Understanding overlapping functional commuting regions with confidence ellipses and social network methods

Working Paper

Author(s):
Killer, Veronika; Axhausen, Kay W.

Publication date:
2011-07

Permanent link:
https://doi.org/10.3929/ethz-a-010189042

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Arbeitsberichte Verkehrs- und Raumplanung 714
Understanding overlapping functional commuting regions with confidence ellipses and social network methods

Appendix: Application of the confidence ellipse in a monocentric perspective

Veronika Killer

Kay W. Axhausen
Understanding overlapping functional commuting regions with confidence ellipses and social network methods

Veronika Killer  
IVT  
ETH Hönggerberg (HIL)  
CH-8093 Zürich  
Telefon: +41-1-633 39 43  
Telefax: +41-1-633 10 57  
veronika.killer@ivt.baug.ethz.ch

Kay W. Axhausen  
IVT  
ETH Hönggerberg (HIL)  
CH-8093 Zürich  
Telefon: +41-1-633 39 43  
Telefax: +41-1-633 10 57  
avhausen@ivt.baug.ethz.ch

July 2011

Summary

Functional regions are becoming more and more prominent in spatial planning policies. In this study we assume, that it will be problematic to clearly delimitate these regions in the future. Functional regions will increasingly overlap, because of increasing linkages of functional flows.

This study explores different approaches of identifying about these overlapping or fuzzy functional regions, where some municipalities belong to more then one region. Methods are developed and evaluated to quantify and qualify these overlaps. Two methods bring together the perspective of statistical and social network analysis. The first method is an enhancement of the bivariate confidence ellipse. The social network method used to sub-divide large social networks, community detection.

The results vary according to the chosen method. Generally, overlaps can be analysed on different hierarchy levels. If the functional overlaps of some neighbouring municipalities are heigh, they will be aggregated to a functional region on a higher hierarchy level. The confidence interval method can be interpreted as the lowest hierarchy level. The methods of social networks show different hierarchies at higher levels.

Keywords

Commuting region, Travel-to-Work-Area, confidence interval, social networks, community detection, Switzerland

Preferred citation style

Figures

Figure 1  Conceptual thinking about overlaps ..........................................................13
Figure 2  Change in distance commuting over time in Switzerland...........................18
Figure 3  Commuting frequency function for long distance correction .................20
Figure 4  Mean regions’ size by different confidence interval ...............................24
Figure 5  Calculation of ellipsoid overlaps (O, R, L) .............................................26
Figure 6  Overlapping communities of Zürich with cFinder – 6-cliques ...................29
Figure 7  Overlapping commuting area based of two hierarchies- social network method..........................................................33
Figure 8  Three alternatives of confidence ellipse’s absolute overlap ...................34
Figure 9  Increasing total overlaps of confidence ellipse over time ......................36
Figure 10 Increasing real overlap with confidence ellipse weighted with commuting intensity over time ...............................................................36
Figure 11 COPRA algorithm: Overlaps and functional regions over time ............37
Figure 12 FLM algorithm with adjusted cities ......................................................39
Figure 13 Incommuting region of large Swiss cities ............................................40
Figure 14 Intensity of commuting within commuting region (see F2/F3) .............42
Figure 15 Distribution possibilities of commuting trips within a commuting region ....43
Figure 16 Relative-distance-index .....................................................................44
Figure 17 Inverse-Gini-coefficient ....................................................................45
Figure 18 Analyses of the additional parameters – angle, shape, and drift ..................48

Figure 19 Identifying commuting linkage with Herfindahl-index .....................................49

Figure 20 Change of commuting region mean size of municipality type 1970-2000 ..50
Tables

Table 1  Approaches to delimit commuting regions .................................................10
Table 2  Aggregation of the mode of transport of commuting.................................15
Table 3  Analysis of uncertainties of the historical commuting data.........................17
Table 4  Different versions of historical commuting dataset...................................20
Table 5  Different versions of commuting regions for all Suisse municipalities (2000) ..........................................................23
Table 6  Top 10 municipalities - confidence ellipse’s overlap ...................................32
Table 7  Case study: Change of commuter flows 1970-2000 .....................................49
1 Introduction

Functional regions are becoming more and more prominent in spatial planning policies around Europe. This study is assuming that it will be problematically to clearly delimitate these regions in the future, because of increasing complexity of functional flows. The overlap of these functional regions will be extended and are not necessary adjacent regions. We focus on the functional commuting flows, which are one of the best investigated functional connections. Several studies clearly highlight a development from separate centre specific commuting flows in 1970 to a more complex and multicentre or peripheral commuting flow structure today (eg. Killer and Axhausen 2000b; Dessemontet et al. 2010). However, which structure produces more functional overlaps? We want to answer questions of quality and quantity of these overlapping areas. How are these overlapping functional regions structured and where are they located in the year 2000 and within the period 1970-2000 in Switzerland?

Why are overlapping functional regions of interest in the future? First, there could be political impacts: Areas of overlaps are attractive places for work and residential location. They profit from the functional closeness to more then one centre. However, they are affected by higher infrastructure costs because they provide efficient connections in different directions. These overlapping areas could be parts of Switzerland that are insufficiently incorporated into the economy if the functional connection is weak in all directions. Finally, we hope to find better functional categorisations that could improve transport or residential choice models.

Recently, different disciplines attempted to explain these overlapping functional regions based on commuting flows. This research brings together the perspective on overlapping functional regions from different disciplines. The analysis of overlapping functional regions is an essential gap, which has been started to be filled up only recently. A literature review about the different disciplines and their concepts of non-overlapping and overlapping functional regions is provided in chapter 2.

Two methods out of these different approaches are tested in more detail. The first method is an enhancement of the bivariate confidence ellipse. The second method is coming from the field of social network analysis and is used for sub-dividing network community detection. These two methods are explained in chapter 4. The method is applied to the historical commuting data of Switzerland for 1970-2000 and results are discussed in chapter 5. In the appendix we test some further possibilities of application of the confidence ellipse.
In this section functional regions are analysed in a monocentric and not overlapping “polycentric” perspective. It discusses the quality of the ellipse method and not the structure of overlapping municipalities.
2 Concepts of functional commuting regions and their overlaps

Several studies investigate thoroughly different concepts of non-overlapping functional regions based on commuting flows. The delimited regions exhibit functional similarities for research and policy-making purposes. The use and practise of functionally defined regions varies between countries. Different spatial categories are the result. The categorisation serves e.g. as labor markets, a statistical and political unit. They should be explicitly comparable. Therefore researcher and politicians prefer clearly separable aggregations. They rarely discuss overlap.

Two main concepts of functional regions exist (Drobne et al., 2010). First, the region is a closed urban system with a high degree of self-containment or maximization of intra-regional flows. The person should live and work within the same region. The second defines a region by a centre with its surrounding “belt” or commuting hinterland composed of all those areas from which people commute to the particular centre. Based on these two concepts different techniques to limit functional regions exists. Table 1 categorizes different calculation techniques of different disciplines. The approaches belong to a category if it is mentioned in the relevant literature or could be extended easily (marked with X). Some algorithms differs only slightly, others have fundamentally different approaches and are founded in different research traditions. However, they are all going back to one of these two main concepts internal maximization of flows or centre-dependence. Some representatives are selected for every approach and at least one approach per category is investigating overlapping or fuzzy functional regions.

Threshold based multistep procedure

Feng (2009) uses explicitly a fuzzy method, where some “belt” municipalities belong to different travel-to-work-areas (TTWA). Travel-to-work-areas are criticised being imperfect because there are always commuting trips crossing them. He investigates the fuzziness of travel-to-work areas by applying fuzzy set theory. So that municipalities belong to more than one functional region to a certain degree depending on their commuting share.
Table 1  Approaches to delimit commuting regions

<table>
<thead>
<tr>
<th>Approach</th>
<th>Literature for commuting analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Threshold based multistep procedure</strong></td>
<td>Coombes et al. (1986); Coombes and Casado-Dias (2000); Papps (2002)</td>
</tr>
<tr>
<td>Travel-To-Work Area by self containment principle</td>
<td>X  X  X  X  X  X  X</td>
</tr>
<tr>
<td>Agglomerations</td>
<td>Prodromidis (2010); Karlson and Ollson (2006); Schuler et al. (2005)</td>
</tr>
<tr>
<td>Fuzzy Travel-to-Work Area</td>
<td>Feng, (2009)</td>
</tr>
<tr>
<td><strong>Cluster and maximisation algorithm</strong></td>
<td>Feldman et al. (2005)</td>
</tr>
<tr>
<td>Multivariate clustering eg. Intramax</td>
<td>X  X  X  X  X  X  X</td>
</tr>
<tr>
<td>Community detection in social networks</td>
<td>Farmer et. al. (forthcoming)</td>
</tr>
<tr>
<td>Overlapping community detection in social networks</td>
<td>Lanchichinetti et al. (2010)</td>
</tr>
<tr>
<td><strong>Simple variable procedure</strong></td>
<td>Baumann et al. (1988)</td>
</tr>
<tr>
<td>Travel-Time distance</td>
<td>X  X  X  X  X  X  X</td>
</tr>
<tr>
<td>Uni- or bivariate commuting variables</td>
<td>Botte (2003); Axhausen and Killer, (2010a/b)</td>
</tr>
<tr>
<td><strong>Urban economic model</strong></td>
<td>Alonso (1964)</td>
</tr>
<tr>
<td>Urban economic monocentric model</td>
<td>X  X  X  X  X  X  X</td>
</tr>
<tr>
<td>New Economic Geography</td>
<td>Gerritse (forthcoming)</td>
</tr>
<tr>
<td></td>
<td>X  X  X  X  X  X  X</td>
</tr>
</tbody>
</table>
All these approaches can be calibrated with different parameters. The delimitations can be adapted flexibly by setting the appropriate threshold. This ability to adapt strengths these approaches and makes them politically acceptable. However, these approaches include often manual steps and are not repeatable and produce their politically acceptance by additional secondary correction.

**Cluster and maximisation algorithm**

Cluster algorithms use different aggregation techniques for the interaction matrices. The different approaches are reviewed and tested for example by Masser and Scheurwater (1980). One is the Markov analytic functional distance approach applied e.g. in Cörvers, Hensen, and Bongaerts (2009). This method is a widely used regionalisation method. The cells of the MFPT matrix represent functional distances between municipalities. Then internal flows are maximised based on this distance. Other statistical clustering approaches are applied by e.g. Brown and Holmes (1971) or Feldman et al. (2005).

These purely statistical and calculative approaches have been criticised by Coombes et al. (1986: 946). They argued that “the disadvantage of purely statistical techniques is that, apart from one or two operational decisions, their application is largely deterministic, and precludes the necessary 'fine tuning' of parameters which is essential if a single global procedure is to produce adequate results across the wide variety of Labour Markets (from coalfields to resorts, conurbations to crafting communities)“.

Community detection is coming from social network analysis, describing the problem of dividing a network into subgroups. Groups should be found that are stronger connected within them than with the rest of the network. In contrast to other partitioning methods the number of groups and their size is not fixed. Traditional methods are e.g. graph partitioning or hierarchical clustering (Fortunato, 2010). Other division are produced by optimising internal to external degree of a cluster e.g. modularity optimization methods (Newman, 2010). Modularity measure quantifies a division of a network into modules or communities. Good divisions, which have high values of the modularity, are those which have dense internal connections between the nodes within modules but only sparse connections between different modules. Recently, new approaches are developed which focus on detection of overlapping community structures (for more detail see chapter 4.2). Some basic methods of social network analysis are already been applied to detect commuting pattern (e.g. Dessemontet et al, 2010). One work especially working with community detection and functional regions will possibly be
published soon (Farmer et. al. forthcoming). The methods for overlapping functional regions are not studied in detail yet. The algorithm of Lancichinetti et al. (2010) has been successfully evaluated tested for the commuting network of Great Britain.

**Simple variable procedure**

This approach is strongly depending on a centre. The commute distance can be drawn for example as a circle around a centre (e.g. Killer and Axhausen, 2010b) or all municipalities within a certain travel time e.g. 45, 60, or 90 minutes (e.g. Baumann et al., 1988). A bivariate statistical approach, the confidence ellipse, is more flexible in shape in a further dimension, which is tested in Botte (2003) and Killer and Axhausen (2010a/b). The technique is explained in more detail in chapter 4.1.

The advantage of these simple approaches is their short calculating times, because there is no iterative procedure. First the centre has to be chosen then the size of the region is strongly dependent on the chosen threshold of the distance variable. This concept cannot be expended to intra-regional maximisation of the flows. The accessibly based approach (Vitins et al., 2010, Johansson et al., 2003) is not strictly centre oriented. The accessibly measure is not only looking at the travel time to one centre but to all surrounding municipalities. The approach of Vitins et al. (2010) explicitly tries to include tangential flows e.g. from one suburban municipality to another. They are not centre directed. Three-step-procedure makes the algorithm more complex and time consuming and could be also characterised as a “threshold based multistep procedure”. Overlaps in all these approaches are determined by belonging to more then one centre (e.g. Killer and Axhausen, 2010a).

**Urban economic model**

The urban economic model is certainly the most abstract model defining a functional region. The functional region is represented regularly by a circular shape, which size is depending on price, transport cost, and total population. By its classical form it has a monocentric not overlapping perspective. This monocentric urban model describes a closed urban system, where every active person is living and working within it. It has one central business district. It assumes the existence of a featureless plain, a homogenous surrounding. The size of the urban region is based on the number of inhabitants and individual travel cost (Alonso, 1964). This model does not take in account interactions between cities. Regional economics is looking at

--

1 The content and title of this work is still unknown; therefore it is not mentioned in the references.
the interaction and spatial organisation of different cities (e.g. Christaller, 1933). However, these classical models are not overlapping. Based New Economic Geography models with a polycentric perspective have been developed recently (e.g. Gerritse, forthcoming\(^2\)).

Non overlapping regionalisation has a long tradition and overlapping concepts are in most disciplines just in development. First approaches and calculation techniques exist. However, no literature is found about the structure of these overlapping areas and how they act the policy and planning process. The idea of overlapping areas can be linked to the overall discussion about polycentricism. This literature describes overlapping areas in this polycentric perspective as commuting towards different centres. This study tries to open up the possible overlaps in defines them as areas where people commute in different directions. Some municipalities belong to more than one centre (see Figure 1 a). Other possibly overlapping areas are characterised by disperse commuting flows which can be independent of a centre – an area between (see Figure 1 b). That means that municipalities are not inside a dominant functional region\(^3\). The categorisation of “Cluster and maximisation algorithm” is rather one of these centre independent approaches and could identify an “area between”.

Figure 1 Conceptual thinking about overlaps

---

2 Michael Gerritse is the “Epainos” prize winner of ERSA 2010. His paper is only a draft version and should not be cited. So this work is not in the reference list.

3 A dominant functional region can also been described as a region with a high quality considering chapter 7.1.
3 Data

3.1 Commuting matrices 1970-2000

This section is bored by Botte, 2003; Fröhlich 2008; Frick et al. 2004. Fröhlich (2008) has been working with a smaller spatial level within large cities. This chapter 3 discusses new and already known findings.

A harmonized dataset of commuting flows exists for the years 1970, 1980, 1990, and 2000 for municipalities as 2000. The collection of commuting data was part of the federal census. Commuting data are available electronically since 1970 census. The commuters’ relations between municipalities are registered where one person at least commutes in one of the four years. Pupils and students are not included in these commuting matrices. The following list discusses changes in data collection and matrices preparation between censuses 1970-2000. This knowledge is necessary for interpretation of the results.

- Originally, the census 1970 and 1980 include people with minimum six hours working per week. In the census 1990 and 2000 this minimum threshold changes to one hour work per week. Because of consistency, only an employee over six hours work per week is considered in these commuting matrices (Fröhlich 2008, Frick et al. 2004).

- Fröhlich (2008) analysis and documents the modes of these commuting matrices. Census asks about the main mode of transportation to commute to work. If more than one mode of transportation is used for that trip the following hierarchy is applied since the census of 1980. Public transport is more important than the individual transport, a fast mode is more important than a slow one. In 1970, the workers undertake the hierarchical assignment themselves. For the aggregation of transport modes Fröhlich (2008) suggests following classification (see Table 2):
Considering the total number of commutes it is simply suggested to add up the three categories non-motorized traffic, private transport, and public transport. Commuters are defined by the Swiss Federal Office of Statistic as workers over 15 years which leave their living place to reach their work place. Other dictionaries suggest that the workplace has to be a minimum or even a long distance away from home. Other focus on the idea that the commuter has to cross a municipality’s or township borders e.g. is commuting from suburban workplace to urban workplace. A last group stress more the fact that a commuter needs to use a transport mode between home and work location. Depending on the time period used for the analysing or commuter’s definition, it could make sense to include more categories such as “no work trip”, “not specified mode”, and commuters without a specified work place into the total number of commute. Therefore, it is worth looking at these categorise in more detail.

- The category "no trip" describes workers living and working in the same place. Accordingly, this person does not need to leave its home for work. These people are often

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Aggregation of the mode of transport of commuting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of transportation of the commuting matrices (names of the original dataset in brackets)</td>
<td>aggregated to</td>
</tr>
<tr>
<td>Pedestrian (<em>walk</em>)</td>
<td>Non-motorized traffic</td>
</tr>
<tr>
<td>Bike and moped (<em>bike</em>)</td>
<td></td>
</tr>
<tr>
<td>Car and car-sharing (<em>car</em>)</td>
<td>Private transport</td>
</tr>
<tr>
<td>Motorbike (<em>mbike</em>)</td>
<td></td>
</tr>
<tr>
<td>Train (<em>railway</em>)</td>
<td></td>
</tr>
<tr>
<td>Train and tramway, bus (<em>railbus</em>)</td>
<td></td>
</tr>
<tr>
<td>Train and car (<em>railcar</em>)</td>
<td></td>
</tr>
<tr>
<td>Train and bike, moped (<em>railbike</em>)</td>
<td></td>
</tr>
<tr>
<td>tramway and bus (<em>bus</em>)</td>
<td>Public transport</td>
</tr>
<tr>
<td>tram, bus, and car (<em>buscar</em>)</td>
<td></td>
</tr>
<tr>
<td>tram, bus, intercity bus und bike, moped (<em>busbike</em>)</td>
<td></td>
</tr>
<tr>
<td>Workbus (<em>workbus</em>)</td>
<td></td>
</tr>
<tr>
<td>Other means of public transport (<em>other</em>)</td>
<td></td>
</tr>
<tr>
<td>No work trip (working at the living place) (<em>no_trip</em>)</td>
<td>deleted</td>
</tr>
<tr>
<td>No specified mode (<em>n_s</em>)</td>
<td>deleted</td>
</tr>
</tbody>
</table>
self-employed e.g. have agricultural activities. These persons should not be included into the volume of commuters, as a commuter is generally defined by its work trip. The average share of workers with no work trip of all people living and working in the same municipality is relatively high (see Table 3: 1970: 28%; 1980: 22%, 1990: 26%, 2000: 22%).

- Commuters with “no specified mode (n_s)” mainly occur with intra-municipality commutes. The shares of intra-municipality commuting trips in the different years are between 3% and 11% (see Table 3). The value 11% is astonishing high in 1980. A miss-categorization possibly happened with the category “no trip” where the value is relatively low or some of the commuters without a specific workplace are included into this category (see next list point). In this work the category "not specified mode (n_s)" is excluded to avoid this uncertainty over the entire period 1970-2000 from the number of intra-municipality commuting.

- For the years 2000 and 1990 workers without any specification of the workplace location are separately indicated. This unknown work location is coded with 0. 100% of these entries are listed in the category “not specified mode (n_s)” in the year 2000. In the year 1990 there are 58% in this category. Using the years 1990 and 2000 it is a useful to add this number of commuters without specified workplace with the same spatial distribution to the current commuting flows. The Federal Statistics Office has already undertaken this correction and the dataset is available under www.pendlerstatistik.admin.ch for the year 1990 and 2000. Everybody should be clearly aware about the version he is working with (see Table 4). For the years 1970 and 1980 person with no workplace location are not reported separately. There are just two vague possibilities what happened with this cases: Firstly, these persons are just deleted or secondly, they are counted as intra-municipal commuter and are most likely taken in the category “no specified mode (n_s)”. Because of consistency 1970-2000, all persons with unknown workplaces are deleted keeping in mind that about 10% of the commuters are lost in all periods.
The historical commuting matrices contain some unrealistically long commuting distances. For most analysis this unrealistic daily long distance commute is unproblematic. For distance based analysis, it is necessary to analyse the change in long-distance or possibly weekly commute. Figure 2 shows that it depends on perception and representation that the relatively small number of long distance commutes can be detected. However, outliers in distance can have an effect on less robust spatial estimation methods. The method tested in this paper is particularly sensitive to outliers with long distances commute. For other analyses this effect should not be underestimated. For example, defining entities below a minimal number of commuters per municipalities can have a much larger effect on distance based methods. It seems worth to think about a correction. How can unrealistic, but not absolutely impossible commuting distances be corrected?

---

As a test the commuting distance based commuting regions size (see chapter 4) has been calculated with and without commutes over 200km first. The mean size of the large urban centres does not differ for the year 1970 and 1980 and is 5% smaller in 1990 and 8% in 2000. In a second test the region is calculated with a threshold value of minimum five commuters per municipalities the difference of the commuting region’s sizes of large urban centres are impressive. The areas are getting 28% smaller in 1970 and the difference is increasing up to 41% in 2000.
The commuting matrices contain generally daily commuters. By definition, weekly commuters are people who have two residences, i.e., the economic residence differs from civil residence (Steinmetz and Pola, 1997). The weekly commute in the years 2000 and 1990 is corrected by the economic place of residence. For the year 1990, 115,000 persons, around 3.8% of all employees are counted as weekly commuters (Steinmetz and Pola, 1997). In the census of the years 1980 and 1970, the question for the second residence was not included into the questionnaire. The weekly commute elimination procedure applied for these years are unknown. In 1980, the dataset
seems to be corrected on the basis of a threshold value “commuting distance” (see Figure 2). In 1970, there are around 0.005% of all commuters traveling more than 200km network distance (in absolute numbers: 104) in 1980 there is one person - 0%, in 1990 absolutely 1641 persons - 0.06%, in 2000 there are absolutely 1693 persons 0.06%. The share of long distance commute is relatively low i.e. is mostly eliminated in advance. A fixed threshold as 200km network distance travelled could be applied for consistency over all years. However, because of the few working hours (about 6 hours), it is theoretically possible to commute daily a distance over 200km.

This section proposes a linear probability function of commuting frequency per network distance to reduce this long distance commute. In a first step the commuting distances are compared with two attributes of the census data 2000 “employment ratio” and “number of days commute per week”. These analyses show that the “employment ratio” of daily commuters per distance does not change significantly. However, the commuting frequency for the full time workers per week is decreasing per distance (see Figure 3). The problem of this attribute is, that it is unknown if somebody is commuting one, two or four ways a day (maybe even returning home for lunch). If somebody stays for one night at his workplace, he has the same commuting frequency as if he returns home every night. This uncertainty is ignored here. However, we assume that values are possibly higher compared to other studies. However, the values seem reasonable because we do not a seperate of weekly and daily commute but a correction by commuting frequency. It is difficult to adapt the function to other national studies and it has no information about the weekly and daily commute. However, the reasonability of the categorization of weekly and daily commute should be questioned because of increasing flexibility and different ways of working.

A linear weighting function of the weekly commuting frequency of full time workers in 2000 is shown in Figure 3 as a corresponding standardized function. A 200km commute trip is assumed on four days instead of five days. That assumes somebody commuting from Bern to St. Gallen (around 2 hours by public transport), working five days, would not return home at one or two days a week. In 2000, the real number of commuters from Bern to St. Gallen is 20 persons 17 use the public transport and three the individual transport. The same function is applied for all year, assuming no change in this behaviour.
Figure 3  Commuting frequency function for long distance correction

Comment: Includes full-time commuters with no second resident location.

### Table 4  Different versions of historical commuting dataset

<table>
<thead>
<tr>
<th>Dataset description</th>
<th>available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commuting matrices 1990-2000 with correction of the commuters without a specific workplace</td>
<td>Swiss Federal Office of Statistics</td>
</tr>
<tr>
<td>Commuting matrices 1970-2000 with a spatial resolution of quarters corresponding to the national traffic model</td>
<td>IVT</td>
</tr>
<tr>
<td>Commuting matrices 1970-2000 with distance and frequency added useful for weekly and daily commuting correction</td>
<td>IVT</td>
</tr>
</tbody>
</table>

### 3.2 Active population and workplaces

The database of Tschopp (2005) can be used for the data on employment and workplaces. However, the data of workplaces does not match in time. The census of enterprises is collected in different years. Working with the total number of workplaces it is more appropriate to use number of workplaces coming from the commuting matrices.
4 Methods of quantifying and qualifying overlapping commuting areas

4.1 Adaption of the confidence ellipse method

Confidence ellipses are an explorative method to investigate the relationship between two variables. They are often used for hypotheses testing and to detect outliers. Confidence ellipses are analogous to the confidence interval of univariate distributions defined as the smallest possible area in which the true value of the population should be found with a certain probability. It is a bivariate analysis applied spatially in several studies (e.g. Turner 1967; Botte 2003, Schönfelder 2006; Buliung and Kanaroglou 2006; Botte and Olaru 2010). The concept is often used to detect activity space size, where the travellers’ home location is taken as a its mathematical centre. That stresses the importance of home for daily life travel and would use a real-world location instead of the artificial mean point of the chosen locations. The calculation of these ellipses for the activity space is regularly tied to the assumption of bivariate normal distributed variables. Botte (2003) adapts this concept for commuting regions or spaces and undertakes two major changes. First, he works with the assumption of a bivariate chi-square distribution with two degrees of freedom, which is more appropriate for the commuting behaviour, because the events are closer to the centre. Secondly, the centre of the ellipse is neither the work nor home location, the arithmetic mean. The approach adapted for commuting regions by Botte (2003) is investigated further in this work for analyzing the spatial effect of commuting behaviour. The confidence ellipse can give more information about the commuting behaviour: The orientation of the main axis of the ellipses or the ratio of the axis. The shift of the arithmetic mean in regard to the home or workplace location can be analysed further.

4.1.1 Necessary adjustment for commuting regions of all municipalities

The analyses of Botte (2003) are producing good results for incommuting regions of the largest cities in Switzerland. The approach should be enhanced in this study to all Suisse municipalities and to the outcommuting direction (see appendix chapter 7). For evaluating the overlaps of functional commuting regions every region should have overlaps. The changes in functional commuting regions size based on the in- and outcommuting distribution of each municipality is tested in different versions of relative and absolute values, with and without commuting of the dataset of 2000 (see Table 4). By doing so, we turn away from the classical
statistical concept of the interval ellipse by making some adjustment which makes the idea application more and acceptable in the planning process.

**Absolute or relative commuting value**

The confidence ellipses are calculated with absolute or relative numbers of commuting frequencies. The calculation of Botte (2003) with absolute numbers of commute produces relatively small variation in size between large urban centre and agrarian municipalities. This effect can be corrected mostly with relative numbers of commute i.e. the commuting intensity. The commuting frequency is related to workplaces and employee according to the incommuting or outcommuting direction. The use of relative number increases the size of the ellipse of the large urban centre immensely (see Table 4). With the relatives a value smoother spatial distribution of the commuting values between the municipality types is obtained. The absolute numbers the long commute in rural municipalities, often goes to larger urban centres, produce relatively large ellipses of rural and agrarian municipalities, however with relative values they disappear. With the application of the method to all municipalities the statistical fundamentals of this interval ellipse are uncertain, because there are different distributions for intensity distributions for every municipality according its location or type. Because of the reasonable changes in sizes of the relative ellipses, we will work with this relative calculation in further analyses.

**With and without local commuting**

A further step includes the intra-municipal commutes into the point dataset of the bivariate distribution. The inclusion of local commuting lets the ellipse contradict, because commuting intensity increases near the arithmetic mean, i.e. near the analysed municipality. The decrease is enormous for rural and agrarian municipalities. For large centres the change is small. Therefore, differences between large urban centre und rural/agrarian municipalities are enlarged. A second positive effect of including the local commutes is that the ellipse has its mean centre closer to the analysed municipality, and it is rather certain that analysed municipality is within the interval ellipse. This documentation presents a closed commuting catchment as all approaches above (chapter 2).
Table 5  Different versions of commuting regions for all Suisse municipalities (2000)

<table>
<thead>
<tr>
<th>Typ of municipalities</th>
<th>out</th>
<th>In</th>
<th>out</th>
<th>In</th>
<th>Mean [km²]</th>
<th>sd</th>
<th>Mean [km²]</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Δ</td>
<td>% Δ</td>
<td>% Δ</td>
<td>% Δ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large urban centre</td>
<td>+399</td>
<td>+125</td>
<td>-4</td>
<td>-1</td>
<td>3197</td>
<td>32</td>
<td>2495</td>
<td>32</td>
</tr>
<tr>
<td>Medium urban centre</td>
<td>+52</td>
<td>+61</td>
<td>-10</td>
<td>-4</td>
<td>1250</td>
<td>22</td>
<td>1112</td>
<td>22</td>
</tr>
<tr>
<td>Small urban centre</td>
<td>+13</td>
<td>+47</td>
<td>-19</td>
<td>-13</td>
<td>848</td>
<td>21</td>
<td>802</td>
<td>19</td>
</tr>
<tr>
<td>Suburban municipalities (first ring)</td>
<td>+14</td>
<td>+54</td>
<td>-23</td>
<td>-15</td>
<td>729</td>
<td>21</td>
<td>1069</td>
<td>30</td>
</tr>
<tr>
<td>Height income municipalities</td>
<td>-6</td>
<td>+32</td>
<td>-39</td>
<td>-29</td>
<td>647</td>
<td>23</td>
<td>787</td>
<td>28</td>
</tr>
<tr>
<td>Suburban municipalities (second ring)</td>
<td>-46</td>
<td>-17</td>
<td>-42</td>
<td>-37</td>
<td>394</td>
<td>18</td>
<td>404</td>
<td>21</td>
</tr>
<tr>
<td>Touristic municipalities</td>
<td>-58</td>
<td>-36</td>
<td>-63</td>
<td>-55</td>
<td>327</td>
<td>18</td>
<td>265</td>
<td>17</td>
</tr>
<tr>
<td>Industrial and tertiary municipalities</td>
<td>-53</td>
<td>-26</td>
<td>-41</td>
<td>-39</td>
<td>376</td>
<td>19</td>
<td>351</td>
<td>18</td>
</tr>
<tr>
<td>Rural municipalities</td>
<td>-69</td>
<td>-49</td>
<td>-45</td>
<td>-46</td>
<td>250</td>
<td>15</td>
<td>212</td>
<td>16</td>
</tr>
<tr>
<td>Semi-agrarian municipalities</td>
<td>-69</td>
<td>-48</td>
<td>-51</td>
<td>-54</td>
<td>246</td>
<td>14</td>
<td>190</td>
<td>14</td>
</tr>
<tr>
<td>Agrarian municipalities</td>
<td>-79</td>
<td>-77</td>
<td>-59</td>
<td>-72</td>
<td>146</td>
<td>15</td>
<td>94</td>
<td>13</td>
</tr>
</tbody>
</table>

4.1.2 The size threshold

Botte (2003) uses the common 95% confidence interval of the ellipse, the area where 95% of the values are. However, the sizes for large urban centres are increasing with this new weighted version based on relative values and even when including intra-municipal commuting. Therefore, this threshold has to be adapted. An exact threshold can not be found. The estimates are based on the in- and outcommuting regions of the year 2000. Only these regions are included which contain at least three municipalities. If the region is too small (one or two municipalities), it can not be considered a region with sufficient functional connection. The
thresholds vary between 95% and 50%. Two criteria of the ellipse’s sizes are considered: The share of commuters within the ellipse (1); the shares of municipalities with any the commuting flow within region (2). Both curves a) and b) in Figure 4 represent the means for all ellipses. The change of threshold value is linear. There is no stair or specific shape of the function which could help to fix the threshold value (see Figure 4).

Figure 4  Mean regions’ size by different confidence interval

No obvious threshold emerges. Therefore, three further considerations lead to the threshold value of 70%:

(1) In Switzerland, two official categorisations of functional regions are available. There are the current official Swiss agglomerations and the metropolitan areas (Schuler et al., 2005). The ellipse approximates more the extent of agglomeration rather than the metropolitan area, because it is linked to a single centre. However, the sizes of the current official Swiss agglomerations is rather limited including no overlapping or fuzzy areas. So, the ellipse for the large urban centre should be larger than the Swiss agglomerations but definitively smaller than the metropolitan areas.

(2) The mean share of both criteria in Figure 4 is over 50% for in- and outcommuting when choosing the 70% confidence interval.

(3) An evaluation with different municipality types shows less variation within the different types the smaller the confidence interval.
4.1.3 Overlaps based on the confidence ellipses

This method is based on the adopted confidence ellipse estimation applying the thresholds explained in the previous section. The results are overlaps of the 70% confidence ellipse of incommuting. It is an overlapping perspective – from a single more to a multi-centre structure. In a first step, the confidence ellipses are calculated for nearly every municipality of Switzerland. Municipalities with less than three municipalities within their commuting region are excluded. The number of overlaps can then be counted for every municipality. Every municipality belongs to a number of commuting regions ($O_i$) (see Figure 5). One municipality belongs to another commuting area if it is inside its ellipsoid. This approach is only looking at the incommuting direction going back to the idea of “overlaps belong to different centres”. The incommuting regions’ overlaps are actually the outcommuting flows of the municipality which the overlaps are calculated.

This measure involves inaccuracy. First, the generalisation of the ellipsoid shape inherent in municipalities within the region, which do not provide a commuting flow to the region’s centre. The number of realised overlaps ($R_i$) is a correction of the total overlaps ($O_i$).

Second, the frequency of commute within the overlapping regions is not considered. It can be assumed that intensity of absolute overlaps is increasing if there are more commuters traveling from the analysed municipality to the centres of the overlapping commuting region. The two components “number of commute to centre of overlapping commuting region” and “number of linked centres within the overlapping areas” are multiplied ($L_i$).
4.2 Social networks methods

In this section we evaluate a method where the dimensionality of commuting flows can be included into the calculations of commuting regions. The commuting flows are translated into a network. A network consists of nodes or vertices and links or connections between the nodes. In this network the nodes are the municipalities of Switzerland and links are the commuting flows in directions. The link’s or edge’s weight is its relative commuting frequency. The relative frequency is calculated as the total flow on the edge divided by the total possible flow on the edge. The possible flow depends on the number of workplaces and employees of either node $i$ and $j$ of the edge. The total possible flow is calculated on the sum of minimum number of workplaces in $i$ or $j$ and the minimum number of employee in $i$ and $j$.

The methods applied in this section come from the field of social network analysis. It describes several algorithms used for community detection. The question of community detection is how to interpret the organization of complex networks as the mixture of structural sub-
units associated with more highly interconnected parts. Most networks typically contain parts in which the nodes are more highly connected to each other than to the rest of the network. The subsets of such nodes are clusters, communities, cohesive groups or modules. There is no unique definition. Generally, these functional connections are first unknown at the beginning. Therefore, different techniques are applied to subdivide the network. The number of groups and their size is not fixed. Looking at overlaps each node of a network can be characterized by a membership function, which is the number of communities that the node belongs to.

Three measures of these overlapping or non-overlapping communities are important (Palla, Derényi, Farkas, and Vicsek 2005):

- The size of any community can be defined as the number of its nodes.
- The number of links within a community can be called its community degree or in-degree and their out-degree, the number of links that connect the cluster with remaining network.
- Communities can share nodes. This number of nodes defines the overlap size between these communities. A share of overlapping non overlapping nodes can be calculated.

In the beginning community detection algorithms were applied to detect non overlapping communities and this technique has been extended to explore overlapping communities. These communities are also called as fuzzy overlapping communities (Gregory, 2010b). The differences are described as follow (Gregory, 2010b): “With non-fuzzy or crisp overlapping, each individual (network vertex) belongs to one or more communities with equal strength: an individual either belongs to a community or it does not. With fuzzy overlapping, each individual may belong to more than one community but the strength of its membership to each community can vary.” In this study the nodes of the network are not individuals but municipalities. A municipality belongs to different communities or commuting regions with different strength (the total of commuters commuting to this commuting region) reflecting more the fuzzy idea.

Different methods to find overlapping communities were discussed e.g. in Fortuna (2010). More and more algorithms were developed recently, allowing overlapping communities. To confirm the algorithm benchmark experiments are often used to analyse the efficiency of algorithm and to learn about its strengths and weaknesses. This work does not discuss the advantage and disadvantage of these algorithms in detail. For application some overlapping algorithm are chosen out of this set for the following properties: Popularity, robustness for large and weighted networks, considering hierarchy structure of the network, and algorithm’s structure should be available and easily applicable. For application it should not be necessary be an
expert in the field. The first algorithm tested in the following is older but rather popularly often used as a reference for testing other algorithms. Three of the newer algorithms are tested subsequently for identifying overlapping commuting clusters. Two of them were successfully tested to the identification of fuzzy overlaps in Gregory (2010b). The last one was developed recently; little experience in application was not be available yet (Lancichinetti et al. 2010). The last one seems particularly interesting because it looks at relative values of connectivity comparing the network against a random network while the other algorithms focus on absolute values.

4.2.1 Percolation method

One of the first algorithms and a rather popular one is the clique percolation algorithm (Palla et al., 2005). The community definition relies on the observation that a typical community consists of several fully connected subgraphs called cliques. Each node is connected with all others nodes in a clique (where the number of connections is k). A community is defined as a k-clique community if each of these k-cliques shares at least k–1 vertices with another k-clique in the community. This method is convenient for undirected and not weighted networks. Weights of links can only be considered by putting a threshold for the inclusion of the linkes. Changing the threshold changes the resolution of community. By increasing the threshold the communities start to shrink and fall apart. A similar effect can be observed for the value of k as well: increasing k makes the communities smaller but more cohesive linked.

The calculation could not be conducted for the whole year 2000 commuting network because of memory problems. Therefore the threshold for the edges of 0.03 was applied for this test to avoid a crash and a reasonable calculation time of the cFinder Software (http://www.cfinder.org/). This algorithm is not especially applicable for large networks. Therefore, the results are only discussed and not in the results section (section 5). Analysing the regions and their overlaps the results of 3-cliques showed that nearly all of the flatland of Switzerland including Jura and Wallis are one region, because of relatively weak connectivity. Only the canton Tessin and some parts of Graubünden are separated into independent commuting regions. 184 separated regions are formed at the higher level of 4-cliques. These regions are really small. With higher k many municipalities are no longer part any community because of the required strong connectivity is missing. In Figure 6 these relatively strong connected functional regions with k=6 are mapped. This algorithm seems not be able to produce reasonable commuting regions. However ideas about the functional behaviour of overlaps can be gained. The highest numbers of overlaps are large urban centres as Zürich, Lausanne, Bern followed by some medium and small centres.

28
4.2.2 The COPRA, LFM and OSLOM algorithm

The Community Overlap Propagation Algorithm (COPRA) were tested first (Gregory, 2010a). This algorithm is applicable for very large and weighted networks. The algorithm is based on the label propagation technique of Raghavan et al. (2007). This algorithm employs an iterative process. Every node gets a unique label first. Then this label is replaced iteratively by the label of the node to which the neighbouring vertices belong most frequently. It is assumed that each node in the network chooses to join the community to which the maximum number of its neighbours belongs. If the highest frequency is equal then the node is allocated randomly. Nodes with the same labels belong to the same community at the end. This algorithm is extended to overlapping communities, so that each node can now belong to at least $v$ communities. $v$ is the only parameter to be set initially. This parameter influences also the hierarchy level, the larger $v$ the smaller is the number of communities. Due to the stochastic procedure in the first iterations, the algorithm runs a number of times and the “best” commu-
Community assignment is chosen according to the highest modularity measure. The available modularity in this algorithm moreover goes back to special modularity measure for overlapping communities (Nicosia et al., 2009).

Second, the LFM algorithm of Lancichinetti et al. (2009) is detecting overlapping community and especially exploring the hierarchical structure of a network. The basic assumption of this algorithm is that communities are local and often nested. A community is identified by the maximization of the fitness of its nodes (see Formula F1).

\[
F_1 = \frac{k_{in}^G}{(k_{in}^G + k_{out}^G)^\alpha}
\]

$k_{in}^G$ and $k_{out}^G$ are the total internal and external degrees of the nodes of module $G$, and $\alpha$ positive real-valued parameter, controlling the size of the communities.

An iterative calculation over all neighbouring nodes of a subgraph chooses the neighbour with the largest fitness. This one is added and produces a larger subgraph. The fitness of each node of in the new subgraph is recalculated. If a node turns out to have negative fitness is removed. The process stops when the nodes examined in the first step all have negative fitness. For the overlaps a procedure of randomly selection of nodes is applied and their possible membership to other communities is explored iteratively. The parameter $\alpha$ changes the resolution of the resulting communities. Large values yield very small communities; small values instead deliver large modules. Often used values for $\alpha$ are between 0.5 and 2. However a modularity measure of the “best” clustering in not implemented in this algorithm.

The third algorithm is the OSLOM (Order Statistics Local Optimization Method) method of Lancichinetti et al. (2010). This method goes back on the principle on the LFM algorithm. The method is capable to detect clusters in large networks with edge directions, edge weights, overlapping communities and hierarchies. Hierarchies are explicitly defined and evaluated as a community containing other subgroups and not just as different community’s sizes. The size of communities is no longer that flexible, there are a specific number of hierarchies based on a lowest one. It is also based on the local optimization of a fitness function. However in contrast to the FLM method this algorithm works with the statistical significance of clusters. The community structure is compared to the probability of finding cluster in random graphs without any community structure. This is the a main improvement to the previous algorithm. Lancichinetti et al. (2010).give a good overview of the calculation technique.
5 Results of overlapping functional commuting region

5.1 Results for the year 2000

In a first step, the two methods “confidence ellipse” and “social networks” and their alternatives are tested with the commuting flows of the year 2000. These calculations use relative commuting values with a distance correction. The first method is based on incommuting intensities, the second method on the share of possible commutes on an edge (possible commuters: All employees living in place A work in place B if enough workplaces are available in B and vice versa). The commuting network is processed differently. The confidence ellipse method works with a directed network and the community detection methods with an undirected network. The confidence ellipse method considers all incommuting flows. The social network method applies a very low threshold possible commute on an edge (0.0000001). At least 0.00001% of possible commutes really undertake this daily trip. The threshold is applied to reduce the enormous size of the network under the assumption that some shares of possible commutes are just too low to be relevant.

We start with the interpretation of the overlapping commuting regions produced by the different variation of the confidence ellipse (see chapter 4.1.4). In Figure 8 the calculation of the total overlap (a) can be interpreted as home locations being connected with a number of potential workplaces. However, the actual home location’s attractiveness can be relatively small, when municipalities have small numbers of inhabitants. These municipalities have a large potential to increase in future, because they have access to other municipalities and possible workplaces within an expedient functional distance. The real (not potential) dependence of the functional region is considered in Figure 6b where the overlaps considers commuting flows only that actually exist. The real overlap intensity (see Figure 8c) tries to express the current situation of including intensity, accounting the size of municipalities.

The results (see Figure 8a-c) show strong overlaps in the flatland of Switzerland. The language border between the French speaking and German speaking part of Switzerland leads to an area of weaker overlaps. The region of Canton Aargau (near city Aarau) is a particularly high potential overlapping region (see Figure 8a). The really linked and the intense overlaps (see Figure 8b,c) are more disperse.

Table 6 lists the top ten municipalities. The list shifts from mainly small municipalities (a) to small or medium centre (b) to large urban centres (c). The top ten of the first two approaches
are mainly situated in the Canton Aarau and while in the last method (c) the overlapping cities are distributed all over Switzerland.

All social network algorithms produce different hierarchies. The LFM and Copra algorithm are relatively flexible setting the hierarchy level by changing the relevant parameter. The OSLOM algorithm has a certain number of hierarchical levels. For the commuting network of Switzerland that are three levels in the year 2000, where the highest hierarchy (minimal number of regions) contains three large regions. A flexible hierarchy setting can be an advantage when commuting regions need to be adapted current perception and political demands. But it is difficult to find a best hierarchy. It is helpful to test number of hierarchies of these algorithms. This study evaluates the regions only due to reasonability and tries to understand what happened with the overlaps when hierarchy and community size is changing. In Figure 7 we fixe the level of hierarchy relatively arbitrarily, so that all algorithms produce a similar number of communities. The algorithm a) and b) have a slightly different behaviour considering their overlaps and community generation. The algorithm c) produces results of commuting regions and their overlaps which have to be interpreted differently.

Table 6  Top 10 municipalities - confidence ellipse’s overlap

<table>
<thead>
<tr>
<th></th>
<th>(a) Total overlaps</th>
<th>(b) Real overlaps</th>
<th>(c) Real overlaps weighted with intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Villmergen</td>
<td>Wohlen (AG)</td>
<td>Zürich</td>
</tr>
<tr>
<td>2</td>
<td>Wohlen (AG)</td>
<td>Lenzburg</td>
<td>Winterthur</td>
</tr>
<tr>
<td>3</td>
<td>Fischbach-Göslikon</td>
<td>Villmergen</td>
<td>Lausanne</td>
</tr>
<tr>
<td>4</td>
<td>Hilfikon</td>
<td>Wettingen</td>
<td>Geneva</td>
</tr>
<tr>
<td>5</td>
<td>Waltenschwil</td>
<td>Baden</td>
<td>Bern</td>
</tr>
<tr>
<td>6</td>
<td>Büttikon</td>
<td>Fislisbach</td>
<td>Basel</td>
</tr>
<tr>
<td>7</td>
<td>Niederwil (AG)</td>
<td>Dottikon</td>
<td>St. Gallen</td>
</tr>
<tr>
<td>8</td>
<td>Dintikon</td>
<td>Dietikon</td>
<td>Dietikon</td>
</tr>
<tr>
<td>9</td>
<td>Remetschwil</td>
<td>Mellingen</td>
<td>Uster</td>
</tr>
<tr>
<td>10</td>
<td>Seengen</td>
<td>Bremgarten (AG)</td>
<td>Luzern</td>
</tr>
</tbody>
</table>
Figure 7 Overlapping commuting area based of two hierarchies - social network method
Figure 8 Three alternatives of confidence ellipse’s absolute overlap

(a) Total overlaps
- 1 - 13
- 14 - 23
- 24 - 35
- 36 - 48
- 49 - 60
- 61 - 74
- 75 - 128

(b) Real overlaps
- 0 - 8
- 9 - 14
- 15 - 21
- 22 - 27
- 28 - 35
- 36 - 47
- 48 - 116

(c) Real overlaps weighted with commuting intensity
- 0 - 1572
- 1573 - 4054
- 4065 - 9486
- 9487 - 20633
- 20634 - 53235
- 53236 - 102583
- 102584 - 12229440
The main differences of commuting regions and their overlaps:

**Regional difference of commuting region’s size:** Are small regions in the mountain, possibly separated by a mountain, really independent even if the absolutely commuter flow is relatively low? The first two algorithms produce more small commuting regions in the mountain area. The last algorithm concludes that these small communities are not significantly different. Therefore, e.g. nearly all of Canton Graubünden is one commuting region. The overlapping areas match the topography.

**Clear and weak boundaries:** All algorithms represent the morphological and the language barriers clearly. The areas of overlaps or uncertainty are generally larger in the first two algorithms.

If we think about the traditional concept of Labor Market enclaves are generally undesirable. The OSLOM algorithm shows that in a functional perspective the regions are not always closed. Especially cities are often enclaves with relatively high number of overlaps. The number is increasing likewise with the number of communities. The boundaries in the flatland of Switzerland are slightly fuzzier in all algorithms whereas the boundaries in the mountains separated are clearer. The overlaps are strongly connected to the chosen hierarchy level and functional region’s size.

Comparing the confidence ellipse and social network methods the overlap by the confidence ellipse procedure described the lowest hierarchy level of functional regions, what is more a commuting linkage than a region. A larger functional commuting region will be produced on a higher functional level, where the commuting linkage is highly overlapping. The Zürich-Area has relatively high overlaps in the ellipsoid method, and a relatively large functional region in the cluster detection algorithms. The mountain region is a no overlapping area in the ellipsoid method and produces rather small commuting regions in the social network methods.

### 5.2 Increasing overlaps 1970-2000

In the period 1970-2000 a significant increase of commuting linkage and overlapping areas of commuting regions is expected. Figure 10 and Figure 11 highlights clearly the increasing overlapping of functional commuting links in the flatland of Switzerland over time. It is reasonable that overlaps start growing near the agglomerations of Switzerland. Particularly, the Greater-Zürich-Area has grown along the main roads. Figure 10 shows that mainly the cities and their adjacent communities produce overlaps in many functional regions because of their large number of work places.
Figure 9  Increasing total overlaps of confidence ellipse over time

Figure 10  Increasing real overlap with confidence ellipse weighted with commuting intensity over time
All social network algorithms show an expected increasing number of commuting when going back in time. However, under the assumption that the chosen parameter over the years are stable and not the number of region (see Figure 11, as an example). The functional regions in the mountain are more stable over the decades than in the flatland. Overlaps are influenced by the changing size of the commuting regions. Large functional regions or communities produce rather large “fuzzy” functional overlapping areas, smaller functional regions rather disperse overlays. The increasing number of functional regions detected by the OSLM algorithm increases the overlaps of mainly the cities in the past.
6 Conclusion and Outlook

The presence of overlapping areas of commuting regions changes according to the chosen algorithm. The confidence ellipse approach and community detection has one main difference. The overlaps of the confidence ellipse can be described as the lowest hierarchy level, where the overlaps are high they will produce a large functional region at a higher level found by the community detection algorithm. The community detection method also allows non-overlapping functional areas. That helps to understand the overlaps. Obviously, they are not that frequent but larger in size when the functional regions are large. However, it is also visible that high degrees of overlap happen more often in the flatland of Switzerland.

In chapter 2 we made two assumptions about the characterisation of overlapping areas of functional commuting regions. Overlaps are municipalities that belong to more than one centre (1) or where the commuting flows are disperse and equally distributed (2). It is difficult to conclude about these two patterns. Overlaps of the first type are found in two social network algorithm. These overlaps are dominated by several centres and are identified in Switzerland at the border of functional regions. These areas can be discussed as uncertainties in allocation. An “area of between” not influenced by a centre produces no overlaps but a rather small functional region of its own or no functional region. This pattern is detected in the mountain region of Switzerland.

Additionally, structural overlapping patterns are large cities associated with their number of overlaps is relatively large not their spatial extension affecting only one community. However, they are not depending on the hierarchy level and the local distance size a city to the next functional region. The cities are particularly the driving power. They still have an special spatial function as enclaves in overlapping functional regions.

What could be the political consequences today of the detected functional regions? This study can conclude that overlapping areas exists and their structure and political impact should be studied in more detail. Overlapping regions are possibly locations of higher infrastructure cost because they provide efficient connections in different directions. This assumption could be true looking at the cities’ overlaps. However, weak functional regions are rather small rather than overlapping. If overlapping regions are attractive work and home location then one should look at their population growth over the recent years.

The findings could be followed up in future work: Two important aspects of overlaps are detected in this study. Firstly, the fuzzy areas are at the functional region’s border. Secondly, the
cities close to a functional region’s border have a high degree of overlap. Unfortunately, there is no algorithm which could detect both structures at once. All algorithms give a good first impression. Still several municipalities are not correctly classified, when looking at them in more detail. However, regarding the speed of improvement and development of social network algorithm, this gap could soon be filled up.

For further analysis it is proposed to use the geographically adjacent and closed morphological structure of the FLM algorithm which is corrected with the overlapping city structure in 20km distance to a functional region’s border (see Figure 12). This correction could also be done by travel time and is particularly relevant for commuters travelling by public transport. Finally, it should be noted, that is difficult to find an algorithm which produces everywhere reasonable functional regions as already stated in Coombes (1986).

Figure 12  FLM algorithm with adjusted cities
7 Appendix: Validation the “monocentric” ellipse

This chapter focuses on a single commuting region produced by the confidence ellipse method described in chapter 4.1. We look at a single region in a monocentric perspective. This method is satisfying by its cartographic quality (see Figure 13). However, what is their quantitative use? Firstly, the quality of different commuting regions is evaluated and then application of the confidence ellipse is tested. Because of the monocentric and not overlapping perspective this chapter 7 is added as an appendix.

Figure 13 Incommuting region of large Swiss cities

7.1 Quality of the functional region

In this section the quality of the functional regions is assessed going back to the assumptions of the closed monocentric economic urban model (Alonso, 1986). This monocentric urban model describes a closed urban system, i.e. every active person is working and living in. It is a monocentric system with one central business district. The central business district (CBD) is the municipality for which the commuting region is calculated. The residential location is determined by the trade of between land rent and transport cost. Higher land rents near the CBD produce an increasing housing and population density in direction of the centre. The
higher density can be explained with a comparative advantage i.e. the agglomeration effect. In other aspects e.g. land use, the surrounding area is homogenous, so that home locations are equally distributed in respect to the distance from the CBD.

The “surrounding area” is founded by all municipalities within the commuting area. The CBD is the municipality for where the ellipse is calculated. The concept corresponds to the incommuting region. However, the measures are adopted for both directions, the incommuting and outcommuting perspective. The outcommuting perspective just changes the direction describing a model where the homes are in the centre and the workplaces are homogenously distributed within the commuting region depending on the distance to the analysed home place municipality.

The following indices are chosen (see formula F2 – F5) to describe the quality of the domination of the commuting regions (see Figure 14 and Figure 15). These indexes give more insight about the internal structure of an ellipsoid commuting region. Quality-measures are used descriptively to assess the commuting region. However, this quality measures could be included in statistical models as weights.

These first indices check how self-contained the region is (see formula F2, F3):

Incommuting region:

\[
II_i = \frac{\sum_{j \in T_i} c_{ji}}{\sum_{j \in T_i} a}
\]

Outcommuting region:

\[
IO_j = \frac{\sum_{i \in T_j} c_{ji}}{\sum_{i \in T_j} w}
\]

\(II_i\) incommuting intensity of the workplace centre (CBD) inside the commuting region (TTWA)

\(IO_j\) outcommuting intensity of the home location inside the commuting region (TTWA)

\(i\) home municipality

\(j\) workplace municipality

\(c_{ji}\) commuting from home \(j\) to workplace \(i\), including intra commuting municipality

\(a\) size of active population

\(w\) number of workplace

\(T\) municipalities within the ellipsoid commuting region (TTWA)
There is a quite similar picture for the in- and outcommuting perspective. Relatively closed commuting regions are in the mountain area (e.g. Poschiavo for both directions). 43% of all incommuters are living within the commuting area, and 53% of all outcommuting are working within its outcommuting area. Poschiavo is a small municipality in an isolated mountain valley with around 3500 inhabitants in 2009, this result is reasonable. You have to cross the Bernina pass to get from Poschiavo to other parts of Switzerland. It takes at least half an hour by car to the next (touristic) centre Pontresina. The train takes much longer. In winter time this road is even closed sometimes. The most urban centres represent a next category of closed urban systems. The closest city in incommuting perspective is Chur with a value of 38%. Geneva is the most self-contained among the large cities in in- and outcommuting perspective. This result is due by the border situation of Geneva. Incommuting and outcommuting of people abroad are not considered in these analyses. However, the size of outcommuting is known and relatively small.

In a monocentric urban model the number of commuters is equally distributed over the municipalities within the commuting region in distance of the CBD. There are two indices looking at this characteristic. The index of relative distance (formula F4 and F5) shows whether relatively short or long distances are commuted within a commuting region. The Gini-index (formula F4 and F5) analyses the equality of distribution of the commuting frequency to the
municipalities within the commuting region. These differences in spatial distribution of size and distance are schematically represented in Figure 15. These two indexes are calculated relative to the number of population correcting the fact that the municipalities are not of the same size. So the share of commuting should be equal according the municipalities’ size.

Figure 15 Distribution possibilities of commuting trips within a commuting region.

(F4) Relative distance of incommuting:  

\[ ID_i = \frac{\sum_{j \in T} (d_{ij} * c_{ij})}{\sum_{j \in T} \sum_{i \in I} a_i} \]

\[ d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \]

(IDi) relative distance of incommuting of the workplace centre

(ODj) relative distance of outcommuting of the home centre

(xi, yi) x, y coordinate of the home municipality

(xj, yj) x, y coordinate of workplace municipality

(i) home municipality

(j) workplace municipality

(cij) commuting from home j to workplace i, including intra commute of analysed municipality

(ai) size of active population

(w) number of workplaces

(T) municipalities within of the ellipsoid commuting region (TTWA)
Figure 16  Relative-distance-index

The relative distance index is looking at the absolute distance commuted proportionally to the total distance commuted under the assumption that all commuters of the analysed municipality are equally distributed over all other municipalities. Only a few municipalities have an absolutely longer commuting distance to the relative equally distributed distances. These municipalities are mainly in mountainous areas. In general, the municipalities have rather shorter commuting distances. Particularly short are the incommuting distances in the Zürich and Basel region. However, this effect is not detected in other agglomerations.

The Gini-coefficient is based on the Lorenz curve, which plots the cumulative proportion of the commuting intensity of all municipalities (y-axis) sorted by intensive (x-axis). The line at 45 degrees represents perfect equality of relative commuting distribution. The Gini-coefficient is the ratio of the area that lies between the line of equality and the Lorenz-curve.

The Gini-coefficient ranges usually zero to one. Zero corresponds to complete equality, while higher Gini-coefficients indicate more unequal distributions. However, in this study the Gini-index is applied inverse. The total equality has the value of 1. It corresponds to the other index, where the values respect to the Alonso model, describing a good quality of a commuting region is high. In contrast to the Herfindhal-index (applied in chapter 7.2) the Gini-coefficient takes into account all municipalities within the commuting region, also the ones with zero percent of commutes.
The results of the inverse Gini-coefficient show an equally distribution for large and medium urban centres of incommuting flows. That corresponds to the theoretical assumption of a monocentric urban model with homogenous surrounding. The result is quite diverse for the outcommuting flows.

Generally, the results of these tests are not surprising: The incommuting direction corresponds quite well to a closed urban monocentric system. However, some separated areas in the Alps also correspond to this theoretical idea except for the relative distance criteria. Additionally, the incommuting direction shows better results than the outcommuting direction.

7.2 The shape and direction of the ellipsoid commuting region

National analyses of the additional parameters such as shape and orientation of the ellipse were undertaken by Botte (2003). The findings of this study are: The orientation of the ellipse is relatively sensitive (1). The orientation of the catchment areas decreases over time in the flatland area of Switzerland (2).
Further we question: Can shape, orientation, and shift in the mean give more information about the commuting linkage on a regional scale? In the regional context e.g. in Zürich-Area commuting ellipses are investigated by the ratio of outcommuters of the suburban communities to its central city. If the criterion of "angle towards the city of Zurich" has a degree of zero, then the ellipse is perfectly oriented towards its centre. No coherent pattern is detected in Figure 18. It is quite disperse. The small centers around the city, such as Winterthur and Zug are more oriented to Zürich in 2000 as in 1970. Moreover, particularly the left side of the shore of Lake Zurich in 1970 was relatively strong facing the city in the 70’s, but no more in 2000. These trends can not be described as significant.

The shape is calculated as the ratio of the longer to the shorter radius, a volume of one indicates a circle. The visualisation of the shape parameter is brighter in 2000 (Figure 18) than 1970. Generally, the ellipses are rounder in 2000 which reflects less interdependence.

The shift of the arithmetic mean of the ellipse to the centre produces the clearest results. This index is the ratio of two distances. The first distance is the arithmetic mean to the nearest centre (here, Zurich). The second one is the distance of the residential municipality to the city of Zurich. In Figure 18 shows, that the eastern side of the city shifts more the centre of Zurich to the west away or towards. The municipalities to the west of Zürich tend to move to the centre of Switzerland. This effect can be explained by the larger attractiveness in the east, where just more workplaces are. However, that shows more a national than a regional pattern.

At the national level more commuter ellipses of a round shape are in the plains than in the mountains, which is reasonable due to the topographical situation. The ellipses appear not to be clearly oriented to the centres, but more the main transport routes. Similar results on national level are presented in Botte (2003). Here, we tested the application at the regional level. The explorative nature of confidence ellipse is useful for analysing large scale patterns on national level. However, at the regional level it looks the appropriate accuracy and a certain arbitrariness of to concept leads to outliers which can not be explained reasonable.

Are other measures measuring the dependence of suburban municipalities on their centres more accurately? To answer this question, the Herfindahl-index was tested to learn more about the decoupling or increasing linkage in commuting patterns. The Herfindahl-index is a concentration index similar to the Gini-coefficient. In contrast, it is absolute to the number of municipalities where somebody is commuting to. For example, if two home municipalities with 50 commuters each are connected to a centre this centre’s Gini-coefficient is zero, because they are equally distributed. However, the values are absolutely high, because there are
just two market shares. Therefore the Herfindahl-index is 0.5. The quality measures we tested in chapter 7.1 explicitly ignore the size of the commuting region, so the Gini-coefficient is applied. Here we are interested in linkage and decoupling of municipalities, therefore the number of participants in the commuting linkage is important, and the Herfindahl-index is tested.

In this example three components influence the Herfindahl-index particularly: (1) local commute, (2) commute to the next centre, and number of connections of commuting flows (3) (so called degree in social network terms). The results show increasing equality nearly everywhere (see Figure 19). The main reason is the increasing size of commuting connections in general. But what are the characteristics influencing this measure in more detail? We do a case study of three municipalities of medium or large cities which are getting more equally distributed communities over time and three municipalities of the second ring in Zürich, where distribution did not occur more equal (see Table 7). The local commute strongly decreased for medium and small centres. This fact could increase the equality (higher the Herfindahl-index). The change of commuting in both categories towards a centre is similarly increasing over the years. The change of the Herfindahl-index highlighted the municipalities with increasing intensity and complexity of linkage. However, a decoupling from an urban centre can neither be clearly identified with this index. Possibly, no decoupling process happened in this area.
Figure 18  Analyses of the additional parameters – angle, shape, and drift

- Angle to Zürich
  - Cities
  - Lakes
  - 0 - 15
  - 16 - 30
  - 31 - 45
  - 46 - 60
  - 61 - 75
  - 76 - 90

- Shape of the region
  - Cities
  - Lakes
  - 1.01 - 1.39
  - 1.40 - 1.65
  - 1.66 - 2.00
  - 2.01 - 2.60
  - 2.61 - 149.30

- Drift of the region's mean
  - Cities
  - Lakes
  - 0 - 90%
  - 91 - 95%
  - 96 - 100%
  - 101 - 105%
  - 106 - 467%
Figure 19  Identifying commuting linkage with Herfindahl-index

![Maps showing commuting regions with Herfindahl-index](image)

Table 7  Case study: Change of commuter flows 1970-2000

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Embrach</td>
<td>+59.68 %</td>
<td>-9.30 %</td>
<td>+5.82 %</td>
<td>+66.88 %</td>
</tr>
<tr>
<td>Buchs</td>
<td>+71.83 %</td>
<td>-23.31 %</td>
<td>+7.55 %</td>
<td>+76.52 %</td>
</tr>
<tr>
<td>Herrliberg</td>
<td>+39.19 %</td>
<td>-11.78 %</td>
<td>+0.20 %</td>
<td>+57.41 %</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Centre around Zürich</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Winterthur</td>
<td>+6.90 %</td>
<td>-28.57 %</td>
<td>+9.54 %</td>
<td>+57.97 %</td>
</tr>
<tr>
<td>Aarau</td>
<td>-3.49 %</td>
<td>-29.98 %</td>
<td>+7.43 %</td>
<td>+57.05 %</td>
</tr>
<tr>
<td>Baden</td>
<td>+19.42 %</td>
<td>-33.40 %</td>
<td>+9.07 %</td>
<td>+62.68 %</td>
</tr>
</tbody>
</table>

7.3 Change of commuting region’s size over time

From 1980 to 1990 the largest increase in region’s size over the years is visible (see Figure 20). This increase of municipalities’ size can be caused by the change in data collection, ex-
plained in chapter 3.1. That outcommuting regions of large urban centre in 1980 are smaller than in 1970 could be also an effect of data correction of long distance commute in 1980 (explained in detail chapter 3).

A restructuring of the suburban municipalities (first ring) can be detected after 1980. The size of incommuting regions is growing faster than the outcommuting regions achieving similar mean size as medium urban centres. The outcommuting regions for large, medium and small urban centres are larger than their incommuting regions in 1980 and 1990. Same tendencies of concentrated incommuting and disperse outcommuting are seen in other municipality types such as touristic, industrial or tertiary municipalities, which are characterised by lots of workplaces. For both types of suburban municipalities characterised as more residential amount, and rural municipalities, the outcommuting region is smaller than their incommuting. For the year 1970 and 1980 this difference between in- and outcommuting regions are ambiguous.

Figure 20  Change of commuting region mean size of municipality type 1970-2000
7.4 Conclusion

The ability of confidence ellipse, as applied in Botte (2003) and Killer und Axhausen (2010a), for visualising the growth of catchment areas of medium and large urban centres remains impressive. The simple shape is a useful generalisation to trace the change over different years, keeping the information visible and procuring a well-designed attractive map.

In this study the method is tested qualitatively not only for the incommuting region (catchment areas) but also for the outcommuting direction. However, the incoming flows are spatially more concentrated near the analysed central municipality and long commuting trips are less visible (e.g. outliers). Therefore, the incommuting region (catchment area) is generally more stable than the outcommuting region.

The method is generally sensitive to data issues. Historical changes in data collection can have a large impact on size and shape of the commuting region. The assumption of a chi-square distribution of the commuter’s flows will not be true for every (especially small) municipality. Therefore, the quality of these methods differs in time and space.

The application on regional level is not satisfactory. Differences between neighbouring municipalities can not be explained adequately. The explorative method of confidence ellipses is useful for analysing large scale pattern on national level. However, it misses the appropriate accuracy for the regional level.

The large mean sizes of incommuting and outcommuting commuting region of the large urban centre is realistic and the size gradient between the municipality types is plausible, but not surprising. It is the first known study, assuming that functional regions exist for all municipalities not only for urban centres and the regions can easily be calculated for around 3000 municipalities. Functional regions in the mountain areas exist. Particularly, even where the areas are relatively small and where no near by centre is existing. A great advantage of this method is its short computing time. The calculation is done with an R script (The R Project for Statistical Computing). This script is relatively simply adaptable and flexibly expandable. A possible extension of this concept was shown in the previous chapter where the perspective of the monocentric model changes to a polycentric one. Taking into account the overlaps of the ellipses and thinking about that a municipality belongs not exclusively to one but to several functional regions.
8 Acknowledgement

We would like to thank the Swiss National Science Foundation (SNF) and the German Research Foundation (DFG) for funding this work as part of the project “Spatial accessibility and the dynamics of commuting in Germany and Switzerland 1970-2005”. This project is jointly carried out at the Institute for Transport Planning and Systems at ETH Zurich and the Department of Transport Planning at TU Dortmund University.
9 References


