zones (called TAZ), and a transportation network in which each zone is represented by a centroid. TAZs represent areas from which and to which trips are allocated in a transport demand model. A centroid represents the ‘centre of activity’ of a zone and the origin and destination for all trips to and from the zone (Chang, Khatib et al. 2002). Trips generated between zones are assigned to the transportation network through connectors joining centroids and the physical network. Because zones and centroids are defined and used at the beginning of the modelling process, they affect the subsequent outputs, particularly the trip assignments on the network (Chang, Khatib et al. 2002).

In most Transportation Planning studies, a lot of effort is put in data collection, estimation of parameters and sophistication of models, but the issue of zoning rarely merits similar attention, normally being done on top of administrative units or ‘by common-sense’.

Dividing the territory in zones (for the purpose of Transport Modelling, and namely for the definition of O/D matrices), is a process of discretisation of space:

- With great advantages for the simplicity of the models and for the interpretation and communication of results.
But with insufficiently recognized (and managed) problems of loss of information in the process.

**Geographical information** is lost in the process of substituting (concentrating) the real origins and destinations of trips (which occur over a continuous space) by an artificial point in each zone, the centroid. In doing this, two types of loss occur:

- **Intra-zonal trips** cannot be processed (as they now start and end in the same point).
- The travelling paths close to the real origins and to destinations of the trips are subject to large errors, implying that traffic load estimates in the lower hierarchy links of the transportation network are highly unreliable.

If loss of information was only geographical, the solution would be simplified in terms of possible solutions: Adopting a high number of small zones. But this leads to a head-on collision with the loss of **statistical precision**, due to the fact that O/D matrices are always estimated through sampling processes, most frequently through direct survey of travellers and/or by traffic counts on network links (Viegas, Martínez et al. 2008). Adopting many small zones would lead to two problems:

- Width of Confidence intervals of the flow estimates.
- Percentage of matrix cells with zero flow.

At the same time, the TAZ concept has not been clearly defined, and their multiple uses sometimes are confused. Different zoning systems can be defined for different uses. Clear examples of that are the zoning system developed for sampling and computation of the expansion coefficient in a Mobility Survey, the zoning system developed for Transportation and Urban Planning analysis, and the zoning system defined for transportation gas emissions in Environmental assessments, which do not necessarily have to be exactly the same, but are often considered in the scientific literature as only one.
Lack of rigor about the TAZ concept led to a non uniform terminology in the scientific literature. Most authors use the Traffic Analysis Zones term, but the domain of this terminology is not clearly defined. Some authors use the Transportation Analysis Zones term, usually associated with the Transportation and Land Use Planning. The term of Traffic Assignment Zones is also often used in scientific literature and technical studies linked to traffic forecasting models (e.g. transportation 4 step classical model). Other terms as Traffic Zones, Travel Analysis Zones or just zones are also used by some authors, normally associated with the O/D matrixes establishment and traffic assignment in the transportation 4 steps model (Martínez, Viegas et al. 2007; Martínez, Viegas et al. 2009).

Through four decades of research, the scientific literature has established some guidelines and constraints to the definition of TAZ. Table 1.1 presents a complete set of the constraints discussed in the literature. Some contradictions and difficulties in their implementation are discussed in detail.

<table>
<thead>
<tr>
<th>Constraints</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trip generation/attraction homogeneity</td>
<td>(Baass 1981; O'Neill 1991; Ding 1994; 1998; Ortúzar and Willusen 2001; Chang, Khatib et al. 2002)</td>
</tr>
<tr>
<td>Contiguity and convexity of zones</td>
<td>(O'Neill 1991; Ding, Choi et al. 1993; Ding 1994; 1998)</td>
</tr>
<tr>
<td>Compactness of TAZ shapes</td>
<td>(Baass 1981; O'Neill 1991; Ding, Choi et al. 1993; Ding 1994; Chang, Khatib et al. 2002)</td>
</tr>
<tr>
<td>Exclusiveness (no doughnuts or islands) of zones</td>
<td>(O'Neill 1991; Ding, Choi et al. 1993; Ding 1994; 1998)</td>
</tr>
<tr>
<td>Equity in terms of trip generation (small standard deviation across zones)</td>
<td>(O'Neill 1991; Ding, Choi et al. 1993; Ding 1994; 1998)</td>
</tr>
<tr>
<td>Adjustment of TAZ boundaries to political, administrative, or statistical boundaries</td>
<td>(Baass 1981; O'Neill 1991; Ding, Choi et al. 1993; Ding 1994; 1998; Ortúzar and Willusen 2001; Chang, Khatib et al. 2002)</td>
</tr>
<tr>
<td>Respect of physical separators</td>
<td>(O'Neill 1991; Ding, Choi et al. 1993; Ding 1994; 1998)</td>
</tr>
<tr>
<td>Decision makers’ preferences are considered in determining the number of TAZ</td>
<td>(Ding, Choi et al. 1993; Ding 1994; 1998; Ortúzar and Willusen 2001)</td>
</tr>
</tbody>
</table>
Constraints

<table>
<thead>
<tr>
<th>Constraints</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoid main roads as zone boundaries</td>
<td>(Ortúzar and Willusen 2001)</td>
</tr>
<tr>
<td>Zone size is selected such that the aggregation error caused by the assumption</td>
<td>(Ortúzar and Willusen 2001)</td>
</tr>
<tr>
<td>that all activities are concentrated at the centroid is not too large (geographical precision)</td>
<td></td>
</tr>
<tr>
<td>Minimization of intra-zonal trips</td>
<td>(Baass 1981; Crevo 1991)</td>
</tr>
<tr>
<td>Maximization of the statistical precision of the estimation of the OD matrix cells</td>
<td>(Openshaw 1977)</td>
</tr>
</tbody>
</table>

It is difficult to consider and implement all of the above criteria in a single process of TAZ design because some rules contradict others (O’Neill 1991; Ding, Choi et al. 1993).

Some of these rules are a consequence of the use of a fixed zoning scheme over time, and of different study scales and purposes. In the scientific literature, it has been stipulated that transport demand modelling requires constant spatial data aggregation in TAZ along the entire process, from data collection to the trip assignment step (Chang, Khatib et al. 2002).

The number of zones problem can be solved through the development of a hierarchical zoning system, as in London Transportation Studies (Ortúzar and Willusen 2001), where subzones are aggregated into zones, which in turn are combined into districts, traffic boroughs, and finally a sector. This facilitates an analysis of different types of decisions at the appropriate level of detail (Ortúzar and Willusen 2001).

Unfortunately, predetermined zoning systems do not take into account the ongoing changes of land use (spatially and temporarily), which can deeply affect TAZ homogeneity and compactness, producing significant misestimates of trip generation and OD matrices (Edwards 1992).

These ongoing spatial and temporal land use changes demonstrate that current traffic management models, which are sometimes directly based on the results of the data collection process, and suppress the trip generation and distribution steps of the classical
travel demand model (Ding 1998), should use a different zoning scheme from the transportation demand forecasting models (Openshaw and Rao 1995). A better solution to the data collection zoning system constraint is the development of survey processes, in which the trip ends are geocoded. This data collection process requires the establishment of an initial zoning system, as with all other data collection processes, for the determination of expansion coefficients of each trip (Ortúzar and Willusen 2001); however, after the conclusion of this process, travel data is not attached to this or any other zoning system, but only to their geospatial coordinates (Chapleau 1997). Each study that uses this database can then develop a new zoning scheme that better fits the study scale and goals, resulting in more flexibility for the transportation analyst and a higher utility of the available database (Chapleau 1997; Trepanier and Chapleau 2001).

The complexity involved in TAZ definition underlies the need of further research on this topic. This report, building on previous research by the authors, tries to develop a new TAZ delineation methodology and a GIS based application, which does not only suits to one purpose (e.g. traffic assignment or 4 steps model definition), but that can also integrate under the same tool, different zoning schemes for different types of studies, allowing a more comprehensive analysis of mobility within Transport Planning Studies.

The new TAZ delineation approach is determined not only by the trips generation and distribution patterns, but is also constrained by other indicators that characterise other mobility and accessibility components. The integration of these indicators allows an holistic picture of mobility within a region and to identify at the very precise level the main zones of activities and traffic exchanges.

In order to be effective for mobility analysis and policy purposes, this kind of approach cannot only rely on the original cell grid unit but a hierarchical aggregation should be set up to ensure integration between the several resulting zoning schemes.
This new methodology will be tested on the Mobility Survey for the Paris Metropolitan Area in 2001 (EGT 2001), which will allow and easier illustration of the capabilities of this new modelling and planning tool.

1.1. Research Goals

The main objective of this project is to develop an algorithm that can automatically define a good set of TAZ, and then to export its results so that GIS and Transport Planning programs can be run on it. This action will grant that the results of TAZ can be used interactively in Land Use and Transportation Studies.

The expected results are the following:

- Develop tools and indicators to easily assess the performance of a particular zoning scheme (computation of information loss generated by the space discretisation).
- Present a new methodology and a set of guidelines on how to delineate TAZ.
- Devise a process that can routinely define TAZ in a way that can be used by GIS packages.
- Use a case study that can be set as reference in other studies when using the TAZ algorithm.

1.2. Structure of the Report

In order to present the developed research, the report was divided in three parts. This structure intends to show the research path undertaken, stressing the important choices taken along the way.

The first part (Chapter 2) of this research presents the methodological approach to TAZ definition, and presents the resulting TAZ Delineation Algorithm and its indicators of compliance.
The second part (Chapter 3) presents a data analysis of the EGT 2001, which will identify the main variables and indicators that will be used for the betterment of the TAZ Delineation Algorithm.

The third part (Chapter 4) presents the new approach to TAZ delineation for a comprehensive analysis of mobility. In this section is presented the path that led to the establishment of a new zoning methodology and application, presenting some specific applications of the new procedure to the Ilê-de-France using EGT 2001 data.

In the last part (Chapter 5), some conclusions and future developments of the research are presented, focusing in the value-added of a careful zoning and the new integrated approach and their implications over the usual Transportation Planning practice.
2. TAZ DELINEATION ALGORITHM

2.1. Problem Formulation

A methodology for zones delineation will be defined to reduce the noise level of the data for traffic modelling, and at the same time, to minimise the geographical error of the trip end location.

The methodology defines zones such that:

- Zone boundaries correspond to places with very low generation of trips (reducing the probability of misallocating trips to zones near zones boundaries);
- intra-zonal trips are minimised;
- the definition of zones with a very low number of trips or very large area (high geographical error) is avoided;
- the density of trip production inside a zone should be as homogeneous as possible.

Some of the constraints to definition of TAZ presented in the preceding chapter were not included in this methodology. These constraints were considered as local improvements to TAZ borders (e.g. adjustment of TAZ boundaries to political, administrative or statistical boundaries), and were introduced only at the end of the TAZ definition process. These local improvements are presented below.

2.2. Methodology of analysis and selection of zones limits

The methodology for the determination of zones starts by the aggregation of the geocoded trip ends (origin and destination) into a (relatively fine) cell grid. The cell grid can be variable and depends on the size of the study area and the precision intended for
the study (Viegas et al., 2008). For the Lisbon Metropolitan Area, a square cell grid a 200 m side length was used.

A thin plate spline\(^1\) was used to smooth the resulting surface and interpolate for cells without observations.

The result of this analysis is presented in Figure 2.1, where it is possible to identify the high concentration of trip ends in the Lisbon city centre and near some locations at the Lisbon municipality boundary.

The TAZ delineation algorithm uses the results of this analysis as background data, where each peak is the centre of a zone, the limits of which should be defined by the valleys surrounding it. The algorithm starts by identifying the local “highest peaks” and their surrounding area, sorts them by decreasing magnitude, and then uses a local search algorithm for the design of the zones. A search is performed for each peak considering a defined set of rules. These search rules, as presented in Figure 2.2, were developed to avoid the delineation of zones with complex spatial structures, which could undermine the applicability of the model in complex urban structures, as well as the assessment of its results, even for an experienced transportation analyst.

---

\(^1\) The thin plate spline is the two-dimensional analog of the cubic spline in one dimension. It is the fundamental solution to the biharmonic equation. The name “thin plate spline” refers to a physical analogy involving the bending of a thin sheet of metal.
These rules were translated into equations considering a matrix composed for each local “highest peak” located in the centre of the search matrix. In order to codify these rules, a set of variables were used. An example with the explanation of these variables is presented in Figure 2.3.

The search rules equations present two different cases for the same level of aggregation, as presented in Figure 2.3: Case 1, where $|x - d|$ is different from $|y - d|$ (except for when $x = d$ and $y = d$, seed cell of aggregation process); and Case 2, where $|x - d|$ is equal to $|y - d|$. This differentiation inside each level of aggregation is caused by the
need of aggregation of cells of the same level, prior to a Case 2 aggregation (see Figure 2.2).

The search rules equations contain two different statements: the first statement is used for the cell identification in the search matrix (e.g. \( n = 0 \) and \( m \neq 0 \) and \( m \neq 2 \times l \)), and the second is the requirements for aggregation of this cell in the search matrix (vicinity conditions) (e.g. If matrix \((d + m - l, d - l + 1) = 1\)). There are nine different equations that restrain the search matrix, five for the Case 1 and four for Case 2. These rules and its application are not presented in detail in this paper, but are described in Martínez (2006).

The TAZ delineation algorithm is defined by five different constraints and an objective function with two variables. These constraints can be divided in two groups: Those derived from the algorithm (4 constraints) and the geographic constraint for the TAZ border delineation (avoiding overlapping between zones). The constraints derived from the algorithm are:

1. The total origins \((O_i)\) or destinations \((D_i)\) of trips of each TAZ should be greater than 70\%\(^2\) of the average origins or destinations of trips by zone (total origins or destinations of trips divided by the number of zones). This is a requirement for quasi-homogeneity of trip quantities across zones, which indirectly controls the relative statistical error of the resulting zones.\(^3\)
2. Each TAZ area should be at least 70\%\(^4\) of the size of the influence area\(^5\) of local predefined “highest peaks,”\(^6\) which avoids the formation of zones with very low geographic precision.

---

\(^2\) This value was obtained for the case study after calibration using different size grid cells.
\(^3\) The relative statistical error is presented in Equation 2.
\(^4\) This value was obtained for the case study after calibration using different size grid cells.
\(^5\) The surrounding area size is a parameter of the algorithm that is defined by the user.
\(^6\) The concept of influence area of local “highest peaks” is defined as the minimum size that an analyst establishes for the specific modelling problem.
3. The average statistical (relative) error in the estimation of OD flow matrix cells should be lower than 50%, which directly controls the statistical precision of each TAZ.

4. The number of zones should fall within the range previously defined by the analyst, which forces the algorithm to follow the analyst’s preferences.

These are the main constraints of the algorithm. The other constraint (geographic constraint) is only used when a TAZ has already been delineated and forces TAZ limits to reach a frontier of an already formed TAZ (finding another boundary and stopping the search by defining the outer boundary of that specific zone) in order to obtain a zoning scheme that covers the entire study area and produces no overlaps.

The first constraint is defined to avoid the definition of zones with a very low and heterogeneous statistical precision in the resulting zoning schemes, as this is strongly correlated with the number of trip origins or destinations of trips per zone. The second constraint is intended to avoid the formation of very small zones that could have good geographic and statistical precision, but would lead to heterogeneous geographic precision from a global point of view (large zones to compensate for other small zones). These two indicators work as an indirect control over the statistical and geographic precision of the TAZ delineation. Of course, the 70% thresholds proposed in this application could be revised for a different application.

The third constraint attempts to directly control the statistical precision in the estimation of the OD matrix cells of each TAZ through the average statistical error of each matrix cell, assuming a distribution of trips from each TAZ as the origin, to all TAZs, proportionally to the total flow in those zones as destinations. This constraint does not guarantee that all of the cells of the OD matrix, in which the zone is the origin, are statistically significant, but considerably increases the likelihood that it is so (Martínez 2006).

The fourth constraint ensures the previously defined range for the number of zones. This constraint is also used as one of the stopping criteria of the algorithm when it cannot
be satisfied because the given range of is too high for the available number of cells and data.

The objective function of the TAZ algorithm contains two different components: The density of trips and the percentage of intra-zonal trips of each zone. This objective function tries, simultaneously, to optimise these variables by minimizing the standard deviation of the density of trips (across cells inside each TAZ), thus leading to more homogeneous zones and minimizing the sum of the percentages of intra-zonal trips across all zones.

These components can have minimum values at different points, with trade-offs solved through the use of a ranking function, which minimises the sum of the rankings of the two variables (see Figure 2.4).

If the ranking function retrieves the same result for different cases, the objective function considers the result with the lower trip density standard deviation to be the “most suitable.” The decision tree used for the objective function is presented in Figure 2.4.

Figure 2.4 – Decision tree for determining the optimum of the TAZ algorithm objective function

The algorithm requires setting some parameters, which are important for the establishment of constraints, as well as for the optimization of some additional features of the algorithm (e.g., the optimal number of zones for a given range, which depends on some macroscopic indicators that will be presented below).
These parameters are:

- The definition of the range of number of zones (searching for the most suitable solution) – compulsory input;
- minimum size of the influence area of local “highest peaks” for TAZ delineation;
- definition of a core problem area\(^7\) – compulsory input;
- percentage of zones belonging to the core problem area;
- maximum proportion of areas between zones in the core problem area and in the “rest of the world;”
- some parameters of the indicators (macroscopic indicators) used to define the optimal number of zones of the given range. This optimal number is found with the help of a multi-criteria additive function, which is defined below.

If the user omits some of these parameters, and they are not compulsory, the algorithm uses the default values of those parameters in the algorithm (e.g., percentage of zones belonging to the core problem area equal to 50%).

After an overview of the TAZ delineation algorithm, the mathematical and data flow of the algorithm are presented to facilitate a deeper understanding of its mechanics.

### 2.3. TAZ Delineation Algorithm Structure

This section describes all of the input data needed for the algorithm, the algorithm (functioning and data flow), and the equations used in the model (constraints and objective function).

The model needs as basic input:

---

\(^7\) The core problem area parameter allows the analyst to establish different levels of relevance inside the modelling area, which does not need to be used, and considers this parameter to be equal to the entire modelling area.
• The square grid cell, which will be used to form the TAZs (this input has a very strong impact on the delineation quality of the zones and the running time of the algorithm);
• the geocoded trip ends to be aggregated into spatial grid cells;
• the total number of trips of the survey, which is used to determine the relative statistical error of estimation of the OD matrix;
• the definition of the core problem area, which is a subset of the study area on which the study is focused; its size relative to the entire study area could have a significant impact on the result of the zone delineation process due to the parameters associated with it, e.g., percentage of zones in the core problem area);
• the definition of the local “highest peaks” (seed of the TAZ formation) influence area;
• the definition of the parameters of the multi-criteria function to be used for selecting the optimal number of zones based on the macroscopic indicators produced by the algorithm.

Using all of the inputs above, particularly the square cell grid and the range of the number of zones, the algorithm starts by ordering all of the cells in a decreasing order of their value of T (total origins and destinations of trips per cell). The algorithm is then structured with three different cycles (see Figure 2.5). The first cycle is responsible for the calculation of the zoning scheme for the different numbers of zones contained in the given range. This cycle iterates from L, the lower bound of that range, to U, the upper bound of that range, and contains the two other cycles that are responsible for the delineation of each TAZ.

The second cycle delineates each TAZ. The cycle starts by defining the local “highest peak,” the initiator of the cell aggregation for the TAZ delineation, and its spatial search. This cycle iterates until the set of zones has a dimension of N (the number of zones given by the first cycle), or until the number of analysed cells (possible local “highest peaks”) reaches the total number of cells available.
The third cycle is responsible for the aggregation of cells for each TAZ. In the beginning of each TAZ delineation, the local “highest peak” search is empty. In the first iteration, the local “highest peak” is inserted at the centre search, thus being the seed of the delineation process. The following iterations aggregate cells to the local “highest peak” using the search rules, where the objective function determines the most suitable cell to aggregate to the TAZ from the set of cells that respect the spreading rules in each iteration (minimization of the standard deviation of the trip density of TAZ and the intrazonal trips).

After the determination of all of the zoning schemes for the available range of number of zones \((L)\), the algorithm uses a compensatory rule to determine the optimal number of zones for the given range (see Figure 2.5). The compensatory rule uses the macroscopic indicators for each zoning scheme that are presented below as attributes. All the attributes \((M_i)\) were previously scaled to values between 0 and 1 using a linear function between the maximum and the minimum available values of the attributes. The weights of the attributes \((w_i)\) of the compensatory rule are specified by the user, as indicated above. The global compensatory function that should be maximised is presented in (1), where \(N\) is tested between \(L\) and \(U\) (lower and upper bounds of the range of number of zones in the input) for the different indicators \((MI) – \text{number of macroscopic indicators})

\[
\max_{L-U} \sum_{i=0}^{MI} \left( \frac{\max(M_i) - M_i}{\max(M_i) - \min(M_i)} \right) \cdot w_i
\]  

(1)

The algorithm selects a number of zones within the available range that maximises the global compensatory function and stores the result (numeric and spatial data) in a warehouse that can be accessed with GIS software. All of the microscopic and macroscopic indicators of the analysed zoning schemes are stored in a spreadsheet, making it possible for the user to quickly assess the results.
L – Lower bound of the range for the number of zones
U – Upper bound of the range for the number of zones

Input Grid Cell (with the trip ends aggregated into cells)

Decreasing Order (T)

N=L
Do Until N=U
Next L
Calc Macroscopic Indicators

i=0
Do Until i > Z
Next i
K=N
Or i=Z

Spatial search of Cell i

Cell i

j = 0
Do Until j > Z
Next j

Spatial search of Cell i
Searching Rules

Step 1: Initialization for aggregation for each local highest peak (future TAZ)
Cycle for the aggregation of each cell to a TAZ, for the N number of zones

Continue until constraints satisfied

Goal Function

\[ \min \sum_{k=1}^{k-1} \left( A_j - \frac{T_{ij}}{A_j} \right)^2 \]

Compensatory Model for the determination of the optimal number of zones for the given range

Macroscopic Indicators – M_i
MI – Number of Macroscopic Indicators

\[ \max \sum_{i=1}^{n} \left( \max(M_i) - M_i \right) / \left( \max(M_i) - \min(M_i) \right) \]

Figure 2.5 – Delineation algorithm functioning and data flow
2.4. TAZ indicators of compliance

The TAZ delineation algorithm presented above uses several indicators of compliance. Some of these are used to verify the constraints or calculate the objective function, while others are only used to characterise each TAZ or zoning scheme. These indicators are either microscopic or macroscopic, depending on the unit they are describing: Microscopic indicators are calculated for each TAZ, while macroscopic indicators are calculated for a zoning scheme. These indicators are dynamically recomputed during the algorithm run after the aggregation of a new cell to a TAZ, as stated above (see Table 2.1).

The macroscopic indicators of the TAZ delineation algorithm were developed to evaluate the resulting zoning schemes and to determine the optimal number of zones within the defined range using a compensatory rule. These indicators are only obtained at the end of running the algorithm, using in some cases the final TAZ microscopic indicators for their calculation (see Table 2.1). All of the indicators in Table 2.1 are presented in detail and discussed by Martínez (2006).
## Table 2.1 - Microscopic and Macroscopic Indicator Summary

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Microscopic Indicators</th>
<th>Geographic Error</th>
<th>Equivalent radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAZ AREA (m)</td>
<td></td>
<td></td>
<td>$R_E \approx \sqrt{\frac{TAZ_AREA}{\pi}}$</td>
</tr>
<tr>
<td></td>
<td>Statistical Error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>$\left( \frac{SWCI}{p} \right)<em>ij \approx \sqrt{\frac{Z^2</em>\alpha}{n} \left( I - p_{ij} \right)}$ (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Global average</td>
<td></td>
<td>$G_{AVER_STAT} = \frac{\sum_{i=1}^{N} X_{STAT} _i}{N}$ (9)</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>$\sigma_{TD} = \frac{\sum_{i=1}^{N} \left( x_{TD} - \left( \frac{O_i + D_i}{A_i} \right) \right)^2}{N - I}$ (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>$CV_{TD} = \frac{\sigma_{TD}}{x_{TD}}$ (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td></td>
<td>$MAX_{STAT} = \max_{ij} \left{ \frac{SWCI}{p} \right}_ij$ (4)</td>
</tr>
<tr>
<td></td>
<td>75th percentile of the equivalent radius</td>
<td></td>
<td>$R_E (75%) = P_{75}\left[ \text{Sort Ascending} \left{ R_E \right} \right]$ (10)</td>
</tr>
<tr>
<td></td>
<td>Trip Density</td>
<td></td>
<td>$G_{C_VTD} = \frac{\sum_{i=1}^{N} \left( \frac{\sum_{i=1}^{N} x_{TD} _i}{N} - x_{TD} _i \right)^2}{\sum_{i=1}^{N} x_{TD} _i}$ (11)</td>
</tr>
</tbody>
</table>

---

*Where $i$ is the origin zone and $j$ the destination zone in the resulting OD matrix, $p_{ij}$ the probability of a trip between $I$ and $j$, SWCI is the semi-width of the confidence interval for cell $ij$ in the resulting OD matrix and $Z_\alpha$ is the value of a standard normal distribution for a level of confidence $\alpha$.**
2.5. Local Improvements to the TAZ Delineation Algorithm

Some local improvements to the TAZ delineation algorithm were also developed in order to improve the results and turn the algorithm more flexible to different study areas. Below is a list of the primary developed improvements.

- A border conversion algorithm that adjusts the borders of TAZ (obtained from the square cell grid) into administrative or statistical land units, which allows a better understanding of the zoning systems results, easier TAZ data assessment, and a more feasible way to use the results as a transportation and urban planning tool. This algorithm takes as its objective function the maximisation of the overlapping areas between the borders of TAZ and the administrative or statistical land units, and as constraints, the preservation of the number of zones of the zoning system and the continuity of zones (split zones not accepted).

- A bi-criterion optimization algorithm that locally optimises the statistical precision of the transport demand estimates associated to the zoning scheme by local reassignment of frontier cells from one zone to its neighbour.

- An urban barriers correction algorithm that respects physical geographic separators of the territory, such as railways, rivers, and common urban and mobility barriers. This algorithm determines the splitting areas generated by the intersection of each TAZ with the defined urban barriers, identifies the TAZ members (cells or institutional borders), and reassigns the cells/institutional borders located “outside” the urban barriers to the neighbouring TAZs.

- A major flow corridors correction algorithm that adjusts the TAZ borders to the main study area flow corridors in order to maximise the intersection area of TAZ with these corridors (making only some frontier cell rearrangements), ensuring that these major flows have as small an intra-zonal representation as possible. The algorithm considers two types of constraints: The preservation of the number of zones and the continuity of zones (split zones are not accepted).
These four local improvement algorithms lead to zoning systems with better performances over the variables and constraints considered relevant for TAZ definition. A more detailed assessment of these algorithms is given by Martínez (2006).
3. ASSESSMENT OF EGT DATA

This chapter presents a characterisation of the Ille-de-France Region based on the data collected by the EGT 2001. First it will be performed a geographical assessment of the region followed by the analysis of the data included on this survey:

- Household characterization
- Mobility characterization
  - Trip length, start time, duration and transport mode
  - Tour purpose, destination, etc.
  - Energy consumption and emissions
- Accessibility and land-use characterization
  - Accessibility to Public Transport station
  - Population density (computed with the resulting zones)
  - Work and study density (computed with the resulting zones)

After this initial assessment, a group of variables or indicators will be selected to integrate the new TAZ Delineation Algorithm. This selection will be performed based on the analysis of their geographical impact over mobility and transport and land use planning.

For this purpose, several spatial analysis techniques will be implied as Cluster Analysis among other Classification Methods. For this spatial analysis two types of spatial units will be used:

- The communes of the Ille-de-France region

---

9 Energy consumption and emissions were estimated using the reference values of the EC TREMOVE Project.
• A zoning scheme resulting from the TAZ Delineation Algorithm with 1500 zones, which main concern is on mobility and it is mainly oriented for traffic assignment models within the 4 steps model.

3.1. The region of the Île-de-France

Île-de-France is one of the twenty-six administrative regions of France, composed mostly of the Paris metropolitan area with the size of 12,211 Km$^2$ (0.86% in Paris Municipality).

Created as the "District of the Paris Region" in 1961, it was renamed after the historic province of "Isle de France" in 1976, when its administrative status was aligned with the other French administrative regions created in 1972 (see Figure 3.1).

Figure 3.1 – Map of Île-de-France location within France
The Île-de-France corresponds also to a NUTS2 EU statistical division of the territory, and it is formed by 8 NUT3, which do also correspond the department administrative division. The Île-de-France region is formed by 1281 communes (20 of then inside Paris Municipality – Arrondissements), which is the French smallest administrative division (see Figure 3.2).

Figure 3.2 – Map of the Île-de-France communes

With 11.7 million inhabitants Île-de-France is the most populated region of France\(^ {10}\). It has more residents than Austria, Belgium, Greece, Portugal or Sweden, and a comparable population to the US state of Ohio or the Canadian province of Ontario. It is the fourth most populous country subdivision in the European Union after England (of the UK), North Rhine-Westphalia and Bavaria (both of Germany).

\(^{10}\) The data from EGT 2001 indicates a population of 10,964,044 inhabitants (19% in Paris Municipality)
Economically, Île-de-France is the sixth richest region in the European Union: in 2006 its total GDP as calculated by Eurostat was €462 billion at market exchange rates, with a per capita GDP of € 40,100 the same year (at market exchange rates, 170% of the European Union average)\textsuperscript{11}.

After this initial assessment of the Île-de-France region, it is relevant to make a short summary of the mobility survey that will be the base of the following analysis: the EGT 2001.

This mobility survey comprehends all the regions of the Île-de-France and was developed in 2001. This mobility survey contains 23,656 respondents from different households, sampling a total of 81,386 trips (0.23% of the total trips). The trip ends, and the residential location of the respondent are geocoded using a 300 square meters grid cell.

The survey includes information of the composition of the household of the respondent, the respondent characterization, and the mobility of the respondent, discretising all the steps within the daily mobility chain of the respondent, having all the transport mode transfers geocoded and characterised.

Figure 3.3 presents the spatial distribution of the origins of trips of the EGT 2001 mobility survey. It is easy to notice a higher concentration of trip origins within the central departments of the study area and mainly within the Paris Municipality.

It is also possible to identify from the figure, the existence of four main axes of mobility to reach Paris centre, corresponding to locations with better accessibility to the public transport network and the road network.

There also a significant number of communes that do present a low number of trips generated, reflecting locations with lower population density and activities concentration with the Île-de-France region.

\textsuperscript{11} Eurostat data.
3.2. Household characterization

This section presents an assessment of the household characterization of the Ille-de-France region using the EGT 2001 data. The main variables that will be analysed are:

- Home ownership and dwelling characterisation
- Household composition and income distribution
- Characterization of the respondent
- Motorisation and private car use of the household

The assessment of the data will be performed at the commune and department geographical level, trying to identifying some spatial pattern with the region. It is relevant to state that just 49% of the communes within the region are represented at the EGT.
2001 sample with households that reside there, which can limit the spatial assessment. However, this fact reveals a greater population concentration within some more urban communes of the region, and very low population density within 50% of the communes, which should be mainly rural.

3.2.1. Home Ownership and Dwelling Characterization

Using the data provided by the EGT 2001, some statistics about the home ownership were computed. The average percentage of home ownership for the Ilê-de-France (weighted by the number of households within each commune) is of approximately 34%. The spatial distribution of this indicator can be assessed in Figure 3.4, where it is presented the percentage of home owners for each commune. The results suggest a greater percentage of home owners outside the Paris city centre.

![Figure 3.4 – Spatial distribution of the percentage of home owners within Ilê-de-France region](image)

Relatively to the type of dwelling that households live in, approximately 66% of them live in an apartment. The spatial distribution of this indicator can be assessed in Figure 3.5, where it is presented the percentage of households living in apartments for each commune. The results show that the communes located within the Petite Couronne
(Inner Ring) and in the department of Val-d'Oise present higher percentages of households living in apartments, which suggest denser urban concentrations.

Figure 3.5 – Spatial distribution of the percentage of households living in apartments within Île-de-France region

3.2.2. Household Composition and Income Distribution

The household composition and income, is also a relevant feature of the households for its mobility assessment. The average size of the household for the study area is of 2.47 members, which 0.21 are averagely children under 6 years old. The spatial distribution of these indicators can be assessed in Figure 3.6 and Figure 3.7. The spatial distribution suggests the existence of bigger families outside Paris and with a greater number of small children.
Figure 3.6 – Spatial distribution of the average household size within Ille-de-France region

Figure 3.7 – Spatial distribution of the average number of children under 6 years old within Ille-de-France region
The average income per household for all the study area is 25,647 €. Figure 3.8 shows that higher income households do not necessarily live in Paris city centre. There are several communes far from Paris that present a significantly high average income. This indicator has to be considered carefully because there is a significant statistical error linked with respondents far from the Paris city centre, due to reduced number of households from those areas which can introduce a significant bias.

![Spatial distribution of the average income of the households within Île-de-France region](image)

**Figure 3.8 – Spatial distribution of the average income of the households within Île-de-France region**

The number of active persons in the household is also a significant determinant of the mobility patterns. The average number of active persons per household for the Île-de-France is 1.12, which represent less than 50% of the average size of the household. The spatial distribution of this indicator is presented in Figure 3.9, where it is possible to notice that households with higher number of active persons are located outside the Petite Coronne.
3.2.3. Characterization of the respondent

Some of the individual attributes of the respondent to the mobility survey are also analysed in this point. The indicators analysed are the age and the profession, which can be determinants to the observed mobility patterns. The average age of the respondents in the EGT 2001 survey is approximately 49 years old. The spatial distribution of this indicator shows younger respondents at Paris city centre (see Figure 3.10) and the highest average respondents’ age communes are located far from Paris.

The overall profession distribution or activity status of the EGT 2001 respondents is presented in Table 3.1. The results show a very significant percentage of retired persons answering the survey, followed by liberal and intermediate professionals. The rate of unemployed persons answering the survey is also considerable; however the students’ rate is quite low.
Figure 3.10 – Spatial distribution of the average age of the respondents within Ille-de-France region

Table 3.1 – Aggregate statistics of the profession distribution or activity status of EGT 2001 respondents

<table>
<thead>
<tr>
<th>Profession</th>
<th>Average Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>0.18%</td>
</tr>
<tr>
<td>Craftsman</td>
<td>3.75%</td>
</tr>
<tr>
<td>Liberal Professional</td>
<td>17.55%</td>
</tr>
<tr>
<td>Intermediate Professional</td>
<td>17.03%</td>
</tr>
<tr>
<td>Employer</td>
<td>15.34%</td>
</tr>
<tr>
<td>Operative</td>
<td>11.89%</td>
</tr>
<tr>
<td>Student</td>
<td>1.70%</td>
</tr>
<tr>
<td>Retired</td>
<td>24.98%</td>
</tr>
<tr>
<td>Unemployed</td>
<td>7.58%</td>
</tr>
</tbody>
</table>

In order to analyse the spatial pattern of respondents’ residence towards their professions or activity status, it were computed for each commune the average percentage of respondents attached to each profession or activity status. Figure 3.11 presents the percentage of retired respondents, where a higher rate of retired respondents is observed in the Outer ring area (Grande Couronne). Figure 3.12 does not show a strong pattern of location of liberal professional, which higher rates are scattered around the Illé-de-France.
Figure 3.11 – Spatial distribution of the average percentage of retired respondents within EGT 2001 database

Figure 3.12 – Spatial distribution of the average percentage of liberal professional respondents within EGT 2001 database
3.2.4. Motorisation and private car use of the household

This point analyses another group of variables relevant for the mobility assessment: the number of motorised vehicles available in the household and the number of kilometres travelled using these vehicles. The average number of motorised vehicles per household is 1.07 and the average number of kilometres travelled is 15,910. Figure 3.13 presents the spatial distribution of the average number of motorised vehicles, which tends to be higher outside the Petite Couronne. Figure 3.14 presents a similar trend for the average number of kilometres travelled, which can be explained by good motorway accessibility and worse public transport accessibility of communes far from Paris.

Figure 3.13 – Spatial distribution of the average number of motorised vehicles per household within Ilë-de-France region
3.3. Mobility characterization

In this point will be assessed some main mobility features that can computed from the EGT 2001 survey. The main mobility variables of trips and tour that will be assessed are:

- Trip start time, length, duration and transport mode;
- Tour purpose and destination;
- Energy consumption and emissions (based on TREMOVE model).

The first indicator that will be analysed is the distribution of trip starting times during the day in order to measure the peak concentration of trips. Figure 3.15 shows the trip starting time distribution along the day, showing a morning peak hour between 7 and 9 am of 19.85% of the trips and an afternoon peak hour between 16 and 18 of 20.60% of
the trips. This information shows a slightly more concentrated peak on the afternoon, which is linked to more flexible arrival to work times.

Figure 3.15 – Trips start time distribution of EGT 2001 survey

Figure 3.16 illustrates that 50% of the trips within the survey are smaller than 2 km and 95% of the trips are smaller than 30 km. Analysing the distribution for different modes, a tendency to smaller trips for walking and metro modes can be observed, opposed to the RER and private car. This fact shows the existence of a large quantity of trips that might be performed walking/biking due its influence on the global curve. This fact indicates the relevance of the proximity mobility within a large metropolitan area as the Ille-de-France.

The trip duration analysis is presented in Figure 3.17. This figure shows that 50% of the trips are shorter than 20 min., and that 95% of the trips are shorter than 75 min. Analysing the different modes, the walking trips and the private cars tend to be shorter than the average, and the RER and Metro take more time. This fact reflects the time consumed in public transportation tends to be higher than in private car for the same
distance, which leads to a significant disadvantage of these modes for the mode share distribution.

**Figure 3.16 – Trip length distribution of the EGT 2001 survey**

**Figure 3.17 – Trip duration distribution of the EGT 2001 survey**
The mode share can be assessed in Figure 3.18. The values show that walking and the private car are the main transportation modes within the Ilê-de-France, followed by the main public transport modes: the Metro and the RER.

![Figure 3.18 – Mode share distribution of the EGT 2001 survey](image)

Another relevant mobility indicator is the trip purpose distribution. The assessment of this indicator is presented in Figure 3.19. The most relevant trip purpose is to return home, followed by the personal matters and going to work. The percentage of non-commuting trips is also very significant (approximately 38%), which is nowadays a common feature of developed societies.

The assessment of the spatial distribution of trip destinations is presented in Figure 3.20. The results show a great concentration of trips within Paris and some locations inside the Petite Couronne, and some new agglomerations more peripheral to the region, mainly in the west area of the Ilê-de-France.
Figure 3.19 – Trip purpose distribution of the EGT 2001 survey

Figure 3.20 – Spatial distribution of destinations of trips in the EGT 2001 survey
In order to perform an environmental assessment of the EGT 2001 data, it was used 
the data and coefficients defined on the EC TREMOVE model. TREMOVE is a policy 
assessment model, designed to study the effects of different transport and environment 
policies on the emissions of the transport sector. The model estimates for policies as road 
pricing, public transport pricing, emission standards, subsidies for cleaner cars etc., the 
transport demand, modal shifts, vehicle stock renewal and scrappage decisions as well as 
the emissions of air pollutants and the welfare level.

TREMOVE models both passenger and freight transport, and covers the period 1995-
2030. TREMOVE is in fact 2 models: a land transport model, and a maritime model. The 
land transport model has been set up to model all transport within 1 country. At the 
moment, input databases are calibrated to feed the model for 31 countries.

The TREMOVE 2.52 model has been developed by Transport & Mobility Leuven in a 
service contract for the European Commission, DG Environment. The most recent 
TREMOVE 2.7 version includes further developments made in the 6th Framework 
Programme GRACE project.

The coefficients used from the TREMOVE model are presented in Table 3.2. This table 
presents different emissions coefficients of the different pollutants for the different 
transport modes and vehicles. The coefficients present average values per 
Millions.Passkm, which are independent of the congestion level of the system. In order to 
introduce this factor, it was considered a time dependent correction coefficient based on 
the distribution of Pass.km along a regular week day. This correction coefficient 
introduces emissions variations that result from congestion, leading to higher emissions 
coefficients during peak hours and smaller coefficients during the rest of the day. This 
coefficient \(Ce\), for a given hour, follows hipergeometric standardization as presented in 
(12). The maximum variation is set to 25% of the average value (0.75 to 1.25).

\[
Ce_i = \frac{\max(\text{Passkm}) - \text{Passkm}}{\max(\text{Passkm}) - \min(\text{Passkm})}
\]
Table 3.2 – Coefficients of TREMOVE models for different transport modes and vehicles

<table>
<thead>
<tr>
<th>Code</th>
<th>Mode</th>
<th>Sum of pkm (Millions passkm) or vkm (Millions vkm)</th>
<th>Sum of CO2 exhaust (tons/Millions passkm)</th>
<th>Sum of CO well to tank (tons/Millions passkm)</th>
<th>Sum of NOx well to tank (tons/Millions passkm)</th>
<th>Sum of PM well to tank (tons/Millions passkm)</th>
<th>Sum of VOC well to tank (tons/Millions passkm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Passenger Railcar Electric</td>
<td>13485.7822</td>
<td>6.1738</td>
<td>0.0009</td>
<td>0.0192</td>
<td>0.0018</td>
<td>0.0106</td>
</tr>
<tr>
<td>2</td>
<td>Passenger Locomotive Electric</td>
<td>15194.5725</td>
<td>2.6269</td>
<td>0.0004</td>
<td>0.0082</td>
<td>0.0007</td>
<td>0.0045</td>
</tr>
<tr>
<td>3</td>
<td>Bus - Diesel</td>
<td>42995.9322</td>
<td>60.1264</td>
<td>0.1801</td>
<td>0.7160</td>
<td>0.0321</td>
<td>0.0834</td>
</tr>
<tr>
<td>4</td>
<td>Car &gt;2.0L - Diesel</td>
<td>35197.1771</td>
<td>291.1052</td>
<td>0.3813</td>
<td>0.8196</td>
<td>0.1223</td>
<td>0.1883</td>
</tr>
<tr>
<td>5</td>
<td>Car 1.4-2.0L - Diesel</td>
<td>130114.3173</td>
<td>215.5595</td>
<td>0.3001</td>
<td>0.7110</td>
<td>0.0930</td>
<td>0.1168</td>
</tr>
<tr>
<td>6</td>
<td>Car &lt;1.4L - Diesel</td>
<td>17485.1463</td>
<td>136.8588</td>
<td>0.0932</td>
<td>0.5403</td>
<td>0.0326</td>
<td>0.0714</td>
</tr>
<tr>
<td>7</td>
<td>Car &lt;1.4L - Petrol</td>
<td>118470.2063</td>
<td>210.7459</td>
<td>6.5875</td>
<td>1.1241</td>
<td>0.0237</td>
<td>1.1904</td>
</tr>
<tr>
<td>8</td>
<td>Car 1.4-2.0L - Petrol</td>
<td>105440.0163</td>
<td>245.7696</td>
<td>5.4035</td>
<td>1.2604</td>
<td>0.0272</td>
<td>1.0753</td>
</tr>
<tr>
<td>9</td>
<td>Car &gt;2.0L - Petrol</td>
<td>13502.5311</td>
<td>320.7008</td>
<td>4.4202</td>
<td>1.1300</td>
<td>0.0349</td>
<td>0.7341</td>
</tr>
<tr>
<td>10</td>
<td>Car &lt;1.4L - Cng</td>
<td>3.8880</td>
<td>136.1137</td>
<td>0.1587</td>
<td>0.0517</td>
<td>0.0028</td>
<td>0.3467</td>
</tr>
<tr>
<td>11</td>
<td>Car 1.4-2.0L - Cng</td>
<td>4.2720</td>
<td>162.1231</td>
<td>0.1451</td>
<td>0.0606</td>
<td>0.0040</td>
<td>0.4129</td>
</tr>
<tr>
<td>12</td>
<td>Car &gt;2.0L - Petrol</td>
<td>1.0900</td>
<td>202.4018</td>
<td>0.1376</td>
<td>0.0706</td>
<td>0.0037</td>
<td>0.5138</td>
</tr>
<tr>
<td>13</td>
<td>Mopeds &lt; 0.05L</td>
<td>4307.2370</td>
<td>71.9323</td>
<td>8.9414</td>
<td>0.0665</td>
<td>0.1337</td>
<td>7.5901</td>
</tr>
<tr>
<td>14</td>
<td>Motorcycle 2 Stroke</td>
<td>500.3190</td>
<td>114.2609</td>
<td>22.9383</td>
<td>0.1457</td>
<td>0.1776</td>
<td>7.5495</td>
</tr>
<tr>
<td>15</td>
<td>Van</td>
<td>64078.3832</td>
<td>231.2544</td>
<td>0.6838</td>
<td>1.0344</td>
<td>0.2020</td>
<td>0.1891</td>
</tr>
<tr>
<td>16</td>
<td>Plane &lt;500Km</td>
<td>228.7250</td>
<td>94.0820</td>
<td>0.2874</td>
<td>0.3744</td>
<td>0.0235</td>
<td>0.2625</td>
</tr>
<tr>
<td>17</td>
<td>High Speed Rail</td>
<td>33252.9341</td>
<td>6.3656</td>
<td>0.0009</td>
<td>0.0198</td>
<td>0.0018</td>
<td>0.0109</td>
</tr>
<tr>
<td>18</td>
<td>Metro/Tram</td>
<td>9522.0000</td>
<td>2.0750</td>
<td>0.0003</td>
<td>0.0064</td>
<td>0.0006</td>
<td>0.0035</td>
</tr>
<tr>
<td>19</td>
<td>Walk</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
After this explanation of the procedure used for the emissions estimation, Table 3.3 presents the weekly assessment of emissions of the different pollutants of the trips of the EGT 2001 survey.

The results show that 50% of the trips produce less than 1.51 kg of CO₂ per week, considering that the trip is repeated 5 times per week. This value is equivalent to perform during a week 10 km in a gasoline car with a motor <1.4l., which is significantly low, showing the weight of the walking and public transport trips in the survey. The 95% percentile is equivalent to 170 km for the same type of vehicle, which shows the huge variation of emissions for different types of trips (trip length and transport mode).

<table>
<thead>
<tr>
<th>Percentile</th>
<th>CO₂ [kg]</th>
<th>CO [kg]</th>
<th>NOₓ [kg]</th>
<th>PC [kg]</th>
<th>COV [kg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>0.010</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>5%</td>
<td>0.050</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>10%</td>
<td>0.100</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>15%</td>
<td>0.190</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>20%</td>
<td>0.340</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>25%</td>
<td>0.480</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>30%</td>
<td>0.620</td>
<td>0.001</td>
<td>0.003</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>35%</td>
<td>0.790</td>
<td>0.001</td>
<td>0.004</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>40%</td>
<td>0.980</td>
<td>0.001</td>
<td>0.005</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>45%</td>
<td>1.210</td>
<td>0.002</td>
<td>0.006</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>50%</td>
<td>1.510</td>
<td>0.002</td>
<td>0.007</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>55%</td>
<td>1.920</td>
<td>0.002</td>
<td>0.009</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>60%</td>
<td>2.470</td>
<td>0.003</td>
<td>0.011</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>65%</td>
<td>3.190</td>
<td>0.004</td>
<td>0.014</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>70%</td>
<td>4.170</td>
<td>0.005</td>
<td>0.017</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>75%</td>
<td>5.550</td>
<td>0.006</td>
<td>0.022</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>80%</td>
<td>7.550</td>
<td>0.009</td>
<td>0.028</td>
<td>0.003</td>
<td>0.004</td>
</tr>
<tr>
<td>85%</td>
<td>10.520</td>
<td>0.012</td>
<td>0.039</td>
<td>0.004</td>
<td>0.006</td>
</tr>
<tr>
<td>90%</td>
<td>15.530</td>
<td>0.019</td>
<td>0.056</td>
<td>0.006</td>
<td>0.010</td>
</tr>
<tr>
<td>95%</td>
<td>25.469</td>
<td>0.039</td>
<td>0.092</td>
<td>0.010</td>
<td>0.017</td>
</tr>
<tr>
<td>100%</td>
<td>175.990</td>
<td>11.529</td>
<td>0.499</td>
<td>0.095</td>
<td>3.794</td>
</tr>
</tbody>
</table>
Figure 3.21 shows the distribution of weekly emissions of the households surveyed on the EGT 2001. The results show that 50% of the households produce less than 12 kg of CO₂ per week.

![CO2 Emissions Distribution](image)

**Figure 3.21 – Distribution of weekly emissions of households of the EGT 2001 survey**

3.4. **Accessibility and land-use characterization**

In this point will be assessed some accessibility and land use indicators, which are relevant determinants for mobility generation/attraction. These indicators are computed from the EGT 2001 survey. The indicators that will be analysed are:

- Accessibility to Public Transport station
- Population density
- Work and study density

The public transport statistical distribution is presented in Figure 3.22. The figure shows that 50% of trip origins are located less than 670 m away from a heavy public transport station and 95% less than 3.2 km from a heavy public transport station. This information shows good public transport accessibility near to the main generators of
mobility, having the public transport modes always as a viable option to travel. Figure 3.23 illustrates the spatial distribution of the accessibility, which presents more favourable values within Paris and some agglomerations in the Petite Couronne.

Figure 3.22 – Public transport accessibility distribution of the households surveyed on the EGT 2001

Figure 3.23 – Public transport accessibility spatial distribution on the Ille-de-France region
The population density is one of the key factors for mobility assessment due to its high correlation with trip generation/attraction. The spatial distribution of the population density is presented in Figure 3.24, where a higher density within the Petite Couronne and some agglomerations along new development axis.

The overlap between the population density and the public transport maps shows the relevance of this land use indicator and its relevance for the design of the public transport system.

![Population density spatial distribution of the Ille-de-France region](image)

Figure 3.24 – Population density spatial distribution of the Ille-de-France region

The work and study density is also a key determinant to mobility generation/attraction due to the importance of commuting trips in the total number of trips performed, as it was previously analysed. Figure 3.25 presents the spatial distribution of this indicator, which overlaps significantly the population density map, showing a planning concern of developing mixed-use agglomerations in order to reduce the needs of mobility within the region. The greatest concentrations of work and studies are located inside Paris, mainly in its central Arrondissements.
3.5. Analysis of the Indicators to integrate the new TAZ Delineation Algorithm

After the comprehensive assessment of the data available on the EGT 2001 survey, it will be performed an analysis of these indicators with two main goals:

- Assess the relevance of some indicators for the mobility analysis of the region in order to include them in the new TAZ Delineation Algorithm;
- Compare the resulting spatial structures with the existing planning zoning scheme for the Île-de-France region (IAURIF).

With this purpose, it was developed a cluster analysis of these indicators using zoning scheme resulting from the TAZ Delineation Algorithm, having in mind the future application of this zoning scheme for the hierarchical procedure.
The zoning scheme used is presented in Figure 3.26. This TAZ configuration presents 1500 and was developed for traffic assignment purposes, presenting the characteristics defined in Chapter 2. Table 3.4 presents some indicators of this zoning scheme. The zoning scheme presents overall good quality indicators, with the exception of the statistical precision indicators, derived from the large number of zones used for traffic assignment purposes.

![Figure 3.26 – 1500 zones scheme of the Ilê-de-France region (TAZ Delineation Algorithm)](image)

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of trips inner zone</td>
<td>28.701%</td>
</tr>
<tr>
<td>Average origin or destination per zone (%)</td>
<td>0.067%</td>
</tr>
<tr>
<td>Maximum origin or destination per zone (%)</td>
<td>0.513%</td>
</tr>
<tr>
<td>Maximum flow</td>
<td>61,325,000</td>
</tr>
<tr>
<td>Maximum flow percentage</td>
<td>0.178%</td>
</tr>
<tr>
<td>Average flow</td>
<td>15.347</td>
</tr>
<tr>
<td>Percentage of null flows</td>
<td>97.166%</td>
</tr>
<tr>
<td>Percentage of trips in non significant O/D matrix cells</td>
<td>75.336%</td>
</tr>
<tr>
<td>Average of TAZ average statistical Relative error</td>
<td>41.077</td>
</tr>
<tr>
<td>Global Average Trip Density (trips/ha.)</td>
<td>279.581</td>
</tr>
<tr>
<td>Average Trips Density (Cv)</td>
<td>2.151</td>
</tr>
<tr>
<td>Average zone equivalent radius</td>
<td>958.595</td>
</tr>
</tbody>
</table>
Some of the indicators previously analysed were computed for the new zoning scheme to perform the clusters analysis. The indicators selected to perform the clusters analysis resulted from the authors’ observation of the main determinants to mobility within the available data and that might have transport and land use planning implications. These indicators are:

- Emissions factor (Kg CO2/Passkm.ha.)
- Population Density (Inhabitant/ha.)
- Work or Study Trips Density (Trips/ha.)
- Average PT Accessibility (m)

In order to introduce a more aggregate assessment of the data it was performed a percentile assessment of the different indicators and defined three levels for the different indicators: high, medium and low.

These values resulted from the following percentiles:

- **High** >= Percentile 75% of the indicator
- Percentile 25% of the indicator < **Medium** < Percentile 75% of the indicator
- **Low** <= Percentile 25% of the indicator

The spatial distribution of the 4 indicators with these 3 levels is presented in Figure 3.27, Figure 3.28, Figure 3.29, and Figure 3.30. The spatial patterns of the 4 different indicators illustrate some similarities, where Paris tends to present always high values, followed by the agglomeration around it and some main axis of development in the region.

The indicator that presents greater spatial variation is the emissions factor indicator. This fact derives from the smaller mode share of less intensive emission transport modes far from the centre of the region. Although some new agglomerations present significant land use mixture, it remains a functional dependency of the region towards Paris, which leads to higher mobility requirements in areas with higher car dependency.
Figure 3.27 – Percentile assessment of the Emissions factor indicator for the Ille-de-France region

Figure 3.28 – Percentile assessment of the Population density indicator for the Ille-de-France region
Figure 3.29 – Percentile assessment of the Public transport indicator for the Ille-de-France region

Figure 3.30 – Percentile assessment of the Work/study density indicator for the Ille-de-France region
After this initial classification, a factorial design was performed resulting in 81 different profiles ($3^4$ cases). These profiles result of the combination of the different levels of each zone for the different indicators.

The numeration of the profiles was ordered from high to low for each indicator, resulting that profile 1 includes high levels for the 4 indicators and profile 81 low levels for all the indicators.

The spatialisation of this analysis is presented in Figure 3.31. The figure confirms once again Paris and the agglomerations around it as the main centralities, followed by some secondary agglomerations developed along some new development axis. These results are aligned with the planning and functional zoning developed for the Île-de-France region (IUARIF). This zoning scheme is presented in Figure 3.32, where 8 types of zones are identified, showing the same developing trends identified in Figure 3.31. The main differences between the two zoning scheme results from the more disaggregate description of some mobility characteristics within the macro-zones identified in the IUARIF zoning.
After this global assessment of the region including all the indicators, it was developed and analysis considering in each step one of the indicators as the main determinant of the mobility of the region. This analysis was performed considering a standardization value for the high level of 1, 0.5 for the medium level and 0 to the low level for each indicator.

In order to give a higher importance to a given indicator, it was considered a compensatory model were the main indicator presents a weight of 0.50 and a homogeneous weight distribution among all the other indicators. This model results in a score for each zone between 0 and 1 for each weight composition. The results of this analysis are presented in Figure 3.33, Figure 3.34, Figure 3.35 and Figure 3.36. The different zoning schemes obtained present very similar patterns once again and confirm some of the conclusions stated above.
2010

TAZ Delineation Application for Transport Planning Studies: A New Approach Applied to the Ilê-de-France (Paris)

Figure 3.33 – Cluster analysis of the Ilê-de-France (Emissions factor as main indicator)

Figure 3.34 – Cluster analysis of the Ilê-de-France (Population density as main indicator)
Figure 3.35 – Cluster analysis of the Ille-de-France (Public transport accessibility as main indicator)

Figure 3.36 – Cluster analysis of the Ille-de-France (Work/study density as main indicator)
After this comprehensive assessment of these indicators, it is possible to state their relevance as mobility determinants and relevant transport and land use planning indicators.

For this reason, these indicators will be used in the following steps of this research as the main indicators to measure the homogeneity within a zone in a new TAZ Delineation procedure that tries to go beyond traffic assignment purposes.
4. NEW ZONING APPLICATION: A HIERARCHICAL ZONING APPROACH

The current chapter presents a new zoning delineation application which tries to include the results the previously defined zoning scheme to produce an efficient space discretisation for another transport planning and analysis purpose.

This new approach introduces the concept of hierarchical zoning, due to the fact the zoning schemes result from previously aggregated zones and not from square grid cells. This new methodology allows integrating results from previous stages to make models computing and assessment compatible.

The concept of hierarchical zoning has been already used in literature and in practical application as in the London TAZ system (Ortúzar and Willusen 2001); however a systematic and algorithmic approach has not been already developed.

In the current study, the use of hierarchical zoning is used to build new zoning scheme configurations, initially idealised for traffic assignment models purposes, to compute new transport models that use more aggregate data (e.g. energy consumption and emissions assessment).

This tool will then allow moving easily from original zoning schemes of 1500 zones to more aggregate zoning schemes, compatible with other transport models use and planning assessment, preserving the main characteristics set initially for quality zoning:

- Zone boundaries correspond to places with very low generation of trips (reducing the probability of misallocating trips to zones near zones boundaries);
- intra-zonal trips are minimised;
- the definition of zones with a very low number of trips or very large area (high geographical error) is avoided.
The main difference introduced by the new model is the definition of homogeneity measures within the same TAZ, which initially was designed to measure the density of trip generation, and now presents indicators to assess this homogeneity.

There also some differences relative to the vicinity definition, which for grid cells were defined by same equation and aggregation levels (see Chapter 2), and in the new algorithm, due to the irregular boundaries configuration, has to be defined using spatial queries and boundaries overlap rules.

The following points of this chapter will present the new TAZ delineation methodology, the main differences introduced by the new algorithm, and the application of the new algorithm to the Ilê-de-France region.

4.1. New Problem Formulation

The new methodological approach for zones delineation is defined in order to keep the goals of the original algorithm (reduce the noise level of the data for traffic modelling, and at the same time, to minimise the geographical error of the trip end location), but also include some new indicators relevant of transport policy assessment in order to measure the homogeneity within the same zone.

The new methodology defines zones such that:

- Zone boundaries correspond to places with very low generation of trips (reducing the probability of misallocating trips to zones near zones boundaries);
- intra-zonal trips are minimised;
- the definition of zones with a very low number of trips or very large area (high geographical error) is avoided;
- the homogeneity of the selected indicators within the same zone is maximized.
The selected indicators to use to measure the zone homogeneity were defined in the previous chapter. These indicators are:

- Emissions factor (Kg CO2/Passkm.ha.)
- Population Density (Inhabitant/ha.)
- Work or Study Trips Density (Trips/ha.)
- Average Public Transport Accessibility (m)

At the same time, the new aggregation unit are no longer grid cells, which will result in a different set of rule for the vicinity assessment. These rules, as in the base algorithm, try to avoid the delineation of zones with complex spatial structures, which could undermine the applicability of the model in complex urban structures, as well as the assessment of its results, even for an experienced transportation analyst.

4.2. Changes to the Original TAZ Delineation Algorithm

The main changes introduced by the new algorithm in order to use the new hierarchical approach are:

- The definition of new vicinity rules based irregular boundaries overlap;
- The definition of a new objective function and constraints set for the algorithm.

4.2.1. Vicinity Rules

The introduction of new vicinity rules in the TAZ hierarchical procedure is due to the possibility of use of irregular borders zones as seed for the aggregation process. The algorithm starts, as the original algorithm, by identifying the local “highest peaks” and their surrounding area, sorts them by decreasing magnitude, and then uses a local search algorithm for the design of the zones.
This search is performed for each peak considering a defined set of rules. These search rules, as presented in Figure 4.1, were developed to avoid the delineation of zones with complex spatial structures, which could undermine the applicability of the model in complex urban structures, as well as the assessment of its results, even for an experienced transportation analyst. This local search algorithm presents now a new set of rules, which use as main indicator the percentage of TAZ boundary perimeter overlap between the already formed zones and the candidate zone, considering as minimum overlap percentage 5% of the perimeter of the already formed TAZ.

As presented Figure 4.1, the new rule for vicinity assessment, presents a simple mathematical definition, however the computational burden needed for the dynamic computing of the TAZ boundary perimeter overlap is considerably higher that the original regular grid search rules.

As in the original regular grid search algorithm, the algorithm establishes vicinity levels, which depend on the number of previously aggregated zones in order to be able to aggregate it to the TAZ. An example of these levels is also presented in Figure 4.1.

![Figure 4.1 – New search rules example](image-url)
4.2.2. Objective Function and Constraints

The objective function of the Hierarchical TAZ algorithm contains two different components as original algorithm: the percentage of intra-zonal trips of each zone and an aggregated Euclidean distance measure of the new set of indicators within each zone. These new indicators were identified in Chapter 3 and are:

- Emissions factor (Kg CO2/Passkm.ha.)
- Population Density (Inhabitant/ha.)
- Work or Study Trips Density (Trips/ha.)
- Average Public Transport Accessibility (m)

The Euclidean distance function is computed after a hypergeometric standardization of the indicators \((N)\) as presented in (13). The distance function is then computed by the difference between the standardized indicators \((NS)\) of the already aggregated zones and the zone under analysis. The distance of each indicator is then weighted by multiplying the standardized distance by a weight factor of each indicator \((wn)\) (considered equal among all indicators by default). This distance function \((DN)\) is presented in (14).

\[
NS_i = \frac{\max(N_i) - N_i}{\max(N_i) - \min(N_i)}
\]

\[
DN_j = \sum NS_i \cdot wn_i
\]

This objective function tries, simultaneously, to optimise these variables by minimizing the distance measure of the new indicators (across cells inside each TAZ), thus leading to more homogeneous zones and minimizing the sum of the percentages of intra-zonal trips across all zones.
These components can have minimum values at different points, with trade-offs solved through the use of a ranking function, which minimises the sum of the rankings of the two variables (see Figure 4.2).

If the ranking function retrieves the same result for different cases, the objective function considers the result with the lower trip density standard deviation to be the “most suitable.” The decision tree used for the objective function is presented in Figure 4.2.

![Figure 4.2 – Decision tree for determining the optimum of the Hierarchical TAZ algorithm objective function](image)

All the parameters and constraints of the model remain the same from the original algorithm, with the addition of four new parameters related with the weight of the different indicators on the Euclidean distance function of the objective function.

If the user omits some of these parameters, the algorithm uses the default values of those parameters (equal weight for all the indicators).

### 4.3. Application to the Region of the Ilê-de-France

After the establishment of the new TAZ Delineation hierarchical procedure, an application of the new algorithm will be presented for the Ilê-de-France region based on the EGT 2001 survey.
The hierarchical procedure will use as base zoning scheme the TAZ system developed in Chapter 3 with 1500 zones (see Figure 3.26 and Table 3.4). This TAZ configuration was developed using the original TAZ Delineation Algorithm and had as main purpose the development of a zoning scheme appropriate to traffic assignment models.

In order to test the new hierarchical algorithm, two different tests were performed.

- A 50 zones scheme that could be applied for transport planning and energy consumption and emissions assessment in the transport sector;
- and a 75 zones scheme for the same purposes.

The obtained results are presented in Figure 4.3 and Figure 4.4. The presented results map the spatial distribution of the emission factor, which leads to a configuration where the agglomeration ring around Paris and the some secondary agglomerations present the highest values of this coefficient. These results are coherent with the assessment presented in Chapter 3, which is due to the greater private car use in areas outside Paris.

It is also relevant to notice that the borders of the resulting zones cross frequently the administrative limits of the departments, which indicates urban developments and activity interactions that not always follow the administrative divisions.

The larger zones are located in more rural areas with more disperse land use patterns and lower trip generation/attraction rates. This fact leads to a trade-off between statistical precision and homogeneity of zones and the geographical precision, which for planning purposes is not as relevant as for traffic assignment models.

Comparing the two zoning schemes developed, a significant overlap between zones can be observed. The main differences result from the subdivision of some zones located within the Petite Couronne, which can present a higher geographical precision for a greater number of zones. The global indicators as the average equivalent radius change from 4646 m to 3842 m, and the average statistical error from 0.51 to 0.70.

Both resulting zoning systems seem adequate for transport planning assessment.
TAZ Delineation Application for Transport Planning Studies: A New Approach Applied to the Ilê-de-France (Paris)

Figure 4.3 – TAZ 50 zones (Application of the Hierarchical TAZ Delineation Algorithm)
Figure 4.4 – TAZ 75 zones (Application of the Hierarchical TAZ Delineation Algorithm)
5. CONCLUSIONS AND FUTURE DEVELOPMENTS

In most Transport Planning studies, a lot of effort is put in data collection, estimation of parameters and sophistication of models, but the issue of zoning rarely merits similar attention, normally being done on top of administrative units or “by common-sense”. The results obtained show that zoning is not a trivial matter giving the significant consequences it may have for the generation of statistical and geographical errors. This study introduces a new methodology of zones delineation that deals with the existent trade-offs in the process.

The introduction of the new Hierarchical Delineation Algorithm enhances the capabilities of the original algorithm, leading to more flexible TAZ delineation process that can be used for multiple purposes.

The application developed for this study focuses on new indicators that are considered determinants of mobility. These new indicators are used to measure the homogeneity within a zone, leading to zoning schemes that are more suited for transport planning.

Applying the new hierarchical zoning procedure over the results of the original TAZ Delineation Algorithm, ensures the conservation of the “good traffic assignment properties” of the original algorithm and simultaneously gather areas with similar values in relevant land use, accessibility and energy consumption attributes.

The results show a great potential in the use of this integrated approach that can easily combine previous studies to other layers of analysis and more comprehensive and holistic assessments. This step forward allows simultaneously achieving good zoning schemes for different purposes and models and using their outputs as inputs for different models due to the hierarchical design used.

The use of this algorithm for different case studies (Lisbon Metropolitan Area and the Îl-de-France region) means a step forward on the validation of this procedure. Testing
the in a different database allowed the tuning of the model and the assessment of the sensitivity of the results to different travel patterns and land use distributions.

Four main directions are envisaged for the next steps of this research:

- Look at the consequence of these different zoning strategies on the traffic load estimates on roads of intermediate and lower hierarchies, especially after the implementation of all the local refinements to the base algorithm.
- Investigate the consequences of these findings for the process of matrix estimation from traffic counts, being aware that the assignment of a trip to a zone directly affects the TAZ definition process. For this reason, studies based on traffic counts might have a simultaneous definition of the zoning system and the OD matrix, and not as a exogenous and pre-defined process. This iterative process until the convergence of the zoning systems and OD matrix might have a similar structure to the k-clustering algorithms, where the centroids of the clusters change position with every iteration.
- Economically quantify the improvements in statistical precision and the reduction of percentage of intra-zonal trips obtained by the use of the algorithm. This significant decrease of information loss can lead to the use of smaller samples in data collection or to a significant increase of the quality of the data used in typical transportation planning studies, obtaining more robust results for the same input data and costs.
- Evaluate the possible application of the algorithm with other indicators resulting from zonal data rather travel vector data (e.g., population vs. employment). Both resulting zoning systems seem adequate for transport planning assessment.
BIBLIOGRAPHY


