Visualising postcode data for urban analysis and planning: the *Amsterdam City Monitor*

Karin Pfeffer, Marinus C Deurloo and Els M Veldhuizen

Department of Geography, Planning and International Development Studies, University of Amsterdam,
1018VZ Amsterdam, The Netherlands

Email: k.pfeffer@uva.nl

Public agencies typically collect statistical databases on population, housing, economic activities and various other indicators. In their existing form, these databases cannot be directly adopted for planning and policy formulation, as they need to be tailored to the issue for which information is required. The specific problem addressed in this article is how to process and visualise small-area data, in our case referring to the Dutch 6-digit postcode, to provide detailed information for areas that are normally difficult to target in policy terms because they do not relate to the identifiable administrative entities that planners and politicians are accustomed to. Therefore a methodology is presented that identifies spatial concentrations based on the underlying data using geographical information systems (GIS). Emphasis is placed on the technical characteristics of this method, which has been applied in a number of urban studies, and on the advantages and disadvantages for further utilisation. Using the Amsterdam City Monitor as a case study, this paper illustrates how to produce detailed maps of urban dynamics within a city to support local policy and planning. It also shows that the presented method of mapping spatial concentrations can be applied to a variety of urban issues, and complements the common practice of using intensity classes to compare places.

Key words: geographical information systems, monitoring, postcode, spatial concentrations, target geography, Amsterdam

Introduction

Public agencies typically collect statistical data on population, housing, economic activities and various other indicators. These data are used for publishing formal statistical analysis of standard administrative areas that are meaningful regarding physical boundaries, age of houses, types of economic activities, political representation, etc. (Amrhein 1995; Young *et al*. 2009). However, standard administrative areas usually used for place comparison are insufficient if local policymakers and urban researchers want an image of urban issues and dynamics not bounded to areas normally used in policy formulation (Baud *et al*. 2008). Moreover, because data are now usually registered at a more detailed scale (e.g. postcode or address), there is no longer a compelling reason to only use standard administrative areas.

This article concerns the presentation of a mapping method for visualising small-area data (i.e. postcode data) to support urban analysis and planning. Specifically,

small-area data are aggregated into case-specific spatial units representing spatial concentrations. The spatial configuration of these units changes over time according to the geographic and thematic nature of the underlying data. We can identify and accentuate conspicuous areas that clearly deviate from the average situation of the phenomenon by filtering areas with values considerably higher than the mean and excluding the base amount. These may be areas at risk and in need of special attention, for example because of high levels of unemployment. To date, this method has been used in several urban studies applications that require detailed knowledge on spatial concentrations beyond standard boundaries, but it has never been documented as such.

One of the early applications of the method presented concerned mapping pockets of poverty to investigate whether a lack of employment opportunities concentrated in specific areas within selected cities (Deurloo *et al* 1998), why it was concentrated and what the effects were on the prospects of people living in those areas. Later,

Musterd and Deurloo (2002) analysed the dynamics of spatial concentrations of immigrants originating from non-industrialised countries in Amsterdam to explore whether local and national government concerns over large spatial concentrations of these immigrants was based on the on-the-ground situation. Similarly, Aalbers and Deurloo (2003) compared patterns of spatial concentrations of different ethnic groups, concluding that immigrants from industrialised countries concentrate more in the same areas and are also more satisfied with their neighbourhood. A further application was to map crime reports (Musterd *et al*. 2004) to question an often assumed correlation between the occurrence of crime and local population characteristics, as well as the social and physical structure of the city. Clusters are also typical for economic geography, and Musterd and Deurloo (2006) applied this method to explore possible explanations for the emergence of clusters of cultural industries and to derive preconditions for a creative knowledge city. The examples above illustrate the applicability of our spatial concentration method for various urban issues to support public policy as well as further our understanding of urban phenomena and dynamics.

In the remainder of the article, we will first discuss related approaches of mapping small-area data and determining spatial concentrations. This is followed by a detailed explanation of our method to process small-area data (in our case, the Dutch 6-digit postcodes that comprise about 10 to 20 households) into useful information. Using the *Amsterdam City Monitor* as a case-study, we show how such detailed maps of urban dynamics can support urban analysis, illustrated by visualising concentrations of the Moroccan population in specific parts of the city over the period 1994 to 2009. We conclude the article with our main findings regarding the proposed method, and its opportunities and limitations for supporting urban analysis and planning.

Background: Review of related methods

From 'source' geography to 'target' geography
'Source' geography is the basic spatial unit for which data are collected, while 'target' geography refers to the aggregated area for which data are made available to the user (Martin 2003).

Nowadays, source geography of social data is often an address, or the approximate centre of a small area such as a service zone or an enumeration block. These kinds of spatial indications can easily be converted into point locations in Cartesian space. However, in many situations, mapping the geographical location of raw point data is not permitted, undesirable or not of use (Deurloo and Musterd 2001). First, point data may be subject to confidentiality restrictions because individuals could be identified (Gutmann and Stern 2007). Secondly, point maps are difficult to read and to interpret, especially if a substantial number of points are plotted and if the study area is large; only lower and higher point densities can be recognised in those cases. Thirdly, the information provided by a point map is highly scale dependent. Accordingly, it is more meaningful to map social point data in a wider geographical context, including the direct surroundings of the point, implying further (geo)processing and aggregation of the source data. To do so, various methods can be applied using a geographical information system (GIS), depending on the nature of the data and the desired outcome. Commonly used methods in the social sciences are aggregating source data using the 'point-in-polygon' method to some target geography, such as standard administrative areas, regular structures (e.g. grid, hexagons) (Dorling and Bethan 2004) or constructed zones. More advanced methods utilise computational algorithms, as used for the 2001 UK census to construct zones based on population characteristics and physical features (Martin 2000).

Standard administrative areas are common in urban management to visualise geographical patterns and compare places. However, their appropriateness or the definition of their boundaries is debatable (Rees 1997; Martin 2004), as the rationale for creating the boundaries may not always be clear, and is often a pragmatic compromise of competing considerations. Small areas with large populations are virtually invisible to the viewer, while larger areas with small populations dominate the map. Such areas are also not standardised statistical units and differ in population size and area; the uncertainty associated with the data varies with their size and area. Furthermore, demarcated areas suggest an abrupt change of a phenomenon at the boundary.

Aggregating source data to a regular structure (e.g. grid or hexagon) is often used to account for the continuous variation in data as it approximates a surface of consistently sized units (i.e. fields). The finer the structure, the more detail can be mapped, but low densities may be exaggerated. In some applications, cells or hexagons containing very few 'sources' (as in the case of people) need to be suppressed to protect privacy. The coarser the regular structure, the more smoothing takes place. Fields may be preferred because of the comparability of the spatial units to which data are assigned, the greater potential for spatial analysis and the possibility for linking data referring to cells or hexagons with other data (Burrough and McDonnell 1998; Martin 1996). They also provide a better estimate of the population distribution as it represents the underlying geography of the phenomenon independently of any zoning system except for the regular structure (Martin 1996; Harris and Longley 2001).

A third type of target geography is constructed zones, such as areas built on Thiessen polygons (e.g. Okabe *et al.* 2000) to construct postcode areas (Boyle and Dunn 1991) or other output areas such as those for the 2001 UK census (Martin 2000). One shortcoming of this method is that rather large areas are generated in less densely populated areas, producing a distorted image in thematic mapping. Furthermore, the set of constructed areas completely cover the study territory, including uninhabited areas. Finally, the procedure generates boundaries with which users are not familiar.

All methods of aggregating point data to some target geography mentioned above suffer from the assumption that all variables can be adequately represented using the same geographical division. However, some themes, such as health, might require a different mapping unit (Flowerdew *et al.* 2007). Secondly, the choice for a particular form, size and location of a geographic area has a considerable effect on the representation of the measured attribute, an effect commonly referred to as the modifiable areal unit problem (MAUP) (Openshaw 1984b). Several studies have been carried out to get a better understanding of the effects of aggregation and scale associated with the number of output areas (e.g. Holt *et al.* 1996; Wong *et al.* 1999). Nevertheless, there is not yet a common rule for the choice of mapping unit. Drawing on the experience of the authors, the choice of areal units in urban analysis is – despite the MAUP – largely determined by data availability, information requirements, familiarity with a particular geographical framework and the policy requirements and scale and units of local government structures.

Alvanides *et al.* (2001) consider the MAUP not solely as a problem, but as a potentially useful exploratory spatial analysis tool, providing the basis for a new, more explicit and geographically focused approach for the geographical analysis of thematic data. In conformance with their argument, we consider small-area data to offer opportunities regarding the MAUP, and we will demonstrate how these can be used to construct dynamic spatial objects.

Spatial concentrations
The key concept in our approach is 'spatial concentration' and here we elaborate this concept and its operationalisation in the literature.

Essentially, spatial concentration of a phenomenon can be conceived in two ways: higher relative values (for instance, compared to the city average) at a few locations ('peaks') or the clustering of similar geographical units that are close to each other, like the clustering of neighbouring wards with similar levels of poverty.

One way of identifying (spatial) concentration is threshold analysis. This method illustrates peaks by pinpointing entities (e.g. areas) that are above a particular threshold or cut-off point. Johnston *et al.* (2003) applied a range of thresholds to different spatial scales, providing a detailed tabular representation of the degree of spatial concentration for different ranges or geographical areas. Their approach, however, does not indicate where the spatial concentrations actually are if not visualised as a map, and also ignores within-area variation.

Another way of mapping spatial concentrations using point data is drawing a kernel of variable widths around each point (e.g. a point representing a postcode), aggregating the attribute data of nearest neighbours of the kernel centre point and of the centre itself (Harris and Longley 2004) to the kernel. Harris and Longley determined kernel widths by population size within the kernel, keeping total population constant across all kernels. A similar approach is frequently used in crime and disease mapping to estimate the density of events (Gatrell *et al.* 1996; Ratcliffe and McCullagh 1999). Within a specified search radius, event data, for instance the number of crimes or infections, are aggregated, possibly weighted by the distance to the kernel centre, producing a continuous surface of event distribution. The kernel method can help address much of the MAUP because it deviates from standard geographies and accounts for within-area variation (Ratcliffe and McCullagh 1999); however points or events may be members of several kernels as information is pooled locally.

A third method of exploring spatial concentrations is through local indicators of spatial associations that are 'useful in detecting places with unusual concentrations of high or low values' compared to the average value (Páez and Scott 2005). The commonly used *G* measures of Getis and Ord (1992) have for instance been utilised for identifying clusters of ethnic groups (Poulsen *et al.* 2011) or mapping hotspots of crimes (Ratcliffe and McCullagh 1999) by comparing the local value to the global mean, also taking into account the values of adjacent areas.

Our method for mapping spatial concentrations resembles elements of the described methods and complements them by creating dynamic case-specific spatial units dependent on the underlying data.

Dynamic target geography

Data, methods and approach
To illustrate our approach to determining spatial concentrations, we use social data collected by the municipality of Amsterdam and aggregated[1] from individual residents, households or houses to the Dutch 6-digit postcode by the statistical bureau of Amsterdam (O+S Amsterdam). We also utilise a data file with the Cartesian coordinates of all postal addresses in the city of Amsterdam.

In the Netherlands, postcodes were introduced in 1978 and mapped in 1989 (Raper *et al.* 1992); their use as a spatial referencing system in geographic research started in the 1990s. The Dutch 6-digit postcode is quite a small unit, averaging 50 x 50 metres (or 10 to 20 households) in urban areas. This offers interesting analytical opportunities, but involves a few constraints. First, like other postcode systems, Dutch postcodes are not directly comparable because they comprise very different sizes and importance. For example, the number of inhabitants in a postcode in Amsterdam varies from 1 to over 1300, and therefore postcodes cannot be used meaningfully as individual objects of analysis. Moreover, the area and boundary of a Dutch postcode is not stable over time or clearly defined on the ground, being originally designed for postal delivery at individual addresses. In order to obtain relevant and meaningful information from postcode-level data, we introduce a dynamic target geography that requires further (geo)processing of the original geographical units, dependent on the thematic content of the underlying data at hand. In our method, less interesting postcodes are removed, and relevant postcode areas are clustered into new, larger, spatial units. The new units may vary across themes and within themes across time.

Considering the methods review in the second section, our method combines identifying a concentration at the postcode unit using the threshold method, plus geographical clustering of identified postcodes where clusters of selected postcodes appear as new areal units. Clusters with small amounts of the group considered can be filtered out via interactive queries. This procedure results in maps as shown in Figure 1A. Figure 1A displays all concentrations (aggregated postcode areas) in 2009 in Amsterdam where at least 100 'non-western immigrants' were registered and the share of 'non-western immigrants' is above 49 per cent. A comparison with the common representation of intensity classes at the neighbourhood level (Figure 1B) clearly shows the refinement achieved by the method: the resulting clusters account for the variation within neighbourhoods and cross neighbourhood boundaries (see Figure 1C). In this context neighbourhoods are administrative units of about 7500 inhabitants used by the municipality of Amsterdam to publish population data.

Producing maps of spatial concentrations as illustrated in Figure 1A involves four steps:

1 construction of the postcode geography to which source data are matched
2 selection of relevant postcodes from source data
3 aggregation of postcode areas and associated data into new spatial units according to clustering rules and
4 selection of a minimal size of a cluster for interactive analysis.

We will elaborate these steps further in the next sections, using the group of 'non-western immigrants' in 2009 in Amsterdam as an example.

Relevant postcodes
To identify relevant postcodes, selection criteria need to be formulated. We label an observation unit (i.e. a postcode) as relevant if the variable at hand (the proportion of 'non-western immigrants') for the unit is at least two binomial standard deviations above the city average (binomial distribution is used for proportions calculated from count data). Applying two standard deviations is fairly stringent, emphasising those cases that deviate considerably from the mean.

For our example, the following calculations resulted: The proportion p of the cases 'non-western immigrants' in 2009 in the city of Amsterdam is 34.74 per cent. The average number of residents per postcode in 2009 in the city is 43.84. Accordingly, the standard deviation σ is:

$$\sigma = \sqrt{p*(100-p)/n} = \sqrt{(34.74*65.26)/43.84} = 7.19 \quad (1)$$

and the threshold T for relevancy is:

$$T = p + 2*\sigma = 34.74 + 2*7.19 = 49.12\% \quad (2)$$

In summary, all postcodes where the share of 'non-western immigrants' in 2009 is above 49.12 per cent are selected for further processing, All postcodes with less than 49.12 per cent 'non-western immigrants' are left out and considered to be part of the 'base amount' in the city.

Postcode geography
As mentioned earlier, postcodes do not exactly define a geographic area. Using postcodes for geographical representation requires geographical demarcation of the postcode unit. Across GIS applications, various solutions are used. We create the postcode geography from the point locations of all postal addresses, given as Cartesian coordinates, using a *convex hull* (Preparata and Shamos 1985) operation. This operation places a rubber band around all address locations that belong to the same postcode such that the resulting polygon encloses all locations that lie on or within the boundary of the polygon (Figure 2). To avoid distortion due to possible errors in the address database, the resulting polygons are checked for strange shapes that would hinder further calculations by examining whether the perimeter or area of the constructed postcode areas exceeds a set maximum value. The postcode geography needs to be regularly updated (at the same pace as the attribute data), since the data set of postal addresses changes continuously due to new developments and urban renewal.

The procedure results in objects of different sizes and shapes that may overlap. They neither fit with other geog-
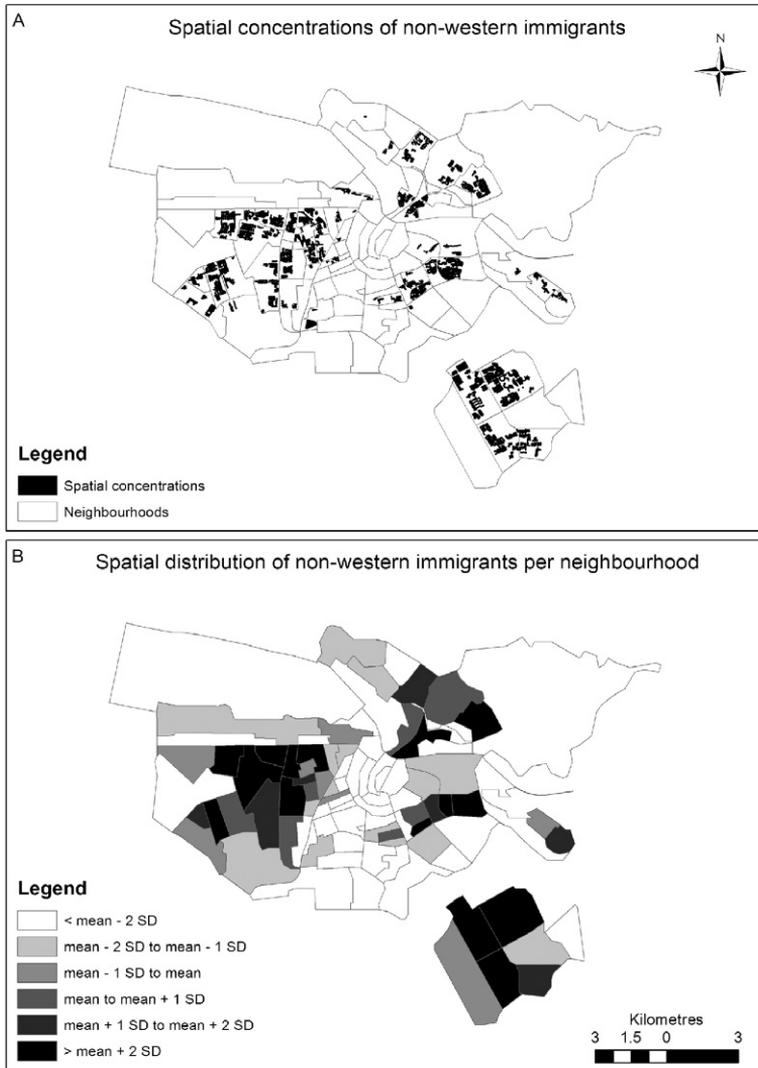
Figure 1   Amsterdam, 2009. A: Concentrations of 'non-western immigrants' (N (count of that group) >100 and percentage of group above 49 per cent); B: Neighbourhood distribution of 'non-western immigrants' using mean and SD for determination of categories
Source: O&S Amsterdam

raphies, nor do they necessarily cover the entire study area. Furthermore, the importance of the exact boundary of an individual postcode should not be over-emphasised; the relative position of the geographical objects compared to each other is more important than the precise demarcation of a postcode area.

There are also other ways to construct postcode areas (e.g. Penninga *et al*. 2005); however, the differences between the boundaries are not essential for our procedure.

Spatial concentration areas

The procedure for creating spatial concentrations from the source data, illustrated in Figure 3, is based on four rules. The first rule is to meet the selection criterion defined in equation 2 (above): an observation unit is considered relevant if the observed percentage of the variable at hand is above the threshold T. Secondly, in order to prevent the predominance of singularities and to protect the privacy of individuals, anomalies are filtered out by removing selected postcodes with a very low density (i.e. fewer than
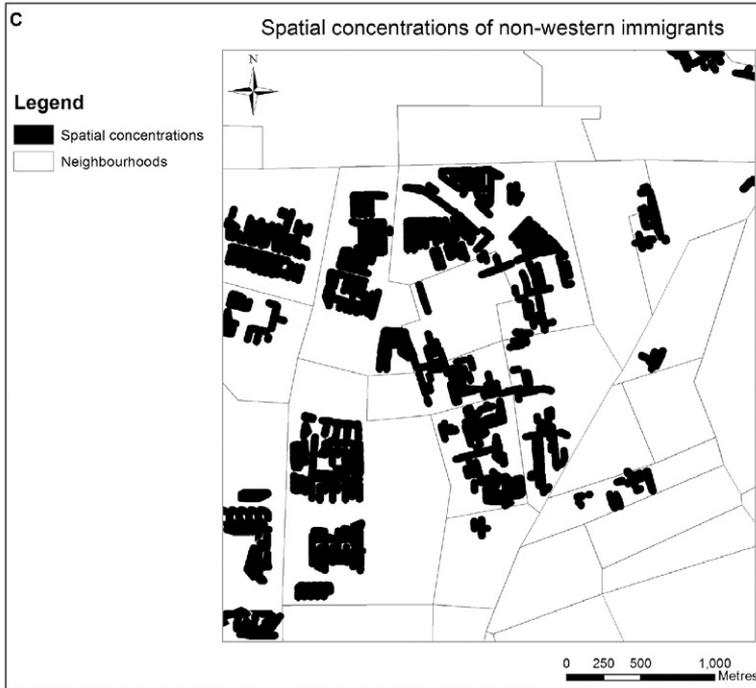
Figure 1 *Continued* **Amsterdam, 2009. C: Zoom to neighbourhoods where spatial concentrations cross boundaries (based on Figure 1A)**

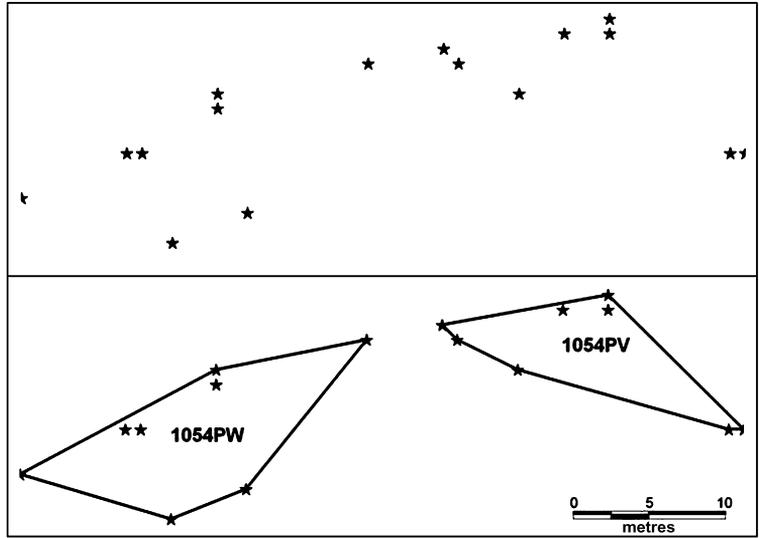| POSTCODE	X	Y
1054PV	119763	486393
1054PV	119758	486396
1054PV	119777	486389
1054PV	119769	486398
1054PV	119778	486389
1054PV	119759	486395
1054PV	119766	486397
1054PV	119769	486397
1054PW	119743	486393
1054PW	119753	486395
1054PW	119745	486385
1054PW	119743	486392
1054PW	119737	486389
1054PW	119738	486389
1054PW	119740	486383
1054PW	119730	486386

Figure 2   Demarcation of postcode areas: table with XY coordinates of postal addresses on the left; on the upper right, point locations of postal addresses; on the lower right, postcode areas after applying the convex hull operation
Source: O&S Amsterdam

20 persons, houses or households per hectare). We use 20 because it is far below the urban population density in the context of Amsterdam. Thirdly, the spatial representation of the postcode area is extended by a buffer of 25 metres to prepare for the next step: combining overlapping areas into new spatial units. The choice of 25 metres is based on experiments with varying buffer widths between 10 and 50 metres; a buffer of 25 metres accommodates the
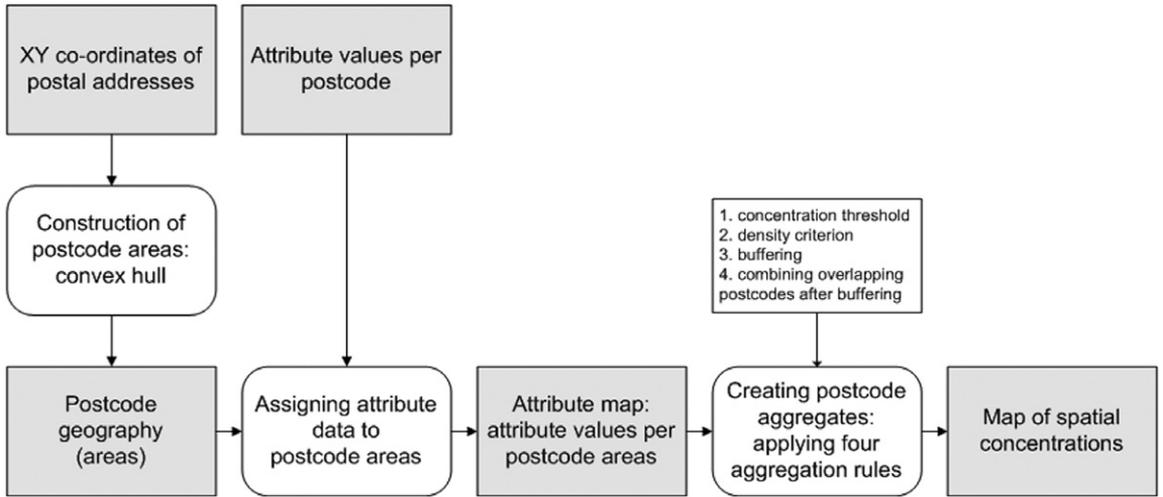
Figure 3   Procedure of generating maps of spatial concentrations (using postcode clusters)

variation in housing density, i.e. high densities in the Amsterdam city centre compared to lower densities in the district Amsterdam South East. Finally, the extended postcode areas are combined if they overlap. The result is the final target geography of concentration areas. Hereafter, relevant thematic data are recalculated for each concentration area.

Querying spatial concentrations

The resulting spatial concentrations, including their attribute data, are stored, forming the 'raw' archive map layer for further analysis within a GIS. It is necessary to apply further selection criteria to these layers to filter small concentrations and tailor the display to the information demand. In Figure 1A, these criteria are concentration areas with more than 100 'non-western immigrants' and more than 49 per cent of that group. The *Amsterdam City Monitor* (cf. the next section) supports the end-user in querying and mapping the archive of spatial concentrations, choosing different cut-off points depending on the purpose.

The *Amsterdam City Monitor*

The proposed mapping method has been applied to the *Amsterdam City Monitor*.[2] This GIS-based intranet/internet application assists in querying interactively more than 1000 map layers of spatial concentrations for a variety of themes and over time (including spatial concentrations since 1994). Because of the GIS functionality, one can apply various filters to exploit the underlying attribute information and fine-tune information needs. It can for instance be used to map the presence and dynamics of 'ethnic neighbourhoods', defined by van Amersfoort

(1992) as residential areas that are predominately inhabited by one ethnic group, but mixed with other groups. Figure 4 displays the dynamics of Moroccan 'neighbourhoods' in Amsterdam. This figure specifically shows an increase in both the number and magnitude of these areas and emergence in new areas. In the displayed concentrations, occupied by at least 200 Moroccans, Moroccans are the largest group, mixed with other ethnic groups such as Turkish, Antilleans, Surinamese or Dutch. The pattern over time reveals a shift of Moroccans' establishment from the inner city towards the more recently built areas, based on a shift in housing career (from lodging houses to council houses), rental type (from private to social) and housing quality (from small sub-standard to spacious dwellings).

This example illustrates that our method can produce meaningful information for studying residential concentrations within and across administrative wards (cf. Musterd and Deurloo 2002). Since the maps of spatial concentrations are annually updated, the *Monitor* informs policymakers about these types of trends, supplemented with a statistical summary of the set of selected areas.

Opportunities and limitations

Methodology

While macro-level research helps to identify problem neighbourhoods, micro-level research helps to identify spots that will be difficult to target in policy terms because they do not relate to the identifiable administrative entities that planners and politicians like to see (Ackerman and Murray 2004; Deurloo and De Vos 2008). We have demonstrated that spatial concentrations highlight areas that deviate from the average pattern and are tailored to the
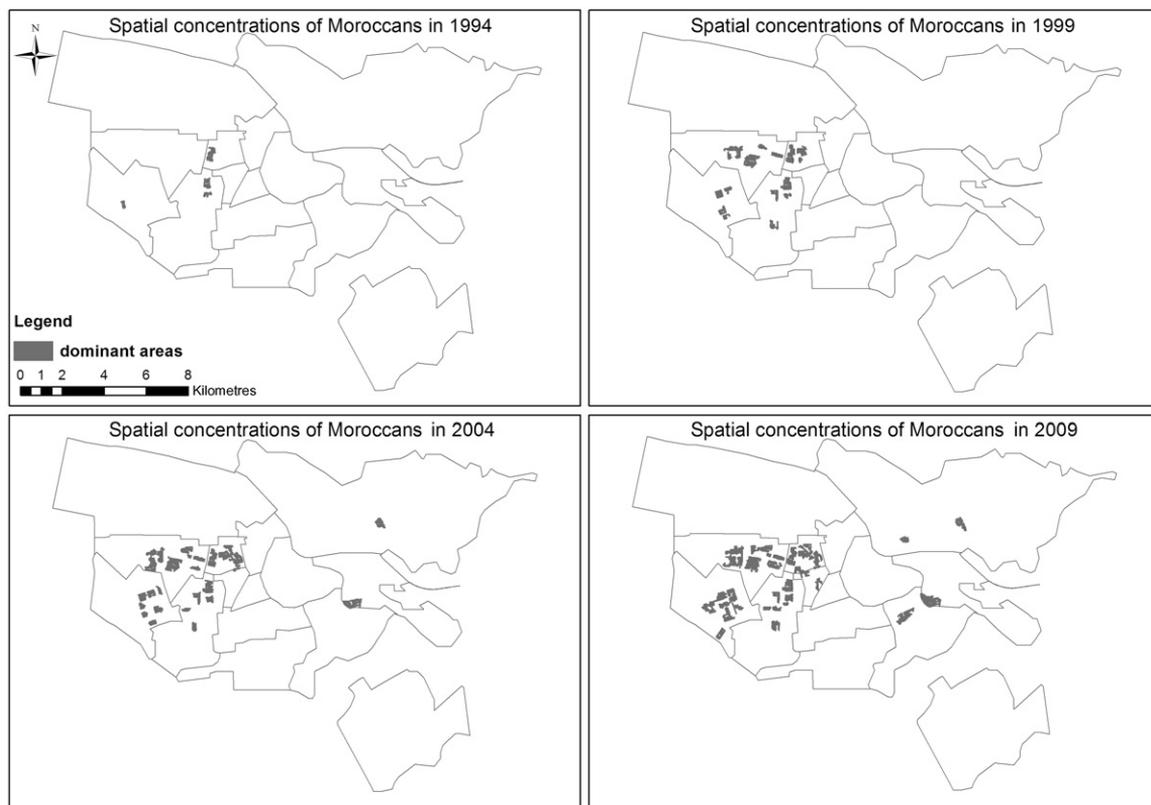
Figure 4   Dynamics of Moroccan neighourhoods; shaded areas display concentration areas that clearly deviate from the city average and that are predominately inhabited by Moroccans (N>200)
Source: *Amsterdam City Monitor*

underlying data; they accentuate the details in a dataset within and across coarser spatial units.

On their own, postcodes are not a useful mapping unit, especially at the urban fringe or in low-density areas. However, our method provides opportunities to aggregate small-area data in such a way that best fits the mapping purpose. Harris and Longley (2001) share the utility of postcode geography in the analysis of residential outlines that are often ignored by choropleth-map based analysis. Very high resolution remote sensing images could help to analyse within- and across-ward variation (Harris and Longley 2001), limited primarily to physical characteristics.

Our approach overcomes the MAUP to a certain extent, because boundaries of postcode aggregates are not predefined but determined by their adjacency to other postcode areas and the underlying statistics. Maps are comparable over time and space because of the consistency in data processing, thus maintaining relevance for monitoring and evaluation purposes. Spatial concentration maps also assist in analysing local spatial associations

between categories by exploring whether spatial concentrations of different categories are in the proximity of each other or whether another category is overrepresented in the same cluster (Deurloo and De Vos 2008).

Ideally, we would like to apply the procedure to individual-level data, but access to these databases is restricted (though not impossible) because of privacy disclosure issues. We have accounted for that also at the postcode level by excluding low densities.

Although the Netherlands appears to be unusual in having detailed social and demographic data from government sources at the postcode scale, the method is applicable to other databases (e.g. census data) and local circumstances. In many other countries, (sample) postcode data can be obtained from private companies.

A weakness of the methodology presented is the particularity of the aggregation rules. For analytical purposes, it would be interesting to tune the buffer width and the amount of deviation from the mean according to the nature of the social data and analysis purpose, for instance by calculations 'on the fly'. With current technology, the

aggregation algorithm takes a few minutes per archive map layer, which makes the procedure impractical for interactive analysis.

A problematic issue of the fine spatial detail is misuse or wrong interpretation. Maps of spatial concentrations may be interpreted to support preconceptions that people wish to see confirmed, for instance an overlap of non-western immigrants and unemployment, and may nourish hostile feelings towards minorities. Another trap is the ecological fallacy (Openshaw 1984a), because the presence of a spatial concentration at a particular location does not mean that all individuals being part of that aggregate can be associated with its characteristics. However, our method provides more detailed and precise information about where particular phenomena are located, and therefore helps to avoid the sorts of misinterpretations and ecological fallacies that the use of other more coarsely aggregated data can bring.

We noticed that visualising only spatial concentrations and not the base amount is the main difficulty in broadly interpreting the maps. For an overall picture we therefore recommend providing a common representation of intensity classes at the same time (see Figure 1), because people are more used to interpreting standard choropleth maps due to their familiar boundaries.

Potential of a GIS-based monitor
In the last decade an enormous wealth of small area data became available. Scientists appreciate such detailed data as an input for data processing and analysis. However, such a high resolution is often not what planners or policymakers are looking for (van Beurden and Douven 1999). They may be overwhelmed by the weight of information available in very fine geographical aggregations (Alvanides and Openshaw 1999). Providing processed maps of spatial concentrations through web-based applications like the *Amsterdam City Monitor* facilitates access to systematic and updated information. Planners and urban managers can easily adopt the variety of small-area information in addressing urban issues over time without getting lost in technical difficulties. The flexibility in fine-tuning the charts interactively enables the analysis of a threshold profile (Johnston *et al*. 2003) and helps to study the nature, dynamics and persistence of spatial concentrations in detail. Concentration maps, however, mainly help to indicate which areas deviate from the average situation and should therefore get further attention.

The adoption of the concept in the *Amsterdam City Monitor* confirms its usefulness in urban management as query results are regularly used in the bi-annual publication *The state of the city of Amsterdam* (O&S 2008) and in other publications with municipal partners in the region of Amsterdam.

Conclusions

The mapping methodology presented here contributes to the methodological debate on spatial concentrations because it takes into account specific geographies within wards, the spatial configuration of spatial units, the nature of a variable and the degree of spatial concentration. The development of the city monitor tool is a useful way to utilise small-area data that are registered on a regular basis and make it accessible. By using the monitor, target areas for different policy applications can be identified. Furthermore, because of its thematic focus, it can play a role in the rising debate on residential integration and segregation in urban areas.

For producing maps of spatial concentrations to support urban management, several aspects need to be considered. To begin with, appropriate data from reliable data sources need to be processed in a systematic way and on a regular basis. Secondly, the data should cover a certain level of detail, both in terms of thematic categories and spatial resolution. Moreover, since planners and policymakers are mostly untrained in handling spatial data, a user should be able to easily retrieve relevant information in a systematic way.

The methodology is based on a complete enumeration. Many databases, for example in the medical or environmental disciplines, are compiled from surveys. Future work will deal with identifying concentration areas in survey samples instead of complete enumerations.

Acknowledgements

Notes

1 The city of Amsterdam currently has about 18 000 postcodes. For privacy reasons, data are only made available as an aggregate at units such as the postcode.
2 Access to the monitoring tool: http://mapinfoserver.fmg.uva.nl. Login details can be requested from M.C. Deurloo.

References

Aalbers M B and Deurloo R 2003 Concentrated and condemned? Residential patterns of immigrants from industrial and non-industrial countries in Amsterdam *Housing Theory and Society* 20 197–208

Ackerman W V and Murray A T 2004 Assessing spatial patterns of crime in Lima Ohio *Cities* 21 423–37

Alvanides S and Openshaw S 1999 Zone design for planning and policy analysis in **Stillwell J, Geertman S and Openshaw S** eds *Geographical information and planning: European perspectives* Springer-Verlag, Berlin 299–315

Alvanides S, Openshaw S and MacGill J 2001 Zone design as a spatial analysis tool in **Tate N J and Atkinson P M** eds *Modelling scale in geographical information science* John Wiley and Sons, Chichester 141–57

Amrhein C G 1995 Searching for the elusive aggregation effect: evidence from statistical simulations *Environment and Planning A* 27 105–19

Baud I S A, Sridharan N and Pfeffer K 2008 Mapping urban poverty for local governance in an Indian mega-city: the case of Delhi *Urban Studies* 45 1385–412

Boyle P J and Dunn C E 1991 Redefinition of enumeration district centroids: a test of their accuracy by using Thiessen polygons *Environment and Planning A* 23 1111–19

Burrough P A and McDonnell R A 1998 *Principles of geographical information systems* Oxford University Press, New York

Deurloo M C and De Vos S 2008 Measuring segregation at the micro level: an application of the M measure to multi-ethnic residential neighbourhoods in Amsterdam *Tijdschrift voor Economische en Sociale Geografie* 99 329–47

Deurloo M C and Musterd S 2001 Residential profiles of Surinamese and Moroccans in Amsterdam *Urban Studies* 38 476–85

Deurloo M C, Musterd S and Ostendorf W 1998 *Pockets of poverty* DGVH/NETHUR, Utrecht

Dorling D and Bethan T 2004 *People and places: a 2001 census atlas of the UK* The Policy Press, Bristol

Flowerdew R, Feng Z Q and Manley D 2007 Constructing data zones for Scottish neighbourhood statistics *Computers Environment and Urban Systems* 31 76–90

Gatrell A S, Bailey T C, Diggle P J and Rowlingson B S 1996 Spatial point pattern analysis and its application in geographical epidemiology *Transactions of the Institute of British Geographers New Series* 21 256–74

Getis A and Ord J K 1992 The analysis of spatial association by use of distance statistics *Geographical Analysis* 24 189–206

Gutmann M P and Stern P C eds 2007 *Putting people on the map: protecting confidentiality with linked social-spatial data* The National Academies Press, Washington, DC

Harris R and Longley P A 2001 Data-rich models of the urban environment: RS, GIS and lifestyles in **Hall P** eds *Innovations in GIS vol 8: spatial information and the environment* Taylor and Francis, London 53–76

Harris R and Longley P A 2004 *Targeting clusters of deprivation within cities* in **Stillwell J and Clarke G** eds *Applied GIS and spatial analysis* John Wiley and Sons, West Sussex 89–110

Holt D, Steel D, Tranmer M and Wrigley N 1996 Aggregation and ecological effects in geographically based data *Geographical Analysis* 28 244–61

Johnston R, Voas D and Poulsen M 2003 Measuring spatial concentration: the use of threshold profiles *Environment and Planning B – Planning and Design* 30 3–14

Martin D 1996 An assessment of surface and zonal models of population *International Journal of Geographical Information Systems* 10 973–89

Martin D 2000 Towards the geographies of the 2001 UK Census of Population *Transactions of the Institute of British Geographers* 25 321–32

Martin D 2003 Extending the automated zoning procedure to reconcile incompatible zoning systems *International Journal of Geographical Information Science* 17 181–96

Martin D 2004 Neighbourhoods and area statistics in the post 2001 census era *Area* 36 136–45

Musterd S and Deurloo M C 2002 Unstable immigrant concentration and integration of newcomers *Housing Studies* 17 487–503

Musterd S and Deurloo M C 2006 Amsterdam and the preconditions for a creative knowledge city *Tijdschrift voor economische en sociale geografie* 97 80–94

Musterd S, Ostendorf W and Deurloo M C 2004 Stedelijke context en onveiligheid: gelegenheid en criminaliteit *Beleid en Maatschappij: tijdschrift voor beleid, politiek en maatschappij* 31 163–72

O & S 2008 *The state of the city of Amsterdam III: Developments in participation and life situation* O+S Amsterdam, Amsterdam

Okabe A, Boots B, Sugihara K and Chin S N 2000 *Spatial tessellations: concepts and applications of Voronoi diagrams* John Wiley and Sons, Chichester

Openshaw S 1984a Ecological fallacies and the analysis of areal census data *Environment and Planning A* 16 17–31

Openshaw S 1984b *The modifiable areal unit problem* Geo Books, Norwich

Páez A and Scott D 2005 Spatial statistics for urban analysis: a review of techniques with examples *GeoJournal* 61 53–67

Penninga F, Verbree E, Quak W and Van Oosterom P 2005 Construction of the planar partition postal code map based on cadastral registration *Geoinformatica* 9 181–204

Poulsen M, Johnston R and Forrest J 2011 Using local statistics and neighbourhood classifications to portray ethnic residential segregation: a London example *Environment and Planning B: Planning and Design* 38 636–58

Preparata F P and Shamos M I 1985 *Computational geometry, an introduction* Springer-Verlag, New York

Raper J F, Rhind D W and Shepherd J W 1992 *Postcodes: the new geography* Longman, Harlow

Ratcliffe J H and McCullagh M J 1999 Hotbeds of crime and the search for spatial accuracy *Journal of Geographical Systems* 1 385–98

Rees P eds 1997 *The debate about the geography of the 2001 Census: collected papers from 1995–6* University of Leeds School of Geography, Leeds

van Amersfoort H 1992 Ethnic residential patterns in a welfare state: lessons from Amsterdam, 1970–1990 *Journal of Ethnic and Migration Studies* 18 439–56

van Beurden A and Douven W 1999 Aggregation issues of spatial information in environmental research *International Journal of Geographical Information Science* 13 513–27

Wong D W S, Lasus H and Falk R F 1999 Exploring the variability of segregation index D with scale and zonal systems: an analysis of thirty US cities *Environment and Planning* 31 507–22

Young C, Martin D and Skinner C 2009 Geographically intelligent disclosure control for flexible aggregation of census data *International Journal of Geographical Information Science* 23 457–82