MODELING AND VISUALIZING POPULATION DENSITY FOR THE FIRE & RESCUE SERVICES IN HELSINKI, FINLAND

Jukka Krisp[1], Riikka Henriksson[1], Anki Hilbig[2]

[1]Helsinki University of Technology (HUT), Department of Surveying, Institute of Cartography and Geoinformatics, P.O.Box 1200, FIN 02015 Helsinki, Finland
jukka.krisp@hut.fi, riika.henriksson@hut.fi

[2]Fachhochschule Bochum, Department of Surveying, Lennershofstr. 140, 44801 Bochum, Germany
anke.hilbig@hut.fi

ABSTRACT

The purpose of this paper is to investigate the calculation and representation of population density in the metropolitan area of the Finnish capital Helsinki. The study of the relationships between the population distribution and the identification of high (or low) population density supports the emergency preparedness planning and resource evaluation for the fire & rescue services. The population density information can be included into risk analysis procedures, which are based on the statistical analysis that consider the spatial distribution of the phenomena and its relevance to the emergency services (e.g. age distribution, workspace distribution etc.). Data is available from the Seutu-CD, a product developed by the Helsinki Metropolitan Area Council (YTV). The SeutuCD includes information on the building register data and the inhabitants. Viewing population as a continues surface, by using kernel density estimations and visualizing it as a “landscape” (using the third dimension), supports the understanding of the density distributions, which are of particular importance to rescue services. Slope calculations give representations of channels along rapidly changing population densities and demonstrate the relationships that physical features have on the distribution for residential land use.

1. INTRODUCTION

The Finnish Ministry of the Interior gave out guidelines on a systematic risk assessment practice, which stated that the preparedness in the fire brigades must be based on the municipal risk analysis (Finnish Ministry of Interior 2000). Such a risk assessment should create a basis for setting the target level of the preparedness for emergencies in each municipality (Lonka, 1999). To assist the municipal fire brigades in making the risk assessments a handbook was published in 1994 in co-operation with the Ministry of the Interior and the Federation for Fire Brigade Chiefs in Finland (Alliniemi, 1994).

Finland is not a country threatened by natural disasters like tsunamis or earthquakes, therefore from our point of view the focus lays on the human vulnerability in the built environments. We consider issues like differences between common and rare accidents, commuting of people, and relations between accidents and networks. In general it is of growing importance to investigate and enhance risk models for the fire and rescue services of the Helsinki Metropolitan Area.

Previous research on the visualization of population and incident densities had been carried out within the project. Throughout this paper we use visualization techniques as a means of exploring spatial data in order to detect features of interest contained in them as suggested by Hearnshaw (Hearnshaw and Unwin, 1994). Research by Krisp (Krisp et al., 2005) investigated the possibility to review the significance of a variable in the context of risk analysis. Population density is still considered as a key variable for the vulnerability model. We don’t want to give up on the hypothesis that people cause the main risk of an incident and that population is the most valuable variable in the risk analysis. The visual comparison between the population density data and the incident data shows that the connection between the population density and the incidents reported by fire and rescue services are not as strong as assumed (Krisp et al., 2005). Nevertheless we consider the visual analysis on population densities, which are introduced further on, as an important assistance to the mitigation process for the fire & rescue services.

1.1 Average population density in Finland

The population density is derived by dividing the number of people by the area of interest in relation to the size of that area in square kilometers (or meters). The definition of “high” and “low” population densities needs to be considered because those values differ greatly between different countries. In Finland typically known values are average density calculations for entire cities or counties. It is difficult to compare population density values between the counties because of the size of the area may differ greatly. This influences to the results of the calculations. The average results for selected example areas are calculated and shown in table 1. Those numbers are based on the population data obtained from the Statistics Finland [http://www.stat.fi] in the end of 2003.

Table 1. Average population densities in Finland

Finland	Inhabitants: 5.219.732, area: 304.473 km2	⇒ 17 inhabitants per km2
Uusimaa (county in southern Finland)	Inhabitants: 1.338.180, area: 6.366 km2	⇒ 210 inhabitants per km2
Helsinki (capital of Finland)	Inhabitants: 559.300, area: 186 km2	⇒ 3007 inhabitants per km2

1.2 Population density as a continues surface

Classic choropleth population density maps may have significant problems. Generally as areas are bigger the level of generalization increases (Baxtor, 1976). There can be a tendency either for increasing the size of the units to provide the lower measures for the population densities (Massey and Stephan, 1977), or to decrease the size of the base areas, so that both the peak values and the variance of the population density increases. Furthermore the location boundaries have in most cases no logical relationship to the density properties that are being mapped, as they take administrative boundaries into account (Langford and Unwin, 1994).

Previous research includes work done by Langford & Unwin (1994). They compare the population densities represented in a conventional choropleth map, in which the population density surface is shown as a continues representation. Unlike in this research the data used in their maps is based on an interpretation of a 30m by 30m pixel gird, which provides a density estimate at the zone centroid. In this paper we use the distribution of individuals at real estate center points available in Finnish census data (more in 2.1).

Population density is one instance of categorical variables. Such variables are measured on discrete scales (nominal and ordinal). For categorical variables there exist four field models and the one, which is used in this research, is the grid model. Here the variation is described by determining the variable’s value within each rectangular cell and the spatial variation within cells is ignored. This causes uncertainties because the spatial variations within cells are generalised. Categorical uncertainty should be modeled along with the underlying categories themselves and with their spatial (Zhang and Goodchild, 2002).

1.3 Scope of this paper

A relevant question for the Fire & Rescue services is “Where and when the peaks in the population density occur?” The focus of this paper investigates the methods to model and to visualize the population densities in the metropolitan area of Helsinki. The intention is to use the visualizations techniques as a means of exploring the spatial data on population densities in order to assist the Fire & Rescue services in their planning and preparation procedures. To support this process we additionally investigate the use of slopes, originally developed to analyze morphological phenomena (like elevation models), which are calculated and visualized by using the population density information.

2. MATERIAL AND METHODS - DETERMINING POPULATION DENSITY

Within the modeling process, visualization plays a crucial role to set the parameters for the intended population density map, that is suitable to be used for the visual identification of hot spots. Figure 1 illustrates this process.

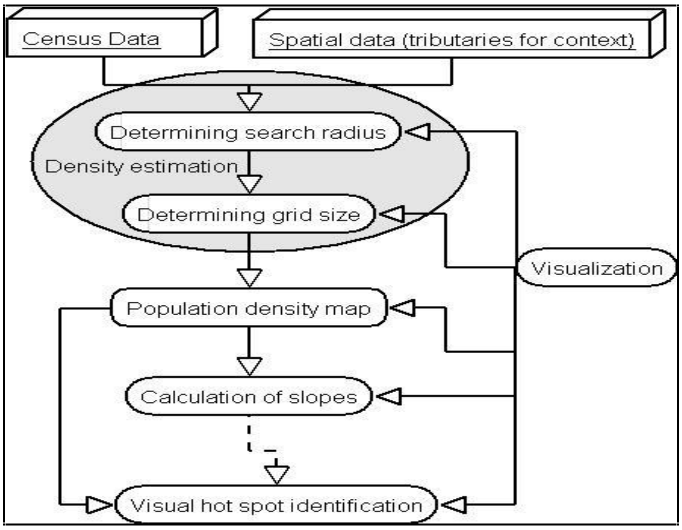


Figure 1. Visualization in the process of density calculations

2.1 Population data

Finland is one of the countries where data collection, digital databases, and information society in general is well developed and organized, but this is no guarantee that information is available and interoperable. Organizational boundaries are clearly visible also in the domain of emergency planning (Jolma et al., 2004).

Generally the unequal quality of spatial data requires individual models and scales for a risk analysis in each community. Municipalities in Finland are obliged to gather register data on their population, buildings and land use plans. Due to the uniform character of the region it is important for the planners and decision makers to have reliable register data on the whole area, irrespective of these municipal boundaries. For this reason that Helsinki Metropolitan Area Council (YTV) has been working since 1997 on the production of a data package, SeutuCD, which covers the whole metropolitan area. It is a data package gathered from the municipalities' registers and other sources. It is updated once a year and includes register data on buildings and land use plans as well as the enterprises and agencies located within the metropolitan area. (YTV, 1999). In addition to register data, SeutuCD includes maps for different scales and metadata software.

Each of the buildings acts as a spatial object, which has a position (X, Y coordinate tuple) and a set of attributes. This enables analysis to be done so that the municipalities' boundaries in metropolitan area are ignored. The attribute we are interested in is the amount of people, which is related to the building. This is partly done to prevent the identification of individual people.

The population information in SeutuCD is originated by the Population Register Center. Their Population Information System contains information for the whole Finland on Finnish citizens and foreigners who are permanently residing in Finland. This data is updated five times a week by the Population Register Center. Completeness omission of the data is said to be 0 %. SeutuCD used in this research is updated in 2003 by YTV but the temporal validity of the data has not been tested.

2.2 Kernel density estimation

Population density is a continues function and in order to present an effective and accurate impression of its distribution, a scheme that recognizes this continuity is needed (Langford and Unwin, 1994). Kernel density estimation allows controlling the degree of smoothing by the search radius (bandwidth) of the kernels.

Given the population data of Helsinki including numbers of humans living in buildings, the probability distribution of these variables at every location needs to be estimated. Replacing each point with a kernel, giving a "spatial meaning" in some sense to it, we obtain, as a sum, a continuous, smooth surface for the variables.

The kernel is defined as a two-dimensional function with two parameters: the shape of the kernel and the bandwidth. In practice only the bandwidth, which represents the search radius, is adjusted. The density λ at each observation point s is estimated by

$$\lambda(s) = \left\{ \sum_{i=1}^{n} K_h (s - s_i) x_i \right\}, s \in U$$

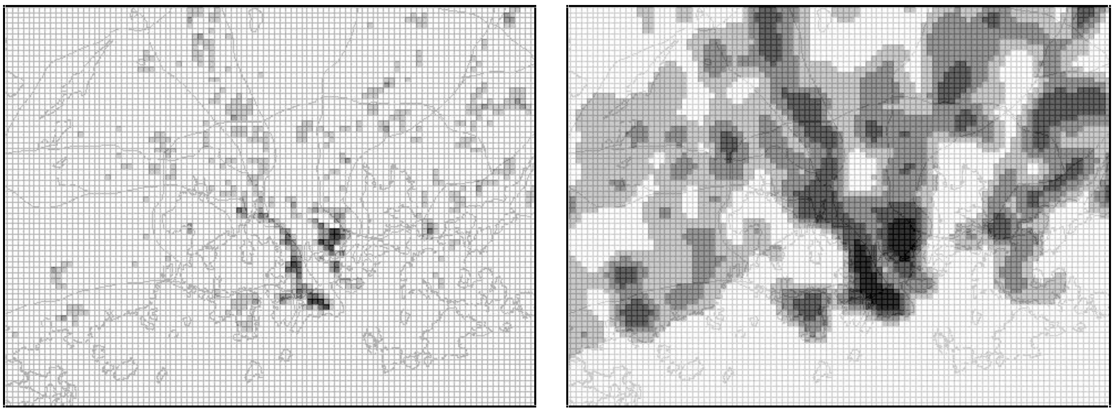
where K is the kernel and h the bandwidth (Silverman, 1986). The choice of the kernel bandwidth strongly influences the resulting estimated density surface. If the bandwidth is too large the estimated densities will be similar everywhere and close to the average population density of the entire study area. When the bandwidth is too small, the surface pattern will be focused on the individual population records. Experimentation is required to derive the optimal bandwidth setting to acquire a satisfactory density surface (O'Sullivan and Unwin, 2003).

2.3 Representation scale

It is important to consider the scale when calculating the population density maps for the fire and rescues services and when setting the search radius / bandwidth for the density estimation. In map of small-scale e.g. 1:100.000 we might aim to have a more general density (see also Figure 6a and 6b). In the city of Helsinki the fire and rescue services aim to use maps with a scale of 1:10.000 and 1:20.000 for their resource planning.

2.4 Kernel search radius & classification

It is important to notice that larger values of the radius parameter produce a smoother, more generalized density raster. Smaller values produce a raster that shows more detail. The following Figure 2 illustrates the different search radius settings with 250m (a.) and 1000m (b.) for a map with a scale of 1 to 100.000 for the same Helsinki area. The grid size in both cases is 250m.



a.) kernel search radius 250m   b.) kernel search radius 1000m

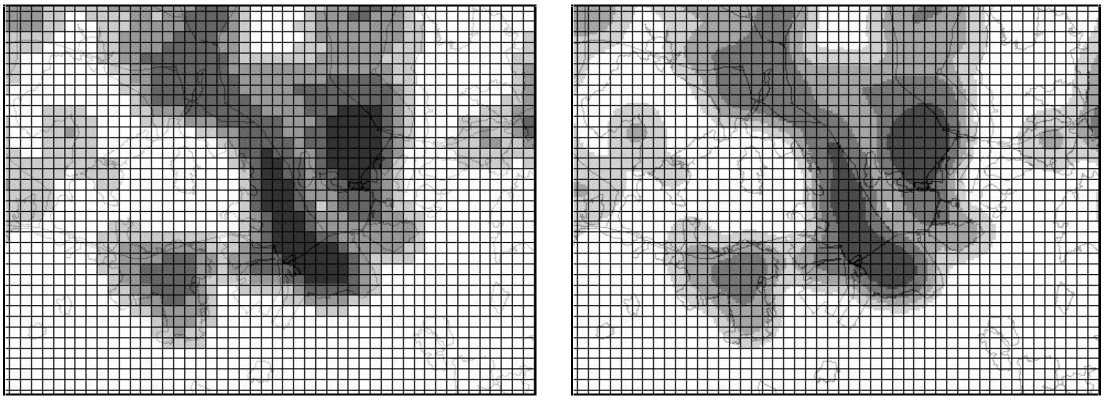
Figure 2. Visual difference in the search radius for the population density with a. 100m and b. 250m

The scale of the map should correspond to the output density. Having a large-scale for example 1 to 10.000 we might want to set a smaller search radius to have a less generalized density surface. For the Helsinki area a density map with a scale of 1 to 100.000 we use a search radius of 1000m.

The classification of the data is an additional problem when representing the population density. Methods used to classify the data can in some cases substantially alter the final appearance of the map. When the data is classified it can introduce an "error", which can be minimized (Unwin, 1981, Jenks, 1977). The number of classes should be limited to a maximum of seven (Gilmartin and Shelton, 1989). In this paper we use generally five classes and the natural breaks (also called Jenks) classification.

2.5 Grid size

The grid size or "cell size" sets size of the raster cells for the output density map. Setting a small grid size provides smoother edges, but requires more computing power. Figure 3 illustrates two different output gird cell settings with 250m (a.) and 62.5m (b.) for a sample map from the Helsinki Metropolitan area. The scale in the example is 1 to 50.000.



a.) grid size 250m   b.) grid size 62.5m

Figure 3. Sample maps showing the grid size with a. 250m and b. 62.5m

When combining the output map visually with a raster grid, it is important to consider this size when setting the output cell size. The output density map and the overlaid gird correspond visually better when the density cell size is the same (or a derivate) with the raster size. For example a density map with a 62.5 m cell size is overlaid with a 250m raster matches well in a visual way.

2.6 Slopes

With the information about the population density it is possible to calculate slopes using the population density values as one would use height values in a Digital Elevation Model (DEM). Applying tools, which are developed for the analysis of surfaces, provide interesting, novel, analytical and visual insights into the population distribution. Wood suggests in previous work to make use of the landscape metaphor in understanding population data (Wood et al., 1999). In the case of population data slopes are the first derivative of the density function. Slopes for population density have the highest value when the changes between the density values are rapid.

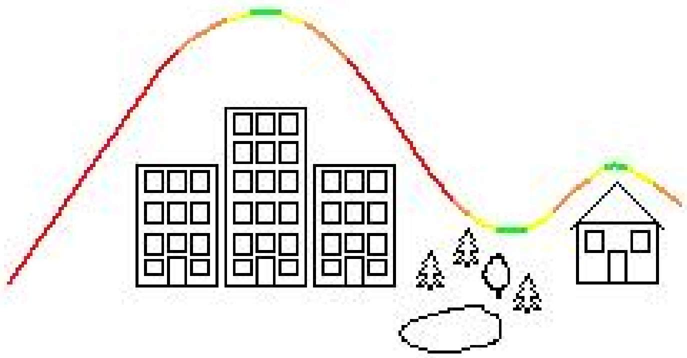


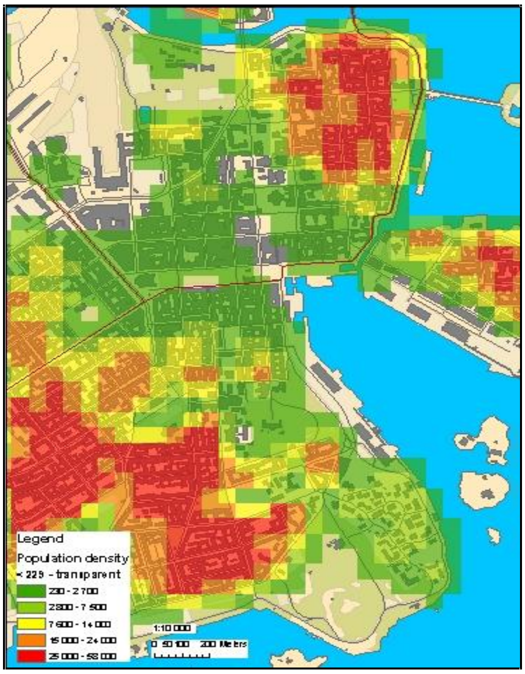
Figure 4. Illustration for density / slope function

The function in Figure 4 describes the population density, dense settled housing areas causes high values in the function. The color of this function shows the slopes. Areas of parks or recreation sites between housing areas have large influence to the density and the slope between these areas is steep (red / dark grey). Shown in Figure 4 the function of population density has maxima and minima with constant slope (green / light grey) when the number of inhabitants stays continuously high or low. It is like the gradient of a bell function. When the differences of the different population density fields are high the slopes are steep (red / dark grey).

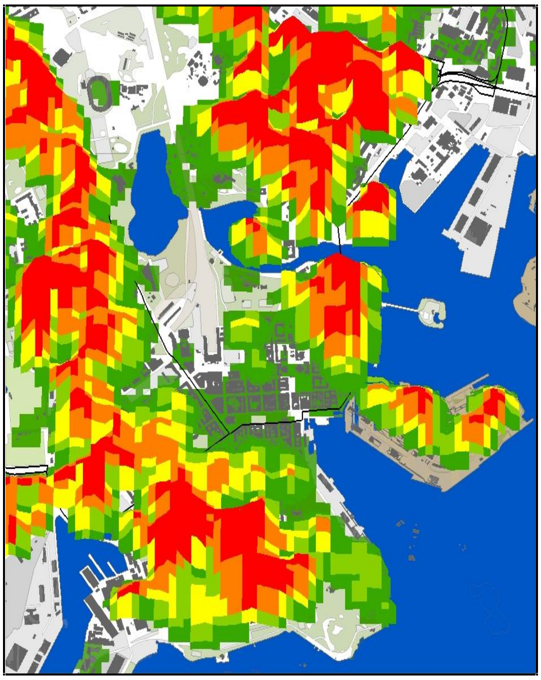
3. RESULTS

Setting the bandwidth for density calculation in the Helsinki area, the value of the population density estimation can rise to improbable big numbers especially in the center of a city when the bandwidth is really small. With a search radius of 10 m, only one building is in the search area and from this building the calculation concludes to a square kilometer. No streets, parks, industrial buildings, other uninhabited areas or smaller buildings are included in the calculation. For example in Kallio, a very densely populated part of Helsinki, the value might increase up to 1.800.000 people per square km. Therefore it is not reasonable to conclude to the average of the population from such a small area, but with this method of changing the search radius to a small diameter, population density differences become visually easier to identify.

For a map with a scale of 1:10.000, a bandwidth with 100m proved to be adequate for the display of population density data as shown in Figure 5 a) and b). In this case the grid cell size with 62.5m. For the visual analysis the population density is displayed with a range of colors from green, light green, yellow, orange and red shows an interval for a specific number of inhabitants. In this map a thirty percent transparency is applied to show the underlying background map.



a.) Population density in 2D (1:10.000)

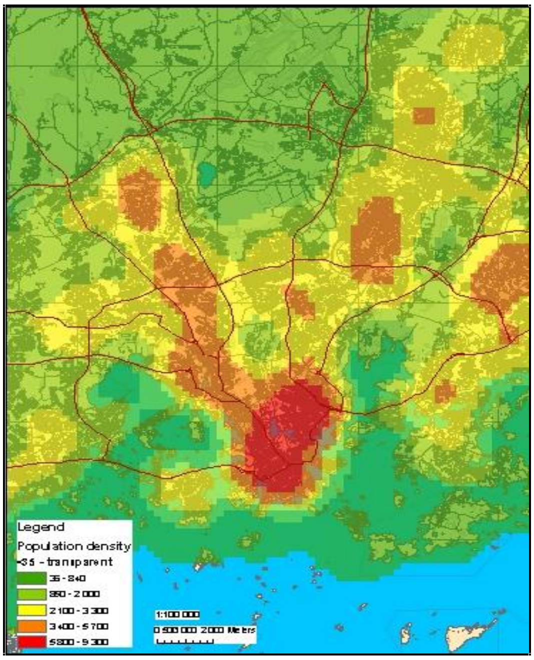


b.) Population density in 3D

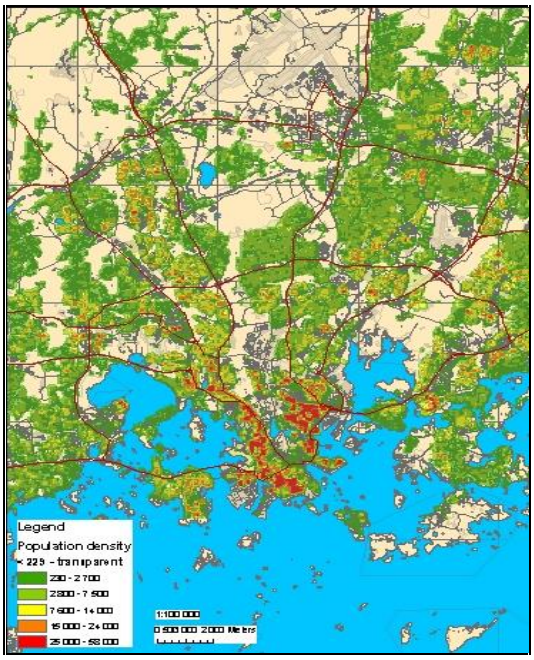
Figure 5. a) Population density map 1:10.000 in a two-dimensional and b) three-dimensional map

In Figure 5b) the densities are visualized in the third dimension by using the same values. In the three dimensional map hills indicate a high while valleys a low density. No shading is applied to this hills, but additionally a color coding from green, indicating a low density, to yellow, orange and red, representing a high density is applied.

Figure 6 a. shows a density map scale of 1:100.000, with search radius 2000m and the cell size 250. In Figure 6 b the search radius is 100m and the cell size 62.5. For a map with a scale of 1:100.000 is might be useful to select a bigger search radius to get a more generalized density surface. In this scale the option is more suitable.



a.) Sample population density map, scale 1:100.000, search radius 2000m, grid size 250m

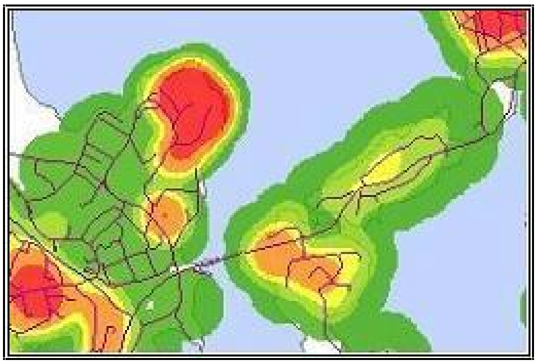


b.) Sample population density map, scale 1:100.000, search radius 100m, grid size 62.5m

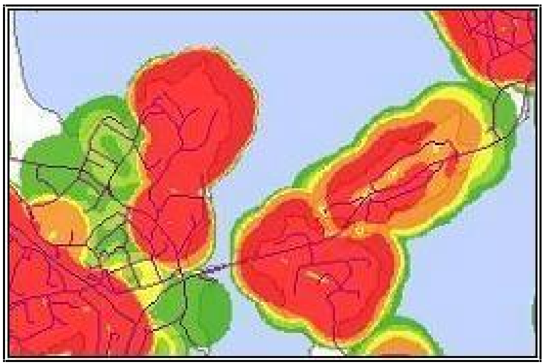
Figure 6. Population density scale 1:100.000

3.1 Using the density surface to calculate slopes

Figure 7 shows the population density (a.) and the slopes (b.) based on the population density for the same area - Otaniemi in Espoo and the two islands Lethisaari and Kuusisaari in Helsinki.



a.) Population Density



b.) Slopes for Population Density

Figure 7. a.) Population Density and b.) slopes based on the population density

It is not necessary to have the highest values in the population density to get the highest values in the slopes and even if there is the highest value in the population density map there could be lower values in the slopes.

4. CONCLUSIONS

The population density is a continues function and to illustrate this, we need a representation that recognizes this continuity. Applying a kernel density estimate gives a smooth continues surface. Assuming that the census data is accurate to a certain extent and that the population is generally sleeping at night in the place where they are registered, this representation has the advantage to provide a continuous surface to the reader, which shows an accurate distribution of the nighttime population densities. These results are open for further interpretation and can be used by the Helsinki Fire and Rescue services in a risk model. It is only one step in the often very complex analysis.

The census data is useful to calculate population densities. The disadvantage here is that we use SeutuCD, which is updated in 2003 and not original data, which is updated five times a week by the Population Register Center. Another disadvantage is that the accuracy of the categorical variable, population density, is not known.

The results can be displayed in various ways. The visualization classified from red to green proves to be useful in visually identifying densely or sparsely populated areas. Using three-dimensional maps can result in an appealing, maybe flashy visualization. However, the usefulness and usability of these combined three-dimensional maps needs to be proven.

Applying slopes somewhat demonstrates the control of physical features on the distribution of residential areas. The calculation and visualization of slopes on population densities shows “channels” between the steep changes in the density of residents. These channels might in some cases assist the planning of emergency routes for fire & rescue services.

5. FURTHER RESEARCH

In this case study we show the population density maps, which present the population densities during the nighttime. During the daytime the values usually change completely effected by commuting to work, shopping, visiting big music or sports events, etc. Further research has to consider individual records in the data, the time of each record and the time of the population density calculations. The integration of a time variable (daytimes, nighttimes etc.) seems to be essential when relating the population density to incident densities. Furthermore we should aim to integrate data from different resources and a longer time period.

Further research will use base station records of cellular phones to calculate the amount of phones locked onto a particular base station at a particular moment in time or over a certain period. Calculating this over a long period and depending on the density of base stations and the amount of cell phones in the population, this might give accurate population density results. This data has to be aggregated, because of the privacy law it would be difficult to have such data which would enable us to follow each individual for example during one day.

The SeutuCD data has been collected since year 1997. It is important to point out that register data in the SeutuCD is always only a snapshot from the history. Because of this it would be interesting and also important to know if there could be found some kind of trends considering especially the population density. Also the urban area development plans could be studied to find more information from the future situation, e.g. how the metropolitan area is going to change during the upcoming years. By doing such data analysis we could take into account the special characteristics of the areas such as those that are populated mostly by elderly people and we could also based our spatial analysis on the forecast information and see if the results are different.

6. REFERENCES

(Finnish Ministry of Interior 2000) Kunnan Pelastustoimen Palvelutasoa Koskevat Päätökset, in Dnro SM-1999-000939/Tu-31.

Alliniemi, J. (1994) Uhat ja mahdollisuudet – tapa tutkia onnettomuuksia ja niiden vaikutuksia.

Baxtor, R. S. (1976) Some methodological issues in computer drawn maps, The Cartographic Journal**,** 13**,** 145-155.

Gilmartin, P. and Shelton, E. (1989) Choropleth maps on high resolution CRTs: the effects of number of classes and hue on communication, Cartographica**,** 26**,** 40-52.

Hearnshaw, H. M. and Unwin, D. J. (1994) Visualization in geographic information systems.

Jenks, G. F. (1977) Optimal data classification for choropleth maps, Dep. of Geography, University of Kansas.

Jolma, A., Virrantaus, K. and Krisp, J. M. (2004) Review of geoinformatical methods for a city safety information system, in Proceedings 9th International Symposium on Environmental Software Systems Harrisonburg.

Krisp, J. M., Jolma, A. and Virrantaus, K. (2005) Using explorative spatial analysis to improve fire and rescue services in Helsinki, Finland, in Proceedings First International Symposium on Geo-information for Disaster ManagementSpringer, Delft, The Netherlands, pp. 1282-1296.

Langford, M. and Unwin, D. J. (1994) Generating and mapping population density surfaces within a geographical information system, The Cartographic Journal**,** 31.

Lonka, H. (1999) Risk Assessment Procedures Used in the Field of Civil Protection and Rescue Services in Different European Union Countries and Norway, SYKE, Helsinki.

Massey, D. S. and Stephan, G. E. (1977) The size-density hypothesis in Greast Britain: analysis of a diviant case, Demography**,** 14**,** 351-361.

O'Sullivan, D. and Unwin, D. J. (2003) Geographic information analysis, New Jersey.

Silverman, B. W. (1986) Density estimation for statistics and data analysis**,** Chapman and Hall, London.

Unwin, D. J. (1981) Introductory spatial analysis, London.

Wood, J. D., Fisher, P. F., Dykes, J. A., Unwin, D. J. and Stynes, K. (1999) The use of the landscape metaphor in understanding population data, Environment an Planning B: Planning and Design**,** 26**,** 281-295.

YTV (1999) Establishments in the Metropolitan Area 1999, Helsinki Metropolitan Area Council (YTV), Helsinki.

Zhang, J. and Goodchild, M. F. (2002) Uncertainty in Geographical Information**,** Taylor & Francis, New York.