Socio-hydrology in the floodplain

Investigating patterns of human response to riverine floods by quantifying the human footprint in flood prone areas using nighttime light data

L.M. Verschuren

Technische Universiteit Delft

TU Delft Delft University of Technology

Challenge the future

Socio-hydrology in the floodplain

Investigating patterns of human response to riverine floods by quantifying the human footprint in flood prone areas using nighttime light data

by

L.M. Verschuren

in partial fulfillment of the requirements for the degree of

Master of Science in Civil Engineering

at Delft University of Technology,
to be defended publicly on 11th of october at 14:30 AM.

Supervisor:	dr. S. Pande	TU Delft
Thesis committee:	dr. J.D. Bricker	TU Delft
	dr. O.A.C. Hoes	TU Delft

TU Delft Delft University of Technology

Preface

This report is the result of my master thesis research in fulfilment of my Master degree Water Management at Delft University of technology. This research has taught me a lot of valuable skills. This result could not have been achieved without the support and guidance of many people.

First, I would like to thank my graduation committee for their guidance and feedback during the process of my thesis. Without their support I would not have been able to complete this research. Special thanks go to Saket Pande, who was my daily supervisor, for his tireless support and close supervision during this entire project. I also want to thank Olivier Hoes and Jeremy Bricker for giving valuable comments on my research, which greatly improved the final product. Next, I want to thank Fernando Nardi and Antonio Annis for all their help during my time in Perugia. I want to thank Serena Ceola for guiding me in the use of nighttime light data.

I would also like to thank Hokje1, for providing support and relaxation in times of stress. Without my colleages of Hokje1 I would have never completed this research. Last but not least, I want to thank my parents, for always believing in me and helping me get through this.

L.M. Verschuren
The Hague, 06-10-2019

Abstract

Flooding is globally one of the most damaging natural hazard. Flood risk will most likely increase in the near future due to increases in flood frequency attributed to climate change and growth in population and wealth in flood prone areas. This growth in wealth and population is increasingly considered as a major driver for the increase in flood losses in the last decades. Floodplains are susceptible to floods, but historically people have always been settling in floodplains. The growth of population in floodplains, which is a substantial cause for increased flood risk, is essential to consider for decision making in floodplain development, as improper development increases flood exposure and aggravates flood risk. The science of socio-hydrology tries to capture the interaction between humans and floods in the floodplain. But, it is necessary to identify these mechanisms on a broader scale. A way of doing this, is to look at the development of floodplain population density over the years, but population data is not available on a long temporal scale. Therefore, Nighttime light data was used to model the gaps in the availability of population data. Nighttime light data data captures the illumination on earth, and is available on a large temporal and spatial scale, and has a high correlation with population data. However, the relationship between Nighttime light data and population data is not straightforward. This study tries to model a population proxy through the use of Nighttime light data data and explains when and why it does or does not work. Validation of the model shows that in some regions the predicted data is relatively precise, but ultimately, due to the lack of data, the accuracy is unknown. This study shows that understanding the behaviour of NTL is valuable, because it has the potential to map Socio-Economic variables in data-scarce areas.

Contents

1

Introduction

Flooding is globally one of the most damaging natural hazards (Ritchie and Roser, 2019). In the last decades, the impact of floods has increased substantially, causing more economic damage and victims than ever before. In fact, in terms of global weather-related disaster losses, 2017 was the costliest year, exceeding 300 billion US dollars (Jongman, 2018). Flood risk will most likely increase in the near future due to increases in flood frequency attributed to climate change and growth in population and wealth in flood-prone areas (Jongman, Ward and Aerts 2012). This growth in wealth and population is increasingly considered as a major driver for the increase in flood losses in the last decades (Bouwer, 2011; Jongman et al., 2012; Kundzewicz et al., 2014; Winsemius et al., 2016). Of the ten largest urban agglomerations in the world, nine are located in floodplains (Di Baldassarre et al., 2013). Between 1970 and 2010, the global population in floodplains increased twofold (Fang et al., 2018b).

A floodplain is the land surface next to a river. The features of a floodplain landscape can be clearly distinguished from neighbouring uplands in terms of shape and dominant hydrologic processes. Floodplains are susceptible to floods, storing excess water in case of extreme hydrological events (Nardi, Vivoni and Grimaldi, 2006). Yet, people have historically been settling in floodplains, as floodplains also bring trade junctions, fertile land and water availability (Tockner and Stanford, 2002).

The growth of population in floodplains, which is a substantial cause for increased flood risk, is essential to consider for decision making in floodplain development, as improper development increases flood exposure and aggravates flood risk (Fowler and White, 1945; Mileti, 1999). Reclaiming floodplains for urban development (and other human interferences) creates increased flood risk, flood impact and decreases the biodiversity (Tockner and Stanford, 2002). Even though much progress is made in quantitatively assessing flood risk, a fundamental understanding of the interplay between physical and social processes is still lacking, according to Di Baldassarre, Martinez, Kalantari and Viglione (2017). As a result, unintended consequences can stem from current policies and measurements to mitigate flood impacts.

The science of socio-hydrology, which was first coined by Sivapalan et al. (2012), tries to capture these bi-directional feedbacks between physical and social processes. One of the mechanisms studied in socio-hydrology is the interaction between human behaviour and floods in the floodplain (Di Baldassarre et al., 2013; Di Baldassarre et al., 2015). Within the conceptualization of human-flood interaction, flood coping strategies are studied (Di Baldassarre et al., 2015; Ciullo, Viglione, Castellarin, Crisci and Di Baldassarre, 2017; Viglione et al., 2014; Barendrecht, Viglione and Bloschl, 2017). Kundewicz and Schnellnhuber (2004) identified three basic strategies of coping with floods:

1. Protection (as far as technically possible and financially feasible, bearing in mind that absolute protection cannot be achieved)
2. Accommodation (living with floods)
3. Retreat (relocation from flood-risk to flood-safe areas). This strategy aims to rectify maladaptation and floodplain development.

In socio-hydrology, these three strategies of coping with floods have been conceptualized by Di Baldassarre et al. (2013) and Ciullo et al. (2017) as technological societies versus green societies. In this conceptualization, technological societies cope with floods through protection and green societies cope with floods by either accommodation or retreat. Di Baldassarre et al. (2015) and Ciullo et al. (2017) propose a modelling framework where they describe societies in terms of being a green or technological society. Societies are differentiated through their use of mainly structural and non-structural measures. This modelling framework proved to be effective in capturing bi-directional feedbacks between social and physical processes in floodplains. These feedbacks are able to describe both the adaptation and levee effect.

The levee effect describes the consequences of a decrease in flood memory that stems from levee heightening. Levee heightening lowers the flood hazard and decreases the possibility of flooding. It also fosters economic growth in floodplains due to the reduced frequency of economic losses. On the other hand, the heightening of levees induce an increased feeling of safety, stemming from the reduced flood memory, which generally leads to an increase in exposure caused by people moving into floodplains in flood-poor periods. This is dangerous as the floods that do occur in these areas are often disastrous ones, and the effects of a flood will be more catastrophic. On the other hand, the adaptation effect describes the way people's flood awareness increases after experiencing a damaging flood event. Because of this increase in awareness, flood-related damages will decrease when a flood event with a similar magnitude occurs.

The modelling exercise carried out by Di Baldassarre et al. (2015) and Ciullo et al. (2017) suggest the occurrence of levee effects in technological systems and adaptation effects in the case of green systems. The impact of the hydrological system is conceptualized through relative losses given a certain flood magnitude. The social system is modelled through flood memory which affects floodplain population density, capturing processes that result in resettlement outside of flood-prone areas, and flood-risk protection level. Furthermore, green societies are modelled as having higher social flood memory and lower flood risk protection, leading to a decrease in floodplain population density. In contrast, technological societies are modelled as having an increase in flood risk protection, leading to a decrease in flood memory and an increase in floodplain population density.

The model was applied for two case studies only (Rome and Bangladesh), while in socio-hydrology, there is a need to move beyond case studies. Therefore, it is necessary to identify these mechanisms on a broader (Pande and Sivapalan, 2017). A way of doing this is to look at the development of floodplain population density over the years, as the mechanisms are easily identifiable through floodplain density. They can then be related to the occurrence of flood events. To capture these mechanisms in floodplains on a larger scale, large temporal and spatial datasets are needed.

Unfortunately, population data is not available simultaneously on a long temporal scale and a large spatial scale. To solve this gap in data availability, the use of Nighttime Light data might be a possible solution. This data is available on a large temporal and spatial scale. Furthermore, it has a high correlation with population density, which is why NTL data has been used extensively to model population density in multiple studies (Sutton, Elvidge and Obremski, 2003; Zhuo et al., 2009). However, population density has mainly been modelled for only one year of data and mostly small study areas. Therefore it is interesting to link NTL and population density on a country and floodplain level for multiple years (1992-2013) and multiple countries. In such manner, a population density proxy might be able to fill the gaps for years no population data is available.

1.1. Research Questions

The objective of this research is to understand, interpret and quantify the spatial and temporal evolution of Nighttime Light linking it to human development in floodplains. This objective will be examined through the following research question:

1. Can NTL be used to accurately quantify and monitor changes in population density in floodplains and therefore be used to identify flood coping mechanisms?

1.2. Document Structure

First, chapter 2 contains a literature review on the use of NTL data. Chapter 3 (Methodology) introduces the data used in linking NTL to population density. It explains the data used and the investigated countries in this study. In addition, the method for linking NTL and population data is explained. This includes intercalibration of the NTL data, the regression methods used and including extra parameters like GDP and urban areas. Chapter 4 (Results) provides the results of the procedure described in chapter 3 and explains the validity of the results. Chapter 5 (Discussion) then addresses the the shortcomings of the model and the pitfalls of using NTL. Lastly, chapter 6 (Conclusions and Recommendations) summarized the answers to the research question and the most important findings in this study. Furthermore, it discusses recommendations for future research.

2

Nighttime Light Data

Nighttime Lights (NTL) Data is a gridded data set that globally shows measured luminosity at night for the years 1992-2013. This chapter explains how NTL data was used in other socio-hydrological studies, what the pitfalls of NTL data are and what parameters can be described with the use of NTL data.

2.1. NTL in Socio-hydrology

Socio-hydrology aims to explain the bi-directional feedbacks between human behaviour and hydrological parameters. For this end, NTL is an easy to use indicator of human activities on earth over time, due to large spatial and long temporal coverage and its easiness to acquire. Furthermore, NTL has the ability to show human signatures on river networks (Ceola, Laoi and Montanari, 2014; Ceola, Laoi and Montanari, 2015; Ceola et al., 2016; Elshorbagy et al., 2017; Fang et al., 2018a; Mard, Di Baldassarre and Mazzoleni 2018). Ceola et al., (2014) shows that increased luminosity at night is associated with flood damage intensification. This relationship shows that higher NTL in floodplains leads to higher economic damages to floods. Furthermore, Ceola et al. (2014) shows that there is an increase in exposure (measured as a proxy with NTL) to floods worldwide, as can be seen in figure 2.1.

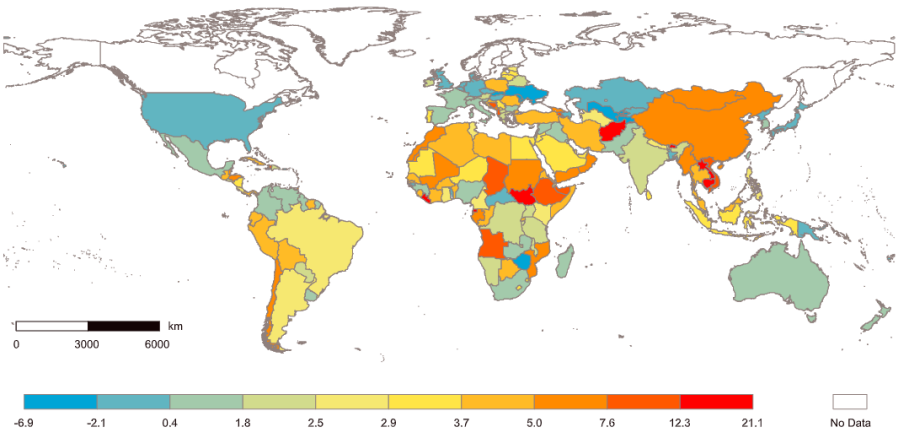

Figure 2.1: Country level percentages of per year variation in nightlight intensity in correspondence to the river network (Ceola et al., 2014)

Ceola et al. (2015) found that in most countries in the world, the amount of NTL has grown in the vicinity of the river network and that people move closer to rivers. Furthermore, Mard et al. (2018) found a correlation between flood protection and human proximity to rivers. In general, countries with a higher flood protection level settle closer to rivers. Furthermore, Mard et al. (2018) shows the average distance of the population to rivers per country and how this develops over the years. Figure 2.2 shows this average distance of the population to the river. The difference in the development

of distance from the river between a green and technological society is visible quite clearly. Here, Mozambique is identified as a green society, and it can be seen that people resettle further away from the river over the last few years.

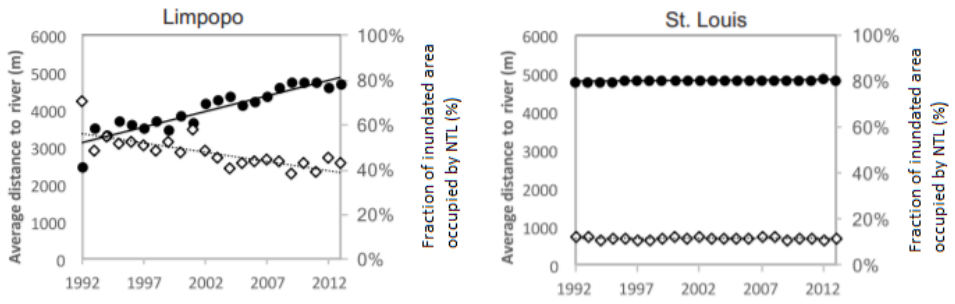

Figure 2.2: Fraction of inundated area occupied by nightlights (empty diamonds, right axis) and Average distance of human settlements to the river (black dots, left axis) between 1992 and 2013 for Limpopo and St. Louis. (Mard et al., 2018)

2.2. Caution in NTL use

Although used frequently in socio-hydrology, NTL data is not an immediate proxy for human behaviour. For example, a 1 percent increase of NTL in Austria does not have the same meaning as a 1 percent NTL increase in Albania. Pinkovskiy (2017) shows that national-level variables such as institutions and policies may have essential effects on growth of NTL. More specifically, both the amount of light and its growth rate are shown to increase discontinuously upon crossing a border from a poorer (or lower growing) to a richer (or higher growing) country and vice versa. For example, figure fig:pinkovskiy shows the differences between Romania (in this case, the more prosperous country) and Ukraine and Moldova (in this case, the less prosperous countries). These differences in growth of light per capita close to the border show the influence of national institutions and that in every country, population and light have a different relationship.

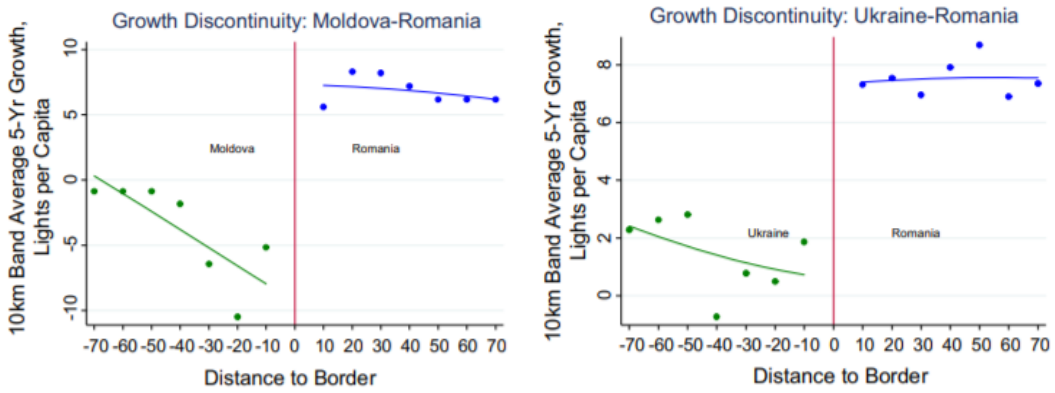

Figure 2.3: Average 5-year NTL per capita growth vs Distance to Border for Moldova and Romania (left) and Ukraine en Romania (right). Blue shows the 5 year average close to the border for the more prosperous country, while green shows the less prosperous country. (Pinkovskiy, 2017)

2.2.1. Pitfalls of Nighttime Light

Next to NTL being incomparable between countries, there are a few pitfalls of using NTL. First, a phenomenon called blooming occurs in the NTL dataset. Blooming is caused by light scattering, which results in bright pixels extending beyond the real bright area. Practically, this means that a large, bright city appears to be much larger than it is (Shen, Zhu, Cao and Chen, 2019). Next to capturing light behaviour itself, there are problems with the six satellites (F10, F12, F14, F15, F16 and F18) that capture luminosity at night. The satellites are not on-board calibrated and therefore are hard to

compare with each other. The inability to compare different years is amplified by the fact that the sensor performance degrades over time, which even makes the comparison of the data captured by the same satellite not straightforward (Zhang and Seto, 2011). Intercalibration tries to solve the lack of on-board calibration, but no optimal calibration method is available at this time. The performance of intercalibration is valued through the phenomenon of convergence between different satellites that were used in the same year. Another indication of a successful intercalibration is the emergence of clear trajectories, such as continuous growth in lighting across the time series. Thus, when convergence happens, the intercalibration is less successful for that country or area for that year (Pandey, Zhang and Seto, 2017). The last problem occurring in NTL data is saturation. When the sensor captures luminosity at night, this data is later transformed into digital numbers (DN) that go from 0 to 63, to represent this luminosity. The low memory and dynamic range causes the sensor to be unable to measure the accurate light intensity higher than the threshold of 63, which leads to many saturated pixels with a DN value of 63 (Imhoff, Lawrence, Stutzer and Elvidge, 1997; Small, Pozzi and Elvidge, 2005). Practically, this means that a very brightly lit area grows in population over the years, but the data cannot capture this change, because it already achieved the highest value of 63 in the first year of measuring.

2.2.2. Behaviour of NTL

The pitfalls of NTL and the differences in what NTL mean per country, make that every country behaves differently. Elvidge et al. (2013) found national-level differences in the behaviour of nighttime lights, when correlating NTL with population and GDP and identifies seven categories of national lighting trends:

1. Rapid and Moderate Growth: describes a positive correlation with NTL and both GDP and population.
2. Population Centric lighting: describes a positive correlation between NTL and population and a negative correlation between NTL and GDP.
3. Economic Centric lighting: describes a positive correlation with NTL and GDP and a negative correlation with NTL and population.
4. Stable and Erratic lighting: describes a low correlation between NTL and both GDP and population.
5. Antipole lighting: describes a negative correlation between NTL and both GDP and population.

These behaviours can be further explained by figure 2.4 below.

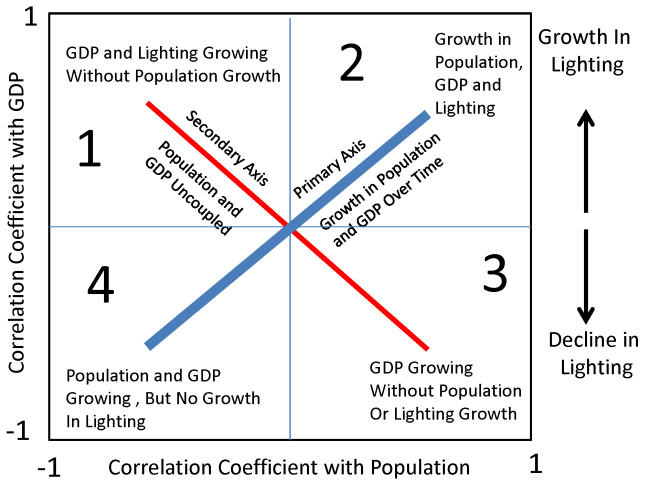

Figure 2.4: In this chart each quadrant gives a different correlation between NTL, population and GDP. Upper right positively correlates with both Population and GDP. Upper left negatively correlates with population, but positively with GDP. Lower left correlates negatively with GPD and population and lower right negatively correlateds with GDP but not with population (Elvidge et al., 2013)

In this study, NTL is linked to population density on a country and floodplain level. In this way, the country-wide institutional effects found by Pinkovskiy (2017) are corrected, which allows for comparison between countries. Furthermore, the pitfalls of NTL will be taken into account when transforming and analyzing the data.

3

Methodology

This chapter describes the study area, the data, and how the data is applied.

3.1. Study Region Administrative Boundaries Data

This study investigates the human-floodplain dynamics in the European continent. The European countries taken into account can be seen below in figure 3.1.

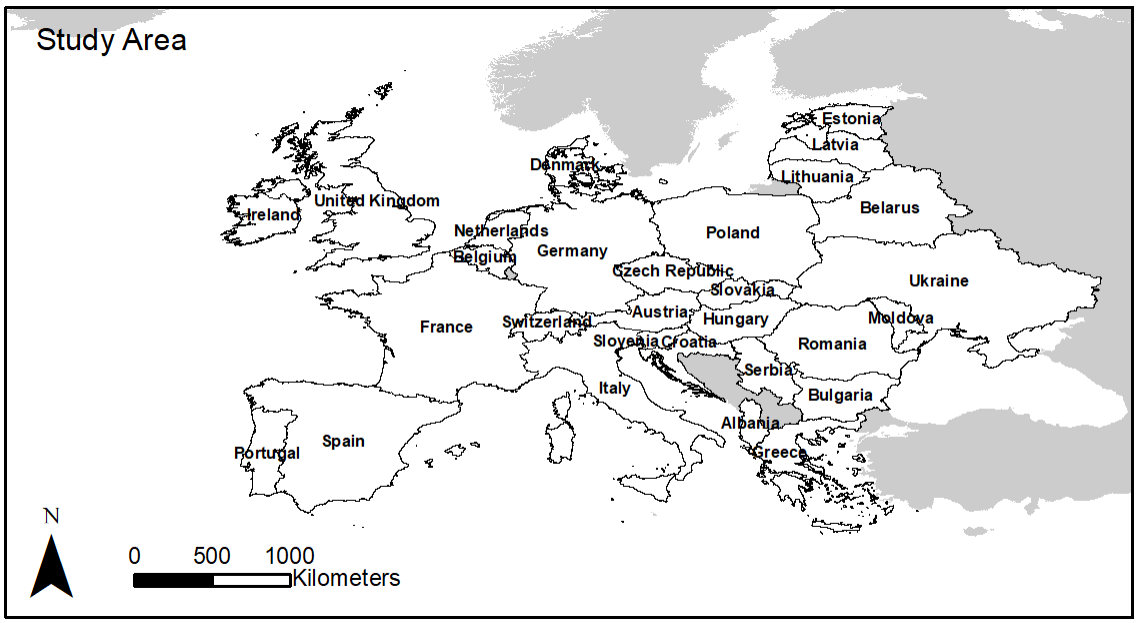

Figure 3.1: Study Area

Not every country on the European continent was included in this research. First, Luxembourg, Cyprus and Malta are excluded, because they were deemed too small to run regressions on. Secondly, Norway, Sweden and Finland were left out of the data, because the data does not cover the high latitudes of these countries. In addition, four Asian and four African countries are added to the study area. They are added to check the development of population against NTL data for countries with different development levels than European countries and different flood risk situations, as flood coping mechanisms tend to be related to development level.

3.2. Floodplain Delineation Data

The Global Floodplain (GFplain) dataset was used to identify the delineation of floodplains in every country (Nardi, Annis, Di Baldassarre, Vivoni and Grimaldi, 2019). The dataset is a vector layer (raster

files are also available) that represents the shape of floodplains globally. It was constructed with the use of the Shuttle Radar Topography Mission (SRTM) Digital Terrain Model and geomorphic algorithms that are able to identify the location of floodplains. River basins are dissected into domains of low lying riparian corridors, separated from their surrounding landscapes. From the stream network, extracted from the DTM, the algorithm assigns each cell with a maximum potential channel flow depth (h), while adopting the power law equation $h \ \alpha \ A^b$, where the contributing area (A) is a scaling parameter. The GFplain algorithm produces a gridded floodplain layer by pointing out the low-lying cells along river networks. The algorithm recognizes the floodplain extent as formed by those low-lying cells. They drain to the selected channel locations, that are characterized by elevations that are lower than the corresponding maximum channel flow level, described as H = z + h. Here, z is the channel cell elevation obtained from the DTM expressed as absolute elevation in meters above sea level. The variation of floodplain flow levels across spatial scales is evaluated employing the dimensionless b exponent to produce a consistent floodplain zoning analysis (Nardi et al., 2019).

3.3. Population data

To model population on a large scale, the Gridded Population of the World, Version 4 (GPW) dataset was used (Center for International Earth Science Information Network (CIESIN), 2016). This dataset is constructed and maintained by the Socioeconomic Data and Applications Center (SEDAC) at the Center for International Earth Science Information Network (CIESIN) at the Earth Institute at Colombia University. The data is available at a resolution of 30 arcseconds, which is approximately 900 m^2 at the equator and was created for the years 2000, 2005, 2010, 2015 and 2020. In this study, only the years 2000, 2005 and 2010 were used. The gridded population layers consist of estimates of population count, which are based on counts consistent with the national censuses and population registers for the year 2010 for small political units (municipalities, census tracts). In every political unit, of which around 12.5 million are used in this global data set, population is distributed evenly. A proportional allocation gridding algorithm is used for assigning a value for every pixel of the dataset. Lastly, by extrapolating the raw census estimates, the population data is derived for the years 2000, 2005, 2010, 2015, and 2020.

3.4. Nighttime Light Data

The NTL Data Set is a gridded data set that globally shows measured luminosity at night for the years 1992-2013. The data is collected by the Defense Meteorological Program (DMSP) Operational Line-Scan System (OLS), which consists of six satellites (F10, F12, F14, F15, F16 and F18) and is maintained and processed by NOAA. Over time the satellites/sensors age and eventually are no longer able to produce data. The degradation is typically gradual enough that a replacement satellite can be deployed to ensure continuity. Thus in most years, two satellites collected data, and two separate composites were produced, as can be seen in Table 3.1.

OLS images are collected from a -65 to 65 degrees latitude at a resolution of 30 arcseconds (NOAA Earth Observation Group, 2018). The DMSP-OLS collects global cloud imagery using visible and thermal bands. At night the visible band signal is intensified with a photomultiplier tube (PMT) to enable the detection of moonlit clouds. The boost in gain enables the detection of lights present at the earth's surface. In the typical annual cloud-free composite most areas have twenty to a hundred cloud-free observations, providing a temporal sampling of the locations and brightness variation of lights present at the earth's surface. Most of these lights are from human settlements (cities and towns) and ephemeral fires. These fires are removed based on their high DN values and lack of persistence (=continued presence across the years). Background noise is removed through thresholds based on values found in areas known to be free of detectable lighting (Baugh, Elvidge, Ghosh and Ziskin, 2010). These steps ultimately result in a readily downloadable stable lights product (NOAA NGDC, 2018). Each pixel in the luminosity data is assigned a digital number (DN), representing its luminosity. The DN's are integers ranging from 0 to 63, with the relationship between DN and luminosity being (Chen and Nordhaus, 2011):

$$Luminosity \ \alpha \ DN^{3/2} \tag{3.1}$$

	Satellites					
Year	**F10**	**F12**	**F14**	**F15**	**F16**	**F18**
1992	F101992					
1993	F101993					
1994	F101994	F121994				
1995		F121995				
1996		F121996				
1997		F121997	F141997			
1998		F121998	F141998			
1999		F121999	F141999			
2000			F142000	F152000		
2001			F142001	F152001		
2002			F142002	F152002		
2003			F142003	F152003		
2004				F152004	F162004	
2005				F152005	F162005	
2006				F152006	F162006	
2007				F152007	F162007	
2008					F162008	
2009					F162009	
2010						F182010
2011						F182011
2012						F182012
2013						F182013

Table 3.1: DMSP/OLS Satellites and overlays in corresponding years.

3.4.1. Intercalibration of NTL

Because the OLS has no onboard calibration, the individual composites are not comparable directly between the years, which poses a challenge for trend analyses. Therefore, before analysing the data, an intercalibration is required. In literature, a wide array of intercalibration methods are available (Pandey et al., 2017). In this case, the method developed by Elvidge et al. (2009) and Chen and Nordhaus (2011) is used because it is relatively robust. It is also the most widely used method in literature (Ceola et al., 2014; Mard et al., 2018).

This method found that F121999 had the highest digital value range and was subsequently used as the reference data set (Elvidge et al., 2013). The island of Sicily of the country Italy has the most favourable characteristics for yearly comparison. Sicily has an even spread of data, the total range of digital numbers and relatively minimal changes in luminosity. These minimal changes can be observed in figure 3.2 in which Sicily is compared to other regions in Italy with roughly the same size. Furthermore, in terms of socio-economic activities over recent decades, Sicily had not changed much (Li and Zhou, 2017). This framework has been extensively adopted over the world due to its easy implementation and robust performance (Zhou, Peo, Haynie and Fan,citeyearMa_2012).

Subsequently, all the other satellites were adjusted to match the F121999 data range of Sicily. The matching of the satellite products to F121999 was done with regressions. A second-order fit, as done in equation 3.2 is drawn on the density distribution of pixels to adjust the digital numbers. The coefficients are empirically derived by comparing images of Sicily with F121999. The derived coefficients that were used can be found in appendix A.

$$DN_{\text{adjusted}} = C_0 + C_1 \times DN + C_2 \times DN^2 \tag{3.2}$$

3.5. Additional Datasets

Besides the datasets mentioned earlier, GDP, Census data, Land use data, and the Hanze dataset were used in this thesis. GDP and population census data per country were taken from the World Bank (2019a; 2019b). GDP data is normalized for the values of 2010.

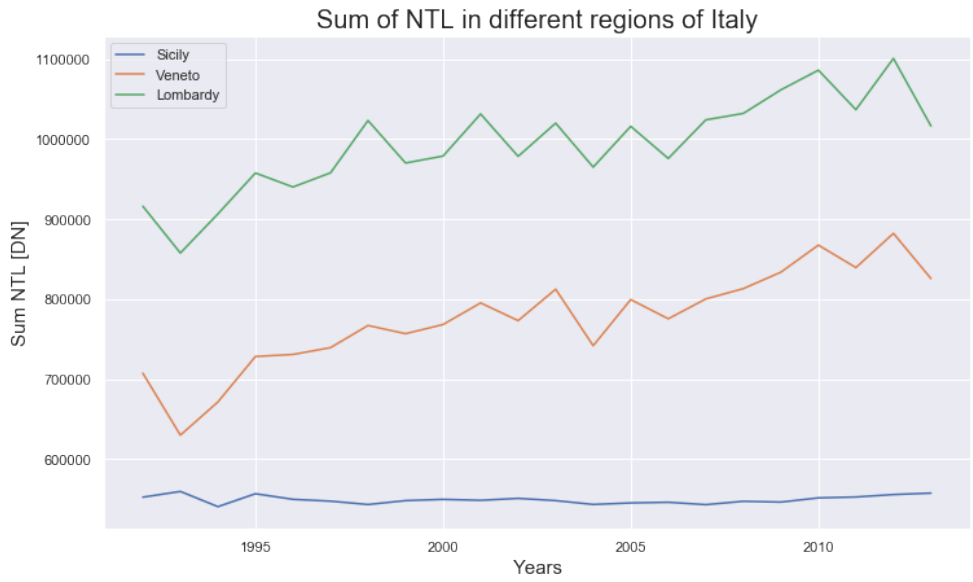

Figure 3.2: Sum of NTL over the years in different regions of Italy. Blue shows Sicily, Orange Veneto and Green Lombardy.

Land use data was retrieved from (ESA, 2018). This is a gridded dataset that shows land use and land cover (LULC) on a global scale. For this research, the years 1992-2013 were used. The dataset consists of multiple LULC categories, but for this study, the dataset was transformed into a binary grid showing 1 for urban areas and 0 for non-urban areas.
The monetary losses caused by floods were retrieved from the HANZE dataset (Paprotny, Morales-Nápoles and Jonkman, 2018). These monetary losses are in euros and all the years are intercomparable, through converting the value of the monetary losses to the 2011 economical standard. The flood losses per country were calculated by summing the monetary flood losses during the period 1992-2013.

3.6. Regression Model

The goal of regressing NTL with population data is to describe a population proxy to fill the gaps for years no population data is available. In doing so, the three years of GPW data are evaluated against the nighttime light data with the use of regression. Thus, two datasets are employed in this model; the population dataset and the NTL dataset as described in the previous paragraphs. The GPW dataset only contains five years (2000, 2005, 2010, 2015 and 2020), of which we use three (2000, 2005, 2010) of global coverage data. The Nighttime Light dataset has 22 consecutive years (from 1992 till 2013). The assumption behind this regression is that the correlation between NTL and population data is very high (R^2 is between 0.7 and 0.9, depending on the country of interest). With the information on the relation between NTL and GPW, interpolation for the missing population years is possible. Three different models are tested on validity through R^2 and RMSE. The first model uses only the relationship between NTL and population, the second model corrects for GDP effects and the third model tries to solve the saturation effect through differentiation between urban and non-urban areas.

3.6.1. Box cox transformations

Using the GPW and NTL datasets, a regression model was built to predict the population in a country based on the lights emitted. The transformed population proxy (y_{new}) is modelled as a function of nighttime light (x).

$$y_{\text{new}} = \beta_0 + \beta_1 x + \epsilon \tag{3.3}$$

To find the values for the parameters β and ϵ, first, the relationship between the actual population data and NTL data must be found for each country and each year, which is done with regression. Some regression assumptions are that x and y are linearly related and that the errors are normally distributed. This is unfortunately not the case for the population and NTL data, and therefore a Box-

Cox transformation is employed of which the aim is to ensure that the usual assumptions for linear models hold. The Box-Cox transformation (Box and Cox, 1964) was first defined for situations in which the dependent variable y is known to be positive and for which the following transformation can be used:

$$y_i^\lambda = \begin{cases} \frac{(y_i^\lambda - 1)}{\lambda}, & \text{if } \lambda \neq 0 \\ log(y_i), & \text{if } \lambda = 0 \end{cases} \tag{3.4}$$

such that for unknown λ,

$$y^\lambda = X\beta + \epsilon \tag{3.5}$$

Where y^λ is the λ transformed response variable, λ a transformation parameter, x is the matrix of explanatory variables, β is a vector of unknown regression coefficients and ϵ is the error term for which $N(0, \sigma^2)$ applies. To simultaneously estimate β, σ and λ, maximum likelihood estimation is used (Carroll and Ruppert, 1985).

3.6.2. Model Steps

In the model, first the parameters β and λ are evaluated through an inverse Box-Cox transformation with formula 3.6. With that information, the NTL data is plotted against the original population data (y) as can be seen in figure 3.3. Subsequently, a relationship between NTL and Population is found. Thereafter, the Box-Cox transformation 3.7, to make the residuals normal is applied, is carried out.

$$y = (((\beta_0 + \beta_1 x)\lambda) + 1)^{\frac{1}{\lambda}} \tag{3.6}$$

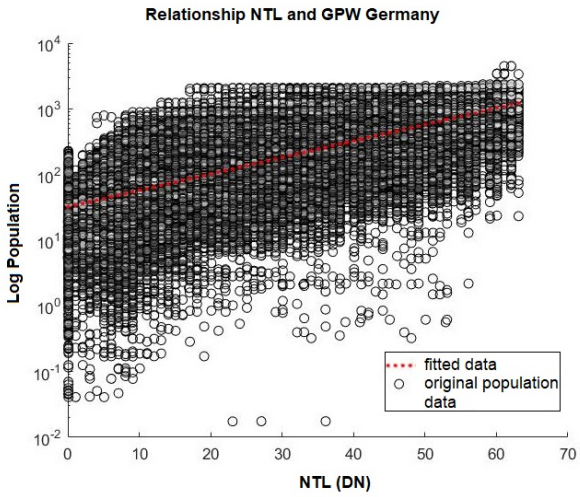

Figure 3.3: Scatterplot which shows the original population data on the y-axis and the NTL data on the x-axis. In red it shows the fit of the transformed data.

$$\begin{aligned} y^\lambda &= \frac{(y^\lambda - 1)}{\lambda} \\ x^\lambda &= \beta_0 + \beta_1 x \\ r &= y^\lambda - x^\lambda \end{aligned} \tag{3.7}$$

Where y^λ and x^λ are the transformed variables and r the residuals. Furthermore, σ and μ are found by fitting a histogram with normal distribution on the residuals. With the λ, β_0, β_1 and σ known, the transformation from nighttime light into a population proxy can be made. This population proxy is made by estimating a transformed value for every value of DN of the nighttime light data for every country with formula 3.11.

$$y_{\text{new}} = \begin{cases} (((\beta_0 + \beta_1 x + \epsilon)\lambda) + 1)^{\frac{1}{\lambda}}, & \text{if } \lambda > 1E-4 \\ \exp(\beta_0 + \beta_1 x + \epsilon), & \text{if } 0 < \lambda < 1E-4 \end{cases} \tag{3.8}$$

The model and the outcomes of using only the model described above is called model 1 from now on.

3.6.3. Model Expansion

Next to the steps explained in the previous paragraph, steps were added to the model to make it more robust.

Population constraints

First, a population count constraint was added. This means that for every country and for every year of data, a line of info on the inserted. The model is calibrated such that the average over the three years is within +/-10 percent of the observed population count.

GDP subtraction

Next to population, NTL also has large correlations with GDP. Therefore, the effect of GDP was added to the model. To model population density of a country without the influence of GDP, a relationship between the NTL and GDP per capita in a country was investigated. First, the relation between NTL (x) and GDP per capita (z) was investigated for every country through regression.

$$log(x) = b1 * log(z) + b2 \tag{3.9}$$

Subsequently, GDP was transformed into an NTL-proxy. The proxy was then subtracted from the raw NTL, creating a new NTL value (NTL-new) with the following formula:

$$x_{\text{new}} = x - exp(b2) * z^{b1} \tag{3.10}$$

Thereafter, new x_{new} was regressed with y (population), with the same steps as in the previous model explained in paragraph 3.5. The model version using the GDP subtraction is from now on called model 2.

Land Use Valuation

When trying to describe Population through NTL, the shortcomings of NTL need to be acknowledged. One of these shortcomings is that the data is constrained on top, which is called saturation. Saturation means in this case that the data does not go beyond the value of DN = 63. This means that cities are often valued as 63, but have way more population than other areas valued 63. Therefore a valuing system was created for urban and non-urban areas and land use data (w) was added to the model. First, a pixel was flagged as either 0 (non-urban) or 1 (urban). Then a new population value was created through the valuation of these pixels.

$$y_{\text{new}} = \begin{cases} (y + 1/y), & \text{if } w = 0 \\ (y), & \text{if } w = 1 \end{cases} \tag{3.11}$$

3.7. Population density through Haversine

With the population proxy, we can identify the floodplain and countrywide population density (PD) for multiple years. This is done by taking the sum of the population proxy (POPP) and dividing it by the area (A).

$$PD = \sum POPP_i / A \tag{3.12}$$

The floodplain area is calculated with the use of the Haversine method. It is applied in this study to accurately calculate the size of a pixel in km^2 (Ivis, 2006). This is necessary because the size of a raster pixel used in this study is approximately 900 m^2 at the equator and differs with latitude. The correction assures an accurate calculation of population density.

4

Results

In this chapter, the results of the NTL-model are shown, the behaviour of NTL is investigated, and growth rates of the population proxy are compared. Lastly, the growth rates are compared to monetary losses caused by flood events.

4.1. Nighttime Light Behaviour

Before converting NTL data into a population proxy for the years no population raster data is available, it is important to understand the behaviour of NTL. In doing so, the model output can be interpreted correctly. First, the average NTL is compared with population density in the floodplain for the year 2010 to see if there is a strong correlation between both datasets. This relationship is shown in figure 4.1.

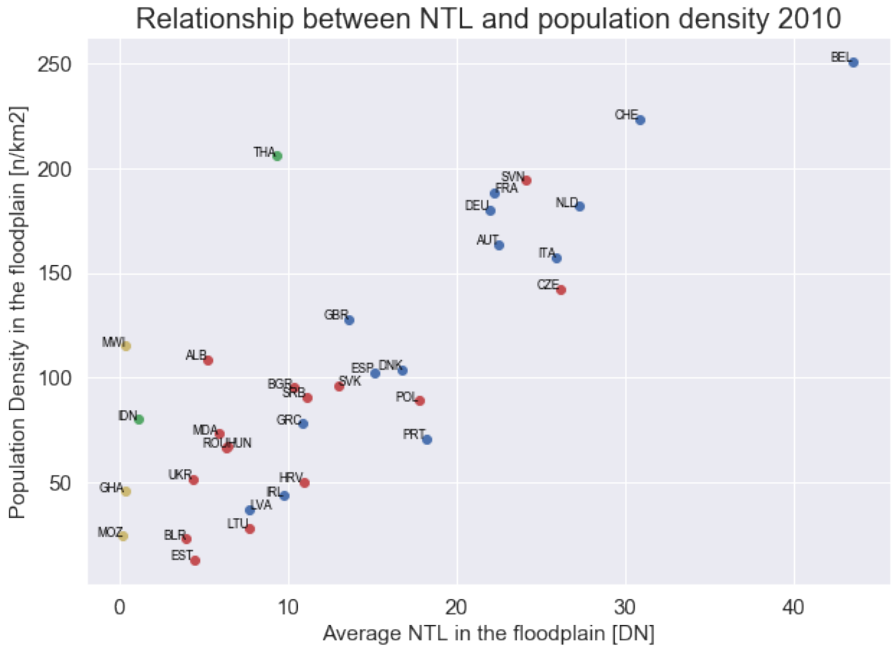

Figure 4.1: Relationship between average NTL and population density in the floodplain for the year 2010. The red dots are Eastern European countries, the blue dots are Western European countries, the green dots are Asian countries and the yellow dots are African countries.

The results in Figure 4.1 show that for most countries, high NTL coincides with high population density. Likewise, low NTL corresponds with low population density. The correlation between NTL and population density is R^2 = 0.900 when only taking European countries into account. When taking the

African and Asian countries into account, the correlation drops to R^2 = 0.227. Only taking Asian and African countries into account, R^2 = 0.589. Bangladesh, Egypt, and Vietnam are not depicted in Figure 4.1, as they have a population density in the floodplain exceeding 500 $people/km^2$ and that would make the figure less readable. The figure with these countries included can be found in Appendix B. The results show that Eastern European countries (depicted in red) tend to have lower population density and lower NTL values in the floodplain than Western countries (depicted in blue). The African and Asian countries tend to have a lower NTL values when relating it to floodplain population density. Egypt, Vietnam, and Bangladesh have a very high population density, but the average NTL in the floodplain in these countries does not reflect that.

To acquire an understanding of patterns of change in both the NTL and population dataset, the relationship between NTL growth and population density growth is illustrated in figure 4.2.

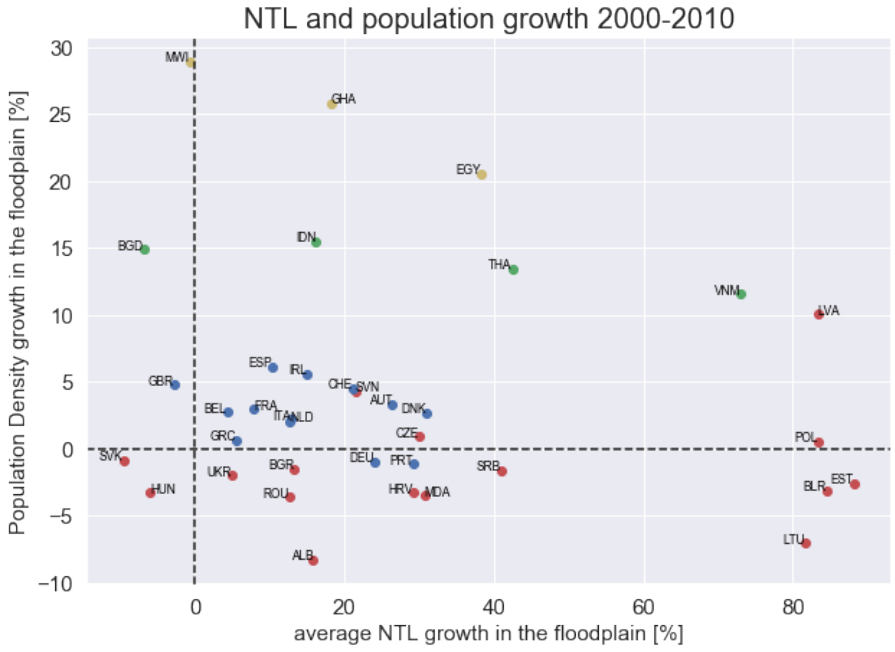

Figure 4.2: Average NTL growth plotted against population density growth in the floodplain for the years 2000 - 2010. In blue, Western European countries are pictured, in red, Eastern European countries, and the Asian countries in green. African countries are pictured in yellow. The horizontal axis shows the average NTL growth in percent change and the vertical axis shows the population density growth in percent change.

The results in figure 4.2 show a less linear pattern than the results in figure 4.1. The R^2 = -0.341 when looking at European countries only. The average NTL seems to grow for all of the countries, except for Hungary, Bangladesh, Great Britain, and Slovakia. The results show an overall pattern of Western European countries experiencing growth in floodplain population density. They have an average NTL growth between -0.02 and 0.31 percent. Eastern/Central European countries mostly experience decline in population density. The change in NTL is more scattered in Eastern Europe. Lastly, the Asian and African countries experience large amounts of population density growth in the floodplain. Except for Malawi and Bangladesh, the African and Asian countries also have a large growth in average NTL. Mozambique is not pictured in this graph, as the average NTL growth in the floodplain exceeds 175 percent. The figure showing the numbers of Mozambique as well, can be found in appendix B.

The correlations of GDP and population for whole countries and in the floodplain are shown in figures 4.3 and 4.4. Figure 4.3 shows the correlation coefficient of NTL with GDP per capita and with population for whole countries. Interestingly, the results cluster in each of the four quadrants. According to the light behaviours discussed by Elvidge et al. (2013) the results can be grouped together and they divide themselves between two axes. The groups are GDP centric in the top left of figure

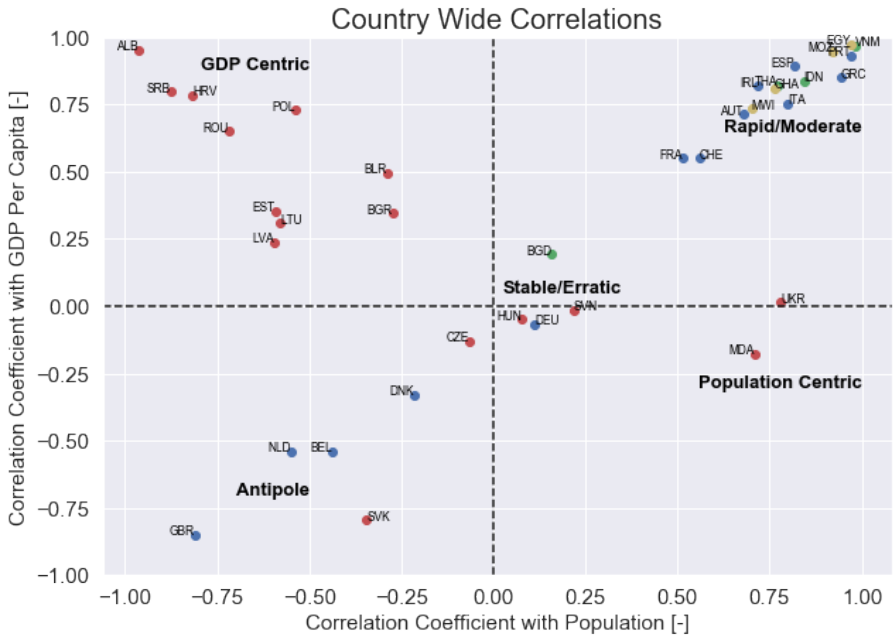

Figure 4.3: Correlations of NTL with population and GDP per country. In blue, Western European countries are depicted. In red, Eastern European countries are depicted and the Asian countries are depicted in green.

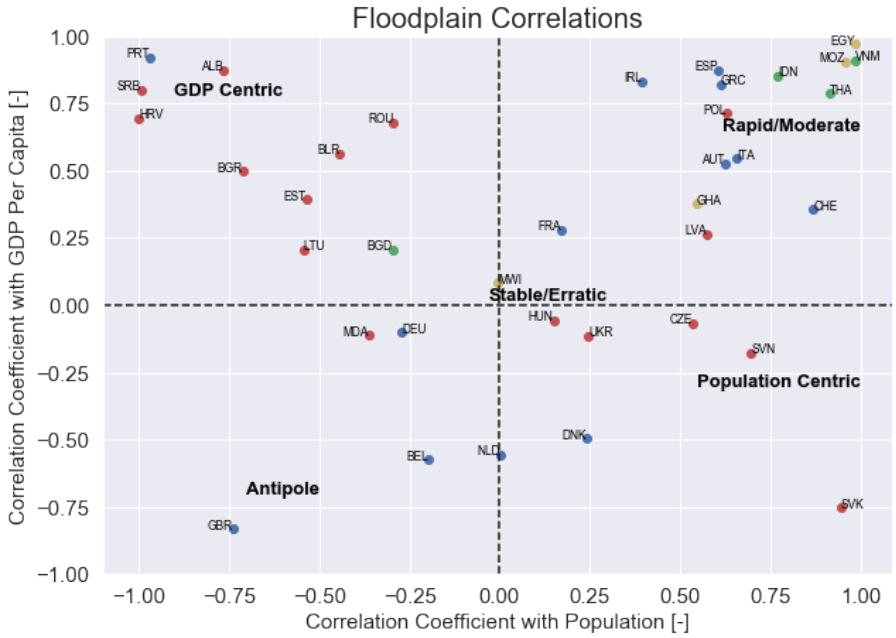

Figure 4.4: Correlations between NTL, population and GDP per country in the floodplain. In blue, Western European countries are pictured. In red, Eastern European countries, the Asian countries in green and the African in yellow.

4.3, rapid/moderate growth in the top right, antipole in the bottom left and stable/erratic or population centric in the bottom right. Figure 4.3 shows a trend as well for Western and European countries, where the Western and Asian countries follow the axis going from (-1, -1) till (1,1) and Eastern European countries follow the opposite axis. The only exception here is Slovakia. The African and Asian countries mostly categorize in the top right corner, showing rapid/moderate growth behaviour. The exception here is Bangladesh.

The results of figure 4.4 are less clustered than the results in figure 4.3. Overall, the countries remain in the same category, with the exception of a few, which will be explained in the discussion.

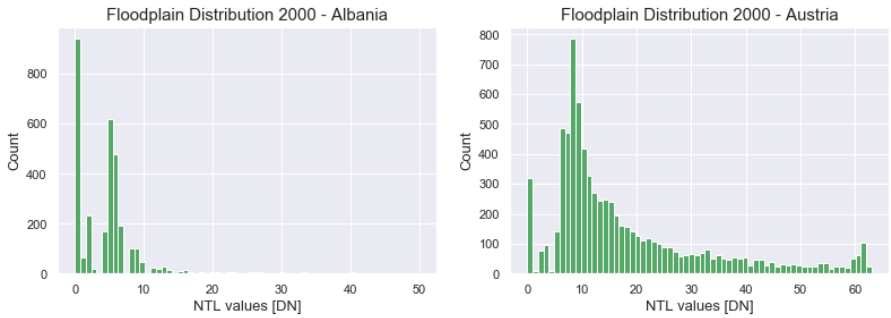

Figure 4.5: Histograms showing the digital number distribution of raw NTL data in Albania and Austria for the year 2000

Focusing on one Eastern (Albania) and one Western European (Austria) country, the results in figure 4.5 show the distribution of raw NTL data across the floodplain. The results also show that the distribution of NTL is different for Eastern and Western European countries. The light is concentrated in low DN values for Eastern countries, and the DN values are more spread for Western European countries.

4.2. Regression Model Performance

A regression model is set up to interpolate NTL data into a population proxy, filling up the gaps for the years population raster data is not available. The regression model performance is validated against the R^2 and root mean square error (RMSE) statistical metrics. The validation results are shown in table 4.1 for three types of model runs. Model 1 refers to the model that interpolates NTL directly to the population proxy. Model 2 refers to the model that first subtracts GDP from NTL and than interpolates the data into the population proxy. Model 3 refers to the model that accounts for differences in population between urban and non-urban areas.

	M1 Country		M1 Floodplain		M2 Country		M2 Floodplain		M3 Country		M3 Floodplain	
Countries	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
Albania	0.73	3.99	0.69	24.34	0.73	4.56	0.69	27.67	0.73	4.35	0.69	23.54
Austria	0.76	2.02	0.80	1.05	0.76	0.95	0.80	1.36	0.76	0.86	0.80	0.87
Belarus	0.71	2.30	0.62	1.11	0.71	0.92	0.63	0.94	0.71	0.72	0.63	0.90
Belgium	0.57	0.95	0.66	2.76	0.58	1.57	0.67	0.53	0.58	1.00	0.66	0.65
Bulgaria	0.52	5.46	0.58	30.81	0.52	3.19	0.58	32.31	0.52	3.26	0.58	31.26
Croatia	0.60	1.75	0.73	2.32	0.60	1.47	0.73	0.86	0.60	1.65	0.73	0.90
Czech Republic	0.78	2.17	0.81	2.34	0.78	0.94	0.81	0.81	0.78	1.04	0.81	0.83
Denmark	0.65	0.95	0.66	2.75	0.65	1.48	0.68	2.88	0.65	0.88	0.68	1.87
Estonia	0.68	1.34	0.76	1.87	0.68	1.34	0.76	2.09	0.68	1.50	0.76	2.45
France	0.61	0.89	0.61	1.97	0.62	2.96	0.61	2.84	0.62	2.01	0.61	1.88
Germany	0.78	0.77	0.80	1.23	0.78	0.78	0.80	0.81	0.78	0.83	0.80	0.85
Greece	0.64	1.13	0.70	2.77	0.66	2.83	0.70	0.98	0.66	1.98	0.70	1.10
Hungary	0.67	2.59	0.70	0.80	0.67	0.87	0.74	3.56	0.67	0.75	0.74	2.89
Ireland	0.70	0.78	0.75	0.92	0.70	1.29	0.75	1.80	0.70	0.80	0.75	1.07
Italy	0.50	2.18	0.48	2.06	0.51	2.32	0.48	2.15	0.51	2.21	0.48	2.06
Latvia	0.79	1.19	0.80	1.89	0.79	0.84	0.80	1.10	0.79	0.90	0.80	1.39
Lithuania	0.60	2.37	0.72	2.36	0.61	1.83	0.72	0.92	0.61	1.66	0.72	0.89
Moldova	0.55	9.42	0.56	9.92	0.62	22.67	0.56	12.62	0.62	20.74	0.56	12.48
The Netherlands	0.73	2.61	0.67	3.41	0.72	0.99	0.67	0.84	0.72	0.76	0.67	0.69
Poland	0.70	2.71	0.68	1.17	0.71	0.70	0.68	0.82	0.71	0.69	0.68	0.79
Portugal	0.44	223.68	0.48	2.46	0.57	2.34	0.48	2.91	0.57	2.33	0.48	3.19
Romania	0.70	1.19	0.68	2.89	0.70	0.90	0.68	1.35	0.70	0.96	0.68	1.48
Serbia	0.66	10.58	0.68	2.33	0.70	1.18	0.68	0.84	0.70	1.33	0.68	0.89
Slovakia	0.66	3.12	0.70	0.83	0.66	0.86	0.70	0.84	0.66	0.95	0.70	0.69
Slovenia	0.76	14.93	0.85	3.68	0.79	2.70	0.85	2.74	0.79	2.65	0.85	2.63
Spain	0.56	2.27	0.60	1.37	0.59	2.69	0.61	2.46	0.59	1.93	0.61	1.84
Switzerland	0.58	7766.71	0.68	1.73	0.74	1.36	0.69	2.67	0.73	1.07	0.66	1.72
United Kingdom	0.68	1.78	0.65	1.51	0.69	2.10	0.65	1.90	0.69	1.57	0.65	1.51
Ukraine	0.64	15.67	0.68	9.38	0.64	6.14	0.68	9.18	0.64	5.65	0.68	8.72

Table 4.1: The value of R^2 and RMSE for every model used in this study.

The results in table 4.1 show that the third model, which incorporates a difference in population between urban and non-urban areas, has a better fit and transformation of the NTL data into the population proxy than the first and second model. These fits can be compared with the way the population density is modelled and how this relates to modelled population values. As can be seen in figure 4.8 on page 22, model 3 also gives a modelled population density for Austria that is closest to the actual population density according to the population data.

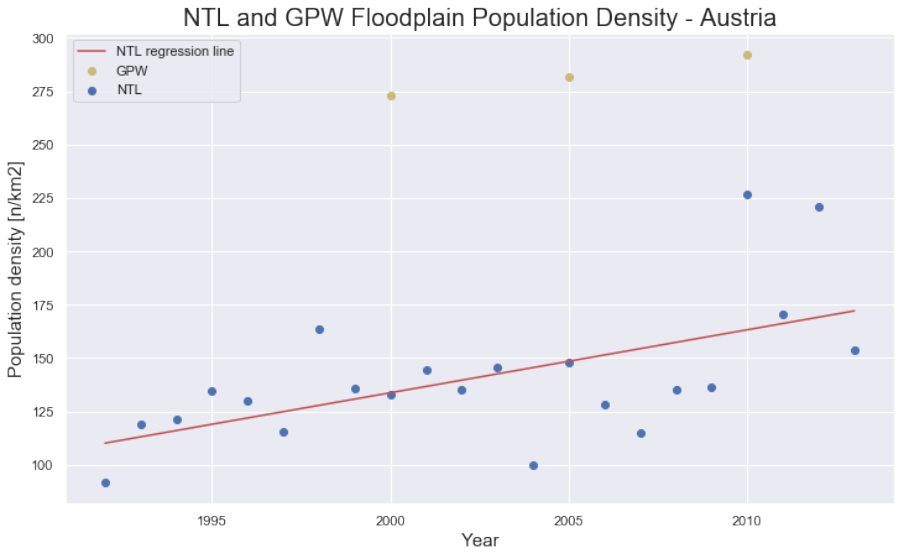

Figure 4.6: Modelled population Density results for Austria from Model 1

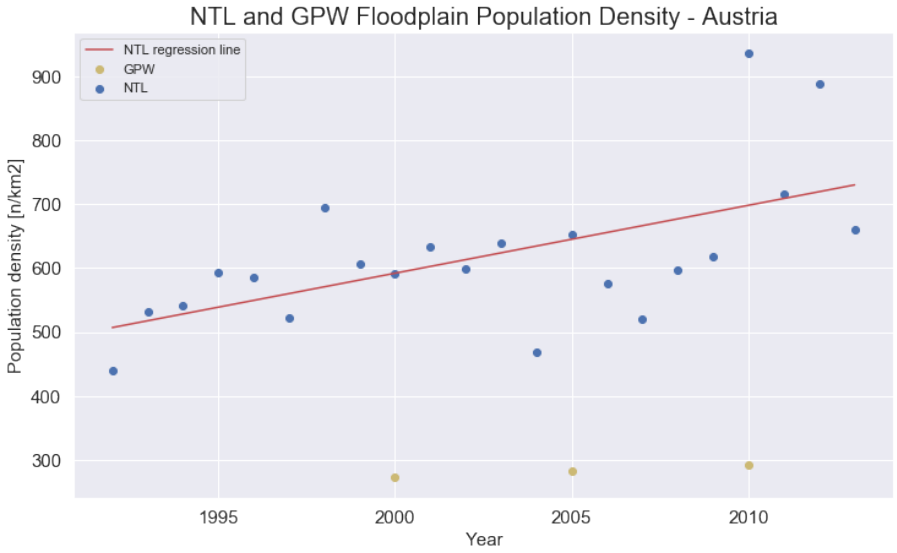

Figure 4.7: Modelled population Density results for Austria from Model 2

Furthermore, the 95 percent interval was examined. In picture 4.9, the example of Austria, you can see that the confidence interval has large bounds. This means that the standard deviation is very high and has a high spread. This is due to the use of a trend amongst a high range of values. This can be observed in 3.3. Because the readability of the figures decreases when the confidence interval is added, the confidence interval will not be plotted hereafter.

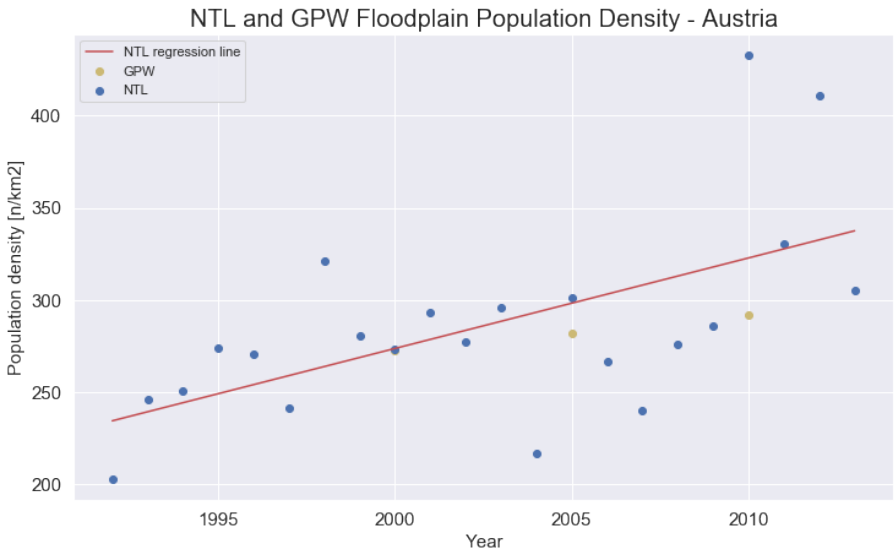

Figure 4.8: Modelled population Density results for Austria from Model 3

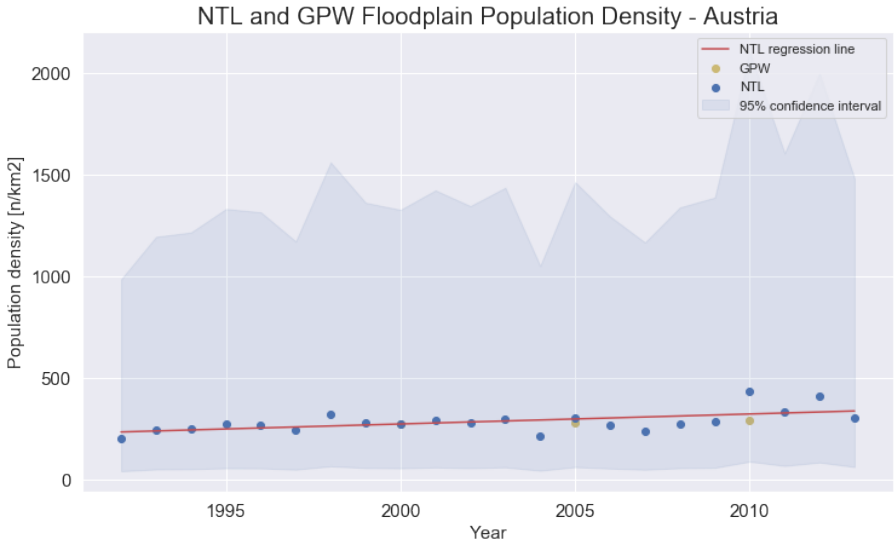

Figure 4.9: The population density proxy modelled with the 95 percent confidence interval.

The model performance is further validated by comparing the absolute differences between the predicted population proxy and the GPW population data. The absolute difference for all countries is below 8 percent and for most countries below 5 percent, as shown in figure 4.10. The difference is smaller when the model transformed the values of entire countries, compared to transforming only in the floodplains. In addition to the correlations, RMSE and the absolute difference between predicted and real value, the Lambda transformation parameter predicts characteristics of light behaviour in countries. The Lambda value determines the division of lights. Eastern European countries tend to have values closer to 1 and Western European countries closer to zero. This can also be seen from the distribution of NTL in the previous paragraph. A value close to 1 corresponds to the distribution found in Eastern European countries earlier, where there are a lots of low NTL values and only a few high NTL values. In Western European countries the value is closer to zero and this means that there are a lot of bright lit pixels, but also an abundance of dim lit pixels. The lambda numbers can be found in appendix C.

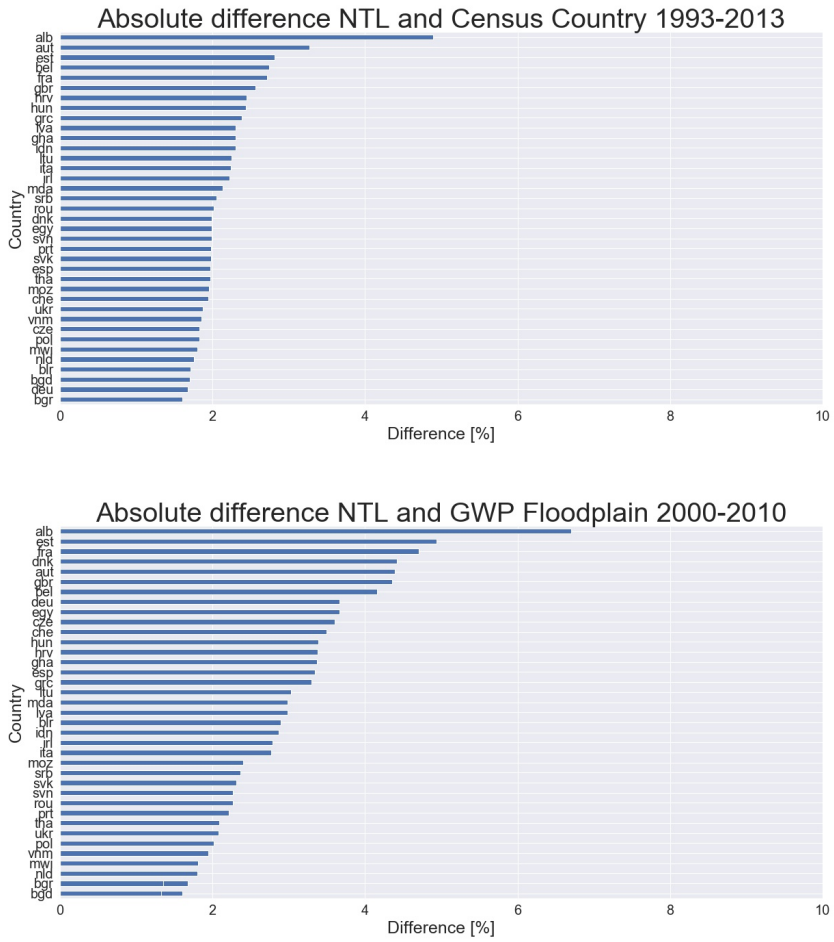

Figure 4.10: Absolute difference between the modelled population through NTL and census data for 1993-2013 (Top) and the absolute difference between NTL and GPW data for the floodplain for 2000-2010 (Bottom)

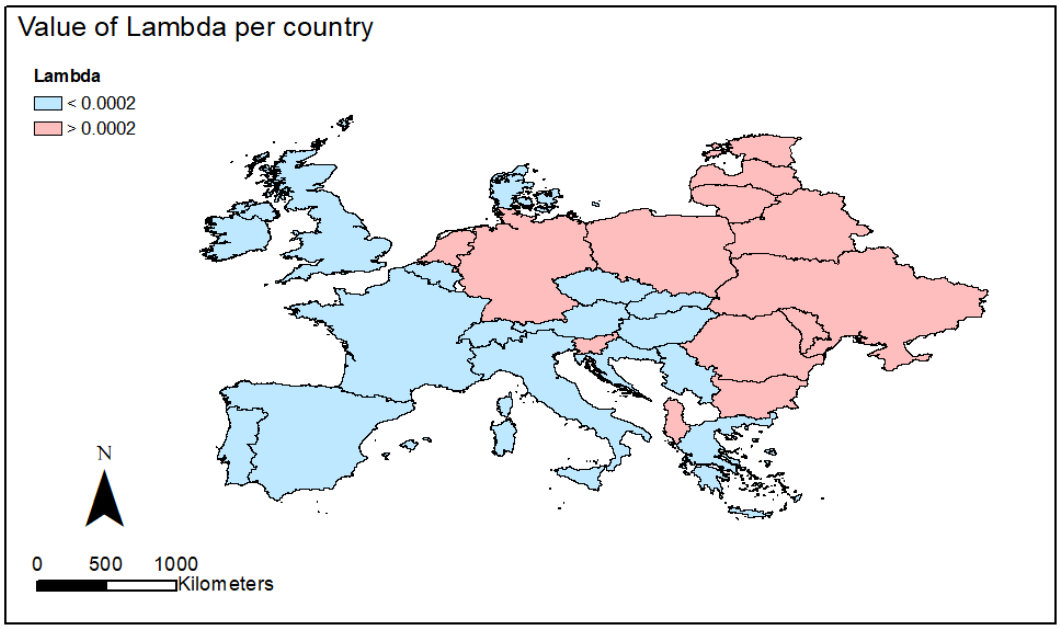

Figure 4.11: Lambda values for the floodplain, obtained by model 3.

4.3. Capturing the population proxy

In this section the spatial and temporal patterns of the population proxy are investigated. For this investigation the results are based on the results of model 3 that accounts for the valuation of urban areas.

4.3.1. Spatial differences in population proxy

The census change rates of figure 4.12 show a division between Western, Central and Eastern European countries. Western European countries tend to grow in population and Eastern European countries either grow moderately in population or have declining populations.

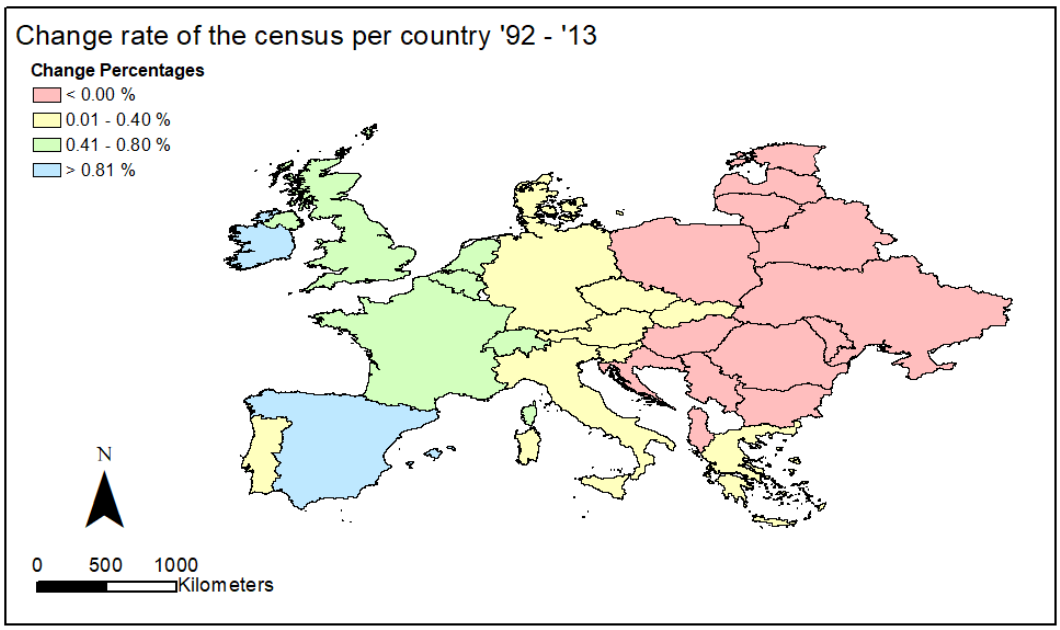

Figure 4.12: Population density change rate per country, based on census results for the years 1992-2013

However, the results of the change rate of the population proxy, shown in figure 4.13 contain different divisions. A clear pattern cannot be observed, the countries that show decline (or very little growth) in transformed population are clustered in either the lower left corner of figure 4.3 which shows antipole behaviour of lights, or in the middle-right area, which shows population centric behaviour of lights. The same can be said of the change rates for the floodplain, shown in figure 4.14. Also, the percentages of population proxy change rate are higher than census change rate. This corresponds with the findings of 4.2

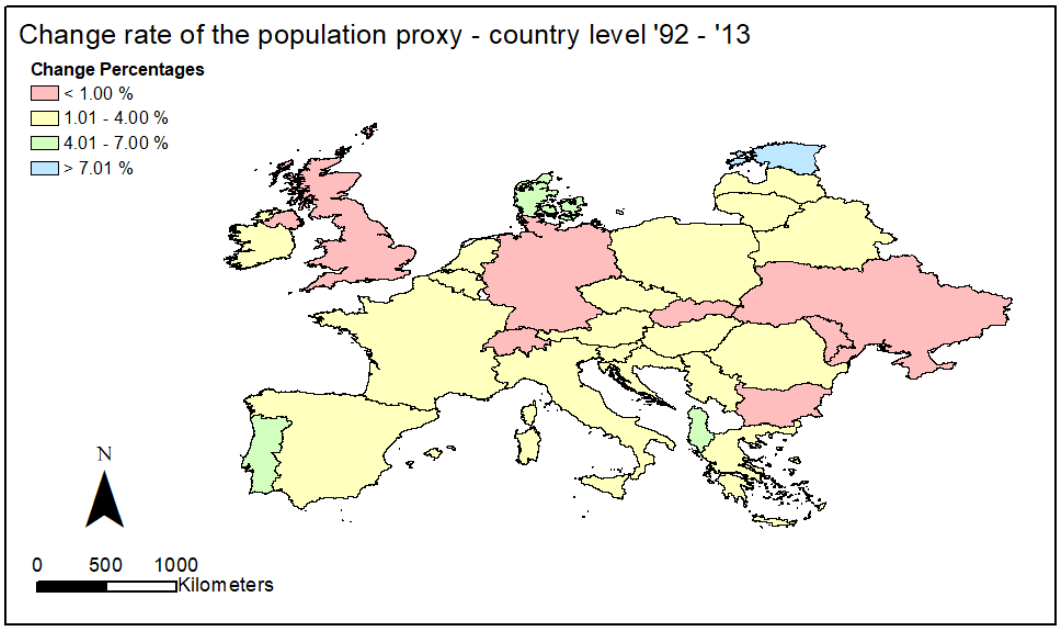

Figure 4.13: Population proxy density change rate per country for the years 1992-2013

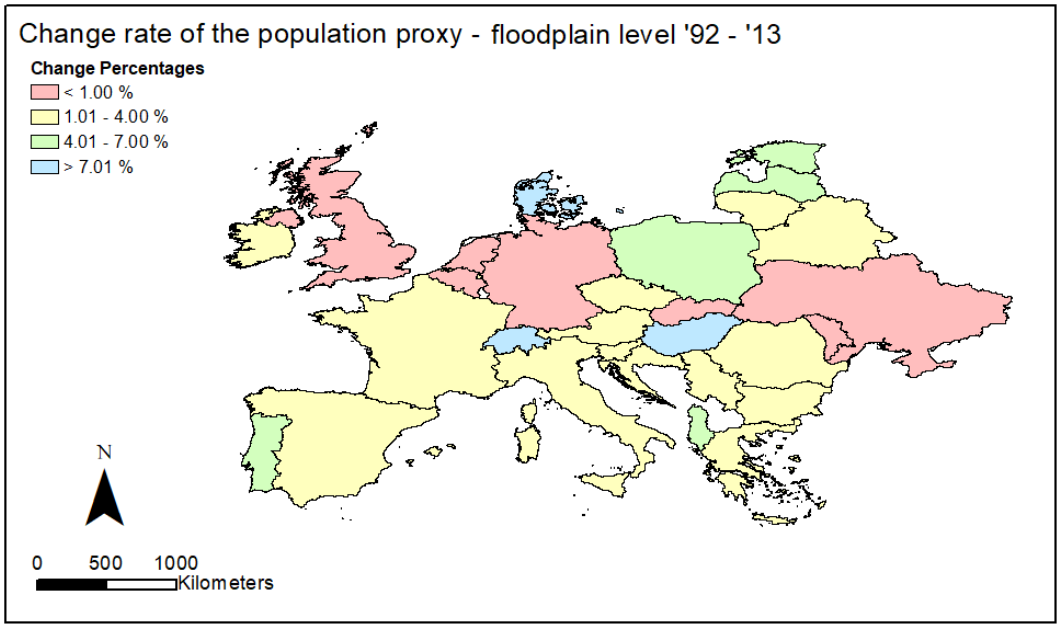

Figure 4.14: Population proxy density change rate per country modelled for the floodplain for the years 1992-2013

4.3.2. Temporal behavior of the population proxy

The growth dynamics of population do not seem to coincide with the lights or the transformed population everywhere, which can be explained by the large confidence interval bounds, even though R^2 showed promising results. However it does seem to coincide with light behaviour divisions, which relate the lights to GDP and population. Therefore, the yearly sum of transformed population of countries and their floodplains are modelled individually.

Rapid/Moderate Lighting growth

One of the countries categorized under Rapid/Moderate growth of light is Indonesia. As you can see in figure 4.15, the growth of the proxy is rapid as well. The results show that the growth of the population proxy contains the same development as the actual population. However, the population proxy underestimates the actual density of people per km^2 living in Indonesia.

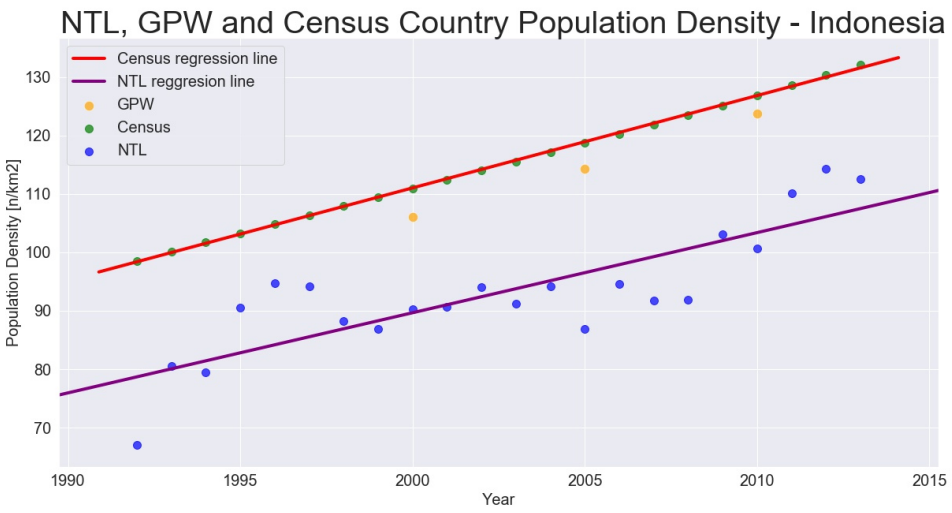

Figure 4.15: Development of the Population Proxy for the years 1992-2013 for Indonesia, compared with census data and GPW data

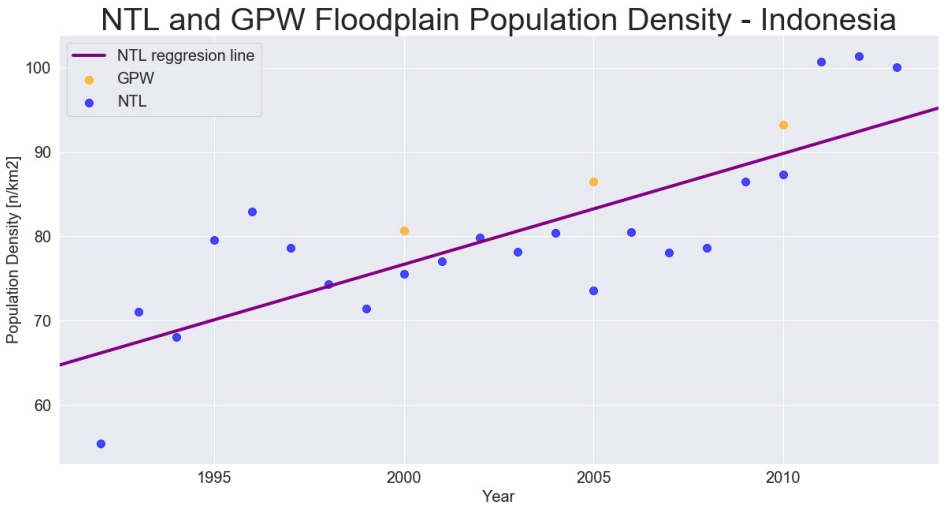

Figure 4.16: Development of the Population Proxy for the years 1992-2013 for the floodplain of Indonesia, compared with GPW data

Figure 4.16 shows the development of the population proxy in the floodplain. It was established in the last figure that the modelled population density follows the same trajectory as the actual population density. Furthermore, figure 4.16 shows that the modelled population density is close to the population density given by the population data. Another country categorized under rapid/moderate growth is Austria. The results indicate a growth in population. The population density modelled by the proxy for the whole country is underestimated, similar as for Indonesia. Also similar is that the floodplain population density is modelled more accurately than for the whole country.

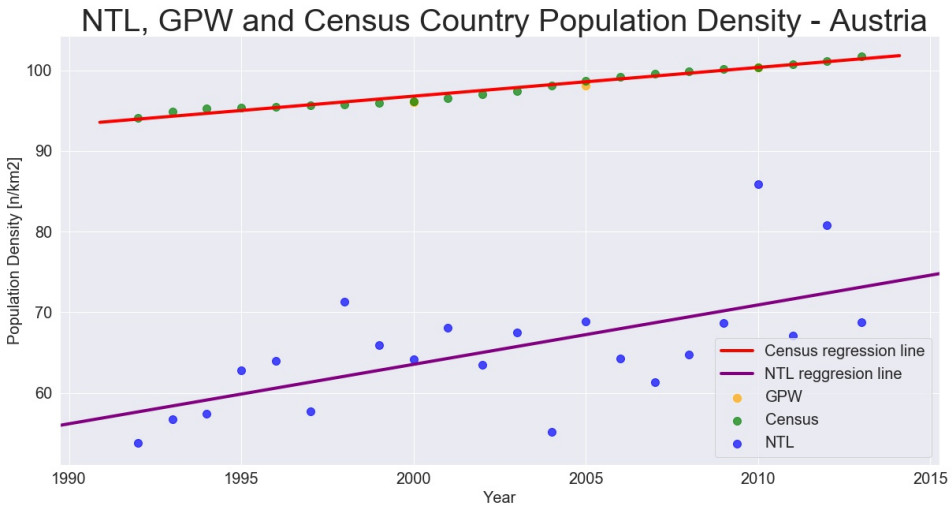

Figure 4.17: Development of the Population Proxy for the years 1992-2013 for Austria, compared with census data and GPW data

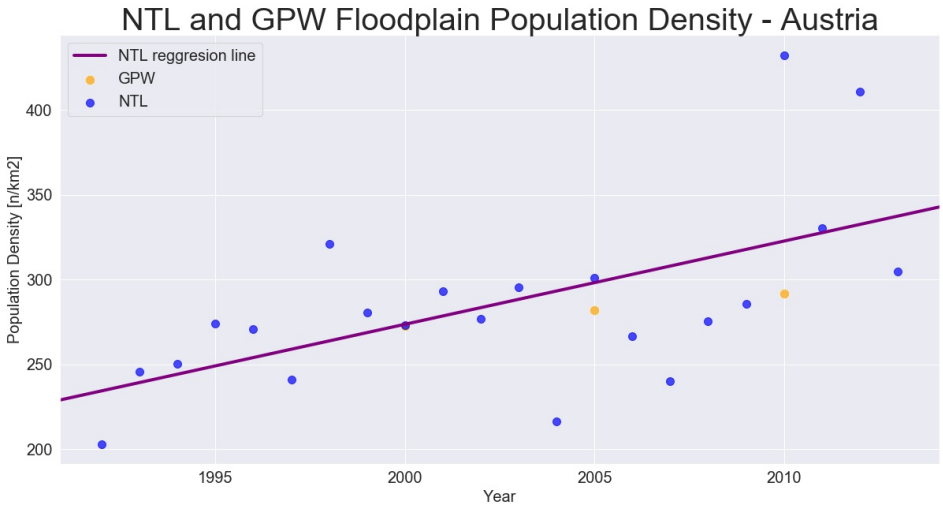

Figure 4.18: Development of the Population Proxy for the years 1992-2013 for the floodplain of Austria, compared with GPW data

Stable/Erratic Lighting growth

Bangladesh is categorized under stable growth. The results in figure 4.19 show that the dynamics of the population proxy and the actual density are counter intuitive. The transformed population density seems to be stable and not growing, but in reality, the population density in Bangladesh is growing at a rapid rate. The same is happening in the floodplain.

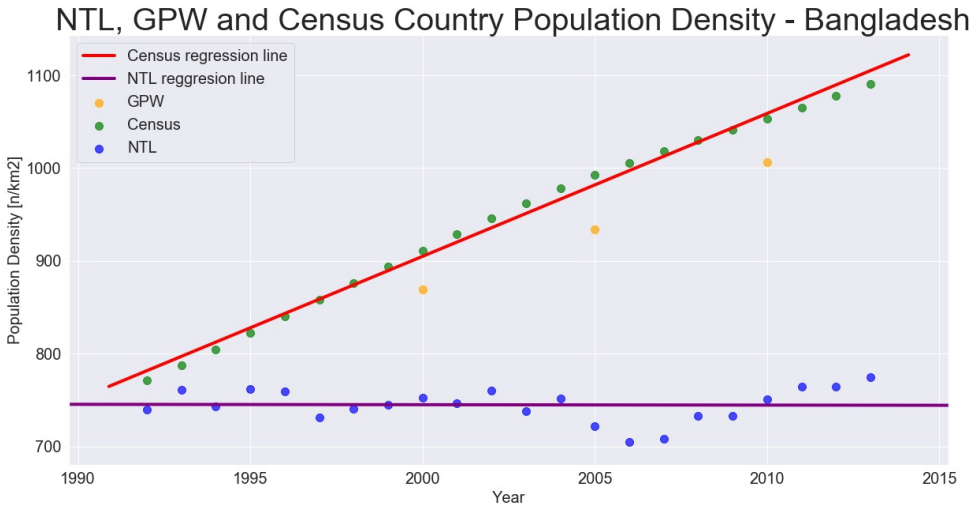

Figure 4.19: Development of the Population Proxy for the years 1992-2013 for Bangladesh, compared with census data and GPW data

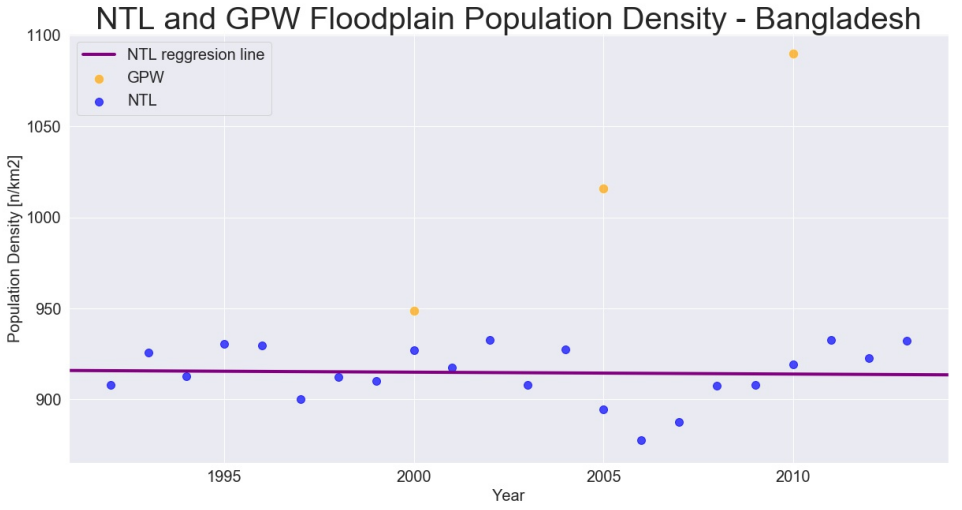

Figure 4.20: Development of the Population Proxy for the years 1992-2013 for the floodplain of Bangladesh, compared with GPW data

Germany is another example of stable light behaviour and the data shows that the trajectory followed by the population proxy is stable. The actual population is stable, but not declining. Furthermore, the population proxy is underestimated when viewing the census and GPW data, both for the whole country and for the floodplain

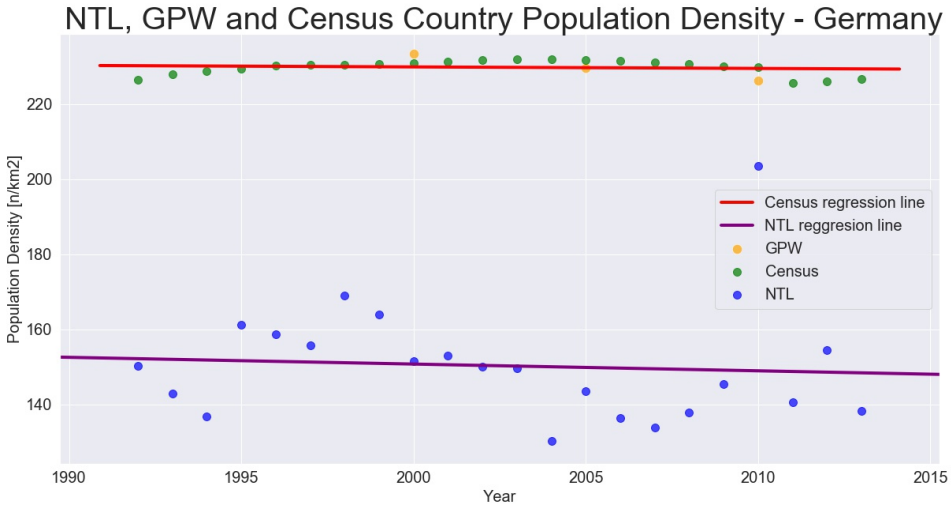

Figure 4.21: Development of the Population Proxy for the years 1992-2013 for Germany, compared with census data and GPW data

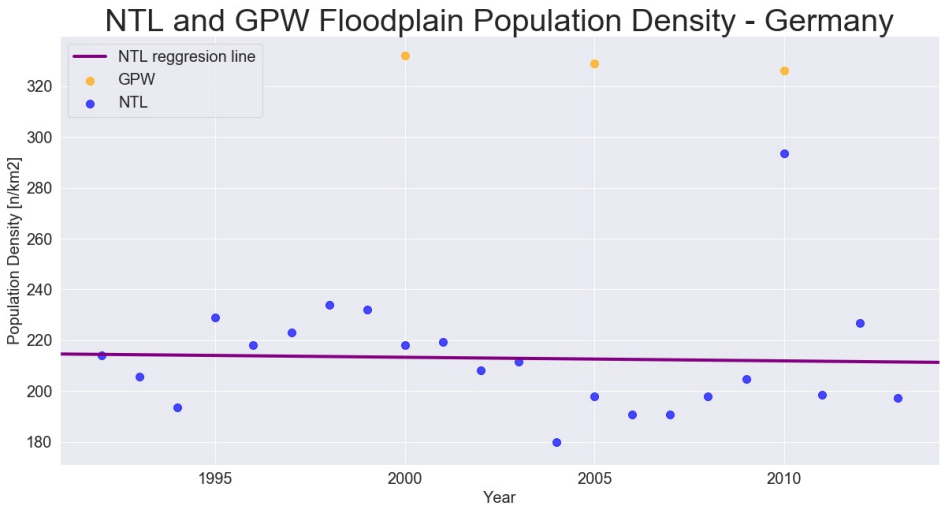

Figure 4.22: Development of the Population Proxy for the years 1992-2013 for the floodplain of Germany, compared with GPW data

Antipole Lighting behaviour

An example of antipole lighting behaviour is Great Britain. The Nighttime Lights decrease both in the floodplain and in the whole country, though the census data shows a growing trend.

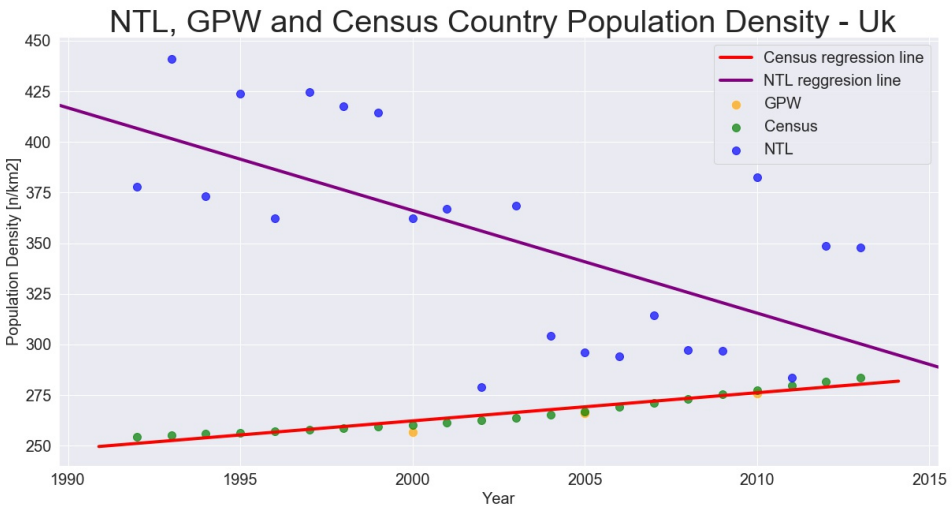

Figure 4.23: Development of the Population Proxy for the years 1992-2013 for Great Britain, compared with census data and GPW data

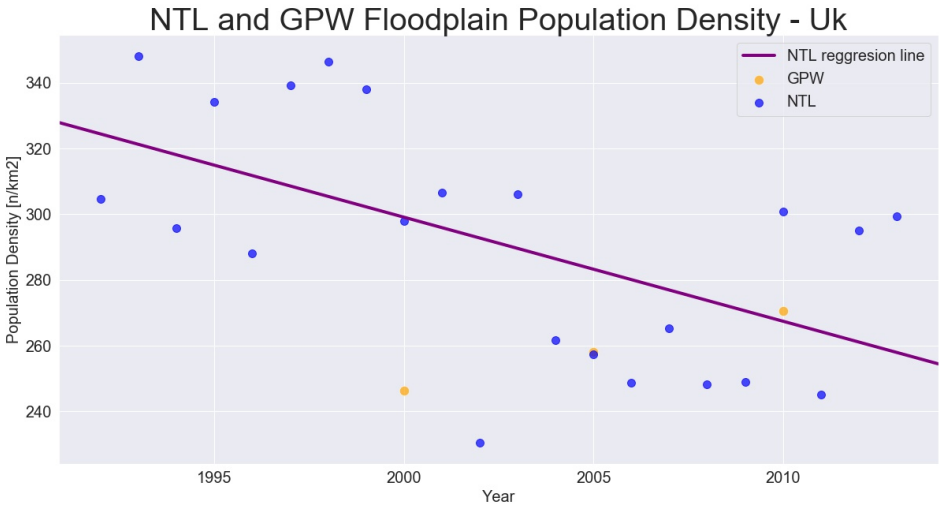

Figure 4.24: Development of the Population Proxy for the years 1992-2013 for the floodplain of Great Britain, compared with GPW data

Population centric

An example of population centric lighting is Ukraine. This means that the correlation with population is positive and the correlation with GDP is negative. The population proxy also shows this, as both the census data and the transformed data show a declining trend. The transformed population proxy has a steeper declining trend than the census data, but other than that the population density seems to be captured quite good.

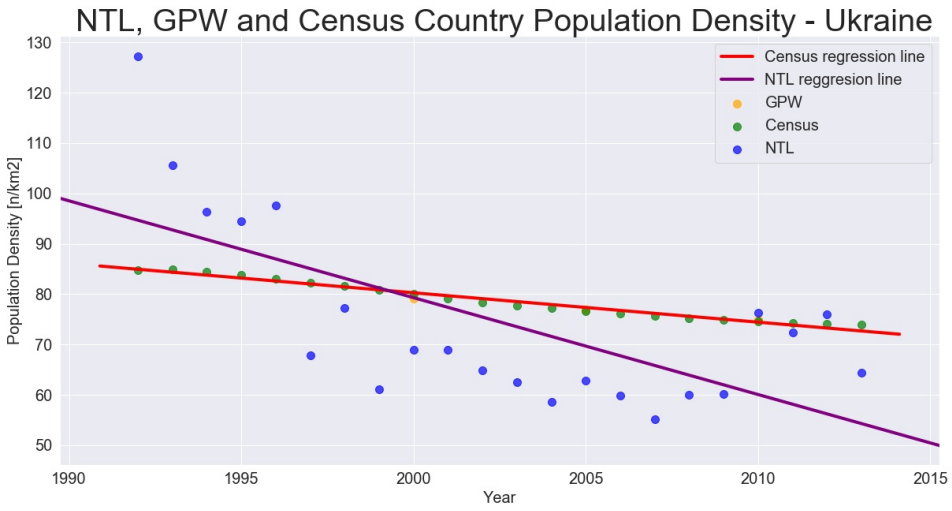

Figure 4.25: Development of the Population Proxy for the years 1992-2013 for Bangladesh, compared with census data and GPW data

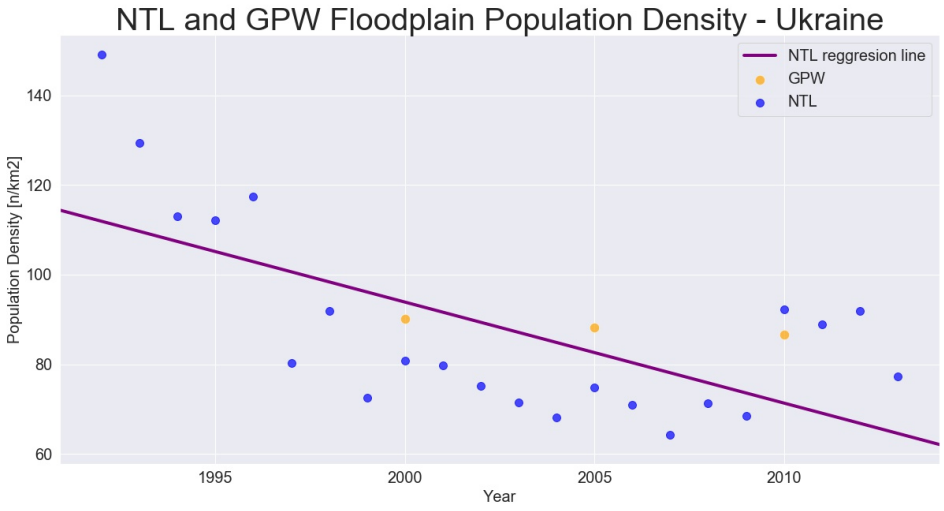

Figure 4.26: Development of the Population Proxy for the years 1992-2013 for the floodplain of Bangladesh, compared with GPW data

GDP centric

Albania is a country with GDP centric lighting behaviour. A lot of countries in Eastern Europe show this behaviour, as it shows that the GDP is increasing, but population is declining. This behaviour can also be seen in the transformed population, which is increasing steeply, though the census shows decline. The value of the population density in the floodplain is captured quite well for the year 2005, even though the lights primarily describe GDP in Albania.

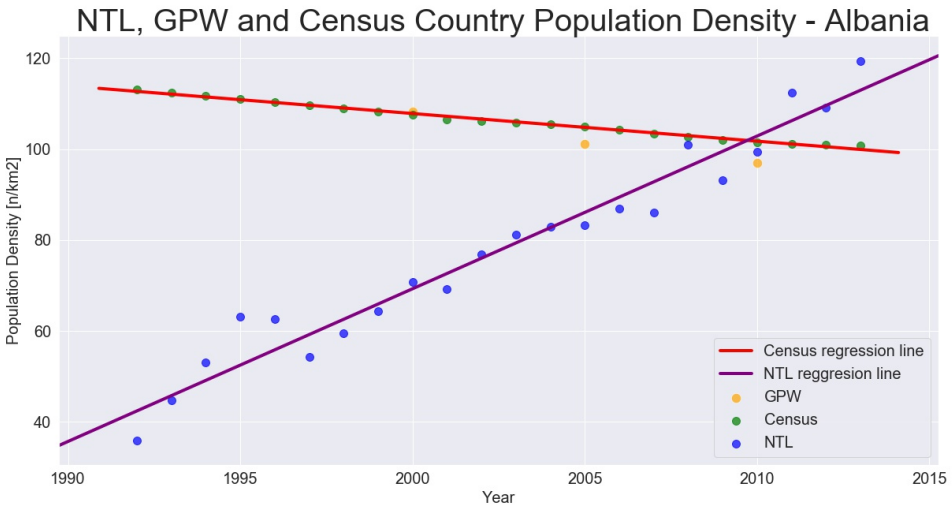

Figure 4.27: Development of the Population Proxy for the years 1992-2013 for Bangladesh, compared with census data and GPW data

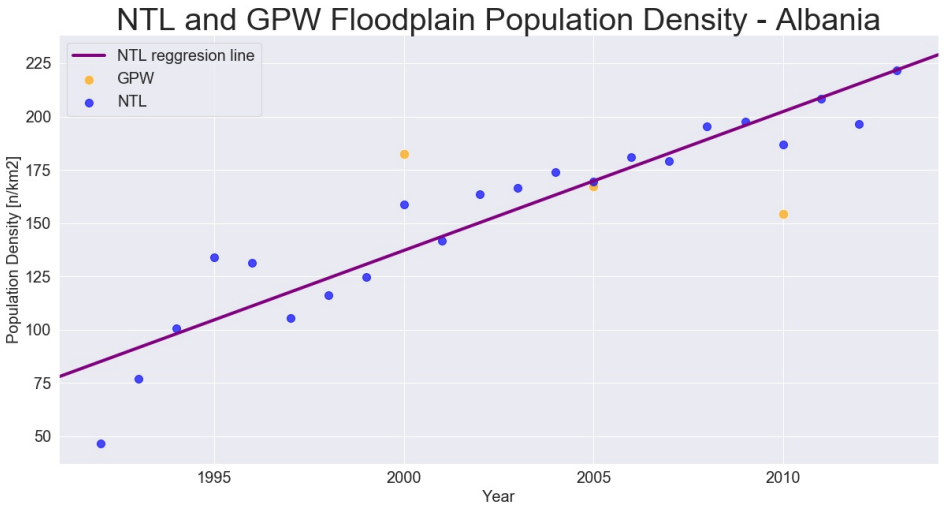

Figure 4.28: Development of the Population Proxy for the years 1992-2013 for the floodplain of Bangladesh, compared with GPW data

4.4. Linking NTL change to floods

The coping patterns defined in the introduction are technological societies and green societies and can be distinguished through, amongst others, growth or decline in population density in areas that have been hit by river floods (which are located in floodplains). To see if there are coping patterns recognizable in areas of big historical disasters, the spatial extents of these disasters were clipped with the population proxy files. Because the chance of success in modelling the population behaviour of countries is the highest for countries with a high correlation with GDP and population, countries in the category of rapid or moderate lighting growth have been selected: Austria, Italy and Mozambique. The disasters were selected through the historical flood database of the Dartmouth Flood Observatory. For Austria the 2002 European flood was chosen, as it was the most costly flood event in the period 1992-2013. Furthermore, the flood event of Italy in the Lombardia region for 2000 was chosen as it also made the top 10 of costliest flood events (Choryński, Pińskwar, Kron, Brakenridge and Kundzewicz, 2012). For Mozambique a flood of 2000 was chosen, as this was the most damaging flood in the years between 1992-2013. The expectation is that Austria and Italy are technological societies and that Mozambique is a green society.

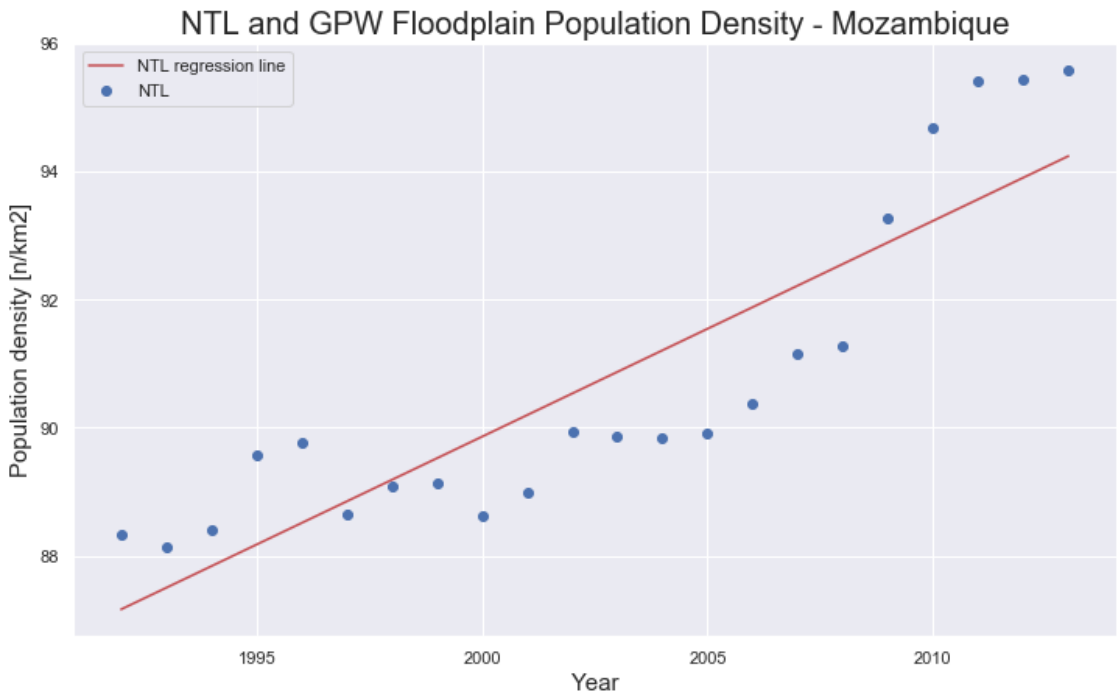

Figure 4.29: Proxy Population Density modelled for Mozambique for the disaster area of the flood of 2000

	Change rate	Change rate before flood	Change rate after flood
Italy 2000 flood	3.474	3.919	3.200
Austria 2002 flood	4.689	7.310	2.306
Mozambique 2000 flood	0.378	0.045	0.583

Table 4.2: Change rate of population proxy growth for three different area's in Mozambique

The figures show that in all three cases, the populations keeps growing, before and after the flood events and no drop in NTL is visible. Also the datapoints are fluctuating too much to know if a drop in NTL is caused by an error in the satellite capturing luminosity or due to a disaster. The change rates before and after the flood show different behaviour than was expected beforehand. The results show that the NTL grows in Mozambique and that a resettling effect that sometimes manifests in green societies, cannot be seen. Italy and Austria can also not be identified as technological societies, as their change rate declines.

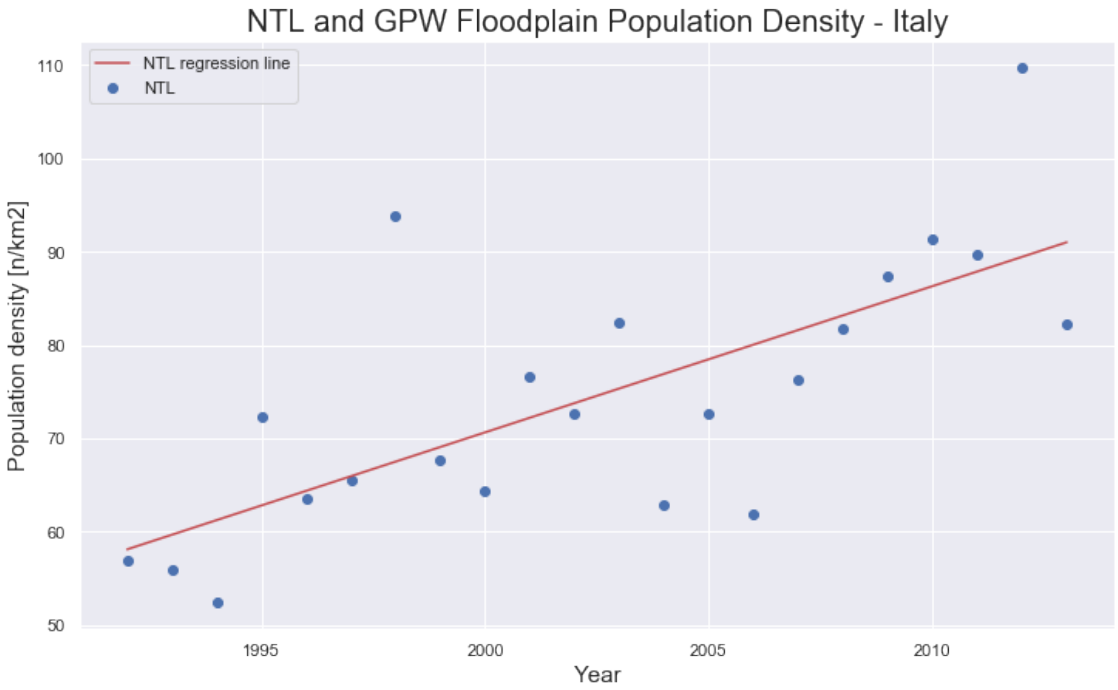

Figure 4.30: Proxy Population Density modelled for Italy for the disaster area of the flood of 2000

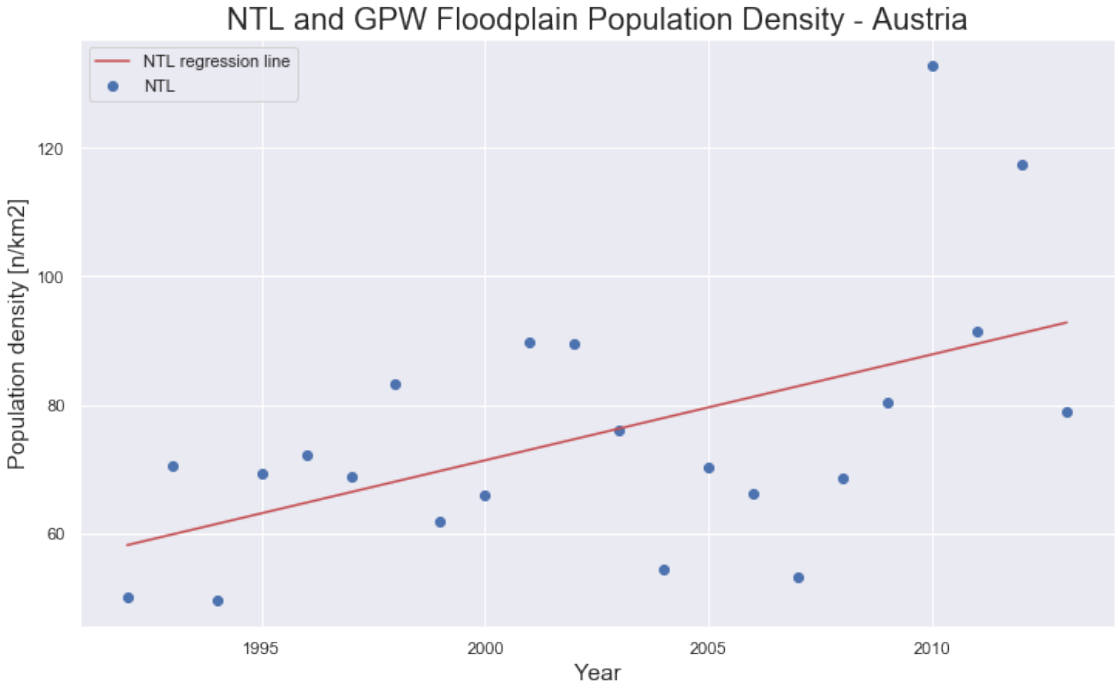

Figure 4.31: Proxy Population Density modelled for Austria for the disaster area of the flood of 2002

4.4.1. Relating Monetary Loss to NTL

The hypothesis is that green societies have a decreasing population density and that technological societies an increasing population density, but this assumption might be too harsh. Therefore it was tested if countries with large monetary losses develop differently from countries with smaller monetary losses. This was also researched by Ceola et al. (2014), but did not take change into account.

The figures show that the initial assumptions found in Ceola et al. (2014) are right, as countries with

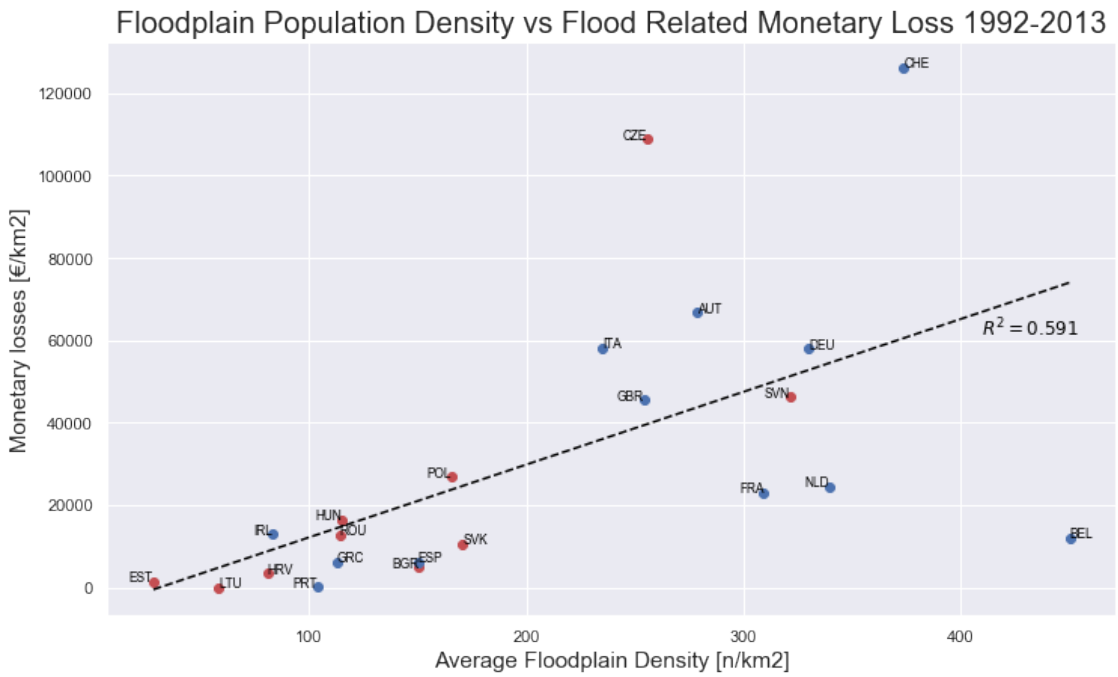

Figure 4.32: Flood related monetary losses per km2 for the year 1992-2013 correlated with the average floodplain density for several European countries.

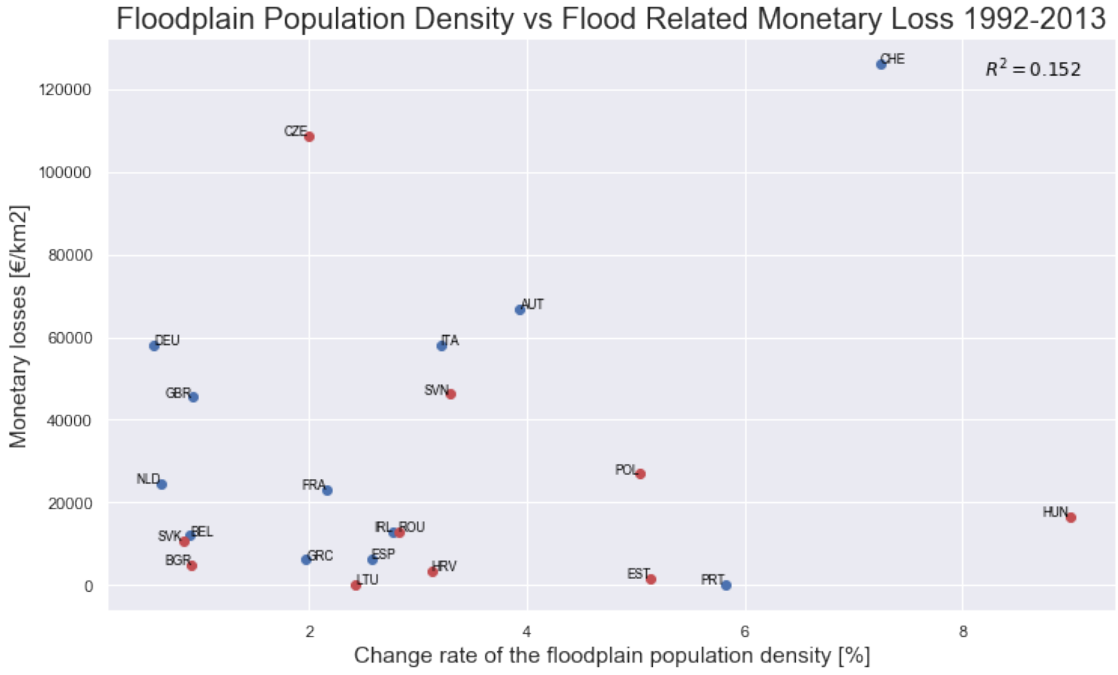

Figure 4.33: Flood related monetary losses per km2 for the year 1992-2013 correlated with the change rate in floodplain population density for several European countries.

high NTL in the floodplain experience larger flood losses. This is different from looking at change rates of population density. The correlation is too low to find a relation between change rate and monetary losses, which indicates that every country has different ways of coping with floods.

5

Discussion

In this chapter, the results of the population density quantification with NTL data and the behaviour of NTL data in different countries are discussed.

5.1. Quantifying Population Density with NTL

Although the preliminary validation of the regression models showed promising results, NTL has to be used with caution due to multiple reasons. Firstly, the confidence intervals of the regression results are very high due to the use of a trend amongst a high range of values. This can be observed in 3.3. Furthermore, it is difficult to verify the floodplain population density results against real data, because only three points of gridded population data are available for the floodplain. The countrywide data is checked against census numbers from the worldbank. Also, the relationship between average NTL and population density in the floodplain is not constant. Here, the model picked a relation based on the interpolation of three years, but this relation may have changed substantially between 1992 and 2013. Next to the lack of verification data, the pitfalls of NTL data cause urgency to use the data with enough precaution. Why it is difficult to compare NTL change over the years (and thus population proxy change) is mainly due to blooming, saturation, and the satellites not being on-board calibrated. Blooming and saturation can be counteracted, but there is no proper solution for the lack of on-board calibration. A diminished saturation effect can be observed in the validation results of the third model, which shows better behaviour and a better fit of population density when light emittance is related to a higher population density in urban areas than in rural areas. Next to problems that lie within the NTL dataset, the relationship between population and NTL in a country is of interest. When modelling a population proxy with the use of NTL it seems that a one on one relationship between NTL and population is applied, but this is not the case. Nighttime Lights are more than just population, as there is also a high correlation with GDP. Areas like highways, recreational parks, sports fields, and airports light up, even though no people are living there.

5.1.1. Behaviour of lights in different countries

The behavioral categories defined by Elvidge et al. (2013) show that the influence of GDP and population is not evenly spread in every country. While the rapid, moderate and population centric lighting categories show results that follow the same trajectory as the population data and might be good fits for measuring population density in the floodplain, the other categories have completely different trajectories than their population counterpart and give unreliable results. Hereafter, it is explained why this lighting behaviour is occuring in countries. Antipole lighting behaviour is defined as a country of which the NTL negatively correlates with both GDP and population. Countries that have antipole lighting behaviour are Great Britain, the Netherlands, Belgium and Slovakia. Antipole lighting can easily be explained through changes in lighting policies. These countries all have made improvements to their lighting efficiency to either be less polluting for the environment or save money. For example, The Netherlands have made a switch to LED light which emits less luminance than regular light. Therefore, the Netherlands experience a decline in NTL. On the other hand, in Slovakia they have shut of lights at night to save money. GDP centric lighting behaviour is defined as a positive correlation with GDP

and a negative correlation with population. Many Eastern European countries are part of this category. This is because these countries experience a growth in GDP, but experience a population decline. The ten worlds fastest shrinking populations are all in Eastern Europe, and some of these ten countries fall under the GDP centric lighting behaviour: Bulgaria, Latvia, Croatia, Romania, Serbia, Hungary and Poland. Ukraine and Moldova are part of these ten countries, but their lighting behaviour is explained by Population centric behaviour. This doesn't mean the opposite of GDP centric behaviour though. This behaviour is explained by a growing GDP and a decline in NTL. This is due to governments being unable to provide enough electricity. The last category is erratic and stable lighting. Erratic behaviour occurs in some countries, and the main assumption behind this is that the erratic behaviour is due to noise within the NTL dataset. When countries experience stable lighting, they actually just lack strong correlations between GDP and NTL data. Next to nighttime lights containing information on GDP and population, the electrification of a country has to be taken into account. For example, in Myanmar there is no universal access to electricity. So not all population will be visible in the data that is used to identify population behavior in Myanmar. Furthermore, when population density declines in an urbanized area, the lighting does not always decrease. Thus, NTL should be used with caution. For example, Ceola et al. (2014) shows the growth rates of NTL near river networks. However, a country like Bangladesh is overlooked here as a danger zone for floodplain development as the lights are not growing as fast as the population is. The categories derived by Elvidge et al. (2013) give insight in which countries the population density can be predicted relatively accurate and for which for which countries making a fit based on population density does not create reliable results.

5.2. Relating NTL to floods

The work done by Mard et al. (2018) seemed promising to be able to tell apart different flood coping mechanisms of different countries, but the question is if NTL is the right dataset to research this. At first it was hypothesized that drops in NTL in the floodplain of a certain country could have been the effect of devastating flood events. However drops and rises in NTL cannot be explained that easily. First, the erratic behaviour (due to the satellites capturing NTL not being on-board calibrated) can cause drops. Furthermore, there can be more than one reason for a drop in NTL, as NTL is not only the result of GDP and population. There are probably a multitude of explanations. In general, the human behaviour in floodplains is a result of socio-economic and hydro-climatic variables and these occur over longer timescales and not instantaneously, which would make the timescale of NTL too short to see shifts in behaviour. Thus, instantaneous drops cannot be observed by NTL, but change rates over the timescale of NTL (1992-2013) might give more insight. Therefore, the monetary losses of the floodplains were modelled against the change rate of the population proxy. A relation between the proxy of average population density in the floodplain was found, which was also found by Ceola et al. (2014). A higher population density in the floodplain also means a higher flood exposure and probably more assets in the floodplain. When a flood occurs, a flood is more damaging in countries with high floodplain density than in countries with low population density. The relationship between monetary loss and change rate is therefore less evident. This could be due to every country having a different way of coping with floods. Also, the change rate of some countries do not change a lot for example, because the floodplain is already very densely populated, or because the electrification is growing very rapidly.

In addition, the change rate of some countries does not change a lot because the floodplain is already very dense and it changes a lot in some countries because the electrification rate is very high.

6

Conclusion and Recommendations

6.1. Conclusion

The objective of this research was to understand, interpret and quantify the spatial and temporal evolution of Nighttime Light, linking it to human development in the floodplain. This was done by modelling a population proxy based on the relation between NTL and population data in the floodplain. In this way, the behaviour of humans in floodplains could be monitored, while also accounting for institutional effects in the NTL dataset, allowing comparison of results between countries.

The research question of this thesis is: "Can Nighttime Light be used to accurately quantify and monitor changes in population density in floodplains and therefore be used to identify flood coping mechanisms?" The first part of this question investigates if NTL is a good enough proxy of population density. Although validation of the model, in terms of R^2 and RMSE seemed promising, there is not enough data available on population in the floodplain to know for certain that the data over the longer timescale is accurate. Furthermore, the behaviour of NTL differs in every country, thus it probably cannot be described with one regression relation alone. The categories derived by Elvidge et al. (2013) give insight in which countries the population density can be predicted relatively accurate (which are countries whose growth in population and GDP change in the same ratio as NTL) and for which countries a fit based on population density will be less accurate. The latter countries cannot be modelled well due to the lighting not being affected by GDP and Population in the same ratio. This is due to a decline in population growth not being noticeable in declining light, thus having a NTL growth that is dominated by GDP growth. Furthermore, important factors are the governments not being able to provide enough lighting or due to regulations in lighting efficiency for environmental reasons. Secondly, the NTL in itself poses limitations. The inner noise of NTL itself is high due to satellites capturing NTL not being on-board calibrated, resulting in satellite images that are hard to compare. The inter-calibration looks promising as well, but an optimal calibration is not achieved yet (and can probably not be achieved). This results in convergence in satellite images. All together, there is a multitude of explanations for the behaviour of light and it is therefore hard to accurately predict population through this proxy.

The second part of the research question investigates whether the population proxy constructed with the use of NTL can be used to identify flood coping mechanisms. Although, the population proxy has its limitations right now, future datasets are promising in being able to capture the population proxy in a more reliable way. For now, the NTL population proxy is too noisy in capturing human reaction to events. In addition, drops and growth in NTL cannot be explained by single events and can be explained by a multitude of reasons. Furthermore, the time scale of the NTL data is too short to measure flood coping mechanisms because they are not a result of a single event, but of multiple events over multiple years and are influenced by culture and institutions. Instead of looking at single events, the change rate was captured. Here correlations with flood losses are low, thus a relationship between change rate and flood losses is lacking.

6.2. Recommendations

Some datasets can be improved with the use of NTL data, as it has a good distribution of socio-economic variables in a country. Understanding behaviour of light is valuable, because it has the potential to map Socio-Economic variables in data-scarce areas. Especially with the new NTL dataset of Visible Infrared Imaging Radiometer Suite (VIIRS), which is a component of the Suomi National Polar-orbiting Partnership (NPP) (Zhao et al., 2019). This NTL dataset improves the capturing of luminosity compared to DMSP-OLS and has on-board calibration. Understanding how light behaves can improve socio-economic variable modelling in the future. Therefore it is important to have understanding about the meaning of light in different countries. The first recommendation is to refrain from NTL data when researching historical flood coping mechanisms. As the DMSP/OLS NTL data can not accurately compare countries on a large temporal and spatial scale. The second recommendation is that modelling the population density through the use of NTL can be done, but not the most accurate with the DMSP/OLS. One of the main setbacks of NTL data is the noise that is present in NTL data due to the lack of on-board calibration. Fortunately, a new dataset is released from 2012 onwards. This is the VIIRS dataset and it has on-board calibration, thus solving a lot of problems experienced with the DMSP-OLS dataset. VIIRS is not useful for the time period of 1992 - 2011, but very useful for future research. Which is why understanding the behaviour of light is important, so in the future socio-economic variables can be modelled more accurately for data scarce areas. The VIIRS dataset also has a higher resolution and has monthly datasets, which will also capture change in population better, and give even more possibilities for using NTL. For example, capturing the behaviour of people after a single flood event could be more successful, as the changes before and after an event can be monitored one month apart. Furthermore, regional economic development could be modelled for regions in data scarce areas. The country-based regression proved to be able to give a relationship based on institutional effects, but modelling regional effect might also be interesting. Lastly, the migration of population due to climate change, conflicts and economic development could be researched with the use of VIIRS.

A

Intercalibration Coefficients

The graph shows the coefficients found for the intercalibration using equation A.1 using the images of Sicily F121999 for intercalibration.

$$DN_{\text{adjusted}} = C_0 + C_1 \times DN + C_2 \times DN^2 \quad \text{(A.1)}$$

Satellite	Year	C0	C1	C2	R2	Number
F10	**1992**	-2.0570	1.5903	-0.0090	0.9075	35720
F10	**1993**	-1.0582	1.5983	-0.0093	0.9360	38893
F10	**1994**	-0.3458	1.4864	-0.0079	0.9243	36494
F12	**1994**	-0.6890	1.1770	-0.0025	0.9071	34485
F12	**1995**	-0.0515	1.2293	-0.0038	0.9178	37571
F12	**1996**	-0.0959	1.2727	-0.0040	0.9319	35762
F12	**1997**	-0.3321	1.1782	-0.0026	0.9245	37413
F12	**1998**	-0.0608	1.0648	-0.0013	0.9536	37791
F12	**1999**	0.0000	1.0000	0.0000	1.0000	39157
F14	**1997**	-1.1323	1.7696	-0.0122	0.9101	36811
F14	**1998**	-0.1917	1.6321	-0.0101	0.9723	36701
F14	**1999**	-0.1557	1.5055	-0.0078	0.9717	38894
F14	**2000**	1.0988	1.3155	-0.0053	0.9278	37888
F14	**2001**	0.1943	1.3219	-0.0051	0.9448	38558
F14	**2002**	1.0517	1.1905	-0.0036	0.9203	36964
F14	**2003**	0.7390	1.2416	-0.0040	0.9432	38701
F15	**2000**	0.1254	1.0452	-0.0010	0.9320	38831
F15	**2001**	-0.7024	1.1081	-0.0012	0.9593	38632
F15	**2002**	0.0491	0.9568	0.0010	0.9658	38035
F15	**2003**	0.2217	1.5122	-0.0080	0.9314	38788
F15	**2004**	0.5751	1.3335	-0.0051	0.9479	36998
F15	**2005**	0.6367	1.2838	-0.0041	0.9335	38903
F15	**2006**	0.8261	1.2790	-0.0041	0.9387	38684
F15	**2007**	1.3606	1.2974	-0.0045	0.9013	37036
F16	**2004**	0.2853	1.1955	-0.0034	0.9039	36856
F16	**2005**	-0.0001	1.4159	-0.0063	0.9390	38984
F16	**2006**	0.1065	1.1371	-0.0016	0.9199	37204
F16	**2007**	0.6394	0.9114	0.0014	0.9511	37759
F16	**2008**	0.5564	0.9931	0.0000	0.9450	37469
F16	**2009**	0.9492	1.0683	-0.0016	0.8918	33895
F18	**2010**	2.3430	0.5102	0.0065	0.8462	36445
F18	**2011**	1.8956	0.7345	0.0030	0.9095	36432
F18	**2012**	1.8750	0.6203	0.0052	0.9392	37576

Table A.1: Calibration Coefficients Elvidge et al. 2013

B

Relationship GPW and NTL

Relationship GPW and NTL with Bangladesh, Egypt and Vietnam added.

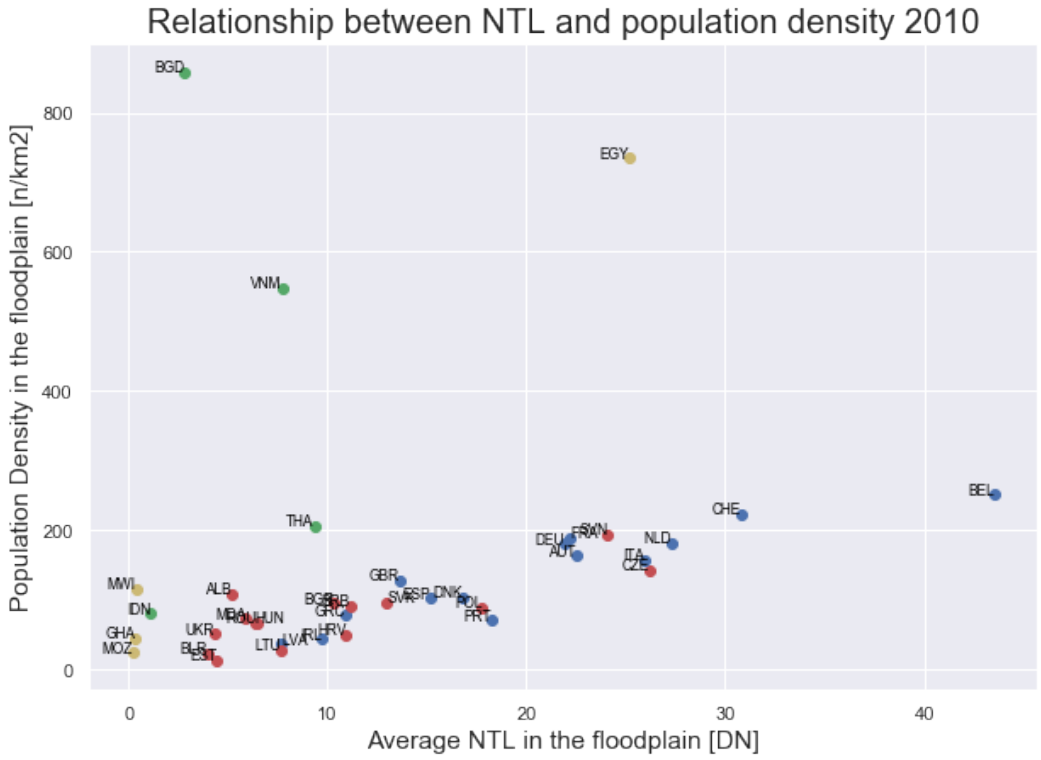

Figure B.1: Relationship between average ntl and population density in the floodplain for the year 2010

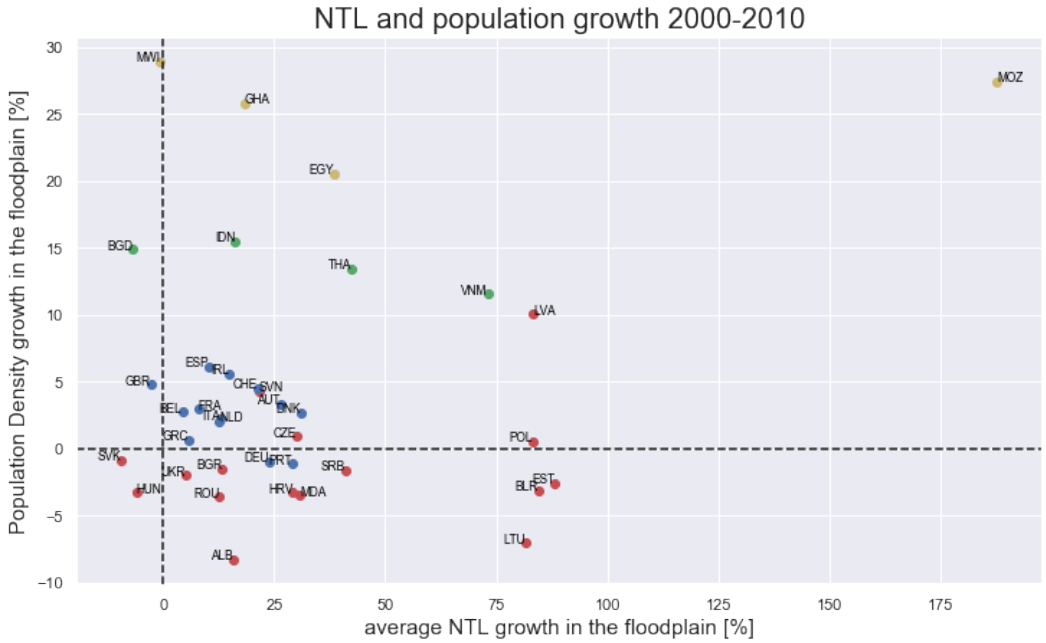

Figure B.2: Relationship between average ntl and population density in the floodplain for the year 2010

C

Lamda Table

Countries	Lambda
Albania	0.67822
Austria	0.00001
Belarus	0.06728
Belgium	0.00006
Bulgaria	0.70678
Croatia	0.00018
Czech Republic	0.00007
Denmark	0.00013
Estonia	0.01723
France	0.00009
Germany	0.00022
Greece	0.00010
Hungary	0.00018
Ireland	0.00004
Italy	0.00018
Latvia	0.08855
Lithuania	0.05651
Moldova	0.61222
The Netherlands	0.00024
Poland	0.00029
Portugal	0.00007
Romania	0.00021
Serbia	0.00015
Slovakia	0.00008
Slovenia	0.25963
Spain	0.00020
Switzerland	0.00005
Great Britain	0.00008
Ukraine	0.50689

Table C.1

Bibliography

[1] Barendrecht, M. H., Viglione, A., and Blöschl, G. (2017). A dynamic framework for flood risk. *Water Security*, 1:3–11.

[2] Baugh, K., Elvidge, C. D., Ghosh, T., and Ziskin, D. (2010). Development of a 2009 stable lights product using dmspols data. *Proceedings of the Asia-Pacific Advanced Network*, 30:114–130.

[3] Bouwer, L. (2011). Have disaster losses increased due to anthropogenic climate change? *Bulletin of the American Meteorological Society*, 92:39–46.

[4] Box, G. and Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, 26:211–252.

[5] Carroll, R. J. and Ruppert, D. (1985). Transformations in regression: A robust analysis. *Technometrics*, 27(1):1–12.

[6] Ceola, S., Laio, F., and Montanari, A. (2014). Satellite nighttime lights reveal increasing human exposure to floods worldwide. *Geophysical Research Letters*, 41(20):7184–7190.

[7] Ceola, S., Laio, F., and Montanari, A. (2015). Human-impacted waters: New perspectives from global high-resolution monitoring. *Water Resources Research*, 51(9):7064–7079.

[8] Ceola, S., Montanari, A., Parajka, J., Viglione, A., Bloshl, G., and Laio, F. (2016). Human signatures derived from nighttime lights along the eastern alpine river network in austria and italy. *Proc. IAHS*, 373:131–136.

[9] Chen, X. and Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *PNAS*, 108(21):8589–8594.

[10] Choryński, A., Pińskwar, I., Kron, W., Brakenridge, R., and Kundzewicz, Z. (2012). Catalogue of large floods in europe in the 20th century. In Kundzewicz, Z., editor, *Changes in Flood Risk in Europe*, chapter 1, pages 27–54. CRC Press.

[11] CIESIN (2016). Documentation for the gridded population of the world, version 4 (gpwv4). data retrieved from NASA Socioeconomic Data and Applications Center (SEDAC), http://sedac.ciesin.columbia.edu/gpw.

[12] Ciullo, A., Viglione, A., Castellarin, A., Crisci, M., and Baldassarre, G. D. (2017). Socio-hydrological modelling of flood-risk dynamics: comparing the resilience of green and technological systems. *Hydrological Sciences Journal*, 62(6):880–891.

[13] Di Baldassarre, G., Martinez, F., Kalantari, Z., and Viglione, A. (2017). Drought and flood in the anthropocene: Feedback mechanisms in reservoir operation. *Earth System Dynamics*, 8(1):225–233.

[14] Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Salinas, J. L., and Blöschl, G. (2013). Socio-hydrology: conceptualising human-flood interactions. *Hydrology and Earth System Sciences*, 17(8):3295–3303.

[15] Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Yan, K., Brandimarte, L., and Blöschl, G. (2015). Debates—perspectives on socio hydrology: Capturing feedbacks between physical and social processes. *Water Resour.Res.*, 51:4770–4781.

[16] Elshorbagy, A., Bharath, R., Lakhanpal, A., Ceola, S., Montanari, A., and Lindenschmidt, K.-E. (2017). Topography- and nightlight-based national flood risk assessment in canada. *Hydrology and Earth System Sciences*, 21:2219–2232.

[17] Elvidge, C., Sutton, P., Baugh, K., Ziskin, D., Ghosh, T., and Anderson, S. (2013). National trends in satellite observed lighting: 1992-2012. In Weng, Q., editor, *Global Urban Monitoring and Assessment Through Earth Observation*, chapter 6, pages 97–119. CRC Press.

[18] Elvidge, C., Ziskin, D., Baugh, K., Tuttle, B., Ghosh, T., Pack, D., Erwin, E., and Zhizhin, M. (2009). A fifteen year record of global natural gas flaring derived from satellite data. *Energies*, 2:595–622.

[19] ESA (2018). Cci land cover product user guide version 2.0. data retrieved from European Space Agency (ESA) at http://maps.elie.ucl.ac.be/CCI/viewer/download.php.

[20] Fang, Y., Ceola, S., Paik, K., McGrath, G., Suresh, P., Rao, C., A., M., and Jawitz, J. (2018a). Globally universal fractal pattern of human settlements in river networks. *Earth's Future*, 6:1134–1145.

[21] Fang, Y., Du, S., Scussolini, P., Wen, J., He, C., Huang, Q., and Gao, J. (2018b). Rapid Population Growth in Chinese Floodplains from 1990 to 2015. *International Journal of Environmental Research and Public Health*, 15(8).

[22] Imhoff, M. L., Lawrence, W. T., Stutzer, D. C., and Elvidge, C. D. (1997). A technique for using composite dmsp/ols "city lights" satellite data to map urban area. *Remote Sensing of Environment*, 61:361–370.

[23] Ivis, F. J. (2006). Calculating geographic distance: Concepts and methods.

[24] Jongman, B. (2018). Effective adaptation to rising flood risk. *Nature Communications*, 9(1):1986.

[25] Jongman, B., Ward, P. J., and Aerts, J. C. (2012). Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change*, 22:823–835.

[26] Kundzewicz, Z. and Schnellnhuber, H.-J. (2004). Floods in the ipcc tar perspective. *Natural Hazards*, 31:111–128.

[27] Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G. R., Kron, W., Benito, G., Honda, Y., Takahashi, K., and Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal*, 59(1):1–28.

[28] Li, X. and Zhou, Y. (2017). A stepwise calibration of global dmsp/ols stable nighttime light data (1992-2013). *Remote Sensing*, 9:637.

[29] Mileti, D. (1999). *Disasters by Design: A Reassessment of Natural Hazards in the United States*. The National Academies Press.

[30] Mård, J., Di Baldassarre, G., and Mazzoleni, M. (2018). Nighttime light data reveal how flood protection shapes human proximity to rivers. *Science Advances*, 4(8):eaar5779.

[31] Nardi, F., Annis, A., Di Baldassarre, G., Vivoni, E. R., and Grimaldi, S. (2019). GFPLAIN250m, a global high-resolution dataset of Earth's floodplains. *Scientific Data*, 6(1):309.

[32] Nardi, F., Vivoni, E. R., and Grimaldi, S. (2006). Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. *Water Resources Research*, 42(9).

[33] NOAA-NGDC (2018). Version 4 dmsp-ols nighttime lights time series. data retrieved from NOAA-NGDC, https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html.

[34] Pande, S. and Sivapalan, M. (2017). Progress in socio-hydrology: a meta-analysis of challenges and opportunities. *Wiley Interdisciplinary Reviews: Water*, 4(4):e1193.

[35] Pandey, B., Zhang, Q., and Seto, K. C. (2017). Comparative evaluation of relative calibration methods for dmsp/ols nighttime lights. *Remote Sensing of Environment*, 195:67–78.

[36] Paprotny, D., Morales-Nápoles, O., and Jonkman, S. N. (2018). Hanze: a pan-european database of exposure to natural hazards and damaging historical floods since 1870. *Earth Syst. Sci. Data*, 10:565–581.

[37] Pinkovskiy, M. L. (2017). Growth discontinuities at borders. *J. Econ. Growth*, 22:145–192.

[38] Ritchie, H. and Roser, M. (2019). Natural disasters. *Our World in Data*. https://ourworldindata.org/natural-disasters.

[39] Shen, Z., Zhu, X., Cao, X., and Chen, J. (2019). Measurement of blooming effect of dmsp-ols nighttime light data based on npp-viirs data. *Annals of GIS*, 25(2):153–165.

[40] Sivapalan, M., Savenije, H. H. G., and Blöschl, G. (2012). Socio-hydrology: A new science of people and water: INVITED COMMENTARY. *Hydrological Processes*, 26(8):1270–1276.

[41] Small, C., Pozzi, F., and Elvidge, C. D. (2005). Spatial analysis of global urban extent from dmsp-ols night lights. *Remote Sensing of Environment*, 96:277 – 291.

[42] Sutton, P., Elvidge, C., and Obremski., T. (2003). Building and evaluating models to estimate ambient population density. *Photogrammetric Engineering Remote Sensing*, 69(5):545–553.

[43] Tockner, K. and Stanford, J. A. (2002). Riverine flood plains: present state and future trends. *Environmental Conservation*, 29(3):308–330.

[44] Viglione, A., Di Baldassarre, G., Brandimarte, L., Kuil, L., Carr, G., Salinas, J. L., Scolobig, A., and Blöschl, G. (2014). Insights from socio-hydrology modelling on dealing with flood risk – Roles of collective memory, risk-taking attitude and trust. *Journal of Hydrology*, 518:71–82.

[45] White, G. (1945). Human Adjustments to Floods: A GEOGRAPHICAL APPROACH TO THE FLOOD PROBLEM IN THE UNITED STATES. Research Paper No. 29.

[46] Winsemius, H. C., Aerts, J. C. J. H., van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., Kwadijk, J. C. J., Ligtvoet, W., Lucas, P. L., van Vuuren, D. P., and Ward, P. J. (2016). Global drivers of future river flood risk. *Nature Climate Change*, 6:381–385.

[47] World Bank (2019a). Gdp per capita (constant 2010 usd). data retrieved from World Bank Open Data, https://data.worldbank.org/indicator/SP.POP.TOTL.

[48] World Bank (2019b). Population (total). data retrieved from World Bank Open Data, https://data.worldbank.org/indicator/NY.GDP.PCAP.KD.

[49] Zhang, Q. and Seto, K. C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal dmsp/ols nighttime light data. *Remote Sensing Environment*, 115:2320–2329.

[50] Zhao, M., Zhou, Y., Li, X., Cao, W., He, C., Yu, B., Li, X., Elvidge, C. D., Cheng, W., and Zhou, C. (2019). Applications of satellite remote sensing of nighttime light observations: Advances, challenges, and perspectives. *Remote Sens.*, 11:1971.

[51] Zhuo, L., Ichinose, T., Zheng, J., Chen, J., Shi, P. J., and Li., X. (2009). Modelling the population density of china at the pixel level based on dmsp/ols non-radiance-calibrated night-time light images. *International Journal of Remote Sensing*, 30(4):1003–1018.