Sina Marie STARMANS

Spatial disaggregation of population data onto Urban Footprint data

MASTERARBEIT

eingereicht an der

LEOPOLD-FRANZENS-UNIVERSITÄT INNSBRUCK
FAKULTÄT FÜR GEO- UND ATMOSPHÄRENWISSENSCHAFTEN

zur Erlangung des akademischen Grades

MASTER OF SCIENCE

Beurteiler:
Prof. Dr. Johann Stötter

Institut für Geographie

Innsbruck, 30. Oktober 2014

Sina Marie Starmans
Matrikelnummer: 11109205
sinastarmans@googlemail.com

Zusammenfassung

Laut Schätzungen des Wirtschafts- und Sozialrats der Vereinten Nationen wird die Weltbevölkerung zwischen 2013 und 2050 um 2,4 Milliarden Menschen wachsen. Während die Bevölkerungszahl in den entwickelten Ländern nahezu unverändert bleiben wird, wird die Bevölkerungszahl in den weniger entwickelten Ländern hingegen von 5,9 auf 8,3 Milliarden ansteigen. Das größte Bevölkerungswachstum ist in den am wenigsten entwickelten Ländern zu verzeichnen.

Neben dem hohen Bevölkerungsdruck sind die am wenigsten entwickelten Länder am vulnerabelsten gegenüber Naturgefahren. So leben zwar lediglich elf Prozent der dortigen Bevölkerung exponiert zu Naturgefahren, aber 53 Prozent aller Opfer von Naturgefahren weltweit werden in eben diesen Ländern verzeichnet.

Die Daten, welche für Schutzmaßnahmen vor Katastrophen notwendig sind, sind meist von schlechter Qualität oder unvollständig.

Im Fall einer Naturkatastrophe ist ein gutes Krisenmanagement unabdingbar. Dieses ist abhängig vom Wissensstand über die Verteilung der Bevölkerung. Die Information über Bevölkerungsverteilung - zur Verfügung gestellt von statistischen Ämtern - ist normalerweise auf ein administratives Level aggregiert. Jedoch ist dieser Detailgrad im Katastrophenfall meist nicht ausreichend. Um eine detailliertere Bevölkerungsverteilung zur Verfügung zu stellen, wurden bereits einige Modelle zur Verteilung von Zensusdaten auf Siedlungsflächen entwickelt. Die Auflösung der bestehenden Modelle „Gridded Population of the World“, „Global Rural Urban Mapping Project“, „LandScan“ sowie „WorldPop“ ist meist zu grob, um eine genaue Aussage über die von einer Naturkatastrophe betroffene Bevölkerungsregion zu machen. Aus diesem Grund ist ein neues Verteilungsmodell mit einer hohen räumlichen Auflösung und weltweiter Anwendbarkeit notwendig.

In dieser Forschungsarbeit wurde ein Bevölkerungsverteilungsalgorithmus entwickelt, welcher auf Zensusdaten und Urban Footprint Daten basiert. Die hohe Auflösung der Urban Footprint Daten ermöglicht eine genaue Bevölkerungsverteilung innerhalb einer Pixelgröße von zwölf Metern. Genannter Algorithmus wurde in Bayern, Deutschland entwickelt und später auf Namibia übertragen. Exemplarisch wurde ein Bevölkerungsverteilungsmodell für die Cuvelai-Etosha Region, eine Region im Norden Namibias, erstellt und mit einer Flutmaske vom Flutevent 2009 kombiniert, um die Anwendbarkeit des neuen Modells für die Abschätzung von Hochwasserschäden aufzuzeigen.

Abstract

According to the Department of Economic and Social affairs of the United Nations, the world population is likely to grow by 2.4 billion between 2013 and 2050. While the population in the developed countries will remain largely unchanged, the population in the less developed countries rises from 5.9 billion to 8.3 billion. In the 49 least developed countries the fastest population growth is recorded.

Besides the high population pressure, the least developed countries are the most vulnerable ones to natural hazards. Only 11% of the population live exposed to hazards, but 53% of all victims of natural hazards are documented in these countries. The data, required for disaster risk reduction, is often of poor quality or lacking.

In case of a natural disaster a proper post disaster management is depending on the knowledge about the quantity and distribution of population. The information on population distribution provided by statistical agencies, is commonly aggregated to administrative units. However, this level of detail is mostly not sufficient enough in case of a disaster. To provide a more detailed population distribution, several models, disaggregating population counts on settlement areas, were developed. The resolution of the existing models Gridded Population of the World, Global Rural Urban Mapping Project, LandScan, and WorldPop is mostly to coarse to facilitate a precise statement for affected population during a disaster. Therefore, a new disaggregation model with a high spatial resolution and worldwide applicability is necessary.

In this research work a population distribution algorithm was developed, based on Census data and Urban Footprint data. The high resolution of the Urban Footprint facilitates a precise population disaggregation within a pixel size of 12 m. The algorithm was developed in Bavaria, Germany and later transferred to Namibia. Exemplary, a population distribution model of the Cuvelai-Etohsa, a region in Northern Namibia, is generated and combined with a floodmask of the flood of 2009 to show the applicability of the new model for flood loss estimation.

Acknowledgments

First of all, I want to thank Prof. Dr. Johann Stötter of the Leopold-Franzens-University of Innsbruck for the supervision of my thesis and the support during the last months.
This thesis was carried out at the German Remote Sensing Data Center (DFD) of the German Aerospace Center (DLR). Therefore, first I want to thank the head of the Department for Geo-Risks and Civil Security Prof. Dr.-Ing. Günter Strunz and the head of the Team for Modeling and Geostatistical Methods Dr. Hannes Taubenböck for the possibility of writing my thesis at the German Aerospace Center. In this context, I want to thank the whole Team for Modeling and Geostatistical Methods, especially Franz Hummel and Michael Wiesner, for their helpful suggestions and constructive criticism.
My special thanks go to my direct supervisor at the DLR, Matthias Mück. Thank you for the support during the last nine month, the technical advice and your optimism.
Finally, I want to thank my parents, my sister Julia, and Max for supporting me during the work on this thesis and my whole studies.

Contents

List of Figures

List of abbreviations

AM	Area-Weighted Mean
AVHRR	Advanced Very High Resolution Radiometer
BKG	Bundesamt für Kartographie und Geodäsie
CIAT	International Center for Tropical Agriculture
CIESIN	Center for International Earth Science Information Network
DCW	Digital Chart of the World
DFD	German Remote Sensing Data Center
DI	Dispersion Index
DLR	German Aerospace Center
GDP	Gross Domestic Product
GLCC	Global Land Cover Characteristics
GPW	Gridded Population of the World
GRUMP	Global Rural Urban Mapping Project
GUF	Global Urban Footprint
HR	High Resolution
IFPRI	International Food Policy Research Institute
MAE	Mean Absolute Error
MR	Medium Resolution
NCGIA	National Center for Geographic Information Analysis
NSA	Namibia Statistics Agency
OLI	Operational Land Imager
SADC	Southern African Development Countries
SAGE	Center for Sustainability and the Global Environment

SAR Synthetic Aperture Radar

SASSCAL Southern African Science Service Center for Climate Change and Adaptive Land-use

TDM TanDEM-X Mission

TDX TanDEM-X

TIRS Thermal Infrared Sensor

TPC Tactical Pilotage Charts

TSX TerraSAR-X

UNDAC United Nations Disaster Assessment and Coordination

UF Urban Footprint

USGS U.S. Geological Survey

VHR Very High Resolution

1 Introduction

1.1 Motivation

The United Nations World Population Prospects of 2012 recorded a world population growth of 81 million people annually between 2005 and 2013. While the population in the developed countries of the world remain largely unchanged between 2013 and 2050, the population in the less developed countries rises from 5.9 billion to 8.3 billion. The 49 least developed countries record the fastest population growth within this period with a population that doubles between 2013 and 2050. In 2100, 88.2% of the population will live in the less developed countries, including 27% in the least developed countries. Only 11.8% of the population is proposed to live in the more developed regions. The population increases particularly in high-fertility countries, mainly in Africa. (United Nations Population Division, 2012)

Besides the strong population growth, the number of natural disasters increase as well. Hotspots for natural hazards are located worldwide, but they have a totally different influence on the population, depending on the country they occur. While in the most developed countries, natural hazards cause high economic losses, whereas in the less developed countries they cause a high mortality rate (Dilley et al., 2005). In the least developed countries, 11% of the population live exposed to hazards. Nevertheless, 53% of all victims of natural hazards are documented in these countries. In contrast, in the most developed countries, where 15% of the population live exposed to hazards, only 1.8% of the victims are registered (Peduzzi et al., 2009).

The high population growth in less developed countries and their high vulnerability, lead to a rising number of victims of natural hazards. Furthermore, increasing population pressure forces especially poor people to settle in areas exposed to natural hazards.

In this context, knowledge about the quantity and distribution of population affected by natural hazards, is crucial for a proper risk and post disaster management (Wurm et al., 2009). It is also the basis for planning disaster risk reduction and important for vulnerability assessment (Setiadi et al., 2010). Therefore, there is an increasing interest in creating spatially-explicit, large-area gridded population distribution datasets (Gaughan et al., 2000).

In high-income countries, extensive mapping resources are often available, in contrast to low-income countries, where such data normally is of poor quality or lacking (Tatem et al., 2007). Existing maps show population on the level of certain administrative regions, or - in the case of distributed data - in a coarse spatial resolution (Hay et al., 2005). Data is mostly outdated, generalized, not area-wide, not reliable, or not existent. To receive accurate and detailed information about

the distribution of population in relation to existing hazards, interdisciplinary integration of social science and remote sensing is necessary (Setiadi et al., 2010). Since high degrees of individual and local heterogeneity within geographic regions or administrative units can be discovered, an accurate mapping of settlements is important, in order to identify where population reside within census units (Tatem et al., 2012).
Area-wide and up-to-date information is needed, in order to know how many people are affected, in case of a natural disaster and where those people live (Wurm et al., 2009). The advances in remote sensing technologies, achieved in the last years, make an accurate identification of urban areas possible (Cheriyadat et al., 2007). On basis of very high resolution remote sensing data, census data can be distributed directly on settlement areas, which are not confined by administrative units.

1.2 Research Objectives and Questions

The aim of this research work is the development of a worldwide applicable and accurate instrument for the disaggregation of census population data, on fully distributed spatial data, showing settlement areas in a spatial resolution of 12 m.
The resulting population disaggregation algorithm, is evolved from studies in two study areas, one in Namibia and one in Bavaria, Germany.
In a next step, the spatial population map is used to assess the population affected by a natural hazard, within the task of risk assessment and post disaster management.
The achieved results are applied on flood risk assessment in Northern Namibia, a work package integrated in a SASSCAL (Southern African Science Center for Climate Change and Adaptive Land-use) Project.

The following research objectives shall be reached within this research work:

- Review the state-of-the-art of population distribution modeling approaches;
- Derivation of spatially detailed population data in German and Namibian test sites, developing a disaggregation method based on urban footprint classifications, derived from Earth observation data;
- Workflow development for Global Urban Footprint (GUF) based population distribution modeling;
- Result validation and method adjustment, based on in-situ data;

- Application of the GUF based population distribution modeling for local flood loss estimation within the SASSCAL project.

The outcome of these research objectives are the following research questions:

1. What are the challenges to derive population information from the GUF for a specific test site?
2. How can we derive a solid workflow for a global application of GUF based population information?
3. What is the benefit of the derived detailed and rapid population information for specific demands and problems in flood risk management?

1.3 State of the Art

There are several different methods of population distribution modeling. The first distinction is between a bottom-up approach and a top-down approach (Figure 1). The bottom-up approach, is the most accurate way to produce a map of population density. It is an upscaling method, where the research area is split into similar grid cells and the number of people living in those grid cells is counted. In addition, dwelling data with geographic coordinates is collected. A dasymetric map is generated, based on the population and dwelling data for each grid cell. The bottom-up approach is often not feasible, because the required detailed data is not available or existent. (Gallego, 2010; Khomardin, 2010)
An alternative solution for the bottom-up approach is the top-down approach, a downscaling method, where global information is disaggregated to a more detailed level to develop a dasymetric map (Gallego, 2010). Downscaling methods for population distribution often orientate on land cover classes and distribute population counts due to the land use, like the methods of Langford et al. (1991), Yuan et al. (1997), Briggs at al. (2007), and Mennis (2003). Xie (1995), Mrozinski and Cromley (1999), as well as Reibel and Bufalion (2005) have produced a more precise downscaling population density layer by additional using road networks, with the setting that population concentrates win a buffer around roads. Nelson and Deichmann (2004) describe the usage of the road network to distribute the population due to accessibility modeling, because people tend to live and move towards areas that are well connected with urban centers.
The top-down approach can be distinguished in choropleth mapping and areal interpolation. Choropleth mapping divides the total population of each enumeration unit by its area and displays the result, using spatial shading schemes. Population is uniformly distributed within an enumeration unit without considering that

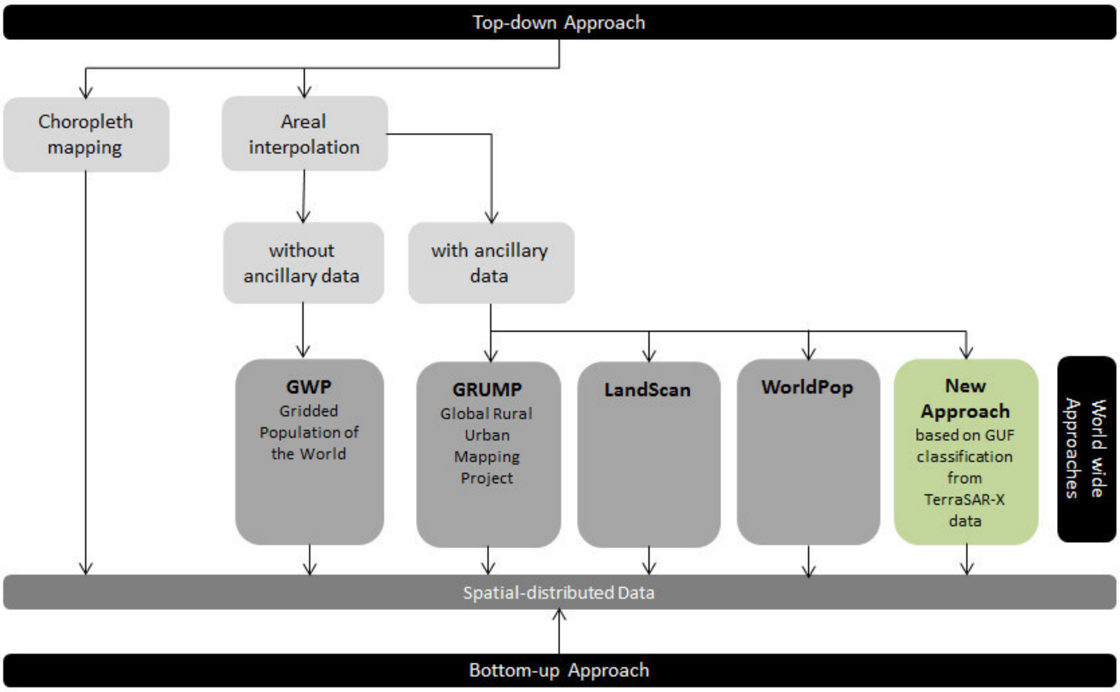


Figure 1: Overview of Chapter 1.3: the State of the Art, which shows different approaches to generate a population distribution dataset with worldwide applicability. The green marked box illustrates the methodology, which is applied in this research work.

within an unit also non-populated areas or different population densities exists. Thus, the choropleth mapping approach removes the spatial heterogeneity and does not reflect the reality of population distribution. (Holt et al., 2004; Langford & Urwin, 1994)
Areal interpolation facilitates to transform data from a source zone to a finer spatial unit, the target zone, which is not confined by an areal unit (Langford et al., 1991; Wu et al., 2005). The areal interpolation methods can be distinguished in two categories, areal interpolation without ancillary information and areal interpolation using ancillary information (Wu et al., 2005).

The approach which has been developed in this master thesis, is a world wide top-down appendage. Due to this, in the following passages only earlier appendages with the same purpose are considered.
The first efforts to develop a global scale grid, linked with population data, started in the mid-1990s at the National Center for Geographic Information and Analysis,

at the University of California, Santa Barbara (Tobler et al. 1997). In 1994, at the Global Demography Workshop, held at the Center for International Earth Science Information Network (CIESIN), experts came together and created the first global population grid (Balk et al., 2006).

The Gridded Population of the World (GPW) is the first major effort to generate a consistent global georeferenced population dataset (Balk et al., 2005). It has been produced at the National Center for Geographic Information Analysis (NCGIA) in 1995 (Tobler et al., 1997) and was updated 2000 by CIESIN (Deichmann et al., 2001). It is the only explained approach, using areal interpolation without ancillary data.
The guiding principle of the approach, is to achieve the best available data products representing the distribution of human population. To accomplish this, two basic input variables are utilized: population estimates, associated with administrative units, and spatially explicit administrative boundary data, collected from many different data providers at the highest available resolution. Population data is transformed from their native spatial units, that are usually administrative divisions of irregular shape and resolutions, to a global grid of square cells. They are proportionally allocated from administrative units to these grid cells, with the assumption that the population is distributed equal within each administrative unit. In first attempts, a resolution of 10 km was achieved. Later, it was improved to a resolution of 5 km. (Balk et al., 2005; 2006)

In 2000, the Global Rural Urban Mapping Project (GRUMP) was developed by the CIESIN, the International Food Policy Research Institute (IFPRI), the World Bank, and the International Center for Tropical Agriculture (CIAT), based on the GPW approach (Balk et al., 2006).
The data input are population data and data of urban areas. The population data is elevated with a global database of cities and towns of 1,000 persons or more, where each settlement is spatially represented as a point and has associated tabular information on its population from official statistically offices. The physical extent of urban areas is collected from raster and vector layers, like the World Stable Lights dataset as nighttime layers, the digital chart of the world (DCW)'s populated places, an product of ESRI, and the Tactical Pilotage Charts (TPC), produced by the Australian Defence Imagery, which is used where nighttime lights and DCW data did not adequately delimit urban areas. All sources of urban extent are combined to obtain the maximum possible coverage for every country. By combining population data and the Urban mask, a 1 km population distribution raster dataset results. (Balk et al., 2005; 2006)
The aim of GRUMP is not only to construct an improved population grid but also

to accomplish a globally consistent database, which shows the percentage of the population in urban areas and in rural areas. To validate the approach, the sum of each grid cell population cannot exceed the value for the administrative unit in which they fall. (Balk et al., 2005; 2006)

LandScan is one of the most popular and accurate approaches of population distribution modeling. It has been developed in 1998 as a worldwide population database with a resolution of 1 km for estimating ambient populations at risk.
The best available census counts are distributed to cells, based on probability coefficients. The probability coefficients are calculated on the basis of several different data types: road proximity, slope, land cover, nighttime lights, coastlines, and high resolution satellite imagery. The input variables depend on freely available databases, offering worldwide coverage with a resolution of 1 km or finer. (Dobson et al., 2000)
The aim of the LandScan model is a population distribution, which considers the likely ambient locations, integrated over a 24-hour period for typical days, weeks, and season, and not like most census counts, where people are at their nighttime residence. This is reasoned by the fact that a disaster, like natural hazards, nuclear, biological, and chemical accidents, or others, not only happen during nighttime but also occur when people are at work or travel. (Dobson et al., 2000)
The chosen variables, describing population distribution, have been selected in order to where the best settlement conditions are, where people stay during daytime and which ways they travel. The transportation network is an indicator for the traveling ways and slope an indicator for settlement areas, because most human settlements occur on flat to gently sloping terrain. The land cover is the best single indicator of population density. It determines the average population density per unit, for each land cover type. Nighttime lights show the housing areas and partly confirm land cover classes. The high resolution satellite imagery is used to identify settlement patterns and building characteristics, to evaluate the accuracy and precision of different spatial layers used in the model. (Dobson et al., 2000)
The probability coefficients are devolved for each value of each input variable, and a composite probability coefficient is calculated for each LandScan cell. The coefficients vary from country to country and also in smaller scales. The generic model is the same for all regions, but the probability weights of individual variables must be customized for each area, due to economic, physical, and cultural factors. (Dobson et al., 2000)
The LandScan is a multivariate dasymetric modeling approach, which works with key indicators of population distribution. The verification of the population distribution appendage is limited by the difficulty of having a suitable reference database. LandScan is updated annually to reflect changes in global political

boundaries and population distribution. (Dobson et al., 2000)
LandScan offers a temporal coverage, while GRUMP and GPW are limited to a single snapshot (Balk et al., 2006).

The last world wide grid-based top-down approach is WorldPop, which arose from the three approaches AfriPop, AsiaPop, and AmeriPop. The early appendages of these population distribution models are related to the LandScan approach. Population is redistributed to gird cells, due to the relationship between land cover and population density. A typical regional per-land cover class population is identified to allocate census counts. (Gaughan et al., 2013; Tatem et al., 2007; Worldpop, 2014)
Recently, this approach has been replaced by a random forest approach, a new semi-automated dasymetric modeling approach that incorporates census and a wide range of public ancillary datasets in a flexible random forest estimation technique. Input data are remotely-sensed and geospatial datasets, like settlement locations, nightlights, topography, land cover, roads, building maps, and health facility locations, and must be widely available to generate a gridded prediction of population density at 100 m spatial resolution. (Gaughan et al., 2013; Tatem et al., 2007)
The result, a prediction layer, is used as the weighting surface to redistribute census counts. The census data and its corresponding unit boundaries are derived from the most recent census data at the highest available resolution. Settlements and their mapping extent are identified with satellite imagery, at the level of a Landsat tile extent (30 m resolution). (Gaughan et al., 2013; Tatem et al., 2007)

GRUMP, LandScan, and WorldPop are approaches, using areal interpolation with ancillary data.

The additional benefit of the new developed approach, based on the GUF classification of TerraSAR-X data, is described in Chapter 2.

1.4 Thesis Structure

This research work is composed of eight chapters. In Chapter 1 the motivation for this research work, the research objectives and the resulting research questions, the state of the art, relating to population distribution modeling, and the structure of this thesis is delineated. Chapter 2, the research framework, describes the basis of the thesis, taking into account the GUF, the SASSCAL Project, the project in which this research work is incorporated, and the study area of the method development. The method development of the new population disaggregation approach is depicted in Chapter 3, with the subsections methodical framework, data requirement, and GUF population modeling. Chapter 4 shows the results of the new approach in the test area and its transferability to another area of interest. In Chapter 5 the applicability of the new population disaggregation approach in reference to the flood event 2009 in Namibia and a rapid flood loss estimation is considered, taking into account the results of other disaggregation models. The results and the application of the new approach are discussed in Chapter 6. Chapter 7 comprises the conclusion of the research work and Chapter 8 gives an outlook.

2 Research Framework

Spatial disaggregation signifies that data of an administrative unit is distributed within the administrative unit through spatial parameters. Required is a dependence between the spatial parameters and the aggregated data. Spatial disaggregation of census data is a combination of socio-economic and spatial data (Figure 2). Areal census data is transformed from an arbitrary zonation (e.g. administrative unit) to a true geography (e.g. settled pixel), using an accurate delineation of non-residential and residential areas. (Chen, 2002)

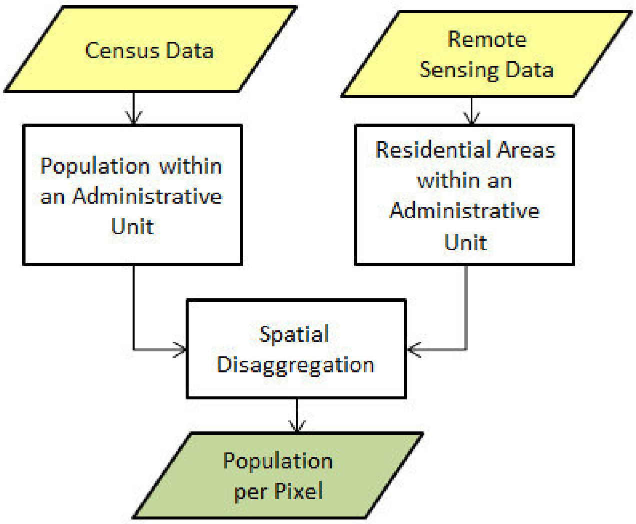


Figure 2: Population Disaggregation

By the disaggregation of population from a Census level (Region or District) to settled areas, it is possible to know really fast how many people live in a certain area, which is not confined by administrative borders. The knowledge about the population count in such a small spatial level is an important background knowledge for the hazard research and an important factor for a proper post disaster management.
There are several different approaches to disaggregate population data from an administrative level to single settlement areas, as already discussed in Chapter 1.3. LandScan and GRUMP reach the best spatial resolution of existing worldwide population distribution models with a approximately 1 km resolution at the equator. Worldpop, which is only available for South-east Asia, Africa, and South America, offers data with a spatial resolution up to 100 m.

Natural hazards can be distinguished according to their spatial extent - three scale classes: micro, meso, and macro - as well as spatial dispersion - concentrated to diffuse. For disasters with an extensive areal dimension and a low dispersion, the spatial resolution of the established population distribution models, like LandScan, is adequate. For smaller and diffuse forms, the affected population cannot be captured properly. Furthermore, population distribution models, like LandScan, are especially suitable for analyzing large-scale settlement patterns, due to their high spatial coverage. Regardless, their coarse resolution cannot appropriately display disperse settlement structures. The requirement for the new algorithm is to bridge these gaps. Population affected by small-scale hazards or living in disperse structures should be detected.

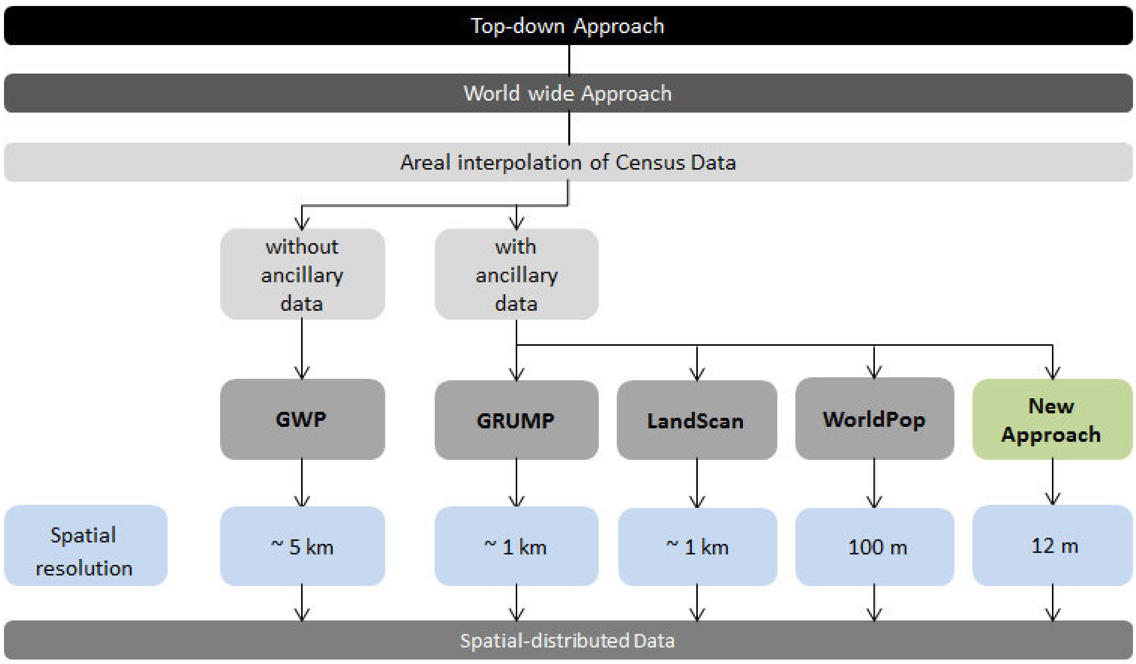


Figure 3: Research Framework

This study uses the GUF (Chapter 2.1), showing worldwide and fully distributed sealed areas, for an approximation of settlement areas on a raster with a spatial resolution of 12 m. The high spatial resolution of the GUF facilitates a considerable higher resolution of the developed population disaggregation algorithm compared to other population models (Figure 3). Small settlement structures, which are not included by further distribution models, are also captured and considered for the population disaggregation. With the new algorithm, a fast localization of the

population concerned by hazards should become possible, with a very high spatial resolution, where also small settlement structures and small-scale hazards are considered. Since the GUF is available worldwide, the new algorithm features global potential.
A GIS-based method for disaggregating census data is applied to provide a more detailed population distribution model. The Census counts of an administrative unit are redistributed to the settled areas, located in the administrative unit. The developed method is generated in two study areas, Bavaria and parts of Namibia, and can be transferred to other countries or regions to provide worldwide population density data. The basic idea of the method development was to add values to the risk research in the Cuvelai-Etosha River Basin, an area in the north of Namibia affected by sudden flood risks.

2.1 The Global Urban Footprint

The GUF, a recently compiled layer for urban areas, is one of the required data sets for the development of the new population algorithm. It is based on the new high resolution radar satellites TanDEM-X (TDX) and TerraSAR-X (TSX).
The concept of building a new urban area layer arose from the phenomena of the urban century. With the beginning of the 21st century, the urban population exceeded the rural population for the first time, with a concentration of population growth especially in urban transition zones. The steady growth and increasing dynamics of urban areas lead to a consistent transformation from natural or rural areas to a cultivated and managed landscape and requires a sustainable development to reduce the increasing loss of land resources. To accomplish this challenge, policy and environmental and spatial planers need effective instruments for the observation, analysis, and control of urbanization. In an increasing urbanized world, it is important to recognize where human settlements develop and how they affect urban form and growth patterns. (Esch et al., 2012)
For regional and global analyses of urban growth patterns, Earth observation is a key technology, because it is an independent source of information with a flexible repetition rate. Furthermore, it is applicable area-wide in an arbitrary scale and allows automated methods of data processing and image analysis. Earth observation data supports decision makers in different research and management areas with up-to-date geoinformation. (Esch et al., 2012; Taubenböck et al., 2011)
With the GUF mask, a geoinformation product which submits the detection and delineation of built-up areas, a worldwide analysis of urban areas should become possible.
The GUF is based on the TanDEM-X Mission (TDM), a German Earth observation mission with the two radar satellites, TSX and TDX, that were launched in June

2007 and June 2010, respectively. They operate as twin satellites and are equipped with a synthetic aperture radar (SAR), which enables Earth observation also during the night and cloud coverage. Because of its very high resolution (VHR), which makes spatial resolutions up to one meter possible, the usability of spaceborne SAR imagery in the context of urban applications increased. In the years 2011 and 2012, two global coverages where collected and a unique dataset could be provided, in addition to already existing global coverages from medium (MR) or high resolution (HR) optical imagery. The aggregated data supports global programs, aiming at the mapping of world-wide settlement patterns. Furthermore, the TDM enables the analysis and monitoring of spatio-temporal dynamics of settlement development, population estimation, vulnerability assessment, and global change. (Esch et al., 2010; 2012; DLR, 2011)

There have been already several attempts towards the global mapping of urban areas. One of the first approaches to generate a global land cover map, is the global land cover characteristic database (GLCC), which was developed on a continent-by-continent basis with unique elements for each continent, based on the geographic aspects of the continent. They display different landscapes with a resolution of 1 km. The database, available since 1997, was generated by the U.S. Geological Survey (USGS). The images have been collected between 1992 and 1993 by an advanced very high resolution radiometer (AVHRR). (Esch et al., 2010; USGS, 2014)

Another database showing the extent of settlements is CORINE land cover. It has been developed since 1985 to capture the influence of human land use on the natural environment but is only available for Europe. CORINE land cover differentiates between 44 land use classes and presents them in a cartographic product at a scale of 1:100,000. (Commission of the European Communities, 1995)

One of the newest attempts to detect urban areas was invented by the Center for Sustainability and the Global Environment (SAGE) at the University of Wisconsin-Madison. Their product is a global urban area map based on 16 urban ecoregions with a spatial resolution of 500 m. (Schneider et al., 2009; 2010)

All of these approaches are still deficient in their spatial and thematic resolution. They fail to provide detailed analysis of the human environment urbanization dynamics that can satisfy the request of science, politics, and planning (Esch et al., 2012).

In order to obtain a global map of human settlements from remote sensing data, the built-up areas need to be extracted. Built-up areas are places, dominated by human-constructed elements, like mainly buildings, but also roads, runways, industrial facilities, etc. The spectral and the spatial resolution of the remote sensing data are the basis to discriminate between different land cover types and surface materials. The spatial resolution defines the smallest feature that can be

determined. Respectively, the smallest distance between two objects. It shows the dimension of the area, which is covered by one resolution cell of the imaging system. With the spectral resolution, which is defined by its number of spectral bands and their width, it is possible to distinguish between different surface materials and land cover types. SAR systems, like the TSX satellite only operate with one wavelength. In opposite to optical sensors, their data is defined by physical and geometrical properties of the objects and not by the biophysical and chemical characteristics. For the identification of urban areas, the textual features and the grey tone are crucial. Man-made structures show specific geometrical properties, which make them appropriate to be identified with radar data. (Esch et al. 2010; 2012)
The method to generate the GUF, consists of two main phases to identify urbanized areas and non-urbanized areas. It is a pixel-based approach with the aim to identify and map built-up areas automatically, based on a single-polarized TSX intensity image in the strip map mode. The first step, is the analysis of the local speckle behavior, a grainy pattern that accrues by superposing true radiometric and textual information of SAR images. By analyzing the speckle characteristics, a texture layer that highlights highly textured image regions, like built-up regions, can be developed. The second step, is the automated image analyzes procedure, where the resulted texture layer of the first step is combined with the original image intensity information to extract settlements. The end product is a thematic, binary mask that points out urban and non-urban areas. The process is a step-by-step image analysis that tries to identify human settlements, due to their high amplitude and their highly heterogeneous neighborhood, with a threshold-based technique. (Esch et al. 2010; 2012)
The high spatial resolution of the GUF (12 m) allows a small scale analysis of urbanization patterns. During the GUF process even small sprawled settlements and single homesteads were detected . (Esch et al. 2010; 2012; Taubenböck et al. 2011)

2.2 SASSCAL Project

The SASSCAL has arisen from several science initiatives in Southern Africa. It is a joint initiative of Angola, Botswana, Namibia, South Africa, Zambia, and Germany with the aim of becoming a regional driver for innovation and knowledge exchange to enhance adaptive land use and sustainable economic development under the global change conditions. The mission of the SASSCAL Project is the establishment of a network of science service centers in the southern African region. Therefore, regional scientific capacity and efforts for coping with climate change and land use change are strengthen. Furthermore, information is provided for policy and development planners to improve livelihoods. The general service of

the project is to provide an appropriate range of information, data and knowledge based services, and products to a broad range of users and practitioners, and to enable capacity building. It is assisted by international and local professionals, scientists, and experts. (SASSCAL, 2011)
The DLR (German Aerospace Center) - DFD (German Remote Sensing Data Center) contributes with the subject area "Remote Sensing applications for flood risk management", aiming the development of novel, applicable, and transferable methodologies to provide information for effective flood management, such as flood monitoring and detection, emergency response and preparedness, and prevention (SASSCAL, 2011).
Five thematic areas are considered by the SASSCAL project: climate change, water, forestry, agriculture, and biodiversity. The thematic area water contains 27% of the investments for the whole project. This shows the high importance of the water sector in the arid areas of the SADC (Southern African Development Countries) region, which are especially threatened by further desertification. (SASSCAL, 2011)
The aim of the thematic area water is to develop a common water resources information base to implement integrated water resource management strategies for improved transboundary river management, and resource use. It is divided into four work packages (Figure 4). (SASSCAL, 2011)

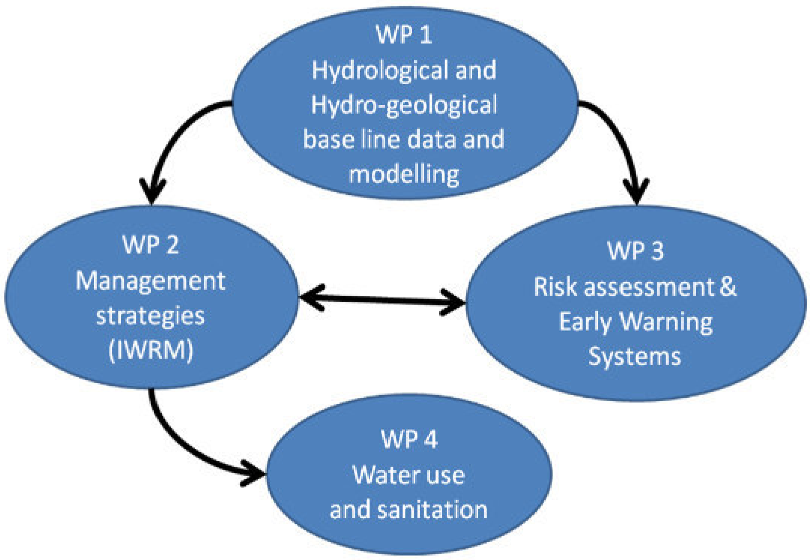


Figure 4: Work packages Water (SASSCAL Business Plan, 2011)

The research of the DLR is grounded in water work package 3 "Risk assessment and Early Warning Systems", in the task "Remote Sensing applications for flood risk management". The focus is on the development of transferable methods for flood detection and monitoring, low resolution flood mapping, and development of a rapid loss assessment concept. (SASSCAL, 2011)
The research of this master thesis was guided by the context of the task "Remote sensing applications for flood risk management". Southern Africa is a semi-arid region with a variable climate and hydro-meteorological extremes, like droughts and floods. In 2008, 2009, and 2010, heavy rains in Northern Namibia led to exceptionally high floods, the highest in living memory. Besides disruption of agriculture and other economic activities, displacement of people, destruction of infrastructure, and losses of life were the consequences of the sudden appearing floods. The rising availability of VHR satellite images offers the opportunity to observe the whole extent of floods and their spatio-temporal evolution. In this context, the DLR develops automatic flood detection methods and a risk assessment concept to support flood management. The master thesis ties in with the risk assessment concept. The population disaggregation algorithm, developed in the framework of this master thesis, shall enable a fast appraisal of how many people are affected by a flood. (SASSCAL, 2011)

2.3 Study Area

In this master thesis, two study areas are processed. The beginning of the method development was executed with a database of four constituencies and three cities in the north of Namibia and the eight constituencies of the capital city of Namibia, Windhoek. During the development process it became visible that the current database is not adequate for the development of a population disaggregation model. Therefore, 14 communities in Bavaria were added to improve the model development. The developed population dissagregation approach was later applied in the first study area, Namibia, testing its transferability. The case study is located in the Cuvelai-Etosha region, in the north of Namibia, where the majority of the Namibian population lives and sudden floods occur.

2.3.1 Namibia

Geography

Namibia, one of the countries in Southern Africa, has been a German colony, called South-West Africa, from 1884 until it was occupied by South Africa during World War 1. After being administrated by South Africa as a mandate, Namibia reached its independence in 1990. Namibia is surrounded by Angola and Zambia in the North, Botswana in the East, and South Africa in the East and South (CIA, 2014; bpb 2014). On the West coast, it borders to the Atlantic Ocean. Windhoek, the capital of Namibia, lies in the center of the country and is next to the Cuvleai-Etosha region, one of the most populated areas in Namibia (Atlas of Namibia, 2014). Because of the limited natural freshwater resources, desertification, wildlife poaching, and land degradation, Namibia incorporates the protection of the environment, as the first country in the world, into its constitution. About 14% of the land is protected (CIA, 2014).

In Namibia, only 0.97% of the country is arable land, due to its climate conditions (CIA, 2014). The south of Namibia, the subtropical part, is dominated by desert and half dessert climate, with a constant danger of droughts. This part is occupied by the Namib Desert, along the coast in the west and the Kalahari Desert in the east. The North of Namibia lies in the xeric shrubland climate, where a permanent water availability is also not secured. Droughts alternate with devastating floods, caused by the sudden appearance of torrential rainfalls and the drainage of the rivers coming from Angola. In Namibia, the average annual rainfall lies between 0 and 600 mm and the rainfall gradient increases from the desert regions in the south and west to the north-east of the country. The whole country is affected by water deficit (Figure 5), which often leads to prolonged periods of drought (Atlas of Namibia, 2014).

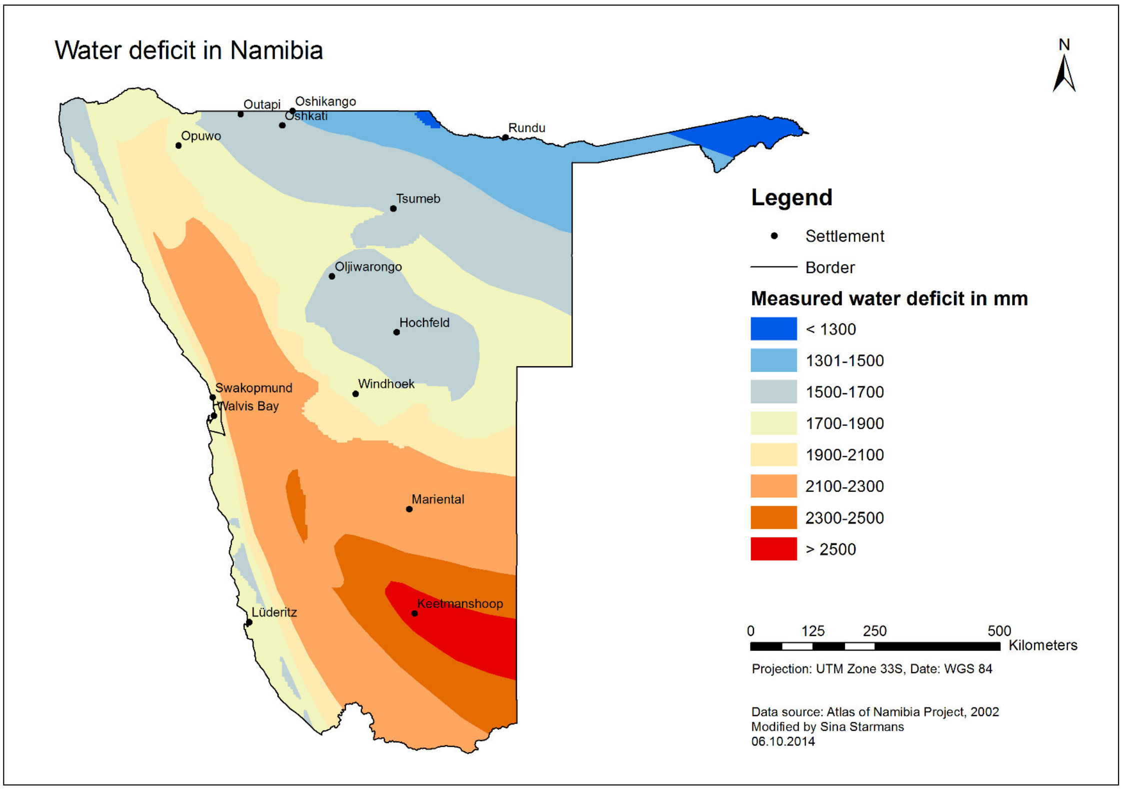


Figure 5: Water deficit in Namibia (Atlas of Namibia Project, 2002)

Demography

The total population of Namibia in the year 2011 constitute 2,113,077 people with an annual growing rate of 1.4% (Namibia Statistics Agency, 2012). Still over the half of the population (57%) lives in rural areas, but an urbanization trend is visible. Ten years before, 67% of the population lived in rural areas (Namibia Statistics Agency, 2012). Namibia has a young population with a median age of 22.8 years and a high fertility rate (CIA, 2014). In 2010, the infant mortality rate was 45.64 deaths for every 1,000 birth (CIA, 2014).
The literacy rate in Namibia reaches 89% with a rising tendency and also the number of students increased from 35% in 2001 to 52% in 2011 (Namibia Statistics Agency, 2012).

Economy

Namibia is rich in natural resources, like diamonds, copper, uranium, gold, silver, lead, tin, and more. Therefore, the economy depends on the extraction and processing of minerals for export. The mining provides more than 50% of foreign exchange earnings, although it only accounts for 11.5% of the GDP (gross domestic product). To counter the vulnerability of Namibia's economy to volatility in the price of uranium and other commodities, the Namibian authorities try to increase higher value raw materials, manufacturing, and services. (CIA, 2014)
Only 1.8% of the population is employed in the mining sector. Especially in the northern part of Namibia, the population lives of subsistence economy, which can be seen in the landscape structure. The rural areas are characterized by a lot of small settlement structures or single dwellings. (CIA, 2014)
50% of the required cereals are imported to Namibia, because of the high drought risk in the entire country and a small proportion of arable land. In drought years, food shortages are one of the major problems in rural areas. (CIA, 2014)
Namibia has one of the world's most unequal income distributions, which is shown by a Gini coefficient of 59.7 in 2010 and place six on the list of the Gini Index in a worldwide comparison (CIA, 2014). The Gini coefficient measures the aberration between the distribution of income among individuals within an economy and a perfectly equal distribution of income. Therefore, a Gini coefficient of 0 implies total equality, while a coefficient of 100 represents total inequality (The World Bank, 2014).

Case study area

The case study area is the Cuvelai-Etosha-Basin, which is located in the northern-central part of Namibia, next to the border of Angola. It is separated into four regions: Ohangwena, Omusati, Oshana, and Osikoto. The landscape is characterized by a system of shallow water courses, called Oshanas, which are not perennial. Most of the water in the basin originates from heavy rainfall events or is importet from the Kuene River in Angola. Ephemeral water courses arise and sudden floods can occur. The water courses end in the Omadhiya lakes and several pans southwards, especially in the Etosha Pan (Figure 6). (Research and Information Services of Namibia, 2013)

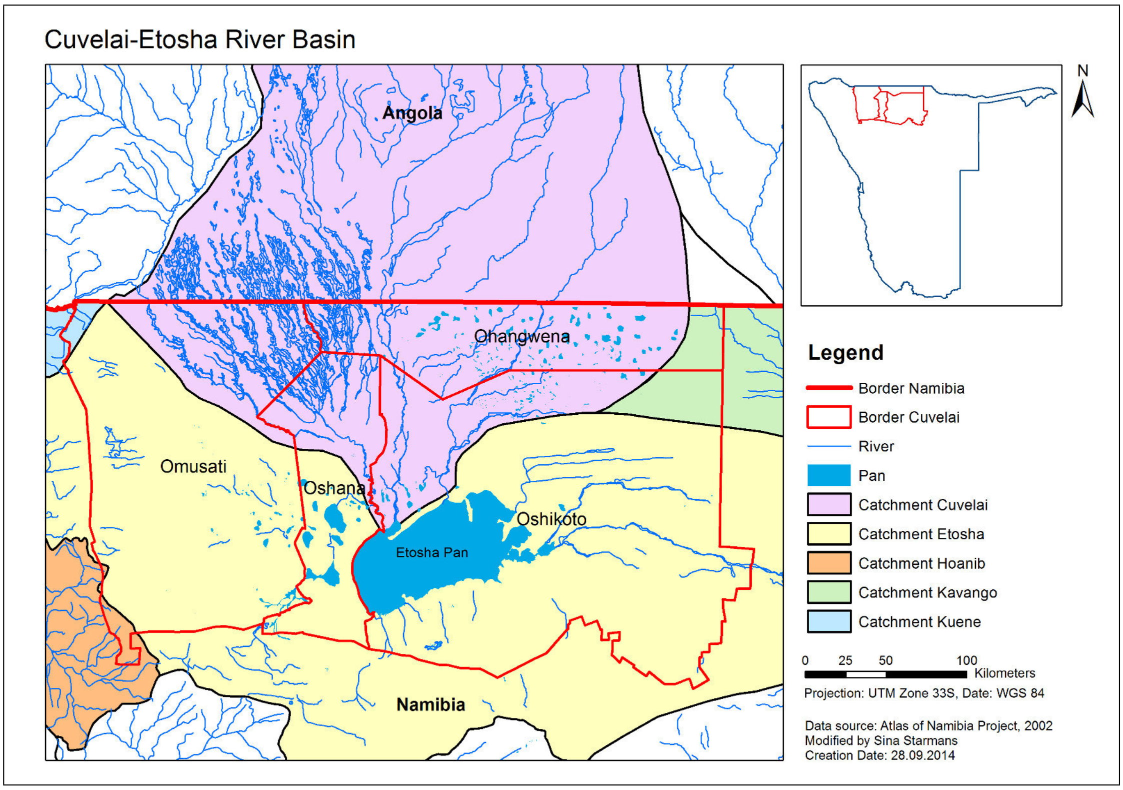


Figure 6: Cuvelai Etosha River Basin

The Cuvelai-Etosha is the most densely populated area in the country (Figure 7). The region is highly dependent on the varying precipitation and the resulting runoff in ephemeral rivers. Another watersource, besides collecting seasonal surface water, are the aquifers, like the Ohnagwena Kalahari Aquifer, the Discontinuous Perched Aquifer, and the Karst Aquifer, a vast underground water system. Through a network of pipelines and water points, the population of the Cuvelai-Etosha regions and their livestock is supplied with water. (Research and Information Services of Namibia, 2013)

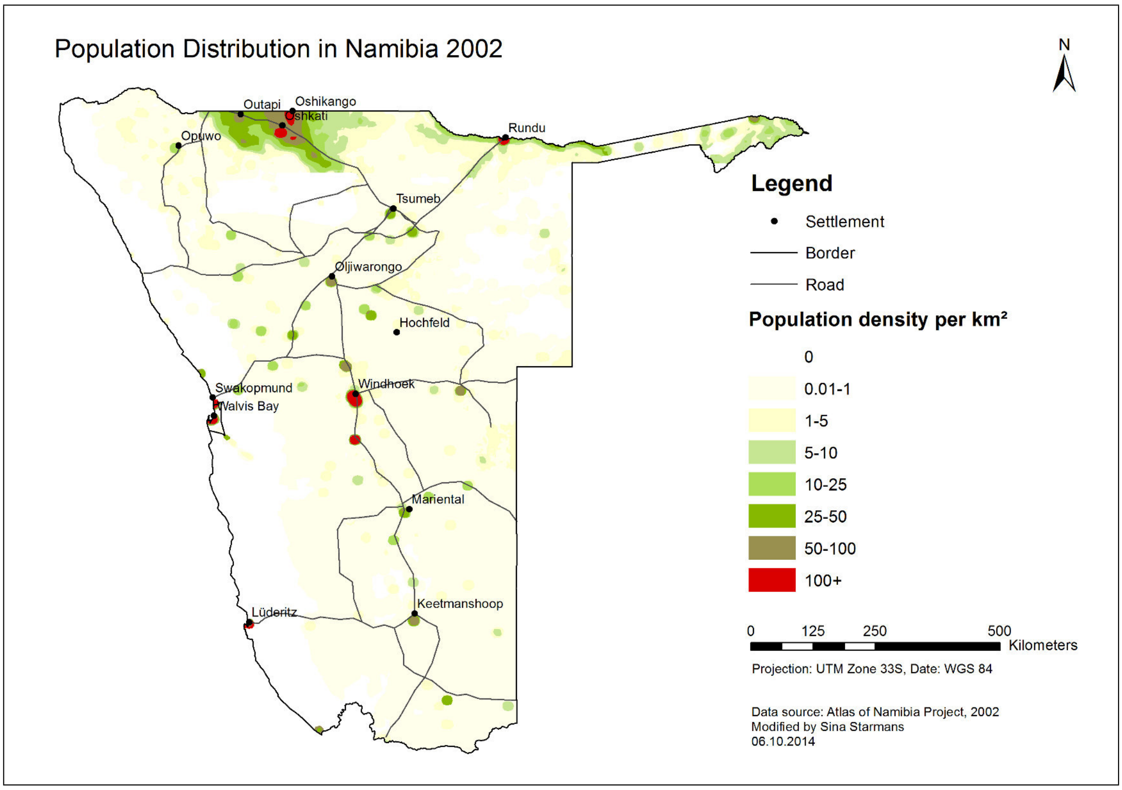


Figure 7: Population Density in Namibia 2002 (Atlas of Namibia Project, 2002)

The Cuvelai-Etosha has a high influx, because it is one of the arable countrysides in Namibia, although water availability is not guaranteed. The region is affected by droughts, which can appear several years in a row and sudden appearing floods, caused by torrential rainfalls in Angola and a high count of water draining into the Cuvelai-Etosha Basin. Dried out rivers reach water heights up to seven meters and huge areas, also build-up areas, are flooded. (Gilau et al., 2011; Research and Information Services of Namibia, 2013)
The land use structure in the Cuvelai-Etosha is characterized by single dwellings and small settlement structures, spread across the rural areas, where many people are living of subsistence farming. More than 50% of the population live in rural areas and 43.5% of the population live in traditional dwellings, which are also located in flood prone areas but not appropriate to survive a flood. (Gilau et al., 2011)

The Cuvelai-Etosha is especially suitable as a case study area. The newly developed population disaggregation algorithm captures the small settlement and housing structures, as well as enables the assessment of the affected population in the case of natural hazards, especially for floods featuring a diffuse structure.

2.3.2 Bavaria

The main part of the method development was acquired in Bavaria, one of the 16 federal states, situated in the south-eastern part of Germany. Bavaria is separated in seven administrative districts, 71 rural districts and 25 urban municipalities, and 2056 communities (Bayerische Staatsregierung, 2014). 14 rural districts have been chosen for the algorithm development, two in each administrative district, in order to represent different settlement structures in urban and rural areas. The comprehensive database in Bavaria allows to select an adequate sample for the method development.

3 Method development

3.1 Methodical framework

Census data, provided by statistical agencies, is normally limited to administrative units, like districts. Usually the smallest available census data is census level 3 'community'. In many countries, census data for small administrative units is not available, mainly due to political reasons, like concealment of results or missing elicitations. The level of detail of available datasets is usually not sufficient enough for post disaster management. A disaggregation method is required to produce more detailed information about population distribution, which is not constricted to borders.

The approach, developed in this research work, is a pixel-based top-down approach, where census data of an administrative unit is disaggregated to the settled areas, situated in the considered unit. As shown in Figure 8, the disaggregation method is based on two input variables, Census Data on a district level and the GUF. During the algorithm development, three different grades of population disaggregation have been reached: the Linear Population Distribution, the Population Distribution based on Built-up Density, and the Population Distribution based on Built-up Density and Settlement Dispersion. The census data must be disaggregted at an appropriate spatial scale, compatible with raster-based datasets. Hence, the population data is distributed onto a pixel size of 12x12 m, the size of a GUF Pixel. The result of the population distribution is a count of people for each settled 12x12 m pixel within a district. The pixel-by-pixel population values (Homogeneous Population Distribution of a GUF Pixel, Density-dependent Population Distribution on a GUF Pixel, and Combined Population Distribution on a GUF Pixel) are evaluated by summarizing them on community level, the next smaller level to district level.

The aim of the developed algorithm is the computation of pixel based population values for every settlement area of interest, independent of administrative units.

The algorithm was developed, based on a worldwide applicability. Hence, only input data which is available worldwide is applicable. The minimum required input variables are information about the extent of settlement areas and information on the population count.

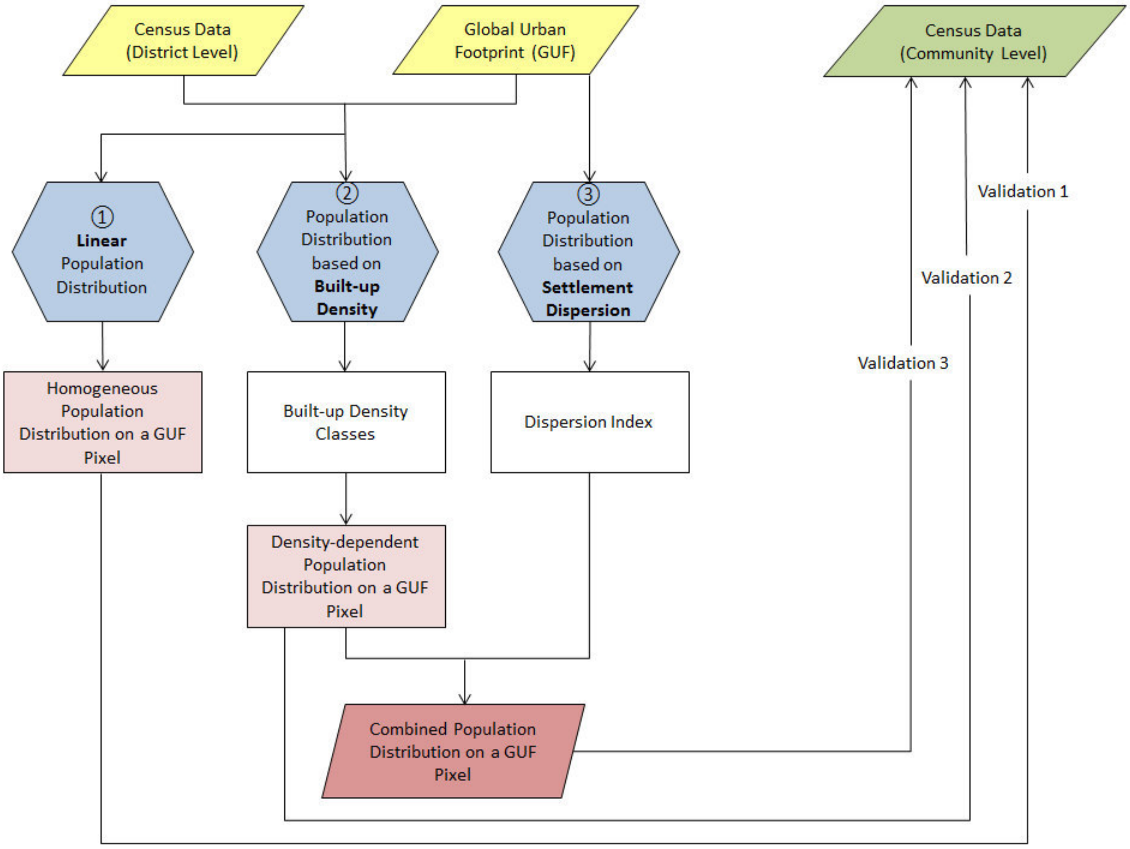


Figure 8: Workflow of the developed population distribution modeling

3.2 Data requirement

According to Figure 8 and previously described, two different types of input data are necessary for the development of the new population algorithm: data of the population within an administrative unit and information about the populated area. Population counts are usually available at national statistic agencies and the populated areas can be extracted from the GUF (Chapter 2.1).

Since its independence in 1990, Namibia conducts their Population and Housing Census every ten years. The official Namibia 2011 Population and Housing Census was carried out in accordance to the UN Principles and Recommendation for 2010 Round of Population and Housing Censuses and is the current basis for population disaggregation in Namibia. It is used by the government to support the socio-economic efforts in Namibia and provides information about the population size and number of households at three levels: national, regional, and constitutional. In addition it allocates information about the differences between urban and rural areas. The Population and Housing Census is conducted by the Namibia Statistics Agency (NSA) and is also provided for public and private sectors, training and research institutions, non-governmental organizations, the media and the general public. Besides numbers of the population and households for different administrative units, the Census also covers demographic issues, like health, labour force, and further social and economic characteristics. In the framework of the Population and Housing Census 2011, the NSA has developed a Census Atlas, "Atlas of Namibia", which shows the results of the Census in maps. (Namibia Statistics Agency, 2012)

The second input variable, the GUF, features a spatial resolution of 12 m, which makes it possible to identify the borders of settlements, larger green zones in settlements, and separate dwellings. Therefore, the GUF is especially appropriate for the application in the Cuvelai-Etosha Basin, which is composed by small cabin settlements. Since the GUF is not yet finally validated in the study area, it is required to gather the needed data in a different way.
Based on the available satellite scenes, two different ways have been chosen to elevate the required settlement data.

The first, is an automatic and object oriented segmentation of images, which may be conducted by several different algorithms, like the Chessboard Segmentation, the Quadtree-Based Segmentation, and the Multiresolution Segmentation.
The Chessboard Segmentation algorithm and the Quadtree-Based Segmentation algorithm are methods with a low calculation expense. The Chessboard Segmentation Algorithm splits an image object domain into square image objects, while the Quadtree-Based Segmentation algorithm segments an image into a quadtree grid, formed by square objects. (Trimble, 2011)
The Multiresolution Segmentation is a more sophisticated algorithm and was used to derive the settlement segments. It is a bottom-up segmentation that merges image objects with the same spectral characteristics to a new image object level. The segmentation starts with each pixel composing one image object. In the next steps, pairs of adjacent image objects are merged into larger objects, depending on their homogeneity. The process is concluded, when no more merges are possible. By merging two image objects, the spectral average heterogeneity increases. Due to this, it is important to set a degree of fitting, describing the change of heterogeneity when two single image objects are merged. The adjacent image objects with the smallest increase of heterogeneity are merged. (Trimble, 2011; Baatz &

Schäpe, 2000)
The scale parameter determines the maximum allowed heterogeneity of two objects that shall be merged. It is calculated by two criteria, the color, which describes the spectral value of the resulting image object, and the shape of an object, which describes the textual homogeneity of the resulting image object. The shape is subdivided into smoothness, the degree of the fractal extent and into compactness, the similarity to the ideal body, the quadrat. (Mott, 2005; Stoinski, 2007)

The automatic data acquisition was conducted with a Landsat 8 scene. The Landsat 8 satellite was launched in February 2013, equipped with a Operational Land Imager (OLI) and a Thermal Infrared Sensor (TIRS). Landsat 8 images consist of nine spectral bands with a spatial resolution of 30 m and two thermal bands with a spatial resolution of 100 m. (USGS, 2014)
The automatic data acquisition could only be applied in Windhoek, because the spatial resolution of the Landsat 8 images is too coarse to detect the small settlement structures in the Cuvelai-Ethosa Region. In Windhoek, the biggest city of Namibia, a compact settlement structure is common, while single settlement structures are rather unusual. The applied Landsat 8 scene is from January 2014.

The second way to elevate settlement data, is the manual data acquisition of settlement areas. Since the spatial resolution of a Landsat 8 scene is not fine enough and no satellite data with a higher spatial resolution is available, the residential areas in the Cuvelai-Etosha Region are digitalized manually, in order to the ArcGis Base Map. The satellite scenes of the used Base Map extract are images of the Worldview 2 satellite from March 2011. The Worldview 2 was launched in October 2009. It is equipped with one panchromatic sensor band and eight multispectral bands. Images from the Worldview 2 satellite have a spatial resolution of 0.5-1 m, which makes it possible to detect the small settlement structures in the Cuvelai-Etosha. (DigitalGlobe, 2010)

The two described ways for collecting data of settlement areas, without having any special layer for it, have been accrued by taking the spatial resolution of the GUF into account.

The employed data for Bavaria are population data from the Bundesamt für Kartographie und Geodäsie (BKG) of 2014, including the census level 'community' and 'district'. Settlement data was provided by the DLR, in form of the GUF (Chapter 2.1).

Generally, the spatial data is required to have an adequate resolution to facilitate an identification of building structures in different sizes. The census data must be accessible in two different administrative units. One for the calculation of the population distribution and one for the validation of the results. Both input variables need to be available area-wide.

3.3 GUF population modeling

The development of the population disaggregation algorithm is divided into three methods. The first method, the Linear Population Distribution, shows the kind of correlation between population and settlement structure. The second method, the Population Distribution based on Built-up Density, includes the findings of the first method and a hypothesis is proposed to improve the algorithm in order to achieve better population disaggregation results. The third method, the Population Distribution based on Settlement Dispersion, involves a further hypothesis to improve the results of the second method.
The aim of the population algorithm is the disaggregation of census data of an administrative unit to the settlement areas (GUF areas) within the administrative unit. The algorithm was developed on the basis of two study areas, Namibia and Bavaria (Chapter 2.3). In Bavaria, the administrative unit 'rural district' has been chosen for the disaggregation. For the validation of the results, a smaller unit than the one at the diasggregation is needed. 'Rural district' is the second smallest census level in Bavaria and the results of the disaggregation can be evaluated by the unit 'community', the smallest census level in Bavaria. In Namibia the smallest census level is 'constituency', which is similar to the 'community' in Bavaria. The next bigger level is 'region', similar to the Bavarian 'rural district'. In Namibia, the database for settlement areas is poor. Due to this, a self-created census level, consisting of seven constituencies in the Cuvelai-Etosha Basin and 8 constituencies in Windhoek, is used as disaggregation level.

In the following chapter, the three methods of the population diasaggregation algorithm are described.

3.3.1 Population algorithm development

For the population disaggreagtion, areal census data is physically linked with actual residential areas. The algorithm development starts with an equal distribution of census count over the living spaces. In the further process of the population disaggregation, the algorithm is improved by parameters, which influence the dis-

tribution of population.
In advance, different influence factors have been analyzed:

- the built-up density;
- the dispersion of the GUF area;
- the position of a GUF Pixel within a grid cell, relative to the observed grid cell within the whole GUF area;
- a combination with land use classes, because the GUF also includes industrial areas;
- a country factor.

The first and the second factor are taken into consideration for the development of the new population algorithm.

Linear Population Distribution

The first method of the population distribution, the Linear Population Distribution, is based on the assumption that an equal count of people is distributed on each GUF Pixel within the observed unit. There is no differentiation, whether the single pixel is situated in a highly densified area or in a sparsely populated area, or if there is any other influence on population distribution.
The result of the Linear Population Distribution is an equal population value for each settled pixel (Figure 9).

Eq. 1 describes the calculation of a population distribution, due to an equal population value per pixel. The population of the area of interest (Pop_i), e.g. a district, is calculated by multiplying the population value per pixel ($\frac{Pop}{Pixel}$) by the area of the settled pixels, lying in the area of interest (A).
The population value per pixel is the population of the whole study area divided by the settlement area of the whole study area.

$$Pop_i = \frac{Pop}{Pixel} * A \tag{1}$$

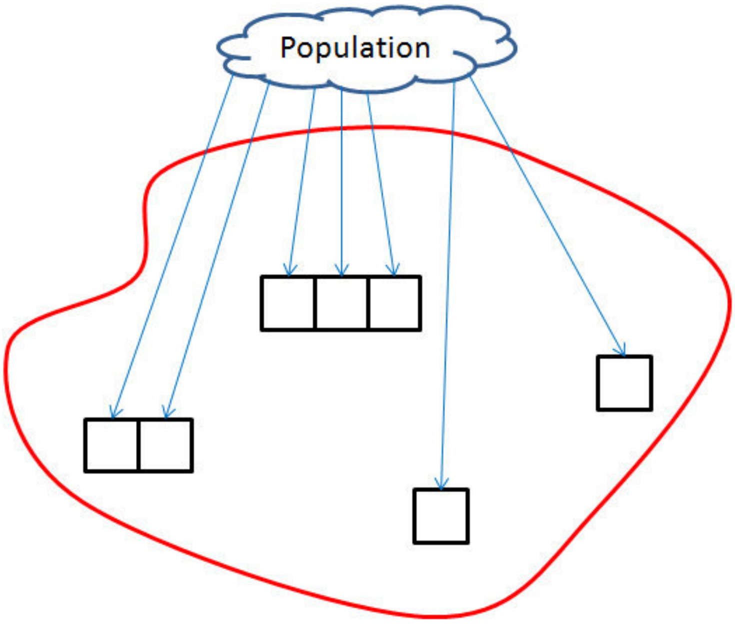


Figure 9: Linear Population Distribution - every GUF Pixel receives an equal population count.

Population Distribution based on Built-up Density

The second method of the population distribution aims at building a relationship between different residential densities and census data. There are several different approaches using built-up density as an influence factor for population distribution.
Chen (2002) linked three levels of residential density with areal census data and tried thereby to show a correlation between areal census data and residential densities.
Taubenböck et al. (2007) disaggregate the population data spatially, according to homogeneous urban structure types. The homogeneous spatial units are derived from three structural elements, the built-up density, the building heights, and the land use. The population distribution correlates directly with the structural characteristics of the 24 urban structure types resulting from the three elements. They calculated four built-up density zones: high density, medium density, low density, and open space.

In this approach it is assumed that densely built-up areas are an indicator of city structures, featuring a higher count of population per sqm than sparsely built-up areas. Cities, in comparison to rural areas, are characterized by higher buildings, which signifies that more people live in the same spatial pixel than in areas, which are described by low building heights. These building heights cannot be considered by the input variables, because the GUF shows the built-up areas only in two dimensions without the height information. Therefore, the density of built-up areas is used to distribute the population on settled pixels.

The basic assumption for the method is the following hypothesis:

The population density behaves directly proportionate to the build-up density.

The hypothesis signifies that the higher the built-up density, the higher is the population count per pixel (Figure 10).

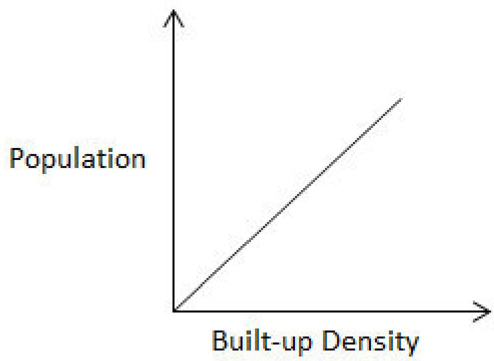


Figure 10: Correlation of the population density and the built-up density

Two approaches, that facilitate a distribution of the population according to the built-up density of the settlement areas, have been developed and tested.
Precondition for the disaggregation, due to residential density, is that the census data and the built-up density correlates. With linear regression it can be tested, if a correlation between areal census dwelling counts and the underlying density levels exists.

First Approach
Based on the data of the study areas in Namibia, the first approach was developed. The study areas, Windhoek and the Cuvelai-Etosha Region (Chapter 2.3), have been selected due to its different settlement structures: dense, medium, and low populated. The constituencies in Windhoek are characterized by a dense settlement area , the cities in the Cuvelai-Etosha Region by a intermidtiate settlement area, and the rural areas in the Cuvelai-Etosha Region by a thin settlement area. The three different settlement structures form the base for the density classes, used in the calculations.
It is assumed that a higher built up density correlates with a higher population density. Due to this, the equation of the Linear Population Distribution (Eq. 1) is amplified by a density factor k (Eq. 2).

$$Pop_i = \frac{Pop}{Pixel} * A * k \tag{2}$$

The density factor is calculated for each of the three density classes by linear regression. The population of a class (x-axis) is related with the GUF area of the same class (y-axis).
Eq. 3 shows the linear equation ($y = a * x$), with y = Pop, x = Pixel, and a = gradient of the straight line, of the linear regression. This equation can be converted to calculate a (Eq. 4).

$$Pop = a * Pixel \tag{3}$$

$$a = \frac{Pop}{Pixel} \tag{4}$$

By combination Eq. 2 and Eq. 4, k can be calculated (Eq. 5).

$$k = \frac{a}{\frac{Pop}{Pixel}} \tag{5}$$

This approach requires a classification into one of the three categories without analyzing the single settlement polygons. The polygons are allocated to a density class, according to the population count instead of the real population density. Therefore, this approach is not discussed further.

Second Approach
The second approach is based on the population disaggregation approach of Steinocher et al. (2005; 2011). It is assumed that data of a region can be distributed locally within the observed region, by using auxiliary parameters, which are related to the aggregated data. In this case, the auxiliary parameter is the built-up density.

The approach of Steinocher et al. (2005; 2011) is based on three assumptions:

- the population density is proportional to the built-up density;
- population only appears within built-up areas;
- within a region the relationship between popualtion- and built-up density is constant.

Before the population count of the administrative unit 'rural district' can be disaggregated in relation to built-up density, it is necessary to define smaller areal units within the administrative unit, in which the built-up density is estimated. The built-up density is calculated in comparison to the open space area. In order to get different densities, the initial spatial unit must be divided into smaller units. Therefore, the research area is fragmented into a grid with a tile size of 250 m, which is also in agreement with Steinocher et al. (2005; 2011), who use 250 m and 500 m tiles. After the built-up density is computed for each grid tile, it is allocated to one of three built-up classes. The built-up classes describe three classes of residential densities:

- built-up class I, low residential density: 0 - 50%
- built-up class II, medium residential density: 50 - 80%
- built-up class III, high residential density: 80 - 100%

Built-up class I defines areas with a low residential density, where only 0 to 50% of the grid tile is settled. Built-up class II defines grid tile with 50 to 80% settlement area, an intermediate residential density, and built-up class III defines areas with a high residential density, where 80 to 100% of the grid tile is settled (Figure 11).

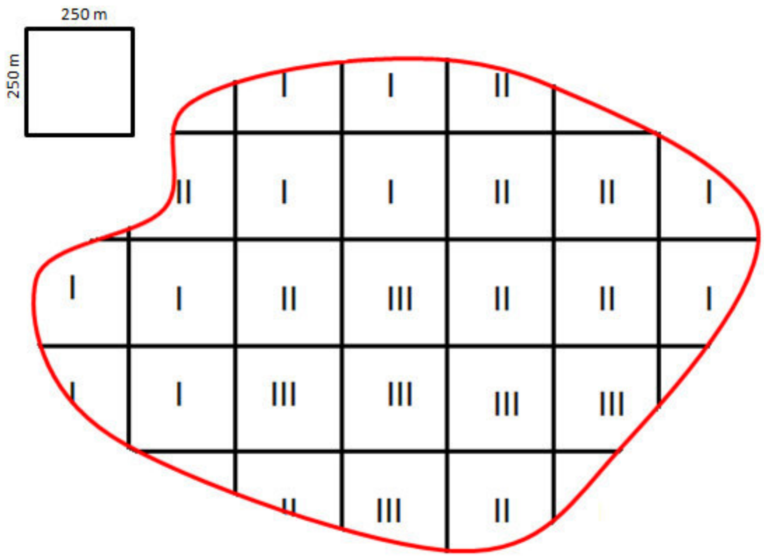


Figure 11: Population Distribution based on Built-up Density - every GUF Pixel receives a population count depending on the density class within the grid tile.

Steinocher et al. (2005; 2011) describe the population density (PopDen) as a function of the built-up density (BUpDen) in the area of interest and the parameter k, which represents the relation between population density and built-up density (Eq. 6).

$$PopDen = k * BUpDen \tag{6}$$

The total population of a region is calculated by a weighted sum function, in which the areal proportion of a density class is multiplied with the weighting parameter k and the average built-up density of the belonging density class (Eq. 7).

$$Pop = \sum_i A_i * k * BUpDen_i \tag{7}$$

The parameter k is constant within a region but varies between different regions. Steinocher et al. (2005; 2011) assume that within a region, the correlation between population density and built-up density remains unchanged.
Two input variables are needed for the calculation of k: census data and the GUF. There are three different built-up classes (low, medium, high) with three different

GUF areas, according to the number of pixels lying in a built-up class. k is calculated by dividing the total population of a region by the sum function of the built-up classes multiplied with the according GUF area (Eq. 8).

$$k = \frac{Pop}{\sum_i A_i * BUpDen_i} \tag{8}$$

As soon as the parameter k is known, the population density of the single built-up density classes can be calculated with Eq. 6.

The result of the Population distribution based on Built-up Density is a population value, which differs in each pixel, depending on the built-up class.

The abbreviations of the single parameters of the calculation used in the equations above, have been adapted to the English names used in the continuous text and therefore differ to the abbreviations in the paper of Steinocher et al. (2005; 2011), which is written in German.

Population Distribution based on Settlement Dispersion

The third method of the population distribution, the Population Distribution based on Settlement Dispersion, describes the influence of different settlement patterns on the disaggregation of population data.
A highly disperse structure of settlement areas allude to a rural area. In comparison, an urban area is characterized by a highly compact structure of settlement areas. Therefore, it is assumed that a high disperse pattern, which signifies rural areas, has a smaller population count per pixel than a compact area like a city.

The following hypothesis is the simplified basic assumption for the approach:

The population density behaves inversely proportionate to the settlement dispersion.

The hypothesis signifies that the higher the settlement dispersion, the smaller is the population count per pixel (Figure 12).

To identify different settlement patterns within an administrative unit, the observed unit is covered by a grid with a cell size of 2000 m. Besides the 2,000 m

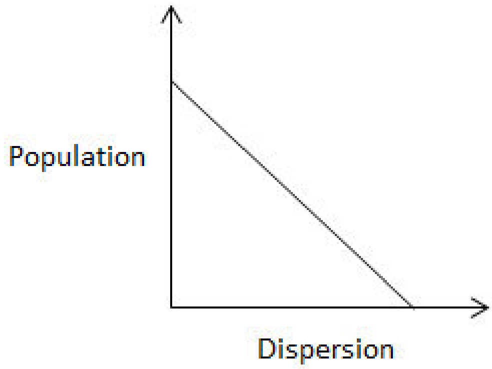


Figure 12: Correlation of the population density and the settlement dispersion

cell size, two additional cell size levels (1,000 and 5,000 m) have been tested. The spatial comparison of the three cell sizes has shown that the 2,000 m cell size has an ideal extent. It is small enough to differentiate an adequate number of tiles within an unit but big enough to identity single settlement patterns.
In each grid tile, a divergent degree of settlement dispersion is visible. The settlement structure is represented by the settlement patches, which are several related pixels in a grid tile. One grid tile comprises a different count of patches (Figure 13).

Two approaches have been tested to identify the dispersion of settlement patterns and to transfer the results on the population distribution. Both approaches assume that the dispersion is described by the size of the settlement patches lying in one grid tile. If a grid tile comprises small patches it is a rural structure, whereas a grid tile comprised of big patches or only one big patch, it is an urban structure. The Population Distribution based on Settlement Dispersion is an expansion of the equation of the Population Distribution based on Built-up Density and should improve the population distribution. It influences the results of the population count of the second method in a positive or negative direction. This means that the result of the dispersion (f_{DI}), affecting the population distribution, must be located between 0.5 and 1.5.

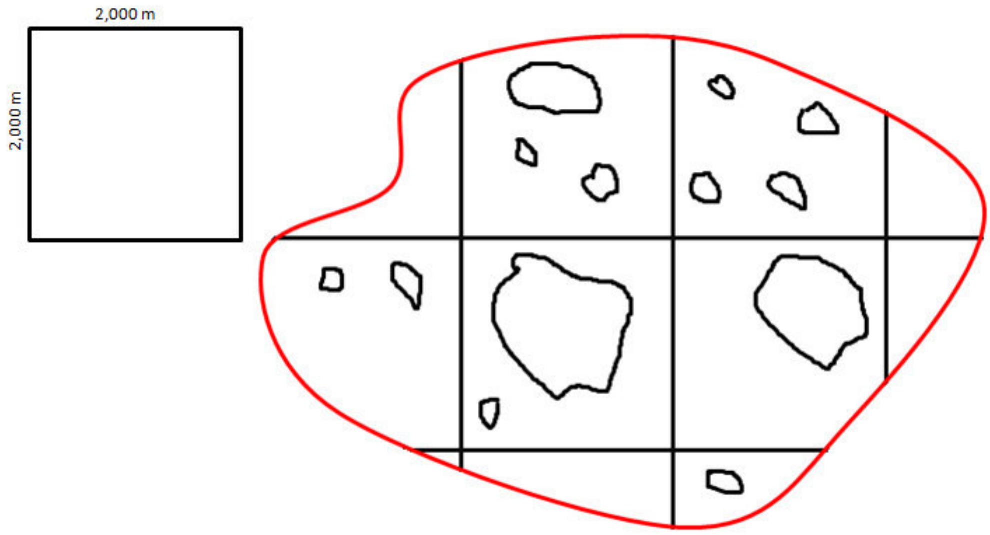


Figure 13: Population Distribution based on Settlement Dispersion - every GUF Pixel receives a population count depending on the degree of dispersion within the grid tile.

First Approach
The first approach, the Dispersion Index (DI), is calculated according to Taubenböck et al. (2014), who developed a concept that describes classified settlement patterns relative to their degree of spatial dispersion using spatial metrics. The aim of the concept is to provide comparability of spatial patterns among each other. The model of Taubenböck et al. (2014) allows, measuring the spatial compactness or the spatial dispersion of two-dimensional settlement patterns with only two spatial input parameters.

The first parameter is the number of Patches (NP), calculated by summing up the number of individual patches (P_i) within a grid tile (Eq. 9).

$$NP = \sum P_i \tag{9}$$

The second parameter is the Largest Patch (LP). It is the size of the biggest patch ($max a_i$) in relation to the size of all patches within a class ($\sum_{i=1}^{n} a_i$), in this case within a grid tile (Eq. 10).

$$LP = \frac{max a_i}{\sum_{i=1}^{n} ai} * 100 \tag{10}$$

The combination of the two parameters describes all possible forming of spatial patterns without considering the shape of the patches and the size of all patches, except the biggest one.
The parameters NP (Eq. 11) and LP (Eq. 12) must be standardized to facilitate a quantification of the degree of dispersion.

$$NP_{norm} = \frac{NP - 1}{\sum_{i=1}^{n} a_i} * 100 \tag{11}$$

$$LP_{norm} = \frac{LP - \frac{1}{\sum_{i=1}^{n}}}{100 - \frac{1}{\sum_{i=1}^{n} a_i}} * 100 \tag{12}$$

The results range between 0 and 100%.

The DI is calculated by the results of NP_{norm} and LP_{norm}, and therefore lies between 0 and 100% (Eq. 13).

$$DI = \frac{NP_{norm} + (100 - LP_{norm})}{2} \tag{13}$$

The calculation of the DI for every settlement pattern is calculated with only two input parameters.

The model of Taubenböck et al. (2014) was developed in order to analyze the settlement pattern and the temporal change of a whole city or similar spatial structures. In this research work, the model is transferred to a much smaller spatial unit and must be aligned to the new requirements.
The studied grid tiles are never totally filled with settlement patches. The highest DI value for one grid tile in Bavaria lies beyond 50%, which makes it impossible to receive a significant indicator for the population distribution, because a DI beyond 50% always influences the population distribution negatively, no matter if the grid

tile is located in urban or rural areas. Therefore, a further standardization is necessary that takes the smallest and the biggest DI value as minimum and maximum borders of the standardization. Furthermore, like already described, the result of the dispersion is an additional parameter in the equation of the Population Distribution based on Built-up density and must be located between 0.5 and 1.5. On account of this, the DI is standardized between 0 and 1 (DI_{norm}) (Eq. 14).

$$DI_{norm} = \frac{DI_i - DI_{min}}{DI_{max} - DI_{min}} \tag{14}$$

The DI is the converse to the population density. A high DI signifies a low population count and in contrast a low DI can be equated with a high population count. Therefore, the result of DI_{norm} must be deducted from 1 (Eq. 15).

$$DI = 1 - DI_{norm} \tag{15}$$

Additionally, the influencing factor (f_{DI}) must be located between 0.5 and 1.5. Hence, Eq. 15 must be amplified with 0.5 to ensure that the results can not be smaller than 0.5 (Eq. 16).

$$f_{DI} = (1 - DI_{norm}) + 0.5 \tag{16}$$

In order to receive a new population disaggregation result, the equation of the Built-up Density must be combined with the influencing factor of the DI (Eq. 17, Eq. 18).

$$Pop = Pop_{PopDen} * f_{DI} \tag{17}$$

$$Pop = (\sum_i A_i * k * BUpDen_i) * (DI_{norm} + 0.5) \tag{18}$$

Second Approach
The second approach, the Area-Weighted Mean (AM), is based on the calculation of landscape metrics. A landscape is defined by an interacting mosaic of patches, which can be distinguished in size, shape, composition, number, and position. These landscape structures are characterized by landscape metrics, which are algorithms quantifying spatial characteristics of landscape patches, classes of patches, or the entire landscape mosaics. Thus, landscape metrics indicate spatial pattern reflecting structure and heterogeneity of the landscape. Primarily, landscape metrics applications originated from the field of landscape ecology, where the relationship between spatial patterns and ecological processes is studied. (Uuemaa, 2009; Zaragozí, 2012)
In this research work, the landscape structure that should be investigated with a landscape metric, are the settlement patches within a grid tile.
It is assumed that in one grid cell only one settlement pattern occurs. Therefore, every grid tile has a different class, which represents an aggregation of patches of the same type. The patches within a grid cell need to be summarized to receive one unique value for each tile. The patches can be summarized by several first- and second-order statistics, like the mean, the area-weighted mean, the median, the standard deviation, etc. (McGarigal, 2014)
Since it is common that the patches in one grid tile do not have exactly the same size or sometimes even differ in a bigger range, the area-weighted mean is used as the summarizing tool. The area-weighted mean calculates, like the normal mean, the average of the settlement patches within one grid tile and additionally considers the different sizes of the patches. The patches are weighted due to their size, which means that a patch with a bigger size has a higher influence on the averaged result than a small patch. Thereby, it is assured that single small patches, which can also occur in city structures, do not have a proportionally influence on the result.
Eq. 19 shows the calculation of the AM of one grid tile. The AM for one grid tile is computed by the sum of the area-weighted patches. One area-weighted patch arises from the multiplication of the area of the patch (a_i) with the area of the patch (a_i), divided by the sum of the area of all patches ($\sum a_i$) within the calculated grid tile.

$$AM = \sum[a_i * (a_i / \sum a_i)] \tag{19}$$

The result of the AM must be standardized to values between 0 and 1 (Eq. 20).

$$AM_{norm} = \frac{AM_i - AM_{min}}{AM_{max} - AM_{min}} \tag{20}$$

In contrast to the DI, the AM is not the converse of the population density. A high AM_{norm} signifies a small dispersion and consequently a high population density. A small AM_{norm} implies a high dispersion and thus a small population density. As already applied at the DI, the influence factor (f_{AM}) is calculated by adding 0.5 to AM_{norm}, ensuring results between 0.5 and 1.5 (Eq. 21).

$$f_{AM} = AM_{norm} + 0.5 \tag{21}$$

The results of the AM are in the next steps combined with the results of the Built-up Density (Eq. 22, Eq. 23).

$$Pop = Pop_{PopDen} * f_{AM} \tag{22}$$

$$Pop = (\sum_i A_i * k * BUpDen_i) * (AM_{norm} + 0.5) \tag{23}$$

3.3.2 Validation

The results of each method of the population distribution algorithm are validated by using a smaller administration level than the one used in the algorithm development (Figure 8).
For the development of the population algorithm, the population count of a district is distributed on the populated pixels within the same district. For the validation of the results, a smaller administrative unit is used. A district can be separated in several communities and therefore the validation unit is the community level. The results of the three algorithm methods are allocated to the communities, which reside in the analyzed district.
The aberration between the census count and the results of the population algorithm methods I, II, or III for one community is calculated to evaluate the result.

The smaller the aberration, the better the algorithm.

For a better illustration, the aberration results of all communities, situated in one district, are summarized to determine the mean absolute error (MAE) for one district. The MAE is the arithmetic average of all aberration results (Eq. 24). The aberration results must be absolute to prevent negative prefixes from influencing the results.

$$MAE = \frac{\sum_{i=n}^{n} |x_i - x_z|}{n} \tag{24}$$

On basis of the MAE, the quality and the improvement of the different population algorithm methods can be measured.

4 Results

During the algorithm development different population distribution accuracies have been reached. In the following chapter, the results of the three developed methods, for the test areas in Bavaria, are compared. Furthermore, the generated algorithm is tested in the study areas in Namibia to analyze its transferability.

4.1 GUF population distribution and validation - Test areas

In Method I (Linear Population Distribution) the population is disaggregated homogeneous on a GUF Pixel. Each populated pixel receives the same population value. Since an equal population distribution does not reflect the reality, in Method II (Population Distribution based on Built-up Density) the algorithm is amplified by the built-up density. To improve the results of Method II, in Method III (the Population Distribution based on Settlement Dispersion) a further factor is added. Besides the density, also the settlement structure is considered.

In Bavaria, 14 rural districts have been chosen to develop the population algorithm. Two in each of the seven regional districts: Oberbayern, Niederbayern, Schwaben, Oberpfalz, Mittelfranken, Oberfranken, and Unterfranken (Appendix A). The selected districts present different settlement and population structures. In the district Munich, one of the communities has the biggest population count (1,388,308 POP) in Bavaria. While in the district Haßberge, only communities with small population counts, between 590 and 13,090 people occur. In Mittelfranken, as well as in Unterfranken, one of the rural districts is composed by the main cities of the region. In Mittelfranken the cities Ansbach, Fürth, Schwabach, Erlangen, and Nuremberg are considered and in Unterfanken the cities Aschaffenburg, Schweinfurt, and Würzburg.

Like already described in Chapter 3, the population is distributed on the basis of the district level. The results of the disaggregation on district level are evaluated by summarizing all settlement pixels situated within a community, which is located within the analyzed district. The calculated population result of the community is then compared with the census count of the regarded community.

In the following, the results of the three population disaggregation methods are visualized. Based on the aberration between the summarized community results and the census data, the MAE (Chapter 3.3.2) is calculated for each district. Furthermore, the development of the population disaggregation in the main city of each district and the diverse outcome for urban and rural areas are investigated. For a better understanding all results are computed in percent.

4.1.1 Linear Population Distribution

The results of the Linear Population Distribution facilitate a comparison of the census data and the calculated population count of an equal population distribution. The comparison features a significant trend with a population overestimation in rural communities and a population underestimation in urban communities.
This trend applies to 13 of the 14 investigated districts in Bavaria. It is especially visible in districts, where larger cities are one of the communities. In the district Fürth for example, the population count of the same named community and the biggest city in the district is underestimated by 25% (30,000 people). In contrast, the population count in 13 of the 15 communities in the district Fürth, is overestimated (Figure 14). The district Landshut and Neu-Ulm feature a similar pattern.

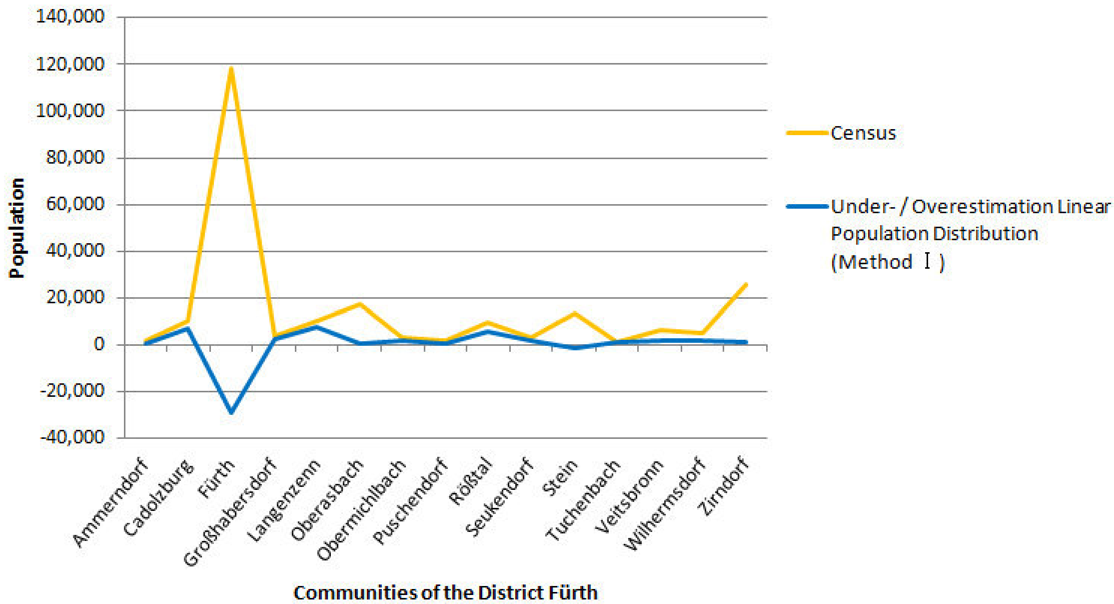


Figure 14: Population over- and underestimation in relation to census data in the communities of the district Fürth.

Especially in the rural district Munich, where the same named community and largest city of Bavaria is located, the trend of urban underestimation and rural overestimation can be observed (Figure 15). In each of the 30 communities the population count is overestimated, except in the community Munich, where the population is underestimated by 26% (365,000 people).

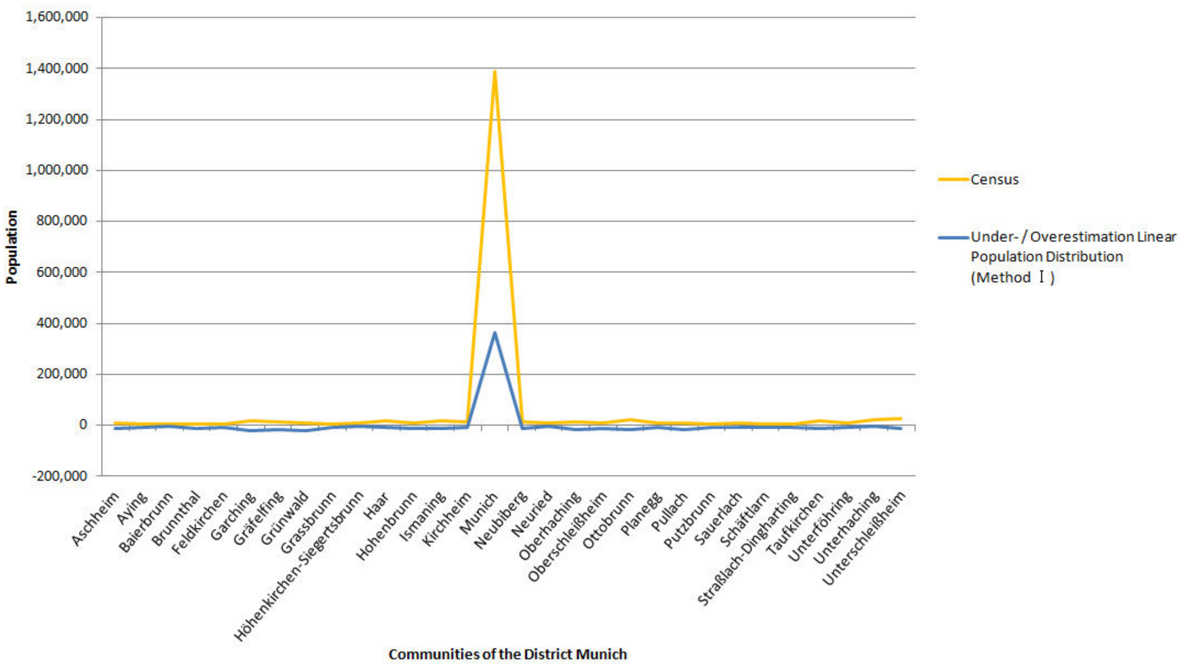


Figure 15: Population over- and underestimation in relation to census data in the communities of the district Munich.

The only exception of the 14 districts, which is not directly following the trend, is the district Haßberge. The over- and underestimation is evenly distributed between the communities.

The identified trend indicates that the population is underestimated especially in areas, where a high built-up density is suspected. In the community with the biggest population count within a district, the predicted population count is always underestimated in the linear population distribution model (Figure 16).

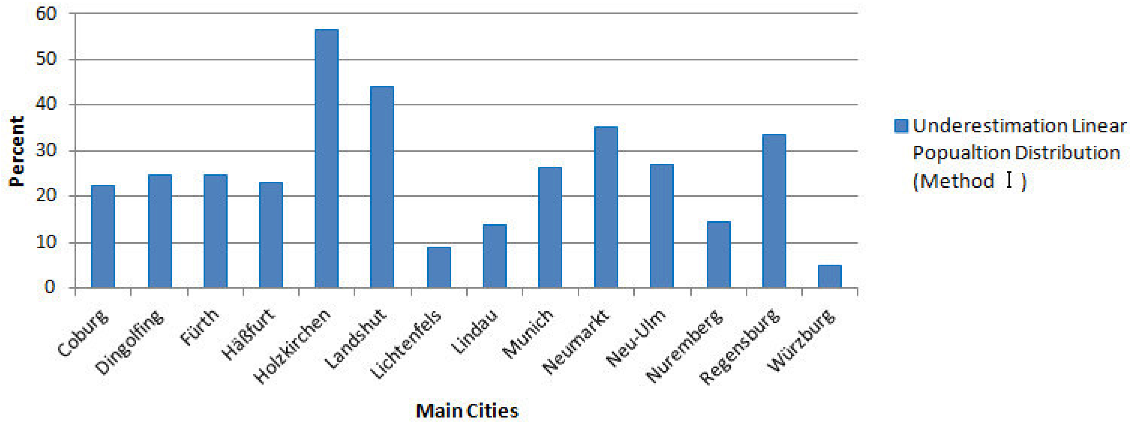


Figure 16: Population underestimation of population counts by the linear distribution in urban areas.

4.1.2 Population Distribution based on Built-up Density

The population distribution model based on built-up density is an improvement of the linear population distribution model. The results of the disaggregation will show if the hypothesis ***"The population density behaves directly proportionate to the built-up density"*** can be accepted.

The improvement of the population distribution with the second method is illustrated by the MAE of the results of the built-up density in comparison to the MAE of the linear distribution. Figure 17 shows the MAE of Method I (blue line) and the MAE of Method II (green line) in relation to the census data in percent. Considering the built-up density leads to an enhancement of population counts in 13 of the 14 observed districts.

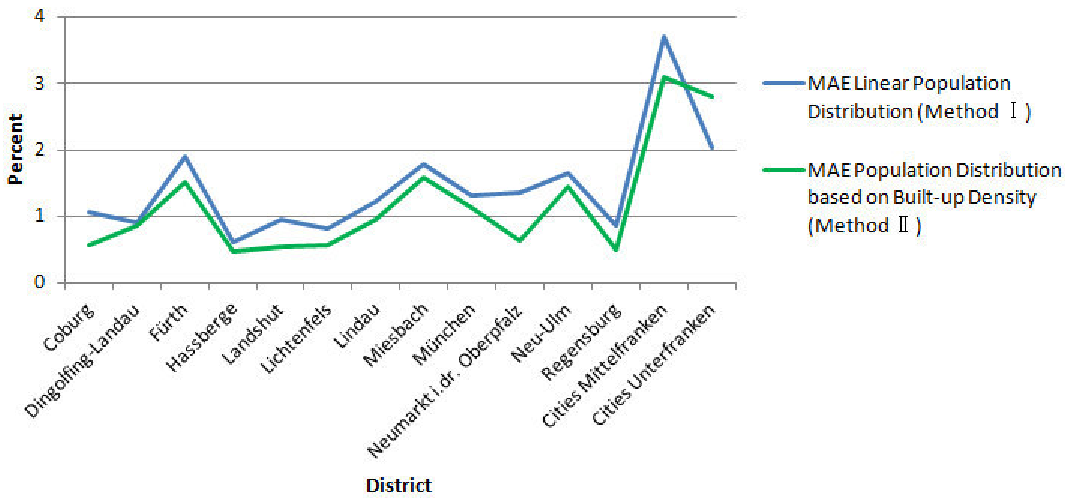


Figure 17: Comparison of the MAE of Method I and Method II for the 14 districts in relation to census data.

The underestimation in urban areas and the overestimation in rural areas is also less pronounced in the population distribution model using built-up density (Figure 15). In each of the analyzed main cities, except Würzburg, the aberration between census data and the population distribution results become smaller by taking the degree of built-up density into account.

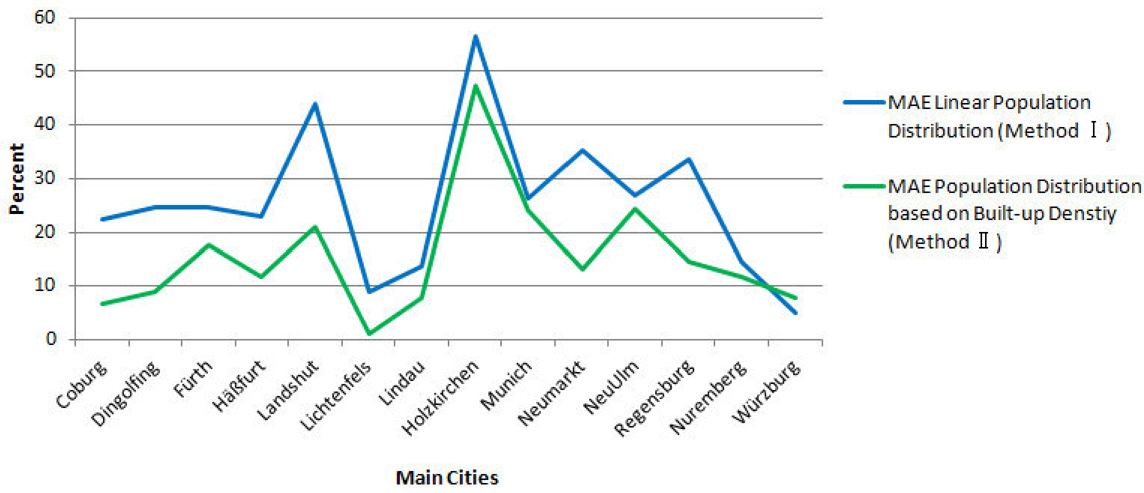


Figure 18: Comparison of the MAE of Method I and Method II for the main cities in relation to census data.

Despite the improvement of the population distribution, taking the built-up density into account, the approach still has deficiencies in properly assessing population counts in urban areas (Figure 15). There is still an underestimation of population in urban areas, which ranges between 1% in Lichtenfels and 47% in Holzkirchen. The most common aberrations to census data lie between 10% and 20%.

4.1.3 Population Distribution based on Settlement Dispersion

A further enhancement of the population disaggregation algorithm is made by taking the built-up density and dispersion of settlement structures into account, in order to minimize the deviation of the MAE to the census count. With the third method, the influence of a dispersion parameter is calculated.
Two different approaches, the DI and the AM, have been tested to improve the diasaggregation.
The results of the disaggregation will show if the hypothesis ***"The population density behaves inversely proportionate to the settlement dispersion"*** can be accepted.

Dispersion Index

The results of the DI show a decline of the population disaggregation. In Munich, one of the districts in the test-area characterized as urban, every community is overestimated. The calculated population count deviates from the census data, of the single communities, between 24% and 308%. In the district Coburg, one of the districts characterized as rural, similar results are predicted. Every community is overestimated. Based on the poor outcomes of the DI, the results are not further listed.

Area-weighted Mean

The results of the AM approach feature two trends. In districts with a community characterized by a highly compact settlement structure, the consideration of the dispersion parameter AM leads to an improvement of the results. In five of the seven districts, characterized as urban, the MAE of Method III is lower than the MAE of Method II (Figure 19). The positive influence of the parameter is particularly high in the districts Munich and the Cities of Mittelfranken.

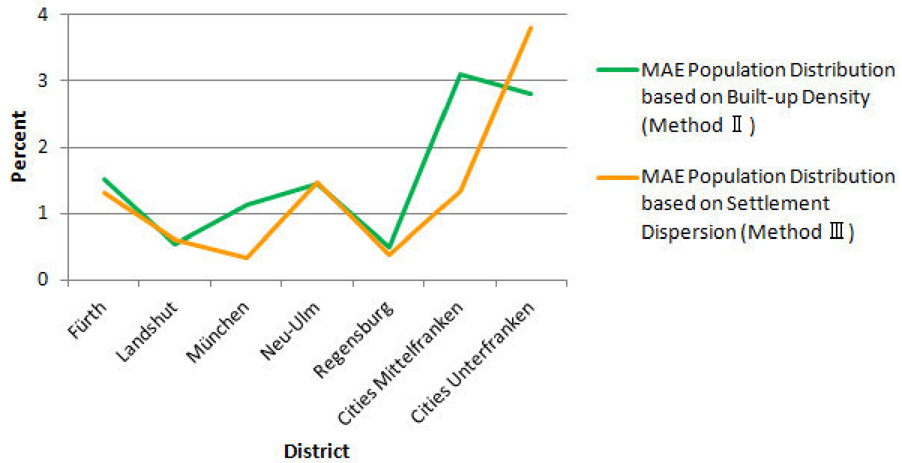


Figure 19: Comparison of the MAE of Method II and Method III for urban characterized districts in relation to census data.

In contrast, the AM leads to a decline of the MAE in the districts, which are mainly characterized by communities with a disperse settlement structure (Figure 20). The total population count of each rural district is underestimated.

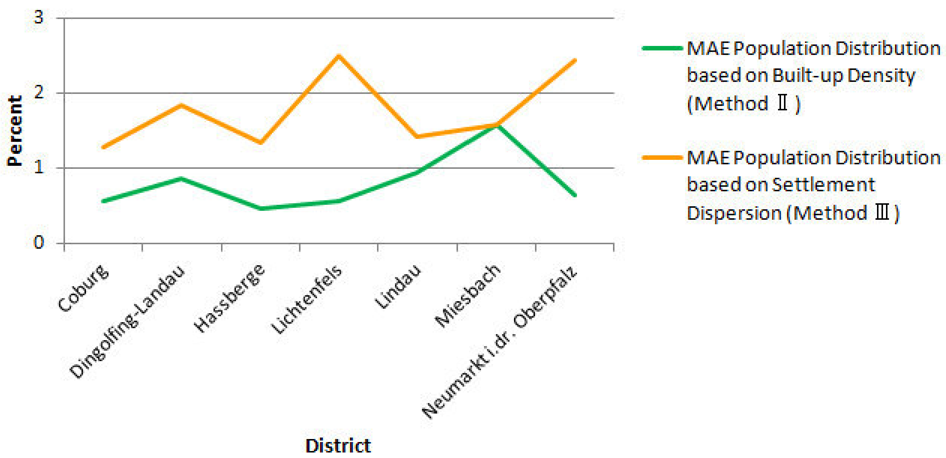


Figure 20: Comparison of the MAE of Method II and Method III for rural characterized districts in relation to the census data.

The comparison of the disaggregation results of Method II and Method III for the main cities illustrates that the dispersion parameter usually only leads to an improvement of the aberration in very dense populated areas (Figure 21). In the largest cities in the region, Munich (1,388,308 POP), Nuremberg (495,121 POP), Regensburg (138,296 POP), and Würzburg (124,577 POP), the disaggregated population count, calculated by the AM, deviates little from the census population count. The difference between the census count and the calculated count in Munich, the city with the most inhabitants in the considered study area, is only 0.57%. In the second biggest city, Nuremberg, an even smaller difference (0.47%) is reached.

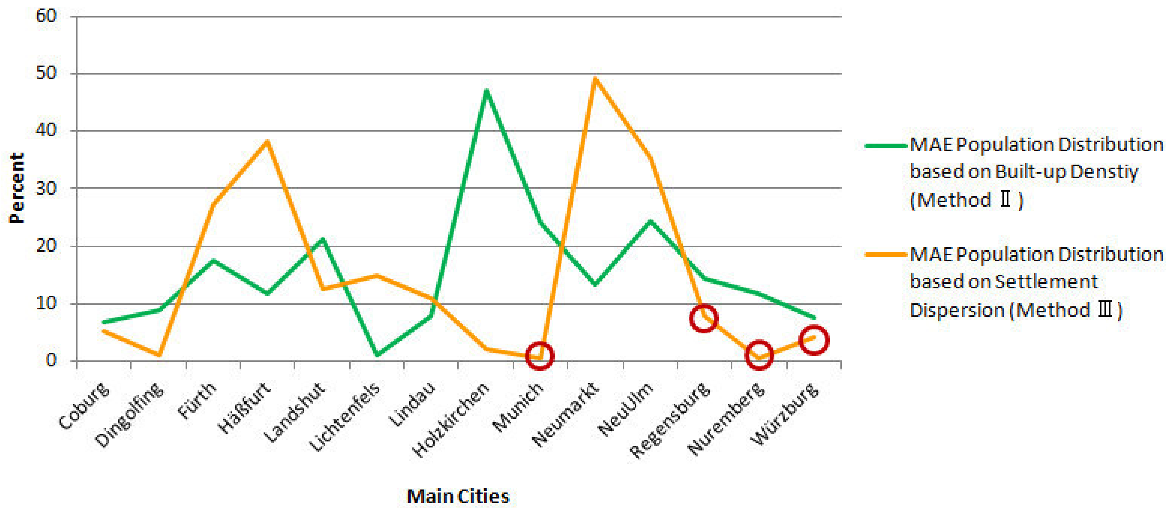


Figure 21: Comparison of the aberration of Method II and Method III for the main cities of each district in relation to the census data.

The results of the single communities within a district clarify that the consideration of the settlement dispersion leads, in each district, characterized as rural, to a population underestimation in almost every community. If only the built-up density is considered, the under- and overestimation of population counts are evenly spread within a district.
Extending the calculation by a dispersion parameter causes an underestimation of the total population of each rural district. In districts, characterized as urban, the total population count also decreased in districts with smaller cities, like Fürth, Landshut, Neu-Ulm, and Regensburg. Only in the district Munich, where the city Munich is one of the communities, and in the non-official districts (Cities Unterfranken and Cities Oberfranken) the total population count was overestimated.

In conclusion, the results display a significant improvement of the population disaggregation from Method I to Method II. From Method II to Method III only an improvement of the population disaggregation can be achieved in areas with a very high residential density. Since the consideration of the dispersion parameter causes a significant improvement in the highly dense urban structures, the result of Method III for the whole test area is better than the result of Method II for the whole test area. The MAE of Method II improves by over 15% in comparison to the MAE of Method I. Whereas the MAE of Method III improves by over 35%.

4.2 Transferability

Based on the results of Chapter 4.1, the population disaggregation algorithm is transferred to the study areas in Namibia. Since the results of the DI were significantly worse than the results of the AM in the test area Bavaria, the AM approach is applied.

In Namibia, it is difficult to identify a trend with the Linear Population Distribution. Two spikes (Windhoek West and Windhoek East) distort the results of the other 13 constituencies. In Windhoek West and Windhoek East the population count is extremely overestimated, wherefore in every other constituency the population count is underestimated, independent of the built-up density (Figure 22).

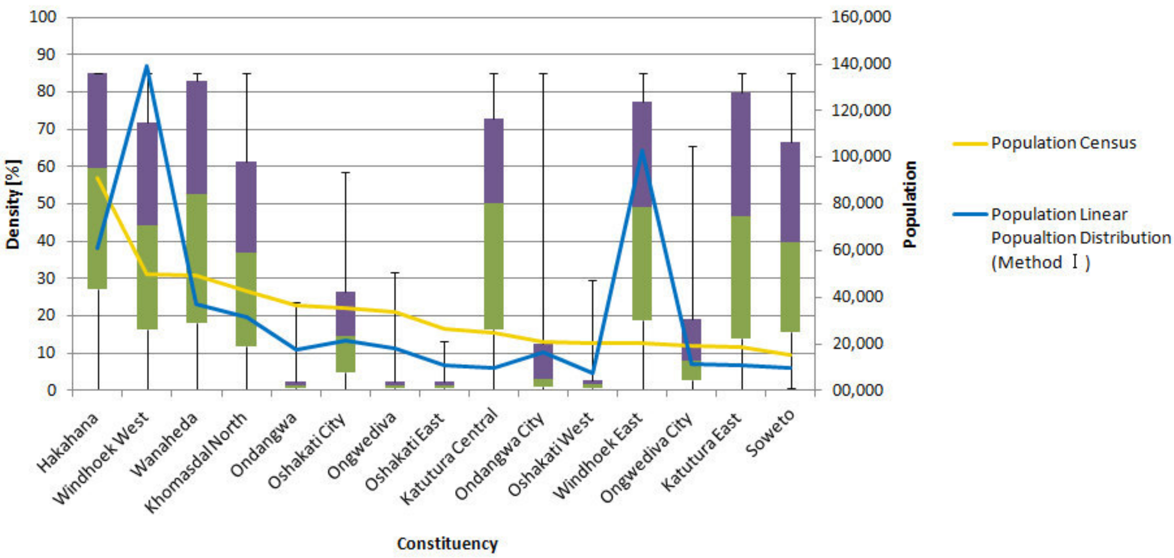


Figure 22: Comparison of the population distribution of the official Census and the Linear Population Distribution in the study area Namibia in relation to the built-up density, which is visualized by the boxplots.

Since both spikes are located in Windhoek, the population disaggregation of Method II and III is calculated separately for the study area 'Windhoek' and the study area 'Cuvelai-Etosha' to reduce the distortion of the results. The constituencies of the respective study area are listed in Appendix B.

Windhoek

The comparison of the MAE of Method I, Method II, and Method III, for the study area 'Windhoek', shows that the aberration to the census is not influenced by the built-up density. Method III displays a slight deterioration in comparison to Method I and III (Figure 23). Taking the settlement dispersion into account deteriorates the population disaggregation.

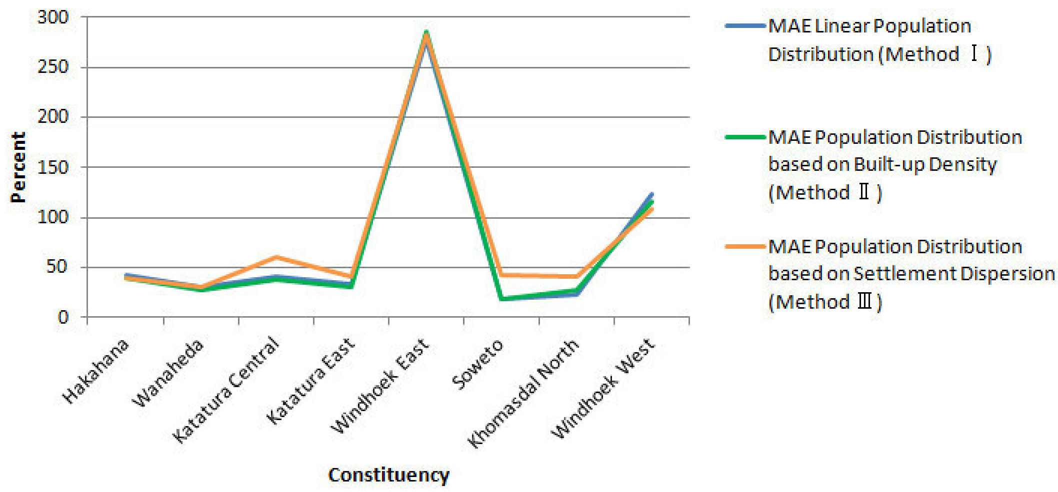


Figure 23: Comparison of the MAE of Method I, Method II, and Method III for the constituencies of the study area Windhoek in relation to the census data.

In conclusion, in the constituencies of Windhoek, an improvement of the population disaggregation based on built-up density and settlement dispersion is not possible.

Cuvelai-Etosha

In the Cuvelai-Etosha the consideration of the built-up density causes a significant improvement of the population disaggreagtion (Figure 24), particularly in the rural areas (Ondangwa, Ongwediva, Oshakati East, and Oshakati West). In the urban areas (Ondangwa City, Ongwediva City, and Oshakati City) different results are visible, like a small worsening in Ondangwa City and a considerable improvement in Oshakati City, where the MAE reaches a deviation from the census count of only 2%.

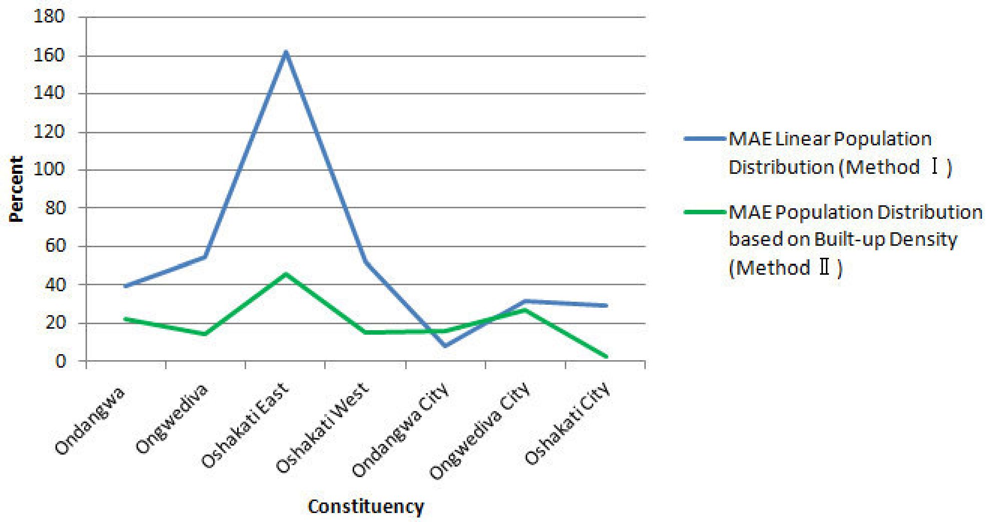


Figure 24: Comparison of the MAE of Method I and Method II for the constituencies of the study area Cuvelai-Etosha in relation to census data.

The comparison of the disaggregation results of Method II and Method III show a sharp deterioration of the results, under the consideration of settlement dispersion (Figure 25). Only in the constituency Oshakati East an improvement of the disaggregation was possible.

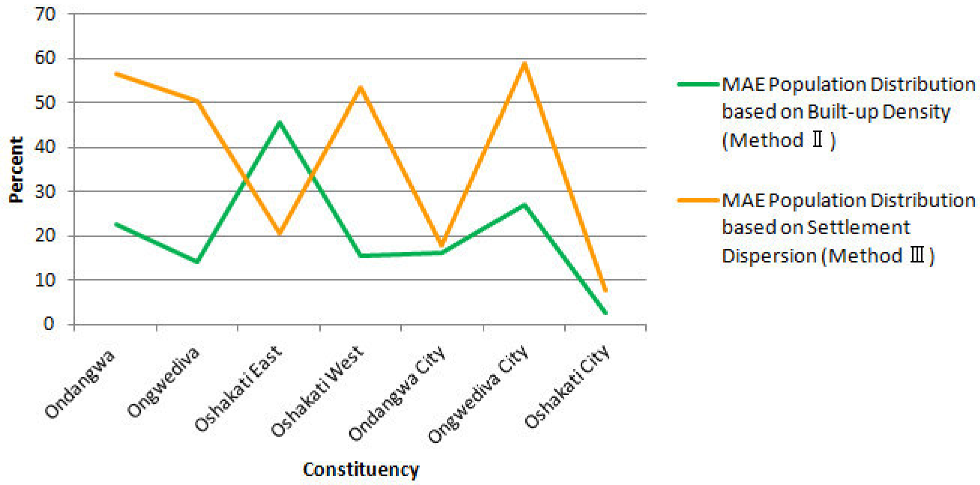


Figure 25: Comparison of the MAE of Method II and Method III for the constituencies of the study area Cuvelai-Etosha in relation to census data.

Method II improves the population disaggregation, while the dispersion parameter degrades the result of the built-up density. However the results of Method III are better than the results of the linear distribution (Figure 26).

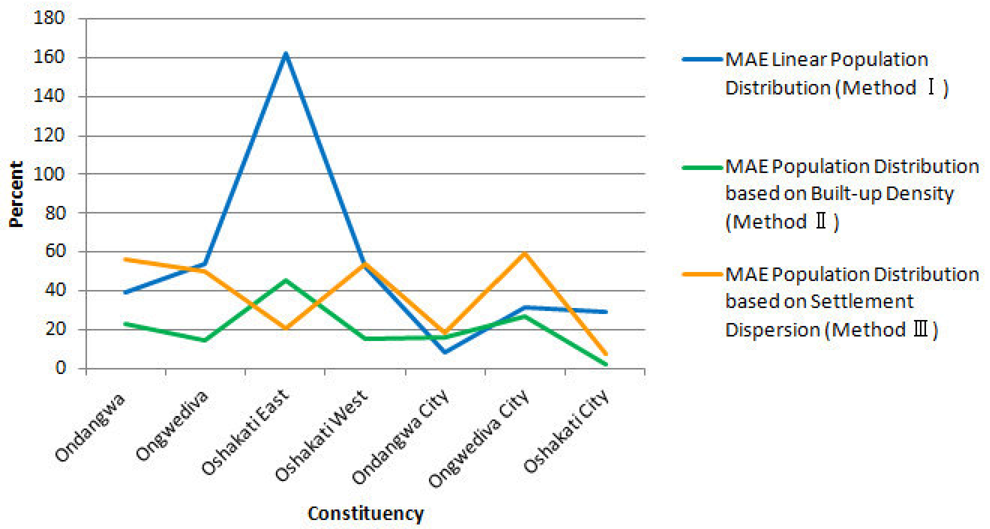


Figure 26: Comparison of the MAE of Method I, Method II, and Method III for the constituencies of the study area Cuvelai-Etosha in relation to the census data.

The results of the GUF population distribution and its transferability are discussed in Chapter 6.

In the following chapter the applicability of the new developed approach is tested, in view of the flood event of 2009 in the northern parts of Namibia.

5 Case study

Floods are a regular phenomena in Namibia, depending on the annual rainfall. They have devastating effects on infrastructure and people located in flood prone areas. The impacts of floods are not evenly distributed among regions. Different socio-economic groups are affected in distinct levels. For decision makers it is important to understand the recurrence of flooding to be able to accomplish an appropriate risk assessment and to prepare mitigation plans. (Gilau, 2011)
In Namibia, six of the the thirteen regions of the country are affected by flooding and half of the Namibian population lives in flood prone areas. Between 2006 and 2009, floods have occurred almost every year and influenced up to 677,500 people. The floods of 2004, 2006, 2009, and 2011 have been the most notable ones in the last 50 years. Floods in Namibia are mainly caused by a high rainy season. Especially the northern part, where the occurrence of wet conditions mostly cause flooding, is exposed to a high risk. (Gilau, 2011)

5.1 Flood event 2009 Namibia

In March 2009, torrential rainfalls in Angola, Namibia, and Zambia appeared and caused an increase of the water levels in the Chobe, the Kuene, the Kavango, and the Zambia River. This led to the worst flooding in the north-central and the north-eastern regions of Namibia since the 1960s and 1970s. The water reached heights up to 7.85 m, which have not been recorded since 1963. The area damaged by inundation was home to 60% of the population.
The flood of 2009 highly affected six regions in the northern parts of Namibia: Caprivi, Kavango, Ohnagwena, Omusati, Oshana, and Oshikoto. It made roads impassible and blocked access to local health facilities and schools. Furthermore, water supply and sewage stations were inundated, electricity provision was compromised, and the entire economy was interrupted for three months. In addition, most of the affected population had not fully recovered from the rainfalls in 2008 and was therefore particularly vulnerable. (Government of the Republic of Namibia, 2009)
The population, which lives in regions affected by floods, mostly inhabits the rural areas. They are dependent on subsistence farming and not only lose their homes but often also their annual harvest and livestock during flood events.
The average income in the six regions is half of the national average and one third of the countries poorest population live there. Especially the vulnerable portion of the population lives exposed to flood events and was affected the most by the flood of 2009. (Government of the Republic of Namibia, 2009)
According to the United Nations Disaster Assessment and Coordination (UNDAC)

350,000 people, 16.5% of the total population of Namibia, have been affected by the flood of 2009, while local authorities count 700,000 people (33%). 85% of the households in the six affected regions are located in rural areas. Most of these are traditional houses, built of wood, wood combined with mud, or sand bricks and are therefore easily destroyable by the deluge of water occurring during flood events. (Government of the Republic of Namibia, 2009)

The Cuvelai-Etosha Basin, which consists of the four regions Ohangwena, Omusati, Oshana, and Oshikoto, is characterized by a dense network of ephemeral rivers (Chapter 2.3), which only contain water during the rainy season (November to March). After the end of the civil war in Angola, trade between Namibia and Angola arose and released a rapid growth in the four regions. Now, the Cuvelai-Etosha is the most populated area of Namibia (800,000 POP) and was also mostly influenced by the flood in 2009. 83% of the population have been affected. Besides destroyed homes and widespread displacement, agricultural land and livestock got lost, and huge parts of critical infrastructure were destroyed. In urban areas the risk of epidemics of waterborne diseases, caused by flood water mixed with sewerage, arose. (Government of the Republic of Namibia, 2009)
In the Cuvelai-Etosha three different types of housing exist: traditional housing, modern (solid) housing, and informal housing. Especially traditional dwellings are located along flood plains and rivers and are mostly built on ground level. Since risk reduction measures have not been included in housing policies, because floods are regarded as a new phenomena in Namibia, none of the three housing types is prepared for floods. 39% of the households have been fully destroyed by the flood of 2009. (Government of the Republic of Namibia, 2009)
Beside the homesteads, also the crop fields and livestock are located in flood prone areas, which makes the population particularly vulnerable to flooding. Poor households are most vulnerable to the impact of floods, because their living conditions can deteriorate very quickly. (Government of the Republic of Namibia, 2009)
The flood of 2009, particularly concerned the private sector and highlighted areas of priority interventions (Government of the Republic of Namibia, 2009).

5.2 Rapid Flood Loss Estimation

In order to facilitate a fast post disaster management, knowledge about the extent of the consequences of a natural hazard is necessary. Therefore, knowledge about the magnitude and intensity of the hazard and about the population distribution is required. A fast data acquisition of the extent of a hazard can be gained by the analysis of remote sensing data, indicating the area of interest shortly after the occurrence of the hazard event. The population distribution within the hazard

zone and consequently the number of affected population is given by population disaggregation approaches.
In this case study, the occurring hazard is a flood and only humanitarian impacts are considered. In order to achieve a count of the affected population, the hazard area must be combined with the spatial referenced population distribution (Figure 27).

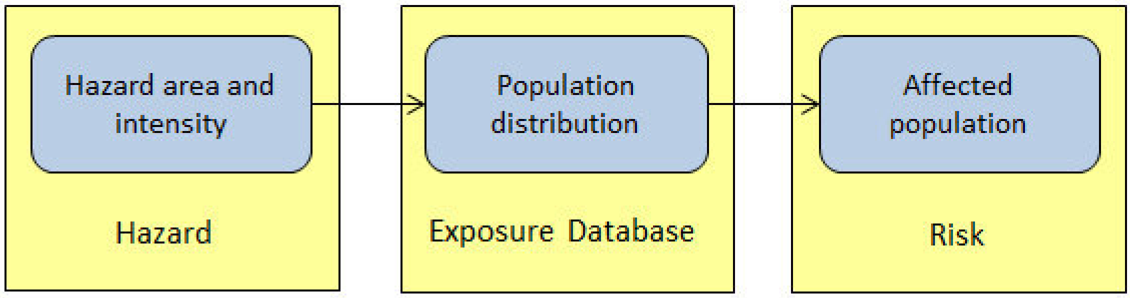


Figure 27: Concept of the Rapid Loss Estimator

The quality of the input (Population distribution) has a huge influence on the accuracy of the outcome (Affected population), which is clarified in the following subsection.

5.3 Comparison of affected population

In this subsection, the population distribution results of the new developed approach for the investigated area in the Cuvelai-Etosha are combined with a flood mask of 2009 to gain an overview of the affected population. Furthermore, the results are compared with the outcomes of the two worldwide population disaggregation approaches WorldPop and LandScan.
The area of interest is located in the region Oshana, one of the four regions of the Cuvelai-Etosha (Figure 28).

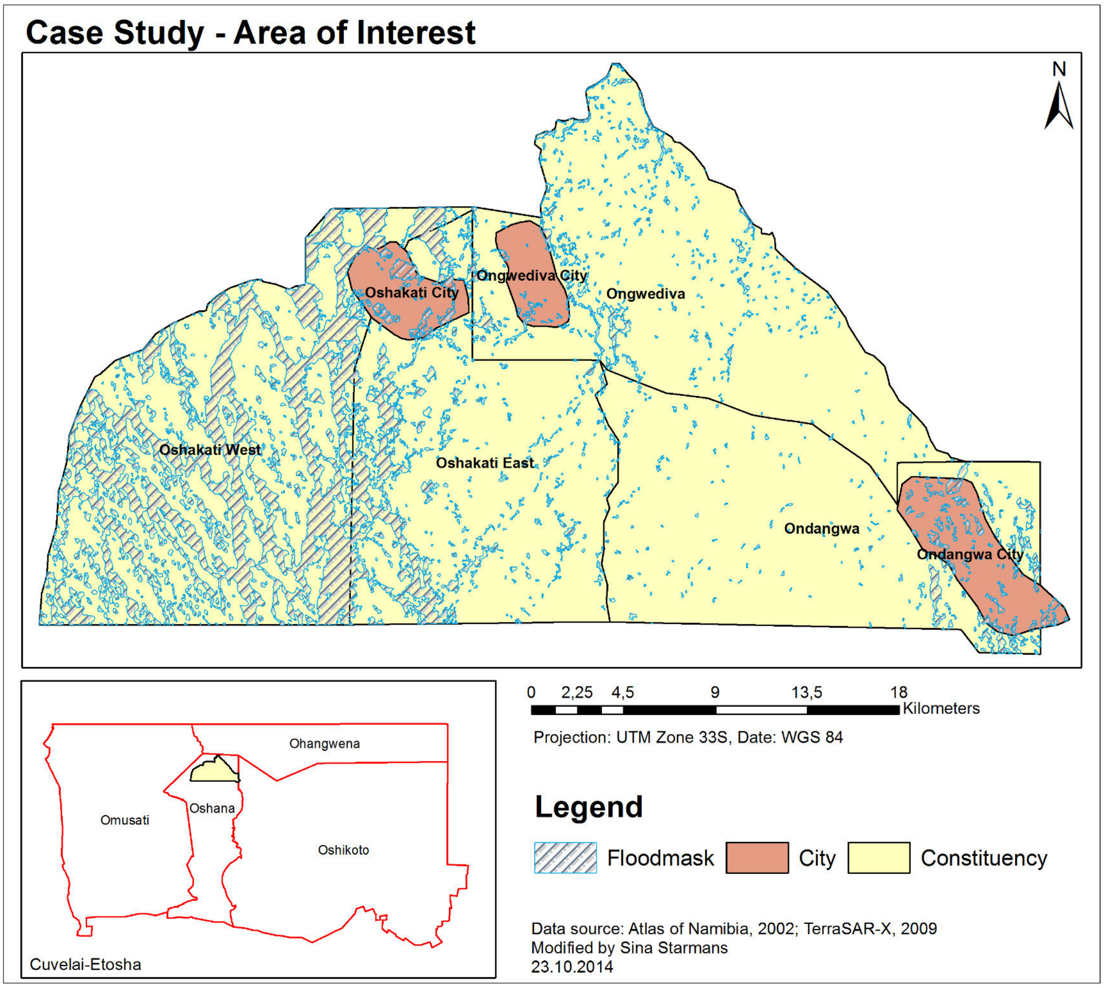


Figure 28: Area of interest in the case study in the Cuvelai-Etosha

5.3.1 Report information

In the region Oshana, 69% of the population live in rural areas, which are characterized by traditional housing. Traditional housing in this region is highly vulnerable to floods, because they generally lie at lower ground level and the applied building materials are not flood resilient. The urban areas are more resistant, because the houses are particularly built with solid materials and furthermore urban areas typically occupy higher ground levels. Urban areas are particularly vulnerable when the drainage system does not work adequately. Besides traditional and solid houses also informal dwelling structures exist, mostly around urban areas. They are highly vulnerable to floods, because of their location and their susceptible con-

struction. (Government of the Republic of Namibia, 2009)
In the region Oshana, around 162,000 people were affected by the flood of 2009. 87% of the households were affected, 9% of them were fully destroyed and 35% were partially damaged. 58% of the affected households were traditional dwellings and nearly all of them lost their grain and food reserves due to inundation. (Government of the Republic of Namibia, 2009)

5.3.2 GUF population tool results in comparison to the results of WorldPop and LandScan

The GUF population tool results are, due to the data basis, only available for four constituencies and three cities in the region Oshana. In the three cities, less people are affected by the flood than in the four rural constituencies. This reflects the situation of the whole region Oshana, where more than half of the affected households are located in the rural areas. The population in the constituencies Oshakati East and Oshakati West is affected the most. In Oshakati East 43% and in Oshakati West 55% of the population is influenced by the flood of 2009. In the other two rural constituencies, Ongwediva and Ondangwa, 28% and 27% of the population is affected. In the three cities Ondangwa, Ongwediva and Oshakati the influenced population ranges between 2% (Ongwediva) and 13% (Ondangwa) (Figure 29). 34% of the population in the four rural constituencies is affected by the flood, while in the three regarded cities only 9% of the population is influenced.

The results of WorldPop show that only 2% of the whole population of the regarded area are considered as affected population. Almost 2% of the urban population and around 5% of the rural population is affected by the flood of 2009. The most affected urban areas are also Oshakati East and Oshakati West. In Oshakati East 11% and in Oshakati West 17% of the population is influenced by the flood event (Figure 29). In the three regarded cities, the degree of impact ranges between 0% (Ongwediva) and 3% (Oshakati).

The calculation according to the population data of LandScan predicts a higher count of affected population than WorldPop. 13% of the population is affected by the flood. LandScan shows a contrary development of the influenced population, compared to the new developed approach and to WorldPop. In rural areas a lower count of population is affected than in urban areas. 11% of the rural population is influenced by inundation, while in urban areas only 14% of the population is affected (Figure 29). The most affected rural areas are Oshakati East and Oshakati West, where 19% and 31% of population are affected. Oshakati City is the most affected urban area (25% of population affected).

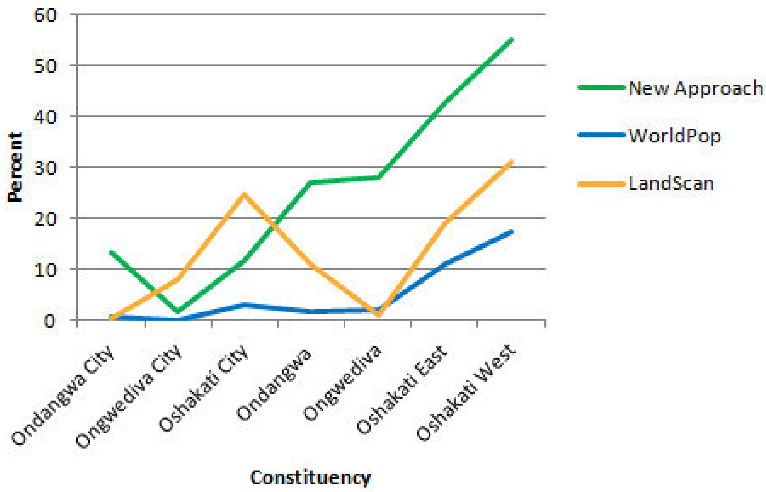


Figure 29: Comparison of the number of affected population per constituency of the New Approach, WorldPop and LandScan.

6 Discussion

The results of the three disaggregation methods in the different study areas feature similar tendencies. The integration of the built-up density improves the population disaggregation, while the dispersion parameter only causes an enhancement in highly dense settlement areas.
In the following Chapters (6.1 and 6.2) the main results, listed in Chapter 4, are analyzed and discussed. Furthermore, in Chapter 6.3 the applicability of the new developed disaggregation approach in comparison to the population distribution approaches LandScan and Worldpop is considered.

6.1 Test-area Bavaria

In this chapter only the outcomes for the Linear Population Disaggregation, the Population Disaggregation based on Built-up Density, and the Population Disaggregation based on Settlement Dispersion for Bavaria are discussed.

Linear Population Disaggregation

The results of the Linear Population Distribution show huge differences between the census data and the calculated count of the linear disaggregation in the test area Bavaria. The population is mainly underestimated in urban areas and overestimated in rural areas. Every regarded district in Bavaria, except the district Haßberge, follows this trend. In Haßberge no distinct trend is visible, the overestimation and underestimation of population counts are scattered on the communities. This is reasoned by the structure of the district, where only a small range of population counts among the communities occur.
A reason for the trend, in the other 13 districts, is the different building architecture of rural areas and urban areas. In rural areas, buildings normally have a small size and a low building height, while in urban ares mostly big buildings like apartment houses or even skyscrapers occur. Since the GUF, the basis for the settlement areas, does not imply height information, the linear disaggregation approach distributes population counts only under allowance of the extent of the settlement area. According to the building height, more people live per GUF Pixel in urban areas than in rural areas. Therefore, further steps are necessary to improve the population disaggreagtion.

Population Disaggregation based on Built-up Density

In the second method of the population distribution, it is assumed that a correlation between areal census dwelling counts and the underlying density levels exists.
Therefore, the following hypothesis is proposed:

The population density behaves directly proportionate to the built-up density.

In the test area Bavaria, in 13 of 14 districts the consideration of the built-up density leads to an improvement of the MAE and consequently to a reduction of the aberration between the census data and the predicted population data. Especially the results of the urban areas, which have been considerable underestimated, are improved by the new calculated population disaggregation.
The district Cities of Unterfranken is the only district where the built-up density leads to a deterioration of the disaggregation results. It is composited by the biggest cities in Unterfranken and was created to analyze the behavior of population disaggregation in a district with similar communities. In the Cities of Unterfranken, the consideration of the built-up density leads to an enhancement of the over- and underestimation of the linear disaggregation. In contrast, in the district Cities of Mittelfranken, the aberration between the census data and the predicted population data considering the built-up density is smaller in each city except one, than in Method I. Nevertheless, the results of the cities in Unterfranken show that the application of the built-up density is particularly useful when different settlement structures exist in the investigated area.
However, an administrative unit, which is one of the bases for the disaggregation, normally is characterized by different settlement sizes. Therefore, the results of the 12 official districts are the essential indicators for the quality of the disaggregation outcomes. In each of these 12 districts the built-up density leads to an improvement of the MAE. Hence, the hypothesis, that the population density correlates positively with the built-up density, can be confirmed.

Population Disaggregation based on Settlement Dispersion

The built-up density leads to an improvement of the disaggregation, but nevertheless in urban areas a huge population underestimation occurs. In order to improve the results of the built-up density, the dispersion of settlements is considered as a further influence factor on population distribution.

Therefore, the following hypothesis is proposed:

The population density behaves inversely proportionate to the settlement dispersion.

In the third method of the population disaggregation, two approaches have been developed and tested. The results of the first approach, the DI, deviates strongly from the census data. During the calculation of the DI, the number of patches within a grid tile and the size of the biggest patch in relation to the other patches within the same tile are considered. As soon as only one patch occurs within a grid tile, it is assumed that the settlement area within this tile is highly compact. In this case, the size of the patch has no influence and even a patch, consisting of only one Pixel (12*12m), signifies high compactness. A highly compact tile implies a high population count within the proposed hypothesis. Every tile involving only one small patch, at least of the size of a pixel, has a small DI and therefore is regarded as a grid tile in a high populated area. A small DI implies a highly positive influence factor on the population distribution. Therefore, the population is overestimated in every community without being able to distinguish between rural and urban communities.
Actually, the DI was developed to predict the dispersion of a whole city and not the dispersion within a 2,000 m grid tile. The areas calculated in this research work are to small for the approach of the DI.

The results of the second approach show that the influence of the AM as dispersion parameter only leads to an improvement of the results in high compact settlement patterns, like in the district of Munich. In districts characterized by a more rural structure, the AM leads to a deterioration of the results.
In the district Munich, the population count of the community Munich is approximately 125 times larger than the average population count of the other communities. If only the built-up density was considered, in each community except Munich, the population count is still strongly overestimated. In Munich it is underestimated by 24%. Taking into account the settlement dispersion as a second influence factor, the population count for the city Munich is overestimated by 1%. The consideration of the AM leads to a significant improvement of the results of a community, where the population count and the area of settlement differ widely. This shows the huge influence of the settlement dispersion and explains why the total population count of almost every other analyzed district was underestimated in a huge scale. The calculated settlement parameter ranges between 0.5 and 1.5, whereas 0.5 signifies a huge dispersion and 1.5 signifies total compactness. Huge settlement parameter values (1.5) in a large count mainly only occur in really dense

settlement areas, while small values occur in a large scale. This causes a population underestimation, where no extensive highly dense settlement area exists.
The settlement dispersion factor has a positive influence on highly dense populated areas but otherwise causes a deterioration of the results. Therefore, the hypothesis that the higher the settlement dispersion, the higher the population density, cannot be confirmed. The calculation of a settlement parameter cannot be inversely proportionate to the population density. It has to be non-linear inversely proportionate, because it only leads to a significant improvement in highly dense areas, while in medium and low dense areas it leads to a deterioration of the results which increases with the dispersion.

In conclusion, if all districts are summarized, the results of the settlement dispersion have a lower MAE than the results of the built-up density. This is reasoned by the much better outcomes for the huge urban areas, like Munich, whose MAE have a strong influence on the summarized MAE for all districts together. An extension of the population distribution based on built-up density with a settlement dispersion parameter causes an improvement of the results in highly dense areas, because the higher the compactness of an investigated area, the better the result of the dispersion parameter. Though, in rural areas its the opposite, the higher the dispersion, the worse the result of the dispersion parameter. In order to achieve an enhancement of all results, taking a dispersion parameter into account, the equation for the settlement dispersion must be conformed to these findings. An improvement of the equation of the dispersion parameter is meaningful in order to enhance the results of built-up density in all areas.

6.2 Transferability-area Namibia

Study area Windhoek

In Windhoek an influence of the different methods of population disaggregation is not visible, because the population counts are distorted by two spikes. Already the linear distribution of the census data shows that the population in Windhoek East and Windhoek West is extremely overestimated. In Windhoek East, four times the real population count is predicted and in Windhoek West twice of the census population count. This development cannot be determined by the disaggregating influence factors of the new algorithm, because it is caused by social reasons.
The comparison of the settlement structure of Windhoek East and Windhoek West with the settlement structure of the constituency Hakahana, a further quarter of Windhoek, where the population was underestimated the most, show a significant reason for the spikes.

Figure 30 shows parts of a satellite scene of Windhoek. The left picture visualizes the typical settlement structure of Windhoek East and the right picture the settlement structure of Hakahana. Both pictures have been extracted from the same satellite scene in a scale of 1:3,000.

Figure 30: Comparison of the settlement structure of Windhoek East (left) and Hakahana (right) in a scale 1:3,000.

In Hakahana much more houses are placed in the identical spatial extent than in Windhoek East. Furthermore, based on the types of buildings and their size, one could assume that Windhoek East is one of the rich quarters in Windhoek, while Hakahana is one of the poorest. This causes the assumption that although the houses are much bigger in Windhoek East than in Hakahana, more people live in a house in Hakahana than in a house in Windhoek East.
In reality, the number of persons per GUF Pixel must be much higher in Windhoek East and West than in the other six constituencies of Windhoek. Since the developed algorithm only considers physical influence factors, while social structures are not incorporated, the phenomena in Windhoek cannot be captured.
Furthermore, in Windhoek East and West also a lot of industry is located, which capture huge parts of the GUF area. These two findings explain both spikes and their large dimension, as well as why they cannot be detected by the developed population algorithm.

Study area Cuvelai-Etosha

In the Cuvelai-Etosha a similar development of the results to the results of the districts in Bavaria is visible. The consideration of the built-up density causes an improvement of the results, particularly in the four rural constituencies.
The dispersion parameter leads to a deterioration of the results except in the rural constituency Oshakati East. The deterioration of the results can be explained as in Chapter 6.1. The influence of the dispersion parameter, as it was calculated, is only useful in highly dense settlement areas, which cannot be found in the Cuvelai-Etosha. The only exception, where the dispersion parameter leads to an improvement is in the rural constituency Oshakati East, which has been considerable overestimated by far compared to all other constituencies in Method I and II (Figure 24).

In conclusion, the results of the transferability feature similar tendencies as the results of the test area. This indicates that the developed algorithm can also be assigned to other study areas. An exception are areas were social factors have a high influence on the number of population, like in slums, or areas which are particularly characterized by large industrial plants.

6.3 Case study area Cuvelai-Etosha

In the Cuvelai-Etosha Basin, the results of the new developed approach, in comparison to the results of the already existing population disaggregation approaches LandScan and WorldPop, feature different accuracies in order to the applicability to a flood event.

The new approach only considers the affected houses in the four constituencies and in the three regarded cities in the Cuvelai-Etosha, while the agricultural land is not included. The population, whose dwelling is not directly affected by the flood, but whose harvest or livestock is destroyed, is not implied in the results for concerned population. This explains the lower count of concerned population (34%) in the case study area, calculated by the new algorithm, than the count of the concerned population of the whole region Oshana, listed in the report. Summarized, the calculated count of affected population features a similar trend as the report information, where the rural areas are more affected than the urban areas.

Taking into account the population data of WorldPop in relation to the flood of 2009, only 2% of the population is affected in the study area, while according to the report, in the whole region Oshana, 87% of the households are impacted by

the flood. However rural areas are more affected than urban areas as well.
The low number of affected population is reasoned by the classification of settlement areas by WorldPop. Most of the rural area in the case study area is regarded as non populated. The majority of the captured settlement structures is located in the three regarded cities. Especially Oshakati East and Oshakati West, the two most affected constituencies of the regarded rural areas, are only very low populated according to the population disaggregation approach of WorldPop. The population distribution in the rural areas is accumulated on a few single spots, which does not reflect the reality. The dwellings in the rural areas in the Cuvelai-Etosha generally consist of a group of small independent constructions or single rooms that interrelate through a system of corridors or fences (Government of the Republic of Namibia, 2009). These traditional housing types are spread dispersed within the regarded area and do not occur in bigger groups concentrated in only a few spots, like WorldPop locates them. In Appendix C the population disaggregation of WorldPop within the study area is visualized.
In conclusion, WorldPop has one of the finest resolutions of all worldwide population disagregation methods, but it is still to coarse to illustrate the real population distribution within the Cuvelai-Etosha. One traditional dwelling structure in the Cuvelai-Etosha is smaller than a pixel of WorldPop. Therefore, the population cannot be exactly spatially linked.

The affected population count calculated with LandScan, displays a higher count of concerned population (13%) than WorldPop but still differs a lot to the value of the report for the whole region. Furthermore, the percentage of concerned population in urban areas is higher than in rural areas, which is the contrary to the results of the report.
LandScan distributes the population more extensive than WorldPop, but the huge pixel size of one kilometer complicates a combination with the disperse flooding structure and also it cannot reflect the actual settlement areas. The arms of the seasonal rivers have a high disperse structure, often smaller than a pixel of LandScan. The settlement structure in the rural areas, like described before, is also highly disperse and the single dwelling structures have a really small size in comparison to a LandScan pixel, like Appendix D visualizes.
During the consideration of a flood event in the Cuvelai-Etosha, a detailed information about single housing spots and their count of population is important to know how many people are affected. The huge pixel size of LandScan cannot supply this.
In conclusion, LandScan pixels are mostly only partly impacted by inundation, which makes it impossible to determine the right number of persons, affected by the flood. Furthermore, the coarse spatial resolution of LandScan could not depict

the disperse settlement structure in the Cuvelai-Etosha Region.

In general the new developed disaggregation approach locates the affected population the best. Due to the high spatial resolution of the GUF, the population can be distributed quite accurately to small areas. In case of a flood event in the Cuvelai-Etosha, this can be a benefit, because the non-perennial river arms do not cover a huge coherent area but are spread within the region. Furthermore, in the investigated rural areas, a disperse settlement structure is common.
A population disaggregation, based on a coarse spatial resolution (WorldPop: 100 m, LandScan: 1 km), complicates the classification of the affected population.
The GUF Pixel, with a pixel size of 12 m, displays every single dwelling and facilitates an accurate assignment of the persons affected by inundation (Appendix E).

Accurate results by rapid flood loss estimation are dependent on their input data. Besides a high resolution of satellite scenes, exact population data with a spatial reference is necessary. The new developed approach is, due to its high spatial resolution, most appropriate to determine the humanitarian consequences of a flood event. It facilitates a fast assessment of population affected by floods and due to the high resolution of the modeling the disperse structure of a flood event can be considered. This allows a precise estimation of influenced population and therefore relieves the post flood management.

7 Conclusion

The main aim of this research work was the development of a population disaggregation method, based on urban footprint classifications, which should be validated and adjusted, based on in-situ data. Furthermore, the applicability of the new method for local flood loss estimation within the SASSCAL project should be tested.
A review about the state-of-the-art for population distribution modeling approaches has been performed (Chapter 1.3) and based on that, a research framework was elaborated (Chapter 2). Subsequently, a method was developed to obtain spatially detailed population data by disaggregating census data onto GUF data. The workflow for a GUF based population distribution modeling was developed on the basis of GUF and census data of Bavaria. Later, it was transferred to study areas in Namibia (Chapter 3). During the workflow development the population disaggregation method was adjusted according to the outcomes and compared to in-situ data (Chapter 4). Besides that, the newly developed algorithm was tested in the context of local flood loss estimation in Namibia (Chapter 5). The results of the population modeling in Bavaria, its transferability to another test area and the applicability for a flood event are discussed in Chapter 6.

Based on the research objectives, three research questions were phrased:

1. What are the challenges to derive population information from the GUF for a specific test site?

The new developed population disaggregation algorithm shall facilitate a combination of GUF data with census data. The GUF data has a high spatial resolution (12 m), while the resolution of census data depends on the country. Normally census data is only available for different levels of administrative units. The spatial extent of the administrative units and consequently the resolution of census data is different in every country. To guarantee a worldwide applicability of the approach, the population has to be distributed on the GUF, independent of its own resolution. To provide the best possible results, the census data with the highest resolution was applied. An equal disaggregation of the population of an administrative unit on the GUF does not reflect the reality of population distribution. Therefore, the main challenge was to identify factors, which can be calculated from the GUF, to influence the population disaggregation. The factors 'built-up density' and 'settlement dispersion' were computed to distribute census data onto the GUF.

2. How can we derive a solid workflow for a global application of GUF-based population information?

To facilitate a worldwide applicability of a GUF-based population information, the input parameters, must be available worldwide. Therefore, the minimum of necessary input parameters was applied to develop a workflow. GUF data is applicable worldwide, as well as census data, although often only in a coarse resolution. Since the developed algorithm is not dependent on country specific factors a facile transferability is possible.

3. What is the benefit of the derived detailed and rapid population information for specific demands and problems in flood risk management?

The disperse structure of a flood event complicates a fast and precise assessment of affected population. The population distribution, reached within this research work, is highly resoluted in comparison to previous disaggregtion approaches, like LandScan or WorldPop. Therefore, it is particularly suitable for disperse structured hazards. Due to the detailed population disaggregation, a precise and fast assessment of affected population is possible.

The small number of input variables facilitates a facile application of the new developed population disaggregation model and submits a transferability to other areas of interest. In the case of a natural hazard the model allows a fast and precise assessment of affected population.
According to the small number of input variables, several factors, influencing population disaggregation cannot be considered. Especially in less developed countries, non-physical factors, like poverty, influence the population distribution.
The quality of the results of the disaggregation is dependent on the quality of the input data. If the census data does not reflect the real population counts, it is difficult to use the calculated population distribution for post disaster management or further applications. In high developed countries census data is mainly of high quality, while especially in the least developed countries it is difficult to evaluate the quality.

8 Outlook

In order to improve the results of the population disaggregation algorithm, a further development of the dispersion parameter would be reasonable. Besides an improvement of the existing calculations, an extension of the algorithm with additional worldwide available influence variables could be advantageous. At the moment, only a 2D-analysis of settlement areas is possible. An enhancement of the GUF with height information would lead to a more precise population disaggregation. In order to improve the GUF as input variable for settlement areas, the removal of industrial areas is necessary. This could be achieved by applying a primary classification based on zoning plans.

In this research work, the transferability of the new disaggregation approach was only tested in one further country. To confirm the transferability of the approach, additional tests are necessary that indicates, if the algorithm must be extended with a country factor. The country factor shall adjust the actual algorithm to the new conditions, if possible. Moreover, the involvement of social factors (e.g. poverty) influencing population distribution, could be considered. However, this would impede a worldwide and fast application of the population disaggregation.

The application of the GUF based population disaggregation algorithm is not only useful for post disaster management of natural hazards. It can also be used as emergency response to man-made disasters, such as nuclear and chemical accidents or terrorist incidents (Dobson et al., 2000).

Furthermore, a detailed knowledge about population disaggregation is helpful for the protection of civilian populations during regional conflicts, for the coordination of humanitarian aid in famines, and for invective disease modeling (Dobson et al., 2000; Tatem et al., 2012).

Besides disaster management, the GUF based population disaggregation can be used to develop adaptive strategies in relation to climate change or for transport and city planning (Gaughan et al., 2013).

References

Baatz, M. and Schäpe, A. (2000). Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In Strobl, J., Blaschke, T., and Griesebner, G., editors, *Angewandte Geographische Informationsverarbeitung XII*, pages 12–23, Heidelberg. Wichmann.

Balk, D., Deichmann, U., Yetman, G., Pozzi, F., Hay, S. I., and Nelson, A. (2006). Determining global population distribution: Methods, applications and data. 62:119–156. DOI: 10.1016/S0065-308X(05)62004-0.

Balk, D., Pozzi, F., Yetman, G., Deichmann, U., and Nelson, A. (2005). The distribution of people and the dimension of place: Methodologies to improce the global estimation of urban areas. In *Urban Remote Sensing Conference*.

Bayerische Staatsregierung (2014). Bezirke, Landkreise, Städte und Gemeinden. Available online at http://www.bayern.de/Bezirke-Landkreise-und-Gemeinden-.4275/index.htm, checked on 09/28/2014.

Briggs, D. J., Gulliver, J., Fecht, D., and Vienneau, D. M. (2007). Dasymetric modelling of small-area population distribution using land cover and light emissions data. *Remote Sensing of Environment*, 108(4):451–466. DOI: 10.1016/j.rse.2006.11.020.

Central Bureau of Statistics (2010). An atlas of Namibia's popualtion: Monitoring and understanding its characteristics.

Central Intelligence Agency (2014). The World Factbook: Namibia. Available online at https://www.cia.gov/library/publications/the-world-factbook/geos/wa.html, checked on 09/17/2014.

Chen, K. (2002). An approach to linking remotely sensed data and areal census data. *International Journal of Remote Sensing*, 23(1):37–48. DOI: 10.1080/01431160010014297.

Cheriyadat, A., Bright, E., Potere, D., and Bhaduri, B. (2007). Mapping of settlements in high-resolution satellite imagery using high performance computing. *GeoJournal*, 69(1-2):119–129. DOI: 10.1007/s10708-007-9101-0.

Commission of the European Communities (1995). CORINE land cover. Available online at http://www.eea.europa.eu/publications/COR0-landcover, checked on 10/21/2014.

Deichmann, U., Balk. D., and Yetman, G. (2001). Transforming population data for interdisciplinary usages: From census to grid.

Deutsches Zentrum für Luft- und Raumfahrt e.V. (2011). TerraSAR-X: Das deutsche Radar-Auge im All - The German radar eye in space.

DigitalGlobe (2010). The benefits of the eight spectral bands of Worldview-2. Available online at http://www.digitalglobe.com/sites/default/files/DG-8SPECTRAL-WP.pdf, checked on 10/20/2014.

Dilley, M., Chen, R. S., Deichmann, U., Lerner-Lam, A. L., and Arnold, M. (2005). Natural disaster hotspots: A global risk analysis.

Dobson, J., Bright, E., Coleman, P., Dirfee, R., and Worley, B. (2000). LandScan: A global popualion database for estimating popualtions at risk. *Photogrammetric Engineering & Remote Sensing*, (Vol. 66, No. 7):849–857.

Esch, T., Taubenböck, H., Roth, A., Heldes, W., Felbier, A., Thiel, M., Schmidt, M., Müller, A., and Dech, S. (2012). TanDEM-X mission - New perspectives for the inventroy and monitoring of global settlement patterns. *Journal of Applied Remote Sensing*, 6(061702):1–21. DOI: 10.1117/1.JRS.6.061702.

Esch, T., Thiel, M., Schenk, A., Roth, A., Müller, A., and Dech, S. (2010). Delineation of urban footprints from TerraSAR-X data by analyzing speckle characteristics and intensity information. *IEEE Transactions on Geoscience and Remote Sensing*, 48(2):905–916. DOI: 10.1109/TGRS.2009.2037144.

Gallego, F. J. (2010). A population density grid of the European Union. *Population and Environment*, 31(6):460–473. DOI: 10.1007/s11111-010-0108-y.

Gaughan, A. E., Stevens, F. R., Linard, C., Jia, P., Tatem, A. J., and Pappalardo, F. (2013). High resolution population distribution maps for Southeast Asia in 2010 and 2015. *PLoS ONE*, 8(2):e55882. DOI: 10.1371/journal.pone.0055882.

Gilau, A. M., Dayo, F. B., Abraham, L. Z., and Mundia, L. (2011). African adaption project Namibia: Drought and flooding risk assessment tool for gender specific decision-making.

Government of the Republic of Namibia (2009). Post-Disaster Needs Assessment: Floods 2009.

Hay, S. I., Noor, A. M., Nelson, A., and Tatem, A. J. (2005). The accuracy of human population maps for public health application. *Tropical Medicine and International Health*, 10(10):1073–1086. DOI: 10.1111/j.1365-3156.2005.01487.x.

Holt, J., Lo, C. P., and Hodler, T. (2004). Dasymetric estimation of population density and areal interpolation of census data. *Cartography and Geographic Information Science*, 31(2):103–121. DOI: 10.1559/1523040041649407.

Khomarudin, M. (2010). *Tsunami risk and vulnerability: Remote sensing and GIS approaches for surface roughness: Determination, settlement mapping and population distribution modeling*. PhD thesis, Ludwigs Maximilians Universität, München.

Kößler, R. (04.01.2014). Namibia. Available online at http://www.bpb.de/themen /NXLXT2,0,Namibia.html, checked on 10/03/2014.

Langford, M., Maguire, D. J., and Unwin, D. J. (1991). The areal interpolation problem: Estimating population using remote sensing in a GIS framework. In Masser, I. and Blakemore, M., editors, *Handling geographic information: Methodology and potential applications*, pages 55–77. Longman, London.

Langford, M. and Unwin, D. J. (1994). Generating and mapping population density surfaces within a geographical information system. *The Cartographic Journal*, 31(1):21–26. DOI: 10.1179/000870494787073718.

Masser, I. and Blakemore, M., editors (1991). *Handling geographic information: Methodology and potential applications*. Longman, London.

McGarigal, K. (2014). Fragstats v4 Help. Available online at http://www.umass.edu/landeco/research/fragstats/fragstats.html, checked on 10/09/2014.

Mennis, J. (2003). Generating surface models of popualtion using dasymetric mapping. *The Professional Geographer*, (55, 1):31–42. DOI: 10.1111/0033-0124.10042.

Ministery of Environment and Tourism (2002). Atlas of Namibia. Available online at http://209.88.21.36/Atlas/Atlas_web.htm, checked on 10/20/2014.

Mott, C. (2005). *Objektortientierte Klassifikationsstrategien zur Erfassung der Landnutzung aus hochauflösenden Fernerkundungsdaten*. PhD thesis, Technische Universität München, München.

Mrozinski, R. D. and Cromley, R. G. (1999). Singly- and doubly-constrained methods of areal interpolation for vector-based GIS. *Transactions in GIS*, 3(3):285–301. DOI: 10.1111/1467-9671.00022.

Namibia Statistics Agency (2012). Namibia 2011: Population and housing census main report.

Nelson, A. and Deichmann, U. (2004). African population database documentation. Available online at http://na.unep.net/siouxfalls/globalpop/africa/africa_index.html, checked on 10/20/2014.

Peduzzi, P., Dao, H., Herold, C., and Mouton, F. (2009). Assessing global exposure and vulnerability towards natural hazards: The disaster risk index. *Natural Hazards and Earth System Science*, 9(4):1149–1159. DOI: 10.5194/nhess-9-1149-2009.

Reibel, M. and Bufalino, E. (2005). Street-weighted interpolation techniques for demographic count estimation in incompatible zone systems. *Environment and Planning A*, 37(1):127–139. DOI: 10.1068/a36202.

Research and Information Services of Namibia (2013). A profile and atlas of the Cuvelai-Etosha Basin.

Rumor, R., Coors, V., Fendel, E. M., and Zlantanova, S., editors (2007). *Urban and regional data managment*. Taylor & Friends Group, London.

SASSCAL (2011). Business plan for the Southern Africa Science Service Centre for Climate Change and Adaptive Land-use (SASSCAL): Version 5.

Schneider, A., Friedl, M. A., and Potere, D. (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, 4(4):044003. DOI: 10.1088/1748-9326/4/4/044003.

Schneider, A., Friedl, M. A., and Potere, D. (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sensing of Environment*, 114(8):1733–1746. DOI: 10.1016/j.rse.2010.03.003.

Schrenk, M., Popovich, V., and Zeile, P., editors (2010). *REALCORP 2010*.

Schrenk, M., Popovich, V. V., and Elisei, P., editors (2014). *Proceedings REAL CORP 2014*, Vienna.

Setiadi, N., Taubenböck, H., Raupp, S., and Birkmann, J. (2010). Integrating socio-economic data in spatial analysis: An exposure analysis method for planning urban risk mitigation. In Schrenk, M., Popovich, V., and Zeile, P., editors, *REALCORP 2010*, pages 367–374.

Steinnocher, K., Köstl, M., and Weichselbaum, J. (2011). Kleinräumige Bevölkerungsmodellierung für Europa - räumliche Disaggregation auf Basis des Versiegelungsgrades. In Strobl, J., Blaschke, T., and Griesebner, G., editors, *Angewandte Geoinformatik 2011*, pages 513–520, Berlin. Wichmann.

Steinnocher, K., Petrini-Monteferri, F., and Tötzer, T. (2005). Räumliche Disaggregation von sozi-ökonomischen Daten. In *Angewandte Geographische Informationsverarbeitung XVII*, Berlin. Wichmann.

Stoinski, S. (2007). *Möglichkeit der objektorientieren Klassifizierung von Waldgebieten aus CIR-Luftbildern*. PhD thesis, Technische Universität Berlin, Berlin.

Strobl, J., Blaschke, T., and Griesebner, G., editors (2000). *Angewandte Geographische Informationsverarbeitung XII: Beiträge zum AGIT-Symposium Salzburg 2000*, Heidelberg. Wichmann.

Strobl, J., Blaschke, T., and Griesebner, G., editors (2011). *Angewandte Geoinformatik 2011: Beiträge zum 23. AGIT Symposium*, Berlin. Wichmann.

Tatem, A. J., Adamo, S., Bharti, N., Burgert, C. R., Castro, M., Dorelien, A., Fink, G., Linard, C., John, M., Montana, L., Montgomery, M. R., Nelson, A., Noor, A. M., Pindolia, D., Yetman, G., and Balk, D. (2012). Mapping populations at risk: Improving spatial demographic data for infectious disease modeling and metric derivation. *Population Health Metrics*, 10(1):8. DOI: 10.1186/1478-7954-10-8.

Tatem, A. J., Noor, A. M., Hagen, C. v., Di Gregorio, A., Hay, S. I., and Gething, P. (2007). High resolution population maps for low income nations: Combining land cover and census in East Africa. *PLoS ONE*, 2(12):e1298. DOI: 10.1371/journal.pone.0001298.

Taubenböck, H., Esch, T., Felbier, A., Roth, A., and Dech, S. (2011). Pattern-based accuracy assessment of an urban footprint classification using TerraSAR-X data. *IEEE Geoscience and Remote Sensing Letters*, 8(2):278–282. DOI: 10.1109/LGRS.2010.2069083.

Taubenböck, H., Roth, A., and Dech, S. (2007). Linking structural urban characteristics derived from high resolution satellite data to population distribution. In Rumor, R., Coors, V., Fendel, E. M., and Zlantanova, S., editors, *Urban and regional data managment*, pages 35–46. Taylor & Friends Group, London.

Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., and Dech, S. (2014). New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. *Applied Geography*, 47:137–153. DOI: 10.1016/j.apgeog.2013.12.002.

The World Bank Group (2014). Gini index. Available online at http://data.worldbank.org/indicator/SI.POV.GINI, checked on 10/16/2014.

Tobler, W., Deichmann, U., Gottsegen, J., and Maloy, K. (1997). World popualiton in a grid of spherical quadrilaterals. *International Journal of Popualtion Geography*, (3):203–225.

Trimble Germany GmbH (2011). eCogniton Developer 8.7: Reference Book.

United Nations, Department of Economic and Social Affairs , Population Division (2013). World population prospects : The 2012 revision: Volume I: Comprehensive tables.

USGS (June 23, 2008). Global land cover characterization background.

USGS (June 24, 2014). Landsat 8. Available online at http://landsat.usgs.gov/landsat8.php, checked on 09/24/2014.

Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., and Mander, Ü. (2009). Landscape metrics and indices: An overview of their use in landscape research. *Living Reviews in Landscape Research*, 3. DOI: 10.12942/lrlr-2009-1.

Wiesner, M. and Taubenböck, H. (2014). It's not big, it's large: Mapping and characterizing urban landscapes of a different magnitude based on EO data. In Schrenk, M., Popovich, V. V., and Elisei, P., editors, *Proceedings REAL CORP 2014*, Vienna.

Wu, S., Qiu, X., and Le W. (2005). Population estimation methods in GIS and remote sensing: A review. *GIScience & Remote Sensing*, 42(1):80–96. DOI: 10.2747/1548-1603.42.1.80.

Wurm, M., Taubenböck, H., Krings, S., Birkmann, J., Roth, A., Dech, S., Michel, U., and Civco, D. L. (2009). Derivation of popualtion distribution for vulnerability assessment in flood-prone German cities using multisensoral remote sensung data. *Remote Sensing for Environmental Monitoring, GIS Applications, ans Geology IX*, 7478(74781K):1–12. DOI: 10.1117/12.830318.

Xie, Y. (1995). The overlaid network algorithms for areal interpolation problem. *Computers, Environment and Urban Systems*, 19(4):287–306. DOI: 10.1016/0198-9715(95)00028-3.

Yuan, Y., Smith, R. M., and Limp, W. F. (1997). Remodeling census population with spatial information from LandSat TM imagery. *Computers, Environment and Urban Systems*, 21(3-4):245–258. DOI: 10.1016/S0198-9715(97)01003-X.

Zaragozí, B., Belda, A., Linares, J., Martínez-Pérez, J. E., Navarro, J. T., and Esparza, J. (2012). A free and open source programming library for landscape metrics calculations. *Environmental Modelling & Software*, 31:131–140. DOI: 10.1016/j.envsoft.2011.10.009.

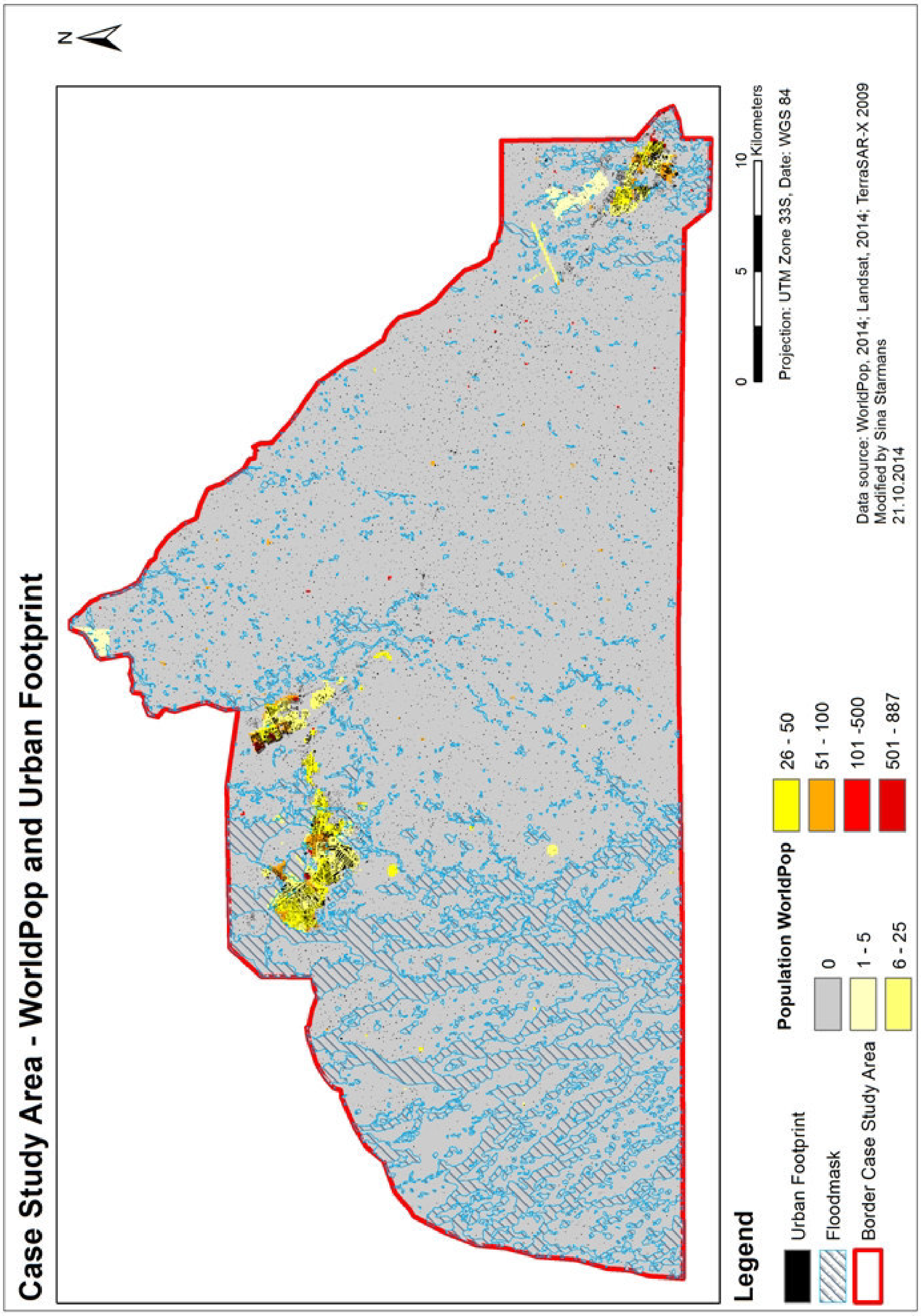
Appendix A

Administrative District	Rural District	Main City
Oberbayern	Miesbach	Holzkirchen
	Munich	Munich
Niederbayern	Dingolfing-Landau	Dingolfing
	Landshut	Landshut
Schwaben	Lindau	Lindau
	Neu-Ulm	Neu-Ulm
Oberpfalz	Neumarkt	Neumarkt
	Regensburg	Regensburg
Mittelfranken	Fürth	Fürth
	Cities of Mittelfranken	Nuremberg
Oberfranken	Coburg	Coburg
	Lichtenfels	Lichtenfels
Unterfranken	Haßberge	Haßberge
	Cities of Unterfranken	Würzburg

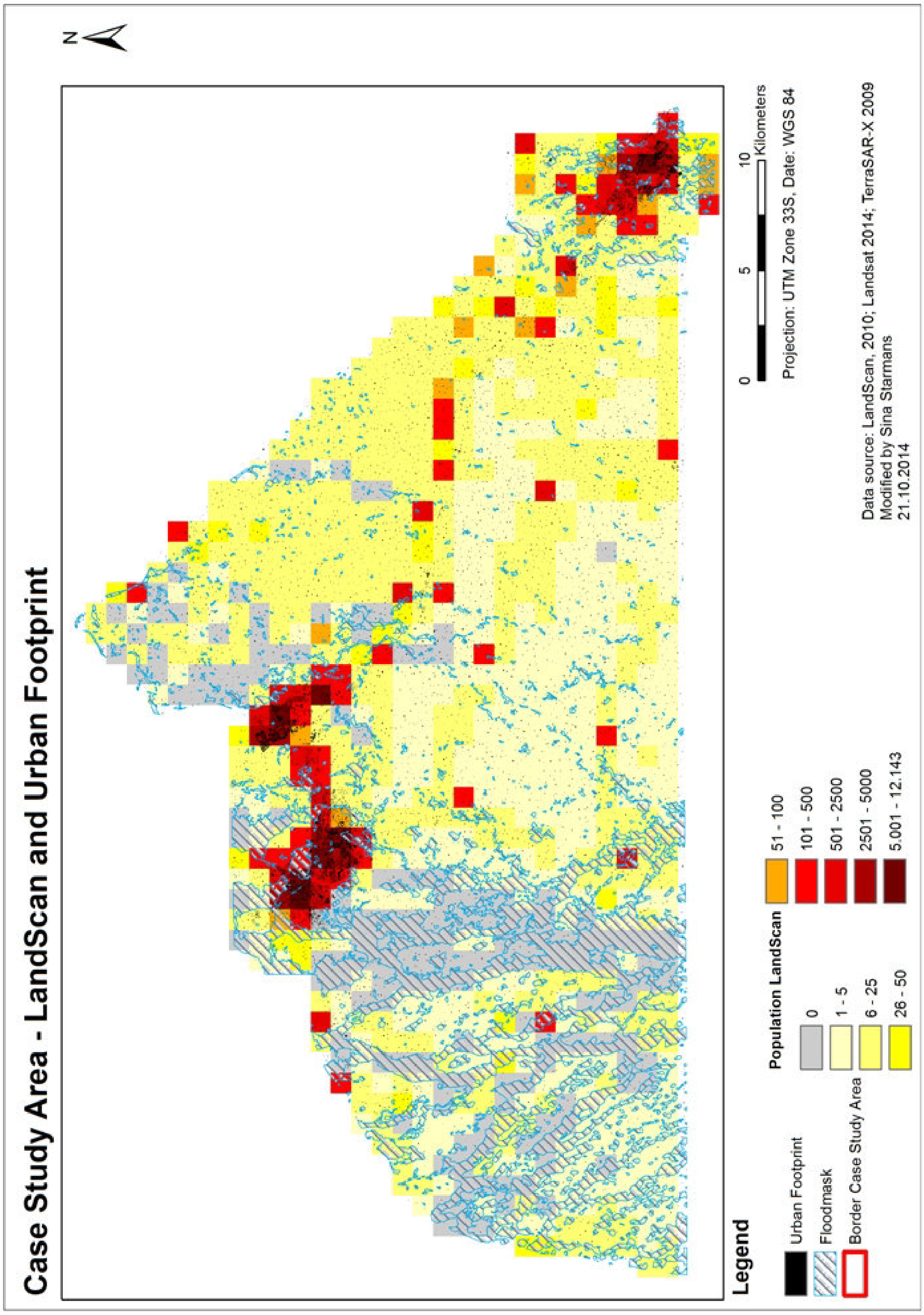
Appendix B

Region	Constituency
Cuvelai-Etosha	Ondangwa
	Ongwediva
	Oshakati East
	Oshakati West
	Ondangwa City
	Ongwediva City
	Oshakati City
Windhoek	Hakahana
	Katatura Central
	Katatura East
	Khomasdal North
	Soweto
	Wanaheda
	Windhoek East
	Windhoek West

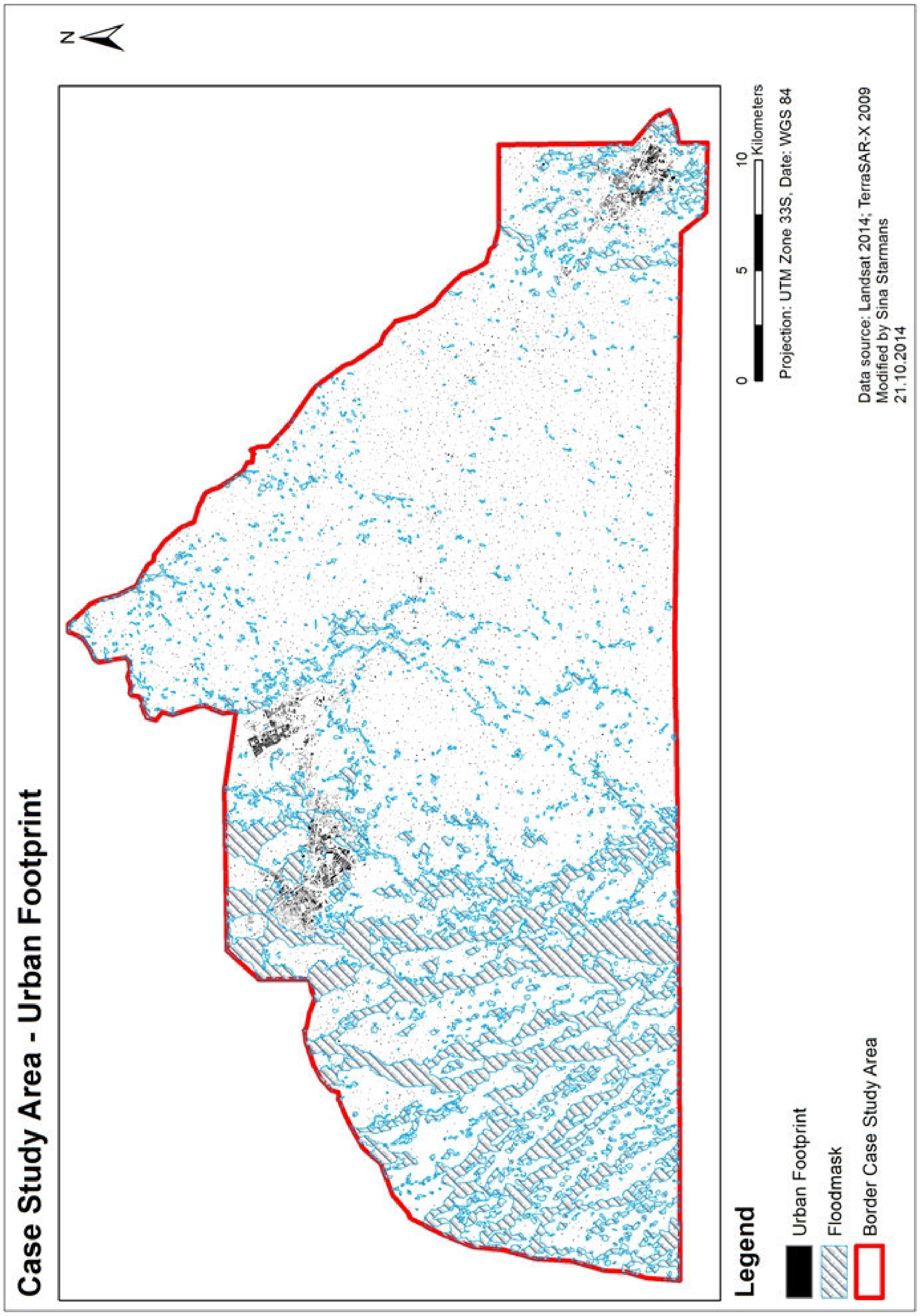
Appendix C


Case Study Area - WorldPop and Urban Footprint
N
Legend
Urban Footprint
Floodmask
Border Case Study Area
Population WorldPop
0
1 - 5
6 - 25
26 - 50
51 - 100
101 -500
501 - 887
0
5
10
Kilometers
Projection: UTM Zone 33S, Date: WGS 84
Data source: WorldPop, 2014; Landsat, 2014; TerraSAR-X 2009
Modified by Sina Starmans
21.10.2014

Appendix D


Case Study Area - LandScan and Urban Footprint
N
Legend
Urban Footprint
Floodmask
Border Case Study Area
Population LandScan
0
1 - 5
6 - 25
26 - 50
51 - 100
101 - 500
501 - 2500
2501 - 5000
5.001 - 12.143
0
5
10
Kilometers
Projection: UTM Zone 33S, Date: WGS 84
Data source: LandScan, 2010; Landsat 2014; TerraSAR-X 2009
Modified by Sina Starmans
21.10.2014

Appendix E


Case Study Area - Urban Footprint
N
0
5
10
Kilometers
Projection: UTM Zone 33S, Date: WGS 84
Data source: Landsat 2014; TerraSAR-X 2009
Modified by Sina Starmans
21.10.2014
Legend
Urban Footprint
Floodmask
Border Case Study Area

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst und alle in ihr verwendeten Unterlagen, Hilfsmittel und die zugrundegelegte Literatur genannt habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.
Die Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Masterarbeit eingereicht und auch nicht veröffentlicht.
Ich nehme zur Kenntnis, dass auch bei auszugsweiser Veröffentlichung meiner Masterarbeit das Institut, an dem die Masterarbeit ausgearbeitet wurde, und der Betreuer zu nennen sind.

Innsbruck, 30. Oktober 2014

Sina Starmans