Working Paper 2.6

Descriptive and Geographical Data for European Cities

Alain PHOLO BALA

Université Catholique de Louvain, Belgium

Revision: 1

FP7-244557

24/06/2010

EUROPEAN COMMISSION
European Research Area
SEVENTH FRAMEWORK PROGRAMME
Funded under Socio-economic Sciences & Humanities

Contents

Descriptive and Geographical Data for European Cities

Alain PHOLO BALA
Université Catholique de
Louvain, Location
Voie du Roman Pays, 34
1348 Louvain-la-Neuve

Teleph.:+32 10 47 4352
Telefax: +32 10 47 4301
Alain.Pholobala@uclouvain.be

24/06/2010

Abstract

UrbanSim is a software-based simulation system designed to support planning and analysis of urban development. Applied to several US metropolitan areas, it has proved to be a very powerful and efficient tool in this context. Nevertheless, recent UrbanSim applications in Europe outline the necessity to adapt the available modelling platform and to develop new tools consistent with the characteristics of European cities. The project Sustaincity (www.sustaincity.eu) will develop such tools, based on the UrbanSim modelling platform.

European cities display strong contrasts with respect to US agglomerations. Firstly, they are denser and more compact than Anglo-american cities. Secondly, they are concentric, while in US agglomerations have developed according to grid plans. Finally, they show up a different social geography. The latest UrbanSim version (version 4.2.2) enables parcel based applications that are more suitable for European cities. Those parcel based applications have data re-

Keywords

UrbanSim; European cities; US agglomerations, Parcels, Grid Cells, Data

Preferred citation style

Pholo Bala, A. (2010) Descriptive and Geographical Data for European Cities, *SustainCity Working Paper*, **2.6**, Université Catholique de Louvain, Belgium.

1 Introduction

UrbanSim is a software-based simulation system designed to support planning and analysis of urban development. It accounts for the interactions between land use, transportation, the economy, and the environment. It primarily intended for use by Metropolitan Planning Organizations (MPOs), cities, counties, non-governmental organizations, researchers and students interested in exploring the effects of infrastructure and policy choices on community outcomes such as motorized and non-motorized accessibility, housing affordability, greenhouse gas emissions, and the protection of open space and environmentally sensitive habitats.

Metropolitan areas have come under intense pressure to respond to mandates to link planning of land use, transportation, and environmental quality; and from citizen concerns about managing the side effects of growth such as sprawl, congestion, housing affordability, and loss of open space. The planning models used by Metropolitan Planning Organizations were generally not designed to address these questions, creating a gap in the ability of planners to systematically assess these issues. UrbanSim has been developed to address these emerging requirements (Waddell, 2002). It has been applied to several US metropolitan areas and has proved to be a very powerful and efficient tool in this context (Waddell *et al*, 2007, Waddell and Borning, 2004). However, recent UrbanSim applications in Europe outline the necessity to adapt the available modelling platform and to develop new tools consistent with the historical, political and social characteristics of European cities.

In this report, we identify the differences between the geographical features of US and European cities and we further show up the distinctions between the descriptive and geographical data required to properly model US and European agglomerations.

In the next section, we describe the geographical differences between US and European agglomerations. Then, in a third section we identify the required and available data to model European cities.

2 Geographical Differences Between US and European Cities

Uncovering distinctive features between US and European cities may lead to understate differences between European cities. European cities differ by their administrative meshing and their urban design. Moreover, there are significant differences across Europe in terms of building and population densities. Those differences may have been triggered by the fact that Europe has experienced different waves of urbanization at different stages of history and techniques evolution. However, careful inspection of existing definitions and statistics as well as observations on European urban agglomerations seem to outline a number of points that render European cities distinct from other cities of the World, especially North American cities (Le Gléau *et al.*, 1996; Kaplan *et al.*, 2004). European cities are denser and more compact than North American agglomerations. Moreover, they are concentric while Anglo-American cities feature a grid plan. Finally, European and North American cities have contrasting social geography.

2.1 Density and Compactness

European cities are more compact than their North American counterparts. For a comparable population they occupy less total area than American cities. Most of their residents are apartment dwellers. Residential streets of old sections tend to be narrow, and front, side, or rear yards or gardens are scarce (Fellman *et al.*, 2003).

European cities were initially designed for pedestrians and they have the dimension appropriate with walking distances; most of them initially established as "**walking cities**" (Kaplan *et al.*, 2004; Allain, 2004). Those "**walking cities**" currently correspond to the central parts of European agglomerations characterized by the highest density of built-up areas (Allain, 2004).

Despite that the evolution of transportation technology allowed to achieve tremendous urban growth, it has not jeopardized this design. Rather, it seems to have ensured its persistence. Compactness, high densities have favored the development of public transportation, including well developed subway systems which allowed cities centers to keep their prominence.

Indeed, the expansion of automobile ownership and the construction of highways are not as pervasive as in US. Therefore, it has not eased the development of extensive periurban areas and European cities mark a strong contrast with the sprawl of American Suburban Areas. Indeed, residence and work are closer in Europe, often within walking and bicycling distance. Most sections of cities have first floor retail and business establishments below upper-story apartments, bringing shopping and employment places within convenient distance of residences (Fellman *et al.*, 2003).

Being more compact, European cities have a higher urban density as Table 1 clearly shows.

Table 1 Urban Residential Density per Acre

City	Country	Density per Acre
Bucharest	Romania	669
Tiranë	Albania	489
Sofia	Bulgaria	403
Amsterdam	Netherlands	375
Budapest	Hungary	304
Prague	Czech Republic	274
Warsaw	Poland	274
Paris	France	269
Chişinău	Moldova	259
Ljubljana	Slovenia	252
Marseille	France	242
Belgrade	Serbia	237
Athens	Greece	235
Toronto	Canada	227
Talinn	Estonia	180
New York	U.S.	178
Bratislava	Slovakia	175
Cardiff	U.K.	133
Stockholm	Sweden	59
Seattle	U.S.	47
Atlanta	U.S.	22

Source: World Resources 1998-99: *A Guide to the Global Environment*. New York: Oxford University Press quoted by Kaplan *et al.* (2004)

Three basic factors may explain this compactness (Kaplan *et al.*, 2004):

- The costs of private transportation are much higher in Europe than in United States. Therefore, European city dwellers cannot spread out as far as American urbanites. As Ta-

ble 2 shows gasoline prices are very high comparatively to United States. Furthermore, because costs and congestion make long commute less advantageous, European countries are less inclined to build new roads or even to expand the carrying capacity of existing roads. European urbanites rely mostly on mass transportation (table 3) and that is far more extensive and cheaper than United States thanks to more important government subsidies;

Table 2 Average Gas Prices

Country	Price per Gallon
U.K.	$5.23
Finland	$5.03
France	$4.58
Belgium	$4.48
Germany	$4.21
Austria	$3.87
Ireland	$3.84
Switzerland	$3.48
Spain	$3.36
U.S.	$1.74

Source: Akron Beacon Journal, September 8, 2000 quoted by Kaplan *et al.* (2004)

- The second factor explaining compactness of urban cities is the higher costs of home ownership and home financing. Housing prices are generally higher in Europe as a result of greater rarity and because European builders use sturdier materials. In Europe the cost of financing a house is higher since most Europeans do not enjoy the 30-year, tax deductible, and amortized mortgage available to Americans. Higher costs of home ownership and financing have led to lower homeownership rates in Europe;
- The third factor explaining high densities in European cities is that planning is much more rigorous in Europe. European states benefit from a variety of growth control mechanisms intended to contain urban population within existing built-up areas. Across Europe, policy makers have sought to concentrate urban growth in existing centers or in a selected number of growing cities.

Table 3 Mass Transit and Car Ownership

City	Country	Work Trips by Transit	Cars per 1000 People
Belgrade	Serbia	64	30
Chişinău	Moldova	48	60
Nizhniy Novgorod	Russia	78	69
Riga	Latvia	57	104
Moscow	Russia	85	138
Zagreb	Croatia	52	215
Vilnius	Lithuania	49	215
Copenhagen	Denmark	27	223
New York	U.S.	51	232
Bratislava	Slovakia	72	282
Budapest	Hungary	66	288
Sofia	Bulgaria	75	310
Cardiff	U.K.	13	350
Athens	Greece	34	354
Stockholm	Sweden	37	390
Liepzig	Germany	33	396
Paris	France	40	426
Toronto	Canada	30	430
Atlanta	U.S.	20	473
Prague	Czech Republic	67	500
Seattle	U.S.	16	654

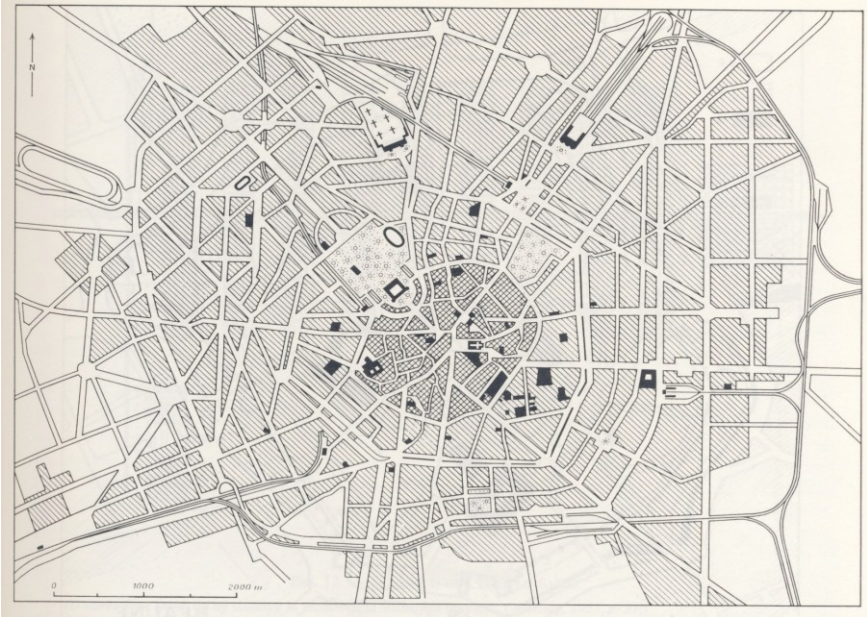
Source: World Resources 1998-99: *A Guide to the Global Environment*. New York: Oxford University Press quoted by Kaplan *et al.* (2004)

2.2 Concentric cities

Most of European cities are characterized by their concentric development (Le Gléau *et al.*, 1996). Indeed, the typical urban entity in Europe includes an urban core, old and very dense. This urban core is surrounded by a concentric outer area showing up continuity of built-up area. In turn this outer area is surrounded by far less dense peripheral areas.

Urban centers of European cities are generally of great historical and cultural value. It hosts historic buildings and are therefore often the preferred attraction for tourists.[1] While the urban core lies often in the oldest section of the city, it still functions as the nerve center of even its most modern parts. Hosting the administrative, the financial (major banks and brokerage houses), and the most important retail functions, European cores act as the central place that seems to have vanished in North American agglomerations. Figure 1 shows the example of a concentric urban development illustrated by the plan of the city of Milan.

Figure 1 An example of concentric plan: Milan

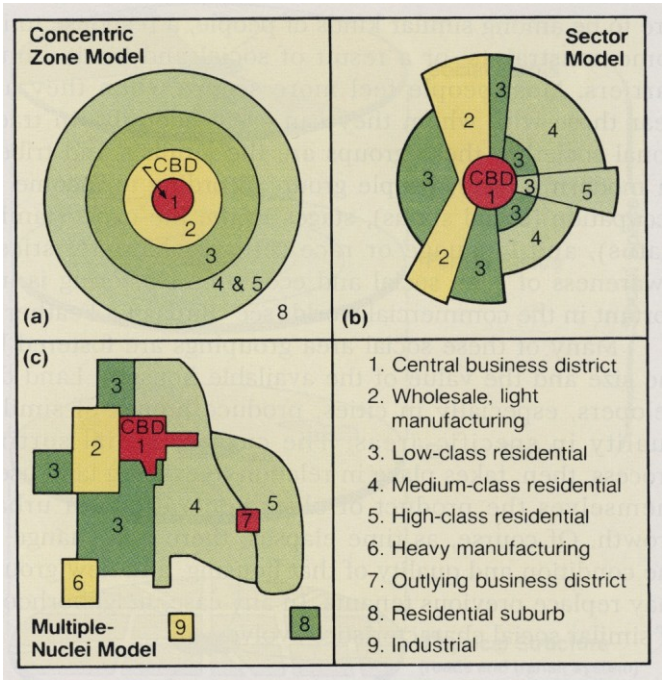


Source: Beaujeu-Garnier (1995), p. 67.

While the concentric model (Figure 2(a)) is still somewhat pertinent the describe development of European cities, it has become less relevant to characterize metropolitan areas in US.

[1] The epicenter of most of European cities lies in the oldest section of the city. This is termed the historical core and is often surrounded by the ancient or Medieval walls (Kaplan et. al, 2004). It hosts several cultural amenities like ancient churches, townhalls etc.

Figure 2 Three classic models of the internal structure of cities.[2]



Source: Fellman *et al.* (2003), p. 419

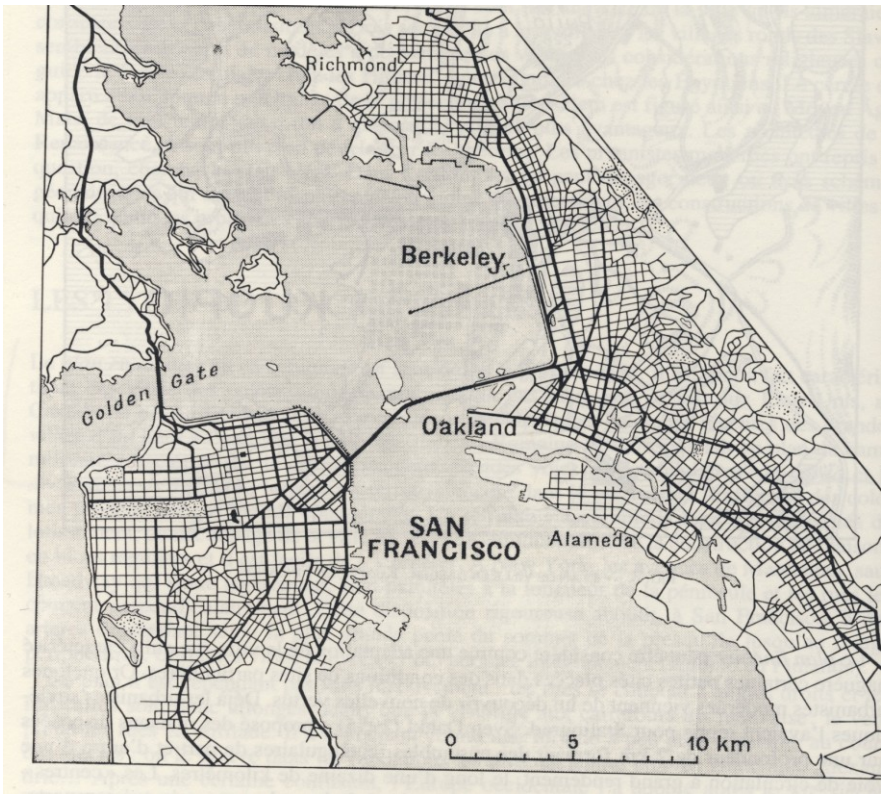
Deprived of a long urban tradition, contrary to Europe, United States experienced the emergence of new metropolitan land use and functional patterns that could no longer be satisfactorily explained by the classic ring, sector, or multiple-nuclei models. Urban development in US is characterized by increasing sprawl. No longer dependent on the central city, the suburbs were reborn as vast, collectively self-sufficient outer cities. Indeed, many, perhaps most, suburbanites have no connection with the core city, feels no ties to it and satisfy almost all their needs within the peripheral zone. New suburbs began to outperform older central districts in size and even as generators of employment and income. So, they have outgrown their former role as bedroom communities and have emerged as chain of

[2] The concentric zone model considers a metropolitan area as a set of nest rings. It acknowledges four concentric circles of mostly residential diversity at increasing distances in all directions from the wholesaling, warehousing, and light industry border of the high-density CBD core:

- A zone in transition marked by the deterioration of old residential structures abandoned, as the city expanded, by the former wealthier occupants and now containing high-density, low-income slums, rooming houses and probably ethnic ghettos;
- A zone of "independent working people's homes" occupied by industrial workers, perhaps second-generation Americans able to afford modest but more ancient homes on small lots;
- A zone of better residences, single family homes, or high-rent apartments occupied by those wealthy enough to exercise choice in housing location and to afford the longer, more costly journey to CBD employment;
- A commuters' zone of low-density, isolated residential suburbs, just beginning to emerge when this model was proposed (Fellman *et al.*, 2003).

independent, multinucleated urban developments. Qualified as "edge cities",[3] those outer cities now exist in all regions of urbanized Anglo America. The declining influence of Central Business District in US metropolitan areas is consistent with the dominance of grid plans in US urban agglomerations. In such plans, streets run at right angles to each other, forming a grid. With grid plans the urban fabric is neutral without any hierarchy and congruence to site topography. Figure 3 shows the example of the grid plan of the city of San Francisco.

Figure 3 San Francisco: A grid plan at the West of the bay



Source: Beaujeu-Garnier (1995), p. 66.

[3] Garreau established five rules for a place to be considered an edge city:

- It must have more than five million square feet (465,000 m²) of office space. This is enough to house between 20,000 and 50,000 office workers, as many as some traditional downtowns.
- It must have more than 600,000 square feet (56,000 m²) of retail space, the size of a medium shopping mall. This ensures that the edge city is a center of recreation and commerce as well as office work.
- It must be characterized by more jobs than bedrooms.
- It must be perceived by the population as one place.
- It must have had no urban characteristics 30 years earlier.

http://en.wikipedia.org/wiki/Edge_city visited the 4/05/2010.

2.3 Social geography

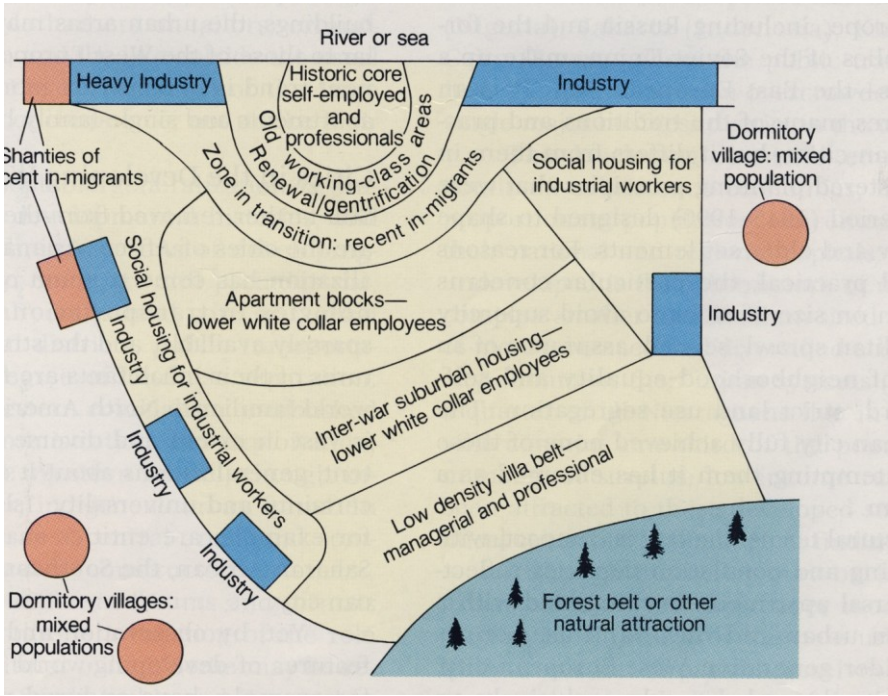
Europe shares with United States the common feature that older structures are located at the center of the city. However, unlike United States where the richest people prefer newer, more spacious peripheral housing, in Europe wealthy people are likely to live and stay in the historical core, while many of the poor are found in the city's outskirts[4]. This trend is reinforced by the greater tradition of providing public housing. In most European countries provision of such housing is far more comprehensive than in United States. Much of this housing is located near the city's edge, where land is available. Paris, for instance, is surrounded by a series of lower status apartment complexes where are located many of the city's poor as well as several immigrant populations. Therefore, rather than facing inner-city blight like most of Anglo-American agglomerations, Paris suffers from a crisis of suburbs in which public housing is linked in the public mind with social deprivation and minority overrepresentation (Kaplan *et al.*, 2004; Paulet, 2000).[5]

Figure 4 presents a schematic of a Western European city. While it cannot account for all the differences among Western European cities, it illustrates some of the features just discussed, in particular the persisting attractiveness of its historical core. The urban core includes a large number of affluent people. Social housing is found along the edge of the city, close to industrial areas. Suburbanization of some of the more affluent people is also evidenced.

[4] Brueckner *et al.* (1999), provide an appealing explanation of this contrast by proposing an amenity-based theory of location by income. They explain the relative location of different income groups by the spatial pattern of amenities in a city. When the center has a strong amenity advantage over the suburbs, as in European cities, the rich are likely to live at central locations. When the center's amenity advantage is weak or negative (the case of US agglomerations), the rich are likely to live in the suburbs.

[5] Another major point of divergence generally outlined between Europe and North America is that, in Europe, ethnic segregation did not emerge as a major dimension of urban morphology. This conclusion refers to studies that were done several decades ago. However, in the meantime cultural diversity and ethnic segregation have become more prominent (Kaplan, 2004).

Figure 4 A diagrammatic representation of the West European city.

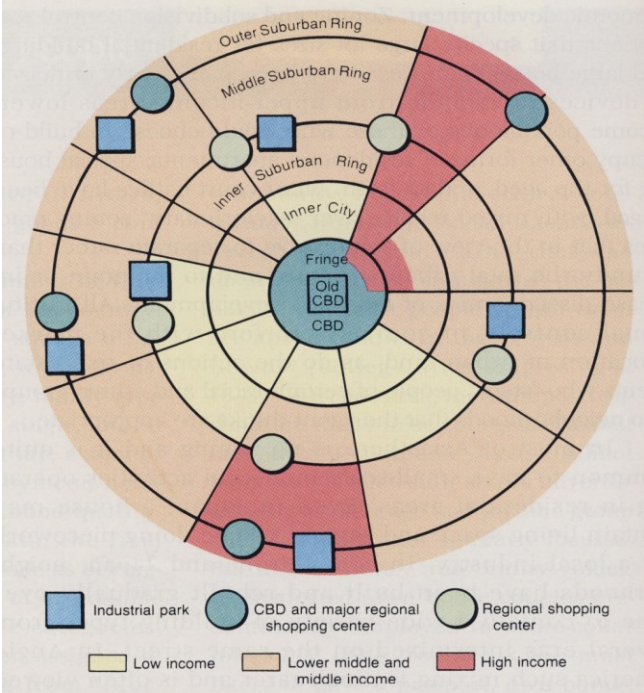


Source: Fellman (2003), p. 433.

As Figure 5 shows, things are rather different in US metropolitan areas. Sectors of high-income residential use expand beyond the central city limits, usurping the most desirable suburban areas and segregating them by price and zoning restrictions. Middle, lower-middle, and lower income groups found their own income-segregated portions of the fringe. Ethnic minorities are frequently relegated to the inner city and to some older industrial suburbs.

To summarize, European cities display strong contrasts with respect to US agglomerations. Firstly, they are denser and more compact than Anglo-american cities. Secondly, they are concentric, while in US agglomerations have developed according to grid plans. Finally, they exhibit contrasting social geography: while in European cities rich people seem to be attracted by the strong amenity advantage of historic urban cores, US agglomerations experience inner city blight and the relegation of low income group and ethnic minorities in the CBD. Such differences may entail contrasting definitions of urban agglomeration in Europe and in US

Figure 5 San Francisco: A grid plan at the West of the bay



Source: Fellman (2003), p. 424.

3 Limits of Urban agglomerations

The delineation of agglomerations is an old issue. Since the 19th century statisticians, economists, geographers and civil servants have realized the importance of space delineation. Several kinds of agglomeration definition have been provided. Among them, we may distinguish firstly analytical and functional definitions that consider agglomerations as an area to analyze regarding a specific function of interest. Secondly, there are morphological definitions based essentially on the continuity of built-up area criterion. Thirdly, there are political and administrative delineation of agglomerations that reduce the scope of agglomerations to fixed boundaries. The last definition of urban agglomerations has raised concerns. Indeed, it has been argued that administrative borders do not capture the essence of economic phenomena that often spill over boundaries. The shortcomings of the administrative delimitation of agglomeration have been the main incentive to the creation of geographical units based either on morphological or functional criteria.

As urbanization is a very complex phenomenon, agglomeration definition and delineation may differ with respect to the goal pursued: transport management, infrastructure design, urbanism specifications, tax receipts collection and allocation (Dujardin *et al.*, 2007). Therefore, there is not any best or optimal criterion. The choice of the proper criterion depends on the objective.

The goal of the Sustain-city project is to advance the state-of-the-art in the field of micro-simulation of prospective integrated models of Land-Use and Transport (LUTI). For such a project a consistent agglomeration definition should be a functional region capturing its economic influence.

The “Core Based Statistical Area” nomenclature proposed by the US Office of Management and Budget is clearly a functional definition. It is consistent with US multinucleated urban developments marked the expansion of suburbs and edge cities and the regression of former CBD.

In Europe, where historical urban cores are still dominant, nomenclatures have put more emphasis on morphological criteria.

3.1 USA

The Office of Management and Budget has published an updated nomenclature of US metropolitan areas in December 2000 (Federal Register, 2000). This nomenclature is based on the concept of **Core Based Statistical Area** (**CBSA**). A **CBSA** is a statistical geographic entity consisting of counties associated with at least one **Core** (**Central counties**), plus adjacent counties having a high degree of social and economic integration with the **Core** (**Outlying counties**). Counties are qualified as central if they have at least 50 percent of their population in urban areas of at least 10,000 population, or if they have within their boundaries a population of at least 5,000 located in a single urban area of at least 10,000 population. The degree of interaction between central and outlying counties is measured through commuting ties. Indeed, an outlying county meets the following requirements: at least 25 percent of the employed residents of the county work in the central counties of the CBSA or at least 25 percent of the employment in the county is accounted for by workers who reside in the central counties of the CBSA.

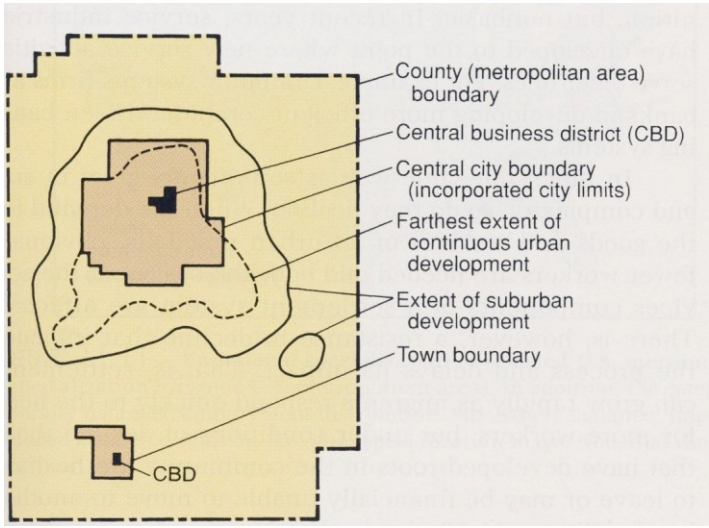
A **Core** is a densely settled concentration of population, comprising either an **urbanized area** (of 50,000 or more population) or an **urban cluster** (of 10,000 to 49,999 population).

There are two categories of **CBSA**: **Metropolitan and Micropolitan Statistical Areas**. They differ by the size of their **Core**. The **Core** of a **Metropolitan Statistical Area** is an **urbanized area** populated by at least 50,000 inhabitants, whereas for a **Micropolitan Statistical Area** it is an urban cluster having a population of at least 10,000 but less than 50,000.

The **urbanized area** designates a continuously built-up landscape defined by building and population densities with no reference to political boundaries. It may be considered as the **morphological agglomeration** (or the **physical city** according to Fellman *et al.*, 2003) and may include a central city and many contiguous cities, towns, suburbs, and other urban tracts.

On the other hand, a **Metropolitan Statistical Area** refers to a large scale functional entity, perhaps including many urbanized areas, discontinuously built-up but nonetheless operating as well integrated economic whole. Figure 6 shows these areas in a hypothetical American county.

Figure 6 A hypothetical spatial arrangement of urban units within a metropolitan area.



Source: Fellman *et al.* (2003), p. 405.

3.2 France

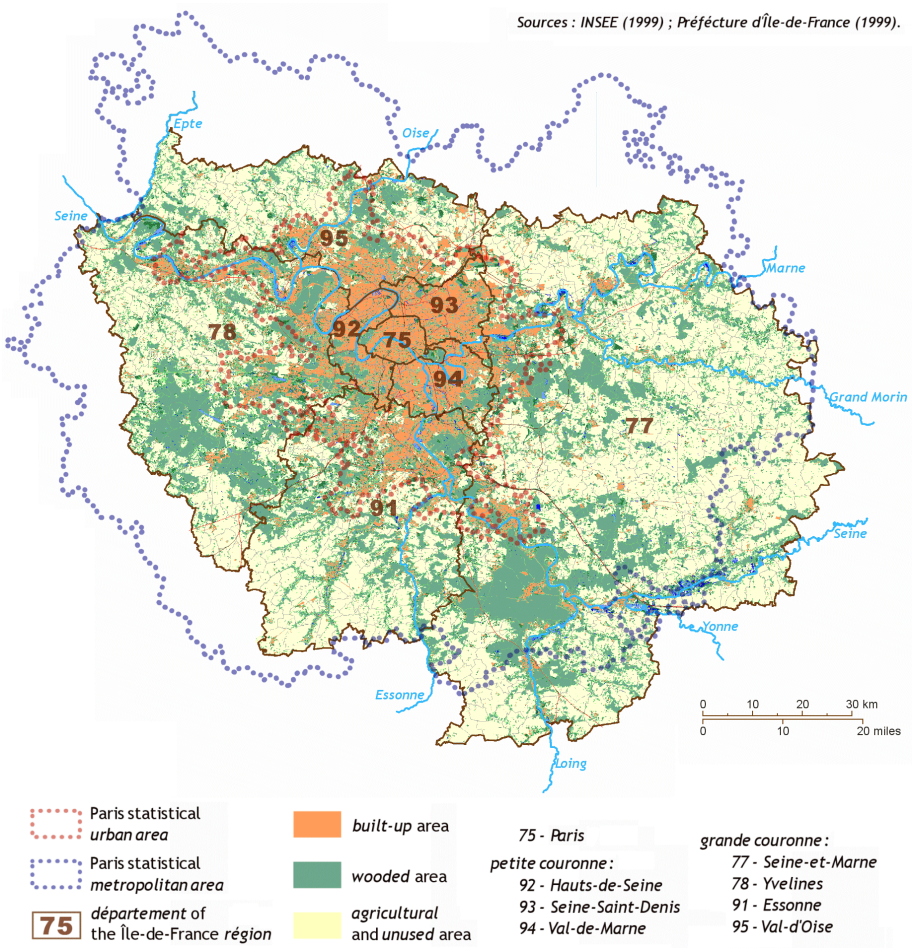
A functional definition has also been proposed by the French National Statistics Institute (INSEE). In 1997, INSEE has defined a new spatial nomenclature of French territory, the "Urban Areas" zoning in order to capture cities' influence beyond administrative borders. Regularly updated during the subsequent census, this zoning is based on two nested spatial entities: the urban unit and the urban area.

Urban units are the basic spatial units. It is defined as a set of administrative districts whose territory hosts a built area with at least 2000 inhabitants. Within an urban unit, the maximal distance allowed between buildings is 200 meters. Whenever an urban unit reaches the threshold of 5000 jobs, it is qualified as urban pole if it does not depend of a bigger urban pole i.e. if less than 40 % of the workers residing in that urban unit work in that huger urban pole.

An urban area is a set of contiguous administrative districts without any enclave. It is composed by an urban pole, and a periurban ring constituted by rural administrative districts or urban units where at least 40% of the dwellers work in the urban pole or in the administrative districts closely connected to it. The INSEE 2009 "Urban Areas" nomenclature counts 354 urban areas in Mainland France.

INSEE has proposed another interesting nomenclature, the "Employment Area" zoning. The purpose of this zoning is to analyze the working of the French labor market and commuting. An Employment Area is a geographical space where most of the workers reside and work[6]. Therefore, the characteristic of "Employment Area" delineation is clearly to minimize daily cross-boundary commuting, or equivalently to maximize the coincidence between residential and working areas (Briant et al., 2009). According to the INSEE 2008 zoning, there were 348 Mainland Employment Area.

Figure 7 Departments inside Paris urban area



[6] http://www.insee.fr/fr/methodes/default.asp?page=definitions/zone-emploi.htm visited the 9/03/2010.

3.3 Switzerland

The Swiss Federal Statistical Office defined in 1980 a new formula of agglomeration delineation putting more emphasis on functional criteria than on morphologic ones. According to this new nomenclature, the **urban area** is composed of, on one hand, agglomerations, and, on the other hand, **isolated cities**.

An **agglomeration** is defined as a set of adjacent administrative districts having at least 20,000 inhabitants. It is composed of a **central area**, and of **outlying administrative districts** fulfilling the following criteria:

- At least a sixth of the workers residing in those administrative districts work in the central area;
- at least 3 of the 5 following criteria: continuity of the built-up area, a combined inhabitants/workers density higher than 10 by hectare, population growth rate exceeding the national average by more than 10% during the last ten years, not less than a third of the workers residing in those administrative districts work in the central area; the share of dwellers working in the primary sector should not exceed twice the national average.

The central area of an agglomeration includes a **core administrative district** and possibly some other administrative districts that fulfill the following conditions: they have at least 2,000 jobs, the ratio of the number of persons working within their boundaries with the number of workers actually residing in those administrative districts must be greater or equal to 0.85 and the following condition: they should form a continuity of built-up area with the core administrative district or share a border with it, or have at least a sixth the workers residing in those administrative districts working in the core administrative district.

An **isolated city** is defined as an administrative district with at least 10,000 inhabitants.

3.4 Belgium

In Belgium, the "Direction générale Statistique et Information économique"[7] has published in 2009 a spatial nomenclature of Belgian urban regions based on the data of the Social and Economic Survey of 2001 (Van Hecke *et al.*, 2009). This nomenclature is merely an update of the definition of Belgian urban regions published by the Belgian National Statistics Institute in 1979. This definition has been previously updated twice; respectively on the basis of 1981 and 1991 data (Van der Haegen *et al.*, 1996). This nomenclature defines several spatial units strongly nested and uses both morphological and functional criteria.

According to this nomenclature the **urban core** is the center of decisions and of activities that displays the highest concentration of retail trade and services. It is surrounded by densely built urban districts and forms with those surroundings the **city-centre**.

An **urban ring** composed essentially of XXth century buildings encircles the **city-centre**. It is a continuous built-up area that is less dense than the city center. The **urban ring** and the **city-centre** form altogether the **morphological agglomeration** whose boundaries are determined by the continuity of built-up area rather than by the administrative borders. By adjusting the morphological agglomeration to administrative boundaries, we get the **agglomeration** (or **operational agglomeration**).

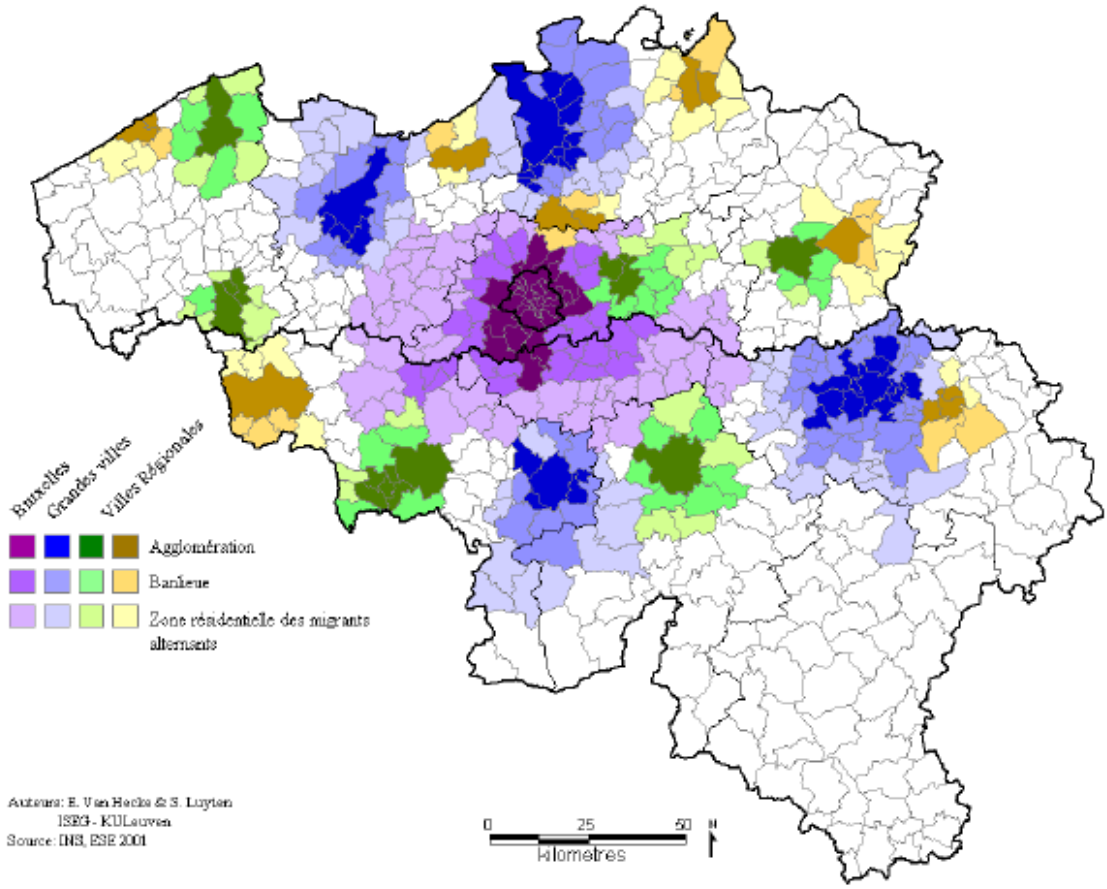
The **suburb** is the outer area beyond the **agglomeration**. Morphologically, it may appear as a rural area, but from a functional view point, it is urban.

The agglomeration and the suburb constitute the **urban region**. It is the enlarged entity where most of the activities undertaken by the urban community take place. Therefore the urban region describes a functional area turned towards the city-centre.

Beyond the urban region, there is the **residential commuting area**. This area depends heavily on the urban region for jobs and forms with the urban region the **residential urban complex**.

[7] It proposes the following English translation of its own designation: Directorate-general Statistics and Economic information. This translation does not seem, however, to be very usual.

Figure 8 Belgian urban regions in 2001.



Source: Van Hecke *et al.* (2009), p.108.

3.5 Comparing Agglomeration Definitions: An Increased Focus on Commuting

All the aforementioned nomenclatures use several indicators that we may classify according to the following typology:

- Morphological indicators: continuity of the built-up area, housing structure, density or level of population as a land use proxy ;
- Functional indicators: median income, employment density, employment threshold, presence of business and services, growth of the population or of the built area; in-commuting and out-commuting;
- Structural indicators: share of the primary sector.

All the definitions given by those nomenclatures perceive urban agglomerations as concentric spatial entities: an agglomeration is roughly defined as an urban core surrounded by an outlying area.

However, the aforementioned indicators have not the same weight in the definition of the inner and outer parts of the agglomeration. Morphological indicators play a more important role in the definition of the urban core, while functional indicators are more relevant for the outlying area.

However, we may distinguish three kinds of definition by the way the outlying area is defined:

- The Swiss delineation retains a mix of functional and morphological criteria to define the outlying administrative districts;
- The Belgian nomenclature retains a mix of functional criteria to delimit the suburb and the commuting criteria to delineate the residential commuting area;
- The US and the French definitions delineate the outlying areas using commuting criteria only.

We can denote that the emphasis on commuting criteria is rising from the Swiss to the US and French definitions[8], while the accent on morphological criteria is decreasing.

Therefore, at the two extreme, there is two different definitions of urban agglomeration: on one hand a definition based on the postulate that an agglomeration is multidimensional phenomenon characterized both by morphological criteria (continuity of the built-up area), and functional features (commuting), and on another hand a functional definition attempting to capture the influence an urban core may have on its catchment area.

Defining an urban entity is a difficult task, because cities and towns can receive a variety of social meanings. Moreover, there is no universal threshold of population density that can be associated to the definition of a city for differentiating urban from rural settlements (Pumain, 2003).

[8] If we consider that the Swiss definition has evolved by raising its emphasis on commuting, we realize the increasing importance of this factor on agglomeration definition.

As stated previously, urbanization is very complex. Agglomeration delimitation differs with respect to the goal pursued. For instance, the aim of the Standards for defining Metropolitan and Micropolitan Statistical Areas is to provide nationally consistent definitions for collecting, tabulating, and publishing Federal statistics for a set of geographic areas. To this end, Metropolitan Areas were designed as statistical representation of the social and economic linkages between urban cores and outlying, integrated areas. The Office of Management and Budget warns clearly that Core Based Statistical Areas (CBSAs) "should not serve as a general purpose geographic framework for nonstatistical activities and may or may not be suitable for use in program funding formulas" (Federal Register, 2000).

The delineation of the outlying area of such functional regions should discard morphological criteria since they do not capture the enlarged range of spatial interaction allowed by the modernization of transport technology (Pumain, 2003). Taking this argument into account, Kammermann (2007), propose to review the current Swiss definition of agglomeration and to replace by a concept perceiving agglomeration as a network.

4 UrbanSim and data structures

There are substantial differences between the features of European and US agglomerations: European cities are denser, more compact than US cities, they show up a concentric design, while most of Anglo-American cities developed according to Grid plans, and they display contrasting social geography. This should have an impact on the choice of geographical units that are the most appropriate to model European cities.

4.1 Geographic Units of Analysis in UrbanSim Models

UrbanSim and the data structures used in it have evolved in the recent years. Until 2005, most UrbanSim applications were based on a widespread use of grid overlaid on study areas. The numerous shortcomings of the grid cell approach led to the recent adoption of a parcel-based data and spatial structure. Referring to UrbanSim User guide and reference manual (Center for Urban Simulation and Policy Analysis, 2009, chapters 17-21), we describe the fundamental differences between these data structures. While both are still used and supported approaches, the advantages of the parcel approach appear as fairly significant. UrbanSim currently supports three data structures: grid cells, parcels, and zones.

4.1.1 Grid Cells

The decision about the resolution to use for a grid to overlay the study area is crucial to implement the grid cell-based approach to developing the data for UrbanSim. The choice of 150 meters by 150 meters was made in early UrbanSim applications, mainly as a compromise between the high level of resolution desired, and the increased computational demands made by higher resolution data.

The principal advantage of using a grid is that it renders possible the use of efficient raster processing as in image processing or raster GIS spatial analysis. For instance, it is possible to compute effectively how much population or employment is within a fixed radius of each cell. Such a computational efficiency was the most important motivation for using the grid cell approach to structuring the input data for UrbanSim.

Each grid cell contains approximately 5.5 acres, at a resolution of 150 meters. In order to prepare the data for UrbanSim, parcel maps are overlaid in the GIS on a vector representation of grid cells, and the contents of the parcel (housing, etc) allocated to the grid cells proportionally to its land area falling within each grid cell. The fragments of the real estate components created in this way are aggregated into a composite at the cell level. UrbanSim then operates on the grid cell-level data. To better reflect the contents of the grid cells, which are clearly heterogeneous in their composition, building objects were created to allow at least different types of real estate in a cell to be represented by different types of buildings. Households and jobs were then associated with buildings, and buildings with grid cells.

However, the grid cell based approach presents several shortcomings. Grid cells bisect parcels; therefore, it is not possible to aggregate parcel information neatly into grid cells. This is an evident outcome of imposing a completely regular shape on a polygonal layer of parcels that vary in size or in shape. Hence, since this approach implies unnaturally splitting the underlying parcel information, recombining it may create artificial representations of the data. Consequently, with this approach applying information on development regulations from general plans is more involved since those are also based on polygons, and apply to parcels.

4.1.2 Parcels

Recent development on UrbanSim has adopted a data structure based on parcels to address some of the limitations of the grid cell-based data structure. The parcel-based UrbanSim application uses a data model that reflects parcels, buildings, households and jobs as the primary objects and units of analysis. Households and jobs choose locations by selecting a specific building, which is associated with a specific parcel. Real estate development is underpinned on development projects occurring on specific parcels.

Parcels data structure appears more suitable to European cities. Indeed, conversely to the US agglomerations that develop according to a grid plan, Europeans are characterized by a concentric and more complex data structure. The shape of the basic geographical units in

Europe cities result is irregular. They are better described by polygons of various size and patterns than uniform grid cells.

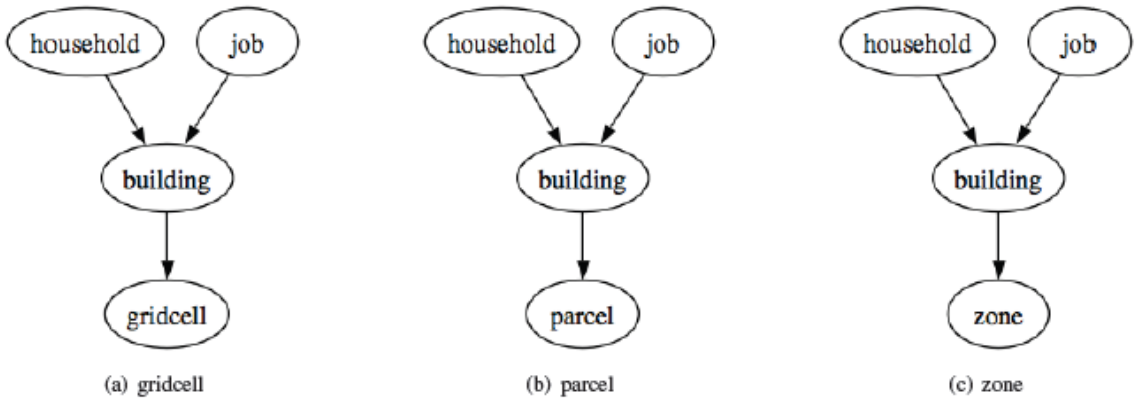
Figure 9 Grid cells and Parcels in UrbanSim.

Source: Waddell (2002).

4.1.3 Zones

Zones constitute alternative data structures that can be substituted for parcel or grid cell based data. The zone based data approach may allow creating a simple model using less geographical detail. The zone based data approach use the same data structure for households, jobs and buildings. The unique change it requires is to assign locations to buildings at a coarser level of detail. Such an approach would keep all the accounting systems in the UrbanSim model: households and jobs are still located in buildings, and buildings can be linked spatially to zones (figure 10). However, it would let aside considerable detail for analyzing development locations and capacity constraints due to zoning or land use plans.

Figure 10 Flexible Geographic Units for UrbanSim.



Source: Center for Urban Simulation and Policy Analysis (2009), p. 85.

4.2 Impact of the Choice of Basic Spatial Unit on UrbanSim Model Components.[9]

The current version of UrbanSim supports three types of database:

- baseyear database: defines the initial state of a simulation in a particular base year.
- scenario database: defines changes to a baseyear (or another scenario) database.
- output database: optional repository for simulation results.

It also supports several database servers: MySQL (which is the most used and tested version), Postgres, SQLite, Microsoft SQL.

Figure 11 describes the overall architecture of UrbanSim model system. The overall logic is essentially the same either for grid cells or parcels based data structures. The only thing that differs is the configuration of particular models.[10] Table 4 summarizes the specifications of models used in the parcel version of UrbanSim. They differ significantly from earlier grid cells versions. For instance, in addition to the substitution of parcels for grid cells as the basic spatial unit, the real estate development model was completely restructured in order to take advantage of the availability of parcel geography in representing actual development projects

[9] More details on the models that compose UrbanSim may be found in Center for Urban Simulation and Policy Analysis (2009).

[10] More details on the UrbanSim model components may be found on Center for Urban Simulation and Policy Analysis (2009).

- which do vary in size and shape in the real world, in ways that were hardly compatible with grid cell geography.

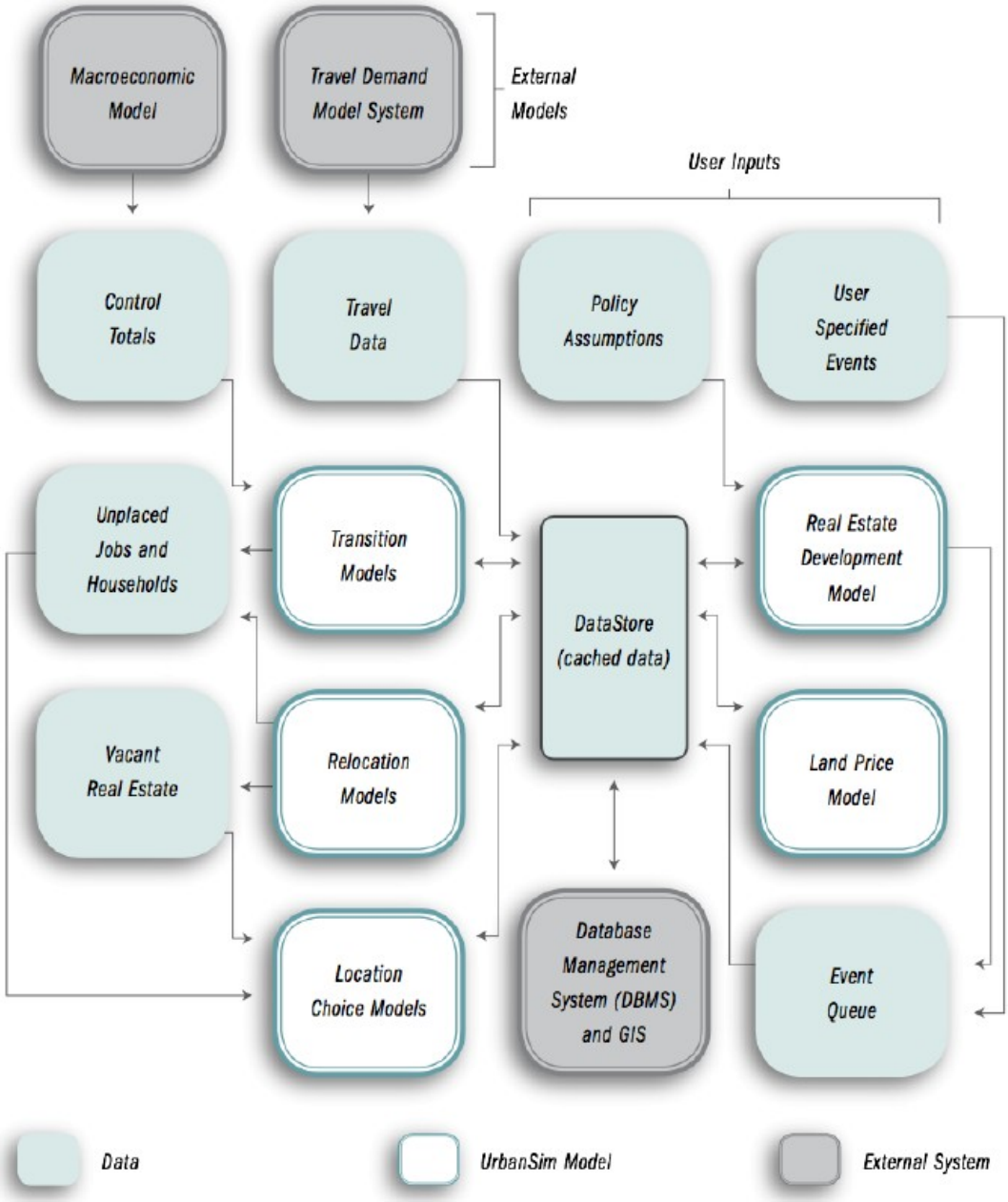
Moreover, the parcel based model specifications also have recently added models to predict the choice of workers to be home-based (normally work from home), and a workplace choice model for workers who are not home-based. This permits a proper handling of the prediction of commuting behavior as a long-term outcome of where a household chooses to live, and where the workers in the household have jobs, and allows the removal of the home-based-work trip distribution model from the set of behaviors predicted by the travel model on a daily basis.

Table 4 Specification of UrbanSim Model Components Using Parcel Data Structure

Model	Agent	Dependent Variable	Functional Form
Household Location Choice	Household (New or Moving)	Residential Building With Vacant Unit	Multinomial Logit
Employment Location Choice	Job (New or Moving)	Non-residential Building With Vacant Space	Multinomial Logit
Home-based Job Choice	Worker (Without Job)	Binary Choice (Work at Home)	Binary Logit
Workplace Choice	Non Home Based Worker (Without Job)	Vacant Job	Multinomial Logit
Real Estate Development	Development Proposal	Parcel (With Vacant Land)	Multinomial Logit Sampler
Real Estate Price	Parcel	Price Per Square Foot	Multiple Regression

Source: Center for Urban Simulation and Policy Analysis (2009), p. 87.

Figure 11 Overview of UrbanSim Model System..



Source: Center for Urban Simulation and Policy Analysis (2009), p. 86.

4.3 Data Requirement for UrbanSim Applications.

UrbanSim is an evolving set of models, some of which have been adapted to different data structures and geographic units of analysis, such as grid cells, parcels, buildings and zones. Each of these models, depending on how the user specifies the model, creates its own data

requirements. Therefore, documenting a universal set of data requirements for all UrbanSim users is impossible.

4.3.1 General Tables Used in UrbanSim.

Most of UrbanSim tables are optional. The required set of tables is determined by the set of models configured for a run. In this section we describe the database tables that are of general use, i.e. not specific to either grid cell or parcel or zone based applications.

Databases Tables About Employment.

a) The annual_employment_control_totals table.

This table gives total target quantities of employment, by sector, by home-based, and by year for each simulated year.

Column Name	Data Type	Description
sector_id	integer	Index into the `employment_sectors` table
year	integer	
total_home_based_employment	integer	Target home based employment for this sector and year
total_non_home_based_employment	integer	Target non-home based employment for this sector and year

b) The annual_relocation_rates_for_jobs table.

This table is only used by the Employment Relocation Model.

Column Name	Data Type	Description
sector_id	integer	Index into the `employment_sectors` table
job_relocation_probability	float	Probability that a job in this sector will relocate within the time span of one year

c) The employment_sectors table.

An EmploymentSector is a logical category of employment, such as “automobile_sales” or “shipping”. Each row defines one EmploymentSector.

Column Name	Data Type	Description
sector_id	integer	Unique identifier
name	varchar	Unique name of the Sector

d) The employment_adhoc_sector_groups table.

Each row defines one EmploymentAdHocSectorGroup, but not the group's membership - the memberships are defined in the employment_adhoc_sector_group_definitions table.

Column Name	Data Type	Description
group_id	integer	Unique identifier
name	varchar	Unique name of the Group

e) The employment_adhoc_sector_group_definitions table.

This table defines the set of employment_sectors in each EmploymentSectorAdHocGroup. Each row defines one "belongs to" relationship (a particular EmploymentSector "belongs to" a particular EmploymentSectorAdHocGroup).

Column Name	Data Type	Description
sector_id	integer	Index into the `employment_sectors` table
group_id	integer	Index into the `employment_adhoc_sector_groups` table

f) The job_building_types table.

This table describes building types for jobs.

Column Name	Data Type	Description
id	integer	Unique identifier
name	string	Name of type, e.g. "commercial"
home_based	boolean	True if home-based

Databases Tables About Households.

a) The annual_household_control_totals table.

This table gives target quantities of households classified by year and an optional set of other user-defined attributes, such as race of head, or size of household. Each attribute is a column.

The table's key is a combination of all attributes other than total_number_of_households. The table must contain a row for each attribute and each simulated year.

Column Name	Data Type	Description
year	integer	Year for the total
age_of_head	integer	*(optional)* Household characteristic bin number of age of head of household
cars	integer	*(optional)* Household characteristic bin number of number of cars in household
children	integer	*(optional)* Household characteristic bin number of number of children in household
income	integer	*(optional)* Household characteristic bin number of household income
persons	integer	*(optional)* Household characteristic bin number of size of household in number of people
race_id	integer	*(optional)* Household characteristic bin number of race of head of household
workers	integer	*(optional)* Household characteristic bin number of employed people in household
total_number_of_households	integer	Target number of households of this household type and year

b) The annual_relocation_rates_for_households table.

The table gives the annual relocation rates for households, by combination of age and income of household. These values are the probabilities that a household with the given characteristics will relocate within the time span of one year. They do not change from year to year.

Column Name	Data Type	Description
age_min	integer	The minimum age for which this probability is valid.
age_max	integer	The maximum age for which this probability is valid, -1 means no maximum
income_min	integer	The minimum income for which this probability is valid.
income_max	integer	The maximum income for which this probability is valid, -1 means no maximum
probability_of_relocating	float	The probability of relocating in a year.

c) The households table.

This table contains only row per household in the region. All people in the region belong to exactly one household. The table below, which is from a gridcell-based application, also works for a parcel-based application with only one exception: the gridcell identifier column should be replaced by a building id.

Column Name	Data Type	Description
household_id	integer	Unique identifier
grid_id	integer	Grid cell this household resides in; zero if currently not residing in a housing unit
persons	integer	Total number of people living in this household.
workers	integer	Total number of workers living in this household.
age_of_head	integer	Age of head of the household
income	integer	Income of this household
children	integer	Number of children living in this household
race_id	integer	Race of head of household
cars	integer	Number of cars in this household

d) The household_characteristics_for_ht table.

This table gives bin definitions for the characterizing households used by the Household Transition Model to produce an N-dimensional partitioning of the households. The names of the characteristics must match attribute names in the households table. If a characteristic is used in the table annual_household_control_totals, the names in both tables must also match.

The table has the following structure:

Column Name	Data Type	Description
characteristic	varchar	See above for examples
min	integer	Minimum value for this bin for this characteristic. Values are placed in a bin iff min <= value <= max
max	integer	Maximum value for this bin for this characteristic; -1 means infinity / no maximum

e) The race_name table.

This table is required in the only case where race related variables are included in model specifications. It has one row per race.

Column Name	Data Type	Description
race_id	integer	Unique identifier
name	varchar	Name of the race
minority	boolean	True if the race is a minority

Databases Tables About Transportation Analysis Zones.

a) The zones table.

Traffic analysis zones are spatial entities. In UrbanSim, these zones are rasterized by the grid cells.[11] In practice, the zones table often includes other columns, depending upon the needs of the models. These data should be updated with the results of any travel model run with whatever attributes are needed.

Column Name	Data Type	Description
zone_id	integer	Unique identifier
travel_time_to_airport	integer	*(optional)* Units: Minutes
travel_time_to_cbd	integer	*(optional)* Units: Minutes
faz_id	integer	*(optional)* Foreign key of the FAZ (forecast analysis zone) containing this zone.

b) The travel_data table.

The travel data can be grasped as the composite utility of going from one location to another given the available travel modes for that household type.[12] Intrazonal travel may have less utility than interzonal travel if mass transit routes or highway options allow for easier travel to an adjacent zone than within a zone. logsum3 often shows lower utility than logsum2 because the logsums represent composite utilities for different household types. So, for example, it may be that 2 car households tend to have a more favorable person- to-car ratio than 3+ car households. Or it may be that 2 car households are more frequently able to combine trips, decreasing the disutility of any individual trip. These data should be updated with the results of any travel model run.

Column Name	Data Type	Description
from_zone_id	integer	"From" traffic analysis zone
to_zone_id	integer	"To" traffic analysis zone
logsum0	float	*(optional)* Logsum value for 0 vehicle households, transit logsum
logsum1	float	*(optional)* Logsum value for 1 vehicle households, transit logsum
. . .	float	. . .
logsumN	float	*(optional)* Logsum value for N+ vehicle households, transit logsum

[11] the zones are distorted to fit to cell boundaries and thus will have rough or stair-stepped edges

[12] Negative values reflect the fact that the time required gives the trip negative utility.

Other Databases Tables.

a) The base_year table.

This table is optional. It is only used if the base year is not defined in the configuration. It has one row only.

Column Name	Data Type	Description
year	integer	Year of base data

b) The cities table.

The table is only needed if you want to create indicators on city level.

Column Name	Data Type	Description
city_id	integer	Unique identifier
city_name	varchar	

c) The counties table.

The table is only needed if you want to create indicators on county level.

Column Name	Data Type	Description
county_id	integer	Unique identifier
county_name	varchar	

d) The scenario_information table.

This table gives the description of the scenario. It has one row only.

Column Name	Data Type	Description
description	varchar	*(optional)* Human readable description
parent_database_url	varchar	The name of the next database in the chain of scenario databases.

e) The urbansim_constants table.

This table gives the constants required for computations made by the various models. It has a single row with one column per constant.

Column Name	Data Type	Description
cell_size	float	Width and height of each grid cell in units
units	varchar	Units of measurement, eg. "meters" or "feet"
walking_distance_circle_radius	float	Walking distance in meters, e.g., 600 m
young_age	integer	Max age for a person to be considered young
property_value_to_annual_cost_ratio	float	Ratio of the total property value to an annual rent for that property
low_income_fraction	float	Fraction of the total number of households considered to have low incomes, e.g., 0.1
mid_income_fraction	float	Fraction of the total number of households considered to have mid-level incomes, e.g., 0.5
near_arterial_threshold	float	Line distance from the centroid of a cell to an arterial for it to be considered nearby, e.g., 300
near_highway_threshold	float	Line distance from the centroid to a highway for it to be considered nearby, e.g., 300
percent_coverage_threshold	integer	The threshold above which a grid cell's percent_*, e.g. percent_wetland, must be to be considered "covered" for that attribute. So, if percent_coverage_threshhold is 50 percent and percent_wetland is 60 percent, the grid cell would be considered "covered" by wetland.
recent_years	integer	Maximum number of years to look back when considering recent transitions. For example, if recent_years = 3, then the value commercial_sqft_recently_added in the `gridcells` table would refer to the number of square feet of commercial space built in the last 3 years.

4.3.2 Data for Grid Cells Based Applications.

Grid cells based applications require several kinds of data:

- Database tables about Grid cells;
- Database tables about Development types;
- Database tables about Employment events;
- Database tables about Development constraints;
- Database tables about Target vacancies.

Databases Tables About Grid Cells.

a) The gridcells table.

This table contains the geographic information partitioned into a rectangular grid of rectangular cells. Attributes that are marked as optional are only required by specific

variables. Thus, the requirement of those attributes depends on model specifications. Attributes not marked as optional are used by various models.

Column Name	Data Type	Description
grid_id	integer	Unique identifier
commercial_sqft	integer	The sum of the square footage of buildings that are classified as commercial (generally including retail and office land uses). This is not a measure of land area.
development_type_id	integer	Index into the Development Types table
distance_to_arterial	float	*(optional)* Units: `urbansim_constants`.units
distance_to_highway	float	*(optional)* Units: `urbansim_constants`.units
governmental_sqft	integer	The sum of the square footage of buildings that are classified as governmental
industrial_sqft	integer	The sum of the square footage of buildings that are classified as industrial
commercial_improvement_value	integer	See description, above
industrial_improvement_value	integer	See description, above
governmental_improvement_value	integer	See description, above
nonresidential_land_value	integer	Units, e.g. dollars
residential_improvement_value	integer	See description, above
residential_land_value	integer	Units, e.g. dollars
residential_units	integer	Number of residential units
relative_x	integer	X coordinate in grid coordinate system
relative_y	integer	Y coordinate in grid coordinate system
year_built	integer	e.g. 2002
plan_type_id	integer	An id indicating the plan type of the grid cell
percent_agricultural_protected_land	integer	*(optional)*
percent_water	integer	*(optional)* Percentage of this cell covered by water
percent_stream_buffer	integer	*(optional)* Percentage of this cell covered by stream buffer
percent_floodplain	integer	*(optional)* Percentage of this cell covered by flood plain
percent_wetland	integer	*(optional)* Percentage of this cell covered by wetland
percent_slope	integer	*(optional)* Percentage of this cell covered by slope
percent_open_space	integer	*(optional)* Percentage of this cell covered by open space
percent_public_space	integer	*(optional)* Percentage of this cell covered by public space
percent_roads	integer	*(optional)* Percentage of this cell covered by roads
percent_undevelopable	integer	*(optional)*
is_outside_urban_growth_boundary	boolean	*(optional)*
is_state_land	boolean	*(optional)*
is_federal_land	boolean	*(optional)*
is_inside_military_base	boolean	*(optional)*
is_inside_national_forest	boolean	*(optional)*
is_inside_tribal_land	boolean	*(optional)*
zone_id	integer	Traffic analysis zone that contains this grid cell's centroid
city_id	integer	*(optional)* City this grid cell belongs to
county_id	integer	*(optional)* County this grid cell belongs to
fraction_residential_land	float	Fraction of residential land in this cell
total_nonres_sqft	integer	*(optional)*
total_undevelopable_sqft	integer	*(optional)*

b) The plan_types table

plan_types are synonymous with Zoning types and with Planned Land Use (PLU) types. The distinction is arbitrary and is to be made by the user. There is only one row per plan type.

Column Name	Data Type	Description
plan_type_id	integer	Unique identifier
name	varchar	Unique name of the Plan Type

Databases Tables About Development Types.

Development types are used to categorize a grid cell according to the "type" of development currently in the grid cell. As an example, grid cells with only a few residential units and no other square footage might be classified as "low density residential" which may be abbreviated as "R1". Other grid cells may be categorized as mixed use, commercial, etc. The set of development types to employ is arbitrary. Development types are grouped by two nested mechanisms: groups and non-overlapping-groups. Each development type may be a member of multiple groups. Each group may be a member of multiple non-overlapping-groups. All of the groups in a non-overlapping-group must be disjoint (i.e., may not share any development types).

Groups and non-overlapping-groups are used in the calculation of the variables in the models, so to fully understand them requires understanding the model definitions.

a) The development_type table.

Each row of this table defines one development type.

Column Name	Data Type	Description
development_type_id	integer	Unique identifier for this row.
name	varchar	Name of the development type.
min_units	integer	Minimum number of units to be in this development type.
max_units	integer	Maximum number of units to be in this development type.
min_sqft	integer	Minimum square feet to be in this development type.
max_sqft	integer	Maximum square feet to be in this development type.

b) The development_type_groups table.

Each row defines one development type group, but not the group's membership - the memberships are defined in the development_type_group_definitions table.

Column Name	Data Type	Description
group_id	integer	Unique identifier for this row.
name	varchar	Unique name of the development type group.
non_overlapping_groups	varchar	Name of the non-overlapping-group or empty for no non-overlapping-group.

The set of required development type groups and non-overlapping-groups is determined by the set of variables used by the models being estimated or simulated. Thus, there is no way to a-priori specify which development type groups will be needed for your application of UrbanSim.[13]

c) The development_type_group_definitions table.

This table defines the set of development_types in each development type group. Each row defines one "belongs to" relationship (a particular development type that "belongs to" a particular development type group).

Column Name	Data Type	Description
development_type_id	integer	Index into the `development_types` table
group_id	integer	Index into the `development_type_groups` table

Databases Tables About Development Events.

These tables represent events in the real estate development model. Events which are planned to take place in the future are stored in the development_events table, events that occured prior to the base year are stored in the development_event_history table. Both tables can contain columns of the pattern "units_change_type". Each value determines a type of change for that type of units. The different values are:

- "A" for Add
- "R" for Replace
- "D" for Delete

If this column is missing for a certain type of units, the default value is "A" for all events.

a) The development_type_group_definitions table.

These development events are changes to grid cells which are planned to occur in the future. For any given year, one may plan any number of changes to the attributes of any number of

[13] There however are two exceptions: first, the model Events Coordinator (Center for Urban Simulation and Policy Analysis, 2009, 25.4.18) is internally using groups 'residential', 'mixed_use', 'commercial', 'industrial', and 'governmental'. Second, the Land Price Model (Center for Urban Simulation and Policy Analysis, 2009, 25.4.1) is using by default a filter that requires a group called 'developable'. Therefore, if the user does not change these settings, he should make sure that the table contains these entries.

gridcells. Each change represents an addition, a subtraction or a replacement of the specified number of sqft, residential units, and improvement values.

For instance, if grid cell 23 is to grow by 200 residential units in 2008 (an apartment building is built), the table would include a row with scheduled_year = 2008, grid_id = 23, residential_units = 200, and residential_units_change_type = 'A'.

The value in the "improvement_value" fields, below, are used to indicate how to change the associated improvement_value for this grid cell. Each event will add/subtract/replace (improvement_value * (number of units [or sqft] being built by this event)) to the current improvement value in this grid cell. The units of the improvement are currency value, e.g. dollars.[14]

Column Name	Data Type	Description
grid_id	integer	Grid cell where the event takes place
scheduled_year	short	Year in which the event will be implemented
residential_units	integer	
commercial_sqft	integer	
industrial_sqft	integer	
governmental_sqft	integer	
residential_units_change_type	char	*(optional)* see 19.3
commercial_sqft_change_type	char	*(optional)* see 19.3
industrial_sqft_change_type	char	*(optional)* see 19.3
governmental_sqft_change_type	char	*(optional)* see 19.3
residential_improvement_value	integer	See description, above
commercial_improvement_value	integer	See description, above
industrial_improvement_value	integer	See description, above
governmental_improvement_value	integer	See description, above

b) The development_event_history table.

The development_event_history table records the development events that occurred prior to the base year. This table uses a subset of the schema used for development_events. It can be considered an extension back in time of the development_events table, though with additional constraints, specified below.

[14] The described procedure is implemented in Events Coordinator (Center for Urban Simulation and Policy Analysis, 2009, 25.4.18).

Column Name	Data Type	Description
grid_id	integer	Grid cell where the event takes place
scheduled_year	short	Year in which the event was implemented
starting_development_type_id	integer	*(optional)* This will be the value of the development_type for this gridcell after "unrolling" this development event.
residential_units	integer	
commercial_sqft	integer	
industrial_sqft	integer	
governmental_sqft	integer	
residential_units_change_type	char	*(optional)* see 19.3
commercial_sqft_change_type	char	*(optional)* see 19.3
industrial_sqft_change_type	char	*(optional)* see 19.3
governmental_sqft_change_type	char	*(optional)* see 19.3
residential_improvement_value	integer	See description, above
commercial_improvement_value	integer	See description, above
industrial_improvement_value	integer	See description, above
governmental_improvement_value	integer	See description, above

Databases Tables About Development Constraints.

a) The development_contraints table.

This table defines rules that constrain the possible development types a developer can create on a particular grid cell. Each row defines one rule. Development is not allowed on any gridcell that matches any of these rules. A grid cell matches a rule if the attribute values for the grid cell match all of the values in the rule (rule columns with the value "-1" are ignored when determining a match).

Column Name	Data Type	Description
constraint_id	integer	Unique rule identification number
name-of-a-gridcell-attribute-1	integer or float	Value for this attribute, or "-1" if this attribute is not part of the constraint (e.g. *don't care*)
...	...	...
name-of-a-gridcell-attribute-N	integer or float	
min_units	integer	Minimum number of residential units for a gridcell. A development project may only be placed on this gridcell if it will result in this gridcell containing at least this number of residential units.
max_units	integer	Maximum number of residential units for a gridcell. A development project may only be placed on this gridcell if it will result in this gridcell containing at most this number of residential units.
min_commercial_sqft	integer	Minimum number of commercial sqft. for a gridcell. A development project may only be placed on this gridcell if it will result in this gridcell containing at least this number of commercial sqft.
max_commercial_sqft	integer	Maximum number of commercial sqft. for a gridcell. A development project may only be placed on this gridcell if it will result in this gridcell containing at most this number of commercial sqft.
min_industrial_sqft	integer	Minimum number of industrial sqft. for a gridcell. A development project may only be placed on this gridcell if it will result in this gridcell containing at least this number of industrial sqft.
max_industrial_sqft	integer	Maximum number of industrial sqft. for a gridcell. A development project may only be placed on this gridcell if it will result in this gridcell containing at most this number of industrial sqft.

Databases Tables About Target Vacancies.

a) The target_vacancies table.

The target_vacancies table gives the model information about acceptable vacancy rates. The table has one row for each year the simulation runs. Each row gives target values for the residential and nonresidential vacancies for that year, which are defined below. Only data after the base year is used.

Column Name	Data Type	Description
year	integer	Year of the simulation for which the vacancy targets apply
target_total_residential_vacancy	float	Ratio of unused residential units to total residential units
target_total_non_residential_vacancy	float	Ratio of unused nonresidential sqft to total nonresidential sqft

4.3.3 Data for Parcels Based Applications.

Parcel based applications require the following types of data:

- Database tables about Parcels;
- Database tables about Buildings;

- Database tables about Development projects;
- Database tables about Development constraints;
- Database tables about Target vacancies;
- Database tables about Refinement of simulation results.

Databases Tables About Parcels.

a) The parcels table.

This table contains attributes about parcels. In general, there will be an identifier in this table for every other level of geography that you may want to aggregate up to. In this example, there are attributes for zones, cities, counties, census blocks, etc. Having these identifiers on the parcel makes it easier to aggregate indicators up to higher level geographies. Any other attributes that one may want to restrict development by, or update throughout a simulation could be stored here as well.

Column Name	Data Type	Description
parcel_id	integer	unique identifier
zone_id	integer	id number for the zone that the parcel's centroid falls within
land_use_type_id	integer	identifies the land use of the parcel
city_id	integer	id number for the city that the parcel's centroid falls within
county_id	integer	id number for the county that the parcel's centroid falls within
plan_type_id	integer	id number that identifies the parcel's plan type
parcel_sqft	integer	square feet of the parcel as an integer
assessor_parcel_id	integer	*(optional)* original tax assessor's id number
tax_exempt_flag	integer	*(optional)* identifies parcel as tax exempt or not
land_value	long	value of the land from the assessor
is_in_flood_plain	integer	*(optional)* indicates whether or not a parcel is in a flood plain
is_on_steep_slope	integer	*(optional)* indicates whether or not a parcel is on a steep slope
is_in_fault_zone	integer	*(optional)* indicates whether or not a parcel is in a fault zone
centroid_x	long	state plane x coordinate of parcel centroid
centroid_y	long	state plane y coordinate of parcel centroid
census_block_id	integer	*(optional)* id number for the census block that the parcel's centroid falls within
raz_id	integer	*(optional)* id number for the raz that the parcel's centroid falls within

Databases Tables About Buildings.

a) The Buildings table.

In recently developed UrbanSim applications, buildings of all kinds are represented in their own table, and linked to the basic spatial unit used for location choice: grid cell, parcel, or zone. This configuration provides simple and flexible means of organizing the data for

UrbanSim. The buildings table is similar for each of the types of applications, whether grid cell, parcel or zone – the only significant difference is the location identifier. In the table below, parcel id is included, but for grid cell or zone applications, the user should substitute grid cell id or zone id.

Column Name	Data Type	Description
building_id	integer	unique identifier
building_quality_id	integer	*(optional)* identified for building quality
building_type_id	integer	identifier for building type; valid id in the `building_types` table
improvement_value	long	value of building (replacement cost)
land_area	long	land area (usually in sqft) associated with building, includes footprint plus associated area such as landscaping and parking.
non_residential_sqft	long	non-residential square footage of building
parcel_id	integer	identifier of parcel in which building is located
residential_units	integer	number of residential units in the building
sqft_per_unit	integer	number of residential square feet per unit in the building
stories	integer	*(optional)* number of stories in the building
tax_exempt	integer	*(optional)* indicator for whether building is tax-exempt
year_built	integer	year of construction of the building

b) The Buildings_types table.

This table gives information about available types of buildings.

Column Name	Data Type	Description
building_type_id	integer	unique identifier
building_type_name	varchar	name of the building type
description	varchar	*(optional)* description of the building type
generic_building_type_id	integer	*(optional)* identifier for generic building type
generic_building_type_description	varchar	*(optional)*
is_residential	boolean	1 if this building type is residential, 0 otherwise
unit_name	varchar	name of units for this building type, e.g. 'commercial_sqft' or 'residential_units'

Databases Tables About Development Projects.

a) The development_project_proposals table.

A record in this table, when combined with one or more records in the development_project_components table, represents a "known" development project. This table would be populated either with projects known to be coming in the future or, during a population run, with projects that are in the middle of their development according to their velocity function.

Column Name	Data Type	Description
development_project_id	integer	unique identifier
development_template_id	integer	indicates the development template that represents the project
far	float	floor to area ratio of the project
percent_open	integer	the percent of the land area of the project accounted for by "overhead" uses such as rights of way or open space
status_id	integer	this represents active, proposed, or planned developments with the following codes: 1: in active development, 2: proposed for development, 3: planned and will be developed, 4: tentative, 5: not available (already developed, 6: refused
parcel_id	integer	indicates the parcel_id on which the development occurs
start_year	integer	the year in which this project is expected to begin building
built_sqft_to_date	integer	the number of non-residential sqft built in the current simulation year
built_units_to_date	integer	the number of residential units built in the current simulation year

b) The development_project_proposal_component table.

A record in this table represents a portion of a development project identified in the development_project_proposals table. In some sense a single record here is meant to represent a single building, or part of a building. Therefore individual records here do not necessarily represent single free-standing buildings, although they are mostly treated that way. This table allows for the flexible representation of mixed uses to occur on a parcel. Examples include multiple free-standing buildings with different uses, a single building with multiple uses inside of it (a single record for each use), or further complex representations of mixed use. This table is not required by UrbanSim, but it is created by the developer model and cached every simulation year.

Column Name	Data Type	Description
development_project_component_id	integer	unique identifier
development_project_id	integer	identifies which development the project belongs to
velocity_function_id	integer	identifies the rate or function by which the project develops over time
percent_of_building_sqft	integer	identifies the percentage of the building that this component takes up. 100% would indicate a free-standing building with a single use, and several records with percent_of_building_-sqft adding up to 100% would indicate a multiple use single building.
construction_cost_per_unit	integer	the per unit construction cost for residential uses only
sqft_per_unit	integer	the square footage per residential unit
building_type_id	integer	indicates the building type of this particular component
land_area	integer	the land area "claimed" by the building component, includ-ing not only the building footprint but also additional land used such as yards, parking lots, etc.
residential_units	integer	the number of residential units in the building component

c) The development_templates table.

This table, along with corresponding records in the development_template_components table, represents development templates that can be used to define virtually any size and configuration of a development project, from a single house on an infill lot to a large subdivision, to a mixed use project with retail on the first floor and condominiums above. The contents of this table are roughly comparable to the development_projects table, since development templates become proposals once they are determined to fit within a parcel and are allowed by development constraints, and then become projects if they are chosen to be constructed.

Column Name	Data Type	Description
development_template_id	integer	unique identifier
percent_open	integer	the percent of the land area of the project accounted for by "overhead" uses such as rights of way or open space
min_land	integer	minimum amount of land in square feet to be utilized for this development
max_land	integer	maximum amount of land in square feet to be utilized for this development
density_type	integer	a readable name that describes the 'density' field: units per acre, FAR
density	integer	indicates the density of the development
land_use_type_id	integer	specifies the land use type for the development template
development_type	integer	a readable name that describes the type of development this record represents (e.g. SFR-parcel, MFR-apartment, MFR-condo, etc.), this field is not used by the model and is there to make the table more readable

d) The development_template_component table.

This table is roughly equivalent to the development_project_proposal_components table and represents buildings or parts of buildings to be included in a particular development template. By breaking development templates into components, development project templates can be configured as hierarchies or combinations of building blocks, providing a very flexible mean of representing a wide variety of development types. Note that the templates can be generated using real or hypothetical data, since they will be compared to regulatory constraints and the size constraints of parcels.

Column Name	Data Type	Description
development_template_component_id	integer	unique identifier
development_template_id	integer	indicates which development template this component belongs to
velocity_funtion_id	integer	indicates the velocity function used by this template
building_type_id	integer	indicates the building type of this particular component
percent_of_building_sqft	integer	identifies the percentage of the building that this component takes up
construction_cost_per_unit	integer	the per unit construction cost
building_sqft_per_unit	integer	the square footage per residential unit

e) The velocity_function table.

This table is designed to hold the velocity functions that specify the rate at which development is built out.

Column Name	Data Type	Description
velocity_function_id	integer	unique identifier
annual_construction_schedule	string	this field will contain a numbered list in brackets of this form: '[25, 50, 75, 100]' indicating with each entry the percentage complete that the project would be in each year from its initiation. A particular development_template_component or development_project_component will have one velocity_function_id attached to it. This could take the form [0, 0, 0, 33, 66, 100] for example, would have no construction in its first 3 years, then in year 4 it would be 33% complete, and 66% and 100% complete in years 5 and 6 respectively.

f) The demolition_cost_per_sqft table.

This table provides information to the developer model about the costs of demolition by building type. These numbers are used to calculate the cost of demolition of existing development so that a more accurate cost of redevelopment can be calculated.

Column Name	Data Type	Description
building_type_id	integer	building type
demolition_cost_per_sqft	integer	cost in dollars per sqft of demolition

g) The building_sqft_per_job table.

This table contains information on the amount of space each job will take in a particular building type, by zone.

Column Name	Data Type	Description
zone_id	integer	the zone the record applies to
building_type_id	integer	the building type the record applies to
building_sqft_per_job	integer	the sqft per job each job will take in a particular building type in a particular zone

Databases Tables About Development Constraints.

a) The development_constraints table.

Since the parcel based real estate development model is fundamentally different than its grid cell based counterpart, this table differs substantially from the grid cell based version application. This table defines rules that restrict the possible development types a developer can create on a particular parcel. Each row defines one rule. Development is not allowed on any parcel that matches any of these rules.

Column Name	Data Type	Description
constraint_id	integer	Unique rule identification number
constraint_type	string (14)	units_per_acre or far (floor-area-ratio)
generic_land_use_type_id	integer	Id of a record in the generic_land_use_types table
maximum	integer	Maximum value for the allowed development, in terms of the constraint type
minimum	integer	Minimum value for the allowed development, in terms of the constraint type
plan_type_id	integer	Id of a record in the plan_types table.

Databases Tables About Target Vacancy Rates.

a) The target_vacancies table.

The target_vacancies table is used by the development proposal choice model. It gives the model information about acceptable vacancy rates. The table has one row for each year the simulation runs. Each row gives target values for the residential and nonresidential vacancies for that year, which are defined below. Only data after the base year is used.

Column Name	Data Type	Description
year	integer	Year of the simulation for which the vacancy targets apply
target_vacancy	float	Ratio of unused space to total space, based on residential_unit or sqft
building_type_id	Integer	Id of a record in the building_types table

Database Tables About Refinement of Simulation Results.

a) The refinements table.

The entries in this table define refinements to make to an existing simulation run. No fields can be null, if the attribute is not needed put a single quote (') in the field.

Column Name	Data Type	Description
refinement_id	integer	unique identifier
agent_expression	string	string expression defining what agents to add or subtract, e.g. households, jobs
location_capacity_attribute	string	defines a capacity attribute such as non_residential_sqft
location_expression	string	expression defining where to add or subtract agents, e.g. 'zone = 123'
amount	integer	number of agents to add or subtract
year	integer	year to which this refinement applies
action	string	add, subtract, or target are the valid entries
transaction_id	integer	if two or more records have matching transaction ids the refinement model will attempt to balance between the refinements

4.3.4 Data for Zone Based Applications.

The zone based modeling is the most recent model system. It may even be considered experimental at this point. As it was modeled after that grid cell model system, it has many tables in common with it. Here are tables unique to the zone based model system.

Database Tables About Buildings.

a) The pseudo_buildings table.

In order to test the zonal-level version of UrbanSim, a pseudo-buildings table has been created. It contains the summary contents of the real estate development model in a zone. Pseudo buildings are meant to represent the amount of commercial, governmental, industrial, and residential space in a zone. There are 4 pseudo building records per zone_id, 1 each for each of the land uses. The attributes are updated during the simulation run by the model system.

Column Name	Data Type	Description
pseudo_building_id	integer	unique identifier
annual_growth	integer	*(optional)* this is the amount that this type of building is allowed to grow per simulation year in terms of floor space or residential units
residential_units	integer	the number of residential units for residential pseudo buildings
zone_id	integer	the zone in which this pseudo building is in
avg_value	integer	the average value per unit or job space depending on the building_type_id
building_type_id	integer	the building type that matches up with the building_types table
job_spaces_capacity	integer	the total number of job spaces allowed in this pseudo building
residential_units_capacity	integer	the total number of residential units allowed in this pseudo building
commercial_job_spaces	integer	the total number of commercial job spaces currently in this pseudo building
industrial_job_spaces	integer	the total number of industrial job spaces currently in this pseudo building
governmental_job_spaces	integer	the total number of governmental job spaces currently in this pseudo building

5 Available Data

For the three cases studies, the available data allow to implement UrbanSim parcel based applications. Indeed for each of the case studies, land use and real estate date contains geographical data on parcels and on buildings. Data on employment, households; and transportation networks are also available.

Stratec and EPFL have drawn up the list of all the available data required to run UrbanSim for Brussels case study. Conversely to the other case studies, the application of Urban Sim to Brussels is fairly recent. All the available data for the Brussels, Zurich and Paris case studies are respectively in Appendix 1, 2 and 3. Brussels data list have been obtained from EFPL and Stratec. Zurich and Paris data have been obtained from Sustaincity Consortium agreement (2010).

6 References

Allain, R. (2004) Morphologie urbaine. Géographie, aménagement et architecture de la ville. Armand Colin, Paris, 2004.

Beaujeu-Garnier, J. (2004) *Geographie urbaine.* Armand Colin, Paris, 1995.

Brueckner, J.K., J.-F. Thisse, Y. Zenou (1999) Why is central Paris rich and downtown Detroit poor? An amenity-based theory. *European Economic Review* 43, 91-107.

Briant, A., *et al.* (2009) Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics* (Forthcoming).

Center for Urban Simulation and Policy Analysis. (2009), The Open Platform for Urban Simulation and UrbanSim Version 4.2.2. Users Guide and Reference Manual. University of Washington.

Dujardin, C., I. Thomas and H. Tulkens (2007) Quelles frontières pour Bruxelles? Une mise à jour, *Reflets et Perspectives de la Vie Economique*, XLVI,156-176.

Federal Register (2000) *Standards for defining metropolitan and micropolitan statistical areas*, Federal Register, vol. 65, n° 249, Wednesday, December 27, 2000/Notices.

Fellmann, J.D, Getis, A., Getis, J., (2003) *Human Geography. Landscapes of Human Activites*. McGraw-Hill, 2003.

Kammermann, M.A. (2007), *Où commence l'agglomération de Genève, où finit celle de Zurich?*, Office fédéral de la statistique OFS, downloaded March, 15th , 2010 at http://www.bfs.admin.ch/bfs/portal/fr/index/news/medienmitteilungen/01.Document.89476.pdf

Kaplan, D.H., Wheeler, J.O., Holloway, S.R., Hodler, T.W. (2004), *Urban Geography*. John Wiley & Sons, 2004.

Paulet, J.-P. (2000), *Géographie urbaine*. Armand Colin, Paris, 2000.

Pumain, D. (2003), Scaling laws and urban systems. Mimeo

Schuler, M., M. Perlik and P. Dessemontet (2009) Agglomérations et aires urbaines dans l'espace transfrontalier Ain - Haute-Savoie - Genève - Vaud, Observatoire Statistique Transfontalier.

Sustaincity Consortium agreement (2010), version 0.6, Tuesday, 16 February

Van der Haegen H., Van Hecke, E., Juchtmans, G. (1996), Les régions urbaines belges en 1991, Institut National de Statistique, *Études Statistiques*, n° 104.

Van Hecke, E., Halleux, J-M, Decroly, J-M, Mérenne-Schoumaker, B. (2009), *Noyaux d'habitat et Régions urbaines dans une Belgique urbanisées*, Enquête Socio-Economique 2001, Monographies, Direction générale Statistique et Information économique.

Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of American planning Association,* 68(3), 297-314.

Waddell, P. and Borning, A. (2004), A Case Study in Digital Government: Developing and Applying UrbanSim, a System for Simulating Urban Land Use, Transportation, and Environmental Impacts. *Social Science Computer Review*, 22 37-51.

Waddell, P., Ulfarsson, G., Franklin, J., & Lobb, J. (2007). Incorporating land use in metropolitan transportation planning. *Transportation Research Part A*, 41(5), 382-410

7 Appendix

7.1 Available data for Brussels case study

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use[15]
PERIMETERS[16]	Maps of the administrative boundaries	GIS – Raster	Belgium	Institut Géographique National - IGN http://www.ngi.be/FR/FR4-1-1.shtm	P
	Administrative boundaries of the regions, provinces, districts and communes	GIS – Vector	Belgium	Institut Géographique National - IGN http://www.ngi.be	
	Limit of the RER area (study area)	GIS – Vector	RER zone	STRATEC	P
	Limit of the Brussels morphological and functional agglomeration	GIS - Vector	Brussels agglomeration	STRATEC / UCL	P
NETWORKS	General maps of the road, rail and hydrographic networks : scales 1/800 000 and 1/200 000	GIS - Raster	Belgium	Institut Géographique National - IGN http://www.ngi.be/FR/FR4-1-2.shtm	P
	Vectorized road networks	GIS - Vector	Belgium	Institut Géographique National - IGN http://www.ngi.be	

[15] P=public, R=restricted, C=confidential

[16] The coordinate system of the geo-referenced maps and GIS files is the Belgian Lambert 1972 projection.

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
	Road network - Navteq data	GIS - Vector	Europe/World	Navteq http://www.navteq.com/	
	Road network - Tele-Atlas data	GIS - Vector	Europe/World	Tele-Atlas http://www.teleatlas.com/index.htm	
	Detailed vectorized road network of Wallonia	GIS - Vector; Interactive map	Wallonia	Mapping Portal of the Walloon Region or via the ArcGIS server of the Walloon Region (cartopro) http://cartographie.wallonie.be/NewPortailCarto/	P
	Detailed vectorized Brussels road network – URBIS	GIS - Vector	Brussels	Région de Bruxelles-Capitale http://www.cirb.irisnet.be/site/fr/departements/services/urbis/	
	Brugis - Interactive charts of Brussels Region	Interactive map - Image	Brussels	BruGIS - Site cartographique de la Région bruxelloise http://www.brugis.irisnet.be/brugis/fr/index.html	P
	TEC - Walloon public transport network maps and schedules	PDF	Wallonia	Transport En Commun - TEC http://www.infotec.be/index.aspx?PageId=633009288182808470	P
	De Lijn - Flemish public transport network maps and schedules	PDF	Flanders	De Lijn http://www.delijn.be/reisinformatie/netplannen/index.htm	P
	STIB - Brussels public transport network map and schedules	PDF - GIS Raster	Brussels	Société des Transports Intercommunaux de Bruxelles - STIB/MIVB http://www.stib.be/netplan-plan-reseau.html?l=fr	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use[17]
	Brussels public transport network - URBIS	GIS - Vector	Brussels	Région de Bruxelles-Capitale http://www.cirb.irisnet.be/site/fr/departements/services/urbis/	
	Vectorized rail networks and stations (SNCB)	GIS - Vector; PDF	Belgium	Institut National Géographique - IGN SNCB http://www.ngi.be www.sncb.be	
	Vectorized hydrographic networks	GIS - Vector	Belgium	Institut National Géographique - IGN http://www.ngi.be	

[17] P=public, R=restricted, C=confidential

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
RELIEF	Relief map (altitude and level lines)	Image	Belgium	Institut Géographique National - IGN http://www.ngi.be	
	Vectorized data (altimetry curves, rating points, geodetic points)	GIS - Vector	Belgium	Institut Géographique National - IGN http://www.ngi.be	
	Digital Ground Model - DGM *(Modèle Numérique de Terrain - MNT*)	GIS - Raster	Belgium	Institut Géographique National - IGN http://www.ngi.be	
LAND USE AND LAND COVER	CORINE land use maps	GIS - Raster	Europe	European Environment Agency - Corine http://dataservice.eea.europa.eu/dataservice/	P
	Land use maps of Belgium - 11 classes	PDF	Belgium	Région wallonne http://developpement-territorial.wallonie.be/Dwnld/Cartes/sd01.pdf	P
	Land use evolution (changes)	Excel	Belgium	SPF Economie - STATBEL - SPF Finances http://statbel.fgov.be/fr/binaries/solhist_fr_tcm326-34198.xls	P
	"Plan Régional de Développement" - PRD General Information and maps	PDF	Brussels-Capital Region	Région de Bruxelles-Capitale http://www.prd.irisnet.be/Fr/info.htm	P
	"Plan Régional d'Affectation du Sol" - PRAS - Prescrip-tions	PDF	Brussels-Capital Region	Région de Bruxelles-Capitale http://www.pras.irisnet.be/PRAS/FR/Frame-menufr.htm	P
	"Plan Régional d'Affectation du Sol" - PRAS - Maps	Interactive map; GIS Server	Brussels-Capital Region	Région de Bruxelles-Capitale http://geowebas1.ci.irisnet.be/PRASAFFECTATIONFR/viewer.htm http://www.brugis.irisnet.be/brugis/framesetup.asp	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
	"Schéma de développement de l'espace régional" - SDER	Image; PDF	Walloon Region	Région wallonne http://developpement-territorial.wallonie.be/pages/Quoi.html http://developpement-territorial.wallonie.be/pages/Cartes.html	P
	"Plan de Secteur wallon" - PS (Sector plan)	GIS - Raster; PDF	Walloon Region	Direction générale de l'Aménagement du territoire, du Logement et du Patrimoine http://mrw.wallonie.be/dgatlp/dgatlp/pages/Observatoire/Pages/DirOHG/Geomatique/PlansSecteurMap.htm	P
	Interactive map of Walloon sector plan	Interactive map; GIS Serveur	Walloon Region	Portail cartographique de la Région wallonne http://carto6.wallonie.be/WebGIS/viewer.htm?APPNAME=PCA&POPUPBLOCKED=true&BOX=1177,08174904939:17998,9999999999:338821,918250951:168000	P
	Land use maps of Wallonia (Cartes d'Occupation du Sol en Wallonie - COSW)	Interactive map; GIS Serveur	Walloon Region	Portail cartographique de la Région wallonne http://cartographie.wallonie.be/NewPortailCarto/index.jsp?page=subMenuCOSW&node=32&snode=321#	P
	"Schémas de Structure Communaux" (SSC) and "Plans Communaux d'Aménagement" (PCA) Land plannings of municipalities	Web; PDF	Walloon Region	Région wallonne http://mrw.wallonie.be/DGATLP/DGATLP/Pages/DAU/Pages/AT/ATLoc01.asp	P
	Interactive map of the PCA localisation PCA (+ and other themes)		Walloon Region	Région wallonne http://carto6.wallonie.be/WebGIS/viewer.htm?APPNAME=PCA&POPUPBLOCKED=true&BOX=1177,08174904939:17998,9999999999:338821,918250951:168000	P
	"Ruimtelijk Structuurplan Vlaanderen" - RSV Land planning of Flanders	PDF	Flemish Region	Région flamande http://rsv.vlaanderen.be/nl/overRsv/downloads.html	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
	"Gemeente Ruimtelijk Structuurplan" - GRS Land planning of Flemish municipalities		Flemish Region	Région flamande http://rsv.vlaanderen.be/nl/overGrs/downloads.html	P
	"Gewestplannen en Ruimtelijke UitvoeringsPlannen" - RUP Regional planning and spatial implementation plans	Interactive map; PDF	Flemish Region	Région flamande http://rsv.vlaanderen.be/nl/ruimtelijkeordening/GWP_RUP.html http://geo-vlaanderen.agiv.be/geo-vlaanderen/gwp/#	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
POPULATION[18]	Total resident population at 1st January of each year	Excel	Belgium	SPF Economie - STATBEL http://www.statbel.fgov.be	P
	Population statistics	Excel	Belgium	SPF Economie - STATBEL http://www.statbel.fgov.be	P
	Population statistics	Excel	Belgium	SPF Economie - STATBEL http://www.statbel.fgov.be	P
	Private households by household size and number of collective households	Excel	Belgium	SPF Economie - STATBEL http://www.statbel.fgov.be	P
	Family nuclei according to the number of children	Excel	Belgium	SPF Economie - STATBEL http://www.statbel.fgov.be	P
	Population forecasts 2007-2060	Excel	Belgium	Bureau fédéral du Plan - BFP http://www.statbel.fgov.be	P
	Income tax revenues - 2007, incomes 2006	Excel	Belgium	SPF Economie - STATBEL http://www.statbel.fgov.be	P
	Survey on household budget	Excel	Belgium	SPF Economie - STATBEL, via FISC http://www.statbel.fgov.be	P
	Rate of household vehicule ownership	Excel	Belgium	SPF Economie - STATBEL http://www.statbel.fgov.be	P
	Living conditions and welfare indicators (EU-SILC)	Excel; Web; PC-Axis; CSV	Europe	Eurostat http://epp.eurostat.ec.europa.eu/portal/page/portal/living_conditions_and_social_protection/data/database	P
	Health and poverty indicators (EU-SILC)	Excel	Brussels	Brussels-Capital Health and Social Observatory http://www.observatbru.be/documents/indicateurs.xml?lang	P

[18] Free STATBEL statistics exist at communal level. Statistics also exist at the "statistical sector" level, but then they are not free and not directly available from Internet

				=en	

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
EMPLOYMENT	Labour market and social protection datawarehouse	Interactive table; CSV	Belgium	Banque Carrefour de la Sécurité Sociale - BCSS http://www.ksz-bcss.fgov.be/fr/statistiques/stats_home.htm	P
	Labour Force Survey (LFS) *(Enquête Force du Travail*)	PDF	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/modules/pressrelease/statistiques/marche_du_travail_et_conditions_de_vie/enquete_force_de_travail_communiques_et_dossiers_2008_-_2009.jsp	P
	Results of the Labour Force Survey (activity rate, employment, education, employees, hours of work, household jobs, ...)	Excel	Belgium	Eurostat http://epp.eurostat.ec.europa.eu/portal/page/portal/employment_unemployment_lfs/data/database	P
	Evolution of the labour market from 1986 to 2006	PDF	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/binaries/pr094_fr%5B1%5D_tcm326-65415.pdf	P
	Labour force	Excel	Belgium	Région de Bruxelles-Capitale http://www.bruxelles.irisnet.be/fr/entreprises/maison/statistiques/analyses_et_statistiques/donnees_statistiques_thematiques/population_active.shtml	P
	Gross fixed capital formation	Web	Walloon Region; Belgium	Région wallonne - Institut wallon de l'évaluation, de la prospective et de la statistique http://statistiques.wallonie.be/dyn/14/article1.ihtml?ID_SITE=14&ID_CATEGORIE=247&ID_ARTICLE=421&NOM_CATEGORIE=0BAH&CAT=1&MODE=MAIN	P
	Firms by municipalities	Excel; Web	Walloon Region	CAP Ruralité (FSAGx) - Cellule d'Analyse et de Prospective en matière de ruralité de la Région Wallonne http://capru.fsagx.ac.be/communes-wallonnes-en-chiffres?view=all&i_ancre=23&depth=2&categorie_1=Economie+et+revenu&categorie_2=Entreprises	P

	Number of workers, number of firms and size of firms in each district	Excel	Belgium	Office National de Sécurité Sociale - ONSS http://www.onss.fgov.be/fr/content/statistics/publications/place.html	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
EMPLOYMENT (continued)	Firm bankruptcies	Excel; PDF; CSV	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/statistiques/chiffres/economie/entreprises/faillites/ans/index.jsp	P
	"Top 150000" of firms	Web	Belgium	TRENDS Tendances (private company) http://www.trendstop.be/	P
	Industry – Leading Walloon companies	Interactive map	Walloon Region	Institut de Conseil et d'Etudes en Développement Durable http://www.icedd.be/atlasenergie/pages/mconfr01.htm	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
REAL ESTATE	Real estate sales	Excel	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/binaries/rest2009_fr%5B1%5D_tcm326-34187.xls	P
	Past evolution in real estate sales and prices (€ /m²)	Excel	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/modules/publications/statistiques/economie/ventes_de_biens_immobiliers.jsp	P
	Renting prices by municipality		Belgium	SPF Economie - STATBEL (Belgian national statistics institute)	P
	Cadastral statistics	Excel	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/binaries/cad2008_fr_tcm326-34175.xls	P
	"Plan de Localisation Informatique" - PLI	Web	Walloon Region	Région wallonne - DGATLP - Division de l'Observatoire de l'Habitat http://mrw.wallonie.be/dgatlp/dgatlp/Pages/Observatoire/Pages/DirOHG/Geomatique/PLI.htm	P
	Building licences *(permis de bâtir)* by commune and year	Excel	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/binaries/bpe_year_fr%5B1%5D_tcm326-63050.xls	P
	Building licences detailed by commune and month	Excel	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/binaries/bpe_month_fr%5B1%5D_tcm326-55958.xls	P
MOBILITY	Origin-destination matrices for home-to-work and home-to-school trips		Belgium	SPF Economie -STATBEL (Belgian national statistics institute)	P
	Annual and monthly statistics on the vehicles [19]	Excel	Belgium	SPF Mobilité - STATBEL - Service d'Immatriculation de Véhicules	P

[19] Some STATBEL data are available online only for 2008 but should also be available for earlier years.

				http://statbel.fgov.be/fr/modules/publications/statistiques/circulation_et_transport/evolution_du_parc_de_vehicules_2009.jsp	

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
	Monthly statistics on plate registrations [20]	Excel	Belgium	SPF Mobilité – STATBEL – Service d'Immatriculation de Véhicules http://statbel.fgov.be/fr/modules/publications/statistiques/circulation_et_transport/evolution_du_parc_de_vehicules_2009.jsp	P
	National mobility survey – MOBEL	PDF; Excel; Web	Belgium	SPF Economie STATBEL ; Groupe de Recherche sur les Transports (GRT) - Facultés Universitaires Notre-Dame de la Paix (FUNDP) http://www.belspo.be/belspo/home/publ/pub_ostc/mobil/rapp18syn_fr.pdf http://www.mobel.be	P
	BELgian Daily Mobility – BELDAM		Belgium	SPF Mobilité http://www.belspo.be/belspo/fedra/agora/agJJ150_fr.pdf	P
	Count of vehicules on highways and national roads, by sense and vehicle type in Walloon Region	PDF; Excel	Walloon Region	MET – Région wallonne http://routes.wallonie.be/struct.jsp?chap=2&page=5	P
	Count of vehicules on highways and national roads, by sense and vehicle type in Flemish Region	PDF; Excel	Flanders	Agentschap Wegen en Verkeer – Region flammande http://www.wegen.vlaanderen.be/documenten/tellingen/	P
	Count of vehicules on principal roads, by sense and vehicle type in Brussels-Capital Region	Excel	Brussels-Capital Region	Bruxelles-Mobilité – Région de Bruxelles-Capitale http://www.bruxellesmobilite.irisnet.be	
	Indicators on transport and mobility of the "BFP"	Excel; Web	Belgium	Bureau fédéral du Plan – BFP http://www.plan.be/databases/database_det.php?lang=fr&TM=30&IS=60&DB=TRANSP&ID=14	P

[20] Some STATBEL data are available online only for 2008 but should also be available for earlier years.

	Travel behaviour survey in Flanders	PDF	Flemish Region	Vlaamse Overheid - Mobiel Vlaanderen – epartment Mobiliteit en Openbare Werken http://www.mobielvlaanderen.be/ovg/ovg1.php?a=19&nav=3	P
	Mobility statistics of Flanders	Excel	Belgium	Vlaamse Regering – Studiedienst http://www4.vlaanderen.be/dar/svr/Cijfers/Pages/Excel.aspx	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
ACCESSIBILITIES	Zone-to-zone accessibilities expressed in "generalised times"		RER area[21]	STRATEC	C
ENVIRONMENT	Results of surveys on noise pollution due to traffic in Brussels	PDF	Brussels-Capital Region	Bruxelles Environnement - Région de Bruxelles-Capitale http://documentation.bruxellesenvironnement.be/documents/Bru_1.PDF	P
	Map of exposure to road traffic noise in Brussels	PDF	Brussels-Capital Region	Bruxelles Environnement - Région de Bruxelles-Capitale http://www.bruxellesenvironnement.be/uploadedImages/Site/Particuliers/Theme_-_Bruit/bruittrafic2001.jpg?langtype=2060	P
	Map of acoustic areas by Lden - Road traffic noise	PDF	Brussels-Capital Region	Bruxelles Environnement - Région de Bruxelles-Capitale http://www.bruxellesenvironnement.be/uploadedImages/Site/Particuliers/Theme_Bruit/zonageacoustique2001.jpg?langtype=2060	P
	Map with the location of road and rail most severe acoustic problems	PDF	Brussels-Capital Region	Bruxelles Environnement - Région de Bruxelles-Capitale http://www.bruxellesenvironnement.be/uploadedImages/Site/Particuliers/Theme_-_Bruit/carte_points_noirs_fr.jpg?langtype=2060	P
	Map of acoustic areas by Lden - Road, rail and air traffic noise	PDF	Flemish Region	LNE - Departement Leefmilieu, Natuur en Energie http://www.lne.be/themas/hinder-en-risicos/geluidshinder/beleid/eu-richtlijn/goedgekeurde-geluidskaarten/goedgekeurde-geluidskaarten	P

[21] area which will be served by the future RER), i.e. the Brussels agglomeration and the surrounding suburban areas

	Interactive map of exposure to road traffic noise	Interactive map	Flemish Region	Agentschap Wegen en Verkeer http://www.wegen.vlaanderen.be/documenten/geluidskaarten/	P
	Noise exposure study for the airport of Zaventem	Pdf	Brussels Airport Zone	Brussels Airport http://www.brusselsairport.be/fr/community/geluid/geluidhinder/geluidscontouren1#	P
	Air quality - Ozone concentrations	Excel	Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/binaries/ozon_tcm326-34200.xls	P
	Atmospheric concentrations and emissions		Belgium	SPF Economie - STATBEL http://statbel.fgov.be/fr/statistiques/chiffres/environnement/air/qualite/index.jsp	P

Category	Name and description	Datatype	Perimeter	Data owner/ Source Availability	Right of use
	Zones in relation to EU air quality thresholds	GIS - Vector	Europe	European Environment Agency - EEA http://dataservice.eea.europa.eu/dataservice/metadetails.asp?id=1095	P
	Environment and energy indicators	Excel; Web; PC-Axis; CSV	Europe	Eurostat http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database	P
SUMMARY	Summary "sheet" of municipalities : network, employment, real estate, mobility, environment,…	Web; PDF	Walloon Region	Région wallonne - Portail Environnement de Wallonie http://environnement.wallonie.be/fiches_enviro/index.htm	P

7.2 Available data for Zurich case study

Category	Name and description	Entity/Accuracy	Right of use
PERIMETERS	**GG25 municipal borders***: Contains the geometry and location of the administrative boundaries defining the municipalities. Municipalities are a very important spatial unit because a lot of data is aggregated to that level	Municipality	P
NETWORKS	**Spatial development plan: Transportation**: Political strategy of spatial development in the canton of Zurich concerning aspects of transportation.	Pixel	P
	Zones of model for individual transportation*: Describes geometry and location of the traffic analysis zones in the aggregate individual transport model. Most important attribute further describing the zones is their accessibility.	Traffic analysis zones	P
	KVM-ZH*: Road network, velocities and distances between traffic analysis zones.	Traffic analysis zones	
	Complete public transportation model: Describes geometry and location of the traffic analysis zones in the aggregate public transport model. Most important attribute further describing the zones is their accessibility.	Traffic analysis zones	
RELIEF	**DHM 25***: Digitial terrain model. In a grid of 25m mesh size each point is assigned an x-, y-,and z-value.	Mesh size 25m / 1.5-3 m	
LAND USE AND LAND COVER	**Spatial development plan: settlement and landscape**: Political strategy of spatial development in the canton of Zurich concerning settlement and landscape aspects.	Pixel	P
	Vector 25*: VECTOR25 is the digital landscape model of Switzerland that is based on the topographic map 1:25000. VECTOR25 includes natural and anthropogenic objects and is most suitable for usage in GIS. The topic layers are: road network, railway network, other transport infrastructure, hydrological network, primary areas, buildings, bushes and trees, facilities and single objects.	3-8 m	P
POPULATION	**Data describing hectar gridcells***: For example vacancy rates, number of single family houses, occupied housing units, total number of housing units, vacant housing units, income of natural persons Entity/Accuracy:	Municipality	P
	Synthetic population: The synthetic population includes roughly 7 millions of agents. Representing the swiss population.	Household	P
	Tax level legal / natural person: Indexes of the tax level in municipalities in percent to a baseyear.	Municipality	P
	Net income and expenses of natural persons: Data describes budget and expenses of single households. Surveyed attributes are: telephone subscription, secondary residence, child care, age of person, employment, aggregate expenses, equipment of households. Data from sampling survey. Entity/Accuracy:	Municipality	P

Category	Name and description	Entity/Accuracy	Right of use
	Mean of population per municipality (P): Total of population in municipalities for the years 2000 until 2005.	Municipality	
	Probabilities of households relocation: Probabilities depending on households attributes.		P
	Population forecast Canton of Zurich*: Number of people living in regions, wards and boroughs of canton Zurich. The age structure of the population is also available. Furthermore migration data is at hand.	Regions, wards, boroughs of Zurich	

7.3 Available data for Paris case study

Category	Name and description	Entity/Accuracy	Right of use
PERIMETERS	**GIS representation (ArcGis), administrative limits**: A Paris map containing the different administrative limits	departement, municipality	R
NETWORKS	**Davisum database (matrices)**: OD (flows) matrices for TC and VP.	606*606 Zones	R
	Davisum database (matrices): Travel time matrices for TC and VP.	606*606 Zones	R
	Zoning (TAZ) (R): 606 centroïds.	606 points	R
	Davisum database (Network, TC): A public transport network composed of 4000 nodes and 5000 links for the whole Paris region. About 400 services (trips) during rush hours (2 hours) for RER, train and metro. Less precise data for buses (within zone more than between zones; see DRE).	Zones (606)	R
	Davisum database (Network, VP): A road network composed of 5000 nodes and 16 000 links for the whole Paris region.	Zones (606)	R
	GIS representation (ArcGis) networks: A Paris map containing the different transport networks. The nodes and links are represented as layers on the Ile de France map.		R
LAND USE AND LAND COVER	**ZAC and ZAE data**: Contains information on creations of ZAE and ZAC (can be used to model municipality choices). Should be associated to information on actual projects (in order to estimate other stake holders behaviour).		R
	Land use (MOS) (R): At ilot level, it contains the development type of each cell in 83 posts. It is available in years 1982, 1987, 1990, 1994, 1999 and 2008	Ilot MOS (520 000)	R
POPULATION	**Dwelling survey (Enquête Logement)**: It is the principal survey of housing; it contains more than 47000 dwellings. Data on moves (actual and contemplated), housing quality, owner income, expenses associated to the house and others.	Municipality	P
	Family Budget Survey: Quality testing of synthetic data (complement to location choice). This table is a French survey; it contains all elements related to the household budget and expenses.	Municipality	P
	Dwelling survey (Enquête Logement): It is the principal survey of housing; it contains more than 47000 dwellings. Data on moves (actual and contemplated), housing quality, owner income, expenses associated to the house and others.	Municipality	P
	population census: An exhaustive census of the population in the Paris region, located at parcel (IRIS) and municipality levels. It contains demographic, spatial, social and economic variables.	Municipality	P
EMPLOYMENT	**Regional Employment Survey (ERE) 1997 & 2001**: Exhaustive database of all firms and plants: detailed sector, number of employees and location (municipality level for all firms and plants in 1997, parcel level for some plants in 2001).	Municipality	R
REAL ESTATE	**Dwelling and offices prices (Cote Callon)**: This table contains prices of m² of a real estate in some cities of the whole France.	Municipality	P
	Notaries' Database: It contains aggregate data on transactions (nb transactions & average prices) of housing sales in Paris region.	Municipality	P