THE THIRD DIMENSION IN URBAN GEOGRAPHY[1]

The urban volume approach

Eric Koomen, Piet Rietveld and Fernando Bacao

Abstract: A new methodology is presented that measures density in urban systems. By combining highly detailed height measurements with, amongst others, topographical data we are able to quantify urban volume. This new approach is tested in two separate case studies that respectively relate to the temporal and spatial dimension of the urban environment. In the first study the growth of the city of Amsterdam over the past century is studied. The urban volume indicator is used to visualise and quantify the urban extension and intensification process. To critically analyse the spatio-temporal development of Amsterdam the self-organizing map approach is applied. Special attention is given to highlighting any signs of recent polynuclear development. The second case study compares the building height frequency and spatial distribution of high-density zones in the four major Dutch cities. Additionally, the presence of built-up areas and the actual urban volume values are simultaneously explained with a Heckman selection model.

Key words: Urban morphology; urban volume; density; indicator; self-organizing maps.

1. INTRODUCTION

The urban landscape is continuously changing. Suburbanisation and urban sprawl have altered the classical monocentric city and given rise to new polycentric urban forms that have, for example, been described as edge-cities (Garreau, 1992), network cities (Batten, 1995), corridors (Priemus, 2001), decentered cities (Stern and Marsh, 1997) and even edgeless cities (Lang and LeFurgy, 2003). It is important to note that different scale levels are considered in these studies on the polycentricity of urban form. These levels range from individual cities, through urban regions to international macro-regions (Dieleman and Faludi, 1998). This inconsistency in the

[1] The definitive, peer-reviewed and edited version of this article will be pubished in Environment and Planning B: Planning and Design.

applied scale levels obscures the ongoing debate on urban form. The discussion is further complicated by the fact that changes in the urban system are increasingly caused by the interdependency of different scale levels (Van der Laan, 1998). Central in all different descriptions of urban form, however, is the notion that the original city centres are losing their importance. Although the decline of traditional city centres in Europe does not nearly resemble the many North American examples, cities here also show a growing importance of its subcentres (e.g. Gaschet, 2002; Martori i Cañas *et al.*, 2002).

The Dutch Randstad area, the constellation of the four biggest cities in the western part of the country, is generally considered to be an interdependent network city (Batten, 1995; Van der Burg and Van Oort, 2001) in which the various urban subcentres are functionally related. Empirical evidence for this claim is, however, hard to find (Ritsema van Eck *et al.*, 2006). The major Dutch cities, in fact, show signs of various opposing processes occurring simultaneously; inner-city redevelopment coincides with ongoing suburbanisation and, at the local scale, the intensification of urban functions is alternated with the demolition of high-rise apartment blocks to provide room for new single family dwellings. All these processes lead to a continuous reshaping of the urban areas and, furthermore, influence the relations with the surrounding suburban and rural areas. The formulation of effective spatial policies related to, for example, open-space preservation, mobility growth limitation and urban regeneration is hampered by a lack of knowledge on the relative importance of the forces that shape urban areas. A thorough understanding of current urban processes is a first step in drafting such policies.

Urban development often leads to changes in the intensity in which the already existing urban fabric is used and is thus difficult to trace with classical geographical analysis that typically focuses on lateral, two-dimensional urban extensions. Typical examples of this type of research compare two subsequent land-use maps and analyse the growth in urban areas, without studying changes in the intensity of urban land use (e.g. EEA, 2006). This omission can, generally, be ascribed to the fact that land-use intensity is difficult to assess. Recent studies on urban density (e.g. Longley and Mesev, 2002; Batty *et al.*, 2004) have applied detailed individual address and postcode point data to characterise intensities in land use. However, as Batty *et al.* (2004) indicate themselves, such approaches fail to incorporate the importance of the third (height) dimension in urban analysis. Without additional data (such as applied by Maat and Harts, 2001) these studies do not recognise the importance of large, tall buildings that characterise high-density zones and that are extremely important in terms of their number of inhabitants, employees or visual dominance. The analysis of the third dimension of urban morphology is scarce however, mainly due to limited data availability. Incidental examples reflect a painstaking data-collection process (e.g. Frenkel, 2004).

This paper presents the results of a detailed analysis of the third dimension of current Dutch cities that makes use of the recently released extremely detailed height information of the Netherlands. This new data set allows for the relatively easy creation of an urban volume layer that effectively captures urban morphology at the level of individual cities. Building volume is taken here as a proxy for urban density and, to our mind, offers the opportunity to properly include the third dimension in studies of urban geography as was previously advocated by Batty (2000). The approach has the advantage of closely resembling the human perception of urban density (Fisher-Gewirtzman *et al.*, 2003) and its results are therefore easily interpreted. To show the potential of the newly developed urban volume methodology for analysing urban form we apply it in two separate case studies that have a temporal and a spatial dimension respectively. Time is the crucial element in the study that deals with historic development of urban density in the city of Amsterdam in the 1900-2000 period. An important element in the analysis of the temporal dimension is the application of the self-organizing map method to help distinguish spatio-temporal relations in our rich data sets. The spatial dimension is the subject of a second application that compares and explains the building height frequency and spatial distribution of high-density zones in the four major Dutch cities. Additionally, the presence of built-up areas and the actual urban volume values are simultaneously explained with a Heckman selection model. In both case studies we seek evidence for polynuclear development at the level of major individual cities.

2. URBAN VOLUME METHODOLOGY

The urban volume indicator that we apply in our analysis is based on the combination of height and topographic data. Figure 1 gives an overview of the methodology that was applied to come to an urban volume layer. This section introduces the data sets that were used and discusses the most important steps in creating the urban volume layer. A full account of the data sets and methodology that were used can be found in: Koomen *et al.* (2004) and Kaufholz (2004).

A crucial data set in this analysis is the newly developed Dutch national elevation data set (*Actueel Hoogtebestand Nederland*) that has become available in 2003. This highly detailed data set was collected over the preceding seven years under the supervision of the Survey Department and is based on laseraltimetric measurements. It has a height precision of about 15 centimetres standard deviation per point and an average point density of 1 point per 16 m^2 or better (Oude Elberink *et al.*, 2003). The provided elevation data has enough spatial detail to distinguish individual houses and gives a detailed account of their heights. Huising and Gomes Pereira (1998) offer a full discussion of all possible errors relating to the laser system, the process of measuring or the target surface. These errors range from 5 to 200 centimetres, but are for a large part corrected before the data is distributed. The remaining inaccuracy does not hamper our analysis, as we are interested in height-differences of several metres. For this study we use a rasterised version of the original point data set with a 5 by 5 metres pixel resolution that provides an average value of all height points within the grid cell. For the rare cases that a grid cell is lacking information (e.g. in the case of a missing overlap in the original data strips) a combination of mathematical techniques is used to fill in the gaps (Vosselman and Maas, 2001). Only the larger water bodies completely lack height information because of their reflecting characteristics. These do not pose a problem in our analysis because we are focussing on the built-up areas.

To select only the heights of buildings an overlay is made with a thematic layer that contains detailed information on the topography (top10vector, see TDN, 1998). This layer allows for the distinction of residential and non-residential building blocks and non built-up land. The latter is important to help reconstruct surface level heights from the original height data set. By subtracting the surface height from the original heights that referred to the national datum level (0 metre or mean sea level) we arrive at the actual building heights. In a second step the occasionally missing extreme high height values are manually added from an additional web source (skyscrapers.com). The grid cell values are then multiplied by their surface area (25 m^2) in order to represent a volume-per-pixel of buildings. This high resolution provides an extremely detailed, but also very heterogeneous and dispersed account of urban volume. To allow for a more straightforward interpretation of the urban volume indicator and, furthermore, speed up subsequent statistical analyses we choose to aggregate the urban volume values to a 25 metres grid in which the total volume of the grid cells that make up these larger units is retained. By using this total, aggregated value we preserve the underlying detailed observations based on the original topographical maps and height information. Small solitary buildings, for example, remain represented in the 25 by 25 metres grid; we only loose their exact position.

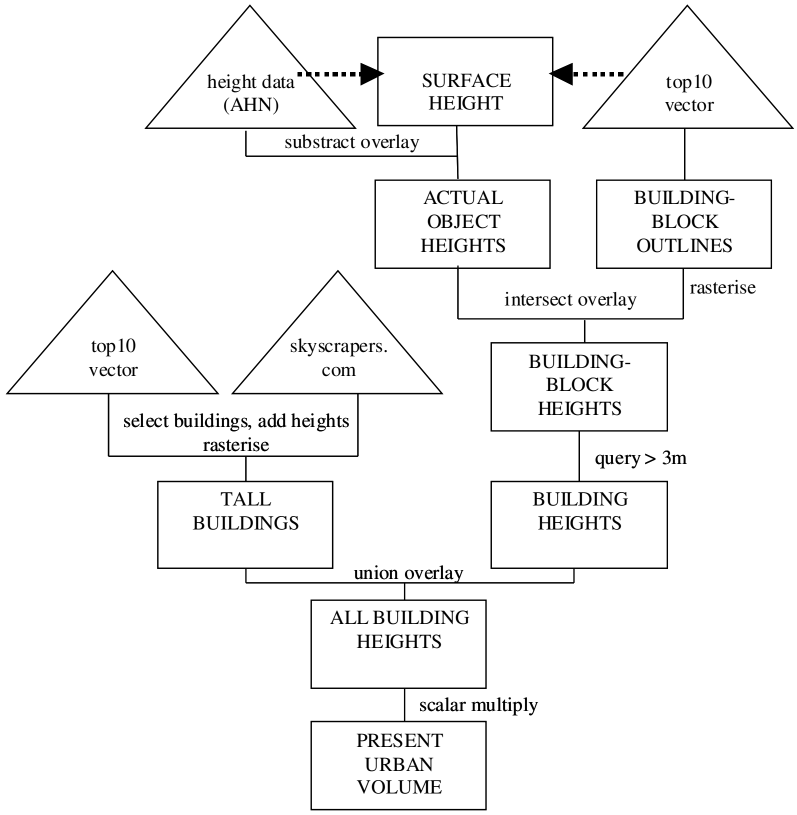

Figure 1. Cartographic model depicting the basic methodology for constructing the urban volume layer.

3. SPATIO-TEMPORAL ANALYSIS OF THE AMSTERDAM URBAN VOLUME 1900-2000

The capital of the Netherlands provides an especially interesting case study area because its urban landscape has changed significantly in the past century. After almost two centuries of stagnation the city started to grow rapidly in the last part of the 19th century, reflecting a late catch-up with the industrial revolution. This period is still notable as an urbanisation ring around the historic centre. From the beginning of the 20th century urban expansion has been steered through municipal town planning, initially resulting in the addition of extensive new neighbourhoods to especially the southern and western edges of town and the first major construction north of the central riverfront. After a disruption during the Second World War, extensive garden villages were added to the western and southern limits of town in the 1950-1970 period, following the 1935 general extension plan (Van der Heiden *et al.*, 1991). The latest major additions to the city layout can be found in the southeast, where a completely new neighbourhood for 100,000 inhabitants was constructed, and attached to the western and northern extremities of town. Large-scale inner-city redevelopment started in the 1980's and consists mainly of residential construction on the former maritime and industrial centre on the southeast shore of the riverfront. Compared to other major cities, Amsterdam was slow to start the construction of tall buildings (Kloos and De Maar, 1995). Since the 1980s, however, small concentrations of office building with maximum heights of up to 150 metres have been constructed near the ring road at the western, southern and southeastern parts of town and around a more centrally located railway station. Amsterdam is thus starting to show signs of poly-nuclear development. Our study aims at visualising and quantifying these urban changes by reconstructing the urban volume of 1900-2000 period.

The historic urban volume is reconstructed by combining the original 2000 urban volume data layer with a detailed data set that includes the year of construction of all individual buildings in the municipality of Amsterdam. The latter point data set is combined with a detailed topographical data set that contains building outlines. This enriched polygon map is then rasterised to allow for the recreation of the urban surface in any chosen time period. By selecting, for example, all cells that relate to buildings that were built in or before 1910 we arrive at a reasonable reconstruction of the historic urban area at that time. This reconstructed historic urban area map allows for the extraction of those grid cells in the urban volume data set that were supposedly built-up in 1910. This rough approach has of course some limitations. Old buildings may have been replaced by newer ones in the past 100 years, as the most recent construction year replaces any previous information on an edifice in our data set. These locations will erroneously be left out of the 1910 analysis, introducing an underestimation of the urban volume in that time step. The opposite may also be true: the applied building outline polygons describe urban blocks that are separated by streets or other open spaces. Especially in the old centre these areas may contain many individual buildings. As the oldest building year is assigned to the total block, recent volumes will be incorrectly related to older edifices, introducing an urban volume that might deviate from the original one. Visual inspection of the historic urban area map however shows the old parts of town as more or less continuous surfaces with a relatively homogenous volume distribution, indicating that the described limitations only affect isolated locations. Moreover, the reconstructed urban area maps correspond well with historic maps of the Amsterdam area (e.g. Knol *et al.*, 2003). Since our analysis is mainly meant to explore the possible use of the urban volume indicator we do not consider these drawbacks to be serious constraints to our analysis.

Historic urban volume maps were created for every decade since 1900. A selection of the most crucial time steps is represented in Figure 2. The figure highlights the above-average urban volumes per grid cell by classifying the volume values according to the standard deviations in the 2000 data set. It shows the exceptionally high values in the darkest colours. The time series reflects the continuous growth of the city in all directions following the large-scale pre-war (1940) and post-war (1970) extensions. It furthermore highlights the recent, erratic spread of high intensity zones throughout the city. The 2000 urban volume map shows an abundance of high volume zones in almost all neighbourhoods of the city, clearly indicating a deviation from the original mono-centric form.

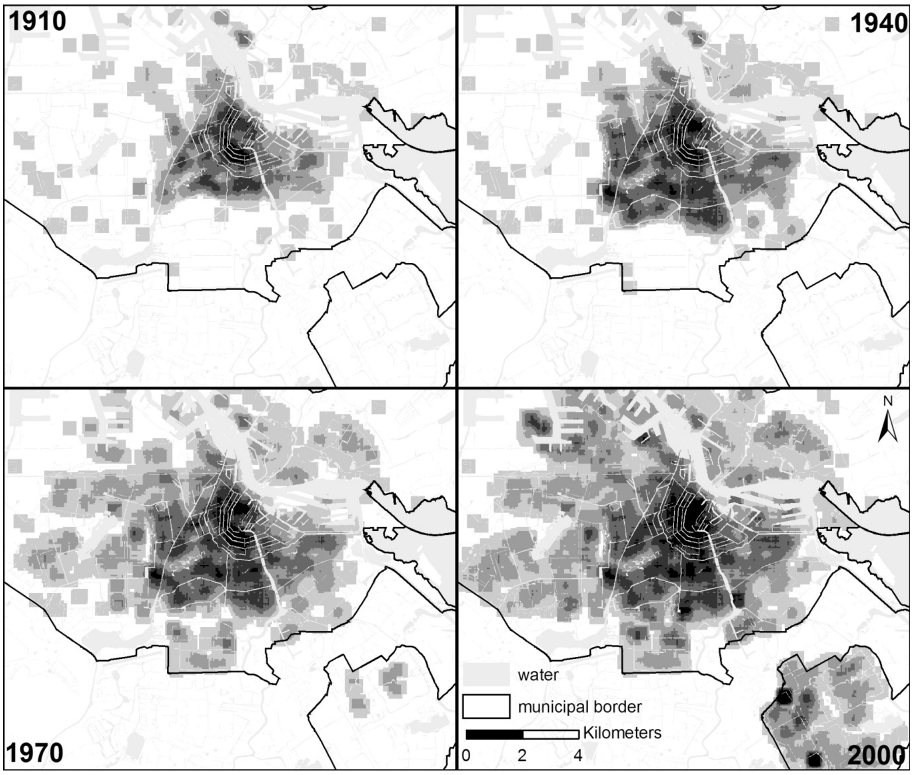

Figure 2. Reconstructed total urban volume ranging from low (grey) to high (black) in the city of Amsterdam for the years 1910-2000. For cartographic clarity the resolution was decreased to 50 metres taking the total urban volume values in a 500 metres moving neighbourhood.

3.1 Applying the self-organizing maps approach

In order to critically analyse the obtained spatio-temporal patterns for the Amsterdam case study the self-organizing map (SOM) approach was applied. This approach can be described as a visualisation and analysis tool for high dimensional data, but it is also applied for clustering (Vesanto and Alhoniemi, 2000), dimensionality reduction, classification, sampling, vector quantization, and data-mining (Kohonen, 2001). The fundamental idea of a SOM is to map the data patterns onto an n-dimensional grid of segments or units. This mapping tries to preserve topological relations, i.e., patterns that are close in the input space will be mapped to segments that are close in the output space, and vice-versa. Each segment, being an input layer segment, has as many weights or coefficients as the input patterns, and can be regarded as a vector in the same space as the patterns. When training or using a SOM with a given input pattern, the distance is calculated between that pattern and every segment in the network. The segment that is closest to the winning segment is selected, and then the pattern is mapped onto that segment. If the SOM has been trained successfully, the patterns that are close in the input space will be mapped to segments that are close (or the same) in the output space, and vice-versa. Thus, SOM is 'topology preserving' in the sense that (as far as possible) neighbourhoods are preserved through the mapping process.

Before training, the segments may be initialised randomly. Usually the training consists of two parts. During the first part of training, the segments are 'spread out', and pulled towards the general area (in the input space) where they will stay. This is usually called the unfolding phase of training (Kohonen, 2001). After this phase, the general shape of the network in the input space is defined, and we can then proceed to the fine tuning phase, where we will match the segments as close as possible to the input patterns, thus decreasing the possible error. The basic SOM learning algorithm may be described as follows. Let:

w_{ij} be the weight vector associated with a segment positioned at column *i* row *j*;

x_k be the vector associated with pattern k;

d_{ij} be the distance between weight vector w_{ij} and a given pattern;

h be a neighbourhood function described below and

A be the learning rate also described below.

For each input pattern then take the following steps:

1. calculate the distance between the pattern and all segments of the SOM with: $d_{ij} = || x_k - w_{ij} ||$ (this is called the calculation phase);
2. select the nearest segment as winner wwinner: $w_{ij} : d_{ij} = \min(d_{mn})$ (the voting phase);

3. update each segment of the SOM according to the update function: $w_{ij} = w_{ij} + Ah(wwinner, w_{ij}) \| x_k - w_{ij} \|$ (the updating phase);
4. Repeat the steps 1) to 3), and update the learning parameters, until a certain stopping criterion is met.

This algorithm can be applied to a SOM with any dimension. The learning rate A must converge to 0 in order to guarantee convergence and stability for the SOM (Kohonen, 2001). The decrease from the initial value of this parameter to 0 is usually done linearly, but any function may be used. The neighbourhood function h assumes values in [0,1], and is a function of the position of two segments (a winner segment, and another segment), and radius. It is large for segments that are close in the output space, and small (or 0) for segments far away. Usually, it is a function that has a maximum at the centre, monotonically decreases up to a radius r (sometimes called the neighbourhood radius) and is zero from there onwards. For the sake of simplicity, this radius is sometimes omitted as an explicit parameter. The two most common neighbourhood functions are the bell-shaped (Gaussian-like) and the square (or bubble), in both cases, we force $r \rightarrow 0$ during training to guarantee convergence and stability. The update of both, the learning rate and the neighbourhood radius, parameters may be done after each training pattern is processed or after the whole training set is processed.

For the Amsterdam case study a relatively large SOM with 60 segments was set up to isolate the areas of growth in volume with a certain degree of precision. Each input data vector, a 25 metres grid cell, was composed of seven variables: the volume values for the years 1910, 1940, 1970 and 2000 and distances to the ring road, the nearest station and the historic city centre. Table 1 gives an overview of the 23 SOM- segments relating to urban development. The missing segments have an average urban volume of less then 1250 m^3 (equivalent to an average building-height of 2 metres in the 25 by 25 metres cell) and are thus considered not to be important for our study. The segments characterise homogenous groups of grid cells that share a common development history and relative location to key features of the city. The analysis clearly distinguishes the subsequent development phases. The first seven rows for example refer to the last stage of urban development in the 1970-2000 period. The low-density developments of segments 29 and 30 can be found far from the original city centre; these correspond with the recent construction of low-density single-family dwellings at the western extremities of town. The high-density developments near the stations of segment 42 represent the recent construction of extremely high office buildings. The 1940-1970 period shows urban developments at 4 to 7 kilometres from the city centre. Several low-density developments (segments 39 and 40) are located near the ring road. The 1910-1940 extensions can be found at an average distance of 2 to 3 kilometres from the centre, with the highest densities near the stations (segment 54). The oldest parts of town are described in the last four segments, with the highest densities in segment 60 within 1.5 kilometres from the Dam Square where the city was founded.

Some of the most notable SOM segments are mapped in Figure 3. This selection consists of the highest densities per building period, each reflecting the different characteristics of the relative high rise developments in that period. The oldest developments (segment 60) only have a medium density but cover an extensive area. Isolated areas of higher density of the 1910-1940 and 1940-1970 period can be found within (segment 54) and outside the ring road (segment 48) respectively. By far the highest densities date back to the last building phase and are found near the stations (segment 42). In the latter part of this paper we will further analyse the relative importance of the spatial factors included in the SOM for explaining the currently observed urban volume patterns in Amsterdam and three other major cities.

Table 1. Selection of SOM analysis results relating to the historic development of Amsterdam

Segment	Volume 2000	Volume 1970	Volume 1940	Volume 1910	Distance to centre	Distance to ringroad	Distance to nearest station
[number]	[m^3]	[m^3]	[m^3]	[m^3]	[m]	[m]	[m]
38	**1769**	118	2	1	4342	**699**	1034
50	**2114**	157	67	13	2156	2506	1531
29	**3064**	0	0	0	**7188**	2429	1336
30	**4449**	1	0	0	**9071**	4003	2019
35	**4922**	5	3	2	3489	1517	1619
36	**10213**	11	6	4	5013	1910	1408
42	**34249**	0	0	0	5901	2077	**862**
39	1533	**1528**	1	1	4361	**918**	**999**
34	2265	**2259**	25	20	6669	2955	2485
40	3017	**3011**	2	2	4250	**924**	1307
41	4650	**4639**	5	5	4387	1385	1424
47	7755	**7714**	1	0	4953	1560	1353
48	15223	**15187**	3	0	4704	1305	1477
45	1792	1785	**1784**	7	3149	1142	1263
51	2400	2368	**2365**	28	2262	2245	1338
46	3095	3093	**3093**	6	3009	1167	1230
52	4291	4288	**4288**	10	2785	1398	1268
53	6326	6321	**6321**	15	2410	1782	1333
54	15324	15324	**15324**	9	1832	2349	**900**
57	2326	2269	2263	**2229**	2266	2080	1457
58	3753	3739	3732	**3714**	1968	2197	1477
59	5629	5625	5622	**5618**	1727	2411	1513
60	10037	10034	10034	**10033**	**1459**	2740	1478

Note: characteristic results are indicated in bold and are discussed in the text.

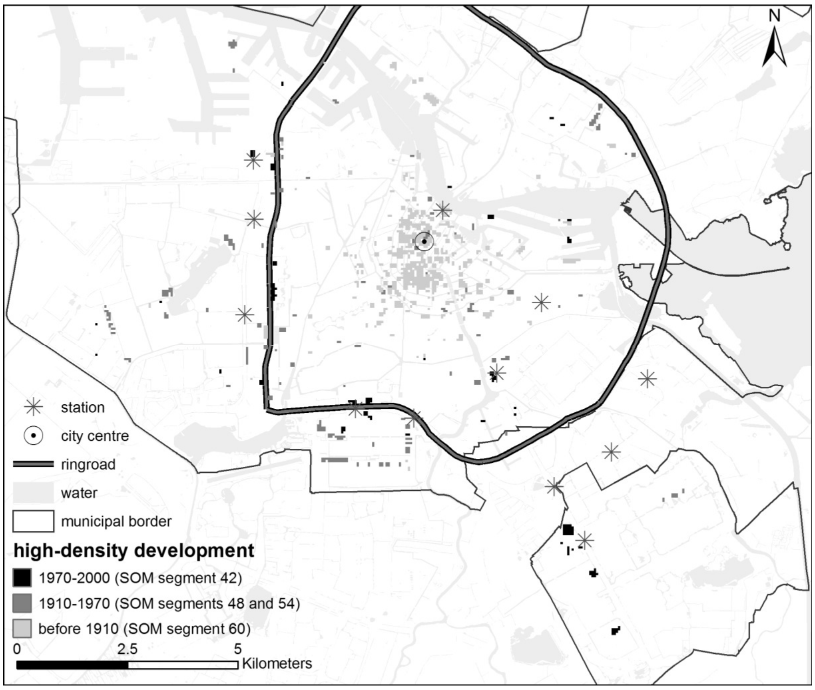

Figure 3. Selected SOM segments reflecting high-density developments in different time periods. For cartographic clarity the grid cell resolution was decreased to 50 metres.

4. SPATIAL COMPARISON OF THE FOUR MAJOR DUTCH CITIES

The second case study in our analysis aims at comparing and explaining the urban density of the four largest Dutch cities: Amsterdam, Rotterdam, The Hague and Utrecht. The urban volume approach is used here to characterise the cities in terms of: 1) their general appearance, 2) their building height-distribution and 3) their spatial, urban density patterns. After this characterisation we will make an attempt at explaining the observed patterns.

The four selected cities are part of the metropolitan Randstad region in the west of the Netherlands, but differ in their history and layout. Amsterdam is the largest city of the country in terms of its number of inhabitants and has a large well-preserved historic centre. Rotterdam covers the largest surface area and has the largest built-up area, mainly as a result of its vast harbour area. Its centre was heavily bombed in the Second World War and it was almost completely rebuilt in the 1950's. The Hague is a relatively new city that houses the Government, most ministry buildings and a large number of offices. Utrecht is the smallest of the four cities, both in terms of its population and size. It is the only city that dates back to before 1000AD and it still retains part of its medieval building history. Table 2 summarizes the key statistics for the selected four cities derived from both the Dutch Central Bureau of Statistics (CBS, 2005) and our own urban volume approach. The latter provides a more detailed account of the area that is actually covered by residential and non-residential buildings then the CBS built-up area statistic that also contains land used for infrastructure and recreation and other building related functions such as gardens and pavements. The buildings in all four cities cover less then a quarter of the total municipal land area. The Hague has the highest building area density (22% of the municipal land surface), Amsterdam the lowest (12%). Interestingly enough the population density per building area is highest in Amsterdam with close to 40,000 inhabitants per km^2 building area. This more intensive use of space is also indicated by the relatively high average building height in Amsterdam. Rotterdam is a special case, since a many of its buildings are voluminous edifices related to commercial functions in the harbour. Its population density is therefore lower, but its average building height is higher than in the other cities.

Table 2. Key statistics for the four major Dutch cities

	Land area	Built-up area	Building area	Population	Population density per land area	Population density per building area	Urban volume	Average building height
	km^2	km^2	km^2	persons	person/km^2	person/km^2	km^3	m
Amsterdam	165	76	19	731,288	4429	37,569	0.216	11.1
Rotterdam	209	102	26	592,673	2841	22,497	0.299	11.4
Den Haag	68	39	15	441,094	6494	28,972	0.162	10.6
Utrecht	61	30	9	233,667	3804	25,991	0.092	10.2

Note: the built-up area includes all types of land use related to residential and non-residential buildings, such as gardens, pavements *et cetera*. The building area only refers to the area actually covered by those buildings.

Source: CBS (2005) for total population, land and built-up area per municipality, other statistics are own calculations based on the methodology described in the text.

4.1 Building height frequency

In order to take a closer look at the base data at hand we first analyse the frequency distribution of the building height data set. By plotting the frequency of all observed building heights for all cities in one graph we can visually compare their full height ranges and related building height distributions, see Figure 4. Please note that the observations relate to the 5 by 5 metres pixels of the original data set; they are thus smaller then individual buildings. Height values have been truncated to full metres to facilitate faster calculations. This figure provides an initial characterisation of the three-dimensional appearance of the cities. In fact, the frequency distribution offers a unique three-dimensional fingerprint for each city. Most striking about the height distribution of the four major cities is that they have the same basic shape. The most common building height is around 7-9 metres, indicating that single and two story houses are less common than houses with three levels. Higher buildings occur less frequently with increasing height[2]. Only Amsterdam and Rotterdam have a considerable number of high buildings, as is shown in the somewhat erratic tail at the right hand side. The maximum building heights are around 100 metres for Utrecht and The Hague and about 150 metres for Amsterdam and Rotterdam. The difference in size of the cities is also apparent: Rotterdam is the biggest city in terms of its building heights, Utrecht the smallest.

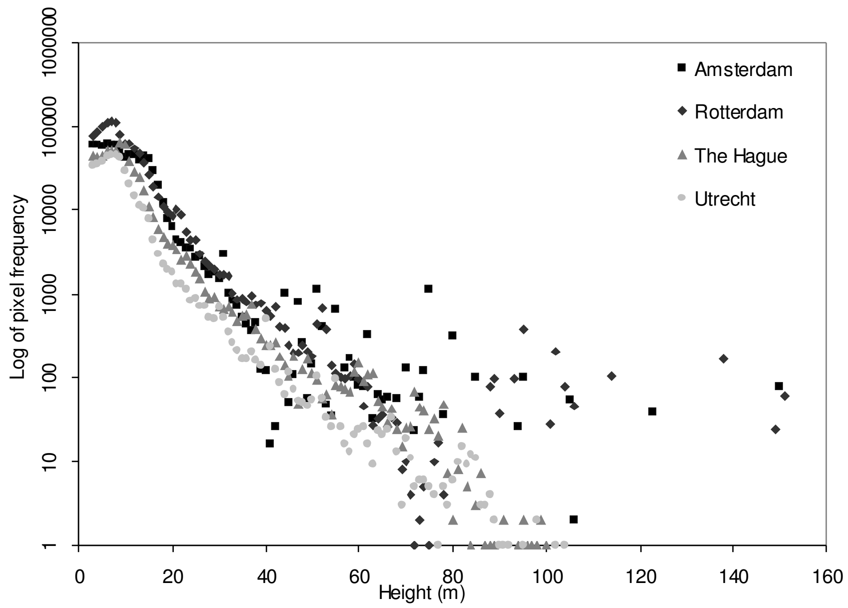

Figure 4. Building height frequency distribution of the four major Dutch cities.

[2] The conspicuously low number of buildings with a height of approximately 40 metres in Amsterdam is probably caused by the processing techniques of the data suppliers. This inconsistency could, apparently, not be fully corrected by the manual addition of missing building heights. Apart from this, no other suspicious values were found.

To facilitate a more quantitative comparison of the city-specific building height distributions we have fitted the observed distributions to a known theoretic one. The gamma distribution was found to provide a very good fit to the frequency distribution of our building height data. This distribution is comparable to the lognormal distribution that was effectively used by Batty (2001) to describe the rank-size population distributions in Great Britain in the past century. The gamma distribution is common in the analysis of, for example, rainfall and flood frequencies (Yue, 2001; Yoo *et al.*, 2005). Table 3 shows the correlation coefficients for the estimated gamma distributions and the related shape (α) and scale (λ) parameters. The fact that the building height distributions of all four cities so closely follow the same mathematical description provides interesting opportunities for further research centred on a number of different research questions. Does this relation also hold true for smaller settlements and other countries? Is it consistent over time? And, perhaps more fundamentally, what processes govern these relatively strict relations? A first attempt at pinpointing some of the relevant factors that explain high building densities is given hereafter. First, however, we look more closely at the spatial patterns of the high-density areas.

Table 3. Estimated gamma distributions of building heights in the four major Dutch cities

	Correlation (R^2)	Shape (α)	Scale (λ)
Amsterdam	0.965	2.506	0.257
Rotterdam	0.988	3.362	0.399
Den Haag	0.954	3.826	0.429
Utrecht	0.979	4.046	0.510

4.2 Density patterns

To visualise the density patterns a filtering operation was applied on the original urban volume layer. By aggregating the original 5 metres resolution to a 250 metres grid using a maximum filter we are able to highlight the areas with highest densities. This approach puts a strong emphasis on the observed maximum values, which is in line with the visual dominance of tall buildings, but it overestimates their actual contribution to the total urban volume. Figure 5 shows the highest density areas per city in black. These areas are defined here as having an urban volume value of one and a halve standard deviations above the city’s average. This representation is, of course, dependent on the (substantial) variance of urban volume values in each city. Thus, the figure does not allow for a comparison between the densities in the cities in an absolute sense, but it provides an interesting view on the local density patterns. These patterns are different for each city. Amsterdam and Rotterdam have the most high-density zones, but the highly erratic pattern of Amsterdam contrasts strongly with the concentrated pattern in Rotterdam. The Hague and Utrecht seem to have a more homogenous spatial distribution of densities and offer less extremely high values. Both cities have a high-density area close to its traditional centre as well as several high-density areas outside that centre. Out of the studied cities only Rotterdam seems to be closest to a classic mono-centric city. Amsterdam offers by far the most varied cityscape.

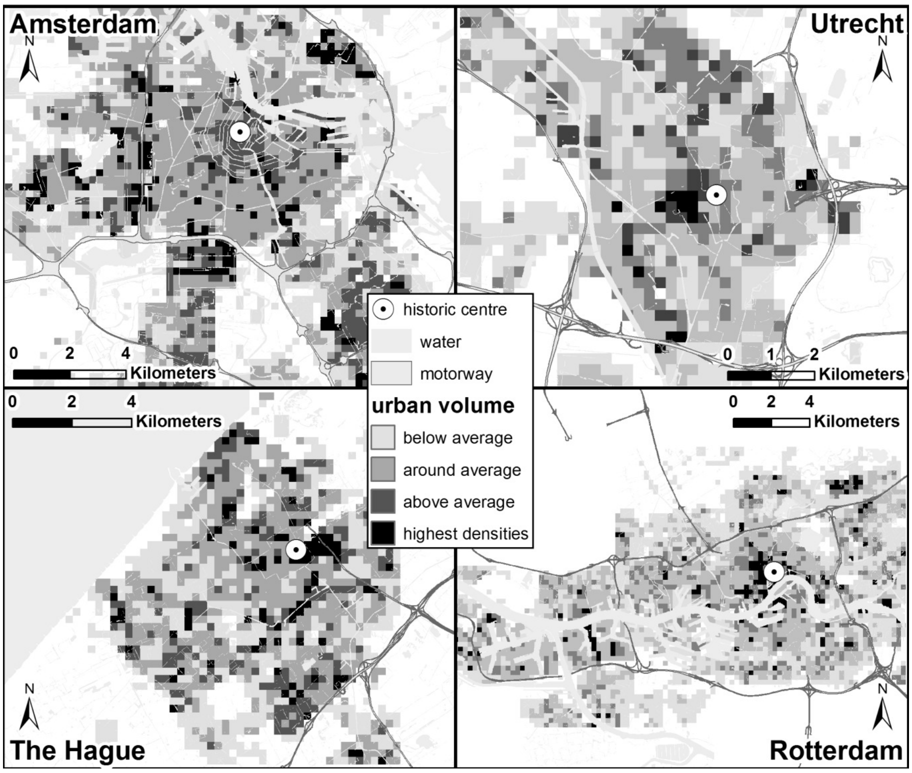

Figure 5. High-density patterns of the four major Dutch cities at 250 metres grid level. Scale varies per city.

4.3 Explaining urban density

After describing the observed urban density distribution and patterns in the previous sections we will now attempt to explain local density values in a statistical analysis with a limited set of explanatory variables. We confine ourselves to such a limited set, as our main objective is to offer an initial indication of several important aspects that explain density patterns, rather than fully explain this process. The analysis of urban density is done in two subsequent steps. First we analyse which factors explain whether a cell is classified as containing a building. In a second step we assess the importance of this same set of independent variables to explain the observed urban volume values.

Since the analysis of urban volume is only possible on the locations where buildings exist, the two steps are logically related. We therefore chose to apply a Heckman sample selection model (Heckman, 1979) that simultaneously analyses the presence of buildings in a binomial logistic regression and the related urban volume in a linear regression (Table 4). This type of analysis controls for possible correlations in the error term that might be present in the separately estimated models that are included in Table 5. The extent of this correlation is expressed in the *rho* parameter. In the case of Utrecht and Rotterdam the values for this parameter differ significantly from zero, indicating that a correlation of the error terms is indeed present and thus underpinning the need for this approach. For the other two cities the more straightforward analysis with two separate statistical models would have sufficed. In fact, in all four cases the two approaches yield very similar results.

In this analysis we use a concise set of spatial explanatory variables related to the proximity of transport facilities and major spatial planning (zoning) regulations. The importance of the selected themes for explaining urban development is widely recognised. Relevant research pointing at their relevance for the Dutch context includes Rietveld and Bruinsma (1998), Verburg *et al.* (2004) and Koomen *et al.* (2008). The transport facilities chosen here are: regular railway stops, Intercity stations and motorway exits (and entrances). For each grid cell we calculated the distance to the nearest of each of these facilities. The negative impact of the proximity of transport infrastructure on urban development that might result from, for example, noise disturbance, is accounted for in two categorical variables that indicate the presence of a railway or motorway area within 500 metres. The remaining spatial variables are also categorical and indicate, where appropriate, a location within a non-central part of town when it is divided by a natural barrier (major river), or a location within a restrictive development zone related to either open-space preservation (buffer zone) or the noise contour of the national airport. For operational reasons we incorporate proximity here as a Euclidean distance to the nearest facility. We thus refrain from using more elaborate measures that, for example, take actual travel time or perceived distances into account, mainly because we are looking at relatively short distances within major urban areas with intricate infrastructure systems. It should, furthermore, be added that from our analyses we exclude the locations that directly refer to water, motorway or railway areas, since these are, by our definition, not built-up.

The results from the part of the analysis that explains whether or not one or more building cells of 5 by 5 metres are present in a 25 by 25 metres grid environment are included in the bottom halve of Table 4. The explanatory power of the estimated statistical model cannot readily be assessed, but the R^2s of the related separate binomial logistic model ranging from 0.28 to 0.35 indicate a reasonable fit (Table 5). The impact of the explanatory variables is, in general, according to the expectations: an increase in distance to regular stations, Intercity train stations and, to a lesser extent, motorway exits leads to a decrease in the probability that a cell is built-up. The proximity of a motorway or railway positively influences the probability on built-up areas. A location in a restrictive development zone (Buffer zones) or the noise contour around the national airport is also less likely to be built-up. Some exceptions to the general pattern can be observed that are probably caused by specific local conditions. In Amsterdam and The Hague the distance to motorway exits has a small positive impact. This may be caused by the fact that the motorways here are located relatively far from the main built-up areas. In The Hague this situation may be partly caused by the city's location at the coast, which forms a natural barrier to the construction of a ring-road type of road infrastructure. The positive impact of the proximity of a railway in Rotterdam can possibly be attributed to the presence of extensive industrial areas surrounding the cargo railway line in the vast harbour area of the city. The fact that this harbour area and related working-class districts are situated on the southern (non-central) shore of the river Rhine, possibly explains why this shore is more likely to contain built-up areas. In Amsterdam the northern (non-central) shore of the river IJ has very few facilities as is reflected in the negative impact of a location here. The relatively low impact of the proximity of Intercity stations in Amsterdam and Rotterdam may be related to the very fine spatial detail of the analysis. The applied data sets clearly show the main (central) stations to be surrounded by sizeable non-built up areas that are usually made up of public squares and clusters of local infrastructure. Initial attempts to also include the distance to the (historic) centres of the four cities produced more ambiguous results, as these sites are normally located close to the (main) intercity station thus leading to colinearity problems. Borzacchiello *et al.* (2007) provide a related analysis on the maximum distances of the accessibility impacts on urban development in the same four cities described here. Their work also contains a more extensive description of the data preparation and statistical analysis process.

Table 4. Heckman selection model for explaining urban volume and presence of built-up areas

	Amsterdam		Utrecht		The Hague		Rotterdam	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
Ln urban volume								
N	88,256		39,340		64,952		129,490	
Constant	7.550	(0.024)	7.506	(0.024)	7.408	(0.016)	7.081	(0.019)
Distance (km) to nearest:								
regular train station	-0.018	(0.006)	0.005**	(0.006)	0.027	(0.006)	-0.099	(0.004)
intercity station	-0.209	(0.005)	-0.150	(0.007)	-0.239	(0.008)	-0.038	(0.002)
motorway exits	0.202	(0.006)	-0.065	(0.008)	0.133	(0.008)	0.103	(0.003)
Location within (1):								
500 m of motorway	-0.013**	(0.014)	0.041*	(0.018)	0.040*	(0.017)	-0.052	(0.012)
500 m of railway	-0.034*	(0.016)	0.082	(0.012)	-0.016**	(0.012)	0.257	(0.007)
buffer zone	-0.787	(0.058)	-1.980	(0.164)	-1.096	(0.095)	-1.706	(0.189)
Amsterdam North/ Rotterdam South	-0.552	(0.039)					0.106	(0.013)
airport noise contour	-0.647	(0.015)						
Built-up area indicator								
N	288,079		117,490		177,786		503,443	
Constant	0.582	(0.010)	1.580	(0.020)	0.746	(0.009)	0.333	(0.005)
Distance (km) to nearest:								
regular train station	-0.152	(0.002)	-0.167	(0.004)	-0.204	(0.004)	-0.146	(0.002)
intercity station	-0.115	(0.003)	-0.352	(0.003)	-0.320	(0.003)	-0.067	(0.001)
motorway exits	0.085	(0.004)	-0.087	(0.006)	0.272	(0.003)	-0.041	(0.001)
Location within (1):								
500 m of motorway	-0.272	(0.008)	-0.555	(0.010)	-0.356	(0.013)	-0.345	(0.006)
500 m of railway	-0.432	(0.007)	-0.067	(0.010)	-0.119	(0.011)	0.021	(0.005)
buffer zone	-1.366	(0.013)	-1.990	(0.052)	-2.337	(0.026)	-2.261	(0.050)
Amsterdam North/ Rotterdam South	-1.020	(0.010)					0.450	(0.006)
airport noise contour	-0.029	(0.010)						
rho	-0.002	(0.043)	0.169	(0.024)	0.020	(0.034)	0.069	(0.029)
sigma	1.098	(0.003)	1.054	(0.005)	0.998	(0.003)	1.124	(0.003)
lambda	-0.002	(0.047)	0.178	(0.026)	0.020	(0.034)	0.077	(0.033)
log likelihood	-279,480		-115,668		-183,762		-444,872	

Note: all variables are significant at the 0.01 level, unless otherwise indicated:
* significant at 0.05 level, ** not significant at 0.05 level.

The upper part of Table 4 shows the results for the part of the model that explains the local urban volume values. These values show a comparable impact of the individual explanatory variables as in the explanation of the presence of buildings. The model, however, only explains a limited part of the variance in urban volume values as is indicated by the low R^2s of the related separate linear regression model explaining urban volume (Table 5).

Table 5. Separate statistical models explaining presence of built-up areas and urban volume

	Amsterdam		Utrecht		The Hague		Rotterdam	
Ln urban volume								
N	88,256		39,340		64,952		129,490	
Adjusted R^2	0.133		0.027		0.065		0.051	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
Constant	7.550	(0.014)	7.477	(0.024)	7.415	(0.011)	7.122	(0.007)
Distance (km) to nearest:								
regular train station	-0.019	(0.004)	0.024	(0.006)	0.029	(0.005)	-0.091	(0.002)
intercity station	-0.209	(0.004)	-0.110	(0.004)	-0.235	(0.004)	-0.034	(0.001)
motorway exits	0.202	(0.006)	-0.058	(0.008)	0.129	(0.004)	0.106	(0.002)
Location within (1):								
500 m of motorway	-0.014**	(0.011)	0.110	(0.015)	0.044	(0.016)	-0.032	(0.009)
500 m of railway	-0.034	(0.009)	0.091	(0.012)	-0.014**	(0.012)	0.256	(0.007)
buffer zone	-0.789	(0.029)	-1.672	(0.158)	-1.058	(0.070)	-1.555	(0.178)
Amsterdam North/ Rotterdam South	-0.647	(0.015)					0.082	(0.008)
airport noise contour	-0.554	(0.018)						
Built-up area indicator								
N	288,079		117,490		188,585		872,266	
Nagelkerke R^2	0.275		0.345		0.335		0.308	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
Constant	-1.790	(0.024)	-0.126**	(0.083)	-1.621	(0.037)	-2.506	(0.080)
Distance (km) to nearest:								
regular train station	-0.243	(0.004)	-0.271	(0.007)	-0.383	(0.007)	-0.270	(0.002)
intercity station	-0.198	(0.004)	-0.588	(0.006)	-0.531	(0.005)	-.0132	(0.001)
motorway exits	0.135	(0.006)	-0.131	(0.011)	0.504	(0.006)	-0.097	(0.002)
Location within (1):								
500 m of motorway	-0.462	(0.013)	-0.927	(0.017)	-0.552	(0.020)	-0.641	(0.009)
500 m of railway	-0.701	(0.011)	-0.105	(0.017)	-0.321	(0.018)	0.320	(0.008)
buffer zone	-2.565	(0.027)	-4.404	(0.152)	-4.628	(0.070)	-6.217	(0.158)
Amsterdam North/ Rotterdam South	-0.062	(0.016)			-	-	0.917	(0.009)
airport noise contour	-1.734	(0.019)						

Note: the presence of built-up areas is explained with a binomial regression; Ln urban volume is explained with a linear regression.
All variables are significant at the 0.01 level unless otherwise indicated:
**not significant at 0.05 level.

The relatively poor performance in explaining in urban volume values is probably related to the limited variability in these values as was also apparent in the building height frequency (Figure 4). The vast majority of the buildings has a similar height (around 10 metres) and high-rise buildings are scarce. The analysis does, however, indicate a number of factors that favour the presence of high, voluminous or closely packed buildings. Especially the proximity of an Intercity station positively influences high volume values. Apparently a location near a main (central) station is crucial for high urban densities. In the case of Amsterdam and Rotterdam the presence of regular train stations seems to matter too, albeit to lesser extent. The proximity of a motorway exit has an opposite impact: an increasing distance is likely to lead to higher volume values. In Rotterdam a strong positive effect is generated by the proximity of the railway itself. These findings contradict the common suggestion that high-rise buildings are generally to be found at the edges of cities near motorways. In this respect, Dutch cities apparently differ from their American counterparts that do show a preference for high-density developments at their edges (Garreau, 1991; Stern and Marsh, 1997). The observed continuing importance of the current (historic) city centres is very much in line with the empirical and simulated evidence presented by Batty (2001). Frenkel (2004) also reports that existing high-rise buildings in the Tel Aviv metropolitan region in Israel have a higher probability of occurring in the core city. Proposed high-rise buildings, on the other hand, are in his research found to have a higher probability of occurrence in the outer rings of the region.

5. CONCLUSION

The proposed urban volume indicator provides an adequate characterisation of the actual physical appearance of the city in time and space. What is more: the quantitative description allows for an objective, highly detailed statistical analysis of urban patterns. The indicator thus helps visualising and quantifying the impact of the different forces that shape our city. In this respect it provides useful input to the ongoing debate on urban (re)development.

The presented spatio-temporal analysis of the urban development of the city of Amsterdam combines the urban volume indicator with other equally detailed base data. This study provides an interesting insight in the making of the city. The gradual, lateral extension is clearly mapped, but the analysis also shows the growing importance of numerous high-density zones throughout the city. This finding is further quantified in the related SOM-analysis. The SOM results also indicate the addition of isolated high-density zones to the historic medium-density city centre in the past century. This approach furthermore proves the recent emergence of small, but extreme high-density developments near stations at a considerable distance from the centre.

The urban volume indicator is also useful for characterising the differences in urban density in the four major Dutch cities. This initial study shows a distinction between cities in which high-density areas are concentrated in the original city centres (Rotterdam and The Hague) and cities that show these areas at a considerable distance from the centre

(Amsterdam and Utrecht). The distribution of the high-density zones in the latter cities clearly suggests a polycentric appearance. The layout of the major Dutch cities thus reflects evidence of opposing centripetal and centrifugal forces. The urban volume indicator can help visualise and quantify the impact of these forces, thus providing useful input to the ongoing debate on urban (re)development.

The statistical analysis that explained the presence of urban areas and their urban volume values puts less emphasis on the extremely high volume values and indicates the importance of intercity and regular train stations in urban development. The distance to motorway exits was found to be less important in this respect, indicating that the urban system in these major cities is still concentrated in the traditional centre served by train infrastructure.

An interesting extension to the current research would be the inclusion of additional information on different types of urban land use. This would allow for a distinction in, for example, residential, commercial and industrial areas and could help disentangle the factors that shape the expected density differences between these types of use. Other interesting additional data sources that could enhance the potential to represent urban density include the number of employees/residents per building.

In more general terms, the presented analysis shows the enormous potential of the highly detailed spatial data sets that are currently becoming available in many countries. This is not only true for the high resolution height data that were used in the current paper, but also for other data derived from sources such as the newest generation of remote sensing satellites, large-scale inventories of cadastral institutes, and the tracking and tracing of mobile phone users. The latter type of data can, for example, be used in the (real time) monitoring of traffic flows based on the movements of individuals. But also for many other socioeconomic phenomena we are now able to use fine-grained data that allows spatial analysis at scales that were unimaginable until recently. From the current analysis it becomes clear that such highly detailed data sets offer considerable challenges in terms of both analysis and visualisation of fine-scaled developments over extensive areas. The increasing level of detail of newly available data sources, in fact, calls for a rethinking of current analysis and presentation methods, leaving many new research roads open to explore.

ACKNOWLEDGEMENTS

The geographical data sets that were indispensable for this research were kindly supplied by the Survey Department of the Dutch Ministry of

Transport and Water Management, the Netherlands Institute for Spatial Research and the Spatial Planning department of the city of Amsterdam. We thank these institutes for their kind co-operation. Furthermore, we are particularly grateful to Richard Kaufholz for his initial work on the urban volume indicator and Maria Teresa Borzacchiello for helping prepare part of the additional data needed for this study.

REFERENCES

Batten, D.F. (1995) Network Cities: Creative Urban Agglomerations for the 21st Century, *Urban Studies* 32 (2): 313-327.

Batty, M. (2000) The new urban geography of the third dimension, *Environment and Planning B* 27: 483-484.

Batty, M. (2001) Polynucleated urban landscapes, *Urban Studies* 38 (4): 635-655.

Batty, M., Besussi, E., Maat, K. and Harts, J.J. (2004) Representing multifunctional cities: density and diversity in space and time. *Built Environment* 30 (4) 324-337.

Borzacchiello, M.T., Nijkamp, P. and Koomen, E. (2007) *Accessibility and urban development: a grid-based comparative statistical analysis of Dutch cities*. Paper presented at NECTAR Workshop "Network Dynamics and Location Choice" October 26-27, 2007, Stuttgart, Germany.

CBS (2005) *StatLine*, the central database of Statistics Netherlands, Centraal Bureau voor de Statistiek (CBS), http:// www.statline.nl, last accessed October 05, 2005.

Dieleman, F.M. and Faludi, A. (1998) Randstad, Rhine-Ruhr and Flemish diamond as one polynucleated macro-region? *Tijdschrift voor Economische en Sociale Geografie* 89 (3): 320–327.

EEA (2006) *Urban sprawl in Europe; the ignored challenge*. European Environment Agency (EEA), Copenhagen, Denmark.

Fisher-Gewirtzman, D., Burt, M. and Tzamir, Y. (2003) A 3-D visual method for comparative evaluation of dense built-up environments, *Environment and Planning B* 30: 575-587.

Frenkel, A. (2004) *Spatial distribution of high-rise buildings within urban areas; the case of the Tel Aviv metropolitan region*. Paper for the 44th congress of the European Regional Science Association, Oporto Portugal.

Garreau, J. (1992) *Edge City: life on the new frontier*. Doubleday New York.

Gaschet, F. (2002) The new intra-urban dynamics: suburbanisation and functional specialisation in French cities, *Papers in regional science* 81 (1): 63-82.

Heckman, J.J. (1979) Sample selection bias as a specification error, *Econometrica* 47: 153-161.

Huising, E.J. and Gomes Pereira, L.M. (1998) Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetery and remote sensing* 53 (5): 245-307.

Kaufholz, R.P. (2004) *Building densities; a physical volume approach*. MSc thesis Faculteit der Aard- en levenswetenschappen Vrije Universiteit Amsterdam.

Kloos, M. and De Maar, B. (1995) The hesitant high-rise of Amsterdam, In: Kloos, M. (ed.) *Amsterdam's High-Rise; considerations, problems and realizations*, Architecture and Natura Press, Amsterdam.

Knol, W.C., Kramer, H., Van Dorland, G.J. and Gijsbertse, H. (2003) *Historisch grondgebruik van Nederland: tijdreeksen grondgebruik Noord-Holland van 1850 tot 1980.* Alterra-rapport 751, Alterra, Wageningen.

Kohonen, T. (2001) *Self-organizing maps.* Springer, Berlin.

Koomen, E., Kaufholz, R., Rietveld, P. and Scholten, H. (2004) 3D-GIS and Urban Volume: applying the third dimension in a morphological study of the Amsterdam Urban Landscape, In: Toppen, F. and Prastacos. P. (eds.) *AGILE 2004 conference proceedings*, Crete University Press, Heraklion Greece, pp: 139-147.

Koomen, E., Dekkers, J. and Van Dijk, T. (2008) Open-space preservation in the Netherlands: planning, practice and prospects, *Land Use Policy* (forthcoming).

Lang, R.E. and LeFurgy, J. (2003) Edgeless Cities: Examining the Noncentered Metropolis, *Housing Policy Debate* 14 (3): 427-460.

Longley, P.A. and Mesev, V. (2002) Measurement of density gradients and space-filling in urban systems, *Papers in regional science* 81 (1): 1-28.

Maat, K., and Harts, J.J. (2001) Implications of urban development for travel demand in the Netherlands. *Transport research record* 1780: 9-16.

Martori i Cañas, J.C. and Suriñach i Caralt, J. (2002) *Urban population density functions: the case of the Barcelona region.* Documents de Recerca. Universitat de Vic, Spain.

Oude Elberink, S., Brand, G. and Brügelmann, R. (2003) Quality Improvement of Laser Altimetry DEMs. In: Maas, H.-G., G. Vosselman and A. Streilein (eds.), *Proceedings of the ISPRS working group III/3 workshop `3-D reconstruction from airborne laserscanner and InSAR data'* Dresden, Germany 8-10 October 2003.

Priemus, H., (2001) Corridors in The Netherlands: apple of discord in spatial planning. *Tijdschrift voor economische en sociale geografie* 92(1): 99-107.

Rietveld, P. and Bruinsma, F.R. (1998) *Is transport infrastructure effective?* Springer-Verlag, Berlin.

Ritsema van Eck, J., Van Oort, F., Raspe, O., Daalhuizen, F. and Van Brussel, J. (2006) *Vele steden maken nog geen Randstad*, NAi uitgevers/ Ruimtelijk planbureau, Rotterdam/Den Haag.

Stern, M.A. and Marsh, W.M. (1997) The decentered city: edge cities and the expanding metropolis, *Landscape and Urban Planning* 36: 243-246.

skyscrapers.com (2004) *Amsterdam, buildings of the city*, http://www.skyscrapers.com, last accessed February 26, 2004.

TDN (1998) *Productbeschrijving top10vector*. Topografische Dienst Nederland, Emmen.

Van der Burg, A. and Van Oort, F.G. (2001) Stedelijke netwerken en de Vijfde Nota over de Ruimtelijke Ordening. In: F. Boekema en Kuijpers, E. (eds.), *Stedelijke netwerken: theoretische en empirische reflecties*. Shaker, Maastricht, pp. 5-18.

Van der Heiden, N. and Wallagh, G. (1991) How the Republic of Amsterdam got the Kingdom of the Netherlands to play its Game, *Built Environment* 17 (1): 34-43.

Van der Laan, L. (1998) Changing Urban Systems: An Empirical Analysis at Two Spatial Levels, *Regional Studies* 32 (3): 235-247.

Verburg, P.H., Ritsema van Eck, J.R., De Nijs, T.C.M., Dijst, M.J. and Schot, P. (2004) Determinants of land-use change patterns in the Netherlands. *Environment and Planning B* 31(1): 125-150.

Vesanto J. and Alhoniemi, E. (2000) Clustering of the selforganizing map, *IEEE Transactions on Neural Networks* 11 (3): 586-600.

Vosselman, G. and Maas, H-G. (2001). Adjustment and filtering of raw laser altimetry data, In: Torlegard, K. (Ed.): *Proceedings of OEEPE workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models*, Stockholm 1-3 March 2001.

Yoo, C., Jung, K-S., and Kim, T-W. (2005) Rainfall frequency analysis using a mixed Gamma distribution: evaluation of the global warming effect on daily rainfall, *Hydrological Processes* 19: 3851–3861.

Yue, S. (2001) A bivariate gamma distribution for use in multivariate flood frequency analysis, *Hydrological Processes* 15: 1033–1045.