

Supporting Information

Seto et al. 10.1073/pnas.1211658109

SI Materials and Methods

Data. Projection and resolution. All data used in the simulations and the subsequent analyses are in Goode's Homolosine Equal-Area projection. The spatial resolution of the maps is 5 km.

Year 2000 urban map. We extract urban extent circa 2000 from National Aeronautics and Space Administration Moderate Resolution Imaging Spectroradiometer (MODIS) land-cover product v5 (1), which provides a conservative estimate of global urban land cover and is significantly lower than those from the Global Rural-Urban Mapping Project (GRUMP) dataset (2). We resample the global urban land-cover map to 5-km resolution from its native resolution of 463 m.

Region maps. We use 16 regions broadly based on the United Nations world regions (Table S1). We deviate from the United Nations regions when one country is economically dissimilar [as measured by per capita gross domestic product (GDP) and total output] to other countries in its assigned region and economically more similar to a neighboring region. We treat China and India as individual regions because of the size of their population, economy, and land area.

Population density. We use population densities from GRUMP to create the population density driver map (3). We first reproject the original GRUMP map to Goode's Homolosine using bilinear interpolation. Next, we carry out a zonal analysis and calculate the mean population density for each pixel using a 5 × 5 window. The window corresponds approximately to the size of a pixel in the final map with 5-km resolution. We resample the resulting mean map to 5 km using nearest neighbor.

Methods. Estimating amount of new urban land (2000–2030). We estimate the amount of new urban land in each region by 2030 in a Monte-Carlo fashion based on present empirical distribution of regional urban population densities and probability density functions of projected regional population and GDP values for 2030. The steps involved are detailed below.

Probabilistic regional population estimates for 2030. We use United Nations World Population Projections 2008 revision as the starting point to derive probabilistic estimates of the total and urban region populations in 2030 (4). We use the uncertainty estimates reported in US National Research Council report *Beyond Six Billion* (5). Although the report uses the United Nations population projections from 1998 revision, this is the only available comprehensive estimate of uncertainty in any world population projections. We considered using the country-level or region-level probabilistic projections from the International Institute for Applied Systems Analysis (6). However, at the time of the analysis, these projections for all of the regions across the world were not yet available. We also considered the approach of the Intergovernmental Panel on Climate Change (IPCC) where future population is bracketed by high and low estimates. We refrain from this approach because: (i) the distribution between the low and high variants and even beyond them are unspecified, (ii) it does not address uncertainty in fertility and mortality, and (iii) the variants may become probabilistically inconsistent when aggregated over countries or regions (5, 7). The method used in the National Research Council report addresses these concerns and generates methodologically consistent estimates across the world.

We use the uncertainty ranges reported for selected country, regions, and the world in tables F.3 and F.5 of the appendix F of the report. Based on the information in these tables, we fit distributions to the resulting future population distributions. We

fitted generalized logistic distribution to urban population in 2030 for each region with different parameters. The fitted probability density function (PDF) and the 2.5% and 97.5% quartiles by region are shown in Fig. S3.

The regions used in the report do not correspond exactly to the regions in our study. However, the differences are minor and, therefore, we apply the uncertainty range estimates for each region/country in the report to the closest corresponding region/country in our study (Table S1). For example, we use uncertainty ranges reported for Latin America/Caribbean region in the report for both South America and Central America in our study.

To get 1,000 urban population estimates for 2030 for each region, we randomly draw 1,000 values from the corresponding PDF of total population projection and multiply them by the corresponding regional estimate of the urban proportion of the population from the United Nations Urbanization Prospects 2009 Revision (8). We assume the urban proportion estimates given in the 2009 Revision are the same across all population estimates.

Probabilistic regional GDP estimates for 2030. To develop region-level probabilistic GDP projections, we use country-level GDP projections used by the IPCC for their four narrative storylines in the Special Reports on Emissions Scenarios (9). We aggregate the country-level projections to obtain corresponding projections for the regions in our study. We use the lowest and highest of the projected region-level GDP estimates as the minimum and maximum values of a uniform distribution in the absence of any information as to the likely distribution of projected GDP estimates. Although this approach suffers from some of the same drawbacks listed above in the context of variant-based population projections, the estimates we use are the best available projections to account for the uncertainty in the absence of more information about the likely patterns of region-level GDP change.

We draw 1,000 values from the resulting region-level uniform GDP distributions for 2030 and divide them by 1,000 values drawn from the region-level population projections for the same year. This process gives us 1,000 estimates for per capita GDP by region for 2030.

Spatial distribution of urban population density by region. We next extract the spatial distribution of urban population densities circa 2000 used in both calculating the average per capita urban land and allocating the projected urban population estimates across each region. We overlay the initial urban extent layer with the population density layer. We then create empirical frequency distributions of population density of urban pixels by region.

Relationship between per capita GDP and per capita urban land. From the spatial distributions, we calculate the average per capita urban land circa 2000 for each region. Because our projection is equal area, we multiply the population density value of each pixel by 25 km² to get an estimate of the population in each pixel. Then we plot the average per capita urban land estimate of each region versus per capita GDP estimates for 2000 from the United Nations Statistics Division (10). We fit a linear model whose slope serves as a globally averaged estimate of the change in per capita urban land for a unit change in per capita GDP (Fig. S4). Although the spread is wide, the slope is positive, indicating an increasing trend in per capita urban land with increasing per capita GDP. We then estimate per capita GDP at 2030 for each region using probabilistic projections of region-level GDP and population at 2030.

Estimates of urban expansion from 2000 to 2030. To calculate the urban-land expansion by region because of only the urban population increase, we take a value from the set of 1,000 urban

population estimates. Then we draw one value from the empirical urban population density and subtract it from the urban population estimate. We continue with this process until all estimated urban population is allocated across space. We repeat this for each 1,000 urban population estimate and eventually obtain 1,000 estimates of urban land expansion needed to accommodate the projected increase in urban population assuming the population density distribution across space will not change substantially through 2030.

To incorporate the increase in per capita urban land because of the increase in per capita GDP we use the relationship in Fig. S4. First, we calculate the region-level estimates for per capita urban land for 2030. We divide the urban expansion estimated solely by allocating the projected urban population across space by the projected urban population. Then, we use these estimates, the region-level estimates for per capita GDP, and the relationship we derive between per capita GDP and per capita urban land to get 1,000 estimated values for the per capita urban land for each region at 2030 (Fig. S5).

Although we incorporate the spatial variability in urban population densities into our analysis, we also assume that the distribution of urban population density will remain constant through 2030. We incorporate this variability in our estimates, which allows us to use more of the spatial information that is available to us. This process allows us to develop a range of likely variability in future urban expansion rather than come up with a mostly likely point estimate. More importantly, it allows us to allocate the estimated urban expansion across space. A complementary approach, adopted by Angel et al. (11), is to ignore the spatial variability in densities but to produce an estimate for urban expansion for different scenarios based on how urban densities are expected to change over time.

Spatial Simulations. The urban land-use change model we use is derived from GEOMOD, a spatially explicit grid-based land-use and land-cover change model. GEOMOD has been fully described in the literature (12) and been applied extensively (13, 14). Our study is the first to use the model's framework in a probabilistic analysis of land change and in a primarily urban land change context. GEOMOD simulates the change between exactly two land categories (e.g., "urban" and "nonurban"). The input maps are an initial land-cover map, and several "driver" maps such as proximity to roads and elevation. Through statistical analysis of the empirical patterns created by the overlay of the initial land-cover map with the driver maps, a map that shows the overall suitability of each grid cell for change is generated. The model also reads from a text file the number of grid cells of each category at a final time. Based on these inputs, the model allocates the net change in each land cover category between the initial and final time points across the study area and, thus, simulates the spatial pattern of land change across the landscape.

We select GEOMOD as the platform to build our land change model because it can be calibrated with input maps from a single year. This functionality of the model is important for this study because the available global urban land-cover maps are from a single point in time only (e.g., GRUMP circa 1995, but MODIS circa 2000) and these maps are not compatible with each other (15). Absent a time-series of global urban land-cover maps, we use GEOMOD framework to generate multiple realizations of global urban expansion patterns out to 2030.

We have significantly modified the original GEOMOD model to create the model we use in this study and call the modified model, URBANMOD. The most significant changes are detailed below:

i) GEOMOD's allocation algorithm prioritizes candidate pixels according to their suitability values. Starting from the pixels with the highest suitability value, it allocates new urban growth until all pixels with that suitability value are con-

verted. It then proceeds to the pixels with the next highest suitability value and so on (12). Although this approach works satisfactorily for relatively small areas when applied at very large regional or continental scales, it does lead to unrealistic clustering of new urban areas in only the locations with the highest suitability. It is more realistic to assume that urban development will take place in proportion to its observed distribution across grid cells with different suitability values rather than allocating as much development as possible in wholesale manner starting with those grid cells with the highest suitability value in strict hierarchical order.

- ii) When a large number of undeveloped grid cells are tied for the selection of the next grid cell to be developed, GEOMOD employs an algorithm in which it allocates development among the tied grid cells in a uniform fashion. When large regions are tied this approach leads to an artificial salt-and-pepper pattern of urban grid cells. We modify this assumption so that the new model allocates urban grid cells randomly among the tied locations.
- iii) We modify the model so that it takes multiple estimates of magnitude of urban expansion and consequently generates a probabilistic urban growth map rather than a single deterministic urban expansion map.

We use URBANMOD to spatially allocate the projected urban expansion derivation of which is detailed in the previous section. We conduct the simulations on IBM iDataplex cluster (Eos) of the Texas A&M University Supercomputing facility.

URBANMOD Data Requirements. URBANMOD requires an initial urban map, a region map, as well as an exclusion mask and driver maps. See *Data* for initial urban maps and region maps. The justification, sources, and derivation of the exclusion masks and the driver maps are detailed below.

Exclusion mask. In addition to driver maps, we use a map that allows masking out areas off limits to urbanization. As a conservative assumption, we mask out protected areas (PAs) to prevent infiltration of urban pixels. To create the mask map, we use International Union for Conservation of Nature (IUCN)-designated PAs from the World Database on Protected Areas 2010 database (16). Some PAs in the database do not have polygons but have centroid coordinates and area information. For those PAs, we create a circular buffer around the centroid approximating its spatial extent. We then use this initial polygon layer to create the final raster IUCN-PA layer with a resolution of 5,000 m in Goode's Homolosine projection. To create the exclusion mask, we combine this layer with the inland water mask we extract from the land cover driver map. A few urban pixels in the initial urban maps may fall into the existing IUCN-PA. We assume there will be no growth around those urban areas that are within an IUCN-PA.

Driver maps. The model takes maps of four factors that are assumed to have primary influences on where urban change is likely to occur.

Slope. Slope is generally accepted as a major factor influencing land-change processes, including urban land expansion. *Ceteris paribus*, gently sloped land is more preferable over land that is steeper. The slope map is derived from new Shuttle Radar Topography Mission (SRTM)-derived digital elevation model (DEM) and The Global Land One-kilometer Base Elevation (GLOBE) dataset (17, 18). The SRTM data covers the globe from 60° N to 56° S and, consequently, these data are not available for most of Scandinavia, northern Russia, and northern Canada. In these regions, we supplement the SRTM data with GLOBE to get the DEM information for these regions as well. We use 1-km resolution versions of both datasets. Once we generate a seamless DEM for all of the global land mass, we reproject it to Goode's Homolosine and we derive the slope map. We resample the resulting slope map to 5-km resolution.

Finally, because URBANMOD takes input maps with discrete categories only, we reclassified the map to discrete categories.

Weighted distance to roads. The global roads map comes from the major road networks layer of the VMAP0 vector-based global coverage product (19). The weight used to create this driver map is the slope map. The calculation of distances requires equidistant projections. However, there is no globally consistent equidistant projection. Therefore, we calculate the weighted distances to roads for each region separately, in each case using a suitable equidistant projection. Because there is no single projection that allows for accurate global distance-related analysis, each continent has to be projected separately. Therefore, we first project each region separately to an appropriately parameterized equidistant conic projection. After calculating the distances, we reproject back to Goode's Homolosine projection and discretize the resulting regional weighted distance maps. We assume the road network remains the same throughout the simulation period. This is likely to be an underestimate of the road network through 2030.

Population density. We use regional discretized population density maps based on the global map derived from the original GRUMP data (see *Data*).

Land cover. We convert the VMAP0 vector-based land cover product (19) to raster, reprojected it to Goode's Homolosine, and resample to 5 km. The reason we use VMAP0 product instead of MODIS v5 as we did for urban extent is that URBANMOD requires the initial data of the land cover to be simulated to be co-located with data from the driver maps. Although the initial urban-extent map overlays with the other three driver maps, this would not be the case for the land cover driver map had we used MODIS v5 product to create it.

- Friedl MA, et al. (2010) MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens Environ* 114:168–182.
- Schneider A, Friedl MA, Potere D (2009) A new map of global urban extent from MODIS satellite data. *Environ Res Lett* 4:44003.
- Center for International Earth Science Information Network (CIESIN) (2010) *Gridded Population of the World, version 3 (GPWv3) and the Global Rural-Urban Mapping Project (GRUMP)*, Available at: <http://sedacciesin.columbia.edu/gpw/>. Accessed June 13, 2010.
- United Nations (2009) *World Population Prospects: The 2008 Revision* (United Nations Department of Economic and Social Affairs/Population Division, New York).
- NRC (2000) *Beyond Six Billion: Forecasting the World's Population* (National Academy, Washington, DC).
- Lutz W, K C S (2010) Dimensions of global population projections: What do we know about future population trends and structures? *Philos Trans R Soc Lond B Biol Sci* 365: 2779–2791.
- Lutz W (1996) *The Future Population of the World: What Can We Assume Today?* (Earthscan, London, UK).
- United Nations (2010) *World Urbanization Prospects: The 2009 Revision* (United Nations Department of Economic and Social Affairs/Population Division, New York).
- Nakicenovic N, et al. (2000) *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. 599, Available at: <http://www.grida.no/climate/ipcc/emission/index.htm>. Accessed September 17, 2010.
- United Nations (2010) *National Accounts Estimates of Main Aggregates*, Available at: <http://unstats.un.org/unsd/default.htm>. Accessed September 17, 2010.
- Angel S, Parent J, Civco DL, Blei A, Potere D (2011) The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Prog Plann* 75: 53–107.
- Pontius RG, Cornell JD, Hall CAS (2001) Modeling the spatial pattern of land-use change with GEOMOD2: Application and validation for Costa Rica. *Agric Ecosyst Environ* 85:191–203.
- Pontius R, et al. (2008) Comparing the input, output, and validation maps for several models of land change. *Ann Reg Sci* 42:11–37.
- Echeverria C, Coomes DA, Hall M, Newton AC (2008) Spatially explicit models to analyze forest loss and fragmentation between 1976 and 2020 in southern Chile. *Ecol Modell* 212:439–449.
- Potere D, Schneider A (2007) A critical look at representations of urban areas in global maps. *GeoJournal* 69:55–80.
- WDPA (2010) *Annual Release*, Available at: <http://www.wdpa.org/>. Accessed April 21, 2010.
- GLOBE Task Team and others (Hastings DA, et al.) (1999) *The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0* (National Oceanic and Atmospheric Administration, Boulder, CO). Available at: <http://www.ngdc.noaa.gov/mgg/topo/globe.html>. Accessed March 31, 2010.
- The CGIAR Consortium for Spatial Information (CGIAR-CSI) (2008) *SRTM 90m Digital Elevation Data*, Available at: <http://srtm.csi.cgiar.org/>. Accessed February 4, 2010.
- NGA (2005) *VMAP0*, Available at: http://geoengine.nga.mil/muse/cgi-bin/rast_roam.cgi. Accessed March 24, 2010.

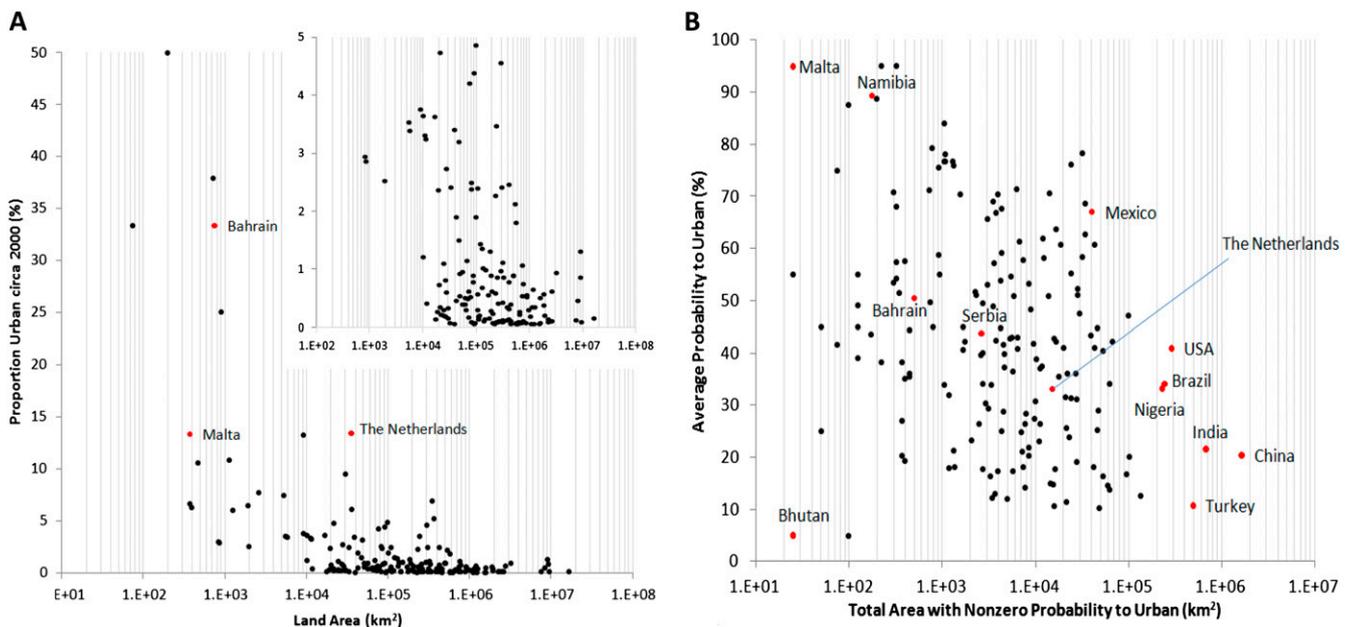


Fig. S1. (A) Land area of country versus proportion of urban land circa 2000. (Inset) Countries with less than 5% proportion urban. (B) Sum of all areas with nonzero probability to urban in each country versus average probability to urban across all areas with nonzero probability. Countries with no forecasted urban growth between 2000 and 2030 are not shown.

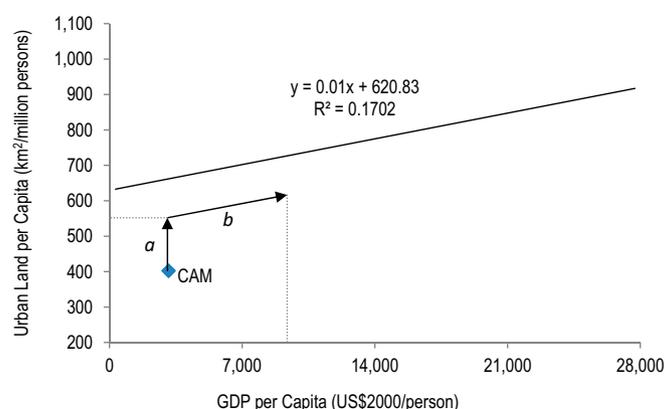


Fig. S5. One realization of per capita urban land in 2030 for Central America (CAM). First, the change in per capita urban land because of increase in urban population is estimated (a), *ceteris paribus*; then, an overall estimate for per capita urban land for 2030 is derived assuming the change in per capita urban land will, on average, be equal to the slope of the regression line (b), *ceteris paribus*.

Table S1. Composition of regions defined in the model

Regions defined in model	Abbreviation	Included United Nations regions	Plus	Minus
Central America	CAM	Central America, Caribbean	—	—
China	CHN	—	China, Hong Kong, Macao	—
Eastern Asia	EAS	Eastern Asia	Taiwan	China, Hong Kong, Macao, Mongolia
Eastern Europe	EEU	Eastern Europe	Kazakhstan, Estonia, Lithuania, Latvia, Albania, Bosnia-Herzegovina, Croatia, Macedonia, Montenegro, Serbia	—
India	IND	—	India	—
Mid-Asia	MAS	Central Asia	Mongolia	Kazakhstan
Mid-Latitudinal Africa	MLA	Western, Middle, Eastern Africa	—	—
Northern Africa	NAF	Northern Africa	—	—
Northern America	NAM	Northern America	—	—
Oceania	OCE	Oceania	—	—
South America	SAM	South America	—	—
Southeastern Asia	SEA	Southeastern Asia	—	—
Southern Africa	SAF	Southern Africa	—	—
Southern Asia	SAS	Southern Asia	—	India
Western Asia	WAS	Western Asia	—	—
Western Europe	WSE	Western, Southern, and Northern Europe	—	Estonia, Lithuania, Latvia, Albania, Bosnia-Herzegovina, Croatia, Macedonia, Montenegro, Serbia

Our regional breakdown broadly follows United Nations (UN) regional categorization. In a few instances, we removed countries from the UN regions they belong to (Minus) and included them in other UN regions (Plus). If a region as defined in the model does not include a UN region or has no removals or inclusions these are signified with an — in the table. See *SI Materials and Methods* for criteria on composing the regions defined in the model.