Supporting Information

Rapid rise in urban sprawl: Global hotspots and trends since 1990

Martin Behnisch[1], Tobias Krüger[1], Jochen A. G. Jaeger[2]

[1] Leibniz Institute of Ecological Urban and Regional Development (IOER), Dresden, Germany.

[2] Department of Geography, Planning and Environment, Concordia University Montreal, Montréal, Québec, Canada.

In: PLOS Sustainability and Transformation (2022)

Contents

A. Definitions of urban sprawl

Many definitions of urban sprawl have been proposed in the literature [1–4]. However, there is still no single agreed upon definition of urban sprawl. Urban sprawl has been vaguely described as “the uncontrolled spread of towns and villages into undeveloped areas” [5] and as “the physical pattern of low-density expansion of large urban areas, under market conditions, mainly into the surrounding agricultural areas” [6].

„Sprawl is sometimes described as the scattering of urban settlement over the rural landscape” ([7] p. 475). It is also characterized as the combination of three features, which are “(1) leapfrog or scattered development, (2) commercial strip development, and (3) large expanses of low-density or single-use developments – as well as by such indicators as low accessibility and lack of functional open space” ([8] p. 108). Sprawl is “a particular type of suburban development characterized by very low-density settlements, both residential and non-residential, dominance of movement by use of private automobiles, unlimited outward expansion of new subdivisions and leap-frog development of these subdivisions, and segregation of land uses by activity” ([9] p. 33). Sprawl means “low-density development beyond the edge of service and employment, which separates where people live from where they shop, work, recreate and educate − thus requiring cars to move between zones” ([10] p.1). Urban Sprawl has also been described as “the process in which the spread of development across the landscape far outpaces population growth. The landscape sprawl creates has four dimensions: a population that is widely dispersed in low-density development, rigidly separated homes, shops, and workplaces, a network of roads marked by huge blocks and poor access, and a lack of well-defined, thriving activity centers, such as downtowns and town centers. Most of the other features usually associated with sprawl − the lack of transportation choices, relative uniformity of housing options or the difficulty of walking − are a result of these conditions” [11]. “Sprawl is low-density, leapfrog development characterized by unlimited outward extension. In other words, sprawl is significant residential or non-residential development in a relatively pristine setting. In nearly every instance, this development is low density, it has leaped over other development to become established in an outlying area, and its very location indicates that it is unbounded” [12]. “Ultimately, what distinguishes sprawl from alternative development patterns is poor accessibility of related land uses to one another... Another characteristic common to all sprawl archetypes is a paucity of functional open space” [13].

Such descriptions are not explicit enough to directly create a corresponding quantitative measure of urban sprawl. As a consequence, different authors have focused on different aspects that are part of various definitions of urban sprawl and selected their own measures to quantify these aspects (e.g., [14]). Thus, the diverse aspects of urban sprawl and the variety of existing definitions of urban sprawl can pose a challenge to its quantification.

A systematic assessment of existing definitions of urban sprawl reveled that a large majority of definitions have three elements in common: 1. the total amount of built-up areas in a given landscape, 2. the spatial dispersion of built-up areas, i.e., how closely clumped or widely scattered the buildings and patches of built-up areas are within the landscape, 3. low-density development (i.e., high land uptake per person) [15]. Taking these most important common characteristics into account, the measurement of global urban sprawl in this article is based on the following definition: “A landscape is affected by urban sprawl if it is permeated by urban development or solitary buildings and when land uptake per inhabitant or workplace is high. The more area built over in a given landscape (amount of built-up area) and the more dispersed this built-up area in the landscape (spatial configuration), and the higher the uptake of built-up area per inhabitant or workplace (lower utilization intensity in the built-up area), the higher the degree of urban sprawl” [16].

Siedentop & Fina ([3] p. 2768) explicitly support these three major components of urban sprawl: They stated that the degree of urban sprawl in a study area increases when more urban land is present in the study area, when the urban land-use patches are more dispersed, and when their urban density is lower.

B. Mathematical formulation of the urban sprawl metrics

To measure urban sprawl, two metrics were applied: Weighted Urban Proliferation [15,16] (*WUP*$_p$, measured in urban permeation units per square meter [UPU/m²]) is the landscape-oriented metric used to quantify urban sprawl, while Weighted Sprawl per Capita (*WSPC*, unit: UPU/inhabitant) quantifies the average contribution of each inhabitant to urban sprawl (cf. Fig 2).

Weighted Urban Proliferation (*WUP*$_p$)

The *WUP*$_p$ method combines three components (Fig 2) [15,16]:

- Proportion of built-up areas (*PBA*) in a reporting unit: $PBA = A_{\text{built-up}}/A_{\text{reporting unit}}$,
- Dispersion (*DIS*), in urban permeation units per square meter [UPU/m²], which is calculated as the average weighted distance between any two points chosen randomly within the built-up area in the landscape investigated, where the second point is chosen within a maximum distance that is called the horizon of perception (*HP*), i.e., the range up to which the distances between locations in the built-up areas are considered in the calculation of *DIS*. A weighting of the distances is necessary, which can be intuitively understood as describing the effort for delivering some service from one of the two points to the other, or for providing some kind of infrastructure between the two points (see detailed explanations in Jaeger et al., 2010b). The calculation of *DIS* can be easily based on a grid as the sum over small cells of built-up area:

  $$DIS(b) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_i}\left(\sum_{j=1}^{n_i}\left(\sqrt{\frac{2d_{ij}}{1\text{ m}}+1}-1\right)\frac{\text{UPU}}{\text{m}^2}+WCC(b)\right),$$

  where n_i denotes the number of built-up cells that are closer to cell i than the *HP*, d_{ij} is the distance between the centers of cells i and j, and $WCC(b)$ is the within-cell contribution, which depends on cell size b and can be calculated by

  $$WCC(b) = \left(\sqrt{\frac{0.97428\,b}{1\,m}+1.046}-0.996249\right)\frac{UPU}{m^2}$$

  (see explanations in detail at Jaeger et al., 2010b).
  For this study b = 19 m was used (with $WCC(b)$ = 3.43 UPU/m^2). The *HP* specifies the scale of analysis of urban sprawl. When the distances between two locations are larger than the *HP*, urban development at the two locations is considered independently. A value of *HP* = 2 km was used (see detailed explanations in EEA, (2016)), consistent with earlier studies.
- Land-uptake per person (*LUP*$_p$), in square meters per inhabitant [m²/inhabitant].

The values of *DIS* and *LUP*$_p$ are combined with *PBA* by two weighting functions in order to allow parts of the landscape in which built-up areas are more dispersed to be more clearly perceived ($0.5 < w_1(DIS) < 1.5$) and to reflect the understanding that dense places like inner cities are not considered to be sprawled ($0 < w_2(LUP_p) < 1$), which is in accordance with qualitative expert opinion (see [16] for detailed explanations). The formulas of the weighting functions are:

$w_1(DIS) = 0.5 + e^{0.294432\ \mathrm{m^2/UPU} \times DIS - 12.955} / (1 + e^{0.294432\ \mathrm{m^2/UPU} \times DIS - 12.955})$ and
$w_2(LUP_p) = e^{4.159 - 613.125\ \mathrm{inhabitant/m^2}/LUPp} / (1 + e^{4.159 - 613.125\ \mathrm{inhabitant/m^2}/LUPp})$.

The product of *PBA* and *DIS* is called urban permeation of the landscape (*UP*), which measures the degree to which a landscape is permeated by built-up areas (Jaeger et al., 2010b). Thus, its value is influenced by how much built-up area there is in a landscape and how it is arranged spatially. Values for landscapes of different size can be directly compared.

The formula for the calculation of WUP_p is: $WUP_p = UP \times w_1(DIS) \times w_2(LUP_p)$. The unit of WUP_p is urban permeation units per m^2 of landscape [UPU/m^2] [16].

For this study, the formulas for the calculation of the urban sprawl metrics regarding inhabitants was used, where LUP_p is defined as land uptake per person based on population data only (i.e., per inhabitant), while the original definition also included workplaces. However, consistent data about workplaces are sometimes difficult to obtain, e.g., in multinational studies. Larger regions are comparable even without workplace data because the ratio between inhabitants and workplaces exhibits less variability among larger regions than among small reporting units, and issues due to the lack of workplace data become relevant only for small reporting units [2].

The combination of three components of sprawl into one measure is an important advantage, compared to earlier studies in which various components of sprawl (e.g., amount of built-up area) were reported. As a consequence, WUP_p measures a rather complex phenomenon in a relatively simple way. The consideration of the values of its three components along with the value of WUP_p is helpful for the interpretation of their combined value in WUP_p.

Various tests of the behavior of WUP_p have shown that this method captures urban sprawl well [17] and is a more suitable method than most approaches used previously (based on 13 suitability criteria for measures of urban sprawl according to Jaeger & Schwick [16]). The method also satisfies the 34 requirements proposed in the literature for indicator selection for environmental reporting [18].

In order to remove any bias due to the boundaries of the reporting units, the cross-boundary connections (CBC) procedure in the calculation of the urban sprawl metrics was applied (i.e., the calculation considers built-up areas within a buffer zone of width *HP* around the reporting unit) [19]. The *DIS*, *LUP*, and *WUP* are intensive measures, i.e., their values remain the same when the analyzed region is being multiplied while maintaining landscape structure. Therefore, their values can be directly compared between reporting units of differing size.

Weighted Sprawl per Capita (*WSPC*)

The metric *WSPC* refers to the number of people in the reporting unit instead of the area of the reporting unit: $WSPC = WUP_p \times A_{\text{reporting unit}} / N_{\text{inhabitants}}$.

Accordingly,

$$WSPC = w_1(DIS) \times w_2(LUP_p) \times DIS \times A_{\text{built-up}} / N_{\text{inhabitants}}$$
$$WSPC = w_1(DIS) \times w_2(LUP_p) \times DIS \times LUP_p.$$

In other words, the measure of Weighted Urban Proliferation (WUP_p) is calculated as

$$WUP_p = PBA \times DIS \times w_1(DIS) \times w_2(LUP_p) = w_1(DIS) \times w_2(LUP_p) \times TS / A_{\text{reporting unit}},$$

where Total Sprawl is

$TS = DIS \times A_{\text{built-up}}$ (Fig 2).

It is useful to also define Weighted Total Sprawl (*WTS*) as

$$WTS = w_1(DIS) \times w_2(LUP_p) \times TS = w_1(DIS) \times w_2(LUP_p) \times DIS \times A_{\text{built-up}}.$$

Then it holds

$$WUP_p = WTS / A_{\text{reporting unit}}.$$

The value of *WTS* answers the question of how much urban sprawl there is in a reporting unit in total (similar to *TS*, but including the weighting functions of *DIS* and LUP_p). The value will usually be larger in larger reporting units (since *WTS* is an extensive measure).

The value of *WSPC* also combines three components (Fig 2), but it answers a different question than the value of WUP_p. The value of WUP_p answers the question of how much urban sprawl there is in a reporting unit per area (i.e., it is an intensive measure that can be compared between reporting units of differing sizes). The value of *WSPC* answers the question of how much on average each inhabitant contributes to urban sprawl in the reporting unit in which the person is living. It is an intensive metric and has the unit UPU/inhabitant. For example, when a new inhabitant moves into a reporting unit and builds a new house at the same LUP_p as everyone else in the reporting unit, in a location where *DIS* stays the same, *PBA* will increase, and WUP_p and *WTS* will increase accordingly, but *WSPC* does not change, because $WSPC = w_1(DIS) \times w_2(LUP_p) \times DIS \times LUP_p$. In contrast, when a new inhabitant moves into a reporting unit and moves into an existing building (i.e., the amount of built-up area and *DIS* stay the same as before), land uptake per person will decrease, and the value of *WSPC* will decrease accordingly. When a new inhabitant moves out of the reporting unit and his/her building is now used by fewer people (i.e., no change in the amount of built-up area and *DIS*), land uptake per person will increase, and the value of *WSPC* will increase accordingly.

In addition to the mathematical formulation, the following list provides a summary of all variables [2,16]:

- ***WUP_p*: Weighted Urban Proliferation** is a landscape-oriented metric used to quantify urban sprawl. It is the product of dispersion (*DIS*), a weighting of *DIS*, the percentage of built-up area (*PBA*), and a weighting of the land uptake per person (*LUP*). It is measured in **urban permeation units per square meter** of landscape (UPU/m^2). It is based on population data only (inhabitants).

- ***WSPC:*** **Weighted Sprawl per Capita** is the inhabitant-oriented metric used to quantify the average contribution of each inhabitant to urban sprawl, measured in **urban permeation units per inhabitant** (UPU/inhabitant).
- ***AUA:*** **Absolute urban area** is the size in **square kilometers** (km^2) of the built‒up areas within a reporting unit.
- ***PBA:*** The **percentage of built-up area** is the ratio of the size of the built-up areas to the size of the total area of the reporting unit and is given as a **percentage** (%).
- ***DIS:*** The **dispersion** quantifies the spatial distribution of built-up areas, expressed as **urban permeation units per m^2** of built-up area (UPU/m^2). The further dispersed the built-up areas, the larger the value of *DIS*. Therefore, more compact built-up areas have lower values of *DIS* than more scattered built-up areas.
- ***LUP_p:*** The **land uptake per person** is the **area of land used per person** (m^2/inhabitant). The values were calculated as the quotients of built-up area and population.

C. Data processing for the computation and mapping of urban sprawl

This section explains the processing steps required to study urban sprawl at multiple spatial scales. S1 Table presents a list of input datasets to conduct the urban sprawl study. The list of input data sets is structured by the following items: label of the dataset, spatial resolution/scale, time reference/period, information on the data availability and the use in this study.

The steps of calculating the urban sprawl metrics worldwide are rather complex and require high computing efforts (see Methods section and mathematical formulation in Part B). All calculations were based on the number and spatial distribution of built-up pixels within the GHSL (Global Human Settlement Layer) [20,21]. Due to the large amount of data, the global settlement grids could not be processed in one go, but tile-wise processing was required. For this purpose, a uniform square grid with 250 km edge length based on the true area Mollweide projection was defined, which was chosen in correspondence to the given projection of the GHSL population layer. This grid was first re-projected into the Mercator projection, which is the original spatial reference system of the GHSL built-up layer, and then used tile by tile to extract and process the corresponding area of the GHSL data set.

However, due to increasing distortions in length and area at higher geographical latitudes the conformal Mercator projection is not reliable for distance measurements, which are needed for the calculation of dispersion [22]. Therefore, in an intermediate processing step, each tile was first transformed into an equidistant azimuthal projection, which was uniquely defined by its tile center as coordinate origin, in order to support the distance measurements for calculating dispersion as close to reality as possible. The coordinate transformations from the original Mercator projection into the tile-specific equidistant azimuthal projections were combined with resampling to a pixel width of 19 m. Thus, the likelihood of a loss of isolated settlement pixels due to the process of coordinate transformation was reduced. Dispersion values were then calculated for each 19×19 m^2 built-up pixel as the mean of the weighted distances between each built-up pixel and all other built-up pixels within the distance of the horizon of perception (HP = 2 km) [4]. In this study, the value of 2 km was used for various reasons, e.g., the typical distances between two settlements in many countries are between 3 and 5 km because the land between them was needed for agriculture to feed the people in the villages (see more detailed explanations in [2,4]). The data were then transformed into the final target coordinate system (Mollweide) and merged into a global dispersion grid of 19 m resolution.

This layer served as input data for the superposition with the administrative boundaries in order to tabulate the absolute urban area (based on the numbers of built-up pixels) and the dispersion of each reporting unit (calculated as the mean of the pixel-based dispersion values). This step was performed separately for all spatial scales of reporting units (subnational units, nations, UN regions, continents, and the world) and for global and local regular grids (cell sizes of 50×50 km^2, 5×5 km^2, and 1×1 km^2).

For the urban sprawl analysis at national and subnational scales, the freely available Global Administrative Areas dataset was used (GADM, 2018, see also S1 Table). This high-resolution dataset comprises the borders of all countries and their administrative subdivisions. At the national scale, all states and state-like entities (autonomous regions, outer territories, etc.) having specific ISO codes have been included in the analysis. In this study, regions that are part of a sovereign state but have autonomous governments, overseas areas, provinces of a patron nation, and several other spatial entities with special legal state (e.g., Greenland, Guernsey, Overseas France) are considered as individual reporting units at the national scale. Therefore, the considered number of 244 countries (or nations) exceeds the current number of 193 member states of the UN, for instance [23]. According to the GADM dataset, 1,764 sub-national units were considered. Based on these territorial boundaries, the higher levels of reporting units have been aggregated. On the scale of UN regions, only the European part of Russia was decided to be included in the “Eastern Europe” region. The Asian part of Russia (Siberia) was defined as a separate region, called “Northern Asia”.

For visualization, boundary data originating from the Natural Earth project was employed.

Since built-up pixels can only occur on land areas, the percentage of built-up area was calculated while excluding large water bodies. The Natural Earth project provides a worldwide database for major lakes and reservoirs which was used to calculate the area of the water surface for each reporting unit [24]. These values were then deducted from the GADM-based administrative areas in order to determine the colonizable land area of each reporting unit. Settlement pixels from the GHSL data being situated beyond the vector coastlines have been assigned to the nearest administrative unit by applying an Euclidean distance allocation. Thus, no built-up pixels had to be omitted from the analysis due to deviations of the input datasets.

Based on the GHSL population layer (a grid at 250 m resolution), the population pixels were assigned to the reporting units and summed up to determine the total population in each unit.

All metrics of interest can then be calculated.

D. Comparison with settlement areas of the Global Urban Footprint dataset

Availability of reference data

The Global Urban Footprint (GUF) maps urban areas worldwide based on the radar satellites TerraSAR-X and TanDEM-X and provides built-up information of 2011 at a resolution of 12 meters. The authors were aware of technology-related uncertainties (optical vs. SAR-based mapping technology). However, the objective of the present study was to generate added value of existing global multi-temporal remote sensing data and to bring them to a globally relevant use case in the best possible way. The authors therefore used the GHSL dataset as a starting point for their urban sprawl calculations and refer to the quality assessments already carried out by the data providers.

Comparison of GHSL and GUF

In 2014, the total global settlement area of the GHSL amounted to approximately 780,000 km², while the GUF layer for 2011 resulted in a total amount of settlement area of approximately 835,000 km². Regional differences between the two datasets were observed: The map in S7 Fig shows the differences in the settlement areas between the two datasets for cells size of 50 × 50 km² as the proportion of the size of these cells. Positive values indicate a higher proportion of settlement areas in the GHSL than in the GUF layer (e.g., in the U.S. and in many countries in Central and Eastern Europe and Southern and Western Africa), while negative values indicate higher amounts of settlement areas captured by the GUF layer (e.g., in many Asian countries such as China, India, Pakistan, Bangladesh, or Vietnam).

The mean value of the differences is -0.046%, the median is -0.005%, while the differences range between -22.7% and +28.3%. The lower and upper 5-percentiles are at -0.83% and +0.59%, respectively, indicating that in 90% of all grid cells containing settlements, the difference in the settlement areas between the two datasets is less than 1% of the size of the grid cell (1% of 2,500 km^2 = 25 km^2).

It is worth noting that the reliability of the urban sprawl values calculated, as in other studies based on geographic datasets, depends strongly on the reliability of the raw data, i.e., the mapped settlement areas from satellite images. It is to be expected that in less densely populated areas (primarily peri- and suburban structures), settlement areas will be underestimated as a result of Landsat's image resolution. In contrast to this, urbanized areas tend to be overestimated due to image resolution. As a result, the urban sprawl values (e.g., dispersion calculated at pixel level) may be influenced by unmapped or overestimated settlement areas. The main differences were observed in the eastern part of the U.S., in several Central European countries (e.g., Netherlands and Belgium), and in Eastern Asia (S7 Fig).

In order to better assess and compare these over- or underestimated settlement areas, the urban sprawl metrics were calculated for both datasets (GHSL 2014 and GUF, see Part E:

Qualitative evaluation of the urban sprawl values based on GHSL in 2014 for various regions and settlement types using GUF for comparison).

Recently, the World Settlement Footprint 2019 (WSF) [25–27] has also become available. A new dataset called WSF Evolution has been also made available, which is promising for future comparative spatio-temporal analyses of urban sprawl. It covers the worldwide growth of human settlement areas on a year-by-year basis between 1985 and 2015 [27].

E. Qualitative evaluation of the 2014 GHSL urban sprawl calculations for various regions and settlement types using GUF-based calculations for comparison

In this section, GHSL 2014 and GUF based urban sprawl results are compared. The absolute differences between sprawl values are displayed at the scale of national and subnational units in the upper maps in S8 Fig and S9 Fig. In several Chinese coastal provinces, GHSL based urban sprawl values are lower than GUF based values. In contrast, higher GHSL based values than GUF based values are found in some Central European and North American coastal subnational units. S9 Table enables the comparison of WUP_p values at five spatial scales based on Global Human Settlement Layer (GHSL, 2014) and Global Urban Footprint (GUF, 2011/12).

The value of WUP_p of all terrestrial areas on the planet combined (except Antarctica) is somewhat higher for the GUF-based results (WUP_p = 0.065 UPU/m²) than for the GHSL based results (WUP_p = 0.050 UPU/m²).

At the continental scale, higher urban sprawl values for Europe (ΔWUP_p = 0.137 UPU/m²), North America (ΔWUP_p = 0.123 UPU/m²), Oceania (ΔWUP_p = 0.016 UPU/m²), and Africa (ΔWUP_p = 0.002 UPU/m²) are observed when using GHSL as input data. In contrast, South America (ΔWUP_p = -0.010) and Asia (ΔWUP_p = -0.024 UPU/m²) showed higher values when using GUF as input data.

At the regional scale, Western Europe (ΔWUP_p = 0.734 UPU/m²), followed by Southern Africa (ΔWUP_p = 0.211 UPU/m²) and Northern America (ΔWUP_p = 0.131 UPU/m²) showed higher values based on GHSL, whereas Micronesia (ΔWUP_p = -0.705 UPU/m²), Polynesia (ΔWUP_p = -0.427 UPU/m²), Eastern Asia (ΔWUP_p = -0.346 UPU/m²), Central Asia (ΔWUP_p = -0.088 UPU/m²), and Western Asia (ΔWUP_p = -0.057 UPU/m²) clearly exhibited lower values when using GHSL as input data.

The bottom maps in S8 Fig and S9 Fig show the classified urban sprawl values using the following class borders: values < 0.1 UPU/m^2 indicate very low levels of sprawl, 0.1–0.5 UPU/m^2: low, 0.5–1.5 UPU/m^2: moderate, 1.5–3.0 UPU/m^2: high, and values > 3.0 UPU/m^2: very high levels of urban sprawl. The class borders are based on the latest study by the European Environment Agency about urban sprawl in Europe[5] and own reflections and calculations. The cross tables of GHSL and GUF class combination frequencies in the abovementioned maps show a strong relationship between the GHSL and GUF based urban sprawl values. At the national and subnational scales, about 70% of the spatial units show corresponding urban sprawl classes for GHSL and GUF. For example, of the 20 countries described by GHSL as very highly sprawled, only four deviating units were identified on the basis of GUF (three are classified as highly sprawled: Germany, Luxembourg, Gibraltar, one is classified as moderately sprawled: British Virgin Islands). On the other hand, eight national units were characterized as less sprawled when using GHSL rather than using GUF (Nauru, Bermuda, Anguila, Guam, Guadeloupe, Cayman Islands, Reunion, Saint Martin).

S10 Fig presents both the absolute differences of the urban sprawl values and the classification consistency between GHSL and GUF based values for the 50 × 50 km² grid. The absolute differences revealed local hotspots of larger deviations (e.g., along the east coast of China and the U.S.), but the classified values in the cross table shows a one-to-one concordance for about 80% of the cells of the 50 × 50 km² grid (sum of the pixel frequencies in the cross table diagonal). Only 4% of all classified pixels differed by more than one class. While deviations between the urban sprawl classes were apparent along the east coast of China (e.g., Shandong Province, Jingjinji), the differences were clearly less pronounced in other regions of the world.

The absolute differences between the two calculations were also investigated. The differences increased for higher classes of urban sprawl at all three spatial scales. Interpretation of the results should take into account that low density peri-urban and sub-urban areas were not mapped in some regions by the GHSL due to the resolution of the Landsat imagery. A critical examination of this problem is illustrated in five case study areas (S11 Fig, S12 Fig, S13 Fig, S14 Fig, S15 Fig). The authors focused on areas exhibiting rather strong differences in settlement pixels between GHSL and GUF (e.g., eastern coastal areas of China and Northern America, the Netherlands, and Belgium). Furthermore, focus was set on several highly sprawled regions and compared high-density mega-cities or regions to low-density peri-urban or sub-urban areas. For each case study area, the figures show the settlement pixels of the mapping products GHSL and GUF. The differences in built-up area percentage were calculated on a 5 × 5 km² grid for a quantitative comparison (map titled: “Difference in built-up area percentage”). Negative (positive) values indicate that the GUF includes a larger (smaller) amount of built-up area per grid cell than GHSL. Each case study also includes maps of the degree of concordance of GHSL and GUF based urban sprawl classifications for the 5 × 5 km² grid and at the subnational scale (titled “WUP_p classification scheme”). Additionally, a map of the absolute differences between the GHSL and GUF based urban sprawl values is displayed for each case study (titled “Difference between WUP_p values”).

In the following, five case study areas are briefly discussed.

- **Jingjinji (Beijing-Tianjin-Hebei, the Chinese national capital region):** The GUF settlement area is visually considerably more extended and more differentiated than the GHSL dataset (S11 Fig). According to the analysis for the 5 × 5 km² grid, the dense agglomerations in the GHSL have significantly larger settlement areas than in the GUF, whereas the less densely populated areas between the large cities are more prominent in the GUF. While in the provinces of Hebei and Tianjin, the sprawl values based on GUF were higher (Hebei: ΔWUP_p = -1.295 UPU/m², Tianjin: ΔWUP_p = -1.856 UPU/m²), the sprawl value of the province of Beijing was higher based on GHSL (ΔWUP_p = +1.428 UPU/m²). However, these differences did not change the assignment of the three provinces to the very-high urban-sprawl class. The difference display of the WUP_p classes showed that both datasets produced largely identical results. However, the classification in the central and south-eastern parts of the map differed.

- **Burkina Faso:** The settlement area according to GHSL is considerably more extended and more differentiated than in the GUF dataset (S12 Fig). A recent systematic accuracy assessment of GUF and GHSL in Africa indicated considerable differences, in particular in low-density semi-desert areas such as Burkina Faso[65]. For that reason, the urban sprawl values of this case study were most likely too low in rural areas. However, settlement areas of cities were still captured with sufficient reliability and the related urban sprawl values can therefore be regarded as valid.
- **Netherlands and Belgium:** In these two very highly sprawled countries, the GHSL showed a larger built-up area than the GUF dataset, resulting in considerable differences in the absolute urban sprawl values (e.g., South Holland (NL): ΔWUP_{p} = +10.312 UPU/m², North Brabant (NL): ΔWUP_{p} = +4.865 UPU/m², Antwerp Province (BE): ΔWUP_{p} = +5.737 UPU/m²). However, the cross table indicated similar classes of urban sprawl, i.e., most areas were labeled as very highly sprawled (S13 Fig).
- **England (UK):** The GHSL showed a higher share of built-up areas in urban areas and a lower share of built-up areas in rural areas than GUF (S14 Fig). At the subnational scale, this resulted in a rather large range of differences in urban sprawl values between ΔWUP_{p} = +0.087 UPU/m² (East Midland) and ΔWUP_{p} = +9.673 UPU/m² (London). Nevertheless, all subnational units (with the exception of South West England) remained in the class of very high urban sprawl. Also in the 5 × 5 km² grid, most areas belonged to the same urban sprawl class notwithstanding differing amounts of urban area.
- **New York Metropolitan Area (U.S.):** Large parts of the U.S. east coast are characterized by larger built-up areas in the GHSL than in the GUF (S15 Fig). In the 5 × 5 km² grid, quite similar urban sprawl classes for the two data products are observed. At the subnational scale, the State of New York is highly sprawled based on GHSL and moderately sprawled according to GUF. The neighboring States of New Jersey and Connecticut are very highly sprawled according to both input datasets.

Overall, this evaluation indicated that the urban sprawl results can to a large extent be regarded as representative and valid across the planet. It is difficult to find a single one data product of settlement areas that would solve all potential issues of settlement analysis. A major advantage of the GHSL is its availability for four points in time and its suitability for temporal analysis of settlement development and urban sprawl for the last 25 years (1990–2015) and, to some degree, even for the last 40 years (including 1975, see [28]). Therefore, GHSL can be used to prepare initial reference data about urban sprawl for comparison in future studies.

F. Changes in urban sprawl values when excluding selected irreclaimable areas from the reporting units

The measurement of urban sprawl can refer to reporting units including or excluding those parts of the landscape that are not suitable for the construction of houses or the establishment of settlement (called “irreclaimable areas”, see [29]). Both values have their respective valid meanings. Since regions differ in the amount of area that is not suitable for construction, a comparison of such regions may be more appropriately done after excluding the irreclaimable parts of the landscape from the reporting units. This means that WUP_p then provides the degree of urban sprawl of the landscape in relation to the area that is, in principle, potentially suitable for settlements. For example, the degree of sprawl of a reporting unit that includes a major lake may be compared more appropriately to a reporting unit without a lake after excluding the lake. The authors already measured urban sprawl after excluding the water areas from the reporting unit areas.

This section provides information about the changes in the WUP_p values when excluding other selected irreclaimable areas, following the approach of the European study ([29] p. 492-494), which compared the results under inclusion and exclusion of uninhabitable parts of the landscape.

In order to evaluate the influence of land cover classes that are generally not suitable for settlements, MDA's BaseVue 2013 Global Land Cover dataset [30] was used to estimate the proportion of irreclaimable land. It is based on Landsat 8 and distinguishes 13 land use/land cover classes globally at a resolution of 30 m. The MDA's BaseVue 2013 Global Land Cover dataset is accessible through ArcGIS online and has been used as input for the identification of the following irreclaimable land-cover classes: Ice/Snow and Barren or Minimal Vegetation for all reporting units. This was realized by tile-wise analysis and subsequent merging of the results. Icy and barren land (including deserts) were subtracted from the land area and the sprawl metrics for the GHSL 2014 dataset were then re-calculated in order to estimate the influence of these areas on the sprawl values.

The types of areas considered as not suitable for settlements included the following land cover classes in BaseVue:

- **Ice/Snow:** Land areas covered permanently or nearly permanently with ice or snow
- **Barren or minimal vegetation:** Land with minimal vegetation (< 10%) including rock, sand, clay, beaches, quarries, strip mines, and gravel pits. Salt flats, playas, and non-tidal mud flats were also included if they were not inundated with water.

The water class from this dataset was not used, since water bodies had already been subtracted from the administrative areas based on the Natural Earth Dataset (see part C, S1 Table).

The largest differences between WUP_p values with and without accounting for irreclaimable areas were expected to find in the reporting units that exhibit a large spatial extent of the excluded land-cover types (e.g., reporting units that include large desert areas, large mountainous regions, large areas covered by ice or snow).

The value of $WUP_{p;excl}$ excluding the irreclaimable areas can be easily calculated from the original value of WUP_p by $WUP_{p;excl} = WUP_p/P_s$, where P_s is the proportion of area suitable for construction in the reporting unit. For example, if WUP_p is 1.62 UPU/m^2 and 10% of the land area of the reporting u nit are irreclaimable, P_s is 90%, and $WUP_{p;excl}$ results in 1.62 UPU/m^2 / 0.9 = 1.8 UPU/m^2. The reason is that the irreclaimable areas usually do not include any built-up areas nor inhabitants, and therefore, *PBA* changes by the factor $1/P_s$, while LUP_p and *DIS* are unchanged. In rare cases, there may be a few buildings located in areas generally unsuitable for construction and there may be a few inhabitants in areas that do not have buildings, but these differences are marginal.

Note that the value of *WSPC* does not differ when the reporting unit includes or excludes the irreclaimable areas, when the irreclaimable areas do not include any built-up areas nor inhabitants, which is usually the case. Therefore, the components of *WSPC* are unchanged (LUP_p and *DIS*).

The value of WUP_p of all terrestrial areas on the planet combined (except Antarctica) is somewhat higher when accounting for irreclaimable areas (WUP_p = 0.061 UPU/m²) than before (WUP_p = 0.050 UPU/m²), (S9 Table).

At the continental scale, North America showed the highest difference in urban sprawl (ΔWUP_p = 0.067 UPU/m²), followed by Europe (ΔWUP_p = 0.026 UPU/m²) and Oceania (ΔWUP_p = 0.002 UPU/m²). Africa exhibited the highest relative difference (ΔWUP_p = +53.8%), followed by Asia (ΔWUP_p = +26.3%).

At the scale of UN regions, Northern America (ΔWUP_p = 0.082 UPU/m²), Southern Africa (ΔWUP_p = 0.069 UPU/m²), and Northern Europe (ΔWUP_p = 0.056 UPU/m²) revealed the highest absolute changes in urban sprawl values. Northern Africa exhibited the highest relative difference (ΔWUP_p = +421.7%), followed by Western Asia (ΔWUP_p = +186.2%) and Western Africa (ΔWUP_p = +75.2%).

Excluding irreclaimable areas resulted in larger WUP_p values for just a small group of reporting units at the national and subnational scales. S16 Fig depicts the 19 countries and 72 subnational units that are most strongly affected by the exclusion of irreclaimable areas. These reporting units showed an increase in WUP_p of more than 0.1 UPU/m². For example, several Swiss subnational units showed higher sprawl values due to their geo-physical conditions, e.g., presence of rocks and glaciers. The total sprawl value of Switzerland increased by ΔWUP_p = 0.269 UPU/m². South Holland in the Netherlands is another example of a higher urban sprawl value (ΔWUP_p = 0.119 UPU/m²), due to areas of barren or minimal vegetation. The largest differences at the subnational scale were observed for regions characterized by deserts or minimal vegetation, respectively, e.g., Galguduud (Somalia, ΔWUP_p = 4.290 UPU/m²), Souhag (Egypt, ΔWUP_p = 1.944 UPU/m²), Absheron (Azerbaijan, ΔWUP_p = 1.883 UPU/m²), Atacama (Chile, ΔWUP_p = 0.936 UPU/m²), and Cyrenaica (Libya, ΔWUP_p = 0.818 UPU/m²). At the national scale, Israel (ΔWUP_p = 1.415 UPU/m²), Libya (ΔWUP_p = 0.863 UPU/m²), and Kuwait (ΔWUP_p = 0.786 UPU/m²) are partly characterized by desert regions and showed the highest changes.

A more detailed and more realistic determination of irreclaimable areas would be possible for every particular reporting unit. For example, wetlands and mangrove areas can be considered as not suitable for settlement development in many regions. Protected areas should also be excluded if the construction of buildings is not permitted. However, the task of identifying all these areas cannot easily be done at the global scale because of limited data availability and because protection regulations differ considerably among countries.

In conclusion, it is important to bear in mind the potential changes in urban sprawl values due to the exclusion of irreclaimable areas from the reporting units. However, no critical differences that would question the general trends found in this study were observed.

References

1. OECD. Rethinking Urban Sprawl: Moving Towards Sustainable Cities. OECD; 2018. doi:10.1787/9789264189881-en

2. EEA (European Environment Agency), FOEN (Swiss Federal Office for the Environment). Urban sprawl in Europe. Joint EEA-FOEN report (EEA Report No 11/2016). Luxembourg: Publications Office of the European Union; 2016. Available: http://www.eea.europa.eu/publications/urban-sprawl-in-europe

3. Siedentop S, Fina S. Who Sprawls Most? Exploring the Patterns of Urban Growth across 26 European Countries. Environ Plan A. 2012;44: 2765–2784. doi:10.1068/a4580

4. Jaeger JAG, Bertiller R, Schwick C, Cavens D, Kienast F. Urban permeation of landscapes and sprawl per capita: New measures of urban sprawl. Ecological Indicators. 2010;10: 427–441. doi:10.1016/j.ecolind.2009.07.010

5. FOEN (Swiss Federal Office for the Environment). Landscape indicator urban sprawl. 2017. Available: https://www.bafu.admin.ch/bafu/en/home/themen/thema-landschaft/landschaft--daten--indikatoren-und-karten/landschaft--indikatoren/indikator-landschaft.pt.html/aHR0cHM6Ly93d3cuaW5kaWthdG9yZW4uYWRtaW4uY2gvUHVibG/ljL0FlbURldGFpbD9pbmQ9TEEwMTImbG5nPWVuJlN1Ymo9Tg==.html

6. EEA (European Environment Agency). Urban sprawl in Europe: The ignored challenge (EEA Report No 10/2006). Luxembourg: Publications Office of the European Union; 2006. Available: https://www.eea.europa.eu/publications/eea_report_2006_10/eea_report_10_2006.pdf

7. Harvey RO, Clark WA. The nature and economics of urban sprawl. Internal Structure of the city readings on space and environment. New York: Oxford Univ. Pr.; 1971. pp. 475–482.

8. Ewing R. Is Los Angeles-Style Sprawl Desirable? Journal of the American Planning Association. 1997;63: 107–126. doi:10.1080/01944369708975728

9. USHUD (U. S. Dept. of Housing and Urban Development O of PD and R. The state of the cities 1999: Third annual report. Washington, D.C.: U.S. Dept. of Housing and Urban Development, Office of Policy Development and Research; 1999. Available: https://files.eric.ed.gov/fulltext/ED438358.pdf

10. Sierra Club (Ed.). The dark side of the American dream: The costs and consequences of suburban sprawl. San Francisco, CA: Sierra Club; 1998.

11. Ewing RH, Pendall R, Chen DD. Measuring sprawl and its impact. Washington, D.C.: Smart Growth America; 2002.

12. Burchell RW, Galley CC. Projecting Incidence and Costs of Sprawl in the United States. Transportation Research Record. 2003;1831: 150–157. doi:10.3141/1831-17

13. Ewing RH. Characteristics, Causes, and Effects of Sprawl: A Literature Review. In: Marzluff J, Shulenberger E, Endlicher W, Alberti M, Bradley G, Ryan C, editors. Urban ecology: An international perspective on the interaction between humans and nature. New York: Springer; 2008. pp. 519–535.

14. Steurer M, Bayr C. Measuring urban sprawl using land use data. Land Use Policy. 2020;97: 104799. doi:10.1016/j.landusepol.2020.104799

15. Jaeger JAG, Bertiller R, Schwick C, Kienast F. Suitability criteria for measures of urban sprawl. Ecological Indicators. 2010;10: 397–406. doi:10.1016/j.ecolind.2009.07.007

16. Jaeger JAG, Schwick C. Improving the measurement of urban sprawl: Weighted Urban Proliferation (*WUP*) and its application to Switzerland. Ecological Indicators. 2014;38: 294–308. doi:10.1016/j.ecolind.2013.11.022

17. Orlitova E, Soukup T, Kopecky M, Vobora V, Gregor M, Schwick C, et al. Urban sprawl typology: Critical analysis of input datasets, indicators calculated and dataflows proposal. Working document, Copenhagen; 2012.

18. Niemeijer D, de Groot RS. A conceptual framework for selecting environmental indicator sets. Ecological Indicators. 2008;8: 14–25. doi:10.1016/j.ecolind.2006.11.012

19. Moser B, Jaeger JAG, Tappeiner U, Tasser E, Eiselt B. Modification of the effective mesh size for measuring landscape fragmentation to solve the boundary problem. Landscape Ecology. 2007;22: 447–459. doi:10.1007/s10980-006-9023-0

20. Pesaresi M, Melchiorri M, Siragusa A, Kemper T. Atlas of the Human Planet 2016 - Mapping Human Presence on Earth with the Global Human Settlement Layer. Ispra: Publications Office of the European Union; 2016. Available: http://publications.jrc.ec.europa.eu/repository/handle/JRC103150

21. Freire S, MacManus K, Pesaresi M, Doxsey-Whitfield E, Mills J. Development of new open and free multi-temporal global population grids at 250 m resolution (Conference short paper). In: Sarjakoski T, Santos MY, Sarjakoski T, editors. Proceedings 2016 The 19th AGILE International Conference on Geographic Information Science Geospatial Data in a Changing World (Helsinki, Finland, June 14-17, 2016).

22. Maling DH. Coordinate systems and map projections. 2. ed. Oxford: Pergamon Press; 1992.

23. UN Visitors Centre. United Nations Member States. 2013. Available: https://visit.un.org/sites/visit.un.org/files/FS_List_member_states_Feb_2013.pdf

24. Natural Earth. Free vector and raster map data at 1:10m, 1:50m, and 1:110m scales. 2009. Available: http://www.naturalearthdata.com/downloads/

25. Esch T, Heldens W, Hirner A, Keil M, Marconcini M, Roth A, et al. Breaking new ground in mapping human settlements from space – The Global Urban Footprint. ISPRS Journal of Photogrammetry and Remote Sensing. 2017;134: 30–42. doi:10.1016/j.isprsjprs.2017.10.012

26. Leyk S, Uhl JH, Balk D, Jones B. Assessing the accuracy of multi-temporal built-up land layers across rural-urban trajectories in the United States. Remote Sensing of Environment. 2018;204: 898–917. doi:10.1016/j.rse.2017.08.035

27. Hennig EI, Schwick C, Soukup T, Orlitová E, Kienast F, Jaeger JAG. Multi-scale analysis of urban sprawl in Europe: Towards a European de-sprawling strategy. Land Use Policy. 2015;49: 483–498. doi:10.1016/j.landusepol.2015.08.001

28. Frye C, Nordstrand E, Wright DJ, Terborgh C, Foust J. Using Classified and Unclassified Land Cover Data to Estimate the Footprint of Human Settlement. Data Science Journal. 2018;17: 20. doi:10.5334/dsj-2018-020

List of Legends (Supporting Figures)

S1 Fig. Development of urban sprawl metrics between 1975 and 2014 at the continental scale. The diagrams are based on all available GHSL time series.1975 values have been calculated and integrated for presentation completeness, but have not been included in the deeper analysis due to data validity reasons of early remote sensing missions. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S2 Fig. Development of urban sprawl between 1975 and 2014 at the scale of sub-continental regions (UN regions). The diagrams are based on all available GHSL time series. 1975 values have been calculated and integrated for presentation completeness. UN Region Eastern Europe was split here into "Eastern Europe" and "Northern Asia" (Asian part of Russia). Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014] _globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S3 Fig. Development of urban sprawl between 1975 and 2014 in large (population > 5 million, area > 20,000 km²) countries and subnational units with high and very high urban sprawl values (WUP_p > 1.5 UPU/m² in 2014). The diagrams are based on all available GHSL time series. 1975 values have been calculated and integrated for presentation completeness, but have not been included in the deeper analysis due to data validity reasons of early remote sensing missions. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000| 2014]_globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S4 Fig. Dimensions of urban sprawl metrics of national and subnational units in 2014 (Filter: POP > 5 million, $Area_{admin}$ > 20,000 km², $WUP_p \geq 1.5$ UPU/m²). For subnational units, the curves of the corresponding nation are displayed as a gray line. Values were normalized using the presented spatial units. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_ 54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S5 Fig. Absolute and relative change of urban sprawl values between 1990 and 2014 at the scale of a 50 × 50 km² grid. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990| 2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_ 38_v1_0. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative

Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Behrmann. All map contents comply with PLOS license CC-BY 4.0.

S6 Fig. Absolute and relative change of urban sprawl per capita values between 1990 and 2014 at the scale of a 50 × 50 km² grid. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_ v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_ r2016a_3857_38_v1_0. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Behrmann. All map contents comply with PLOS license CC-BY 4.0.

S7 Fig. Differences in built-up area between GHSL and GUF at the scale of a 50 × 50 km² grid. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Map made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Behrmann. All map contents comply with PLOS license CC-BY 4.0.

S8 Fig. Absolute differences (upper map; $\Delta \boldsymbol{WUP_p}$) and qualitative relations (lower map; sprawl classes) between GHSL and GUF based urban sprawl values at the scale of national units. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_ v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Hammer Wagner. All map contents comply with PLOS license CC-BY 4.0.

S9 Fig. Absolute differences (upper map; $\Delta \boldsymbol{WUP_p}$) and qualitative relations (lower map; sprawl classes) between GHSL and GUF based urban sprawl values at the scale of subnational units. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_ 54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Hammer Wagner. All map contents comply with PLOS license CC-BY 4.0.

S10 Fig. Absolute and qualitative differences between GHSL and GUF based urban sprawl values at the scale of 50 × 50 km² grid cells. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_ 54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_ r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6,

2018) @ gadm.org. Map Projection: Behrmann. All map contents comply with PLOS license CC-BY 4.0.

S11 hFig. Data comparison GHSL-GUF: Jinjinji (China). Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S12 Fig. Data comparison GHSL-GUF: Burkina Faso. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S13 Fig. Data comparison GHSL-GUF: Netherlands and Belgium. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S14 Fig. Data comparison GHSL-GUF: United Kingdom. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S15 Fig. Data comparison GHSL-GUF: New York (US). Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. Maps made with Natural Earth and GADM: Natural Earth 4.0 @ naturalearthdata.com, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S16 Fig. Changes in WUP_p [UPU/m²] values due to the exclusion of irreclaimable areas (barren or minimal vegetation, ice/snow) from the reporting units (countries and subnational units with $\Delta WUP_p > 0.1$ UPU/m²) in 2014. Data Sources: European

Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. World Land Cover 30m BaseVue 2013 @ MDAUS (https://landscape6.arcgis.com/arcgis/rest/services/World_Land_Cover_30m_BaseVue_2013/ImageServer).

S17 Fig. Patterns of urban sprawl in agglomerations at the scale of a 1 × 1 km² grid: Moscow. Characteristics for the two urban sprawl metrics WUP_p and *WSPC* and their dimensions LUP_p, *DIS* and *PBA*. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015] _globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Maps made with GADM: Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Orientation map: © OpenStreetMap contributors. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S18 Fig. Patterns of urban sprawl in agglomerations at the scale of a 1 × 1 km² grid: Montréal. Characteristics for the two urban sprawl metrics WUP_p and *WSPC* and their dimensions LUP_p, *DIS* and *PBA*. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015] _globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Maps made with GADM: Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Orientation map: © OpenStreetMap contributors. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S19 Fig. Patterns of urban sprawl in Europe at the scale of a 1 × 1 km² grid: "Blue Banana" region. Characteristics for the two urban sprawl metrics WUP_p and *WSPC* and their dimensions LUP_p, *DIS* and *PBA*. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015] _globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Maps made with GADM: Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Orientation map: © OpenStreetMap contributors. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S20 Fig. Coastal areas affected by high urban sprawl at the scale of a 1 × 1 km² grid: Los Angeles. Characteristics for the two urban sprawl metrics WUP_p and *WSPC* and their dimensions LUP_p, *DIS* and *PBA*. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015] _globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Maps made with GADM: Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Orientation map: © OpenStreetMap contributors. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

S21 Fig. Coastal areas affected by high urban sprawl at the scale of a 1 × 1 km² grid: French Riviera. Characteristics for the two urban sprawl metrics WUP_p and *WSPC* and their dimensions LUP_p, *DIS* and *PBA*. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015] _globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Maps made with GADM: Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org. Orientation map: © OpenStreetMap contributors. Map Projection: Mollweide. All map contents comply with PLOS license CC-BY 4.0.

List of Legends (Supporting Tables)

S1 Table. Input datasets to measure urban sprawl at the global scale. The table provides all relevant information about the data products used for the study.

S2 Table. Urban sprawl values from global to national (top 40) scale. WUP_p 1990, 2000, 2014 (sorted descending by 2014 values), change in urban sprawl 1990-2014 (ΔWUP_p, ranking for absolute change), weighted sprawl per capita (*WSPC*, ranking for 2014). Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S3 Table. Values of urban sprawl components Dispersion, Percentage of built-up area, and Land uptake per inhabitant from global to national (top 40 WUP_p) scale. 1990, 2000, 2014 (sorted descending by WUP_p for 2014). Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S4 Table. Highest 50 urban sprawl values at the subnational scale. WUP_p 1990, 2000, 2014 (sorted descending by 2014 values), change in urban sprawl 1990-2014 (ΔWUP_p), weighted sprawl per capita (*WSPC*). Excluding subnational units <100 km². Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S5 Table. Values of urban sprawl components Dispersion, Percentage of built-up area, and Land uptake per inhabitant at the subnational scale 1990, 2000, 2014. Top 50, administrative area > 100 km², sorted descending by WUP_p 2014. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0, Global Administrative Boundaries (GADM 3.6, 2018) @ gadm.org, Natural Earth 4.0 @ naturalearthdata.com.

S6 Table. Chi-Square Independence Test (classes: WUP_p & *HDI*, subnational, 2014). Data Sources: Global Data Lab (GDL) @ Radboud University: https://globaldatalab.org/shdi/, European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Chi-Square Test of Independence in R (©sthda.com 2016).

S7 Table. Chi-Square Test: Residuals (classes: WUP_p & *HDI*, subnational, 2014). Data Sources: Global Data Lab (GDL) @ Radboud University: https://globaldatalab.org/shdi/, European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Chi-Square Test of Independence in R (©sthda.com 2016).

S8 Table. Chi-square test: Contribution of each cell to the total Chi-square score as percentage (classes: WUP_p & *HDI*, subnational, 2014). Data Sources: Global Data Lab (GDL) @ Radboud University: https://globaldatalab.org/shdi/, European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Chi-Square Test of Independence in R (©sthda.com 2016).

S9 Table. Descriptive statistical summary of WUP_p values based on different input datasets. Data Sources: European Commission, Joint Research Centre (JRC): ghs_pop_gpw4[1975|1990|2000|2015]_globe_r2015a_ 54009_250_v1_0, ghs_built_lds[1975|1990|2000|2014]_globe_r2016a_3857_38_v1_0. Global Urban Footprint (GUF) @ DLR 2016. World Land Cover 30m BaseVue 2013 @ MDAUS (https://landscape6.arcgis.com/arcgis/rest/services/World_Land_Cover_30m_BaseVue_2013/ImageServer).

S10 Table. List of countries and autonomously governed parts of a patron nations that were included in the analysis at the national scale. Data source: ISO 3166 country codes, UN member states @ un.org.

List of Legends (Supporting Data)

S1 Data: Data sheet with urban sprawl values (1990, 2000, 2014) for different spatial scales and a data description sheet