(Looking) Back to the Future: A reconstruction of historic land use and its application for global change research

Klein Goldewijk, C.G.M.

2012

document version
Publisher's PDF, also known as Version of record

Chapter 2

Long term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1

This chapter has been published as:

Klein Goldewijk, K., Beusen, A. & Janssen, P. (2010) Long term dynamic modeling of global population and built-up area in a spatially explicit way, HYDE 3 .1, *The Holocene* 20(4), 565-573.

Abstract

This chapter describes a tool for long term global change studies; it is an update of the History Database of the Global Environment (HYDE) with estimates of some of the underlying demographic driving factors of global change. We estimate total and urban/rural population numbers, densities and fractions (incl. built-up area) for the Holocene, roughly the period 10 000 BC to AD 2000. with a spatial resolution of 5 minutes longitude/latitude. With a total global population increase from 2 to 6145 million people over that time span, resulting in a global population density increase of < 0.1 cap/km^2 to almost 46 cap/km^2 and a urban built-up area evolving from almost zero to 0.5 million km^2 (still only <0.5% of the total global land surface, but having tremendous impact through the demand of food, services, building materials, etc), it is clear that this must have had, and will continue to have profound influence on the Earth's environment and its associated (climate) change. We hope that this database can contribute to the Earth System Modelers community to gain better insight in long term global change research. Supplementary material can be found at ftp://ftp.pbl.nl/hyde/supplementary/population.

2.1. Introduction

Humans have played an important role in Earth's history by altering their surrounding landscape and by doing so also changing the composition of the atmosphere (IPCC, 2007). Settlement primarily takes place at the cost of cropland, as people historically settled in the most productive locations (e.g. Maizel et al. 1998; Klein Goldewijk, 2001, 2005). Hence, as infrastructure, settlements, towns and cities grow, the adjacent cropland is reduced to accommodate, roads and housing. Although many studies emphasize the importance of long term scientific analysis of the Earth's system (e.g. Ruddiman, 2003; Betts et al. 2006; Brovkin et al. 2006; Pongratz et al. 2008), there are very few spatially explicit datasets for such efforts readily available. In the field of demography there are many historical statistical datasets concerning (historical) population numbers. One of the leading organizations is the UN (2008) who reports total, urban and rural numbers per country for the 1950 – 2005 period, and presents scenario's up to 2050. However, none of these datasets has a spatial compound to it, which is more and more called for by global change modelers.

In order to accommodate these modelers the HYDE database (version 2) (Klein Goldewijk, 2001) was developed. It was a consistent data set of historical population and land-use data of the 20th century on a spatial resolution of 0.5 by 0.5 degree. It was freely available to the research community on a spatially explicit basis and with a satisfactory resolution for global climate change modelers. An update of HYDE 2 was presented in Klein Goldewijk (2006). HYDE 3 included several improvements compared to its predecessor: (i) the HYDE 2 version used a Boolean approach for allocation of population numbers with a 30 minute degree resolution, while HYDE 3 uses fractional land use on a 5 minute resolution; (ii) more and better sub-national (population) data (Klein Goldewijk, 2005) to improve the historical (urban and rural) population maps as one of the major driving forces for allocation of land cover; (iii) updated historical land-cover data for the period AD 1700-2000; (iv) implementation of different allocation algorithms with time-dependent weighting maps for cropland and grassland used for livestock.

To facilitate global change research we present here an updated and internally consistent revision and extension of the demographic part of the former HYDE 3.0 version of Klein Goldewijk and Van Drecht (2006), which is a dynamic modeling effort of long-term historical population growth. In the current HYDE 3.1 version, we also distinguish urban and rural numbers and compute built-up area as well, all in a spatially explicit manner on a 5 min resolution grid for the whole Holocene (10 000 BC to AD 2000). Built-up area is defined here as artificial areas contiguously occupied by humans (therefore not including vegetative land cover and water, or roads). The land use part of HYDE will be described in chapter 3.

2.2 Methodology and data

Input data for population

Population numbers
Without any doubt human population growth can be regarded as a main driving force of global change during the most recent part of the Holocene. Therefore, it is crucial to get a good insight of the demographic developments which occurred that period. Although not many comprehensive global population studies exist which cover the Holocene, a few important sources are used in this study. Historical population numbers of McEvedy & Jones (1978), Livi-Bacci (2007), and Maddison (2003), Denevan (1992) form the basis of our national historical population estimates. Supplemented with the sub-national population numbers of Populstat (Lahmeyer, 2004, pers. comm.; who provides data for several time periods varying per country), time series were constructed for each province or state of every country of the world. A summary of the population numbers used for HYDE 3.1 is presented in Table 1, and an elaborate summary table can be found in the supplementary material (ftp:/ftp.pbl.nl/hyde/supplementary). For simplicity reasons, current administrative units were kept constant over time, and every historical source was adjusted to match the current sub-national boundaries of HYDE 3 (e.g. by taking fractions of former larger empires). Country and regional totals were checked against other historical estimates (see supplementary material). Table 2 presents regional estimates for total population and population density, and Table 3 presents the resulting population growth rates per year over time for the different world regions.

Table 1. Global historical population estimates (in millions)

	10 000 BC	5000 BC	0	500	1000	1500	1600	1700	1800	1900	1950	2000
Lower literature range	1	5	170	190	253	425	498	410	890	1571	2400	6055
HYDE 3.1	2	18	188	210	295	461	554	603	989	1654	2545	6145
Upper literature range	20	24	330	210	345	540	578	680	1000	1710	2545	6145

Table 2. Total population and population density estimates, selected timesteps

	10 000 BC	5000 BC	0	500	1000	1500	1600	1700	1800	1900	1950	2000
Total population (in millions)												
North America	0	0	1	1	1	2	1	1	7	82	172	316
Latin America	0	1	12	16	24	39	9	12	20	66	168	523
Europe	1	3	29	26	32	68	86	98	149	300	399	519
Africa	0	2	15	22	41	62	73	80	86	141	223	819
CIS	0	1	9	7	7	16	19	23	46	121	177	282
Middle East	0	2	15	18	20	18	22	21	24	37	60	242
Asia	1	9	107	120	168	255	343	366	657	902	1336	3419
Oceania	0	0	0	0	0	1	1	1	1	5	11	26
World	2	18	188	210	295	461	554	603	989	1654	2545	6145
Population density (inh/km2)												
North America	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.4	4.3	9.0	16.5
Latin America	0.0	0.1	0.6	0.8	1.2	1.9	0.4	0.6	1.0	3.2	8.2	25.5
Europe	0.1	0.4	4.1	3.6	4.5	9.5	12.0	13.7	20.7	41.8	55.5	72.2
Africa	0.0	0.1	0.5	0.8	1.4	2.1	2.5	2.7	2.9	4.8	7.6	28.0
CIS	0.0	0.1	0.4	0.3	0.3	0.7	0.9	1.0	2.1	5.4	8.0	12.7
Middle East	0.0	0.3	2.4	2.9	3.3	2.9	3.5	3.5	4.0	6.1	9.8	39.6
Asia	0.0	0.4	4.9	5.5	7.7	11.7	15.8	16.8	30.2	41.5	61.4	157.2
Oceania	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.7	1.4	3.2
World	0.0	0.1	1.4	1.6	2.2	3.4	4.1	4.5	7.4	12.3	19.0	45.8

Urban fraction

Mankind started as a rural society. Just after the domestication of plants and animals people started to have sedentary agriculture and began building farms, houses and small towns/villages. There are quite few historical sources about urbanization numbers and fractions. Urban/rural fractions for all countries were derived from the U.N. after AD 1950 (UN, 2008). Earlier historical urbanization estimates for Europe were derived from (De Vries, 1984; Bairoch, 1988, Chandler, 1987), Canada after AD 1890 from Urquhart and Buckley (1965), China from Rozman (1973) and Maddison (1995), Colombia from Etter et al (2006), USA from Dodd (1993), all other countries were estimated similar to De Vries (1984), which data yielded roughly a factor 10 lower in AD 1700 than the 1950 value of the U.N. Before AD 1700 the urban fraction approaches zero in AD 1 in most countries, with a few exceptions in the older early civilized regions, such as the Mediterranean, the Levant, Central/South America and parts of India and China. Global urbanization was estimated 1% in AD 1, 2.6% in AD 1000, 3.6% in AD 1500, 19% in AD 1900 and 29% in AD 1950. It reached almost 50% in AD 2000 and at present day more people on Earth live in cities than in rural areas. Although regional differences do exist, the overall urbanization levels remained very low for a long time. Europe became relatively more urbanized during the Middle Ages, with China and India as runner up. When the Industrial Revolution was in full swing, not only Europe but also other world regions began to urbanize. Especially after AD 1900 the urbanization fractions increased rapidly. Estimates of urbanization fractions for different world regions over time are presented in the supplementary material.

Table 3. Computed regional population growth rates (in %/yr per time period)

	10 000-0 BC	0-1700	1700-1750	1750-1800	1800-1850	1850-1900	1900-1950	1950-2000	1700-1800	1800-1900	1900-2000	1700-2000
North America	0.03%	0.03%	2.21%	1.40%	2.58%	2.30%	1.49%	1.23%	1.80%	2.44%	1.36%	1.87%
Latin America	0.04%	0.00%	0.45%	0.55%	1.08%	1.32%	1.88%	2.30%	0.50%	1.20%	2.09%	1.26%
Europe	0.04%	0.07%	0.42%	0.41%	0.68%	0.73%	0.57%	0.53%	0.42%	0.71%	0.55%	0.56%
Africa	0.04%	0.10%	0.06%	0.07%	0.37%	0.63%	0.92%	2.64%	0.06%	0.50%	1.78%	0.78%
CIS	0.04%	0.05%	0.76%	0.62%	0.76%	1.19%	0.77%	0.93%	0.69%	0.98%	0.85%	0.84%
Middle East	0.04%	0.02%	0.08%	0.20%	0.31%	0.53%	0.95%	2.83%	0.14%	0.42%	1.89%	0.81%
Asia	0.05%	0.07%	0.75%	0.42%	0.38%	0.26%	0.79%	1.90%	0.59%	0.32%	1.34%	0.75%
Oceania	0.00%	0.04%	0.13%	0.24%	1.11%	2.74%	1.53%	1.69%	0.19%	1.92%	1.61%	1.24%
World	0.04%	0.07%	0.60%	0.40%	0.49%	0.54%	0.86%	1.78%	0.50%	0.52%	1.32%	0.78%

Table 4. Historical urban fraction estimate per region (in %)

	10 000 BC	5000 BC	0	500	1000	1500	1600	1700	1800	1900	1950	2000
Canada	0%	0%	0%	0%	0%	0%	1%	2%	8%	37%	61%	79%
USA	0%	0%	0%	0%	0%	0%	1%	2%	6%	40%	64%	79%
Mexico	0%	0%	0%	0%	0%	0%	1%	2%	6%	22%	43%	75%
Rest Central America	0%	0%	0%	0%	0%	0%	1%	3%	7%	20%	34%	57%
Brazil	0%	0%	0%	0%	0%	0%	1%	4%	9%	23%	36%	81%
Rest South America	0%	0%	0%	0%	0%	0%	1%	4%	10%	30%	49%	78%
Northern Africa	0%	0%	1%	1%	1%	2%	2%	3%	7%	18%	29%	51%
Western Africa	0%	0%	0%	0%	0%	0%	0%	1%	2%	6%	11%	38%
Eastern Africa	0%	0%	0%	0%	0%	0%	0%	0%	1%	3%	5%	18%
Southern Africa	0%	0%	0%	0%	0%	0%	0%	1%	2%	7%	21%	48%
Western Europe	0%	0%	2%	1%	3%	11%	12%	13%	21%	41%	61%	75%
Central Europe	0%	0%	0%	0%	0%	3%	4%	3%	4%	19%	35%	60%
Turkey	0%	0%	1%	1%	2%	2%	2%	2%	6%	16%	25%	65%
Ukraine +	0%	0%	0%	0%	0%	0%	0%	1%	1%	10%	33%	66%
Asia-Stan	0%	0%	0%	0%	0%	0%	0%	1%	2%	12%	34%	42%
Russia +	0%	0%	0%	0%	0%	0%	1%	2%	2%	14%	44%	71%
Middle East	0%	0%	1%	1%	1%	1%	2%	2%	6%	16%	28%	64%
India +	0%	0%	2%	3%	3%	4%	5%	5%	6%	10%	16%	27%
Korea	0%	0%	0%	0%	0%	0%	1%	2%	5%	14%	25%	73%
China +	0%	0%	1%	3%	5%	6%	7%	6%	6%	7%	13%	37%
Southeastern Asia	0%	0%	0%	0%	0%	0%	0%	2%	4%	10%	18%	38%
Indonesia +	0%	0%	0%	0%	0%	0%	0%	1%	3%	8%	12%	41%
Japan	0%	0%	0%	0%	3%	3%	4%	5%	5%	21%	35%	65%
Oceania	0%	0%	0%	0%	0%	0%	0%	1%	8%	35%	72%	82%
Greenland	0%	0%	0%	0%	0%	0%	0%	1%	3%	20%	49%	82%
Antarctica	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
World	0%	0%	1%	2%	3%	4%	5%	5%	7%	16%	29%	47%

Built-up area in AD 2000.

Infrastructure and urban development is increasing rapidly during the last decades (UN, 2008), but illustrative is the fact that even for the present day there is no clear picture of the built-up area in the world. Potere and Schneider (2007) compared six spatially explicit studies which reported global estimates for built-up area. Globally, estimates for the extent of built-up areas in 2000 range from 0.2% - 2.7% of the total land area, with 5 of the 7 estimates below the 0.5%. Most of the differences can be explained by various definitions of built-up area, and differences between satellite derived and inventory based data. All these percentages correspond to ca. 0.3 - 3.5 million km^2 worldwide of land which is at first sight not available for producing food. For comparison, this study estimates a global built-up area percentage of ca.0.3% (0.5 million km^2) in AD 2000.

The maps we use for current built-up areas were created by first combining 'Urban and Built-up' from the DISCover map (Loveland et al., 2000) with the areas of 'Artificial surfaces and associated areas' from GLC2000 (Bartholome et al., 2002). Subsequently, an overlay of this built-up map was made with the Landscan population counts map (Landscan, 2006) to calculate the average population density within built-up areas for each country for the year 2000. The LandScan^{TM} Dataset comprises a worldwide population database compiled on a 30" X 30" latitude/longitude grid. Census counts (at sub-national level) were apportioned to each grid cell based on likelihood coefficients, which are based on proximity to roads, slope, land cover, nighttime lights, and other information. For some countries we took the urban densities from Demographia (2006) as a maximum, to avoid unrealistic (i.e. too high) values for national urban population densities (e.g. Macau, Hong Kong, due to a mismatch in our GIS system of area and total numbers). Since there are hardly any reliable estimates around for historical built-up area, we decided to use historical urban densities as a proxy instead.

Urban density
Also here, few sources of historical urban densities can be found (De Vries, 1984; Baroich, 1976; Demographia, 2008). Supplementary Figure SF1 depicts the urban density over time of some European and North American cities. The figure shows that for cities the urban density has not been constant over time, but always seems first to increase rapidly to very high densities (up to 40 000 inh/km^2) and then, when the standard of living improves, decreasing more slowly over time (compared with its original increase). Income, cultural and human behavior aspects and planning policy play a role here. We therefore assume that the (historical) urban densities follow an asymmetric bell-shaped curve.

Since the shape of the curve is derived from very few available city data, a major assumption is that we believe that they are representative for a whole country and indeed for all countries in the world. Of course this introduces uncertainty, but in the light of this data base (and its intended rather 'crude' use in integrated models of global change) we think this is acceptable. Only the size and the shape of the curve differ between the countries, depending on their development stage in time, see supplementary material figure SF2..

We developed a method to determine the urban population density over time for each country. The asymptotes of the curve (*L* and *K*) are defined as a fraction of the maximum of the curve (*D*max). The parameter *c*, which indicates the ratio between the decrease rate and the increase rate of the curve, is also estimated globally. To determine the final shape of the curve for each country, we need two points on the curve. The first point is derived from an overlay of the Landscan population density maps and satellite imagery for the year 2000. The second point needed is the time and magnitude where the curve reaches is maximum (*D*max and *t*max). We assume that this maximum is reached when the increase in urban population is slowing down for the first time in history. The full methodology is described in the supplementary material Box 1.

We can now compute the built-up area for each country by dividing the total urban population numbers with the time dependent urban densities for each country. Supplementary material figure SF3 presents the scheme to calculate the population numbers, densities and urban areas on a spatially explicit way

2.3 Results

Livi-Bacci (2007) stated that during the Neolithic era – Stone/Iron/Bronze Age era – 'demographic growth took place in varying degrees of intensity and large strategic space, and population numbers were very low and susceptible to climate fluctuations, environmental constraints and warfare'. Shennan and Edinborough (2007) also present evidence that for the late Neolithic period dramatic rises in population were associated with the arrival of farming in central and northern Europe. These increases happened not gradually over time and space, but show fluctuations and large differences per region. However, overall figures remain low for a long time. We have constructed historical maps of urban, rural and total population totals and densities, as well as built-up areas for a 12 000 year period, on a 5 min × 5 min grid resolution. The HYDE 3.1 estimate for 10 000 BC of 2 million is well within the range found in literature (between 1 and 20 with most estimates below 6 million, see Table 1 and tables S1, S2 and S3). We computed a growth rate for the 10 000 BC–AD 1 period of 0.04% per year for that period. This is comparable with Livi-Bacci

(2007) who provides a total population number of 6 million in 10 000 BC. We also estimate the global population at 18 million (literature range 5–24 million) in 5000 BC. Kropelin *et al*. (2008) suggested that a relative small shift in climate patterns in sub-Saharan Africa, resulting from a shift in monsoon regimes, resulted in considerable changes in savanna type biome patterns. It is not clear whether such climate changes triggered large migrations of people (e.g. the great Bantu migration from 1000 BC to AD 500) but it certainly played a role. According to Verschuren *et al*. (2000) such climate changes in Eastern Africa were 'evidence for drought-induced famine, political unrest, and large-scale migration of indigenous peoples' during the ad 1300–1990 period. The scale and magnitude of those fluctuations are now well known, but our estimate of 41 million for Africa in AD 1000, 62 million in ad 1500 and 80 million in AD 1700 is well in range with the literature (see supplementary table ST3).

After the rise and fall of the Greek and Roman Empires, population growth remained low and fluctuated for quite some centuries. Europe gradually faded into the Dark Middle Ages where technological developments almost came to a halt, not being helped by the invasions of the Barbarians, the Huns, and the Mongols. Also, large scale pandemics such as the Black Plague reduced population numbers severely in many parts of the old world. This decimation of the population led to large scale abandonment of agricultural land and subsequently to a gain of forest land e.g. example of Germany after the bubonic plague in early 15th century (Bork et al.,1998).

However, this was not the case in China, especially in the eastern and central eastern parts, where ancient cultivation techniques were perfected to sustain relative high population densities. Here, development did not spread widely because of drought periods, mismanagement and internal warfare and the fact that China became an increasingly inward-looking Empire. We adopted the numbers of McEvedy and Jones (1978) and Liu and Hwang (1979) who show fluctuations between 60 and 160 million from AD 0 until AD 1600.

In contrast to Europe, the Middle Ages and Late Middle Ages (AD 500–1600) were the peak of the Central American civilizations (e.g. the Maya, Aztecs, Inca), with evidence of very densely populated regions, supported by a range of agricultural activities and elaborate trade routes (e.g. Culbert, 1988; DeMenocal, 2001; Etter and van Wyngaarden, 2000). Although the civilizations were highly successful in sustaining relatively high population densities, studies reveal that they were already susceptible to climate changes (DeMenocal, 2001). The Americas experienced further large fluctuations in indigenous population numbers because of pandemics accompying European conquest, and subsequent changes in agricultural practices such as reduction of biomass burning which resulted in significant changes in atmospheric carbon dioxide (Nevle and Bird, 2008). All this is reflected in global population numbers. We estimate the global population at 210 million in AD 500, and at 295 million in AD 1000. Until AD 1400 numbers remained below the 400 million mark, and at the end of the Dark Ages population growth gained momentum again. We estimate global population numbers increased to 555 million in AD 1600, this is well in range with other literature estimates which range between 545 and 578 million, see supplementary table ST1 and supplementary figure SF4.

The decisive increase in world population took place after ad 1600. The start of the Industrial Revolution resulted in the colonization by Europeans of the Americas (e.g. Waisanen and

Bliss, 2002), Australia, parts of Asia and Africa (Houghton, 1999). This was accompanied by a rigorous agricultural expansion, first in the temperate regions, later in the tropics as well. In AD 1800 we estimate that the global population reached the 1 billion point, 1658 million in AD 1900 and 2520 million in AD 1950. Population numbers really exploded after the Second World War to 3681 million in AD 1970 and 6096 million in AD 2000 (UN, 2008a).

Population density.
Evidently, the increase of total population is also reflected in population densities. For thousands of years population densities remained very low (< 1 cap/km^2), reaching an ample 4–5 cap/km^2 around ad 1 in Europe and Asia, and just over 2 cap/km^2 in the Middle East region. The average global population density was estimated at 1.4 person per km^2. The population densities for Europe and Asia remained low for many centuries, reaching 10–12 cap/km^2 around ad 1500, while the other world regions were even less than 3 cap/km^2. It was not until the nineteenth century that densities increased towards 40 cap/km^2 in Europe and Asia, with the latter really taking off after the Second World War towards almost an average of 160 cap/km^2, leaving Europe behind with 72 cap/km^2. Although Africa, the Americas and Middle East have witnessed substantial increases in total population at the end of the twentieth century, the population densities are still less than 40 cap/km^2. See Table 2 and supplementary material for the regional population densities. Figures 1 and 2 present the spatially explicit time series of historical population densities of HYDE 3.1 for the Holocene.

Urban/built-up area
For a long time period the size of many cities in large parts of the world remained rather small. In Europe most cities during the Middle Ages were dependent in size upon their walls and other defensive mechanisms (e.g. canals, rivers, fences), so that many towns hardly could grow. Only in times of the Industrial Revolution, with its colonization of large parts of the world, trade and the emergence of a global economy, generated enough welfare in the Old World to boost the local economy as well. This attracted people from outside the towns and cities who had to be housed so many cities made the 'jump' across the city walls and people started to build houses outside the old historic boundaries.

Supplementary material figure SF5 depicts the non-linearity of city growth. The figure shows the total urban population divided by the built-up area (= urban density) per region, and it clearly shows the bell-shaped curve, not only for a given country (as one would expect since its input) but even for aggregated regional results. It is interesting to see that some regions are already in the 'decline' phase of urban density (e.g. USA, Europe), while other regions are still in the increasing phase (most developing regions).

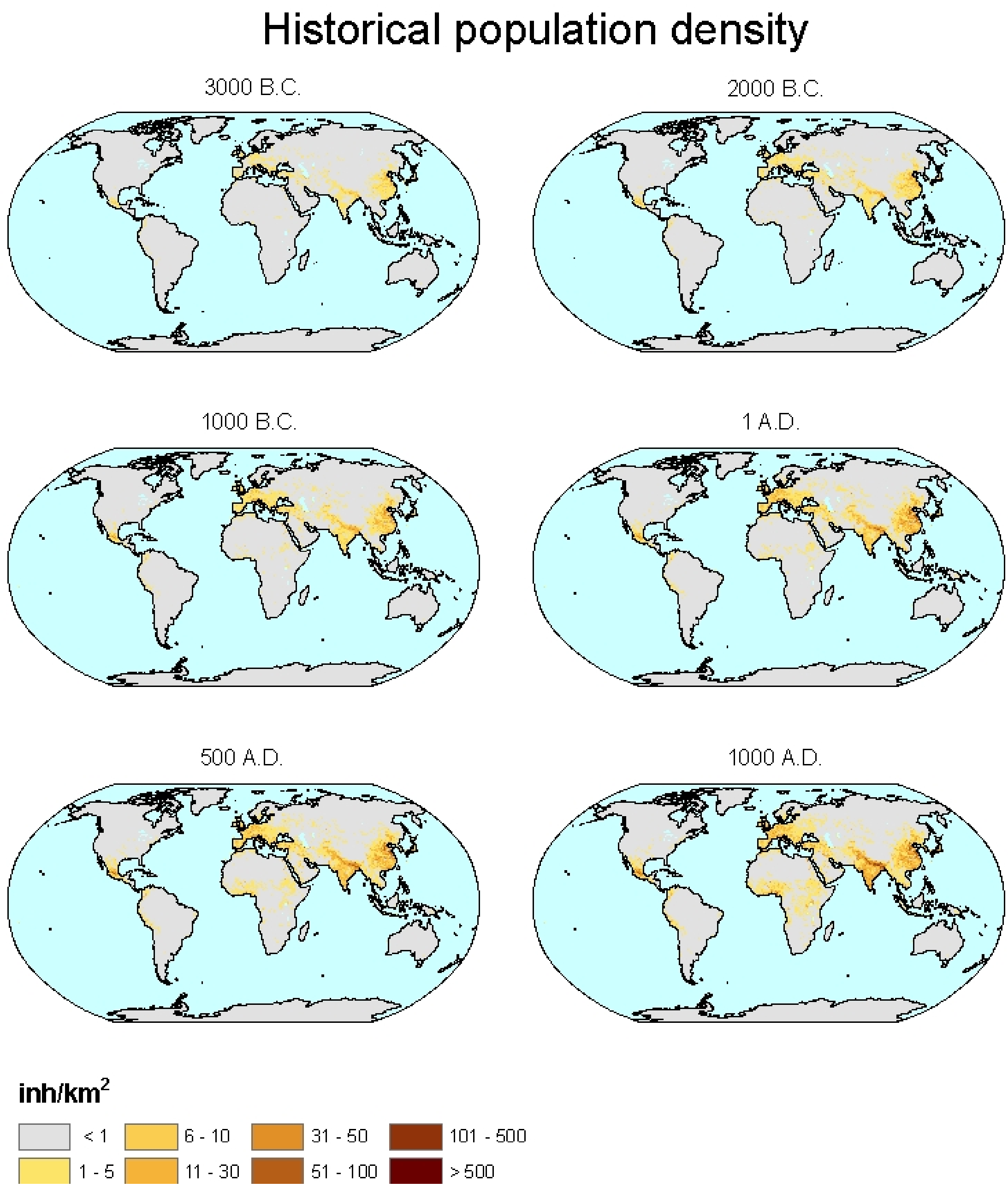

Figure 1. Global maps of population density, selected time steps.

Our estimate for AD 2000 of ca. 535 300 km^2 is well in line with other estimates. Potere and Schneider (2007) presented an overview of six different global estimates of urban built-up area, ranging from 276 000 km^2 to 3 524 000 km^2 (the latter being an extreme, five out of six studies were less than 730 000 km^2). For AD 1900 we computed roughly 47 000 km^2 (± 2300 km^2 uncertainty), for AD 1800 almost 16 000 km^2 (± 2400 km^2) and in AD 1700 it was a merely 13 600 km^2 (± 3500 km^2).Around AD1 we estimate it an ample 4 000 km^2 (± 3200 km^2, see also table 6).

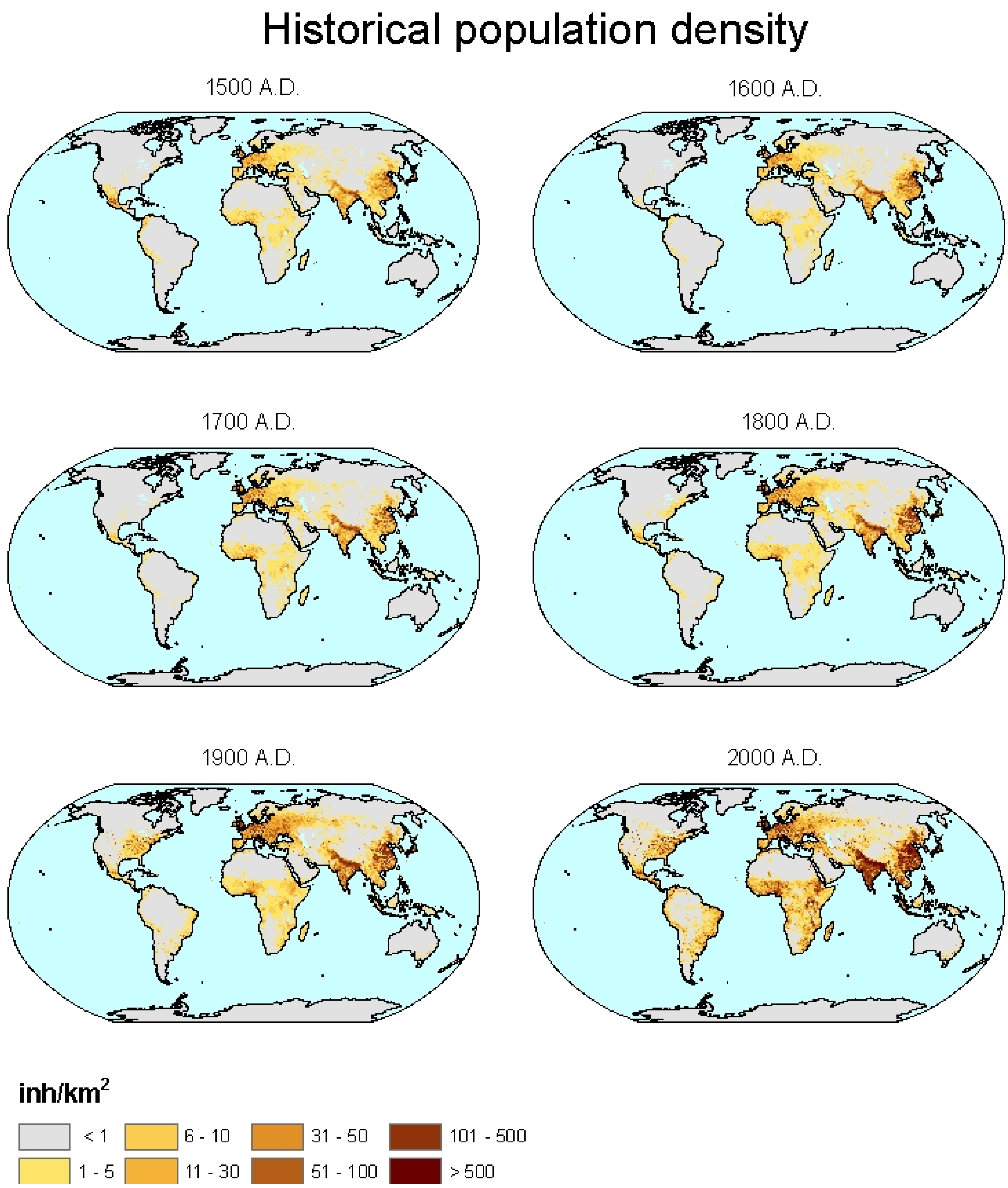

Figure 1. continued. Global maps of population density, selected time steps.

Uncertainties

Obviously, there are large and many uncertainties attached to hind-cast attempts such as this study. We leaned heavily on historical population sources such as McEvedy and Jones (1978) and Livi-Bacci (2007) and especially for the pre-1700 period the numbers have to be treated with care. Although their studies have been reviewed extensively, they always remain ‘guestimates’ or ‘educated guesses’ for the greater part. This is something we have to be aware of and simply have to live with. However, when looking at our results and at the resulting growth rates they seem not to be unfair and rather acceptable as a reasonable reconstruction of historical population trends. As tables S1 and S3 points out, our estimates are well in range with those found in literature.

The data used for determining the urban densities are derived from cities. Our approach has sub-national regions as a spatial scale and therefore our urban population densities are in general lower than in most larger cities. It is clear that our approach is not applicable on a city level, but on a country level it seems a fair estimate.

Potere and Schneider (2007) concluded that even for the present day satellite derived built-up areas do not match exactly the national statistics for almost every country, and since our method reproduces the 2000 satellite derived urban built-up area, this already introduces uncertainty, up to 10-20%. We recognize this uncertainty but unless statistics and satellite information are not better geared to each other, we have to accept the difference.

Table 5. Uncertainty range of the global historical population numbers (in millions)

	10 000 BC	5000 BC	0	500	1000	1500	1600	1700	1800	1900	1950	2000
Lower	0	2	47	84	161	319	399	452	841	1572	2469	6083
Medium	2	18	188	210	295	461	554	603	990	1654	2545	6145
Upper	5	34	329	337	430	604	709	754	1138	1737	2621	6206

Table 6. Total computed urban area per world region (in km^2)

	10 000 BC	5000 BC	0	500	1000	1500	1600	1700	1800	1900	1950	2000
Canada	0	0	0	0	0	0	0	1	7	193	1126	6030
USA	0	0	0	0	0	4	4	15	139	10766	46336	156919
Mexico	0	0	0	0	0	102	11	112	308	669	1884	10520
Rest Central America	0	0	0	0	0	20	41	129	200	596	1927	8383
Brazil	0	0	0	0	0	9	22	72	163	843	3145	22168
Rest South America	0	0	0	0	1	122	125	367	311	1029	3579	16232
Northern Africa	0	5	295	231	239	150	193	170	158	560	1383	13038
Western Africa	0	0	0	0	1	120	362	1096	978	565	1235	13680
Eastern Africa	0	0	0	0	0	8	24	69	73	181	395	5853
Southern Africa	0	0	0	0	0	8	24	79	66	257	1288	9108
Western Europe	0	2	342	230	619	2072	3070	2864	3905	11438	30307	83617
Central Europe	0	0	0	0	10	1657	2286	1419	1403	4182	6495	17088
Turkey	0	0	371	250	220	120	140	134	165	292	663	5203
Ukraine +	0	0	0	0	0	2	6	22	33	691	3096	8126
Asia-Stan	0	0	0	0	0	3	10	29	40	373	1499	5791
Russia +	0	0	0	0	0	19	64	189	203	2416	9177	21064
Middle East	0	99	977	861	666	355	421	417	521	765	1507	15497
India +	0	0	931	1012	1110	942	1154	1419	1729	2615	4953	23854
Korea	0	0	0	0	0	4	13	43	81	144	420	2919
China +	0	52	1310	2300	2861	4013	5762	3916	3923	2846	7005	43199
Southeastern Asia	0	0	0	0	0	24	83	290	398	877	2309	13989
Indonesia +	0	0	0	0	0	11	34	101	134	479	1181	8772
Japan	0	0	12	80	445	468	747	1104	504	1887	5049	13450
Oceania	0	0	0	0	0	3	14	49	801	2035	4433	13871
Greenland	0	0	0	0	0	0	0	0	0	0	0	24
World	0	158	4237	4964	6172	10238	14609	14104	16243	46697	140391	538395

We tried to quantify the uncertainty in the total population estimates by introducing a ‘lower’ and ‘upper’ range beside the HYDE 3.1 estimate, based on different estimates found in the literature. These estimates yield some sort of increasing uncertainty range when going back in time. We assume the uncertainty to be ±1% in AD 2000, ±5% in AD 1900, ±25% in AD 1700, ±45% in AD 1000, ±75% in AD 1 and ±100% in 10 000 BC. Since we chose the high end of the literature estimates, the minimum and maximum results can be regarded as extremes, and the HYDE 3.1 as a reasonable scenario for historical population developments. Table 5 presents the three variants from this study for historical population estimates. As Shennan and Edinborough (2007) and Verschuren *et al.* (2000) pointed out, climate shifts probably have had impact on the spatial patterns of human settlement during the early Holocene. Since we do not use climate as an allocation proxy we therefore could have missed the resulting fluctuations and migrations of people in the past. However, in the light of the sheer magnitude of these fluctuations on a millennial timescale, we feel that it is acceptable to use our allocation by weighing maps approach because the long-term development of total number of people is the most important driving force of global environmental change.

2.4 Concluding remarks

To analyze the effects of anthropogenic activities on long-term global environmental change, the Earth System Modelers community has long lacked a data set that is both temporally and spatially explicit. HYDE 3.1, presented here, fills this gap, providing a data set of global human population and its spatial distribution over the Holocene (10 000 BC to AD 2000). It is designed to serve as an essential tool for global change studies.

Acknowledgements
This research was performed with the support of the Dutch Ministry of Housing, Spatial Planning and the Environment, and is endorsed by the Global Land Project (GLP), a joint research project for land systems for the International Geosphere-Biosphere Programme (IGBP) and the International Human Dimensions Programme (IHDP).

References

Bairoch, P., Batou, J., Chevre, P. (1988). *The population of European cities.Data bank and short summary of results: 800–1850*. Geneva: Librairie Droz, xix, 336 pp. Publications du Centre d'Histoire Economique Internationale de l'Universite de Geneve No. 2.

Bartholome, E., Belward, AS. (2005). GLC2000: a new approach to global land cover mapping from earth observation data. *International Journal of Remote Sensing* 26: 1959–77.

Betts, RA. (2006). Forcings and feedbacks by land ecosystem changes on climate change. *Journal De Physique IV: JP*, 119–42.

Biraben, J-N. (1980). An Essay Concerning Mankind's Evolution, *Population*, Selected Papers, December, Table 2.

Bork H.-R., Bork, H., Dalchow C., Faust B., Piorr H.-P. and Schatz Th. (1998). Landschaftsentwicklung in Mitteleuropa. Wirkungen des Menschen auf Landschaften, 328 pp., Gotha (Klett-Perthes).

Braudel 1981. *Civilization and Capitalism, 15th-18th Century*, New York: Harper and Row, 1981-84, 3 vols.

Brovkin V., Clausen M., Driesschaert E., Fichefet T., Kicklighter D., Loutre M.F., Matthews H.D., Ramankutty N., Schaeffer M. and Sokolov A. (2006). Biogeophysical effects of historical land cover changes simulated by six earth system models of intermediate complexity. *Climate Dynamics*, 26,587-600.

Chandler, T. (1987). *Four thousand years of urban growth: an historical census*, Lewiston, New York, St. David's University Press, 1987. xiii, 656 p.

Culbert, T.P. (1988). The collapse of the Classic Maya Civilization. *In* Yoffee, N. and Cowgill, G.L. (eds.), *The collapse of ancient states and civilizations,* p. 69-101. University of Arizona Press, Tuscon.

De Vries, J .(1984) *European urbanization 1500–1800*. Cambridge MA: Cambridge University Press.

deMenocal, P.B. (2001). Cultural responses to climate change during the Late Holocene, *Science* 292: 667-673.

Demographia (2006). Global built-ups and urbanization, Wendell Cox Consultancy, http://www.demographia.com (visited July 2006).

Denevan, W.M. (1992). *The pristine myth: The landscape of the Americas in 1492*, American Association of Geographers.

Dodd, D.B. 1993, *Historical statistics of the states of the United States: two centuries of the census, 1790-1990*, Westport, Connecticut/London, England, Greenwood Press, 1993. viii, 478 p.

Etter, A, C McAlpine, D Pullar, H Possingham (2006). Modelling the conversion of Colombian lowland ecosystems since 1940: Drivers, patterns and rates, *Journal of environmental management*, 114 (2-4), 369-386.

Etter and van Wyngaarden (2000). Patterns of landscape transformation in Colombia with emphasis in the Andean region, *Ambio* **29** (7) 432–439.

Houghton, RA (1999) The annual net flux of carbon to the atmosphere from changes in land use 1850–1990. *Tellus, Series B: Chemical and Physical Meteorology* 51: 298–313.

Intergovernmental Panel on Climate Change (2007) *Climate change, the scientific basis*. Cambridge: Cambridge University Press.

Klein Goldewijk, K (2001) Estimating global land use change over the past 300 years: the HYDE database. *Global BiogeochemicalCycles* 15: 417–33.

Klein Goldewijk, K (2005) Three centuries of global population growth: a spatial referenced population (density) database for 1700–2000. *Population and Environment* 26: 343–67.

Klein Goldewijk, K, van Drecht, G (2006) HYDE 3: current and historical population and land cover. In: Bouwman, AF, Kram, T, Klein Goldewijk, K (eds) *Integrated modelling of global environmental change. An overview of IMAGE 2.4*. Bilthoven: Netherlands Environmental Assessment Agency.

Kropelin, S, Verschuren, D, Lezine, AM, Eggermont, H, Cocquyt, C, Francus, P, Cazet, JP, Fagot, M, Rumes, B, Russell, JM, Darius, F, Conley, DJ, Schuster, M, Von Suchodoletz, H, Engstrom, DR (2008) Climate-driven ecosystem succession in the Sahara: the past 6000 years. *Science* 320: 765–68.

Lahmeyer, J (2004) Populstat database, Growth of the population per country in a historical perspective, including their administrative divisions and principal towns. Available at: http://www.populstat.info/

Landscan (2006) *Landscan global population database, the 2004 revision*. Oak Ridge National Laboratory. Available at: http://www.ornl.gov/landscan

Liu, PKC, Hwang, KS (1979) Population change and economic development in mainland China since 1400. In: Hou, CM, Yu, TS (eds) *Modern Chinese economic history*. Taipei: Academica Sinica.

Livi-Bacci, M (2007) *A concise history of world population*. Fourth edition. Oxford: Blackwell Publishing.

Loveland, TR, Reed, BC, Brown, JF, Ohlen, DO, Zhu, Z, Yang, L, Merchant, JW (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing* 21: 1303–30.

McEvedy, C, Jones, R (1978) *World atlas of population history*. Hammondsworth: Penguin Books Ltd.

Maddison, A (1995) *Monitoring the world economy*. Paris: OECD.

Maddison, A (2001) *The world economy: a millennial perspective*. Paris: OECD.

Maizel, M (1998) *Historical interrelationships between population settlement and farmland in the conterminous United States, 1790 to 1990*. Biological Science Report USGS/BRD/BSR 1998-0003, U.S. Geological Survey, Biological Resources Division.

Mitchell, BR (1993a) *International historical statistics, Africa, Asia & Oceania: 1750–1993, third edition.* Indianapolis: MacMillan.

Mitchell, BR (1993b) *International historical statistics, the Americas: 1750–1988*. Indianapolis: MacMillan.

Mitchell, BR (1998) *International historical statistics, Europe: 1750– 1993, fourth edition.* Indianapolis: MacMillan.

Nevle, RJ, Bird, DK (2008) Effects of syn-pandemic fire reduction and reforestation in the tropical Americas on atmospheric CO2 During European conquest. *Palaeogeography, Palaeoclimatology,Palaeoecology* 264: 25–38.

Pongratz, J, Reick, C, Raddatz, T, Claussen, M (2008) A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles* 22: 1–16.

Potere, D, Schneider, A (2007) A critical look at representations of urban areas in global maps. *GeoJournal* 69: 55–80.

Rozman, G (1973) *Urban networks in Ch'ing China and Tokugawa Japan*. Princeton NJ: Princeton University Press.
Ruddiman, WF (2003) The anthropogenic greenhouse era began thousands of years ago. *Climatic Change* 61: 261–93.
Shennan, S, Edinborough, K (2007) Prehistoric population history: from the Late Glacial to the Late Neolithic in Central and Northern Europe. *Journal of Archaealogical Science* 34: 1339–45.
United Nations (2008a) *World population prospects, the 2007 revision*. New York: United Nations Population Division.
United Nations (2008b) *World urbanization prospects, the 2007 revision*. New York: United Nations Population Division.
Urquhart, M, Buckley, K (1965) *Historical statistics of Canada*. University Press.
Verschuren, D, Lalrd, KR, Cumming, BF (2000) Rainfall and drought in equatorial east Africa during the past 1,100 years. *Nature* 403: 410–14.
Waisanen, PJ, Bliss, NB (2002) Changes in population and agricultural land in conterminous United States counties, 1790 to 1997. *Global Biogeochemical Cycles* 16: 84–91.

Supplementary material Box 1 (see also ftp:/ftp.pbl.nl/hyde/supplementary).

Calculation of the urban population density and the urban built-up area

We have assumed that the form of the curve which describes the urban population density as a function of time, resembles a bell-shaped like curve (Gaussian probability density function), with a certain asymmetry and different asymptotic behavior in the past and future. (see Figure 2 article). It reflects the idea that the urban population density will increase in the beginning over time. After a certain point the urban population density will decrease, but slower and the urban population density will converge to a constant after a long time. This is expressed in the following form, which adapts and generalizes the Gaussian probability density functional form:

$$D_{urb}(t) = \begin{cases} f_1 \exp\left(-0.5\left(\dfrac{t-t_{\max}}{\sigma}\right)^2\right) + L \ \ for \quad t < t_{\max} \\ f_2 \exp\left(-0.5\left(\dfrac{t-t_{\max}}{c\sigma}\right)^2\right) + K \ \ for \quad t \geq t_{\max} \end{cases}$$

with:
$D_{urb}(t)$: the urban population density on time t (number of inhabitants per km^2)
t_{max}: time when urban population density has reached its maximum
c: indicates how slow the second curve will decrease relatively to the first curve. c is greater than 1.
f_1 and f_2 are factors that determine the maximum height of the curve.
L: the minimum height of the first curve (left asymptote)
K: the minimum height of the second curve (right asymptote)
t: time
σ: determines the width of the bell curve.

We define Land K as a fraction of D_{max} : $L = b * D_{max}$ and $K = d * D_{max}$

We used b = 0.2, c = 1.5 and d = 0.5 for each country.

To determine the shape of the curve for each country, we need two points on the curve. The first point is derived from an overlay of the Landscan population density maps and satellite imagery for the year 2000 and is noted as (t_{2000},D_{2000}). The second point needed is the time and magnitude where the curve reaches is maximum (D_{max} and t_{max}). We assume that this maximum is reached when the increase in urban population is slowing down for the first time in history.

Using the unknown point (t_{max},D_{max}) in both curves leads to the values of f_1 and f_2, which in fact guarantee that the left part ($t<t_{max}$) of the curve continuously connects to the right part ($t>t_{max}$) of the curve

$f_1 = D_{max} – L$ and $f_2 = D_{max} – K$.

Using the known point (t_{2000},D_{2000}) will give the value of σ:

For $t_{2000} < t_{max}$: $\sigma = \frac{(t_{max} - t_{2000})}{\sqrt{-2\ln\left(\frac{D_{2000} - L}{D_{max} - L}\right)}}$

For $t_{2000} > t_{max}$: $\sigma = \frac{(t_{2000} - t_{max})}{c * \sqrt{-2\ln\left(\frac{D_{2000} - K}{D_{max} - K}\right)}}$

Basic assumptions for these functions must be fulfilled to avoid a negative argument of the natural logaritme or square root:

1. D_{2000} must be greater than K for $t_{2000} > t_{max}$
2. D_{2000} must be greater than L for $t_{2000} < t_{max}$
3. D_{2000} must be smaller than D_{max}.

Determination of the point (t_{max},D_{max}) **for a region.**

Given information for each region:

1. population numbers over time
2. urban population fraction over time
3. urban built-up area for the year 2000. From this the urban population density is determined and the point (t_{2000},D_{2000}) is known.

The point (t_{max},D_{max}) is not known for each region.

To find the time t_{max} where the curve has reached is maximum D_{max} we apply the following rules:

1. t_{max} has a default value of 2100. Most of the population scenarios are known to the year 2100.
2. The maximum D_{max} is not reached when the urban population fraction is below 20%.
3. The maximum is reached in the year that the growth of the urban population starts to decrease. The growth of the urban population will be determined on a period of 10 years after 1700. Before 1700 we take a time period of 100 years. Two following time points must be smaller than the maximum D_{max}, to set t_{max}.
4. It is important that (t_{max},D_{max}) is different from (t_{2000},D_{2000}), otherwise the bell curve is not specified. We set t_{max} on 2005 when t_{max} is almost 2000 (1998, 1999, 2000, 2001, 2002)

For D_{max} the following rules are applied:

1. Default value for D_{max} is 0.1% percent of the total land area of the region. Between the year 1900 and 2000, we assume an increase of the default value of 0.001% a year till a maximum of 0.2% of the total land area of the region in the year 2000. After 2000 the default value for D_{max} is 0.1% percent of the total land area of the region.
2. Maximum value for D_{max} = 30000 cap/km^2. This is a high value. Some cities have a maximum of 40000 cap/km^2. We took a lower value because we only have regions which are mixed, some rural and some urban areas.
3. When t_{max} is close to t_{2000} we assume that D_{max} is close to D_{2000}. To realize this we define that t_{max} is close to t_{2000} when the difference abs(t_{max}-t_{2000}) is smaller than 50 years. In that case the difference between D_{max} and D_{2000} must be smaller than D_{2000} * a correction factor F. where F is calculated as 0.2 * abs(t_{max}-t_{2000})/50..When t_{max} and t_{2000} are not close, F will be given a value of 0.2.

The first three steps give a first estimate of D_{max}, but there are more assumptions that must be fulfilled (see calculation of σ):

4. If K > D_{2000} then D_{max} = (1.0 - F) * D_{2000} / d for $t_{2000} > t_{max}$
5. If L > D_{2000} then D_{max} = (1.0 - F) * D_{2000} / b for $t_{2000} < t_{max}$
6. If $D_{max} < D_{2000}$ then D_{max} = (1.0 + F) * D_{2000}

Now we can calculate the urban built-up area for each point in time. For some countries, there occurs a temporary decrease of urban built-up area (U_{area}(t)). We tried to avoid this phenomenon by performing a iterative method to adjust D_{max}:

7. If $U_{area}(t-1) > U_{area}(t)$ and $U_{pop}(t-1) < U_{pop}(t)$ then D_{max} is adjusted with an iterative method, in order to decrease the discrepancy between the rate of increase in urban population and the urban density (sometimes U_{area} decreases which is not logical). Since $U_{area}(t) = U_{pop}(t)/U_{dens}(t)$, D_{max} is decreased in each iteration step with 1% of the orginal difference between D_{2000} and D_{max}, until the calculated built-up (U_{area}(t)) does not decrease during the phase of continuous historical urban population increase.
8. For each time urban population density must be larger than 100 cap/km^2. This avoids large urban built-up area in combination with a low urban population numbers.