

# A spatial model to generate high-resolution projections of U.S. state-level population for coupled human-environment analysis

Hamidreza Zoraghein<sup>a\*</sup>, Brian O'Neill<sup>b</sup>

<sup>a</sup> Social and Behavioral Science Research, Population Council, New York, USA

<sup>b</sup> Frederick S. Pardee Center for International Futures, University of Denver, Denver, USA

\* hzoraghein@popcouncil.org

**Keywords:** spatial downscaling, human-environment analysis, population projection, scenarios

## Introduction

Projections of the spatial distribution of population are critical for modeling land-use/land-cover change (Meiyappan et al., 2014), urbanization (Gao and O'Neill, 2019) and risk assessment to climate change (Jones et al., 2015). We detail and analyze a spatial model that can generate population distributions at 1 km resolution in each U.S. state. The model combined with the proposed semantic framework facilitate creating projections of population distributions according to distinct socioeconomic scenarios for human-environment analysis. This work improves and tailors the gravity-based population downscaling model developed by Jones and O'Neill (2013, 2016), based on earlier work by Grübler et al. (2007) to each state. This model has several advantages. First, it relies on multiple ancillary datasets that make it adaptive to be improved as new data become available. Second, the well-defined structure of the model makes it easy to adapt to different study areas and requirements. Third, its underlying gravitational form employs two historically grounded parameters that characterize different patterns of spatial population change.

The model consists of calibration and projection components. We focus on the calibration component, in which we employ historical urban and rural population grids of each state in 2000 and 2010 to estimate the optimum pair of parameters that represent the state's spatial population change over the decade.

## Method

The model takes a spatial downscaling approach that allocates state-level aggregate urban/rural population change during a period to its constituent grid cells based on estimating urban/rural suitability values (Equation 1). Notably, when the population decreases, the change is allocated proportional to the inverse of suitability.

$$v_i = l_i \sum_{j=1}^n P_j^\alpha \times e^{-\beta d_{ij}}$$

$v_i$  is the suitability value estimated for the focal cell,  $i$ ,  $l_i$  is the mask value,  $P_j$  is the population of the neighboring cell,  $j$ , and  $d_{ij}$  is the distance between the focal cell and its neighboring cells. The summation over  $j$  is performed for  $n$  cells contained in the neighborhood within 100 km of the focal cell.

We derived a mask value ( $l_i$ ) for each cell to exogenously constrain population allocation according to physical barriers such as elevation, slope and land-cover, as well as mandates determined in the form of preserved areas by both federal and state governments. Figures 1 and 2 demonstrate the calculation of mask values.

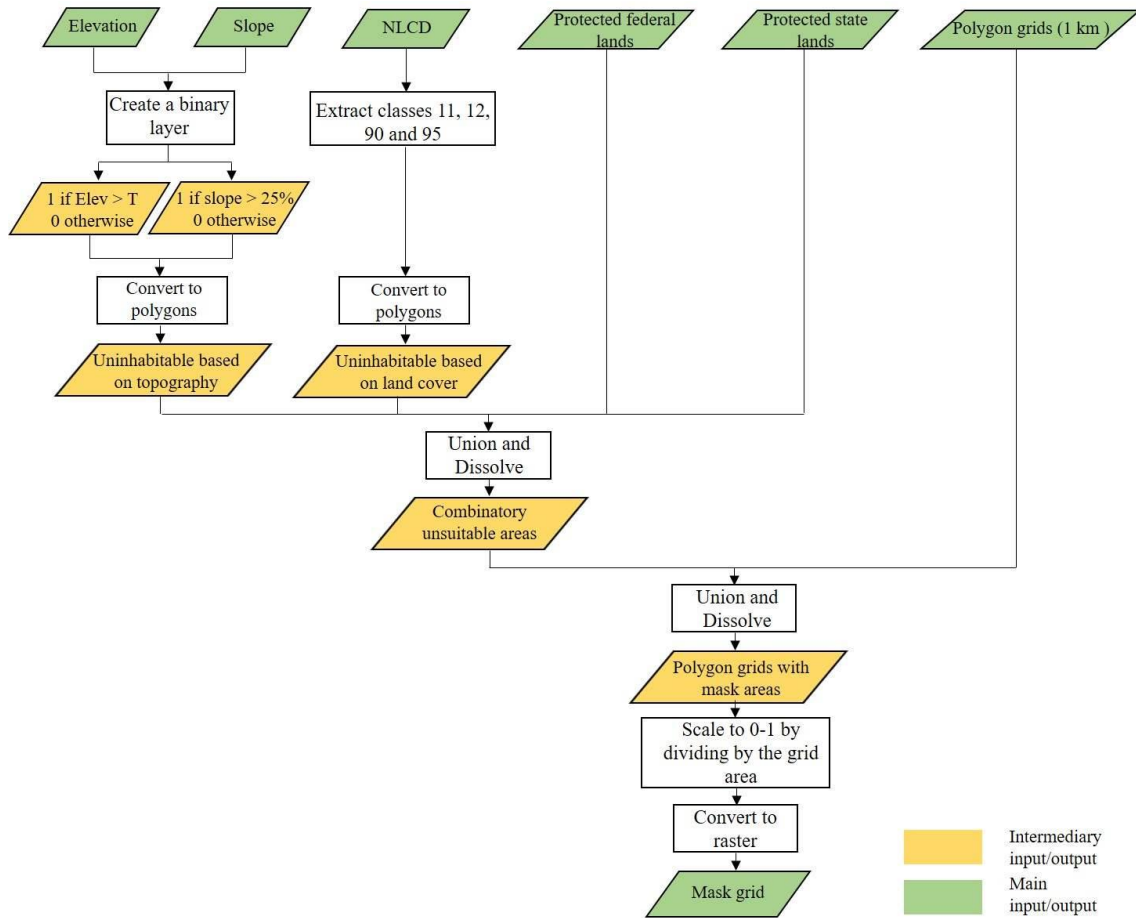


Figure 1. Steps to create the state-level spatial mask layer.

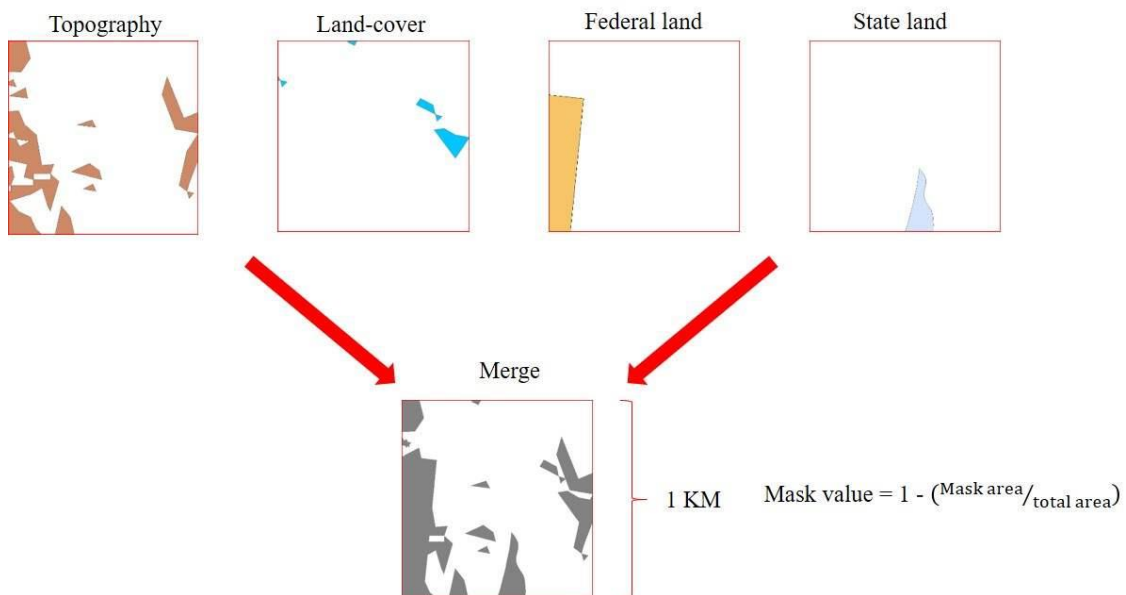


Figure 2. Estimation of spatial mask value for a hypothetical cell.

The  $\alpha$  parameter indicates the degree to which the population size of surrounding cells translates into the suitability of a focal cell. A positive value indicates that larger population within the 100 km neighborhood results in higher suitability of the focal cell (while a negative value of alpha would imply otherwise). The  $\beta$  parameter reflects the significance of distance to surrounding cells on the suitability of a focal cell. The higher positive value of the parameter indicates greater deterrent effect of distance. In that case, local characteristics prevail in attracting more population, and the presence of more distant (but still within 100 km) population centers matters less. In contrast, negative values imply a lower friction of distance. When  $\beta$  is 0, it means each cell contributes to the suitability of a focal cell only proportional to its population raised to  $\alpha$ .

We estimated  $\alpha$  and  $\beta$  parameters for the rural and urban population using an optimization approach (Figure 3). We generated 1 km resolution historical population grids in 2000 and 2010 for each state using its census blocks as the smallest available units.

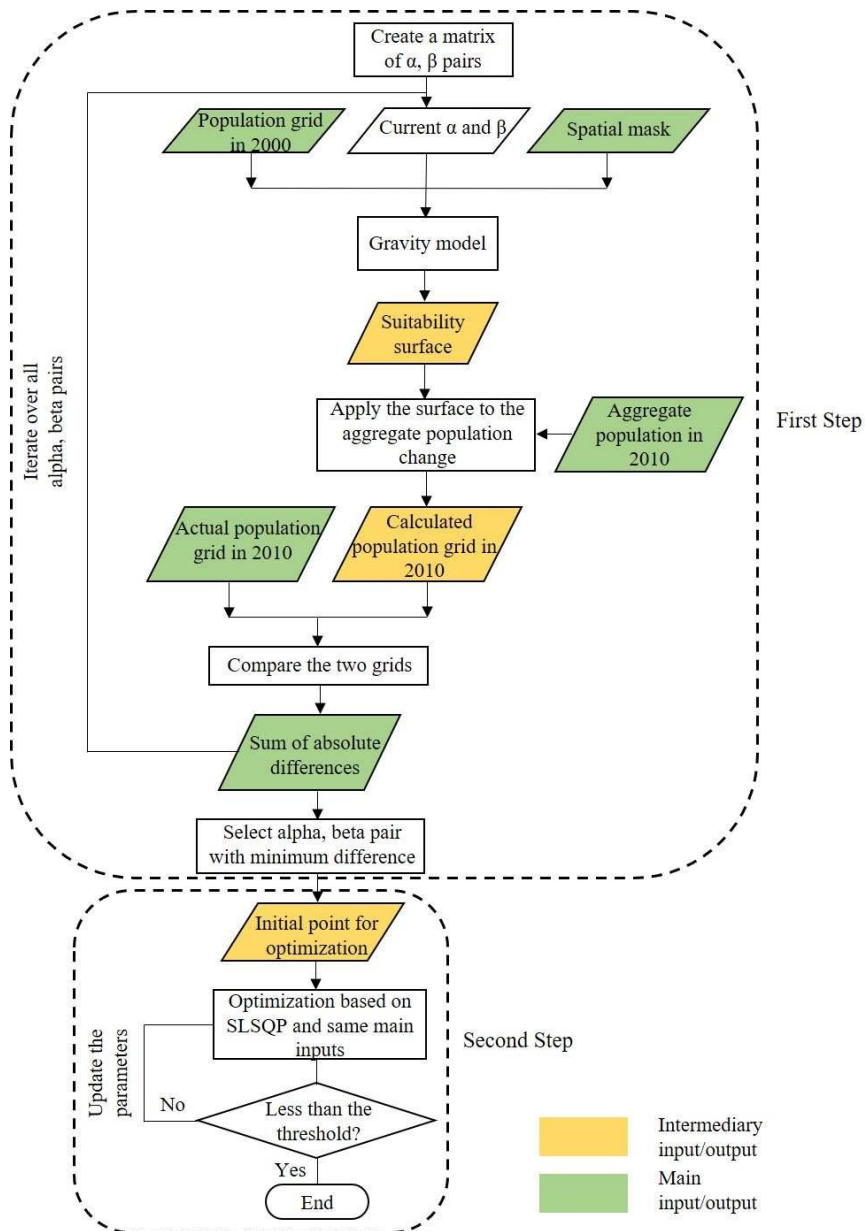


Figure 3. The calibration process of the gravity-based population downscaling model.

We formulated a semantic framework to map output parameters to different types of spatial population change (Figure 4). This important step allows us to modify parameters in compliance with socioeconomic scenarios such as the Intergovernmental Panel on Climate Change’s shared socioeconomic pathways (SSPs) (O’Neill et al., 2017) in the future.

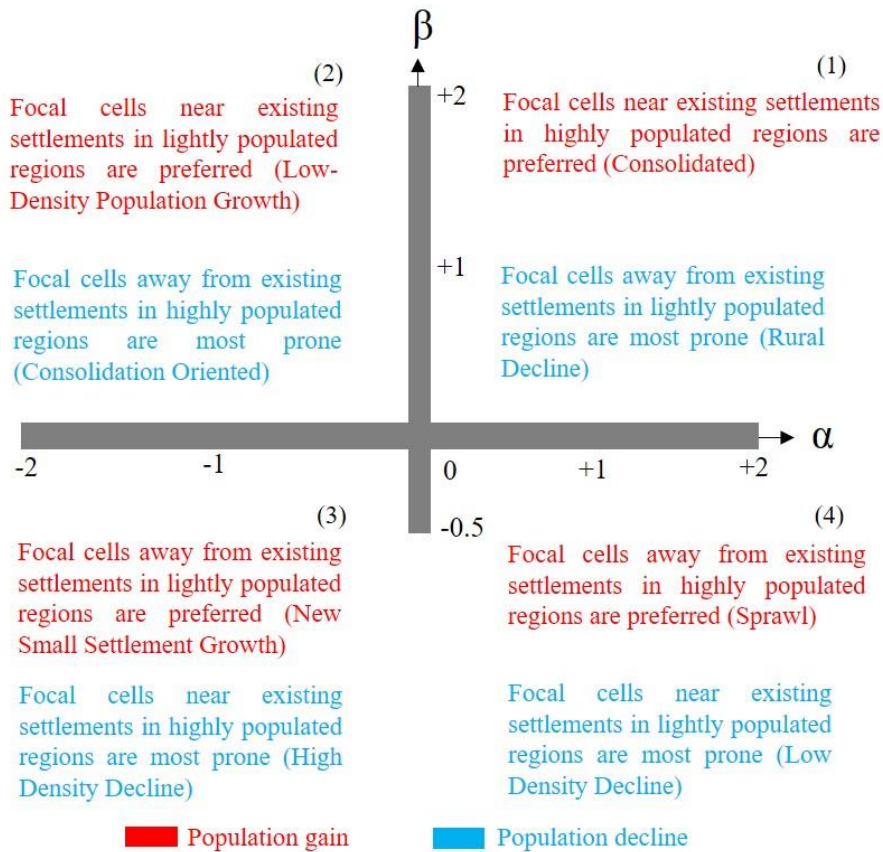


Figure 4. The influence of  $\alpha$  and  $\beta$  on the suitability of a focal cell.

### Results and Discussion

Figure 5 shows state-level choropleth maps of the distribution of estimated parameters. Figures 6 and 7 illustrate scatterplots of those parameters, respectively. We divided the plots into four quadrants consistent with Figure 4. We also differentiated states based on their sign of population change to reflect the points in Figure 4.

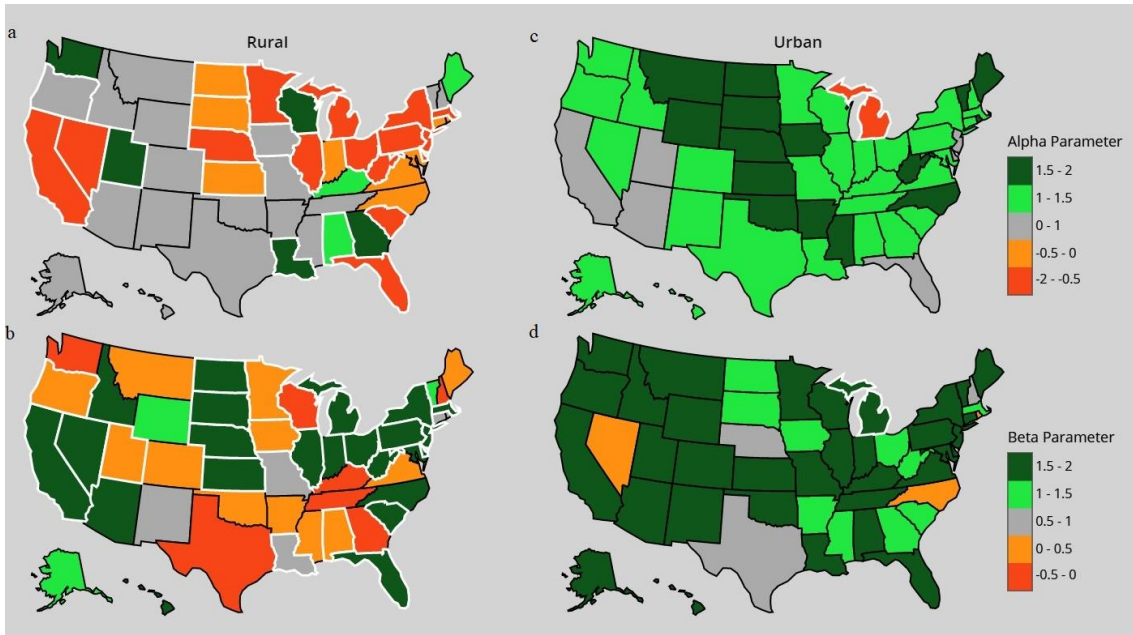


Figure 5. State-level (a)  $\alpha$  for rural population, (b)  $\beta$  for rural population, (c)  $\alpha$  for urban population and (d)  $\beta$  for urban population. Population declining states have white borders.

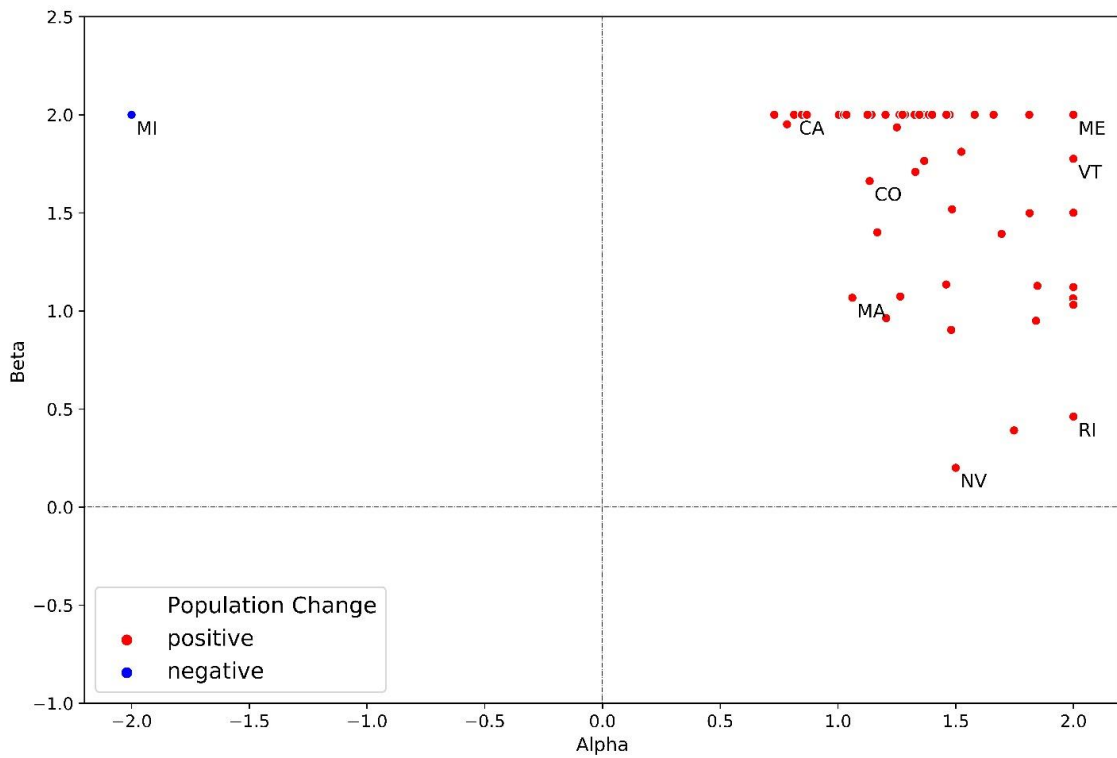


Figure 6. Scatterplot of state-level urban  $\alpha$  and  $\beta$  parameters.

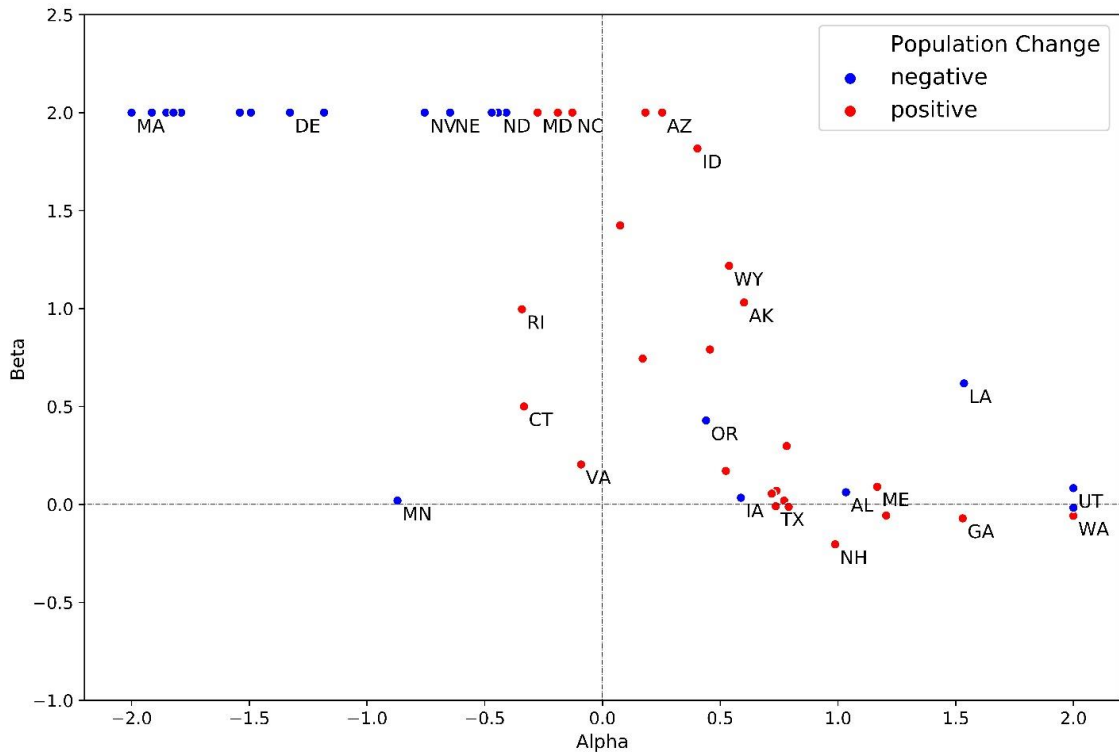


Figure 7. Scatterplot of state-level rural  $\alpha$  and  $\beta$  parameters.

Figure 6 shows that all states with urban population growth follow a consolidated urban growth pattern. Figure 8 shows a negative association between urban  $\alpha$  values and urbanization levels of states with growing urban population. This generally suggests the increasing importance of populous centers in states that have lower urbanization, where urban growth occurs most strongly in the most heavily populated regions. Conversely, in highly urbanized states, urban population growth is less strongly concentrated in the most highly populated regions. Michigan, as the only state with urban population decline lands in the second quadrant, suggesting a consolidation-oriented decline pattern. This is consistent with the historical pathway that Detroit, as the dominant urban center in the state, had experienced (McDonald, 2014).

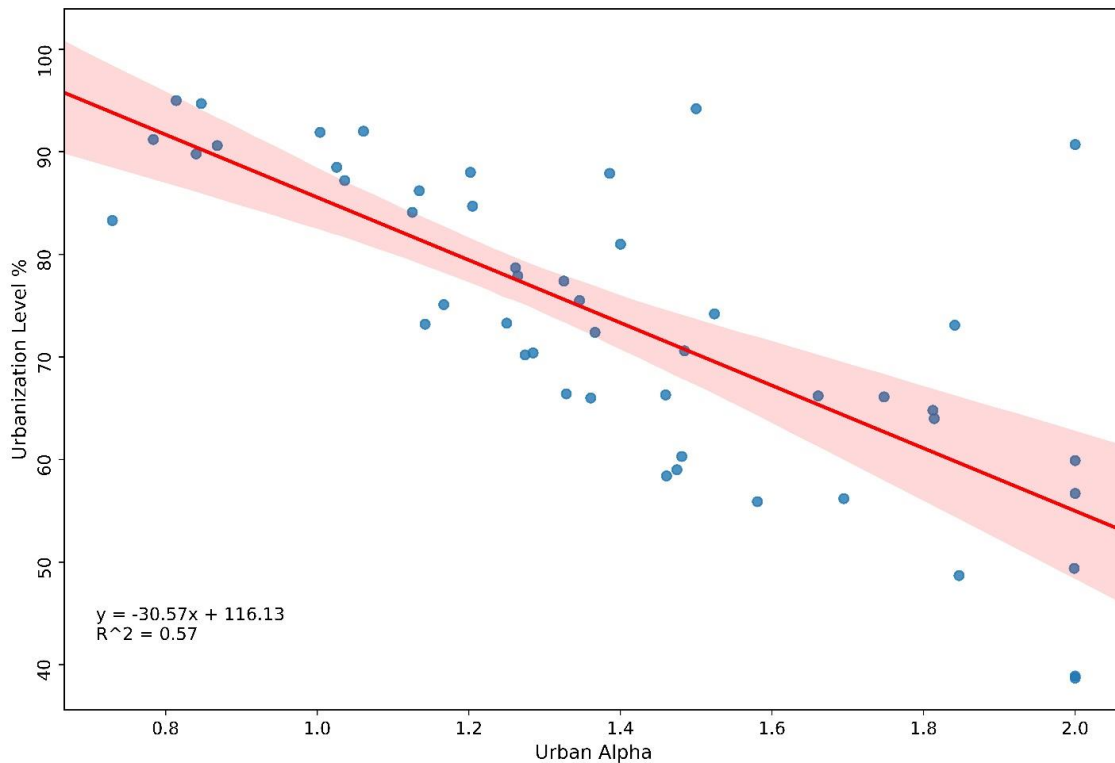


Figure 8. Relation of urbanization levels of U.S. states in 2010 (except Michigan) to their estimated urban  $\alpha$  parameters.

Figure 7 shows that states with the highest rural population growth, such as Alaska, Wyoming, Arizona, Maine and Idaho follow the consolidated pattern to varying degrees. The low rural population growth in Indiana, Maryland, North Carolina, Virginia, Connecticut and Rhode Island tends to settle close to existing population centers in low density regions. Texas, Kentucky, Tennessee, Georgia, New Hampshire and Washington are states with growing rural population with their parameters in the fourth quadrant, pointing to a sprawling development pattern. However, the resulting  $\beta$  values, except for New Hampshire, are close to 0, suggesting no significant preference on distance.

Many states with significant rural population decline such as Nevada, Nebraska, North Dakota, Kansas and Massachusetts are situated in the second quadrant with the maximum allowable  $\beta$  and varying degrees of  $\alpha$ , following the consolidation-oriented decline pattern. Oregon, Iowa, Alabama, Louisiana and Utah are states with rural population decline that lie in the first quadrant, pointing to a pattern of decline in the most remote, low density areas. The rural population decline in majority of these states is small.

The final output of the projection component of this model will be a set of state-level total population grids according to socioeconomic scenarios for human-environment analysis. Thus, we evaluated the calibration component based on total population grids (summation of urban and rural). We compared the estimated population grid in 2010 with the corresponding observed (block-based) grid. We derived cumulative distribution functions (cdf) of errors for each state. The horizontal axis represents absolute values of percentage differences while the vertical axis is the corresponding cumulative



population percentage. Errors are calculated based on mean population values over 10 km by 10 km windows to alleviate the spatial mismatch issue typical at the original 1 km resolution. Figure 9 shows these plots for several states.

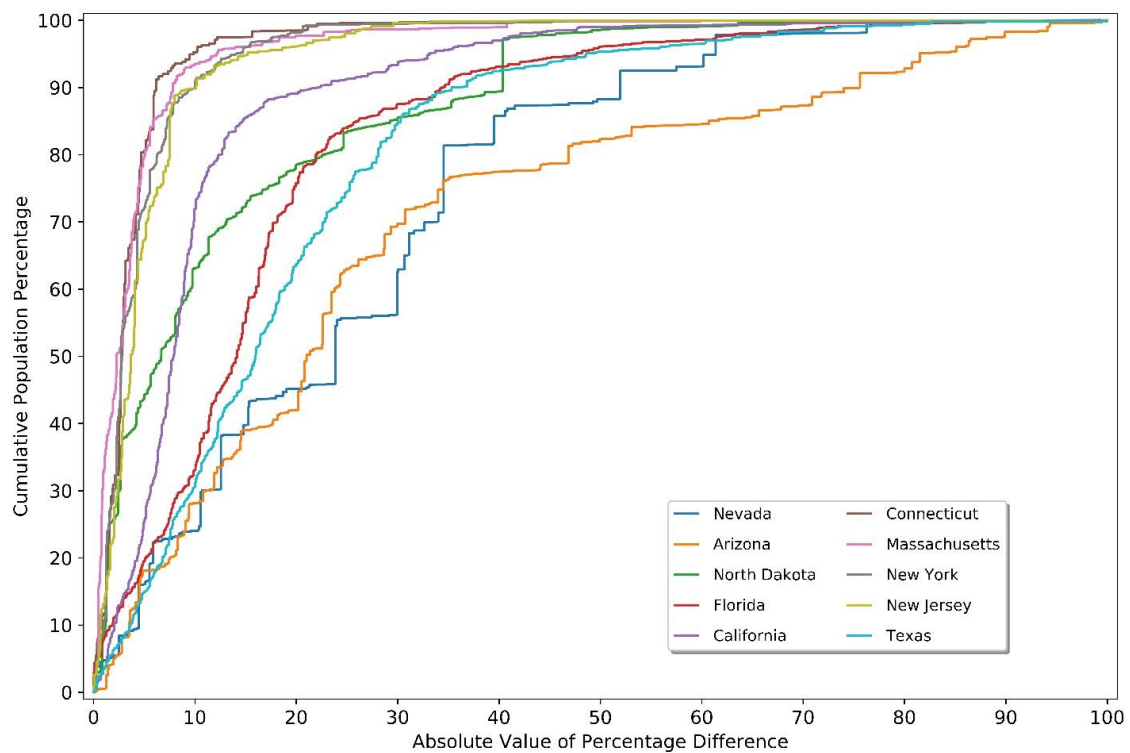


Figure 9. CDF plots based on total population grids.

Plots with narrower distributions indicate better model performance, whereas wider distributions indicate the contrary. Figure 10 shows mean absolute values of percentage differences corresponding to 50% and 90% of states' total population. While the performance of the model is overall satisfactory (e.g. 49 states with errors lower than 20% at 50% of their population and 38 with errors lower than 30% at 90% of their population), both figures demonstrate that the model performed well in some states including New York, New Jersey, Massachusetts and Connecticut while did not so in states such as Texas, Arizona and Nevada, suggesting the model's poor performance in capturing spatial change patterns within them.

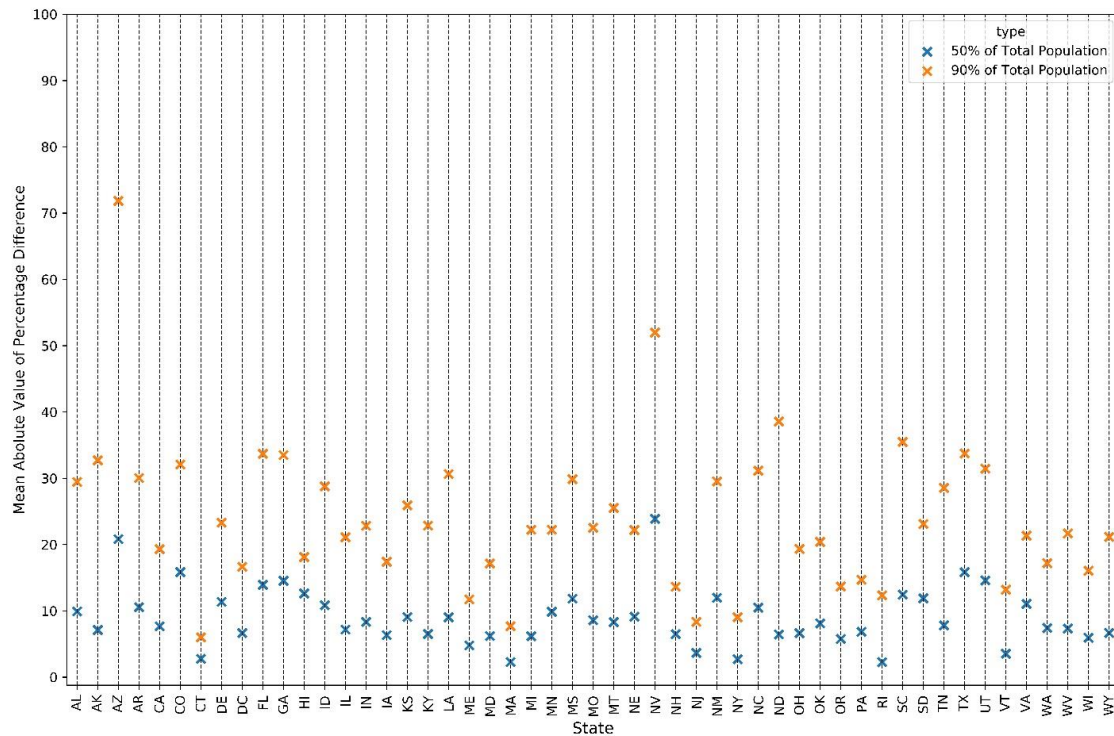


Figure 10. Absolute percentage differences at 50% and 90% of total population in each state.

## Conclusions and Future Direction

In this paper we calibrated and evaluated a spatial population downscaling model. The structure of the model and our semantic framework allow us to modify the model for population projection according to different socioeconomic scenarios. For example, current historically estimated values of Idaho (rural: (0.4, 1.82), urban: (1.48, 1.52)) can be retained for population projection from 2010 on according to SSP2 (business as usual). Based on historical parameters depicted in Figures 6 and 7 and for SSP3, as a scenario with slow urbanization and low population growth, those values can gradually change to rural: (-2, 2) and urban: (0, 2) over 2010 to 2100. For SSP5 with high urbanization and rapid population growth, they can gradually change to rural: (1, -0.5) and urban: (2, 0). We will create population projection grids for all U.S. states based on these suggested parameters and report on their performance in future work.

## References

- Gao, J., & O'Neill, B. C. (2019). Data-driven spatial modeling of global long-term urban land development: The SELECT model. *Environmental Modelling and Software*, 119, 458–471. doi: <https://doi.org/10.1016/j.envsoft.2019.06.015>
- Grübler, A., O'Neill, B., Riahi, K., Chirkov, V., Goujon, A., Kolp, P., ... Slentoe, E. (2007). Regional, national, and spatially explicit scenarios of demographic and economic change based on SRES. *Technological Forecasting and Social Change*. doi: <https://doi.org/10.1016/j.techfore.2006.05.023>

Jones, B., & O'Neill, B. C. (2013). Historically grounded spatial population projections for the continental United States. *Environmental Research Letters*, 8, 4, 044021. doi: <https://doi.org/10.1088/1748-9326/8/4/044021>

Jones, B., & O'Neill, B. C. (2016). Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters*, 11, 8. doi: <https://doi.org/10.1088/1748-9326/11/8/084003>

Jones, B., O'Neill, B. C., Mcdaniel, L., MCGinnis, S., Mearns, L. O., & Tebaldi, C. (2015). Future population exposure to US heat extremes. *Nature Climate Change*, 5, 7, 652–655. doi: <https://doi.org/10.1038/nclimate2631>

McDonald, J. F. (2014). What happened to and in Detroit? *Urban Studies*, 51, 16, 3309–3329.

Meiyappan, P., Dalton, M., O'Neill, B. C., & Jain, A. K. (2014). Spatial modeling of agricultural land use change at global scale. *Ecological Modelling*, 291, 152–174. doi: <https://doi.org/10.1016/j.ecolmodel.2014.07.027>

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., ... Solecki, W. (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change*, 42, 169–180. doi: <https://doi.org/10.1016/j.gloenvcha.2015.01.004>