MODELLING OF POPULATION DYNAMICS:
GIS VERSUS REMOTE SENSING – A CASE STUDY FOR ISTANBUL

Julia Kubanek, Martin Kappas

University of Göttingen, Germany[1]

Eike – Marie Nolte, Friedemann Wenzel

Karlsruhe Institute of Technology, Germany, [2]

Hannes Taubenböck

German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Germany, [3]

Keywords

Population estimation, building inventory, emergency planning, remote sensing, Istanbul

Abstract

Over the last decades, the rapid growth of the world population has led to a large number of emerging megacities. The 1999 Izmit (Turkey) earthquake is a striking example of the impact of natural hazards on megacities. On August 17 1999, a magnitude 7.6 earthquake struck the area of Izmit in Turkey, causing about 20.000 fatalities and US$6.5 billion economic loss. The probability of a magnitude 7 earthquake striking Istanbul within the next 30 years ranges between 30% to 70%. In order to reduce the impact of natural hazards on human lives, emergency management plans are essential. The development of these plans strongly relies on up-to-date population and inventory data. However, existing techniques for population data generation do not meet the requirements of today's dynamic cities. In this context remote sensing has become an important source of information in the last years. However, a rational discourse on the suitability of remote sensing for urban applications is still missing. In this study a quantitative evaluation of the suitability of IKONOS imagery for population modelling using the district of Zeytinburnu (Istanbul, Turkey) is conducted. The results reveal that IKONOS images can be used for complementing existing inventory data set. The automated extraction of single buildings was identified as the major source of error in the population estimation. Further advantages and limitations such as the associated costs are discussed in this present paper.

[1] Julia Kubanek, Center for Disaster Management and Risk Reduction Technology, Karlsruhe Institute of Technology, Hertz Str. 16, 76187 Karlsruhe; Julia.Kubanek@web.de

[2] Eike-Marie Nolte, Center for Disaster Management and Risk Reduction Technology, Karlsruhe Institute of Technology, Hertz Str. 16, 76187 Karlsruhe; eike-marie.nolte@kit.edu

[3] Hannes Taubenböck, German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Münchner Str. 20, 82234 Wessling; hannes.taubenboeck@dlr.de

Introduction

Over the last decades, the rapid growth of the world population has led to dynamic and complex urbanization on a global scale. By 2050, it is predicted that 6 billion people – by then about 70% of the world's population - will reside in urban areas (UNITED NATIONS, 2008). Cities as a habitat play an important role in human life but are often prone to natural hazards: the recent Haiti earthquake has just illustrated the disastrous consequences of an earthquake striking a densely populated area (www.zki.dlr.de).

In the immediate aftermath of a major disaster, the priorities are undoubtedly medical and rescue needs (Coburn & Spence, 1992). The effectiveness of emergency planning strongly depends on reliable information on a cities population and its dynamics. Existing techniques for population data generation do not match the requirements of today’s cities. Additionally, inventory data – especially in developing countries – might be incorrect, outdated or not available at all. In this context, satellite imagery has been increasingly used as an independent and up-to-date source of information over the last years. However, the apprehensible excitement and enthusiasm by the scientific community in the advent of high to very high resolution satellite imagery such as IKONOS and Quickbird, lead to the impression of very high resolution images being the solution to all data generation problems. A rational discourse on the suitability of satellite images as a source of information for urban applications is still missing. The study presented in this paper aims at constituting the first step towards a quantitative evaluation of the suitability of high resolution satellite imagery (IKONOS) for urban population modelling. Based on a statistical comparison between population distribution modelled using satellite images and population distribution modelled using GIS and secondary sources, the limitations and advantages of IKONOS imagery for population modelling are discussed. As a test site for this case study the district of Zeytinburnu in Istanbul (Turkey) has been selected.

Theory and Method

In the past, censuses, surveys, and official registration systems served as the main source for accessing information on population. In many cases, traditionally applied techniques are too time- and cost-intensive to cope with today’s population dynamics and with the increasing spatial extent of urban areas. In addition, census data are usually conducted on a decennial basis and afflicted with the problem of the so called “10-year gap“. Existing population data are often only available on aggregated level and not suitable for developing emergency plans (Hardin, Jackson, & Shumway, 2007; Harvey, 2002; Newell, 1988).

To overcome problems like missing up-to-dateness and unsuitable spatial resolution, the importance of remote sensing for information extraction in urban areas is increasing. A number of previous studies attempt to investigate the capabilities of high resolution satellite imagery for population estimation. For example, Souza, Pereira, & Kurkdjian (2003) analyzed the potential of IKONOS images for estimating population for Sao Jose dos Campus (Brazil) using GIS technologies. Liu, Clarke, & Herold (2006) explored the possible correlation between the population density and textures in IKONOS images for a study site in Santa Barbara County. Further studies using IKONOS images for urban information extraction were conducted by Taubenböck, Roth, & Dech (2007) and Taubenböck, Wurm, Setiadi, Gebert, Roth, Strunz, Birkmann, & Dech (2009). In these studies IKONOS images were used as the basis for classifying patterns of urban morphology and inferring socioeconomic parameters. The main intention of Taubenböck (2008) was to analyse the capabilities of remote sensing to assess earthquake risk and vulnerability in the megacity Istanbul. The vulnerability assessment was conducted for the district of Zeytinburnu (Istanbul) including the development of the building inventory categories which are based on physical parameters such as building density and height. These parameters were then used to estimate the population density and to model its spatial distribution. In the following section, the population estimation methodologies are explained in detail as the results of Taubenböck (2007) and Taubenböck (2008) are used for the statistical comparison in this study.

For the population estimation, two approaches using different external data sources were developed. The first approach included a top-down methodology. First, a combination of physical city characteristics such as land use and building density was used to delineate 24 physically homogeneous structural zones. In order to calculate the residential living space for each zone, the residential built-up area extracted from IKONOS image was multiplied by the number of floors calculated from the remotely estimated building height. Using the district population, the average number of inhabitants for each zone was calculated. Employing other land use categories such as commercial areas extracted from IKONOS image, the residential day and night-time population was calculated. The second approach uses a bottom-up methodology based on information about the inhabitants of 50 sample buildings obtained in a field survey conducted by Taubenböck (2007). Using this information, for each urban zone an average number of inhabitants per building was calculated. In combination with the estimated residential area in a zone, the number of inhabitants per square meter was calculated (Taubenböck, 2007).

The methodologies for population estimation presented by Taubenböck (2007) involve the development of several intermediate data sets which serve as an input for the following methodological steps. As the overall uncertainty and accuracy of the population estimates strongly depends on the level of accuracy of the input data, all intermediate results are quantitatively compared to the results from the GIS-based approach presented in this paper.

Development of an up-to-date single building inventory for Zeytinburnu

The population estimation by Taubenböck (2007) involves the extraction of residential areas in Zeytinburnu. In order to be able to analyse the accuracy of this base data set, an up-to-date single building inventory for Zeytinburnu is developed. A digital building outline data set developed in the *Zeytinburnu Pilot Project* forms the basis in this first part. This data set is primarily complemented using building outlines from the *Earthquake Master Plan for Istanbul*. It is important to note that the building inventory from these two studies showed significant deviations. To eliminate the ambiguities and to ensure the developed inventory data set only includes presently existing buildings, the two above mentioned data sets are compared and complemented by other data sets. Digital administrative boundaries for the district of Zeytinburnu and for the city of Istanbul are available from the study *A Disaster Prevention/Mitigation Basic Plan in Istanbul including Seismic Microzonation in the Republic of Turkey*. For the comparison, the data from Taubenböck (2007) are clipped to the same administrative boundary as used in this study.

Spatial assessment of occupancy categories in Zeytinburnu

Based on the up-to-date single building inventory, the second step includes the assessment of the spatial distribution of different occupancy categories for every single building in Zeytinburnu. The main data sources are the *Zeytinburnu Pilot Project* and the *Earthquake Masterplan for Istanbul*. The resulting data set displays the spatial distribution of different building occupancy categories in Zeytinburnu. Using the building inventory developed with data of Taubenböck (2007), a data set including the occupancy categories of Zeytinburnu was developed. For a better comparison of the two data sets generated in this section, the occupancy categories are limited to the following categories: (1) commercial, (2) residential-mixed, and (3) rest (see figure 1 and 2).

Night-time population distribution modelling for Zeytinburnu

Besides building occupancy, the number of floors is an essential parameter for population modelling. Therefore, the floor numbers included in the *Zeytinburnu Pilot Project* are assigned to the GIS-based building inventory. For the data set provided by Taubenböck (2007) only categorical floor number information is available. In this study, the night-time population modelling is based on the generalizing assumption that most people reside at home at night-time. In a first step, combining residential occupancy and usable space gives the available residential living space.

Figure 1: Distribution of simplified occupancy categories in Zeytinburnu on single building level based on Kubanek (2010)

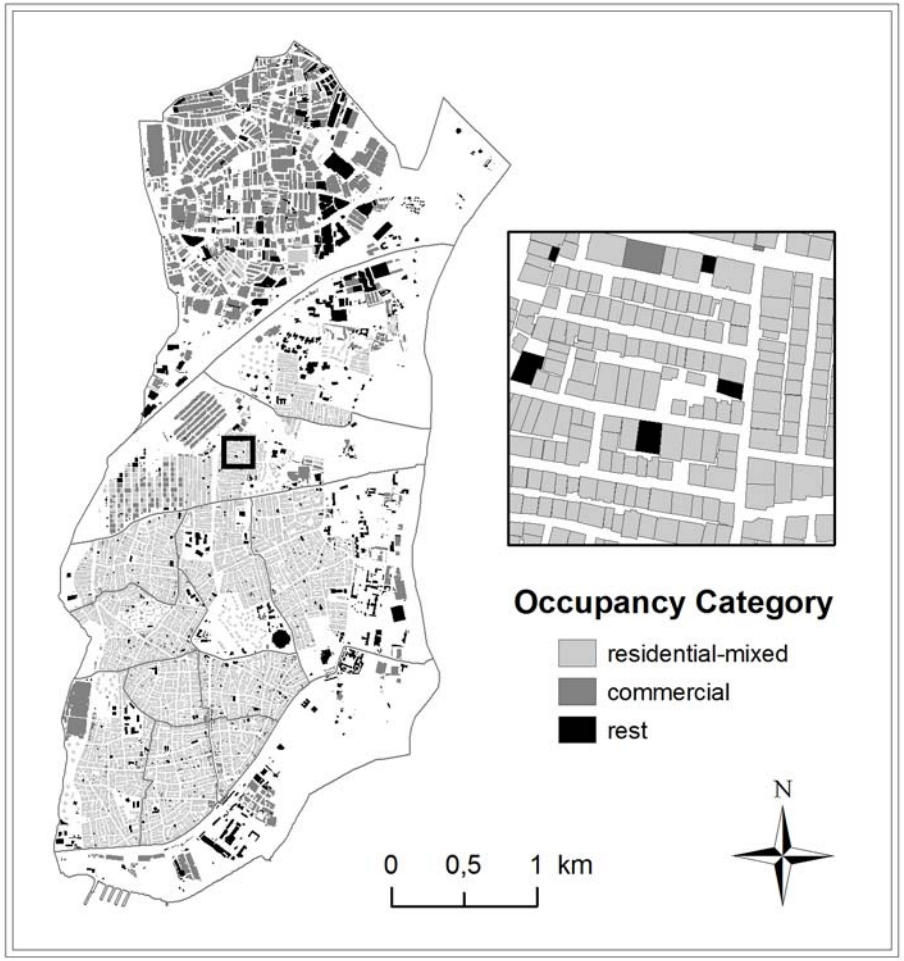
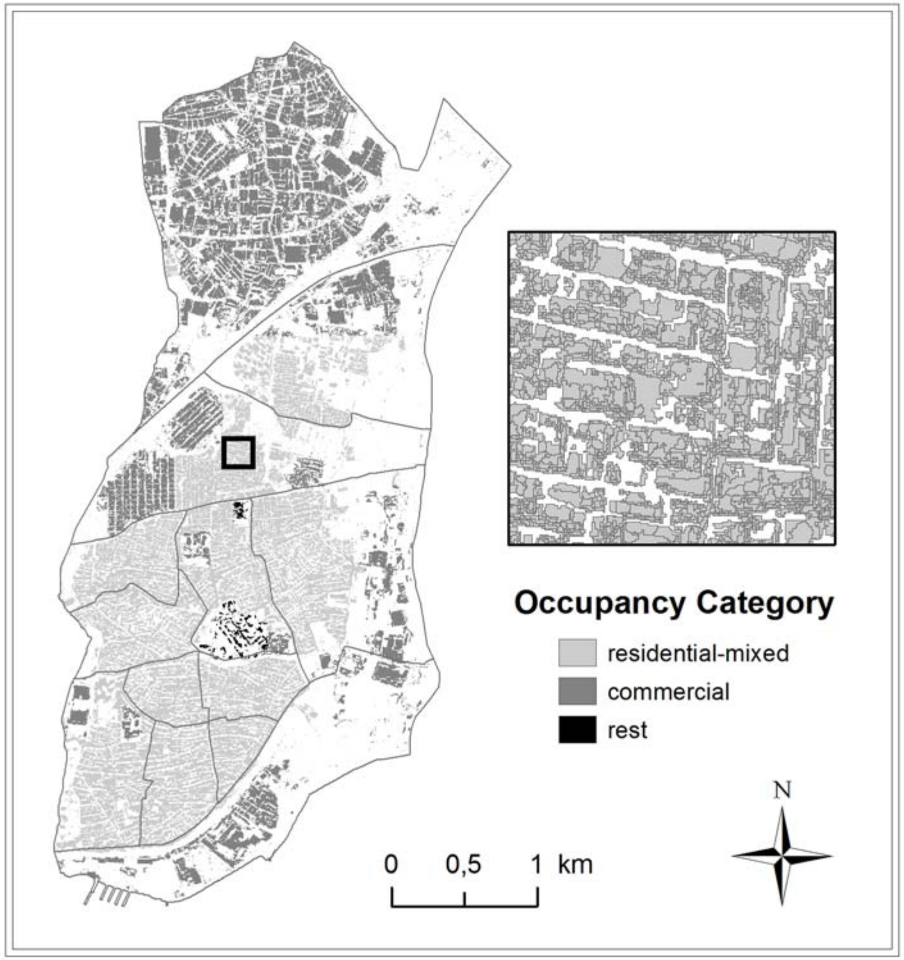


Figure 2: Distribution of simplified occupancy categories in Zeytinburnu based on Taubenböck (2008)



In a second step using a top-down approach, the average residential living space (m^2) per person is estimated for each Mahalle (sub-district) by dividing the total residential living space by the total Mahalle population provided by the *2000 Census*. To estimate how many people live in a single building, the living space of every single residential building was divided by the averaged living space per person. The same methodology was applied to the data from Taubenböck (2007) in order to ensure the comparability between the GIS-based and the remote sensing based data set.

<u>Cost analysis</u>

Population data collection is a very time- and cost-intensive task. Taubenböck (2009) emphasised that the cost and effort can be significantly reduced if remote sensing technique is integrated in the data collection scheme. However, the costs associated with employing cutting-edge technology are still considerably high. In addition, census data generally need to be purchased from census bureau or other official authorities and are only freely available on aggregated level. For the methodology introduced in this paper, the following costs need to be considered:

(1) Satellite imagery
For the presented analysis, a high-resolution satellite imagery (IKONOS) was used. One km^2 of IKONOS imagery with the accuracy required costs US$33. The total area of the Zeytiburnu district is 11.5 km^2. In order to implement the presented techniques for the whole of Istanbul with an area of 1830 km^2, the corresponding image would cost approx. US$45540 (E-geos, 2009).

(2) Software license
The license fees for software packages used have also to be taken into account. The market leader for object-oriented image analysis software Definiens used by Taubenböck (2007) charges approx. US$8000 for their software package Definiens Developer. For data analysis Arctic 9.3 was used in this study. ESRI charges US$3325 for a multi-user educational license and US$670 for a basic single user license plus US$400 for extensions such as spatial analyst.

(3) Data/ Data acquisition
The data used in this study were freely available on the courtesy of the project partners. In general data acquisition costs can be substantial. As mentioned in the introduction, sample surveys are very expensive tasks due to the high number of people involved. As validation data such as census data might be not available - especially in many developing countries, the validation relies on data from time- and cost-consuming surveys.

Particularly less developed countries or NGOs might not be able to afford the costs associated with the presented methodology. To expand the usage of remote sensing technology for population modelling and data generation in general, the use of open source software needs to be encouraged. In addition, the distribution of freely available high resolution satellite imagery needs to be extended. For example, initiatives such as the International Charter: Space and Major Disasters could provide data also in the pre-event phase to foster the generation of reliable data set – not exclusively but also for disaster management measures.

Projects like the Global Earthquake Model (GEM) make significant headway in this direction. GEM aims at providing the first global, open source model for seismic risk assessment at a national and regional scale. This model also includes a model for data generation incorporating the use of satellite imagery and open source GIS software. The authors of this present paper and other researchers associated with the Center for Disaster Management and Risk Reduction Technology (www.cedim.de) at the Karlsruhe Institute of Technology (www.kit.edu) and the German Aerospace Center (www.dlr.de/en) are involved in the data generation work within GEM.

Global applicability

Global applicability of data generation techniques depends on the global availability of the required input data. For the presented approach for population modelling, the following five data sets are essential: (1) up-to-date single building inventory, (2) administrative boundaries of sub-districts or unit of analysis, (3) occupancy of single buildings, (4) height information of single buildings, and (5) population data aggregated on sub-district level.

Results

The presented comparison proves that high resolution satellite images can be successfully employed for urban parameter extraction. However, some constraints were identified in this study. Figure 3 shows the results for extracted built-up area from IKONOS by Taubenböck (2007) and the built-up area calculated in this study. From figure 3 it becomes obvious that Taubenböck (2007) overestimated the built-up area of each Mahalle by approximately 10% compared to the GIS-based inventory. The comparison of the simplified occupancy categories revealed the same deviation. As buildings like other objects are identified from satellite images through their reflective signature, the observed deviations are mainly caused by similar spectral characteristics of land cover categories like built-up area and its surroundings such as streets or bare ground. The similarity strongly depends on the roofing material of the buildings. For example, concrete building with tiled roofs are easier to distinguish from tarmac than concrete roof with the same greyish colour. In the second step of this study, simplified occupancy categories on Mahalle level were compared. Taubenböck (2007) subdivided the buildings into four classes: residential-mixed, commercial, military and hospital; in this study (in the following referred to as Kubanek (2010)) the following classes were used: industrial, commercial, residential, other, mixed, depot, service, under construction, vacant, and unknown. Due to the very broad categorization scheme applied by Taubenböck (2007), almost no buildings are assigned to a "rest" category (see figure 1 & 2). This comparison demonstrated the capabilities of high resolution satellite imagery to delineate occupancy categories. The determined percentage of occupancy types shows the same trend for some Mahalles. For example, Kubanek (2010) identified a large share of residential living space for some Mahalles (figure 1 & 2); the occupancy data generated by Taubenböck (2007) display the same trend. The same can be observed for Mahalles with little residential living space.

Figure 3: Comparison of built-up area in % generated by Kubanek (2010) using a GIS software and based on Taubenböck (2008) extracted from IKONOS imagery

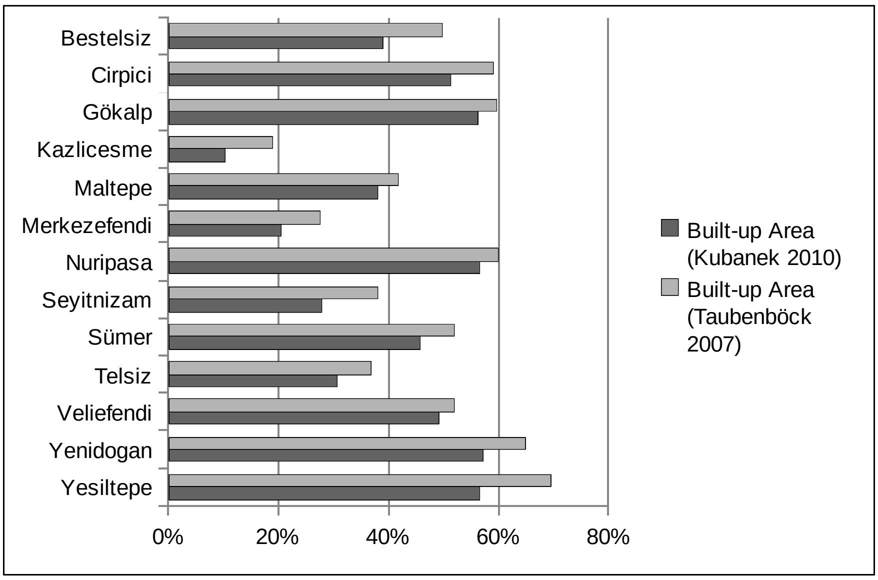


Figure 4: Distribution of night-time population in Zeytinburnu on single building level based on Kubanek (2010)

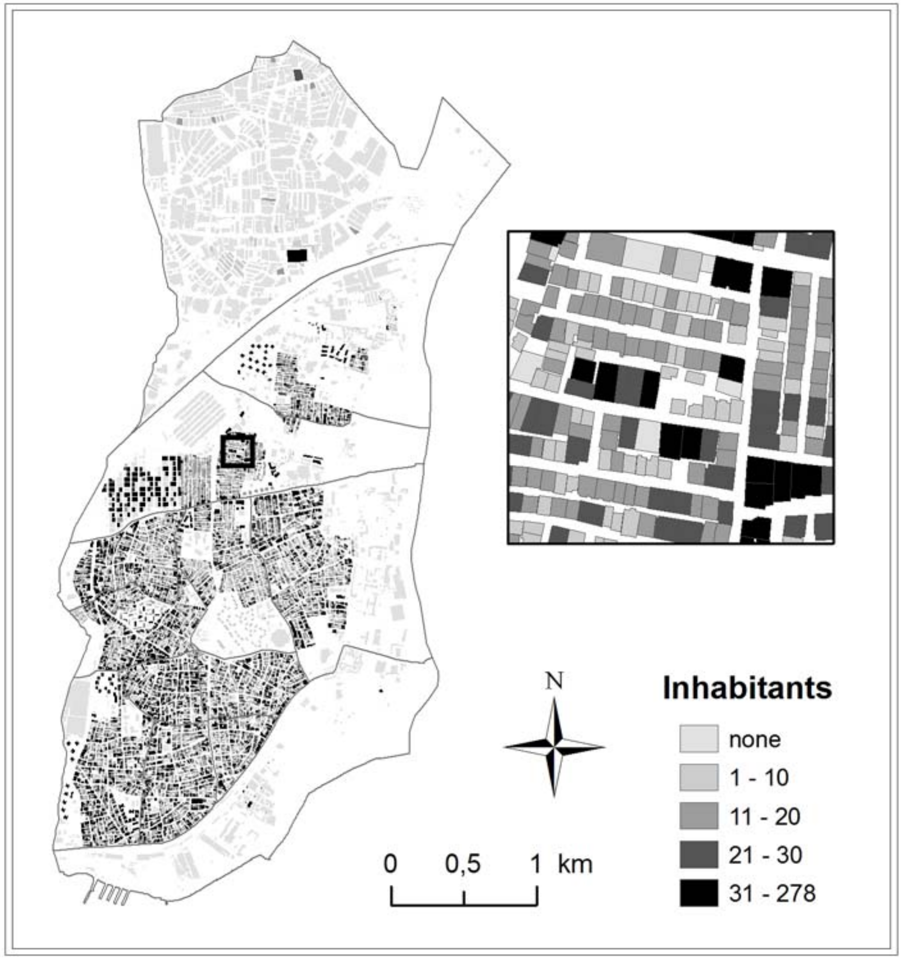
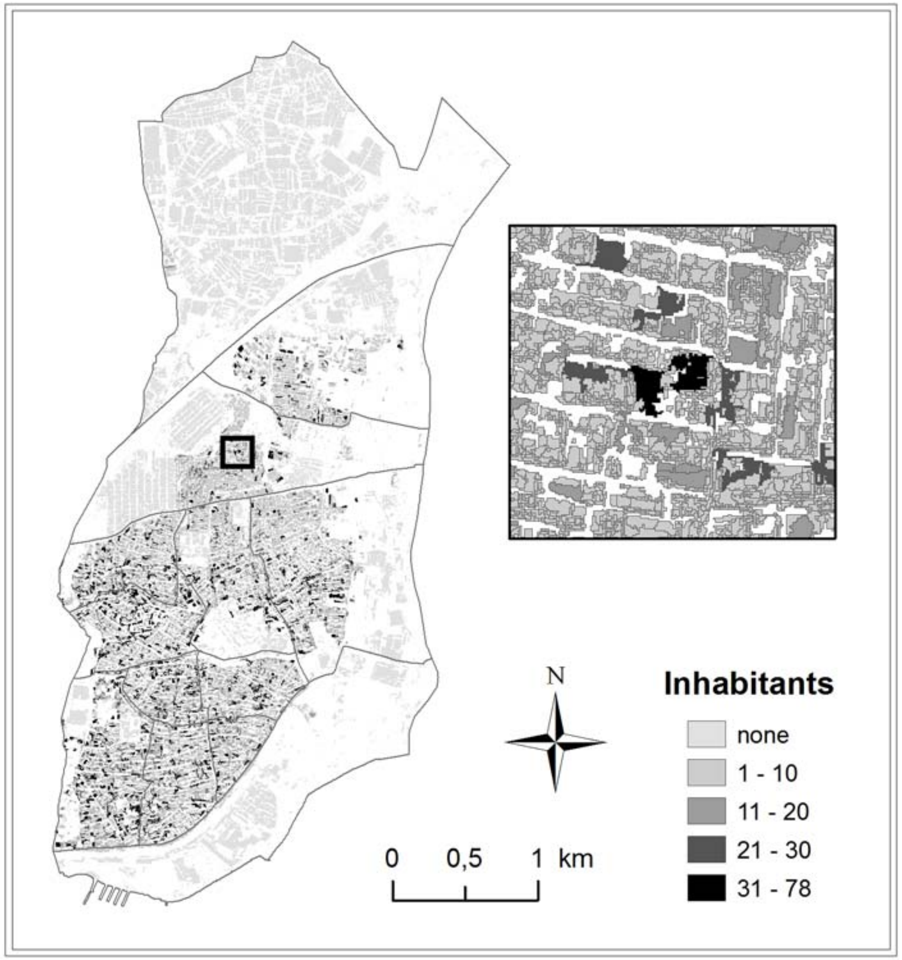


Figure 5: Distribution of night-time population in Zeytinburnu based on Taubenböck (2008)



The comparison of the population distribution in Zeytinburnu constitutes the last step of this analysis. Figure 4 shows the GIS-based population distribution in Zeytinburnu modelled using data from Kubanek (2010), figure 5 shows population distribution based on data from Taubenböck (2007). In general, the number of inhabitants per building is higher in the southwestern part of Zeytinburnu, in the so called "residential districts". A significant lower night-time population is modelled for the industrial districts in the northern part of Zeytinburnu. For a more detailed comparison, a small section of Zeytinburnu from both data sets is displayed in figures 4 and 5. It becomes obvious that the population modelling by Kubanek (2010) is based on realistic single building footprints whereas the built-up area extracted from IKONOS by Taubenböck (2007) does not depict single buildings. This is the main limitation for utilising very high resolution optical imagery for per-building population modelling as the extraction of single buildings proofed to be very difficult for urban areas.

Figure 6: Residential living space per Mahalle in square kilometer. Comparison between Taubenböck (2008) and Kubanek (2010)

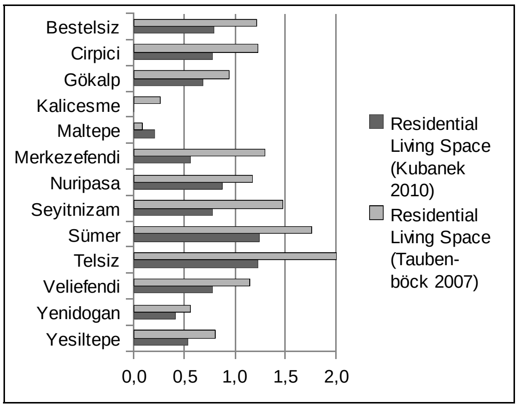
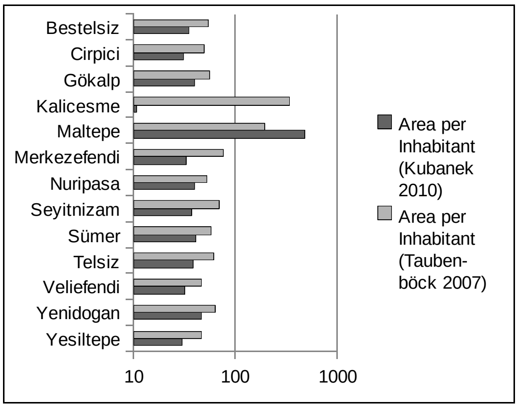


Figure 7: Area per inhabitant on Mahalle level in square meter. Comparison between Taubenböck (2008) and Kubanek (2010)



Discussion

This study focused on the quantitative assessment of the capabilities and limitations of optical high resolution satellite imagery for population modelling in urban areas. The findings confirm that these kind of images are a valuable source of information for modelling population distribution on sub-district level. However, the overestimation of the built-up area due to the limitations of optical images (figure 3) constitutes a major source of error as it results in a constant overestimation of the residential area per Mahalle and consequently in an overestimation of the averaged residential living space per person on Mahalle level (figure 6). Considering an average overestimation of 10%, the question arises whether this methodology is suitable to provide information for risk managers and urban planners. From the present study, we conclude that modelling population distribution on building level is feasible using automatically extracted building footprints only with a certain loss of accuracy. Satellite images prove to be a valuable source to supplement existing building inventory data.

Another source of error is the estimation of the floor numbers from high resolution optical satellite images. As the floor number is estimated from building height which is derived from building shadows, the floor numbers can only be provided as categorical data. Inaccuracy also arises from the use of mixed occupancy categories. Buildings with different occupancies on different floors were classified as mixed. In order to calculate the residential living space in a building, the footprint area should be multiplied with the number of residential floors only and not with the total number of floors. As this presented study is ongoing research, the next step will include an analysis of the buildings to obtain a more precise distribution of the residential floors and thus of the population. A further methodological step may include the consideration of employment and commuting population to model the daytime population. Different approaches on the estimation of building height using remotely sensed data sets have already been completed. Using LaserScan data in combination with IKONOS data showed a significant increase in the building height and floor estimation (Wurm, Taubenböck, Roth, & Dech, 2009).

From an emergency planning perspective, the overestimation in all steps of the analysis is of minor importance as only with underestimation might have severe consequences in case of a disaster. Although population estimates per building level using remote sensing might be not sufficiently accurate for the coordination of rescue teams and fire brigade, the generated population data can provide a valuable insight to the distribution of potentially affected people on district level in case of a disaster.

Although the study site was selected because of the high earthquake hazard, the presented methodology is not limited to densely populated areas and earthquake risk. The knowledge of the spatial distribution of the inhabitants of an entire city including scarcely populated outskirts also plays a critical role for risk reduction and mitigation strategies for other hazards such as tsunamis, floods, or volcanic eruptions. Another important issue related to population modelling is the identification of crowded places and the assessment of the potential for human crowding. This information is for example essential for determining places at risk from terrorist attacks. In addition, the health sector is interested in building level population data, for example for pandemic prevention or for identifying people at risk concerning the contamination e.g. of drinking water.

References

Coburn, A. & Spence, R. (2002). Earthquake Protection (2nd edition), Wiley, England.

E-geos (2009). GeoEye – I & IKONOS, Introductory Price List, July 2009. http://www.eurimage.com/products/ge_docs/ik-ge_price_list.pdf. Last Accessed 09 May 2010.

Hardin, P. J., Jackson, M. W., & Shumway, J. M. (2007). Intraurban Population Estimation Using Remotely Sensed Imagery. In R. R. Jensen et al. (Eds.). Geo-Spatial Technologies in Urban Environments. Policy, Practice and Pixels, Second Edition, p. 47-92. Berlin, Heidelberg, Germany: Springer.

Harvey, J. (2002). Estimating Census District Populations from Satellite Imagery: Some Approaches and Limitations. Int. Journal of Remote Sensing, Vol. 23, No. 10, p. 2071-2095. Taylor & Francis Ltd, United Kingdom.

Kubanek, J. (2010). Comparison of GIS-based and High Resolution Satellite Imagery Population Modeling – A Case Study for Istanbul. Diploma thesis. Work in Progress.

Liu, X., Clarke, K., & Herold, M. (2006). Population Density and Image Texture: A Comparison Study. Photogrammetric Engineering & Remote Sensing, Vol. 72, No. 2, pp. 187-196. American Society for Photogrammetry and Remote Sensing, USA.

Newell, C. (1988). Methods and Models in Demography, The Guilford Press, New York, USA.

Souza, I., Pereira, M. N., & Kurkdjian, M. (2003). Evaluation of High Resolution Satellite Images for Urban Population Estimation. Instituto National de Pesquisas Espacias. Aceito Para Apresentacao No 3° Simposio de Sensoriamento Remoto de Areas Urbanas, Istambul, 11-13 de junho, 2002 – Turquia. Sao Jose dos Campos, Brazil.

Taubenböck, H., Roth, A., & Dech, S. (2007). Linking Structural Urban Characteristics Derived From High Resolution Satellite Data to Population Distribution. In V. Coors et al. (Eds.). Urban and Regional Data Management. p. 35-45. Taylor & Francis Group, London, United Kingdom.

Taubenböck, H. (2008). Vulnerabilitätsabschätzung der erdbebengefährdeten Megacity Istanbul mit Methoden der Fernerkundung, Dissertation, Universität Würzburg, Germany. http://www.opus-bayern.de/uni-wuerzburg/volltexte/2008/2804/

Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J., & Dech, S. (2009). Integrating Remote Sensing and Social Science – The Correlation of Urban Morphology with Socioeconomic Parameters. In Urban Remote Sensing Joint Event, Shanghai, China.

UNITED NATIONS (2008). World Urbanization Prospects: The 2007 Revision, United Nations Department of Social and Economic Affairs/Population Division, New York, USA

Wurm, M., Taubenböck, H., Roth, A., & Dech, S. (2009). Derivation of Population Distribution for Vulnerability Assessment in Flood-prone German Cities Using Multisensoral Remote Sensing Data. In SPIE Europe, Berlin, Germany.

Acknowledgements

The authors gratefully acknowledge the support of Mustafa Erdik for supplying the Study *A Disaster Prevention/Mitigation Basic Plan in Istanbul including Seismic Microzonation in the Republic of Turkey*. We would also like to thank Tamer Soylu for the provision of GIS-data generated during the *Earthquake Master Plan for Istanbul* as well as Dietmar Borst for making available data collected in the context of the *Zeytiburnu Pilot Project.*

Author Biography

Julia Kubanek is student of geography at the University of Göttingen in Germany. She is currently writing her diploma thesis at the Center for Disaster Management and Risk Reduction Technology (CEDIM) in Karlsruhe, Germany. Her research interests include urbanisation, megacities and disaster risk reduction.